Prediction of Full-Scale Propulsion Power using Artificial Neural Networks

Benjamin Pjedsted Pedersen, FORCE Technology, TU of Denmark, bpp@force.dk
Jan Larsen, Technical University of Denmark, jl@imm.dtu.dk

Abstract

Full scale measurements of the propulsion power, ship speed, wind speed and direction, sea and air temperature from four different loading conditions, together with hind cast data of wind and sea properties; and noon report data has been used to train an Artificial Neural Network for prediction of propulsion power. The model was optimized using a double cross validation procedure. The network was able to predict the propulsion power with accuracy between 0.8-1.7% using onboard measurement system data and 7% from manually acquired noon reports.

1. Introduction

As part of the Industrial PhD project "Ship Performance Monitoring" automatic data sampling equipment was installed on the tanker "Torm Marie" in January 2008 and presently data from four different loading conditions are available. By considering the ship as a dynamical system which can be modelled as a general nonlinear state-space model, the ship propulsion performance (referred to as the performance) is a measure of energy consumption which depends on the current state of the ship and a large number of external factors/variables such as speed, loading conditions, ship conditions, weather and sea conditions. Fig.1 shows some factors influencing propulsion performance.

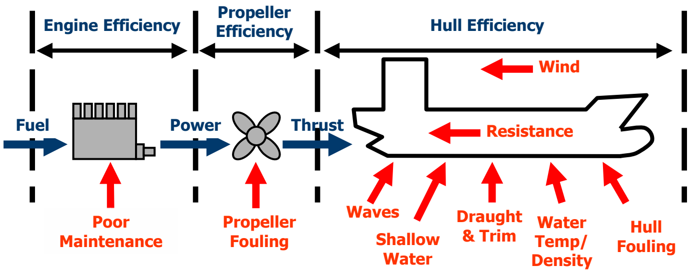

Fig.1: Variables influencing propulsion performance

The variables have different properties. Some of the variables are observable (and measurable with high reliability) whereas others are difficult to observe, e.g. the fouling. Some variables are largely controllable, whereas others are almost incontrollable. For instance, heading is controllable whereas wind conditions and fouling are almost incontrollable. The first goal in performance measurement is to provide a reliable estimation of performance as a function of the state and external variables. The second goal is to optimize performance by manipulating the controllable variables. This paper will focus on the first goal.

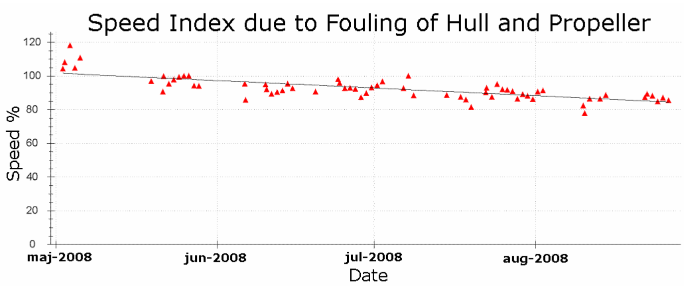

Fig.2: Increase in fuel consumption as consequence of fouling

During the lifetime of the ship the performance will decrease. As an example the fuel consumption will increase at a certain state, or the ship speed will decrease at a certain power setting. This is mainly due to fouling of the hull and propeller. Fig.2 shows a typical trend of the speed reduction.

This work will not consider a full dynamical model of the ship but merely focus on a model which predicts propulsion power in a specific state based on the measurements of a set of significant input variables, which are:

- Ship speed through the water
- Wind speed and direction
- Seawater temperature
- Air temperature
- Water depth
- Wave height and direction

The model can be based on a classical physical/empirical model, e.g. *Harvald (1983)* or *Holtrop (1984),* or a data-driven (non-parametric) approach, e.g. an artificial neural network. Previous work suggests that a data-driven approach is preferable, e.g. *Pedersen and Larsen (2009*). The empirical methods are derived from model tests and sea trials, and since most model tests are carried out in the design condition (even keel) and speed, this is the region where it should be applied. In operation the ship will travel in many other conditions i.e., ballast draught and trimmed conditions. Consequently, these methods give a rough estimate of the propulsion power rather than an accurate reference point. If measured values from model tests or sea trials are available, they can be used to adjust the empirical data and thus give a more accurate result. Fig.3 shows the measured power together with estimated power using respectively *Harvald (1983)* and *Holtrop (1984)*, with a standard setup i.e., without any adjustment. It is obvious that a change in a few percent, which is realistic performance deterioration over a year, is impossible to detect.

Furthermore the traditional methods are based on "Noon Reports" data, which are reports containing information of the ship speed, travelled distance, position, heading, and a number of other measurements and readings. One problem with noon reports are that only one *sample* is collected per day, excluding days in harbour and e.g. travelling in areas with limited water depth. This might leave out 200 observations per year. Many noon reports data are mean values over time from the last noon report, e.g. average logged ship speed, and others are observations at the report time, e.g. current wind speed. This makes it difficult to analyse relations e.g., between the average ship speed and the instantaneous wind speed.

