

Towards Improved Prediction of Ship Performance: A Comparative Analysis on In-service Ship Monitoring Data for Modeling the Speed–Power Relation

 Simon DeKeyser*, Casimir Morobé*,  Malte Mittendorf†

* *Toqua – Ghent, Belgium*

† *Technical University of Denmark – Kgs. Lyngby, Denmark*

Abstract—Accurate modeling of ship performance is crucial for the shipping industry to optimize fuel consumption and subsequently reduce emissions. However, predicting the speed-power relation in real-world conditions remains a challenge. In this study, we used in-service monitoring data from multiple vessels with different hull shapes to compare the accuracy of data-driven machine learning (ML) algorithms to traditional methods for assessing ship performance. Our analysis consists of two main parts: (1) a comparison of sea trial curves with calm-water curves fitted on operational data, and (2) a benchmark of multiple added wave resistance theories with an ML-based approach. Our results showed that a simple neural network outperformed established semi-empirical formulas following first principles. The neural network only required operational data as input, while the traditional methods required extensive ship particulars that are often unavailable. These findings suggest that data-driven algorithms may be more effective for predicting ship performance in practical applications.

Index Terms—Ship Performance, Speed-Power, Added Wave Resistance, Resistance model, Neural Networks

1 INTRODUCTION

ACCURATE ship performance modeling is a powerful decision-making instrument for the shipping industry, offering the ability to save fuel and reduce emissions (Morobé, 2022), e.g., through optimized hull maintenance and voyage optimization. In the last decades, a spectrum of strategies has been developed, ranging from simplified empirical formulas derived from model tests to advanced full-scale 3-D numerical simulations (Tezdogan et al., 2015). However, considering both accuracy and computational effort, Liu and Papanikolaou (2020) argue that for practical applications where only a limited amount of ship parameters is available, a semi-empirical formula seems to be the most efficient method that captures the underlying physics of the problem. Unfortunately, the development of such methods leads to a loss of accuracy, which we classify into three categories.

(1) Approximations are made to generalize the actual hull shape to an efficient number of parameters, i.e., main particulars. The hull form and its displacement distribution significantly impact the magnitude of the wave resistance in calm water. Nevertheless, incorporating the full hull geometry into a semi-empirical model would reduce its practical advantage over full-scale numerical simulations.

(2) Assumptions about the range of validity need to be made to enforce regimes where certain physical relations hold. Extrapolating results outside particular validity ranges may result in poor accuracy and non-physical transitions.

(3) The theoretical expressions are fitted on experimental and numerical data obtained from controlled environments, which are far from

reality. Most of the experiments are performed on scaled models in towing tanks, causing an increase in errors when scaling the measurements back to full size. The problem lies in a lack of Reynolds number similarity between the ship and the model as only Froude number similarity is obtained. There, further assumptions are made on, e.g., hull smoothness and weather conditions, as imitating every possible ship/environment would be infeasible. Hence, we face data scarcity when assessing the in-service ship performance.

The overarching goal of this article is to show that machine learning (ML) approaches may complement or possibly surpass traditional methods in the assessment of in-service ship performance. ML is a subfield of artificial intelligence that involves the use of algorithms and statistical models to enable a system to learn from data and improve its performance on a specific task. It consists of a set of computational techniques to discover patterns and relationships in data and make predictions or decisions based on those patterns. ML algorithms can be thought of as an extension of statistical regression analysis, which is a technique used to model the relationship between a dependent variable (the outcome or response of interest) and one or more independent variables (the predictors or explanatory variables).

Whereas regression analysis typically involves fitting a fixed model to data and then using the model to make predictions or inferences about the data, ML algorithms can be trained on data and adapt as new data becomes available, allowing the system to learn and improve over time without being explicitly programmed for each task.

In the past, naval engineers have pursued finding universal semi-empirical formulas that rely on ship-specific inputs by performing fits (regression analysis) on experimental data to model ship performance. Instead, we propose training ship-specific ML models on real-world data while ensuring physical relevance without needing ship particulars. Nowadays, a lot of in-service sensor data is acquired, allowing us

to leverage the scalability of ML and accurately model at once all of the ship-specific intricacies that semi-empirical formulas cannot.

ML is a natural progression for the shipping industry's practical purposes, as it allows us to predict unobserved outcomes in high-dimensional parameter spaces. In fact, several studies have examined the use of ML in the maritime industry, such as fuel consumption prediction and optimization for Diesel engines (Parlak et al., 2006), condition-based maintenance of naval propulsion plants (Coraddu, Oneto, Ghio, et al., 2016), and modeling marine fouling speed loss (Coraddu, Oneto, Baldi, et al., 2019; Gupta et al., 2022).

In this work, the focus lies on predicting the power delivered by the engine as a function of the ship's speed. This so-called speed-power relationship forms the basis of ship performance modeling as it is the gateway to crucial economic parameters such as fuel consumption. An assessment of the current state of ship performance modeling research is made by comparing the accuracy of older, simpler theories with state-of-the-art (semi-)empirical formulas on in-service monitoring data. Herein, engine power sensor measurements will form our ground truth. We follow this approach to enable the extension to a supervised learning problem, where a learning algorithm is trained on observed data (ship speed, weather, and loading conditions) to predict specific targets (engine power) in unobserved situations.

2 METHODOLOGY

In the following sections, we use the ship resistance model according to the ISO 19030 (2019) industry standard to model the speed-power relationship. This model splits the resistance experienced by a ship into different parts:

$$R = R_{calm} + R_{AA} + R_{AW} + R_{AH} + R_{others} \quad (1)$$

where R_{calm} is the calm-water resistance, R_{AA} and R_{AW} respectively are the added resistances due to wind and waves, R_{AH} is the added resistance due to changes in hull condition such as the

accumulation of marine growth on hull and propeller, and R_{others} combines the effect of all other contributions such as steering and shallow water resistance.

