Application of NARX neural network for predicting marine engine performance parameters

Y. Raptodimos* and I. Lazakis

Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, Glasgow, United Kingdom

* Corresponding author details:

Yiannis Raptodimos

Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde

Henry Dyer Building, 100 Montrose Street, Glasgow G4 0LZ, Scotland UK

yiannis.raptodimos@strath.ac.uk

Other author details:

Iraklis Lazakis

Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde

Henry Dyer Building, 100 Montrose Street, Glasgow G4 0LZ, Scotland UK

iraklis.lazakis@strath.ac.uk

Application of NARX neural network for predicting marine engine performance parameters

Y. Raptodimos & I. Lazakis
Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, Glasgow G4 0LZ, United Kingdom

Abstract

Though the maritime industry is still predominantly reliant on a time-based, prescriptive approach to maintenance, the increasing complexity of shipboard systems, heightened expectation and competitive requirements as to ship availability and efficiency and the influence of the data revolution on vessel operations, favour a properly structured Condition Based Maintenance (CBM) regime. In this respect Artificial Neural Networks (ANNs) can be applied for predictive maintenance strategies that can assist decision makers in selecting appropriate maintenance actions for critical ship machinery. This paper focuses on developing a Nonlinear Autoregressive with Exogenous Input (NARX) ANN model for forecasting future values of performance parameters of a marine main engine. Moreover, a detailed sensitivity analysis is conducted to examine the performance and robustness of the developed NARX model, based on the dataset applied for its training. Through the NARX model, a predictive monitoring approach can be achieved for ship machinery monitoring as high forecasting accuracy was achieved for the case study of a main engine cylinder exhaust gas outlet temperature. The sensitivity analysis overall demonstrated the good performance and generalisation of the NARX model as it successfully considers different values of the time series data for conducting the one-step-ahead output.

Keywords: ANN; NARX; time series analysis; ship condition monitoring; marine engine;

Nomenclature

AI	Artificial Intelligence
ANN	Artificial Neural Network
APE	Absolute Percentage Error
BS	British Standards
CBM	Condition Based Maintenance
DWT	Deadweight Tonnage
INCASS	Inspection Capabilities for Enhanced Ship Safety
ISEMMS	Integrated Ship Energy & Maintenance Management System
MAPE	Mean Absolute Percentage Error
MCR	Maximum Continuous Rating
MINOAS	Marine Inspection Robotic Assistant System

MSE	Mean Square Error
MUNIN	Maritime Unmanned Navigation through Intelligence in Networks
NAR	Nonlinear Autoregressive
NARX	Nonlinear Autoregressive with Exogenous Input
RNN	Recurrent Neural Network
SOM	Self-Organising Map

1. Introduction

1.1 Maintenance and CBM

The marine industry is responsible for the transportation of the vast majority of the merchandise worldwide, as over 80% of global trade by volume is carried on board ships, emphasising the importance of maritime transport for trade and development (UNCTAD 2017). In this context and considering the emerging trends currently shaping the outlook for seaborne cargo flows in combination with technological advancements, the safe, efficient and environmentally friendly operation of ships is extremely important. In this respect, maintenance is an important contributor to reach the intended life-time of technical capital assets and is defined as a combination of all the technical and associated administrative activities required to keep equipment, installations and other physical assets in the desired operating condition or to restore them to this condition (BS, 1993). Furthermore, routine and periodic maintenance accounts for approximately 20% of a ship's operational expenses (Stopford 2009).

The importance of maintenance is also demonstrated by the fact that it is the only shipboard activity to have one whole element assigned to it in the ISM code (IMO 1993). Moreover, technological advances, overburdened crew and high cost of ownership, have resulted in considerable interest in advanced maintenance techniques (INCASS 2015a). Though the industry is still predominantly reliant on a time-based, prescriptive approach to maintenance, the increasing complexity of shipboard systems, heightened expectation and competitive requirements as to ship availability and efficiency and the influence of the data revolution on vessel operations, favour a properly structured Condition Based Maintenance (CBM) regime (Tinsley 2016).

Unlike breakdown and preventive maintenance, CBM focuses on not only fault detection and diagnostics of components but also degradation monitoring and failure prediction (Prajapati et al. 2012). The heart of CBM is condition monitoring which aims in collecting data regarding equipment conditions and is applied through various tools by recording and evaluating different measurable parameters. Data can include vibration, acoustic, temperature, oil and lubricant and current signal measurements. Sullivan et al. (2010) and Pascual (2015) describe such condition monitoring technologies and techniques in depth.

It is only recently that new approaches investigating the enhancement of ship's reliability, availability and profitability have been considered according to Lazakis and Olcer (2015). The outcome of this study indicated that preventive maintenance is still the preferred approach by ship operators, closely followed by predictive maintenance; hence, avoiding the ship corrective maintenance framework and increasing overall ship reliability and availability. Moreover, shifting from scheduled, rule-based maintenance to a data-driven approach can lead to more accurate and timely maintenance, resulting in lower costs, greater availability of ship systems and increased safety. Consequently, the maritime industry is seeking for increased reliability, maximum uptime and optimal operational efficiency, as well as ensuring safe and sustainable environmental performance in harsh environments.

1.2 Artificial Neural Networks

Artificial Neural Networks (ANNs) can readily address modelling problems that are analytically difficult and for which conventional approaches are not practical, including complex physical processes having nonlinear, high-order and time-varying dynamics. Therefore, they can be applied for predictive maintenance strategies that can assist decision makers in selecting appropriate maintenance actions for critical ship machinery. Moreover, Raza and Liyanage (2009) also stated that there has been an increasing demand for testing and implementing intelligent techniques as a subsidiary to existing condition monitoring programs and that artificial neural networks have emerged as one of the most promising techniques in this regard.

According to Nasr et al. (2012) ANNs provide an effective analysing and diagnosing tool to understand and simulate the nonlinear behaviour of complex systems. As more data describing the system condition and its influencing parameters become available, data-based methods are being increasingly applied in the field of fault detection (Tan et al. 2012), fault diagnostics (Tamilselvan and Wang 2013) and for predicting the residual useful life (Tian et al. 2010).

Opposed to the traditional model-based methods, ANNs are data-driven and self-adaptive methods, meaning that there are few a priori assumptions about the models under study. They learn from past examples and capture subtle functional relationships among the data even if the underlying relationships are hard to describe or unknown. Secondly, ANNs have good generalisation capabilities. After being trained on the data presented to them, ANNs have the ability to correctly simulate data they have not seen before. Thirdly, ANNs are universal functional approximators and have more general and flexible functional forms than the traditional analytical and statistical methods.

