POLISH MARITIME RESEARCH 1 (105) 2020 Vol. 27; pp. 70-82
10.2478/pomr-2020-0008

NEURAL NETWORK ENSEMBLE APPROACH TO PUSHED CONVOYS DISPATCHING PROBLEMS

Aleksandar Radonjić
Danijela Pjevčević
Vladislav Maraš
University of Belgrade, Faculty of Transport and Traffic Engineering, Republic of Serbia

ABSTRACT

This paper investigates the use of neural networks (NNs) for the problem of assigning push boats to barge convoys in inland waterway transportation (IWT). Push boat–barge convoy assignmentsare part of the daily decision-making process done by dispatchers in IWT companiesforwhich a decision support tool does not exist. The aim of this paper is to develop a Neural Network Ensemble (NNE) model that will be able to assist in push boat–barge convoy assignments based on the push boat power.The primary objective of this paper is to derive an NNE model for calculation of push boat Shaft Powers (SHPs) by using less than 100% of the experimental data available. The NNE model is applied to a real-world case of more than one shipping company from the Republic of Serbia, which is encountered on the Danube River. The solution obtained from the NNE model is compared toreal-world full-scale speed/power measurements carried out on Serbian push boats,as well as with the results obtained from the previous NNE model. It is found that the model is highly accurate, with scope for further improvements.

Keywords: Neural Network Ensembles, push boat, inland waterway transportation, full-scale speed/power trials

INTRODUCTION

In recent decades there have been significant changes in inland waterway transport (IWT) in Serbia. Safe conditions for inland navigation are being established through the maintenance and development of inland waterways, while communications between ships and operators are being improved. Conditions for the use of different types of both self-propelled and non-self-propelled inland ships are being determined in that way. This has turned out to be important in barge transportation due to the possible large number of combinations of barges in push boat–barge convoys (pushed convoys). At the same time, this has caused changes in the organisation of the IWT companies and their management structure.

IWT in Serbia plays a significant role in the transportation of cargo to and from Serbian river ports. There are roughly 2,500 ships, mostly consisting of barges owned by Serbian and foreign IWT companies. Companies cannot become more competitive by increasing their market share significantly unless they can set up cooperation with a large number of other operators.Severe competition and the resulting low margins in the main market imply that their profits can only be improved by lowering costs and improving their service by improving their operational efficiency [1].Furthermore, the options that are open to all IWT companies to improve their services are limited, as service improvement usually consists of complex organisational and technical actions necessary for the rational use of ships (barges), ports and waterways.

Some of the operational efficiency in IWT derives from the assignment of push boats to barge convoys (assignments). Assignments are made by dispatchers who work for companies on the basis of the availability of push boats and their power at a given moment. Most studies in the literature [1, 9, 20] have only focused on assignments without a precise analysis of the assignment possibilities on the basis of the push boat power and barge convoy speed. However, in order to operate at predefined speeds, there is still a need for calculation of the push boat power in pushed convoys. If the push boat cannotprovidethe necessary power, it will not be paired with the barge convoy. If there is enough power for the given pushed convoy speed and if there is more than one available push boat at a given time, there will then only be a question of which push boat can be more conveniently assigned to the barge convoy from the view point of the power use and energy consumption.

When it comes to dispatcher decision-making, it is assumed that the dispatcher has enough knowledge to make the right decision. Dispatcher decisions arise from previous experience and available full-scale speed/power trials [20]. It is not recommendedfor dispatchers to make decisions on the basis of their own calculations if thesehave not been previously verified. Correct assignments of push boats to barge convoys should be beneficial for the IWT companies.

Barge convoys propelled by push boats offer flexibility in barge scheduling as well as in cargo scheduling. For example, an IWT company owns and operates an inland fleet, including $b+p$ vessels. The fleet consists of b barges and p push boats. There are push boats of different sizes, from the smallest one that can push up to 1 barge to large ones that can push up to k barges ($k<b$). Between the smallest and largest push boats, there are many push boats that can push almost any barge convoy between 1 and k barges. It is assumed that each push boat can move any barge convoy and that it will operate at full power. The question is at what speed it can do it. Since the speeds are determined by the time of arrival of the cargo at the given location, the final question is whether each push boat can satisfy a predefined barge convoy time of arrival at full power. In other words, although more than one push boat can carry out transport with the given barge convoy, only one can do it with the most efficient use of power. The person who should answer these questions is the dispatcher.

Depending on the type of barges, different types of cargo can be transported between any two points where loading and unloading of barges take place. The interest of the IWT company is in transporting barges with the maximum draft. In order to avoid the transport of empty barges, terminals where barges are unloaded are used also for the loading of empty barges at the same time. In that way, barge convoys are disassembled and new convoy barges are assembled at the same terminal. If, for some reason, transport of a barge convoy must be stopped and one or more barges have to be taken out of it and left at anchorage, the new barge convoy without these barges will continue its journey to the next terminal for unloading and loading. While some barges are transferred to the terminal and then loaded and unloaded, the transport of each barge convoy is continuously carried on. And whether the barge convoys are assembled or disassembled, there is also a need for a change of push boats at the terminals if more than one push boat is intended to push them.

The aim of this paper is to propose the bestNNE model among several options that are idenitified in the process of training and testing neural networks(NN). The most important feature of the future NNE model will be the ability to select the best possible push boat for the transport of the given barge convoy with the minimal error. Consequently, minimal errors and variations of predicted data will be the criteria for the selection of various NNE models. The model is intended for use in situations in which the dispatcher makes decisions about assigning one of several available push boats to a barge convoy. Situations that precede or take place after the availability of push boats will not be considered. This makes the NNE model similar to a decision support tool for dispatching ships. After the initial information about the availability of push boats and before dispatching the barge convoys with push boats, the dispatcher should consult the NNE model for his decisions.At that moment, the NNE model should guarantee less work and more accurate information and assistance for the dispatcher. The NNE model's ability to predict the best push boat for a barge convoy could be a benefit for any IWT company.

The contributions of the paper are the following: to offer a completely new way of helping with dispatcher decisions, where the proposed model is a reliable approach for push boat–barge convoy assignment.

For the design and building of the neural network architectures in the NNE model, data from full-scale speed/power trialsare used.Classification of the neural networks in the NNE model is done by AdaBoost Regression and Threshold (AdaBoost.RT), given in [17]. NNs are trained by theResilient backpropagation(RPROP) learning algorithm. Several NNE models are made in the process of finding a final solution before the best model is selected on the basis ofgraphical and numerical methods.Finally, the best NNE model is compared with the previous NNE model given in [11].

As a result oftraining NNs and applying the AdaBoost. RT algorithm over roughly 80% ofthe data from a target dataset [4, 11], it is shown that the NNE model can be used to assign a push boat to a barge convoy and carry out transport in a predetermined period of time.The results have shown that anew NNE model could be made quickly if the new data had to be included in a partly repeated process of modelling the new NNE and training itsNNs. However, the NNE model is not applicable to shallow draft navigation.

LITERATURE REVIEW

Over the past few years there has been a significant increase in the number of papers investigating the use of NNs in IWT. In addition to using NNs, researchers also use regression analysis to solve various problems in IWT. In the past, regression analysis was used to solve every problem, but it isgradually being replaced by NNs.

Today, NNs are being used as an important decision support tool in a variety of applications. Papers related to IWTand in which NNs are investigated as a tool for solving problems in IWT can be roughly divided into the following two groups: papers directly related to IWT [10, 11, 20] and papersin which NNs can be applied indirectly or for solving problems that are beneficial for IWT [5, 12].The second group of papers are usuallythose in which NNs are used to predict the values of resistance of ships or coefficients that are important for shipbuilding.

Couser & Mason [5] investigated the accuracy of NNs as prediction tools for the hull resistance of catamarans. The goal of the paper was to determine a predictive model for residuary resistance based on input values of the Froude number, beam-to-length ratio, draft-to-beam and length-toslenderness ratio. The data used for the investigation originated from a series of tank tests. The authors demonstrated that a combination of genetic algorithms and artificial neural networks could be used as an optimisation tool for catamaran design parameters.