Several (semi-)empirical techniques are elaborated on and critically examined for the different resistance contributions in section 3. To simplify our analysis, we will discuss the calculation of the first three resistance terms, which constitute the most considerable contribution to R (Dalheim and Steen, 2020). The other terms require knowledge of ship-specific and environmental parameters, which are often unavailable in real-world situations.

In this study, we select neural networks (NNs) to prove the value of ML algorithms in predicting ship performance. We chose NNs for their simplicity and effectiveness in demonstrating the potential of ML without adding unnecessary complexity to the comparison. In section 4, we provide a brief overview of the NN architecture, pre-processing techniques tailored for in-service vessel data, and the training and evaluation process.

Section 5 discusses a benchmark of different R_{AW} estimation procedures evaluated on in-service ship monitoring data to highlight the uncertainties of using experimentally fitted formulas for practical purposes. The benchmark, focusing on the resistance due to calm water, wind, and waves, will provide a clear and functional assessment of the accuracy of the (semi-)empirical techniques discussed in section 3.

The goodness of fit between the actual P_i and predicted \hat{P}_i engine powers (for the observations $i = 1, \dots, N$) is assessed with several metrics. We use the well-known mean absolute error (MAE):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |P_i - \hat{P}_i| \quad (2)$$

, the mean absolute percentage error (MAPE):

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{P_i - \hat{P}_i}{\hat{P}_i} \right| \quad (3)$$

TABLE 1: Input and target variables of the in-service dataset.

Category	Variables
Input	Draft aft, Draft forward, Displacement
	Wind direction, Wind speed
	Wave direction, Significant wave height, Wave period
	Current speed, Current direction
Target	Speed-through-water, Ship heading
	Brake power

, the mean bias error (MBE):

$$\text{MBE} = \frac{1}{N} \sum_{i=1}^N (\hat{P}_i - P_i) \quad (4)$$

, and the R-squared score (R2):

$$\text{R2} = 1 - \sum_{i=1}^N \frac{(P_i - \hat{P}_i)^2}{(P_i - \bar{P})^2} \quad (5)$$

with $\bar{P} = \frac{1}{N} \sum_{i=1}^N P_i$ the mean observed power.

The in-service dataset combines high-frequency sensor measurements of ship speed and main engine brake power^a with weather data and loading conditions. Table 1 summarizes the variables into input and target categories. The data is filtered to *steady state* voyages in full sea conditions, where ship accelerations, shallow water effects, and steering maneuvers are eliminated.

To further isolate our analysis to the R_{calm} , R_{AA} and R_{AW} resistance terms, measurements are taken from periods after cleaning events, such that R_{AH} can be disregarded. Note that the relative contribution of R_{AH} to R will increase over time and for practical use cases, accurately modeling R_{AH} is crucial. Although it remains a challenging task to solve using empirical models (Demirel et al., 2017; Guo et al., 2022), recent research has shown that digital twin ML-based methods may bring solace (Coraddu, Oneto, Baldi, et al., 2019; Gupta et al., 2022).

Finally, a comparison is made with a simple ship-specific NN model, which incorporates all

a. The brake power of a ship's engine is the power output of the engine measured at the engine's crankshaft, before any transmission losses. It can be calculated by measuring the torque and angular speed of the crankshaft.

of the resistance contributions while requiring only operational data as input. Here, we also consider some limitations of using ML for ship performance modeling. Section 6 will summarize our findings and use them to propose ML as the natural next step towards improved prediction of ship performance.

3 SHIP RESISTANCE MODEL

The brake power P_B of the ship's main engine is calculated according to:

$$P_B = \frac{RV_S}{\eta_D\eta_M} \quad (6)$$

with V_S the ship's speed-through-water (STW), η_D the propulsive efficiency, η_M the mechanical (shaft + gearbox) efficiency, and R the total resistance. As with R_{AH} , empirically modeling the propulsive and mechanical efficiencies with limited information remains difficult (Shigunov, 2017).

In particular, η_D depends on the wake fraction, thrust deduction, propeller diameter, and even the total resistance experienced by the hull. All these factors but the propeller diameter are speed and seaway dependent. Several curves and empirical models have been proposed to estimate η_D (Breslin and Andersen, 1995; Holtrop and Mennen, 1982; Kristensen and Lützen, 2012). However, these are only valid for calm water, and there is a lack of data in waves. Therefore, we use the default values recommended by ISO19030.

3.1 Calm-Water Resistance

A direct method for calculating a ship's calm-water resistance is to conduct a sea trial, during which the speed-power relation is measured under calm weather conditions. Usually, these measurements are performed around the design speed of the vessel. In recent years, however, operational velocities have been lowered to reduce emissions (Psaraftis and Kontovas, 2014), meaning that operating speeds do not always match the design speeds anymore.

It is widely accepted that the correlation between speed and power follows a relation $P \approx V^c$ with constant exponent $c \approx 3$, which holds around the design speed (MAN Energy Solutions, 2018).

However, previous studies have found that the *cubic law* underestimates power at speeds below the design speed (Adland et al., 2020; Taskar and Andersen, 2020; Tillig et al., 2018). These findings suggest that simply extrapolating sea trial curves to operational speeds lower than the design speed can lead to underestimations of power.