Adjallah et al. (2007) applied ANN for classifying bearing faults based on condition monitoring of bearing acceleration signals in order to ease the burden of decision making on system and equipment integrity. In addition, Zhu (2009) utilised ANN to diagnose faults of a marine diesel engine related to variations in valve clearance and engine cylinder loads. Their research indicated that ANN has a high degree of accuracy when predicting ship main engine faults and can improve the reliability of the engine overall. Furthermore, Zhou and Xu (2010) applied neural networks for the fault diagnosis of a marine engine cooling system by using failure modes as input to the network and failure causes as output based on simulated data.

Basurko and Uriondo (2015) applied ANNs for CBM of medium speed diesel engines in operation for the case study of a fishing vessel. The developed ANN analyses actual monitored data to determine engine faulty conditions. Moreover, Noor et al. (2016) applied ANN modelling on a marine diesel engine in order to predict its performance in terms of output torque, brake power, brake specific fuel consumption and exhaust gas temperature using as input data various engine speeds and loads. Results showed that the prediction error of the ANN model was lower than the mathematical model. Raptodimos and Lazakis (2018) presented the application of the Self-Organising Map (SOM) neural network for data clustering applications and monitoring of a marine diesel engine, utilising the unsupervised nature of the SOM algorithm. Moreover, Cipollini et al. (2018) investigated data-driven models for performing CBM on a ship propulsion system. The results confirmed the possibility to implement regression techniques for CBM and that amongst the machine learning models tested, ANN models generally showed the best performance.

Although ANNs have recently gained importance in time series applications (Aizenberg et al. 2016, Laboissiere et al. 2015, Liu et al. 2015, Szoplik 2015), there is a gap in the literature regarding such ANN applications in the marine industry which could assist in reporting future health status of systems and components and provide users with the earliest warning of potential faults. Aizenberg et al. (2016) performed time series analysis using multilayer neural network for forecasting oil production in the Gulf of Mexico. They concluded that the choice of embedding dimensions from time series data is a challenging and ongoing task requiring additional research effort.

Bearing the above in mind, this paper focuses on developing a Nonlinear Autoregressive with Exogenous Input (NARX) ANN model for forecasting future performance parameter values of a marine main engine that can be utilised for predictive maintenance strategies. Moreover, a detailed sensitivity analysis is conducted to examine the performance and robustness of the developed NARX model, based on specific marine main engine components. A NAR model is also presented for comparison purposes. Through the NARX model, a predictive monitoring approach can be achieved for ship machinery monitoring, that assists in preventing failures and issues advanced warnings of potential faults leading to appropriate maintenance actions and suggestions.

The present research paper is organised as follows: Section 2 presents the suggested methodology while Section 3 presents the case study application through which the methodology is applied and associated results. Finally, the discussion and conclusion of this research study is presented in Section 4 and 5 respectively.

2 Materials and methods

2.1 NARX methodology

A time series is a sequential set of data points, measured typically over successive points in time spaced at uniform time intervals. It is mathematically defined as a set of vectors $y(t)$, $t = 0, 1, 2, \ldots, d$ where t represents the time elapsed with a set of discrete values $y_1, y_2, y_3, \ldots,$ *etc.* The variable $y(t)$ is treated as a random variable and the measurements taken during an event in a time series are arranged in a proper chronological order. In the NARX model, future values of a time series $y(t)$ are predicted from past values of $y(t)$ and another external series $x(t)$. Therefore, compared to the Nonlinear Autoregressive (NAR) model, NARX can consider external (exogenous) input for predicting the time series $y(t)$ and detecting changes in model parameters due to external conditions as seen in equation 1.

$$y(t)=f(x(t-1), \ldots, x(t-d), y(t-1), \ldots, y(t-d)) \quad (1)$$

where $x(t)$ is the observation of the exogenous input at time t

Although the NARX network is applied for short-term forecasting, multi-step-ahead predictions can be acquired if knowledge of the future exogenous inputs is known. This is done by using the output of a one-step ahead prediction as the input for the subsequent prediction in an iterative process as described below and presented in Figure 1.

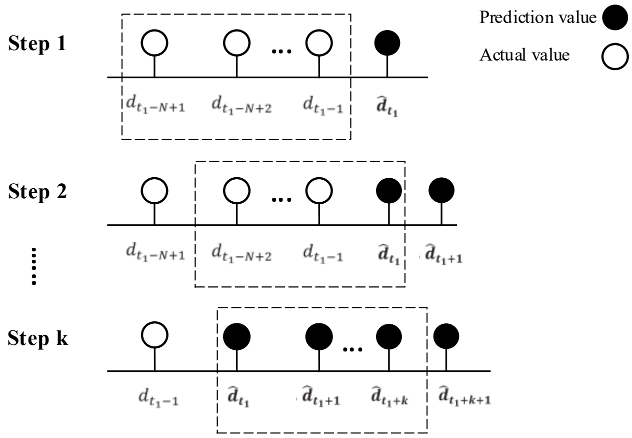

Figure 1: Procedure for multi-step-ahead NARX predictions

Data are normalised for direct use in network training by transforming them to the network's operating range, shaped to meet the requirements of the network input layer and are adapted to the nonlinearities of the neurons, so that their outputs should not cross the saturation limits (Maier and Dandy 2000). The data is pre-processed in the ANN models by mapping data to a matrix row with minimum and maximum values from -1 to 1 for conducting proper analysis and improving efficiency of the network training. Additionally, the time series data is prepared by shifting time by the minimum amount to fill input states and layer states for network open loop and closed loop feedback modes. This allows the original time series data to remain unchanged, easily customising it for networks with different numbers of delays. Tapped delay lines are used to store previous values of the *x(t)* and *y(t)* sequences. Figure 2 presents the NARX model utilised in this paper.

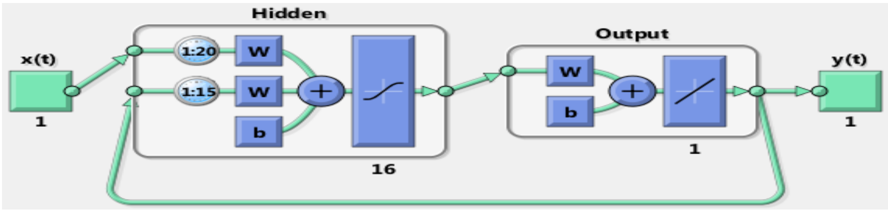

Figure 2: Representation of NARX model

The NARX model is initially designed as feedforward backpropagation networks. During training, the network weights and biases are updated after all the inputs and target values have been presented to the network. The networks are autoregressive as the only inputs are lagged target values and lagged external input values in the case of the NARX model. The Bayesian regularisation backpropagation algorithm is primary used for training the NARX network. This algorithm updates the weight and bias values according to Levenberg-Marquardt optimisation. It minimises a combination of squared errors and weights and then determines the correct combination to produce a network which generalises well and avoids overtraining.