Vukadinovic et al. [20] proposed anNN to be used as the dispatcher's decision support system. The main objective was to apply the NN technique to the assignment of loaded barges to pusher boats for the planned period of one day within inland waterways and to develop the dispatcher's decision support system. They showed that the proposed NN can be used to help the dispatcher's decision making. Their results could be applied to any mode of transportation.

Reich & Berai [12] created an NNE model to predict the propeller thrust coefficient, the propeller torque coefficient and the propeller efficiency. Input variables were the propeller pitch–propeller diameter ratio, expended area ratio, number of blades, advance coefficient and cavitation number. They concluded that NNE model design should be done carefully, starting from data collection, model quality estimation, to solution deployment. They recommended that further improvement of the NNE model quality is possible by using advanced methods.

Radonjic and Vukadinovic [11] proposed an AIC-based NNE model and a single NN with two hidden layers to predict pushboat shaft power. These two models were compared on the basis of their calculated mean absolute error (*MAE*) values, root mean squared error (*RMSE*) values and relative errors. Computational results from this numerical example showed that the NNE definitely outperformed a single NN with two hidden layers.

Parks et al. [10] appliedthe NN technique to train and test data acquired during normal operational service of three sister merchant vessels. A key aspect in their paper was to determine NN architectures that could give close relationships between the input variables and shaft power. The input variables were the following: GPS ship speed, wave height, true wind speed, apparent wind direction, draft and trim.Predictions of shaft power were made by NNs. Their values werecompared to the values calculated using regression analysis on the same dataset.The predictions of shaft powershowed an error of less than 10%, while the NN showed good repeatability of the relationshipsbetween the input variables and the measured shaft power. They concluded that their method and their NN may be capable of providing a baseline for performance monitoring across a wide range of environmental conditions, thus allowing faster decision-making.

PROBLEM DESCRIPTION

Part of the cargo transport in which push boats are assigned to barge convoys has to be solved by dispatchers in operating centres. Fig. 1 shows an example of a dispatcher'sdaily push boat to barge convoy assignment decisions based on experience.

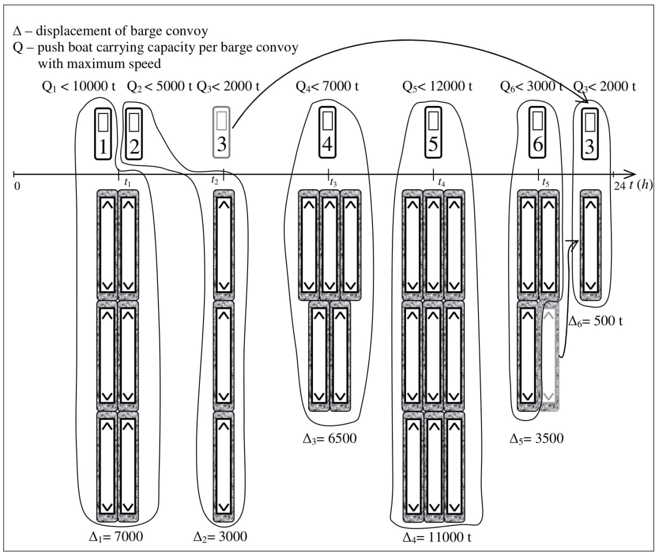

Fig. 1. Example of dispatcher's push boat to barge convoy assignment decisions

Loaded barge convoys are ready to be transported by push boats at certain times. If there is enough power to propel the given barge convoy at maximum speed, a push boat will be assigned to thebarge convoy by the dispatcher. Otherwise, the dispatcher will wait for another more powerful push boat, ora barge convoy will be disassembledin such a way as to satisfy the power constraints of the available push boat. In Fig. 1, two push boats are available for transporting the first barge convoy at time t_1. Since the first push boat is able to transport the first barge convoyon its own ($\Delta_1<Q_1$), it is assigned accordingly and the second push boat is left to wait for another barge convoy. At time t_2, the second barge convoy is formed and is ready to be pushed to its final destination. Like a couple of hours earlier, two push boats are available to transport it, but the second push boat meets the criteria for coupling with it ($\Delta_2<Q_2$ and $\Delta_2>Q_3$). As can be seen in Fig. 1, the third push boat cannot be assigned to any of the upcoming barge convoys until time t_5. At that time,the fifth barge convoy is reduced to three barges in order to be available for the sixth push boat, and the convoy's fourth barge is assigned to the third push boat.

All the dispatcher assignments in Fig. 1 are results of his experience,and are not confirmed bya mathematical model. Therefore, an NNE model is proposed in this paper to help the dispatcher with his decision and improve the existing decision-making process.

Dispatchers should enter the input values of the pushed convoy's slenderness ratio, length-to-beam ratio, draft-to-beam ratio, Froude number, propeller diameter, propulsive efficiency and cavitation number into the NNE model to get the required push boat SHP. On the basis of the calculated

SHP and NNE model errors, the dispatchers decide if the push boat will be able to carry out the transportation of the given barges.

DATA DESCRIPTION

Full-scale speed/power trials are performed whenever an IWT company has to confirm that the newly built ship has met its specification as regards design speed [18], whenever a new propeller is being installed on a push boat, and for scientific purposes [4]. Data from the full-scale speed/power trials are divided into row data and target data in this paper. Row data consist of all data from the trials, while target data encompass only the data that will be used in the process of NN training and the prediction performance of the NNE model. The idea is to separate valid target data from the data that are unnecessary and do not contribute to improvement of the NNE model results, but can increase the NN training time.

For example, all the new data in this paper consisting of pushed convoy speeds lower than the highest speeds (which can be reached at full power of the push boats' main engines) will not be added to the target data. This means that there will be many combinations of push boats and barge convoys with the same SHPs, but different input data in the target data. Thisfurther means that the NNE model is a complex nonlinear prediction model.

Target data are gathered from full-scale speed/power trials of Serbian push boats. Some of the data are listed in [11], while the rest are given in [4]. Together, thesemake upthe single target dataset in this paper.

Input data included the following variables: pushed convoy slenderness ratio (length–displacement ratio or $\nabla^{1/3}/L$), pushed convoy length-to-beam ratio (L/B), pushed convoy draft-to-beam ratio (d/B), Froude number (F_r), propeller diameter (D_p), propulsive efficiency η_D) and cavitation number (σ).The number of blades was not included in the input variables due to the fact that all propellers have the same number of blades.

The output data included the push boat shaft power (*SHP*),making the future NNE modela model for the estimation of the power requirements in pushed convoys. It was assumed that the push boat *SHP* was a function of the geometric characteristics of the pushed convoys,the pushed convoy speed through water,and the propulsive characteristics of the push boats. The geometric characteristics and pushed convoy speed through water were represented in the first four input variables, while the propulsive characteristics included the last three input variables. Pushed convoys with the same number ofbarges arranged in a row and with ahigher draughtrequiremore power to operate than those arranged in a line. The ratios of the main dimensions are incorporated into the model as inputs accordingly [11].The last three input variables are added into the input dataset because the motions of pushed convoys areaffected also by the propulsive characteristics of the push boats, and are expressed in power losses. The push boat *SHP* was taken as an output variable because full-scale speed/power trials were performed on the basis of measurements of shaft power on push boats.

The example set of input values is given in Tab. 1. [next page]

The greatest power losses are expressed via η_D and appear in the segment of the push boat power chain which includes part of the chain from shaft power (*SHP*)to effective power (P_e) (see Fig. 2). η_D is calculated as the ratio of effective power (P_e) per shaft power (*SHP*). Power losses due topropeller geometry and operating conditions are described by propeller diameter (D_p) and cavitation number (σ) [12]. Cavitation causes several undesirable effects [22], among which reduction of the propeller performance is the most important part that affects the value of the pushboat SHP, while the increase of D_p causes an increase in the pushboat SHP.