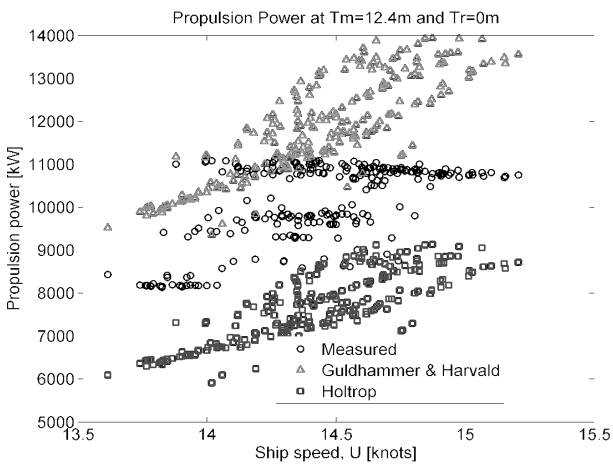

Fig.3: Example of empirical power prediction using *Harvald (1983)* and *Holtrop (1984)***,** compared with measure values for a single loading condition (Mean draught: 12.0m, even keel).

If sufficient data is available it is possible to make a partial or fully data driven model. The first step of such a model is to capture the dynamics of the fastest changing variables, e.g. ship speed, wind speed, etc. where the slowest changing variables are draught and trim. Initially, this leads to a prediction model for data sets where draught and trim are kept constant.

Adequate quality of input data is fundamental to get reliable results from the prediction model. This is a problem for sea and weather information data due to the difficulty of measuring these quantities, but especially to noon report data which are collected manually hence human factors and errors can play a significant role.

2. Data sources

Three data input sources have been used to train and predict the propulsive power:

1. Onboard measured data (4 conditions)
2. Noon report data
3. Hind cast weather and sea information

An overview of all the relevant data set variables is listed in Table 4.

Onboard measured data: The data was collected onboard the 110,000 dwt tanker "Torm Marie" where a number of measurements were continuously logged. Only the relevant data for this problem has been taken. The sampling was split into time series of 10 min with 10 min intervals. The sampling frequency of the time series was 1 second, but many of the measurements had inconsistent and missing signal values. Power and speed were updated consistently, every 13 s. Four independent data sets, with different loading conditions have been sampled so far. The data include samples from non-stationary situations as well as situations with zero forward speed, which are deleted. The variance of the heading is one of the governing figures on the variation of the propulsion power in particular. Even small changes (less than 1°) in the heading, had significant influence on the measured propulsion power. Samples with excessive variance in the heading have thus been excluded. During each data sampling period factors related to the ship performance including the hull and propeller fouling, were assumed to be constant, consequently this effect will not be accounted for in the analysis

Noon report data: The noon reports contain a long array of data and basically the same variables as the measured ones, but with differences in quality and resolution. Due to their nature, noon reports are usually only collected once a day, which gives a smaller resolution and a mix of data with different origins, e.g. logged average speed over ~24 hours and one weather observation at the report time. Noon reports are usually filled in manually and are thus also subject to human factors and errors. In this analysis the noon reports are important for obtaining the draught and seawater temperature.

Hind cast data: Hind cast data has been received from a tool developed for SeaTrend®1 at FORCE Technology based on weather information from NOAAH2. For a given position and time this tool returns wind speed and direction, significant wave height, peak period and direction. Some areas, e.g. the Mediterranean are not included in this database.

Dataset for training and test: Two different configurations of the dataset were used for the analysis. One based on the measured values and one based on noon report data.

Onboard measured dataset: The dataset based on measured values has a high density of data (approx. 72 per day), but there is only a limited amount of this data available: in total 27 days, Table 1.

[1] Performance Monitoring tool developed at FORCE Technology, www.force.dk
[2] National Oceanic and Atmospheric Administration, United States Department of Commerce http://www.noaa.gov/

Table 1: Onboard measured data sets. N represents the number of 10 minute recording windows

Data set	Number of Samples	Start date End date	Number of valid noon reports	Mean draught, T_m	Trim Ta-Tf	U_{min}-U_{max}	P_{min}-P_{max}
M	N			[m]	[m]	[knots]	[kW]
1	236	09-02-2008 14-02-2008	3	7.4	2.4	14.2- 16.2	7573-11283
2	109	22-03-2008 27-03-2008	4	7.85	2.7	13.6- 15.1	7750-9248
3	301	30-01-2008 06-02-2008	7	12.15	0	13.4- 16.0	8138-11216
4	555	01-03-2008 11-03-2008	9	13.0	0	13.0- 15.9	9741-12096

All the measured input data are the mean values over 10 minutes of the time series. In order to justify this, variance of the signal has been analyzed for the ship speed, U, propulsion power, P and apparent wind speed, V_R. The air temperature has been neglected since it is very stable. For every 10 minute period the relative standard deviation, ($\sigma_{x,n}/\mu_{x,n}$) has been found and for every dataset the average of the relative standard deviation $\bar{\mu}_{x,M}$, has been determined:

$$\bar{\mu}_{x,M} = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{\sigma_{x,n}}{\mu_{x,n}}\right|, \tag{1}$$

$\sigma_{x,n}$ is the standard deviation for the *n'th* time series, $\mu_{x,n}$ is the mean value for the *n'th* time series, and x indicates the input/output variables (U, P, V_R, γ_R)

The relative standard deviation of both the measured power and ship speeds are all less than 1, whereas the wind speed has a significantly high variance.

Table 2: Average of the relative standard deviation

M	N	$\bar{\mu}_U$	$\bar{\mu}_P$	$\bar{\mu}_{V_R}$
1	236	0.6%	0.7%	18.0%
2	109	0.6%	0.5%	9.1%
3	301	0.6%	0.9%	9.5%
4	555	0.6%	0.6%	11.4%

The ship speed intervals are approximately in the same region for each sample, Table 2. However inspecting the distributions of the ship and true wind speed, Figs.4 to 7, the actually ship speed range is different. Especially for dataset #2 where most of the ship speeds is in a band of around 14.7 knots. The Beaufort wind force (BF) 5 starts at approximately *16 knots* wind speed. In this condition the wind driven waves are around 2 m high, which is when the sea state starts to influence the power increase in waves. Only a few occurrences are above this level and thus datasets #1 and #2 can be regarded as calm water conditions, Figs.4 and 5. Datasets #3 and #4 on the other hand have a more significant contribution of measurements above BF 5 and the power increase in waves must be regarded as an extra contribution.