The performance of the network is evaluated using the *Mean Square Error* (*MSE*) average sum of square errors and *Correlation Coefficient (R)*. The ANN is then trained based on the above parameters in open loop, as a feedforward network. The trained ANN can then be converted to closed loop mode and the data is reformatted to simulate the network's closed loop response in order to carry out multi-step-ahead predictions. The output *y(t)* is fed back to the input of the network. The trained ANN is converted to closed loop by replacing the feedback input and creating a feedback connection from the network output to the network input, thus making the network a Recurrent Neural Network (RNN) as presented in Figure 2.

The network architecture and training parameters are selected to minimise prediction errors. Furthermore, the forecasting results obtained are evaluated through defined performance measures (Aladag 2017); the Absolute Percentage Error (APE) and Mean Absolute Percentage Error (MAPE). The accuracy of each forecast is expressed as the APE which is calculated as:

$$APE = \left|\frac{(y(t) - \hat{y}(t))}{y(t)}\right| \times 100\% \tag{2}$$

MAPE represents the percentage of average absolute error of forecasted values from original ones, showing the magnitude of overall error occurring due to forecasting and it is calculated as follows:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} APE \tag{3}$$

where *y(t)* is the actual observation, $\hat{y}(t)$ is the forecasted observation for the same time interval *t* and *N* in the number of data points.

2.2 Measurement uncertainty-calculation of prediction intervals

Sensor measurements are often distorted by noise and bias, thereby masking the true condition of the system which can lead to incorrect estimation results (Bocaniala et al. 2006). The prediction intervals are defined as:

$$Prediction\ Intervals = \widehat{y_{pred}} \pm t_{1-\frac{a}{2},n-2} \times SE_{\widehat{y_{pred}}} \tag{4}$$

where $\widehat{y_{pred}}$ is the predicted point, *t* is the value of the two tailed *"t"* distribution for a probability equal to *a*, *n* is the number of samples in the data and $SE_{\widehat{y_{pred}}}$ is the standard error of the prediction (standard deviation of the sampled population). Given a dataset *x*, the standard error of the prediction is calculated through equation 5:

$$SE_{\widehat{y_{pred}}} = s\sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{{S_x}^2(n-1)}} \tag{5}$$

where *s* is the standard error, S_x is the standard deviation of *x* and $\bar{x}$ is the sample average of *x*.

Therefore, a prediction interval forecast consists of an upper and lower limit between which a future value is expected to lie with a prescribed probability. Furthermore, it is important to mention that no generally accepted method of calculating prediction intervals exist (Chatfield 2000).

2.3 Sensitivity analysis methodology

Saltelli et al. (2008) defines sensitivity analysis as the field examining how uncertainty in model outputs can be apportioned, qualitatively or quantitatively, to different sources of uncertainty in the model input. The Monte Carlo method is used as a comprehensive approach for accomplishing the sensitivity analysis (Wei et al. 2013). In the context of Monte Carlo simulation, this is described as the process of approximating the model output through repetitive random application of the model's algorithm. Monte Carlo sensitivity analysis is performed with Latin hypercube sampling which is a sampling scheme designed to ensure that the upper and lower ends of the data used in the sensitivity analysis are well included in the analysis (Firestone et al. 1997). Furthermore, this sampling method is generally recommended over simple random sampling and is one of the most widely used random sampling methods for Monte Carlo based analysis (Shields and Zhang 2016). Figure 3 provides a graphical outline of the sensitivity analysis regarding the inputs and output for the one-step-ahead prediction of the NARX model.

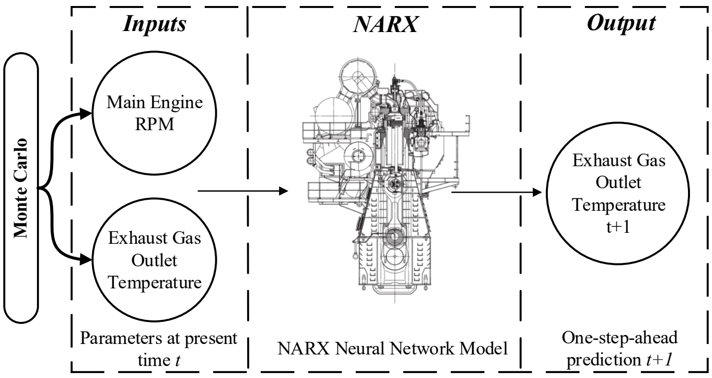

Figure 3: Graphical outline of NARX model inputs and output for sensitivity analysis

The sensitivity analysis is conducted by varying the inputs of the NARX model, the main engine rpm and/or cylinder exhaust gas outlet temperature at timestep *t* representing the present time, used for producing the subsequent one-step-ahead prediction *t+1* output. Based on the above, both the main engine rpm and exhaust gas outlet temperature parameters are varied to examine the output of the NARX model.

3 Case study

3.1 Case study description

The methodology is applied on a two-stroke, eight-cylinder marine diesel engine of a Panamax container ship. The main particulars of the Panamax ship case study are demonstrated in Table 1 below.

Table 1: Case study main characteristics

Principal characteristics	
Year built	2009
Ship type	Cellular container
DWT (summer)	50829 *tons*

Length overall	260.00 *m*
Beam	32.00 *m*
Depth moulded	19.30 *m*
Draft (summer)	12.60 *m*
Engine particulars	
Main engine	HSD MAN B&W 8K90MC-C
Maximum Continuous Rating (MCR)	49680 *BHP* @ 104 *RPM*
Number of cylinders	8
Bore	900 *mm*
Stroke	2300 *mm*
Exhaust gas temperature sensor characteristics	
Sensor type	Type K (NiCr-Ni) Thermocouple
Measuring range	0-800 ^{o}C
Specified accuracy	± 1%
Measurement frequency	Hourly

The NARX model developed forecasts the upcoming 20 hourly values of the exhaust gas outlet temperature of cylinder 5 of the main engine. A dataset consisting of 920 hourly measurements is utilised. Data was collected by utilising existing sensors installed in the various systems and components of the main engine. The measurements for the main engine performance parameter was recorded hourly (one measurement per hour) as stated in Table 1 and saved on the vessel's onboard data acquisition system. The specified sensor accuracy of ± 1% is assumed over the entire sensor measurement range.