In this work, we will omit three of the main issues associated with the use of sea trial curves to estimate R_{calm} :

- Sea trial curve availability
- Operational speed \neq design speed
- Time-dependent performance loss (e.g., fouling)

by using the data-driven method proposed by Berthelsen and Nielsen (2021). This method fits a calm-water curve using in-service data. Their model starts from a simple least-squares fit of (x_1, x_2) to a log-linearized power law:

$$P = x_1 V^{x_2} \quad (7)$$

$$\ln(P) = \ln(x_1) + x_2 \ln(V) \quad (8)$$

They extend the regression model to incorporate the ship's draft T :

$$\ln(P) = \ln(x_1) + x_2 \ln(V) + x_3 T + x_4 \ln(V)T \quad (9)$$

and make the exponent speed dependent by introducing breakpoints that separate different speed intervals where the exponent remains constant. With one breakpoint B_p , the regression is performed on:

$$\ln(P) = \ln(x_1) + x_2 \ln(V) + x_3 T + x_4 \ln(V)T + x_5 (\ln(V) - \ln(B_p))V_d \quad (10)$$

In their paper, V_d is a Heaviside function centered at B_p . We, however, propose a differentiable dummy index to make the speed-power relation smooth:

$$V_d = \frac{1}{2} \left(1 + \tanh \left(\frac{V - B_p}{\delta} \right) \right) \quad (11)$$

with δ the smoothing factor.

In practice, the breakpoints are detected with a binary segmentation algorithm, as implemented in the `ruptures` Python library (Truong et al.,

2020). The algorithm is applied to the measured power of the speed-sorted data, and it finds change points in the signal where the slope of speed and power alters. At last, the regression is performed on the power data, which is first corrected from weather conditions (see next sections) to ensure that we fit calm-water data. As we do not correct the fouling power loss, a part of $R_{fouling}$ will be included in the fitted calm-water curves.

3.2 Added Resistance due to Wind

The industry standard for calculating R_{AA} is: (ISO, 2015)

$$R_{AA} = \frac{1}{2}\rho_A A_{XV} C_{AA}(\theta_{rel}) V_{wrel}^2 - \frac{1}{2}\rho_A A_{XV} C_{AA}(0^\circ) V_G^2 \quad (12)$$

with ρ_A the air density, V_{wrel} the relative wind speed calculated according to ISO 15016, V_G the ship's measured speed over ground, A_{XV} the transverse projected area of the ship above the water line, C_{AA} the wind resistance coefficient and θ_{rel} the relative wind direction. The negative term is the air resistance due to the ship moving forward with a headwind caused by V_G ($\theta_{rel} = 0^\circ$), which is already included in R_{calm} .

C_{AA} is estimated with the regression formula based on wind tunnel tests developed by Fujiwara (2006):

$$C_{AA}(\theta_{rel}) = C_{LF} \cos(\theta_{rel}) + C_{XLI} \left(\sin(\theta_{rel}) - \frac{1}{2} \sin(\theta_{rel}) \cos^3(\theta_{rel}) \right) + C_{ALF} \sin(\theta_{rel}) \cos^3(\theta_{rel}) \quad (13)$$

where the coefficients C_{LF} , C_{XLI} , and C_{ALF} consist of different regression expressions for $\theta_{rel} < 90^\circ$ and $\theta_{rel} > 90^\circ$. Unfortunately, the latter coefficients depend on several detailed ship geometry-related parameters, such as the bridge height and longitudinal projected area of superstructures, which are usually unknown.

However, Kitamura et al. (2017) developed ship-type-specific regression formulas that estimate the input parameters P of Fujiwara's formula and

A_{XV} from the ship's overall length L_{OA} and beam B :

$$\left. \begin{array}{l} P \\ P/L_{OA} \\ P/B \\ P/L_{OA}^2 \\ P/(L_{OA}B) \\ P/B^2 \end{array} \right\} = \begin{cases} aB + bL_{OA} + c \\ aB + c \\ bL_{OA} + c \end{cases} \quad (14)$$

with (a, b, c) the regression coefficients and where the left-hand-side and right-hand-side expressions were carefully chosen for each specific parameter to maximize the accuracy. In the following, this approach is used to calculate the inputs for Fujiwara's expression, and while it introduces additional approximations, it is necessary to keep the required parameters at a feasible level.

Fig. 1 shows R_{AA} calculated according to this method for a bulk carrier ($L_{OA} = 190$ m, $B = 32$ m) with $V_{wrel} = 8$ m/s and $V_G = V_S = 13$ kn, where a distinction is made between laden and ballast loading conditions. The absolute value of R_{AA} is higher for ballast conditions, which is expected as the area above water is larger. Also, the added resistance is more considerable for headwinds and drops to negative values for following winds, meaning that the ship is being pushed forward. The reader may find it counter-intuitive that $R_{AA} = 0$ kN occurs at angles of only $\pm 40^\circ$, but it is stressed that we plot the added resistance relative to the ship's headwind (Eq. 12).

3.3 Added Resistance due to Waves

Modeling a ship's added resistance due to waves is a highly complex, non-linear problem, and most methods rely on simplified assumptions. In practice, R_{AW} is defined as "the unsteady longitudinal force a ship experiences apart from the calm water, and wind resistances in a realistic seaway" (Mittendorf, Nielsen, and Bingham, 2022).

The force is a second-order quantity, which depends on the incident wave's amplitude and speed (Liu, Papanikolaou, and Zaraphonitis, 2011). One

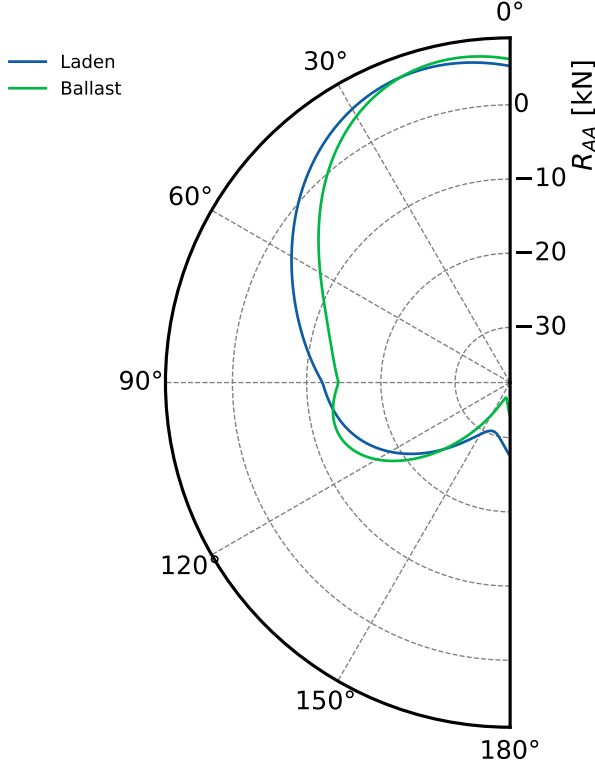


Fig. 1: Polar plot of the added resistance due to wind as a function of the relative wind direction, calculated with Fujiwara's and Kitamura's regression formulas for different ship load conditions.