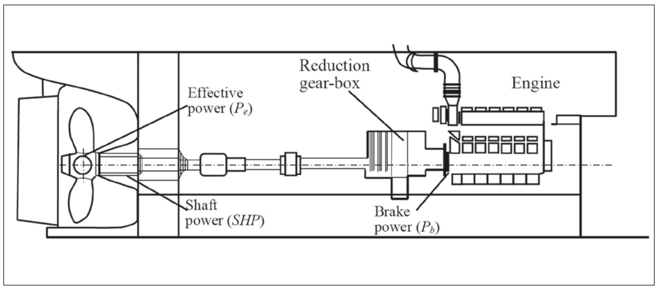

Fig. 2. Schematic overview of the push boat propulsion system (power losses expressed by η_D, D_p and σ) (Source: [4])

The first four input variables have been calculated by applying the methodology presented in [11] to all combinations of push boats and barge convoys in this paper. All the data in the target dataset are normalised by linear transformation to [0,1] [23].

The values of the propeller diameters were provided by [4], while cavitation numbers were calculated on the basis of the Burrill cavitation number [3].Propulsive efficiencies (η_D) are calculated as a ratio of each push boat'seffective power (P_e) and each push boat'smeasured *SHP*. The push boat effective powers (P_e) were calculated from [19].

The process of designing the NNE model was divided into two stages. The first stage included training of NNs, while the second stage included validation of the NNE model. For the training purposes, appropriate data splitting was done. It was assumed that NN training over bigger datasets should output NNE models with smallerroot mean square error (*RMSE*) and mean average error (*MAE*)values, as well as models with smaller deviations and variations.

The target dataset was divided according to combinations of push boat and barge convoysinto smaller datasets with roughly 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% of all the collected target data. The number of datasets is denoted as N ($n = 1, 2,\dots, 9$, $N = 9$). These datasets were divided into test data and training datathat were used in the NN training process. The rest of the target data were included as test data. However, 100% of the target data were used to test the NNE models. Before the testing, all input data were normalised [23] so that they could be processed in the NNE model. Upon the testing, all output data were denormalised and as such they were analysed. The aim was to develop an NNE model that

Tab.1. Example set of input values related to few pushed convoys

Length-to-beam ratio (L/B)	Draft-to-beam ratio (d/B)	Slenderness ratio ($V^{1/3}/L$)	Froude Number (F_r)	Propeller diameter (D_p) in meters (mm)	Propulsive efficiency (η_D)	Cavitation number (σ)
8.941176	0.099020	0.095382	0.065996	1500	0.270571	0.600208
8.219453	0.112414	0.109381	0.06497	1650	0.335688	0.704681
11.49316	0.112414	0.088705	0.046268	1650	0.301756	0.710474
5.481421	0.074967	0.124034	0.058131	1650	0.335095	0.639230
6.055263	0.144737	0.138876	0.134967	1600	0.399969	0.715922
10.13421	0.144737	0.101854	0.087178	1600	0.305555	0.734457
6.75614	0.096491	0.115529	0.071457	1600	0.235865	0.748878
14.21316	0.144737	0.082374	0.059132	1600	0.197835	0.751965
5.067105	0.072368	0.126561	0.062168	1600	0.208157	0.760211
10.47479	0.222743	0.114307	0.128539	1170	0.596787	0.655714
5.237397	0.111372	0.141232	0.10696	1170	0.52969	0.695191
5.915378	0.072877	0.117171	0.084462	1800	0.517824	0.559295
8.873066	0.109316	0.10288	0.100339	1800	0.558304	0.524407
25.02366	0.218631	0.066571	0.089311	1800	0.468053	0.519611
5.232484	0.109316	0.140503	0.129009	1800	0.329565	0.525384
17.74431	0.218631	0.082894	0.117491	1800	0.493101	0.509415
6.832968	0.086881	0.1112	0.086241	1800	0.256204	0.565607
5.121919	0.064661	0.121318	0.077503	1800	0.248758	0.573698
9.542507	0.086374	0.090167	0.0604	1800	0.234184	0.573168
6.77734	0.094015	0.11404	0.076667	1800	0.279898	0.786509
9.487752	0.092946	0.088939	0.056607	1800	0.279624	0.813906
14.77902	0.212222	0.091492	0.109589	1600	0.388818	0.758877
4.342943	0.106202	0.155808	0.135065	1600	0.391985	0.787975
20.89252	0.214087	0.07342	0.079085	1600	0.396488	0.797247
7.398881	0.107136	0.103749	0.0886	1600	0.400375	0.809415

would not fully depend on all the target data during model creation. At the same time, the new NNE model would be valid only if it was tested on all the target data. The validation process of the NNE model is in accordance with the policies of shipping companies, in which cargo is transported with fully loaded ships on rivers.

There was a total of 918 data points that formed the target dataset. The scatter plots of the target dataset with all the measured data are composed using the measured pushed convoy SHPs data (ordinate) and pushed convoy slenderness ratios ($\nabla^{1/3}/L$), pushed convoy length-to-beam ratios (L/B), pushed convoy draft-to-beam ratios (d/B), Froude numbers (F_r), propeller diameters (D_p), propulsive efficiencies (η_D) and cavitation numbers (σ) (abscissas), as shown in Fig. 3 [next page].

After careful observation, the complex nonlinear characteristics of the measurements data are found in Fig. 3.

Division of the target dataset into smaller datasets (n) was done according to the following rule:

1. Number 193 was set as the 0% of data, as these 193 data were initial for the creation of the rest of the target dataset.
2. The difference between number 918 and number 193 was multiplied by 10% or 20% or … or 90%, then added to 193 and rounded down to the integer to get the numbers of data points in each dataset n.

Based on the previous rule, the following numbers of datapoints per datasets n were obtained: 266 for n = 1.339 for n = 2.412 for n = 3.485 for n = 4.557 for n = 5.630 for n = 6.703 for n = 7.776 for n = 8 and 849 for n = 9.

The principle of target dataset partitioning was done in accordance with [21], which found that too many or too few samples in the training set havea negative effect on the estimated modelperformance, and thata good balance between the sizes of the training set and validation set is necessary for reliable estimation of model performance.

METHODOLOGY

NEURAL NETWORK ENSEMBLE METHODOLOGY

Prediction of the push boat *SHP* was done by usingseveral Feed Forward NNs. NNs were trained by using the RPROP learning algorithm [14]. The AdaBoost.RT algorithm [17] was used to combine NNs in an ensemble and to create an