The sea state has a significant influence on the ship resistance and hence the propulsion power. No direct measurements of the sea state have been made, but the wind driven waves can be represented by the true wind speed to a certain extent. Making this assumption the swell is not accounted for.

Hind cast information gives an estimate of the sea state, including significant wave heights, peak period and direction, at the specific position and time, and has been found for all the relevant data. Furthermore the hind casts also give the true wind speed and direction.

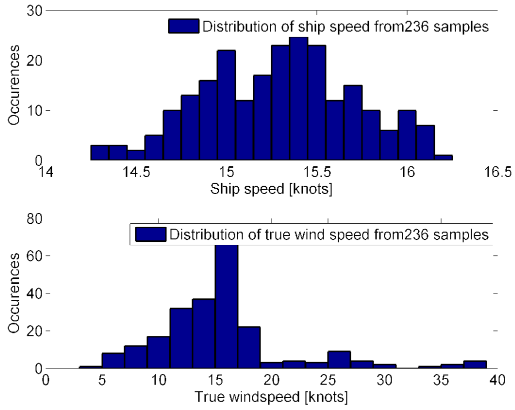

Fig.4: Ship speed and true wind speed distribution of dataset #1

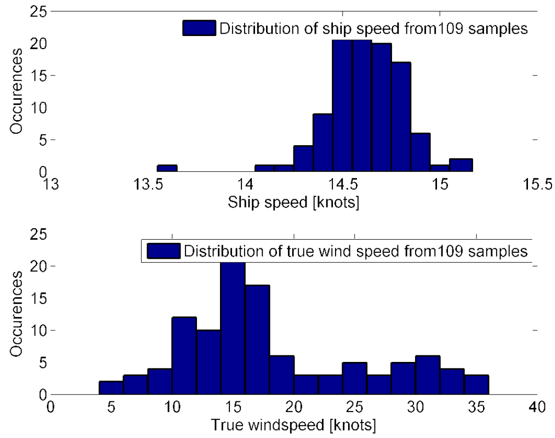

Fig.5: Ship speed and true wind speed distribution of dataset #2

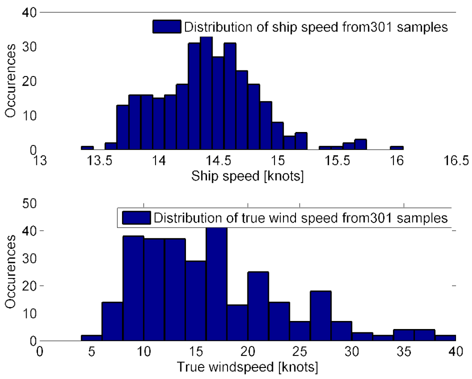

Fig.6: Ship speed and true wind speed distribution of dataset #3

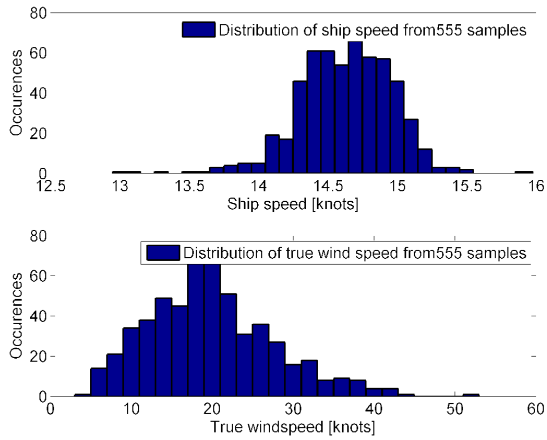

Fig.7: Ship speed and true wind speed distribution of dataset #4

Noon report dataset: The time density of the dataset based on noon reports is much less than for the measure based dataset. There is a maximum of one sample per day and many are invalid due to e.g. anchoring, alongside in harbour. But the time span is much longer, approximately 2 years and the variation in draught and trim has to be taken into account, and possibly also the time, Table 3. In order to give a more representative value of the sea state and wind condition for the noon report data, hind cast has been made for every hour in between each noon report. Afterwards the mean value and variance of the time series (approximately 24 hours) prior to the report time, has been found and are thus ready to use for the analysis.

Table 3: Noon report dataset for analysis

Date UTC	Number of valid samples	Mean draught [m]	Trim, Ta-Tf [m]	Ship speed [knots]	Seawater temp [°C]	Specific HFO [tons/day]
09-12-2006 - 05-12-2008	323	7.35-15.35	0-3.4	9.9-17.5	12-32	1.1-3.9