of the oldest and simplest modeling approaches is that of Kreitner, used in ITTC2005 (2005):

$$R_{AW} = 0.64gH_S^2C_B\rho_w \frac{B^2}{L_{pp}} \quad (15)$$

with g the gravitational constant, H_S the significant wave height, C_B the block coefficient, ρ_w the water density, and L_{pp} the length between perpendiculars. As this expression only holds for head waves, a cosine law can be used to account for waves with an arbitrary heading (Hansen et al., 2011):

$$R_{AW} = 0.64gH_S^2C_B\rho_w \frac{B^2}{3L_{OA}} (2 + \cos(\alpha_{rel})) \quad (16)$$

with α_{rel} the relative wave direction.^b

Another simple, but more recent standard of ITTC 2014 (ITTC, 2014), is the empirical STAwave-1 formula:

$$R_{AW} = \frac{1}{16}gH_S^2\rho_w B \sqrt{\frac{B}{L_B}} \quad |\alpha_{rel}| \leq 45^\circ \quad (17)$$

^b We use $\alpha_{rel} = 0^\circ$ for head waves, and $\alpha_{rel} = 180^\circ$ for following waves.

where L_B is the length of the bow at the waterline.

More advanced (semi-)empirical approaches split the added wave resistance into two contributions:

$$R_{AW} = R_{AWM} + R_{AWR} \quad (18)$$

where the motion induced R_{AWM} and wave reflection R_{AWR} resistances are a function of the wave frequency ω . To accurately model the ship's resistance in irregular sea conditions with waves of varying frequencies, integration is performed over a wave spectrum $S(\omega)$ to calculate the mean added resistance:

$$\bar{R}_{AW} = 2 \int_0^\infty S(\omega) \frac{R_{AW}(\omega)}{\zeta_a^2} d\omega \quad (19)$$

with ζ_a the wave amplitude. Hereby, we assume that the superposition principle holds and that the calculation is valid for long-crested waves. Often, a Pierson-Moskowitz type spectrum is used in fully developed sea states:

$$S(\omega) = \frac{5}{16} H_S^2 \frac{\omega_p^4}{\omega^5} \exp\left(-\frac{5}{4} \left(\frac{\omega_p}{\omega}\right)^4\right) \quad (20)$$

where $\omega_p = 2\pi/T_p$ and T_p is the wave peak period. Using this frequency response framework, ITTC2014 (2014) and ISO 15016 (2015) recommend using the semi-empirical method called STAwave-2, which again holds for $|\alpha_{rel}| \leq 45^\circ$.

From 2016 to 2020, Liu and Papanikolaou (2016; 2020) derived an improved formula with regression analysis on model test data, which holds for arbitrary wave headings. In 2022, Mittendorf et al. (2022) enhanced their formula by performing multivariate regression on the parameter vector with model test data, thereby also including the 90 % prediction interval of R_{AW} . We note that a discontinuity is present in the Liu/Mittendorf formulas, which we found to be caused by the non-linear form factor in the expression for R_{AWR} :

$$\left(\frac{0.87}{C_B}\right)^{(1+4\sqrt{Fr})f(\alpha_{rel})} \quad (21)$$

with Fr the Froude number and:

$$f(\alpha_{rel}) = \begin{cases} -\cos(\alpha_{rel}) & \pi - E_1 \leq \alpha_{rel} \leq \pi \\ 0 & \alpha_{rel} < \pi - E_1 \end{cases} \quad (22)$$

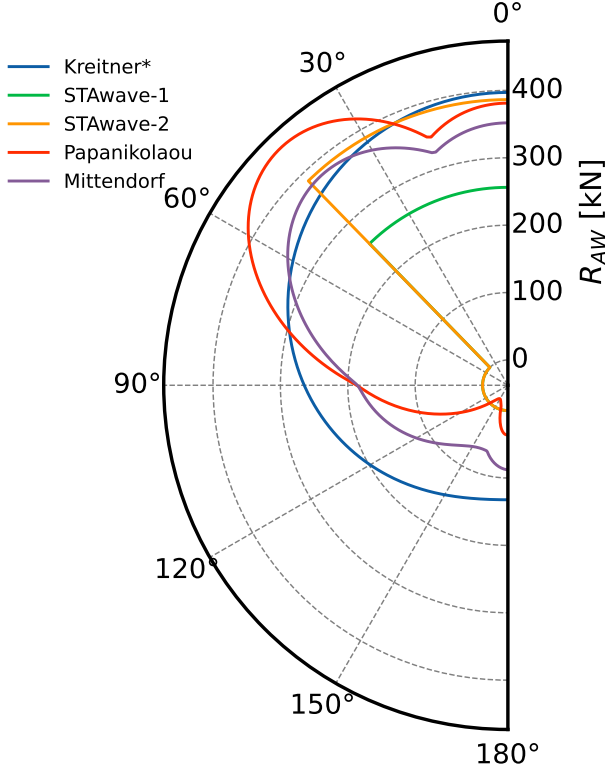


Fig. 2: Polar plot of the added resistance due to waves as a function of the relative wave direction, calculated with different wave resistance theories in laden condition.

where E_1 is an angle that defines where the ship's bow ends. The discontinuity can be solved by shifting the cosine of $f(\alpha_{rel})$ to 0 when $\alpha_{rel} = \pi - E_1$, while still forcing $f(\alpha_{rel}) = 1$ for $\alpha_{rel} = \pi$:

$$f(\alpha_{rel}) = \begin{cases} -\frac{\cos(\alpha_{rel})+1}{\cos(\pi-E_1)+1} + 1 & \pi - E_1 \leq \alpha_{rel} \leq \pi \\ 0 & \alpha_{rel} < \pi - E_1 \end{cases} \quad (23)$$

We will use this new form of $f(\alpha)$ in the R_{AWR} expression of the Liu and Mittendorf methods, as discontinuities will not be present in the measured data.