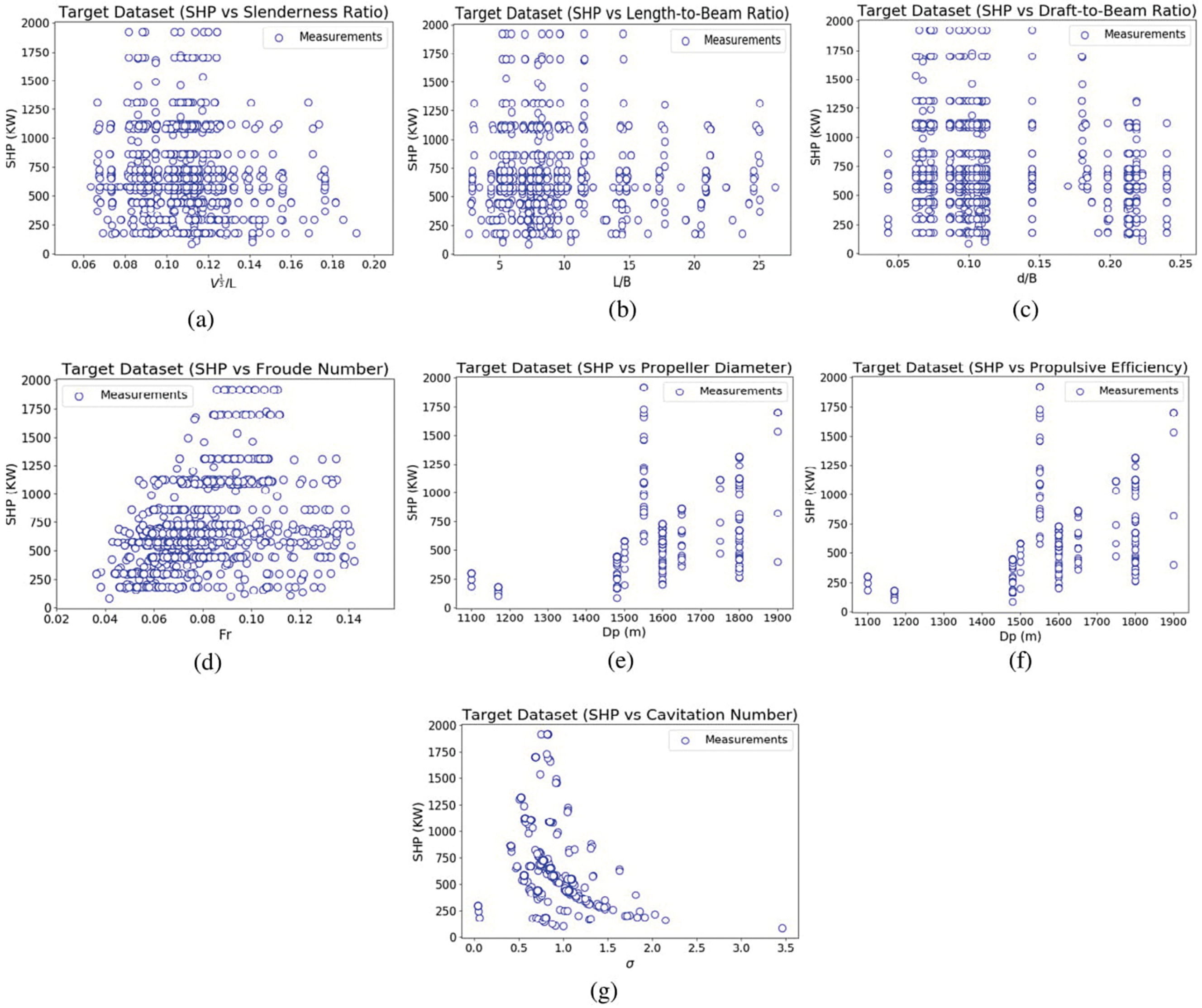

Fig. 3. The nonlinear characteristics of measured full-scale speed/power trials data

NNE model. It has been suggested in [6, 15, 16] that boosting techniques and particularly AdaBoostwork well with NNs.

For the purposes of modelling, Python programming language is used for training and testing as well as for weights (W_{t_n}) determination by the resampling technique [2] within the AdaBoost.RT algorithm, where $t_n = 1, 2, \ldots, T_n$ is the number of potential NNE models per dataset n.

AdaBoost.RT was first outlined as an adaptive ensemble method by [17]. Its main difference from other AdaBoost algorithms is its threshold (θ), which separates correct and incorrect predictions. The main part of the algorithm was to determine weights (W_{t_n}) in order to improve predictions from separate models. The separate models are the NNs in this paper. The AdaBoost.RT presented in [17]does not specifically address the number (T_n) of NNs in any given ensemble model. It is obviously left to the authors to set T_n, depending on the level of difficulty of the problem they encounter during the training of NNs. Therefore, T_n was not fixed before the beginning of the application of the AdaBoost.RT algorithm in this paper.

For the purpose of getting the best possible NNE model, it was decided that the best procedure for this investigation should contain twostages. These consisted of training and testing single-hidden layer NNs (NNs in the following text) on the different datasets described in the section "DATA DESCRIPTION".

In the first stage, K-fold cross-validation [8] and a golden-section search procedure were done to avoid overfitting, to determine the number T_n, to select NNs for the NNE model and to get a more accurate estimate of the NNE model prediction performance. Datasets partitioning was performed by $K = 4$ different datasets divisions. They were split into 3, 4, 5 and 10 consecutive parts. Each part was then used once as a validation and as a test set in the same time, while the other threeparts made the training set. Based on the number of training data (N_{tr}), the number of inputs (N_I) and number of outputs (N_O), and number of hidden nodes (N_h) was calculated by the methodology presented in [13, 20] for every dataset. As N_I and N_O, were fixed, the architecture of the NN per dataset depended only on N_h. Calculation of more N_h per each dataset n meant that more NNs per dataset n were potential candidates for the NNE model.

Following the large number of NNs generated as a result of data splitting, dataset partitioning and $N_h$calculation, only carefully selected NNs were selected to be trained. This was done due to avery possible scenario that could includethe training of NNs that would never enter the NNE model, which would also be a waste of time in the process of creating the NNE model. The resulting numbers of selected NNs and sets of possible NNs that could enter the NNE model per dataset were obtained by the golden-section search procedure [13]. Thiswas the input for the determination of T_n in the AdaBoost.RT algorithm.

Based ongolden-section point calculation, each NN was trained 10 times with 3, 4, 5 and 10 different combinations of trainedparts and tested partsthat were made up with the help of bootstrap sampling [7]. The average *RMSE* value(of 10 NNs of the same architectures) calculated overthe trained datawasinput for the calculation of the AIC numbers [13] of one NN architecture per part. With the help of the AIC numbers, the best NN architecture per part as well as a collection of trained NNs per part were gathered to form a *set of trained NNs per one part* (S_{nK}). Since there were 4 parts per dataset, 4 S_{nK} setswere formedafter the application of the golden-section search procedure. The intersection of four different sets (S_{nK}) gave another set, the final set (S_n) ofpotential NN candidates for the NNE model within a dataset$n$. The first stage is explained in Fig. 4.

In the second stage, the AdaBoost.RT algorithm was applied to create the best NNE model per each dataset n. First, set $P(S_n)$ was created for all n. Each $P(S_n)$ was the collection of all subsets of S_n sets. The number of possible NNE models per each dataset n (C_{T_n}) was equal to the number of different combinations of NNs in each dataset n. It is defined by the following Eq. (1):

$$C_{T_n} = \sum_{r=2}^{T_n} \begin{bmatrix} T_n \\ r \end{bmatrix} \quad (1)$$

Dataset n was split following a 72/28 rule, where 72% of the data were used for training NNs, while 28% of the data were used as test data. Each NN from $P(S_n)$ was trained to get the functional relationship between the inputs and output ($f_{t_n}(x)$). $f_{t_n}(x)$ is used to get the weights of the NNs (W_{t_n}) with the AdaBoost.RT algorithm. The AdaBoost.RT algorithm applied with the previous assumptions is as follows:

START (dataset n, c_{T_n}-th combination)
Input: $(x_1, y_1), \ldots, (x_{i_n}, y_{i_n}), \ldots, (x_{m_n}, y_{m_n}), x, y \in \mathfrak{R}$;
Where:
$x_{i_n} = (\nabla_{i_n}^{1/3}/L_{i_n}, L_{i_n}/B_{i_n}, d_{i_n}/B_{i_n}, F_{ri_n}, D_{pi_n}, \eta_{Di_n}, \sigma_{i_n})$
$y_{i_n} = SHP_{i_n}$
Initialise the distribution for all i_n: $\Lambda_{t_n}(i_n) = \frac{1}{m_n}$
Set the threshold: $\theta = 0.07$

FOR $t_n = 1, 2, \ldots, T_n$
IF $t_n = 1$
Select the training data based on random numbers End IF
IF $t_n > 1$
Select the training data based on distribution $\Lambda_{t_n}(i_n)$ End IF