Table 4: Propulsion performance variables

Data		Unit	Data source
Speed through water	*U*	[knots]	Measured onboard
Relative wind velocity	*Vrel*	[knots]	Measured onboard
Relative wind direction	*grel*	[deg]	Measured onboard
Air temperature	*Tair*	[degC]	Measured onboard
Propulsion power	*P*	[kW]	Measured onboard
Logged mean speed	*NR.U*	[knots]	Noon report
Sea water temperature	*NR.Tsw*	[degC]	Noon report
Air temperature	*NR.Tair*	[degC]	Noon report
Arrival draught fore	*Tf*	[m]	Noon report
Arrival draught aft	*Ta*	[m]	Noon report
Specific fuel consumption	*SpHFO*	[ton/hour]	Noon report
Report time, UTC	*NR.UTC*	[hh:mm:ss]	Noon report
True wind speed	*HC.Ws*	[m/s]	Hind cast
True wind direction	*HC.g*	[deg]	Hind cast
Significant wave height	*HC.Hs*	[m]	Hind cast
Wave period	*HC.Tp*	[s]	Hind cast
True wave direction	*HC.Td*	[deg]	Hind cast
Mean arrival draught	*Tm*	[m]	Derived from noon reports
Arrival trim, *Ta-Tf*	*Trim*	[m]	Derived from noon reports
Relative wind speed	*HC.Vrel*	[knots]	Derived from hind casts
Relative wind direction	*HC.grel*	[deg]	Derived from hind casts

3. Regression models for propulsion power prediction

Three different regression models have been tested and evaluated: a linear model, a (custom) non-linear model and a Artificial Neural Network model

3.1. Linear and Non-linear models

Both a linear and non-linear method based on the general assumption of relation between the ship speed, wind speed and power was developed and presented in *Pedersen and Larsen (2009)*. In short the methods are based on Eq.(2) which can be developed to Eqs.(3) and (4) where $\Delta\eta_D^{-1}$, ΔK and ΔL are adjustable parameters optimized using a "Levenberg-Marquardt" optimization routine, *Madsen et al. (2004), Nielsen (1999)*). If $\Delta\eta_D^{-1}$ is zero the model is regarded as linear.

$$P_D = \eta_D^{-1} U \left(R_{SW} + R_{wind}\right) \tag{2}$$

$$P_D = \eta_D^{-1} U \left(KU^2 + LV_R^2\right) \tag{3}$$

Is it now possible to adjust the three parameters, η_D^{-1} K and L, by introducing the additional weights, $\Delta\eta_D^{-1}$, ΔK and ΔL.

$$P_D = \left(\eta_D^{-1} + \Delta\eta_D^{-1}\right) U \left((K + \Delta K)U^2 + (L + \Delta L)V_R^2\right), \tag{4}$$

where,

$$K = C_{SW}\, \tfrac{1}{2} \rho_{SW} S \tag{5}$$

$$L = C_X\, \tfrac{1}{2} \rho_{air} A_T \tag{6}$$

Both the linear and non-linear methods resulted in a cross validation error, Eq.(9), of 3-12%. This could be improved a bit by using a "Leave One Out" (LOO) routine for training of the linear and non-linear models and by subsequently using the mean of the *N* weights from the LOO training as the final weights. But in order to ensure consistency with the ANN models the data has been split into test and training sets as described in section 5.

3.2. Artificial Neural Network (ANN)

After a brief test of regression with an ANN this method appeared superior to the previously described methods which lead to a thorough exploration of the ANN methods. An ANN is a non-linear method where the so called hidden layer with hidden units is the non-linear link between input and output, Fig.8, as described in Eqs.(7) and (8).

$$y(x) = \sum_{j=0}^{M} w_{kj}^{(2)} z_j \,, \tag{7}$$

$$z_j = g\left(\sum_{i=0}^{d} w_{ji}^{(1)} x_i \right), \tag{8}$$

where x is the measured input data; y is the output, in this case the propulsion power; z_j are the nonlinear basis functions; w are the weights for the hidden units and output.

The network used for this analysis is a flexible non-linear regression model with additive Gaussian noise and is trained with a Bayesian learning scheme. It has a tangent hyperbolic sigmoidal function and is trained using a BFGS (Broyden-Fletcher-Goldfarb-Shanno) optimization algorithm with a soft line search to determine step lengths. The Hessian matrix is evaluated using the Gauss-Newton approximation.

More details into the specific neural network used here can be found in the following references: *DTU toolbox (2002)*, *Larsen (1993)*, *MacKay (1992)*, *Pedersen (1997)*, *Svarer et al.(1993)*. A basic description of neural networks can be found in *Bishop (2006)*.

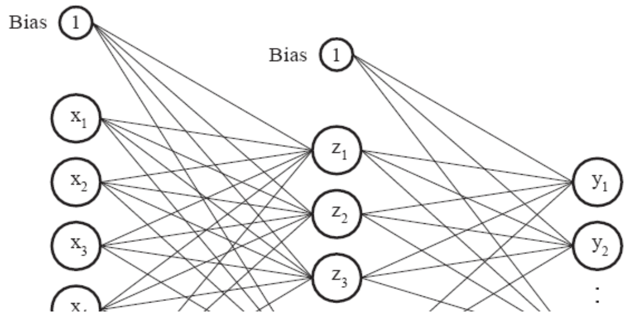

Fig.8: Single hidden layer artificial neural network, with multiple outputs

4. Training and testing

In order to test and validate different variations of input variables and the number of hidden units and to select the best combination it is necessary to split the data set into three parts: a test set, validation set and a training set. A double cross validation method was used in order to find the best combination of number of hidden units and the input variable. Double cross validation consists of two steps.

First training is performed on the training set and tested on the validation set, referred to as the inner-loop, the optimum set input parameters and number of hidden units is decided from the cross validation error of the mean relative error $\overline{\omega}_m$, Eq.(9). Using the optimum setup training is performed on the validation set and training set, and then tested on the test set, referred to as the outer-loop. This procedure is called double cross validation.