The data acquisition system is connected to the main engine sensors and other machinery sensors in the engine room. Measurements from the various sensors are set to be recorded on an hourly basis from the data acquisition system and are stored on a platform developed by the shipping company. Overall, the data stored in the platform is pre-processed by the data acquisition system in order to remove and rectify potential faulty sensor readings. As such, the data used for the cylinder exhaust gas outlet temperature has been pre-processed and further data pre-processing is considered out of scope for this research. The stored data is accessible from the shipping company's platform to crew and onshore personnel and data logs can be extracted in a Microsoft Excel spreadsheet format for the various monitored machinery equipment with the respective hourly timestamps.

The first 900 hourly measurements are used for training the NARX model while the last 20 measurements are used to validate the model accuracy through the APE and MAPE criteria. The exhaust gas outlet temperature parameter was selected for the case study as the results of the NARX model developed for the cylinder exhaust gas outlet temperature parameter varies significantly for different main engine rpm speeds compared to the other main engine modelled parameters. Moreover, the exhaust gas temperature is directly emitted from the engine cylinders and therefore will indicate the operation and condition of the engine and its combustion process and is a valuable source of diagnostic information regarding the technical condition of elements such as the cylinder and piston, scavenging air, fuel supply system amongst other (Taylor 1996). The main engine rpm is utilised as the exogenous input in the NARX model. For the main engine, the performance parameters largely depend on the main engine rpm. Thus, the integration of knowledge or data related to the engine rpm speed is applied as exogenous input in the NARX models. Table 2 presents the training parameters applied for training the NARX model.

Table 2: NARX training parameters

Training conditions	**Value**	**Description**
net.trainParam.epochs	1000	Maximum number of epochs
net.trainParam.goal	0	Performance goal
net.trainParam.lr	0.01	Learning rate
net.trainParam.max_fail	6	Maximum validation failures
net.trainParam.min_grad	1E-05	Minimum performance gradient
net.trainParam.time	inf	Maximum time to train (seconds)
net.trainFcn	trainbr	Training algorithm
net.performFcn	mse	Performance function
net.divideFcn	divideblock	Dataset division
net.divideParam	70%-0%-30%	Training, validation, test set ratio

Training stops when any of the following conditions are met; The maximum number of epochs (iterations) for training is reached, performance is minimised to the goal, the maximum amount of time is exceeded, the validation performance has increased more than the maximum number of validation failures since the last time it decreased. In terms of the training algorithm, the Bayesian regularisation backpropagation algorithm was selected for network training that uses Jacobian derivatives as it updates the weight and bias values according to Levenberg-Marquardt optimisation. This algorithm minimises a combination of squared errors and weights in the network and then determines the correct combination to produce a network which generalises well (Okut 2016). Moreover, this training algorithm includes the validation data in the training set, hence the dataset is split 70% for training and 30% for testing in a time sequence manner. This is done compared to splitting the dataset randomly in order to preserve the correlation relationships of the time series data.

The number of delays is set experimentally. Experimental runs with different number of network feedback delays were performed to obtain an accurate prediction model that performs well. The developed NARX model consists of 16 neurons in the hidden layer and uses a hyberbolic tangent function in the hidden layer and a linear transfer function in the output layer. Moreover, 15 feedback delays of the target time series *y(t)* and 20 input delays of the external input *x(t)* which is the main engine rpm are applied as presented in Figure 2. Once training is complete, the network is converted into closed loop form and its own predictions *y(t)* in the output layer become the feedback inputs for carrying out one-step-ahead or multi-step-ahead predictions alongside the external input *x(t)* time series.

3.2 NARX results

The error autocorrelation function is used to validate the NARX network performance. This function describes how the prediction errors are related in time. For a faultless prediction network model, there should be only one non-zero value of the autocorrelation function occurring at zero lag implying that the forecast errors are entirely uncorrelated with each other. Figure 4 presents the error autocorrelation function results for exhaust gas outlet temperature of main engine cylinder 5.

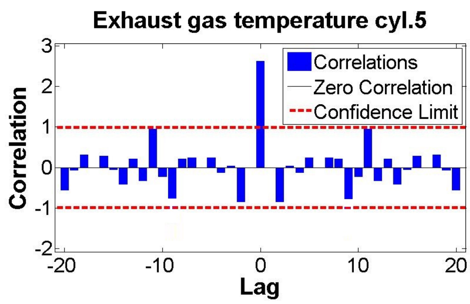

Figure 4: Error autocorrelation function

Satisfactory network training has been achieved as the forecast errors are completely uncorrelated with each other and fall within the 95% confidence limits around zero which are calculated based on the sample size of the time series data generated from the MATLAB autocorrelation function. This implies that the prediction errors are completely uncorrelated with each other. Figure 5 presents the correlation coefficient R results for the training and test datasets.

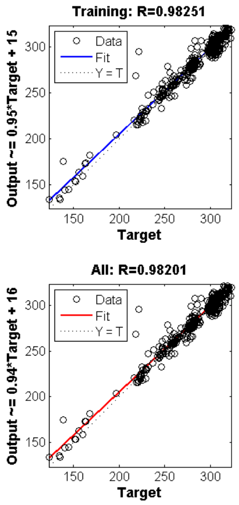

Figure 5: NARX training and test sets R results

Figure 6 presents the resulted forecasted results with their 95% prediction intervals for the exhaust gas outlet temperature.

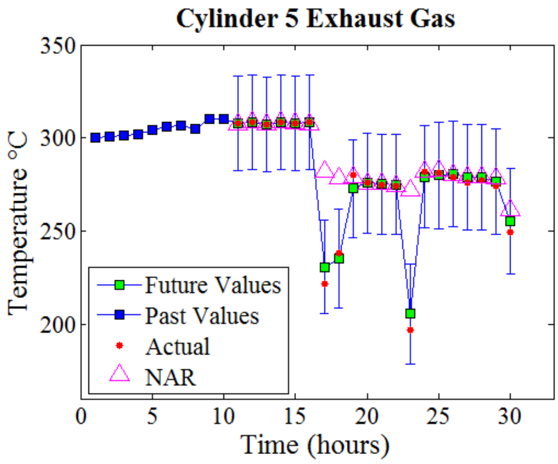

Figure 6: NARX forecast results for exhaust gas outlet temperature of cylinder 5 and comparison with NAR results

The forecasted NARX results are plotted in green for timesteps 11-30 hours (20 forecasted values) while the actual values are plotted as red dots on the graph. Moreover, the magenta triangles represent the associated results for the corresponding NAR model. Compared to the NARX model, the NAR model performs a univariate time series forecast by only considering the time series data of the exhaust gas parameter. This is presented in order to demonstrate the superiority of the NARX model against the NAR model. As observed by examining the graph, compared to the NAR model, the trained NARX model is capable of identifying the fluctuations in the modelled parameter through the addition of the main engine rpm as the exogenous input to the time series model and analysis. Specifically, in the 17th, 18th and 23rd hour mark the exhaust gas temperature forecasted values fall close to the actual monitored values compared to the NAR results.