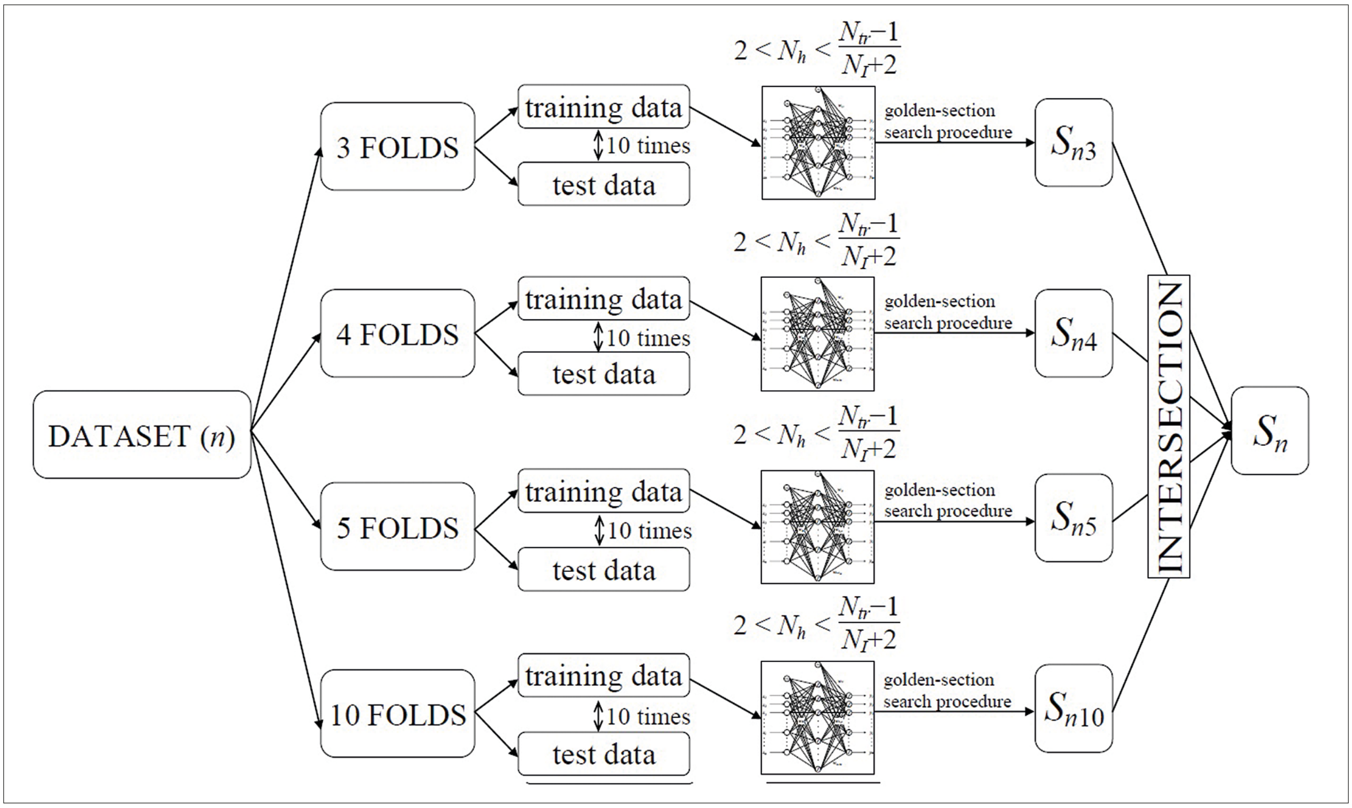

Fig. 4. An example of a derivation of one set (Sn) from K-fold cross-validation and golden-section search procedure applied together with NN training in the first stage

Train one hidden layer NN with a maximum of 100 iterations before calculation of total error of NN. If total error doesnot drop in the next three 100 iterations, stop NN training, get the values of weights where the total error was minimum and obtain $f_{t_n}(x)$ as

$$f_{t_n}(x) = s\left([w_{hO}]^T_{N_h,1}\left[s([w_{Ih}]^T_{N_I,N_h}[x_I]_{N_I,1} + [w_{bh}]^T_{1,N_h}) + w_{bO}\right]\right) \quad (2)$$

where:
w_{Ih} – weights between I-th input neuron and h-th hidden neuron (randomly initialized at the network configuration)
w_{hO} – weights between h-th hidden neuron and O-th output neuron (randomly initialized at the network configuration)
w_{bh} and w_{bO} – bias parameters (always have value of 1)
$f_{t_n}(x)$ – push boat *SHP* function based on $\nabla^{1/3}/L$, L/B, d/B, F_r, D_p, η_D and σ
s – sigmoid activation function
$N_h = N_{ht_n}$

Calculate absolute relative error for each dataset example as:

$$ARE_{t_n}(i) = \left|\frac{f_{t_n}(x_{i_n}) - y_{i_n}}{y_{i_n}}\right| \quad (3)$$

Calculate the error rate of $f_{t_n}(x)$

$$\varepsilon_{t_n} = \sum_{i_n \cdot ARE_{t_n}(i_n) > \theta} \Lambda_{t_n}(i_n) \quad (4)$$

Calculate $\omega_{t_n} = \varepsilon^c_{t_n}$, where c is a power coefficient (e.g. linear, square or cubic)

Update distribution $\Lambda_{t_n}(i_n)$ as

$$\Lambda_{t_n+1}(i_n) = \begin{cases} \dfrac{\Lambda_{t_n}(i_n)}{Z_{t_n}} \times \omega_{t_n},\ ARE_{t_n}(i_n) \le \theta \\ \dfrac{\Lambda_{t_n}(i_n)}{Z_{t_n}} \times 1,\ otherwise \end{cases} \quad (5)$$

Z_{t_n} is a normalisation factor chosen such that Λ_{t_n+1} will be distribution

END loop

Output the final weights

$$W_{t_n} = \frac{\dfrac{1}{\log\omega_{t_n}}}{\sum_{t_n=1}^{T_n} \dfrac{1}{\log\omega_{t_n}}} \quad (6)$$

$$f_n(x) = \sum_{t_n=1}^{T_n} W_{t_n} \cdot f_{t_n}(x) \quad (7)$$

where $f_n(x)$ is the push boat *SHP*function based on $\nabla^{1/3}/L$, L/B, d/B, F_r, D_p, η_D and σ.

The AdaBoost.RT algorithm was applied up to C_{T_n} times in order to form up to C_{T_n} NNE models. Each NNE modelwith the minimum $RMSE_n$ and MAE_n valueswas declared to be the best NNE model among all of the NNE modelsthat were trained on each dataset n. The$RMSE_n$ and MAE_n values are calculated from Eq. (8) and Eq. (9).

$$RMSE_n = \sqrt{\frac{1}{m}\sum_{i=1}^{m} (SHP_i - f_n(x_i))^2} \quad (8)$$

$$MAE_n = \frac{1}{m}\sum_{i=1}^{m} \left|(SHP_i - f_n(x_i)\right| \quad (9)$$

Calculations of the $RMSE_n$ and MAE_n values are performed over the target dataset.

There were a total of 9 potentially best NNE models, each created as a result of training the NNs and applying the AdaBoost.RT algorithm on every dataset n. The procedure for the creation of one NNE model per dataset n is presented in Fig. 5. [next page]

In an effort toreduce the number of iterations in the second stageand to speed up the procedure, the following rule was adopted: If the $RMSE_n$ and MAE_n values are equal to or lower than the threshold *RMSE* and *MAE* values, NN training is interrupted, and the best NNE model at this point is taken as the best NNE model overall and the procedure of getting the best NNE model is stopped. Threshold *RMSE* and *MAE* values were set just after the end of the first stage and they were 70 kW and 30 kW respectively. The pseudo-code of the entire procedure for getting one best NNE model is presented in Fig. 6. [next page]

NNE modelling was done separately by using iterations in which different datasets were used. As soon as an acceptable result had been reached by one of the previous datasets, training and testing of NNs were stopped and the NNE model was created.

This procedure represents an innovative alternative to present datamapping with NNs in waterway transportation. It is believed that this procedure will improve the results obtained by training only single NNs.

DATA EVALUATION

In the S_{NNE} setthere area possible 9 NNE models that are the best for each dataset n. Only those NNE models with *RMSE* and *MAE* values lower than 100 kW and 50 kW respectively will be analysed by graphical and numerical methods. Deviations and variations of push boat *SHPs* from a linear plot and the randomness and unpredictability of eachNNEmodel will be analysed by graphical methods, while the*RMSE*, *MAE* and *R*-squared valuesof each NNE model (R^2_n), presented in Eq. (12),will be the focus of the numerical method. In this way, anNNE model that may have lower *RMSE* and *MAE* values but higher variance of the data than some other NNE model in theS_{NNE} set will be compared to an NNE model with high *RMSE* and *MAE* values but lower variance of the data.