To give the reader a visual understanding of the different wave resistance formulas, Fig. 2 plots R_{AW} calculated with the aforementioned methods for a bulk carrier ($L_{OA} = 190$ m, $B = 32$ m, $C_B = 0.7$) with $H_S = 4$ m, $T_p = 10$ s and $V_G = V_S = 13$ kn in laden condition. Given Fig. 2, it is stated that while the theories' R_{AW} predictions are in the same order of magnitude, their shapes vary differently as a function of α_{rel} .

Although the discontinuity has been resolved, the semi-empirical Liu and Mittendorf methods still exhibit non-differentiable irregularities where the resistance increases or decreases. The reason is that their formula consists of expressions for the bow and stern of the ship, including form factors to incorporate non-linear behavior. Of course, in natural seas, the waves will come from multiple directions at once, and Eq. 19 can be extended to a double integration over an angular wave spectrum. Accounting for short-crested waves will smooth the $R_{AW}(\alpha_{rel})$ curves to behave as one would expect physically. Unfortunately, performing a two-dimensional numerical integration comes at the cost of computational time.

Another observation is that the more advanced methods predict a larger wave resistance for oblique waves compared to head waves. The latter effect is confirmed in model tests, where oblique waves cause more pitch motions dissipating energy (Valanto and Hong, 2015). The magnitude of the wave angle-dependent Kreitner's formula remains similar for the different wave headings and lies somewhere between the STAwave and the Mittendorf methods.

Lastly, to highlight the importance of the peak wave period on the frequency response-type methods, Fig. 3 shows a surface plot of R_{AW} as a function of T_p and the wave heading α_{rel} calculated with the Mittendorf method. The same values were used as in Fig. 2 for B , H_S , etc. and $V_G = V_S = 0$ kn. The ship experiences the highest added resistance in sea states with short waves that arrive obliquely to the ship's heading.

4 AN ML-BASED APPROACH

Neural networks (NNs) are ML models inspired by the structure and function of the human brain. They are composed of interconnected processing nodes, called neurons, which work together to recognize patterns in data and make predictions. Here, we will use one of the most straightforward NN architectures, the feedforward NN (FNN). This section will briefly explain the architecture of the FNN, the data pre-processing steps, and how

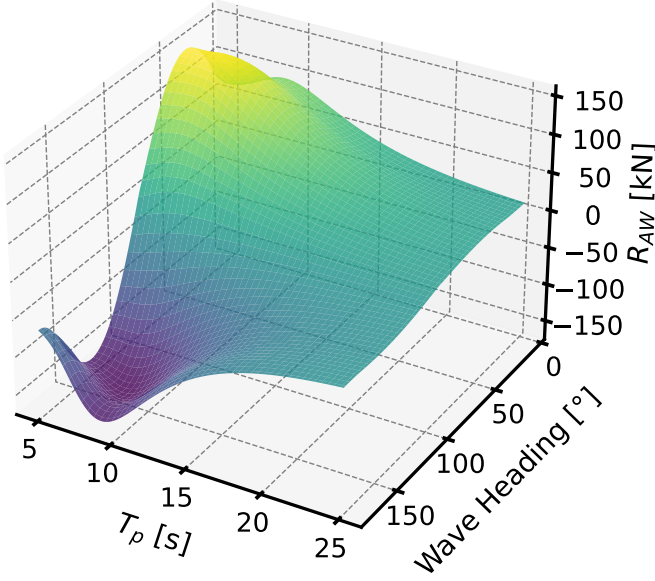


Fig. 3: Surface plot of R_{AW} as a function of T_p and the wave heading α_{rel} calculated with the Mittendorf method in laden condition.

a standard training procedure is performed and validated.

4.1 The Feedforward Architecture

An FNN is organized into three main parts: the input layer, the hidden layers, and the output layer. The input layer receives input data, the hidden layers process that data and the output layer yields the final output of the network. Each layer consists of several neurons, which receive input from the previous layer and produce the output that is conveyed to the next layer.

Mathematically, a neuron can be represented as a function that takes in a set of inputs, x_1, x_2, \dots, x_n , and produces a single output, y . The inputs are multiplied by a set of weights, w_1, w_2, \dots, w_n , and summed with a bias term, b , to produce an intermediate value, z , which is then passed through a non-linear activation function, f , to produce the output:

$$z = \sum_{i=1}^n w_i x_i + b \quad (24)$$

$$y = f(z) \quad (25)$$

A commonly used activation function is the rectified linear unit (ReLU):

$$f(z) = \max(0, z) \quad (26)$$

which has the disadvantage of being non-differentiable at $z = 0$. In this work, we use a smooth approximation of the ReLU function, namely the softplus activation function, defined as:

$$f(z) = \log(1 + e^z) \quad (27)$$

The main premise of activation functions is to replicate the behavior of brain neurons, which fire when the electrical impulses coming from connections with other neurons reach a certain threshold (at $z = 0$ here).

4.2 Feature Engineering

The input features are the variables presented to the network's input layer. The process of selecting and transforming the input features to improve the performance of a NN is called feature engineering. It involves identifying the most relevant features in the dataset and pre-processing them in a way that makes them more suitable for the NN to learn from.