From the dataset 20% was used for test data and 80% validation and training. The validation and training set was then split into 20% for validation and 80% training, corresponding to respectively (20%·80%) 16% and (80%·80%) 64% for validation and training of the entire dataset. This separation is illustrated in Table 5. The best combination of input parameters was selected manually. All the training of the ANN has been carried out using the above mentioned routine where training has been started 10 times to ensure capturing the best solution.

Table 5: Double cross validation illustration

Total data set				
Test set 20%	Val set 20%· 80%	Training set 80%· 80%		
Val set 20%· 80%	Test set 20%	Training set 80%· 80%		
Training set 80%· 80% / Val set 20%· 80%		Test set 20%	Training set 80%· 80% / Val set 20%· 80%	
Training set 80%· 80%			Test set 20%	Val set 20%· 80%
Training set 80%· 80%			Val set 20%· 80%	Test set 20%

Due to limited computational time, it was not possible to train and test all combinations of input variables. The selection was made manually based on initial testing and basic physical assumptions. Table 6 shows the different input combinations for the analysis of the measured dataset. Draught and trim are missing, as the analysis is performed for each loading condition separately. The output variable, propulsion power *P,* is naturally present in all combinations.

Table 6: Input variable setup for the measurement based dataset

ID	U	Vrel	grel	Tair	NR.Tsw	NR.Tair	HC.Ws	HC.g	HC.Hs	HC.Tp	HC.Td	HC.grel	HC.Vrel	P
2	x	x	x	x	x									x
8	x	x	x	x	x		x	x	x	x	x			x
10	x				x	x	x	x	x	x	x			x
11	x				x	x						x	x	x
12	x				x	x			x	x	x	x	x	x

Table 7: Input variable setup for the noon report based dataset

					Mean value of 1 hour intervals during steaming time									
ID	NR.U	NR.U TC	NR. Tair	NR. Tsw	HC.Ws	HC.g	HC. Hs	HC. Tp	HC. Td	HC. grel	HC. Vrel	Tm	Trim	SpHFO
2	x		x	x	x	x						x	x	x
3	x		x	x	x	x	x	x	x			x	x	x
4	x		x	x			x	x	x			x	x	x
5*	x		x	x	x	x	x	x	x			x	x	x
6	x		x	x						x	x	x	x	x
7	x	x	x	x						x	x	x	x	x
8	x		x	x			x	x	x	x	x	x	x	x
9	x	x	x	x			x	x	x	x	x	x	x	x

* The variance of HC.Ws, HC.g, HC.Hs, HC.Tp and HC.Td over the steaming time has also been included as input variables.

Table 7 shows the combination scheme of the input variables for noon report analysis. All the hind cast data are mean values of hind cast data produced for the last steaming time period with one hour intervals. That makes all the input data mean values of the steaming time period, instead of instant values at the report time. No propulsion power is available and thus the specific fuel consumption, SpHFO, is used as an output variable.

5. Evaluation

In order to make a consistent evaluation of the ANN training and testing two cross validation errors have been introduced: One for the inner-loop of the double cross validation, testing on the validation set, $\overline{\omega}_K$ Eq,(9) and one for the outer-loop of the double cross validation, testing on the test set, $\overline{\omega}_M$.

$$\overline{\omega}_K = \frac{1}{K}\sum_{k=1}^{K}\omega_k \tag{9}$$

Where K is the number of cross validation set for the *inner-loop* and ω_k is the mean of the relative error for each of the cross validation sets, Eq.(10).

$$\omega_k = \frac{1}{N_k}\sum_{n=1}^{N_k}\left|\frac{\hat{P}_{test,n} - P'_{test,n}}{P'_{test,n}}\right|, \tag{10}$$

$\hat{P}_{test,n}$ are the predicted values of the test data, $P'_{test,n}$ are the test samples from the validation set, and N is number of test data. The outer-loop cross validation $\overline{\omega}_M$ error is equivalent to $\overline{\omega}_K$, Eq.(9), except that the test set has been used for the mean relative error opposed to the validation set.

6. Results

6.1. Results from measured input/output data

Due to lack of computation time the ANN was only trained for 5 and 20 hidden as these are the extremes. This can be justified by *Pedersen and Larsen (2009),* where training/test were performed with 5,10,15 and 20 hidden units, it was concluded than the number of hidden units are not critical to the solution, although in general 5 hidden units were too little. Table 8 shows the cross validation errors of the inner-loop $\overline{\omega}_K$ for the input variable combinations defined in Table 6 and 5 and 20 hidden units.

Looking at all the cross validation errors for each of the datasets in Table 8 it is clear that some datasets in general have smaller errors. Particularly in dataset #2 and to some extent #4 it is noted that the cross validation errors do not vary no matter what input data variables or number of hidden units are used. In these datasets it is thus difficult to detect what input variables have the most influence on the solutions.

Table 8: Inner-loop cross validation errors, $\overline{\omega}_K$

	Dataset #1		Dataset #2		Dataset #3		Dataset #4	
Number of hidden units	5	20	5	20	5	20	5	20
Input variable combination	$\overline{\omega}_K$	$\overline{\omega}_K$	$\overline{\omega}_K$	$\overline{\omega}_K$	$\overline{\omega}_K$	$\overline{\omega}_K$	$\overline{\omega}_K$	$\overline{\omega}_K$
2	3.93%	2.92%	1.07%	0.81%	3.07%	2.37%	1.72%	1.30%
8	2.63%	1.97%	0.97%	0.89%	2.21%	1.65%	1.49%	1.04%
10	2.14%	1.65%	0.99%	0.95%	2.17%	1.65%	1.52%	0.94%
11	3.77%	2.79%	1.10%	0.90%	2.45%	1.88%	1.42%	1.02%
12	2.25%	1.65%	0.99%	0.94%	1.75%	1.40%	1.28%	0.90%

The cross validation errors of the outer-loop are presented in Table 9, together with respectively the best combination of hidden units and input variable combinations. Datasets #2 and #4 have rather low cross validation errors, which must be due to the nature of the dataset. What is more interesting is to see how the error drops by introducing hind cast sea state information and the best solutions in general are where only the hind cast information has been used for the sea and wind property inputs.