R-squared of each NNE model (R^2_n) measures how close the predicted *SHPs* are to the linear plots. It also indicates the variation of the data and is defined as the ratio of sum of squares regression (SSR), as calculated from Eq. (10) and Eq. (11) and the total sum of squares as in Eq. (12).

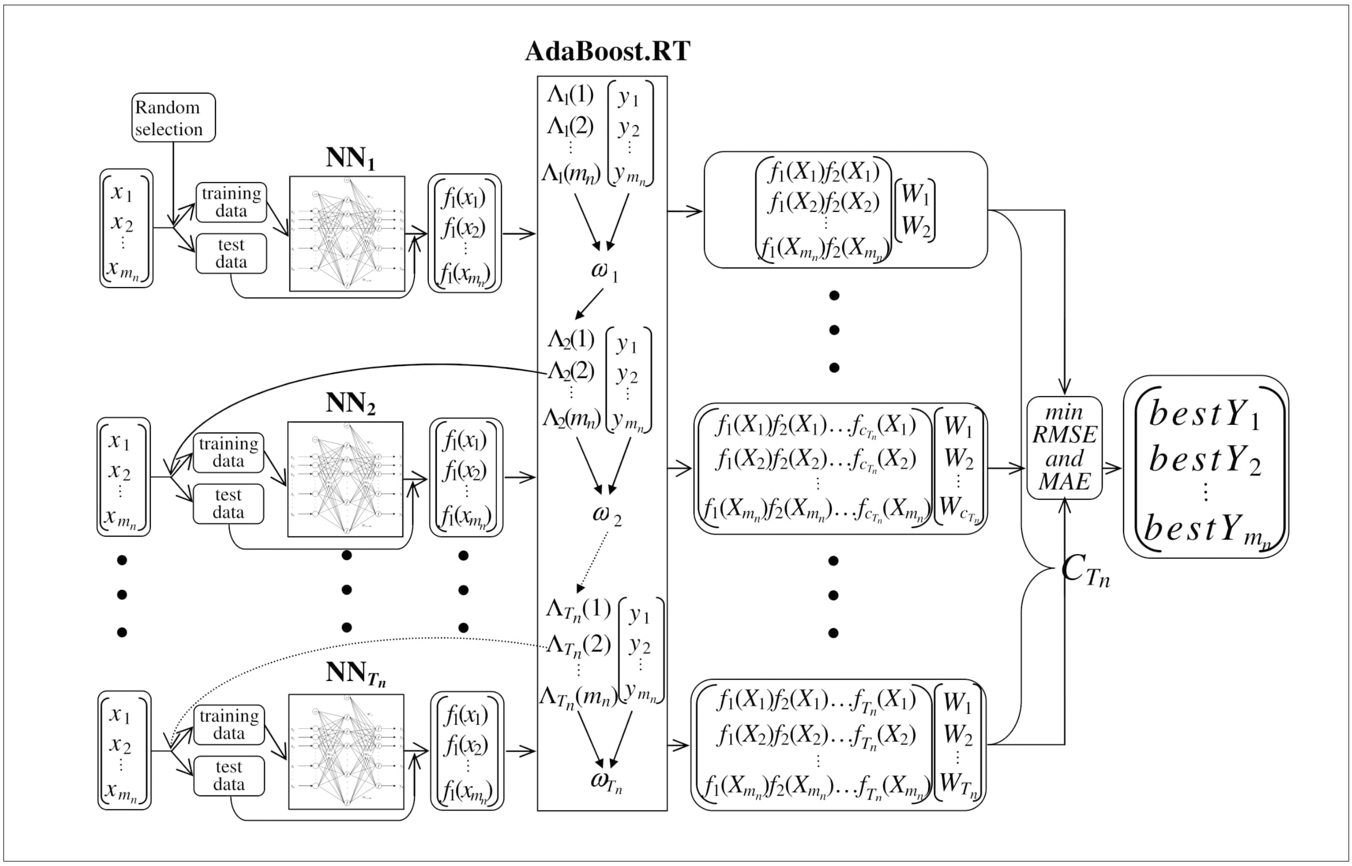

Fig. 5. Procedure for creation of one NNE model per dataset n

```
Set input parameter
Divide dataset to 9 (n = 1, 2, ..., 9) smaller sets with 10%, 20% ,..., 90% of data
CREATE empty set of NNE models S_NNE = {}
WHILE n ≤ 9
    S_Nh ∈ {2, 3, ..., (N_tr–1)/(N_I+2)| N_tr = N_tr_n, N_I = 7}
    SPLIT Dataset n into 3, 4, 5 and 10 FOLDS
    FOR K in FOLDS
        S_nK = {}
        FOR N_h in S_Nh
            DESIGN FF NN with N_h hidden neurons
            TRAIN NN 10 times
            CALCULATE average RMSE on training data
            DO Golden-section search procedure to get new N_h
            UPDATE S_nK with new N_h
            IF new N_h = N_h BREAK
            ELSE N_h = new N_h
        END LOOP
    END LOOP
    S_n = S_n3∩S_n4∩S_n5∩S_n10
    CREATE set P(S_n) of all subsets of set S_n
    SPLIT dataset n following 72/28 rule – 72% for training and 28% for test
    SET RMSE = 0.16 and MAE = 0.07
    FOR subset S_t_n in P(S_n)
        FOR N_h in subset S_t_n
            TRAIN NN with N_h neurons in hidden layer
            DO AdaBoost.RT algorithm
            DO new SPLIT of dataset n based on 72/28 rule and based on ARE(i_n) values
        END LOOP
        CREATE NNE_t_n model based on final weights
        CALCULATE RMSE(NNE_t_n) and MAE(NNE_t_n)
        IF RMSE(NNE_t_n) > RMSE(NNE_t_n-1) OR MAE(NNE_t_n) > MAE(NNE_t_n-1)
            NNE_t_n = NNE_t_n-1 CONTINUE
        ELSE CONTINUE
        IF RMSE(NNE_t_n) > 70kW OR MAE(NNE_t_n) > 30 kW
            CONTINUE
        ELSE
            BREAK
        UPDATE S_NNE
        BREAK
    END LOOP
    UPDATE S_NNE
END WHILE
RETURN best NNE model from S_NNE based on graphical and numerical methods
```

Fig. 6. Procedure for getting the best NNE model from the input and output data

$$SST_n = \sum_{i=1}^{m} \left[SHP_i - \frac{1}{m} \sum_{i=1}^{m} SHP_i \right]^2 \tag{10}$$

$$SSE_n = \sum_{i=1}^{m} (SHP_i - f_n(x_i))^2 \tag{11}$$

$$R_n^2 = \frac{SSR_n}{SST_n} = \frac{SST_n - SSE_n}{SST_n} \tag{12}$$

The graphical method includes comparison based on the prediction plots and residual plots per NNE models.

RESULTS AND DISCUSSION

In this section, the results from training the NNs and the application of the AdaBoost.RT algorithm are presented.

The computational results are presented and discussed in an effort to assess and analyse the efficiency of the best NNE model. The generalisation abilities of all the NNE models are assessed from two points of view: the graphical method, which includes the prediction plot and residual analysis, and numerical methods that include the *RMSE* value (see Eq. (8)),*MAE* value (see Eq. (9)) and the*R*-squared value (see Eq. (12)).

Eight *NNE* models were gathered in theS_{NNE} set. The *RMSE* and *MAE* values of each *NNE* model from S_{NNE} set are given in Tab. 2.