In this work, we deliberately reduce the NN's complexity by using as few features as possible without losing predictive power by combining different input variables from table 1. Table 2 summarizes the engineered features, which mimic theoretical formulas of wave and wind resistance. E.g., the *Wind Product* feature resembles the form of Eq. 12, while the *Wave Power* mimics the power per unit wave crest in deep water:

$$P = \frac{\rho_w g}{64\pi} H_S^2 T_e \quad (28)$$

with T_e the wave energy period.

Although power is not a vector quantity, we separate all directional variables into their longitudinal (long.) and transversal (trans.) direction to the vessel's heading. Doing so results in a correct angular dependence of model w.r.t. the wind direction θ_{rel} and wave direction α_{rel} relative to the vessel's heading.

The final FNN architecture consists of an input layer with six features, four fully-connected hidden layers (64, 32, 16, and 8 neurons, respectively), and an output layer that predicts the brake power.

4.3 Training and Validation

During training, the weights and biases of the neurons are adjusted to minimize a loss function that measures the difference between the predicted output and the true output (e.g., with the MAE (Eq. 2) or mean squared error (MSE)). The optimization is generally conducted with algorithms like stochastic gradient descent, which calculates the gradient of the loss function with respect to the weights and biases and updates the weights and biases in the direction that reduces the loss.

In this study, we employed ADAM (Kingma and Ba, 2014), a frequently used variation of gradient descent that incorporates momentum (Rumelhart et al., 1985) to accelerate convergence to the minimum of the loss function in the high-dimensional space of weights and biases.

To assess the performance of the NN, we use k-fold cross-validation. Here, the data is divided into k *folds*, and the model is trained on k-1 of those folds while leaving the remaining fold as a validation set. This process is repeated k times, with a different fold being used as the validation set each time. The final performance of the model is then calculated as the average performance across all k iterations. This approach ensures that the NN is tested on completely unseen data, giving us a reliable representation of how the model would perform in unobserved real-world conditions.

TABLE 2: Features used as input of the NN.

Feature	Formula
STW	
Draft Average	$(T_{aft} + T_{fwd})/2$
Wind Product (long.)	$V_{wrel}^2 \cos(\theta_{rel})$
Wind Product (trans.)	$V_{wrel}^2 \sin(\theta_{rel})$
Wave Power (long.)	$H_S^2 T_p \cos(\alpha_{rel})$
Wave Power (trans.)	$H_S^2 T_p \sin(\alpha_{rel})$

5 DISCUSSION

In the following, we will use in-service data of different vessels to compare the performance of the resistance models described above and extend the analysis to include a simple NN. Before proceeding, we will first justify the use of a speed-dependent exponent in the calm-water speed-power relation.

5.1 Calm-Water Resistance

Sea trial curves of a tanker, were extrapolated to lower speeds, and weather corrections R_{AA} , R_{AW} were calculated with Fujiwara's and Mittendorf's semi-empirical formulas, respectively. The main engine power is then obtained with Eq. 6, ignoring R_{AH} and R_{others} . Fig. 4 shows the predicted and measured power as a function of STW, along with the extrapolated laden and ballast sea trial curves. Predicted points \hat{P} are colored based on their absolute percentage error (APE) = $|\hat{P} - P|/P$ with respect to the measured values P .

The scatter plot in Fig. 4 supports our earlier hypothesis that extrapolating sea trial curves below the design speed leads to an underestimation of the power, indicating that the cubic law does not hold in this range. Additionally, time-dependent performance losses such as fouling can also cause an underestimation of the power, which increases over time.

To address the limitations of the extrapolated sea trial curves, we perform a least-squares fit of Eq. 10 on calm-water data according to the method by Berthelsen and Nielsen. The calm-water data is obtained by correcting the $H_S \leq 1$ m in-service power data for wind and waves with Fujiwara's and Mittendorf's formulas.^c First, the binary segmentation algorithm was used to find the breakpoint at $V_S = 11.53$ kn. Fig. 5 shows the results in the same way as before, but with the fitted calm-water curves. The speed-power exponent changes (smoothly ($\delta = 0.5$)) at the breakpoint from ≈ 1.8 to ≈ 2.8 for the laden curve.

Table 3 compares the accuracy metrics obtained

c. A histogram of H_S can be found in Fig. 10.

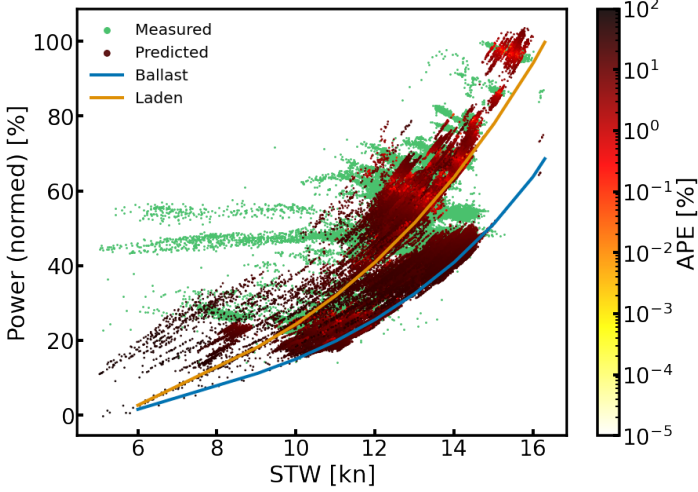


Fig. 4: Scatter plot of the measured and predicted main engine power (normalized w.r.t maximum) as a function of STW, together with the extrapolated laden and ballast sea trial curves. (Log-scaled color map.)