Table 9: Table of outer-loop cross validation errors $\overline{\omega}_M$

	Dataset #1	Dataset #2	Dataset #3	Dataset #4
Optimum number of hidden units	20	20	20	20
Optimum input variable combination	10 / 12	2	12	12
$\overline{\omega}_M$	1.63%/1.74%	0.83%	1.46%	0.80%

Figs.9 to 16 show the test errors for the best combination of the number of hidden units and input variables. The plots on the left show the test errors as a function of the ship speed; there is no apparent correlation between ship speed and the error, which indicates that the ship speed has integrated properly into the model. The plots on the right are a relative histogram of the errors together with a normal distribution based on the mean and the variance of the test errors. Except for dataset #2, all the error distributions are centered approximately around *0* and have a nice distribution. Dataset #2 is a sparser dataset so each bar represents 1-4 counts, but the errors are relatively small.

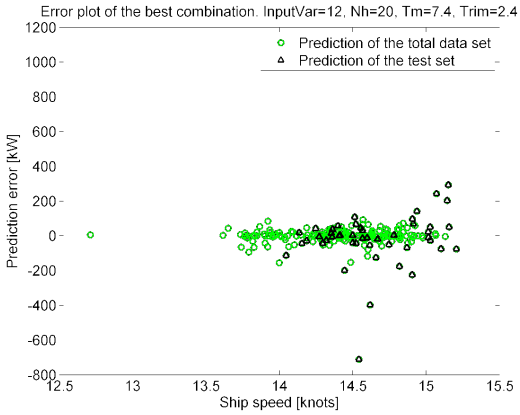

Fig.9: Prediction errors for dataset #1

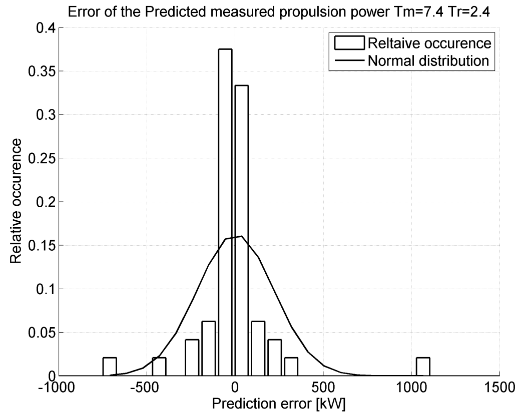

Fig.10: Relative distribution of the predicted errors for dataset #1

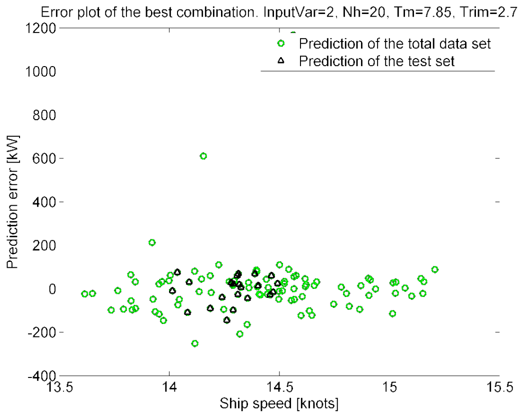

Fig.11: Prediction errors for dataset #2

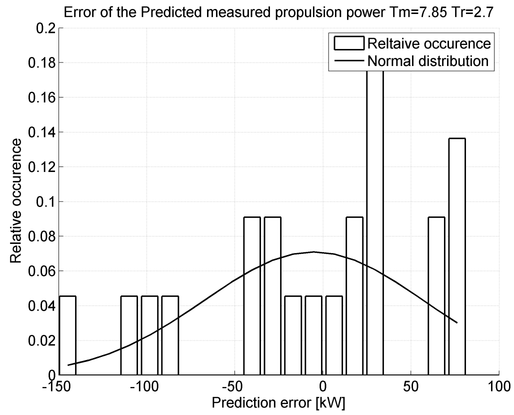

Fig.12: Relative distribution of the predicted errors for dataset #2

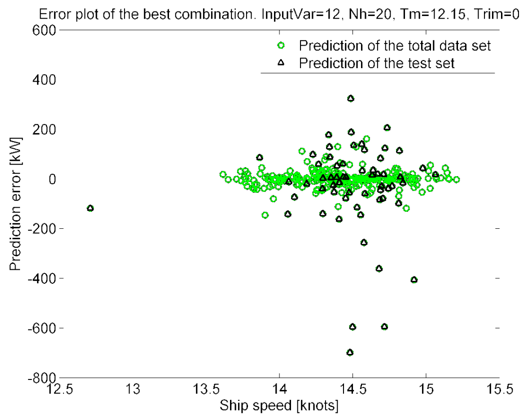

Fig.13: Prediction errors for dataset #3

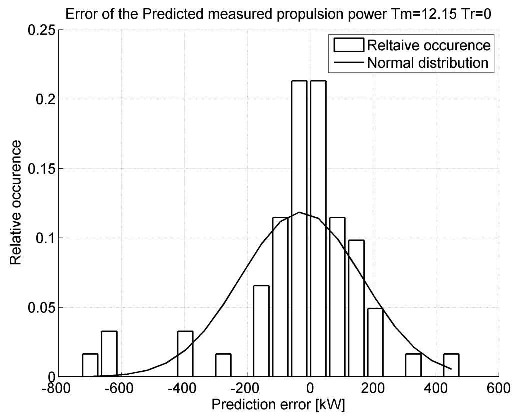

Fig.14: Relative distribution of the predicted errors for dataset #3

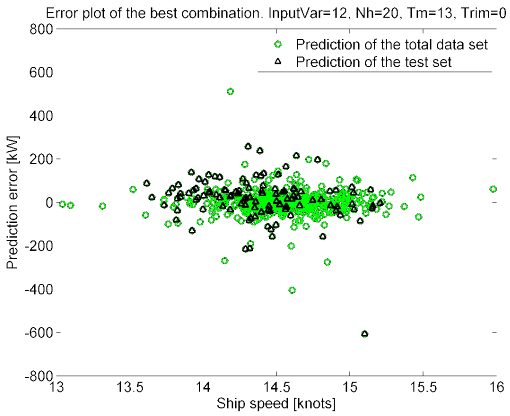

Fig.15: Prediction errors for dataset #4

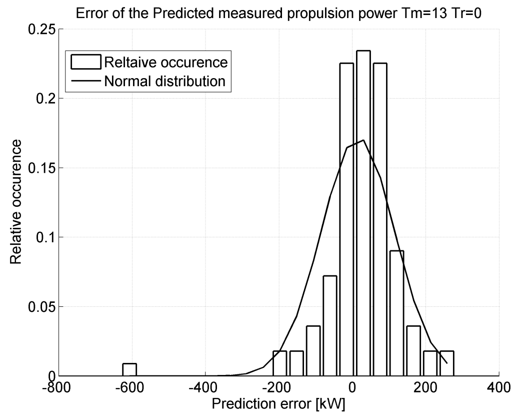

Fig.16: Relative distribution of the predicted errors for dataset #4