Tab. 2. RMSE and MAE values of all 8 NNE models

n	1	2	3	4	5	6	7	8
Number of *NNsin* each *NNE* model	2	2	3	2	2	3	4	2
NNE model (N_hs per n)	(8,9)	(8,10)	(9,10,12)	(13,15)	(14,16)	(16,17,18)	(16,17,18,20)	(19,22)
RMSE(kW)	233.19	180.31	154.08	103.17	99.80	88.53	89.45	45.22
MAE (kW)	148.42	107.05	91.96	57.94	55.50	46.08	34.69	24.95

From Tab. 2, it can be observed that last three NNE models (n = 6, 7, 8) can be subjects forcomparison by the graphical and numerical methods in this paper.

PREDICTION PLOT ANALYSIS

Prediction plots are used to show relationships between the actual push boat *SHPs* and NNE models *SHP* predictions. The desired relationships are assumed to be linear, meaning that the actual *SHPs* are equal to the predicted *SHPs*. They are presented with linear plots in each prediction plot as red lines. The predicted *SHP* values are along the *Y*-axes while the actual *SHP* values are along the *X*-axes. They are visualised as scatter plots. Each dot on the scatter plot represents one actual *SHP* along the *X*-axis and one predicted *SHP* along the*Y*-axis. The predictions from the entire target dataset are plotted and visualisedin each plot.

Fig. 7 representsthe prediction plots that refertothe last three NNE models (n = 6, 7, 8).

In each of the three prediction plots (see Fig. 7), there are dots that deviate more or less from the red lines. Going from Fig. 7 a) to Fig. 7 c), the deviations are smaller, pointing to the already assumed fact that NNE models created over a larger amount of data have smaller deviations. In the case of *SHP* predicton, the NNE model created over 80% of the data has the smallest deviations among all NNE models. The lowest variance of the the data along the desired response red line in Fig. 7 c) indicates that the last NNE model (n = 8) has the highest accuracy of all the NNE models. The variations amongthe data are also the smallest in the NNE model created over 80% of the data. In fact, as the percentage of data for creation of the NNE model rises, the variation between the data reduces. Consequently, the last NNE model has the highest precision of all the NNE models.

RESIDUAL ANALYSIS

Residual plots validate the randomness and unpredictability of the NNE models in this paper. A residual indicates the difference between the actual *SHP* and predicted *SHP*. Fig. 8 illustratesthe residual plots of the last three NNE models (n = 6, 7, 8).

From Fig. 8 [next page], it can be seen thatthe dots, in general, are clustered around the lower single digits of the residual axis. As n increases, the accumulations around the red lines increase. This suggests that the NNE model created over 80% of the data has the lowest residuals among all three analysed. None of the NNE models is symmetrically distributed and they all have outliers, which means that all three NNE models have room for improvement in terms of residuals. Indeed, a specific curve that could fit the residuals cannot be found in Fig. 8 c).The last positive thing about each NNE model is that they do not have clear patterns. From the observation of the residual graphs, it can be concluded that all three residual plots are stochastic because the points are random and unpredictable due to the lack of any observed specific curve, but the residuals are the smallest in Fig. 8 c). As a conclusion, the best fit for the estimation when it is evaluated by residual analysis is the last NNE model.

NUMERICAL EVALUATION METHODS

A summary of the analysis by numerical evaluation methods is presented in Tables2 and 3.

From Tab. 3 [next page], it can be observed that *R*-squared analysis for the n = 4, 5, 6, 7 and 8 NNE models has a good fit with values more than or equal to 0.92. However, if the evaluation is examined further, the last NNE model (n = 8) has a value of 0.98, which is better than the other NNE models.

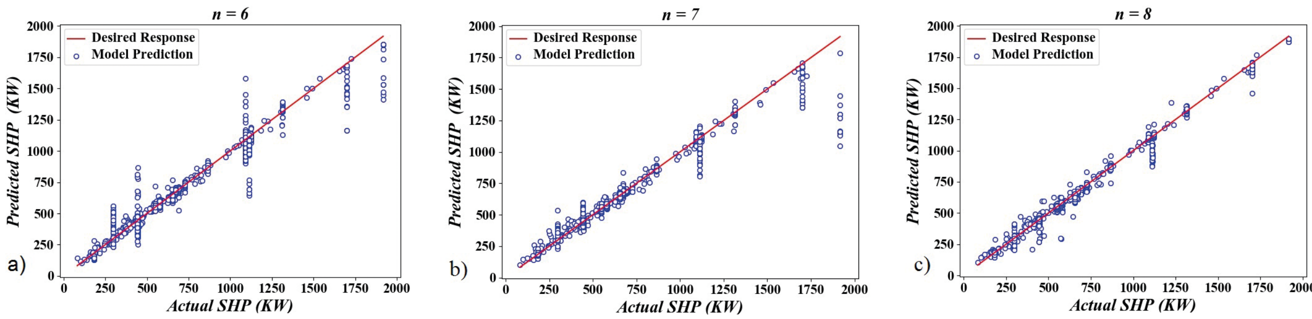

Fig. 7. Prediction plots of pushboat SHP estimation using NNE model created by training NNs over 60%, 70% and 80% of data

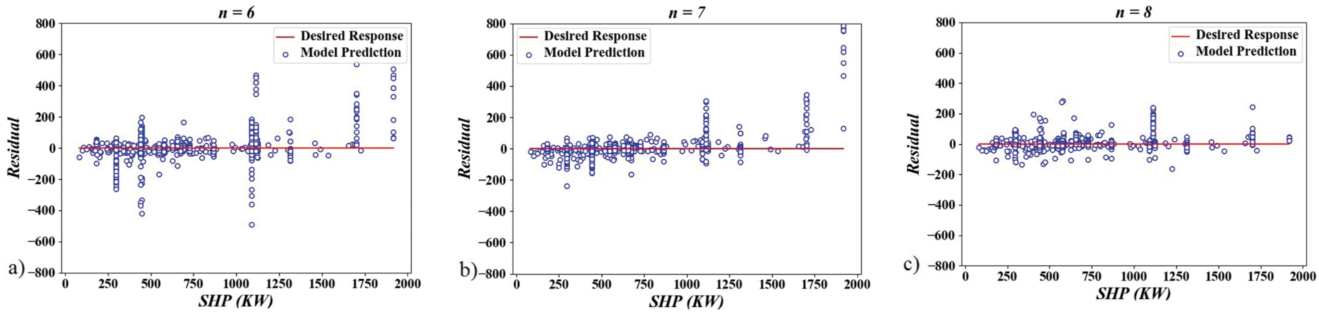

Fig. 8. Residual plots of pushboat SHP estimation using NNE model created by training NNs over 60%, 70% and 80% of data

Tab.3. R-squared values of each NNE model

n	1	2	3	4	5	6	7	8
R-squared	0.58	0.75	0.82	0.92	0.92	0.94	0.94	0.98

For the MAE evaluation (see Tab. 2), the last NNE model (n = 8) has the lowest error of 24.95 kW. This was followed by the other NNE models, among which the first five models didnot satisfy the predefined *MAE* requirements.

The lowest *RMSE* value of 45.22 kW (see Tab. 2)was found in the last NNE model (n = 8). The second lowest *RMSE* value of 88.53 kW was from the NNE model created over 60% of the data from the target dataset, while the third lowest *RMSE* value of 89.45 kW was from the NNE model created over 70% of the data from the target dataset. The other NNE models did not come into analysis as their *RMSE* values were above 100 kW except for the fifth NNE model, whose value of RMSE is 99.80 kW but whose MAE value is higher than 50 kW and which therefore doesnot satisfythe predefined conditions for analysis.

From the analysis performed by graphical and numerical methods, the last NNE model (n = 8) is selected for prediction of pushboat *SHPs* due to having the lowest *RMSE*, *MAE* and *R*-squared values of all the NNE models as well as the lowest variations of predicted *SHPs*and lowest residuals.

However, after a careful analysis of the best NNE model, caution regarding some of the predicted *SHPs* should be taken, particularly considering the data pointsat which the *MAEs* are above 200 kW (see Fig. 8 c). This is the major weakness of the best NNE model.