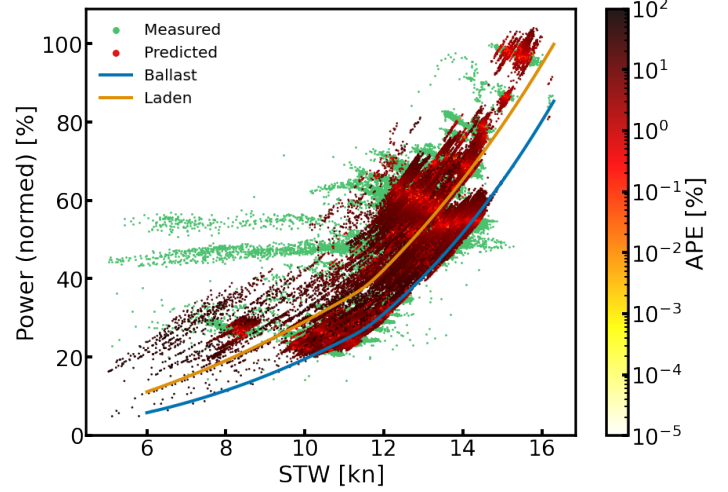


Fig. 5: Scatter plot of the measured and predicted main engine power (normalized w.r.t maximum) as a function of STW, together with the fitted calm-water laden and ballast curves. (Log-scaled color map.)

with the sea trial curves and the fitted calm-water model. Our analysis shows that the data-driven calm-water method significantly improves the accuracy compared to extrapolated sea trial curves in terms of R^2 , MAPE, and MBE. This suggests that the data-driven calm-water model is a more reliable approach for predicting power requirements in realistic operating conditions. However, as shown in Fig. 5, the fitted calm-water model still tends to underestimate the power (negative MBE) at low STW. This error is likely due to the presence of *bands* in the speed-power data, where the vessel encounters heavy weather while operating in constant power or constant shaft rpm autopilot, resulting in large power values at low STW.

5.2 Added Wave Resistance Benchmark

Up to this point, we have only used Mittendorf’s formula to calculate the power corrections due to waves. In the following analysis, we perform

TABLE 3: Accuracy metrics evaluated on the predictions from different calm-water models in combination with Mittendorf’s and Fujiwara’s resistance corrections.

Calm-Water Model	R^2	MAPE [%]	MBE [kW]
Extrapolated Sea Trial Curve	0.72	15	-1522
Fitted Calm-Water Curve	0.89	7.1	-443

a benchmark of different added wave resistance theories to evaluate their performance, along with sea trials and the fitted calm-water regression model, using data from both an oil tanker and a dry cargo carrier. The method is similar to that used to produce Figs. 4 and 5, except that different wave theories are used in combination with Fujiwara’s R_{AA} . Fig. 6 shows the results, including the MAPE scores for each vessel and wave resistance theory, along with a comparison of the fitted calm-water model to sea trials.

Interestingly, we find that the accuracy of all the added resistance theories is very similar, despite significant differences in their complexity. One possible explanation for this finding is that the more advanced approaches are too complex for noisy data. When examining Fig. 2, we see that the Liu theory exhibits abrupt changes (non-differentiable points) in added wave resistance, which are not physically realistic. Based on these results, one might even consider using the wave angle-dependent Kreitner’s formula, which has the advantage of requiring fewer ship-specific parameters and can be evaluated approximately ten times faster than Mittendorf’s method (≈ 0.1 ms versus ≈ 1 ms).

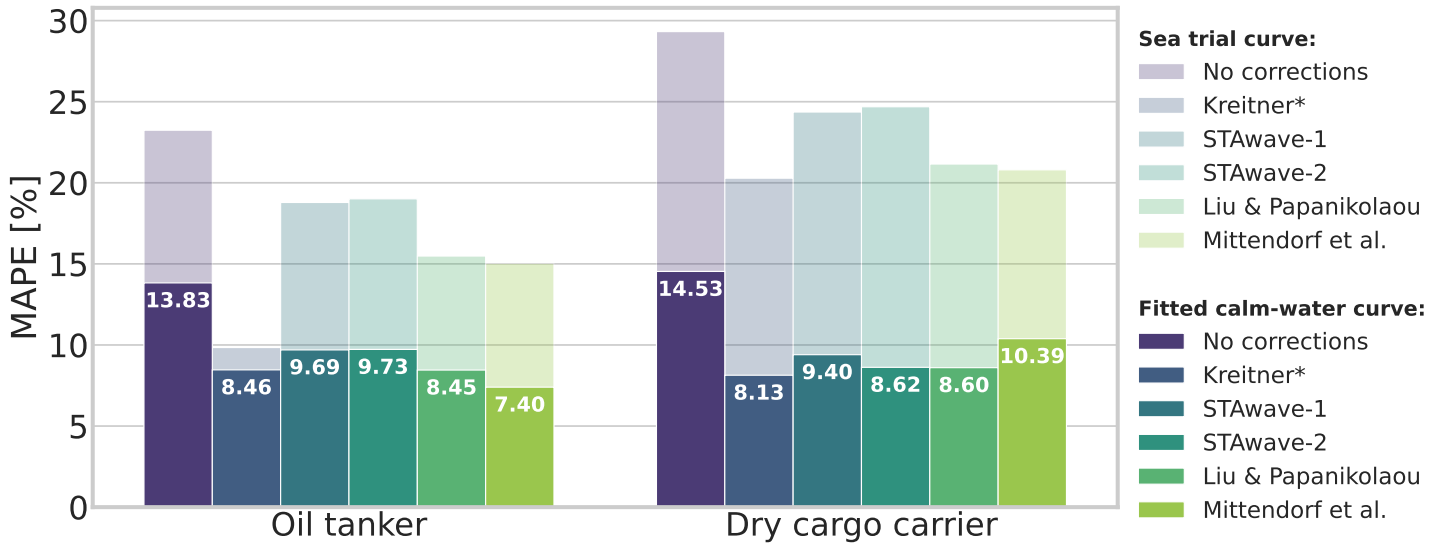


Fig. 6: MAPE obtained with the different combinations of wave resistance theories and calm-water models evaluated on multiple vessels.