Training and testing was performed as described above using a similar cross validation although not double since the input variables are specified by the model.

Table 10: Cross validation errors $\overline{\varpi}$ for the dataset, based on measurement using a linear and non-linear method

	Cross validation error	Dataset #1	Dataset #2	Dataset #3	Dataset #4
Linear method	$\overline{\varpi}$	11.58%	3.63%	12.04%	5.98%
Non-linear method	$\overline{\varpi}$	11.36%	3.58%	10.79%	5.98%

6.2. Results from noon report data

From the inner-loop cross validation errors listed in Table 10 it is noted that the model in general is less sensitive to the number of hidden units. The dependency on certain variables seems not very strong since most errors are in the same region. The error drops significantly when the time is introduced as a variable for input variable combination 7 and 9. Furthermore combination *5* might have been over trained since it has the highest number of input variables but one of the highest errors.

The outer-loop cross validation error is presented in Table 12 and is based on the input variable combination *7* which does not even take into account the sea state.

Table 11: Inner-loop cross validation errors, $\overline{\varpi}_m$ for the noon report dataset

Number of hidden units	5	10	15	20
Input variable combination	$\overline{\varpi}_K$	$\overline{\varpi}_K$	$\overline{\varpi}_K$	$\overline{\varpi}_K$
2	9.05%	9.57%	8.91%	9.13%
3	8.84%	9.91%	10.26%	10.82%
4	8.70%	10.00%	11.22%	13.78%
5*	9.79%	11.13%	8.95%	8.38%
6	8.18%	8.50%	9.07%	9.05%
7	7.18%	7.74%	9.28%	10.06%
8	8.46%	9.06%	9.60%	9.18%
9	7.26%	8.47%	10.58%	9.25%

Table 12: Table of outer-loop cross validation errors, $\overline{\varpi}_M$ for the noon report dataset

Optimum number of hidden units	5
Optimum input variable combination	7
$\overline{\varpi}_M$	7.02%