As observed, although the NAR model for the cylinder exhaust gas forecasts most values successfully, the results obtained have significant errors in timesteps 17, 18 and 23 specifically, owing mostly in the alteration of the main engine's rpm. Specifically, from timestep 1-15 the engine speed is equal to 61-63 rpm while at timestep 16 it is reduced to 40 rpm and is then increased to 51 rpm, thus explaining the shift in the parameters and the reasonable hiatus of the NAR model to capture this change. Specifically, in the 17th hour interval, for the NAR model the APE value between the forecasted and actual exhaust gas outlet temperature value is equal to 27.26%, while the results for the NARX model indicate a major reduction in the APE of 23.12% magnitude, resulting in an APE value for the NARX model of 4.14%. Furthermore, in the 18th hourly mark the APE is reduced from 16.97% to 1.20% and at 23 hours it is reduced from 38.10% obtained in the NAR model to 4.52% by applying the NARX model. Moreover, the overall model MAPE value is reduced from 4.64% down to 1.02%. The APE and MAPE results for the NAR and NARX models are presented in the Appendix.

3.3 Sensitivity analysis results

Both the main engine rpm and the exhaust gas temperature input values are altered progressively from their original value at present time *t* in order to examine the output of the NARX model regarding the one-step-ahead forecast result at time *t+1*. Table 3 presents the different tests conducted for the various input values.

Table 3 Input parameter ranges from their baseline (rpm=61.9, exhaust gas=310.3 ^{o}C)

Reference	**% difference from baseline**	**Engine rpm values (rpm)**	**Exhaust gas values (^{o}C)**
Test -1	[0% -2%]	61.9-60.6	310.3-304.1
Test -2	[-2% -4%]	60.6-59.4	304.1-297.9
Test -3	[-4% -6%]	59.4-58.2	297.9-291.7
Test -4	[-6% -8%]	58.2-56.9	291.7-285.5
Test -5	[-8% -10%]	56.9-55.7	285.5-279.3
Test -6	[-10% -20%]	55.7-49.5	279.3-248.2
Test -7	[-20% -30%]	49.5-43.3	248.2-217.2
Test +1	[0% +2%]	61.9-63.1	310.3-316.5
Test +2	[+2% +4%]	63.1-64.4	316.5-322.7
Test +3	[+4% +6%]	64.4-65.6	322.7-328.9
Test +4	[+6% +8%]	65.6-66.8	328.9-335.1
Test +5	[+8% +10%]	66.8-68.0	335.1-341.3
Test +6	[+10% +20%]	68.0-74.3	341.3-372.4

Test +7	[+20% +30%]	74.3-80.5	372.4-403.4

Figure 7 presents the results for the described 14 test cases regarding the NARX t+1 Monte Carlo results. The results of the Monte Carlo analysis for each test case is presented and the forecasting result for the exhaust gas temperature one-step-ahead prediction (308ºC) from Figure 6 is presented as the baseline reference for comparing the sensitivity analysis outputs.

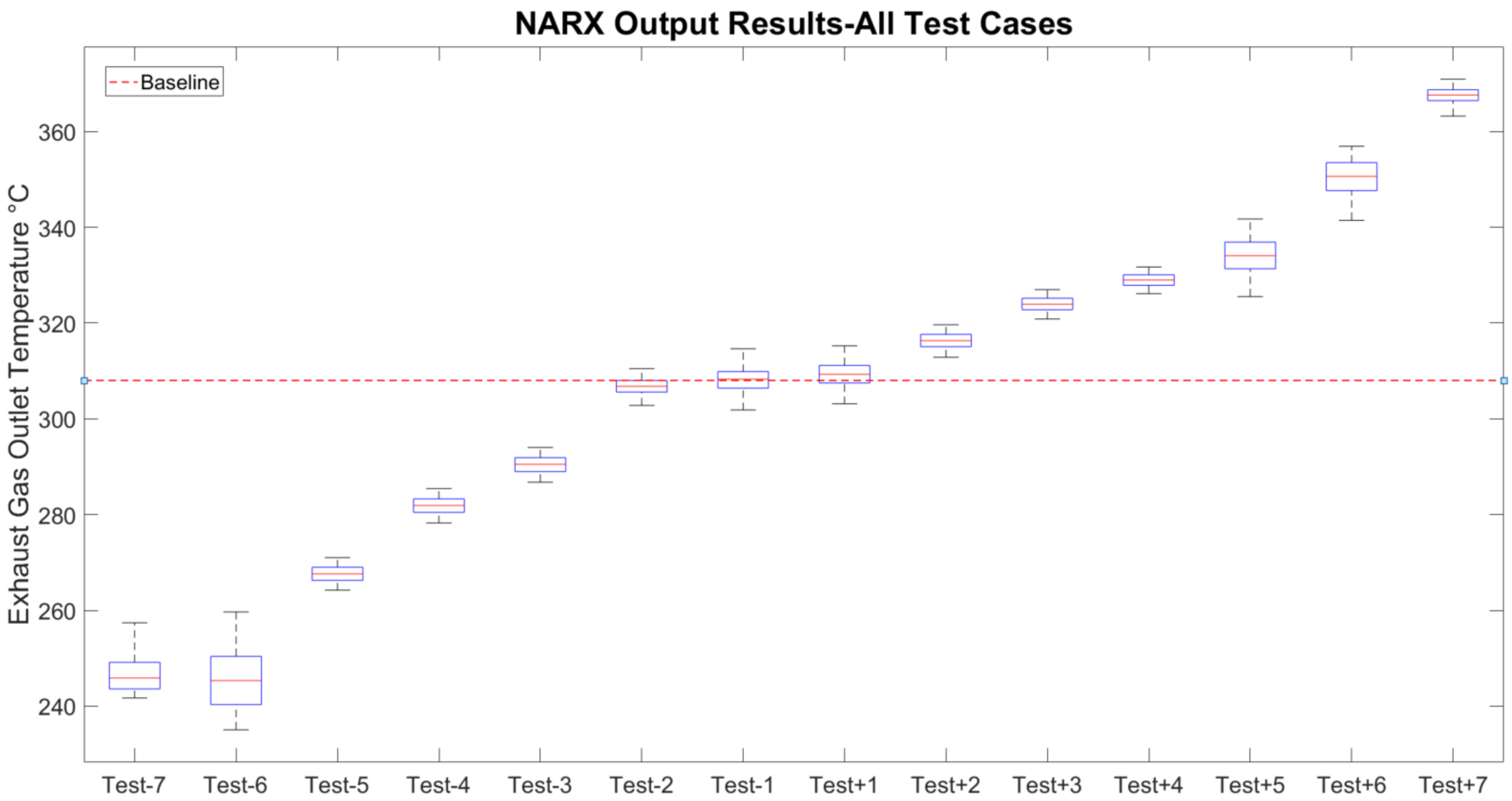

Figure 7: NARX sensitivity analysis results

As observed in Figure 7, for Test-2 to Test+2 the results indicate satisfactory network output for rational reductions and increments of the input parameters around the baseline value. Moreover, in the remaining test cases the error is increased from the baseline as both the exogenous NARX input and exhaust gas input time series data are reduced and increased respectively to protracted ranges. Concurrently, the response of the NARX model to these input parameters also demonstrates its generalisation capabilities and that overfitting manifestations have been evaded. Figure 8 presents the error results obtained from the baseline for the average forecasted results *t+1* of Monte Carlo analysis for each test case.