The other weaknesses refer to the *RMSE* and *MAE* values if these are compared to the *RMSE* and *MAE* valuesof similar NNE models. One such NNE model was found by [11] with *RMSE* and *MAE* values of 20.45 kW and 9.94 kW respectively. Although [11] found lower *RMSE* and *MAE* values than thevalues in this paper, it can be concluded that the predicted *SHPs* from the best NNE model in this paper have reasonable accuracy if the number of data and push boat operation only at full power are taken into account.

CONCLUSION AND FURTHER WORK

In this paper, Neural Network Ensembling with the AdaBoost.RT algorithm for the prediction of push boat *SHP* is presented to improve the operational efficiency of IWT companies and to help dispatchers in theirdaily decision-making processes. Full-scale speed/power trials are arranged in a dataset from which the target dataset is derived. The target dataset is used for training of NNs which are then ensembled in one NNE model with the AdaBoost.RT algorithm.

Eight different NNE models were created based on division of the target datasetinto 8 smaller datasets. Three models were chosen for analysis by graphical and numerical methods. On the basis of variations of *SHPs* in the prediction plots and residual plots and *RMSE*, *MAE* and *R*-squared values, a NNE model created using 80% of the data was declared as the best NNE model.

The model has two major weaknesses, however. One refers to a couple of data pointerrors representing deviations from the predicted *SHPs*, while the other refers to the comparison of the *RMSE* and *MAE* values to the same values obtained by the model found in [11]. Both weaknesses suggest that the limitation of the *RMSE* and *MAE* values might be decreasedbelow the already defined threshold values.

By using the NNE model, effects on the IWT like punctual arrivals of cargo, costs of transport and profit of IWT companies could be investigated. Indirect benefits to transport such asincreased environmental protection, better energy efficiency and better utilisation of barges and push boats should emerge also from further investigation. Overall, it improves customer satisfaction, and eliminates the movement of empty transport vehicles, for example.

The model and solution methodology given in this paper could be a very useful practical tool for dispatchers to make the right decisions about the assignment of push boats to barge convoys. They can use the best NNE model as a decision support tool to solve their daily assignment tasks, test different solutions related to the planning, routing and scheduling of pushed convoys, and choose assignments which are suitable for their own needs at any given moment. Therefore, the part ofthe decision-making related to correct assigning according to the push boat *SHP* in the IWT company could be improved by applying the proposed decision support tool.

Further research should be undertaken in the following two directions: additional data and new methodologies.

Additional data from the full-scale speed/power trials could be incoroporated into the target dataset. These data would refer not only to pushed convoy speedsat full push boat power, but also to any pushed convoy speed. Overall, this would contribute to better knowledge of pushed convoy arrival times.

New methodologies include machine learning algorithms. Important factors like the complex and finite datasets that are contained in this paper are a good starting point for the application of machine learning algorithms.

ACKNOWLEDGEMENTS

This research has been partially supported by the Serbian Ministry of Science and Technological Development, Grants No. TR36024 and No. TR36027.

REFERENCES

1. Alfandari L., Davidovic T., Furini F., Ljubic I., Maras V., Martin S. (2019):*Tighter MIP models for Barge Container Ship Routing*. Omega, 82,38–54.

2. Botta M. (2001): Resampling vs reweighting in boosting a relational weak learner. *In Proceedings of the 7th Congress of the Italian Association for Artificial Intelligence on Advances in Artificial Intelligence*,Springer-Verlag,London, UK, p. 70–80.

3. Burrill L. C. (1943):*Developments in Propeller Design and Manufacture for Merchant Ships*. Transactions, Institute of Marine Engineers, London, Vol. 55. p. 106–136.

4. Colic V. (2006):*Research of navigational, technical, energetic and propulsive characteristics of Danube towboats*. Faculty of Transport and Traffic Engineering (in Serbian), University of Belgrade, Belgrade, p 350.

5. Couser P.R., Mason A. (2004): Artificial Neural Networks for Hull Resistance Prediction. *Computer Applications and Information Technology in the Maritime Industries (COMPIT'04)*, 9-12 May, Siguenza, Spain.

6. Drucker H. (1997): Improving regressors using boosting techniques. *Proceedings of the Fourteenth International Conference on Machine Learning*, 107–115.

7. Efron B., Tibshirani R. (1993):*An Introduction to the Bootstrap*. Chapman & Hall, New York, p.456.

8. Kohavi R. (1995): A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann. Vol. 2 (12), 1137–1143.

9. Maras V., Lazic J., Davidovic T., Mladenovic T.N. (2013):*Routing of barge container ships by mixed-integer programming*. Appl. Soft Comput., 13, 3515–3528.

10. Parks A.I., Sobey A.J., Hudson D.A. (2018):*Physics-based shaft power prediction for large merchant ships using neural networks*. Ocean Engineering, 166, 92–104.

11. Radonjic A., Vukadinovic K.(2015):*Application of Ensemble Neural Networks to Prediction of Towboat Shaft Power*.J. Mar. Sci. Technol., 20, 64–80.

12. Reich Y., Berai S.V. (2000):*A methodology for building neural networks model from empirical engineering data.*Engineering Applications of Artificial Intelligence, 13(6), 685–694.

13. Ren L., Zhao Z. (2002):*An optimal neural network and concrete strength modeling*. Advances in Engineering Software, 33(3), 117–130.

14. Riedmiller M., Braun H. (1993): A direct adaptive method for faster backpropagation learning: The RPROP algorithm. *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco, CA, USA, 28 March-1 April 1993,pp. 586–591.

15. Schwenk H., Bengio Y. (2000):*Boosting Neural Networks*. Neural Computation, 12(8), 1869-1887.

16. *Seiffer C., Khoshgoftaar T.M., Van Hulse J., Napolitano A. (2008*): Resampling or Reweighting: A Comparison of Boosting Implementations. *20th IEEE International Conference on Tools with Artificial Intelligence*, p. 445–451.

17. Solomatine D.P., Shrestha D.L. (2004): AdaBoost.RT: a boosting algorithm for regression problems.*Proceedings of International Joint Conference on Neural Network*, Vol. 2, 1163–1168.

18. Tupper E.C., Rawson K.J. (2001):*Basic Ship Theory*. 5th Edition, Elsevier, p.784.

19. Vaganov G.I., Voronin V.F., Shanchurova V.K. (1986):*Tyaga Sudov: myehtodeka i premyehreh vihpolnyehneya sudovih tyagovih raschyehtov (Ship propulsion: methodology and examples of the ship propulsion calculations)*.Transport, Moscow, p. 201 (In Russian).

20. Vukadinovic K., Teodorovic D., Pavkovic G. (1997):*A neural network approach to the vessel dispatching problem*. Eur. J. Oper. Res., 102, 473–487.

21. Xu Y., Goodacre R. (2018):*On Splitting Training and Validation Set: A Comparative Study of Cross –Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning*. Journal of Analysis and Testing, 2, 249 262.

22. Young Y. L. (2002):*Numerical Modeling of Supercavitating and Surface-Piercing Propellers*. PhD thesis, Environmental

and Water Resources Engineering,Department of Civil Engineering, University of Texas at Austin, Austin, USA.

23. Zhang G., Eddy Patuwo B., Hu M. Y. (1998):*Forecasting with artificial neural networks: The state of the art*.International Journal of Forecasting, 14, 35–62.

CONTACT WITH THE AUTHORS

Aleksandar Radonjić
e-mail: a.radonjic@sf.bg.ac.rs

University of Belgrade
Faculty of Transport and Traffic Engineering
Vojvode Stepe 305, 11000 Belgrade
Republic of Serbia

Danijela Pjevčević
e-mail: danijela@sf.bg.ac.rs

University of Belgrade
Faculty of Transport and Traffic Engineering
Vojvode Stepe 305, 11000 Belgrade
Republic of Serbia

Vladislav Maraš
e-mail: v.maras@sf.bg.ac.rs

University of Belgrade
Faculty of Transport and Traffic Engineering
Vojvode Stepe 305, 11000 Belgrade
Republic of Serbia