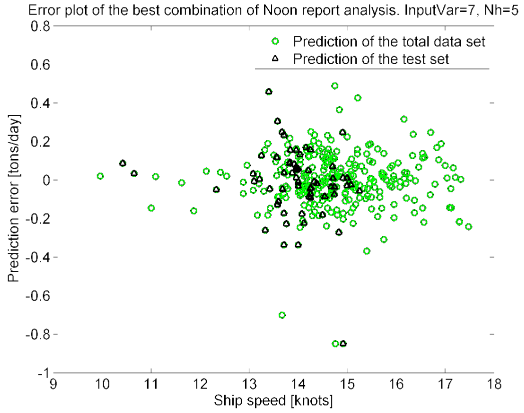

Fig.17: Prediction errors of noon report analysis

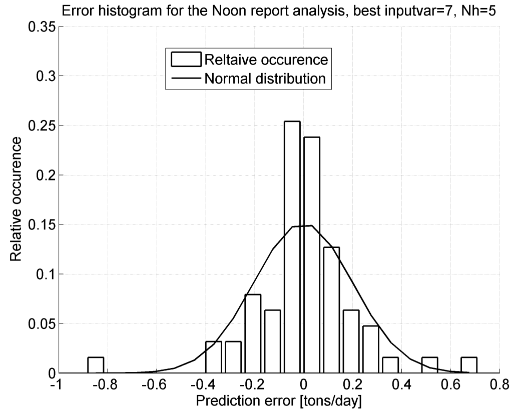

Fig.18: Relative distribution of the predicted errors from noon report analysis

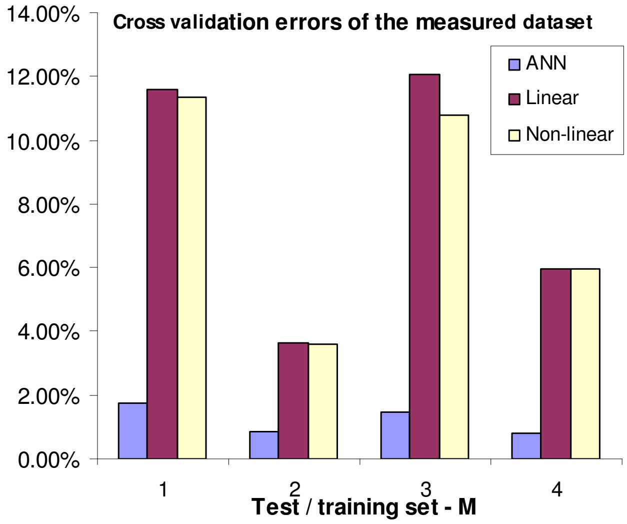

Fig.19: Comparison of different prediction methods

6.3. Comparison of the results

It is only possible to compare the dataset based on measured values, since it is the only one tested by other methods. Fig.19 shows the cross validation error for every condition and predicted by a linear, non-linear or ANN method. As previously mentioned dataset #2 gives significantly lower prediction errors for all methods due to the rather narrow ship speed band, the low wind speed and low number of samples.

7. Conclusion

Artificial Neural Networks (ANNs) can successfully be used to predict propulsion power, given that sufficient data are available. They have a significantly better performance than the linear and non-linear models tested. The propulsion power was predicted with an accuracy of less than 2% for the measured dataset. This accuracy is although of the same order of magnitude as the standard deviation of the propulsion power, so if a confidence interval analysis is introduced it is questionable if the method can get better.

ANNs can also be used with noon report data to predict the specific fuel consumption with an accuracy of about 7%, which is a bit surprising considering the rather rough input/output data. It is noted that this accuracy was obtained using “time” as an input variable, this indicates that it is possible to detect a trend of the fuel consumption over time.

It is shown that by introducing sea states and wind property information from the hind cast, the ANN solutions can be improved significantly, in the best case, from 2.97% to 1.65%. This eliminates the need for onboard measured wind speed and direction.

Unfortunately it was not possible to compare the solutions of the *four* different measured datasets with a solution using the noon report data from the same time, simply because of the lack of a sufficient number of noon reports for each dataset (there is only 3,4,7 and 9, see Table 1). Since the ship is not usually sailing in a single loading condition more that three weeks (21 noon reports), it will always be a problem to acquire enough data for making a reliable comparison of manual data acquisition (noon report) and automatic. If measured data for more loading conditions were available it would be possible to make an analysis similar to the one made for the noon reports.

Acknowledgements

Many thanks to the ship owner Torm and the crew onboard Torm Marie for allowing me to install the equipment and providing noon report data. In particular Chief Engineer Rasmus Hoffman, who has showed great interest in the project and been very helpful monitoring the onboard system and sending data home. Many thanks to Kjeld Roar Jensen from FORCE Technology who has been providing the hind cast data. The work presented in this paper has been carried out during Pedersen’s PhD study at FORCE Technology and the Technical University of Denmark, which is sponsored by The Danish Industrial PhD programme and Danish Centre of Maritime Technology (DCMT)/The Danish Maritime Fund.

References

PEDERSEN, B.P.; LARSEN, J. (2009), *Modeling of propulsion power*, WMTC 2009, Mumbai.

BISHOP, C.M. (2006), *Pattern Recognition and Machine Learning*, Springer.

FORCE (2008), SeaTrend® info sheet,
http://www.force.dk/en/Menu/Products+and+Concepts/Products/080220_seatrend.htm

HARVALD, S.A. (1983), *Resistance and Propulsion of Ships*, John Wiley & Sons.

HOLTROP, J. (1984), *A Statistical re-analysis of resistance and propulsion data*, International Shipbuilding Progress 31, pp.272-276.

ISHERWOOD, R. (1972), *Wind resistance of merchant ships*, RINA.

ITTC (1978), *ITTC – Recommended Procedures - Performance, Propulsion1978 ITTC Performance Prediction Method*, ITTC, 7.5-0203-01.4.

DTU toolbox (2002), *Neural regressor with quadratic cost function*, http://isp.imm.dtu.dk/toolbox

LARSEN, J. (1993), *Design of neural network filters*, PhD thesis, Technical University of Denmark.

MACKAY, D.J.C. (1992), *A practical Bayesian framework for backpropagation networks*, Neural Computation 4, pp.448-472.

MADSEN, K.; NIELSEN, H.B.; TINGLEFF, O. (2004), *Methods for non-linear least squares problems*, IMM, Technical University of Denmark.

NIELSEN, H.B. (1999), *Damping parameter in Marquardt's method,* Report IMM-REP-1999-05, IMM, Technical University of Denmark.

PEDERSEN, M. (1997), *Optimization of recurrent neural networks for time series modeling*, PhD thesis, IMM, Technical University of Denmark.

SVARER, C.; HANSEN, L.; LARSEN, J. (1993), *On design and evaluation of tapped-delay neural network architectures*, IEEE International Conference on Neural Networks, pp. 46-51.