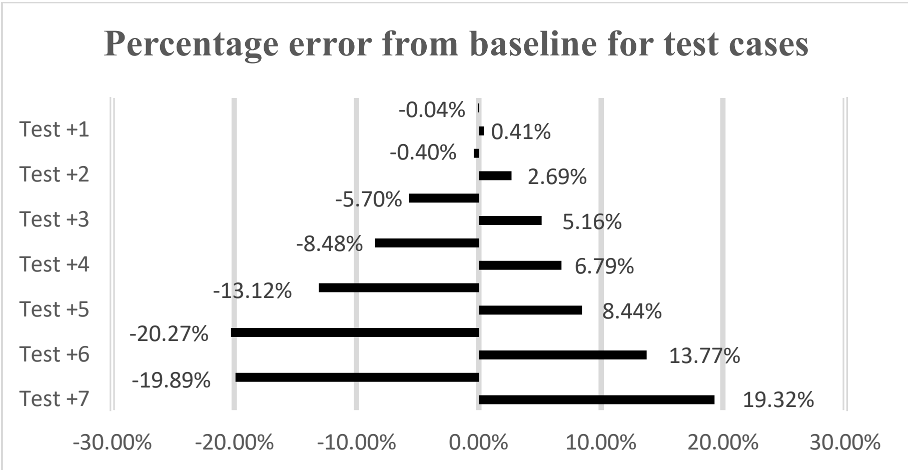

Figure 8: Error results from baseline for each test case

It can be observed that the errors are insignificant for the test cases Test-1, Test+1, Test-2 and Test+2 where reasonable changes in the input parameters were executed. Hence the model is considered robust and the differences in values in these test cases do not cause any significant effects in the outcome. On the other hand, for the remaining test cases the output of the variable and error from baseline increases as the input parameters take on acute values.

4 Discussion

The NARX model provided highly accurate forecast results by evaluating network training and performance, calculated prediction intervals and the APE and MAPE criteria against the actual recorded values. The results obtained for the error autocorrelation function demonstrate the uncorrelation of errors with each other, thus ensuring satisfactory network performance and confidence in the forecasting analysis. Overall, high network fits to the input data were achieved, specifically 98.25% for the training set and 98.02% for the test set. According to Shine et al. (2018), R values greater that 0.90 suggest the strength of agreement between actual and predicted values is excellent, values between 0.80 and 0.90 is considered substantial, while values between 0.65 and 0.80 are considered moderate and poor if lower than 0.65. The forecasted results are demonstrated in two segments of time including the actual data and the forecasted future data. The NAR model was capable of forecasting with small APE values most of the multi-step ahead predictions. Furthermore, it was unable to detect the shift in parameter measurements due to the alteration in the main engine rpm. In order to resolve this issue, the NARX model was developed by introducing the main engine's rpm as exogenous input to the time series model and indicated that the it can identify the fluctuations in the modelled parameters and thus reduced significantly the APE and MAPE values compared to the results obtained from the respective NAR models.

The Bayesian regularisation backpropagation algorithm was used for training the NARX network. This algorithm can provide better solutions for smaller datasets compared to the Levenberg-Marquardt training algorithm, which can also be applied depending on the amount of data available for training (Kayri 2016). When the dataset is small, Bayesian regularisation provides better generalisation performance than early stopping. This is because it does not require that a validation dataset be separate from the training dataset. However, the Bayesian regularisation method takes longer to converge than early stopping (Demuth and Beale 2002).

Although the raw input data undergoes data cleansing and clustering process, it is assumed that adequate data quality exists. It is assumed that no faults or alarms occurred during the onboard measurement campaigns regarding the data collected for analysis. This was confirmed after discussions with the shipping company which provided the data. Moreover, data collection intervals are acquired in uniform time step intervals (hourly), ensuring consistency in the predictions of the developed NARX model.

The NARX model has been introduced for dynamic time series analysis and forecasting, utilising performance parameter data for predicting the future system state, thus enhancing the monitoring framework with predictive capabilities. Research contribution has also been achieved by implementing and integrating sophisticated methods and algorithms for the developed novel framework. This approach can also be applied to other systems in the shipping industry and beyond, thus offering the foundation and opportunity for academics to adapt this approach for applications in other sectors and further validate and enhance the research. The AI tools can provide ship operators with early warnings of upcoming faults to mitigate potential issues in advance, which helps reduce unplanned downtime and increased availability by enabling the crew onboard the ship and the onshore personnel to prepare accordingly and proactively.

The sensitivity analysis was performed by utilising Monte Carlo analysis with the Latin hypercube sampling method ensuring and demonstrating correct network training and generalisation capabilities. The sensitivity analysis was carried out for investigating the change in the NARX model output by shifting systematically both the exhaust gas temperature outlet and main engine rpm parameter. Overall, the sensitivity analysis demonstrated reliable performance of the NARX model output and that it is tolerable to reasonable changes in the input parameters. Therefore, the output is robust to reasonable changes in parameter values within the model. For the extreme test case scenarios, the error from the baseline is increased as expected for the exhaust gas outlet temperature.

The main aim of time series modelling is to carefully collect and rigorously study the past observations of a time series to develop an appropriate model which describes the inherent structure of the series. This model is then used to generate future values for the series. The procedure of fitting a time series to a proper model is termed as time series analysis, while time series forecasting can be termed as the act of predicting the future by understanding the past (Hipel and McLeod 1994). An important task when forecasting a value of y from one or more predictor variables is to obtain an estimate of the likely amount of error inherent in the forecast (Chatfield 1995). A prediction interval is an assessment of this forecast error and is a range that is probable to contain the response value of a single new observation and allows assessment of future uncertainty (Chatfield 1993). The NARX network provided highly accurate forecasted results contained in their calculated prediction intervals. This outcome provides a high level of confidence in the model predictions by addressing uncertainty in the predictions.

The forecast accuracy of the NARX model depends on a variety of factors. These include factors such as the amount of data available for training the NARX model as the data will determine the NARX architecture and thus the forecasting and generalisation capabilities. Moreover, the trendline of the historical data influence the NARX forecasting output as it "*learns from the past to predict the future*".

Paucity of literature in the topic of ANN time series forecasting for marine condition monitoring presents difficulties in comparing the work in this paper with other research. Moreover, various papers utilising data-driven methods use different datasets as input, thus making any head-to-head comparisons subjective in nature as the size and nature of the data will depict the architecture of the data-driven tool. For this reason, this paper also included the NAR model in the case study results in order to achieve a suitable comparison of the NAR and NARX model by using the same exact dataset for training and simulation.

The most recent research projects between academia and industry MINOAS (Caccia et al. 2010), INCASS (INCASS 2015b), MUNIN (Burmeister et al. 2014) and ISEMMS (Gkerekos et al. 2017) indicate the current research trend and exploration of advanced maintenance techniques. These techniques focus on machinery condition monitoring applications onboard, through the utilisation of artificial intelligence, reliability and data-driven methods to provide CBM functionalities and decision support. Moreover, Zaman et al (2017) presented the challenges and opportunities of big data analytics in the shipping sector and potential functionalities with regards to ship safety, energy management and performance monitoring amongst other.

5 Conclusions

In the present paper, the application of a NARX neural network model for the cylinder exhaust gas outlet temperature parameter of a marine main engine was presented. In its wider context, the NARX model can predict various main engine performance parameters. The capabilities and advantages of utilising the NARX neural network have been demonstrated and the methodology is applicable for any performance parameter of the main engine; provided data exists for these performance parameters. The main concluding remarks are:

- The proposed methodology can be applied to other machinery systems such as diesel generators, boilers, turbochargers etc. and can also be extended beyond machinery systems, as it has the flexibility and capability factors to adapt accordingly due to its data-driven nature.
- The AI model can provide a prediction model for CBM in shipping by providing early warnings of upcoming faults to mitigate potential issues in advance.
- High forecasting accuracy was achieved by employing the NAR and NARX models. The case study demonstrated that up to 35% increased forecasting accuracy can be achieved by introducing the main engine rpm as the exogenous input in the NARX models compared to the accuracy of the NAR models. In the NAR model, the APE value between the 23rd hour forecasted and actual exhaust gas outlet temperature value is equal to 38.10%, while the results for the NARX model indicate a major reduction in the APE of 33.58% magnitude, resulting in an APE value for the NARX model of just 4.52%.
- The overall model MAPE value is reduced from 4.64% down to 1.02%.
- The sensitivity analysis overall addressed uncertainty in the predictions and demonstrated the good performance and generalisation of the NARX model as it successfully considers the different values of the exogenous input data for conducting the one-step-ahead output for lower and higher main engine rpm values and exhaust gas temperature values from the baseline.

Future research areas are identified that can extent the research scope. As more data becomes available, the NAR and NARX models should be trained again to fully take advantage of the new information so that the models can capture the developing trend of the monitored system more accurately. When and how to implement such an update should be further examined. Future research also includes investigating NARX capabilities for forecasting various other main engine parameters. Moreover, a set of parallel NARX models can be utilised to forecast multiple performance parameters, and then fuse them in an ANN diagnostic classifier for the whole main engine system in order to develop an overall CBM tool.

Acknowledgments

The work in this paper is partially funded by INCASS project. INCASS has received research funding from the European Union's Seventh Framework Programme under grant agreement No. 605200. This publication reflects only the authors' views and European Union is not liable for any use that may be made of the information contained herein.

References

Adjallah KH, Yang H, Mathew J, Ma L. 2007. Basis pursuit-based intelligent diagnosis of bearing faults. Journal of Quality in Maintenance Engineering.13:152-162.

Aizenberg I, Sheremetov L, Villa-Vargas L, Martinez-Muñoz J. 2016. Multilayer Neural Network with Multi-Valued Neurons in time series forecasting of oil production. Neurocomputing. 175:980-989.

Aladag CH. 2017. Advances in Time Series Forecasting: Volume 2 Sharjah, United Arab Emirates: Bentham Science Publishers.

Basurko OC, Uriondo Z. 2015. Condition-Based Maintenance for medium speed diesel engines used in vessels in operation. Applied Thermal Engineering. 80:404-412.

Bocaniala CD, Jain LC, Palade V. 2006. Computational intelligence in fault diagnosis London, UK: Springer.

Burmeister H-C, Bruhn W, Rødseth ØJ, Porathe T. 2014. Autonomous Unmanned Merchant Vessel and its Contribution towards the e-Navigation Implementation: The MUNIN Perspective. International Journal of e-Navigation and Maritime Economy. 1:1-13.

Caccia M, Robino R, Bateman W, Eich M, Ortiz A, Drikos L, Todorova A, Gaviotis I, Spadoni F, Apostolopoulou V. 2010. MINOAS a Marine INspection rObotic Assistant: system requirements and design. IFAC Proceedings Volumes. 43:479-484.

Chatfield C. 1993. Calculating Interval Forecasts. Journal of Business & Economic Statistics.11:121-135.

Chatfield C. 1995. Model uncertainty, data mining and statistical inference. Journal of the Royal Statistical Society Series A (Statistics in Society).419-466.

Chatfield C. 2000. Time-series forecasting Florida, USA: Chapman & Hall/CRC Press.

Cipollini F, Oneto L, Coraddu A, Murphy AJ, Anguita D. 2018. Condition-Based Maintenance of Naval Propulsion Systems with supervised Data Analysis. Ocean Engineering. 149:268-278.

Demuth H, Beale M. 2002. Neural Network Toolbox-For use with MATLAB User's Guide version 4 United States: The MathWorks Inc.

Firestone M, Fenner-Crisp P, Barry T, Bennett D, Chang S, Callahan M, Burke A, Michaud J, Olsen M, Cirone P. 1997. Guiding principles for Monte Carlo analysis. Washington, DC: US Environmental Protection Agency.

Gkerekos C, Lazakis I, Theotokatos G. 2017. Implementation of a self-learning algorithm for main engine condition monitoring. Maritime Transportation and Harvesting of Sea Resources: Proceedings of the 17th International Maritime Association of the Mediterranean (IMAM).

Hipel KW, McLeod AI. 1994. Time series modelling of water resources and environmental systems Amsterdam, The Netherlands: Elsevier.

IMO. 1993. International Safety Management (ISM) Code, Resolution A741(18). London, UK.

INCASS. 2015a. Deliverable D5.4 'Data exchange'. UK.

INCASS. 2015b. Deliverable D8.1 Exploitation Plan. UK.

Kayri M. 2016. Predictive Abilities of Bayesian Regularization and Levenberg–Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. Mathematical and Computational Applications.21:20.

Laboissiere LA, Fernandes RAS, Lage GG. 2015. Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks. Applied Soft Computing. 35:66-74.

Lazakis I, Ölçer A. 2015. Selection of the best maintenance approach in the maritime industry under fuzzy multiple attributive group decision-making environment. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment.230:297-309.

Liu H, Tian H-q, Liang X-f, Li Y-f. 2015. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Applied Energy. 157:183-194.

Maier HR, Dandy GC. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental Modelling & Software. 15:101-124.

Nasr MS, Moustafa MAE, Seif HAE, El Kobrosy G. 2012. Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. Alexandria Engineering Journal. 51:37-43.

Noor M, MH MY, MM N. 2016. Prediction of Marine Diesel Engine Performance by Using Artificial Neural Network Model. Journal of Mechanical Engineering and Sciences (JMES).10:1917-1930.

Okut H. 2016. Bayesian Regularized Neural Networks for Small n Big p Data. In: Artificial Neural Networks-Models and Applications. London, UK: InTech.

Pascual DG. 2015. Artificial Intelligence Tools: Decision Support Systems in Condition Monitoring and Diagnosis USA: Crc Press.

Prajapati A, Bechtel J, Ganesan S. 2012. Condition based maintenance: a survey. Journal of Quality in Maintenance Engineering.18:384-400.

Raptodimos Y, Lazakis I. 2018. Using artificial neural network-self-organising map for data clustering of marine engine condition monitoring applications. Ships and Offshore Structures. 13:649-656.

Raza J, Liyanage JP. 2009. Application of intelligent technique to identify hidden abnormalities in a system: A case study from oil export pumps from an offshore oil production facility. Journal of Quality in Maintenance Engineering.15:221-235.

Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. 2008. Global sensitivity analysis: the primer New Jersey, USA: John Wiley & Sons.

Shields MD, Zhang J. 2016. The generalization of Latin hypercube sampling. Reliability Engineering & System Safety. 148:96-108.

Shine P, Murphy MD, Upton J, Scully T. 2018. Machine-learning algorithms for predicting on-farm direct water and electricity consumption on pasture based dairy farms. Computers and Electronics in Agriculture. 150:74-87.

Stopford M. 2009. Maritime economics 3e New York, USA: Routledge.

Sullivan GP, Pugh R, Melendez AP, Hunt WD. 2010. Operations & Maintenance Best Practices-A guide to achieving operational efficiency. UDo Energy, translator. Third ed. Washington D.C, USA.

Szoplik J. 2015. Forecasting of natural gas consumption with artificial neural networks. Energy. 85:208-220.

Tamilselvan P, Wang P. 2013. Failure diagnosis using deep belief learning based health state classification. Reliability Engineering & System Safety. 115:124-135.

Tan WL, Nor NM, Abu Bakar MZ, Ahmad Z, Sata SA. 2012. Optimum parameters for fault detection and diagnosis system of batch reaction using multiple neural networks. Journal of Loss Prevention in the Process Industries. 25:138-141.

Taylor DA. 1996. Introduction to marine engineering USA: Elsevier Ltd.

Tian Z, Wong L, Safaei N. 2010. A neural network approach for remaining useful life prediction utilizing both failure and suspension histories. Mechanical Systems and Signal Processing. 24:1542-1555.

Tinsley D. 2016. Dawning of new era in asset maintenance. Marine Power & Propulsion Supplement 2016.

UNCTAD. 2017. Review of Maritime Transport 2017. Geneva, Switzerland: U Nations.

Wei P, Lu Z, Yuan X. 2013. Monte Carlo simulation for moment-independent sensitivity analysis. Reliability Engineering & System Safety. 110:60-67.

Zaman I, Pazouki K, Norman R, Younessi S, Coleman S. 2017. Challenges and Opportunities of Big Data Analytics for Upcoming Regulations and Future Transformation of the Shipping Industry. Procedia Engineering. 194:537-544.

Zhou J, Xu L. The fault diagnosis of marine engine cooling system based on artificial neural network (ANN). Proceedings of the The 2nd International Conference on Computer and Automation Engineering (ICCAE); 2010.

Zhu J. Marine diesel engine condition monitoring by use of BP Neural Network. Proceedings of the Proceedings of the International MultiConference of Engineers and Computer Scientists 2009.

Appendix

NARX APE and MAPE results

Parameter	Results	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
Cylinder Exhaust Gas Temperature no.5	Actual	308.10	308.30	307.50	308.40	307.80	308.30	221.60	238.10	279.80	276.20
	ANN Prediction	308.00	308.30	307.13	308.35	307.86	308.42	230.78	235.25	272.75	275.84
	APE	**0.03%**	**0.00%**	**0.12%**	**0.02%**	**0.02%**	**0.04%**	**4.14%**	**1.20%**	**2.52%**	**0.13%**

Parameter	Results	t+11	t+12	t+13	t+14	t+15	t+16	t+17	t+18	t+19	t+20
Cylinder Exhaust Gas Temperature no.5	Actual	275.00	274.10	196.70	282.10	281.00	278.70	276.00	277.20	274.20	249.20
	ANN Prediction	275.20	274.90	205.60	279.05	280.02	280.67	279.06	279.04	276.65	255.32
	APE	**0.07%**	**0.29%**	**4.52%**	**1.08%**	**0.35%**	**0.71%**	**1.11%**	**0.66%**	**0.89%**	**2.46%**

MAPE=1.02%

NAR APE and MAPE results

Parameter	Results	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
Cylinder Exhaust Gas Temperature no.5	Actual	308.10	308.30	307.50	308.40	307.80	308.30	221.60	238.10	279.80	276.20
	ANN Prediction	307.50	308.30	307.50	308.40	307.80	307.30	282.00	278.50	279.00	275.10
	APE	**0.19%**	**0.00%**	**0.00%**	**0.00%**	**0.00%**	**0.32%**	**27.26%**	**16.97%**	**0.29%**	**0.40%**

Parameter	Results	t+11	t+12	t+13	t+14	t+15	t+16	t+17	t+18	t+19	t+20
Cylinder Exhaust Gas Temperature no.5	Actual	275.00	274.10	196.70	282.10	281.00	278.70	276.00	277.20	274.20	249.20
	ANN Prediction	275.20	273.90	271.65	282.00	282.30	280.13	279.13	279.04	278.41	261.32
	APE	**0.07%**	**0.07%**	**38.10%**	**0.04%**	**0.46%**	**0.51%**	**1.13%**	**0.66%**	**1.54%**	**4.86%**

MAPE=4.64%