5.3 A Simple NN

In this section, we evaluate the performance of the simple NN by comparing it with the fitted calm-water model using different added wave resistance theories on the in-service data of a chemical tanker.^d In contrast to the previous vessels, including sea trial curves in the analysis is not possible as they are unavailable. This further highlights the benefit of using the data-driven calm-water model by Berthelsen and Nielsen.

Fig. 7 compares the MAPE score achieved with the fitted calm-water curve using different wave resistance theories and the simple NN. The Kreitner and STAwave-2 theories perform worse than the model without weather corrections due to multiple data points falling outside the valid ranges for these empirical formulas. The simple NN achieves a MAPE of 7.4 %, which represents a roughly 3 % improvement over the best (semi-)empirical formula. This is a remarkable result considering the simplicity of our NN architecture. It is expected that further increasing the complexity and effort in developing the NN architecture would significantly improve the accuracy.

To further analyze the performance of these

d. The FNN was trained using mini-batches of 32 samples with a learning rate of 0.015 for five epochs on each k-fold.

models, we can perform an error analysis using binned scatter plots. In these plots, the data points are grouped into bins according to a variable of interest. Then, the MAPE is calculated for each bin, which allows us to visualize the relative accuracy of the model as a function of the variable of interest, providing insight into the factors that may influence the model's performance.

Figs. 8 and 9 show a binned scatter plot with respect to the significant wave height and wind speed, respectively. Here, we compare the fitted calm-water model using Mittendorf's

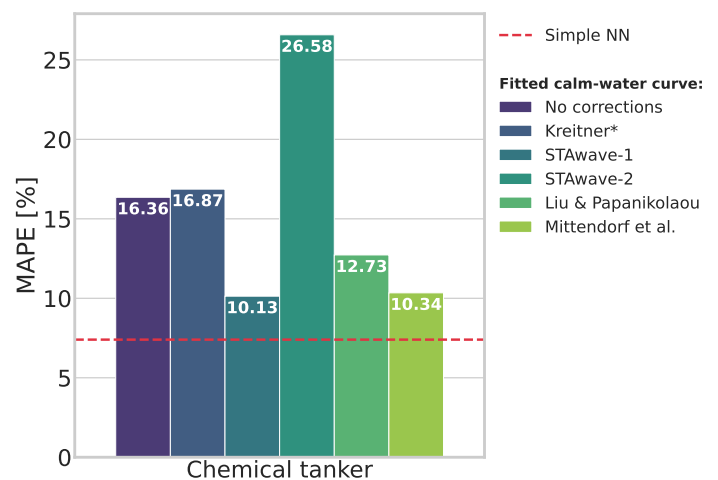


Fig. 7: MAPE obtained with the different combinations of wave resistance theories and the fitted calm-water model. A comparison is made with a simple NN.

added wave resistance corrections with the simple NN from Fig. 7. The histograms show the data spread in the variable of interest, and the error bars denote the standard deviation of the binned MAPE. From both plots, we conclude that the NN achieves a better accuracy *globally*, and thus it generalizes better than the theoretical model.

Taking the significant wave height as the variable of interest, the theoretical model performs particularly poorly in bad weather conditions, with a binned MAPE reaching over 70 %. This effect is not as noticeable when evaluating the overall MAPE because the data spread gives increased weight to low wave heights. Similarly, using the relative wind speed as the variable of interest yields similar results for the NN, while the theoretical model performs worse for a wide range of wind speeds with a binned MAPE of up to 45 %.

It is notable that, despite having less data available for high wave heights and wind speeds, the NN’s accuracy remains approximately in a similar range. This finding is remarkable because NNs generally require considerable amounts of

training data to learn and generalize effectively. The fact that the NN can maintain a similar level of accuracy with fewer data suggests that our choices of feature engineering were effective.

5.4 Machine Learning as the Next Step

Our findings above demonstrate the potential of ML methods to improve the prediction of ship performance in various operating conditions compared to traditional methods. However, ML has previously been proposed as a valuable tool for shipping companies. Alexiou et al. (2022) have compared a number of the latest data-driven models used in papers to model the speed-power relation and concluded that data-driven models were more suitable for shipping companies, provided sufficient representative historical data was available.

The key aspect of this work lies in the direct comparison of the data-driven methods and multiple theoretical models on several vessels with different design types. However, before we propose ML as the natural next step for ship performance modeling, we must carefully weigh

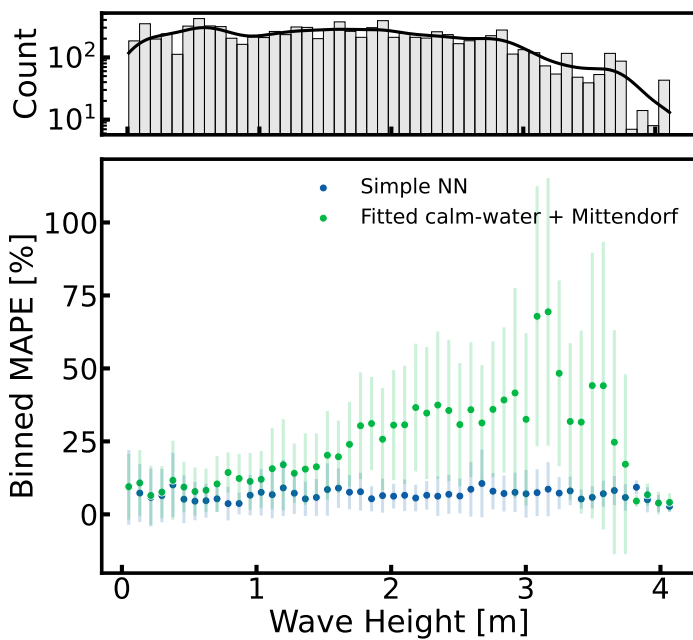


Fig. 8: MAPE binned w.r.t. the wave height, evaluated on a chemical tanker. A comparison is made between the fitted calm-water model in combination with Mittendorf’s wave resistance and a simple NN. The histogram shows the data spread of the wave height.

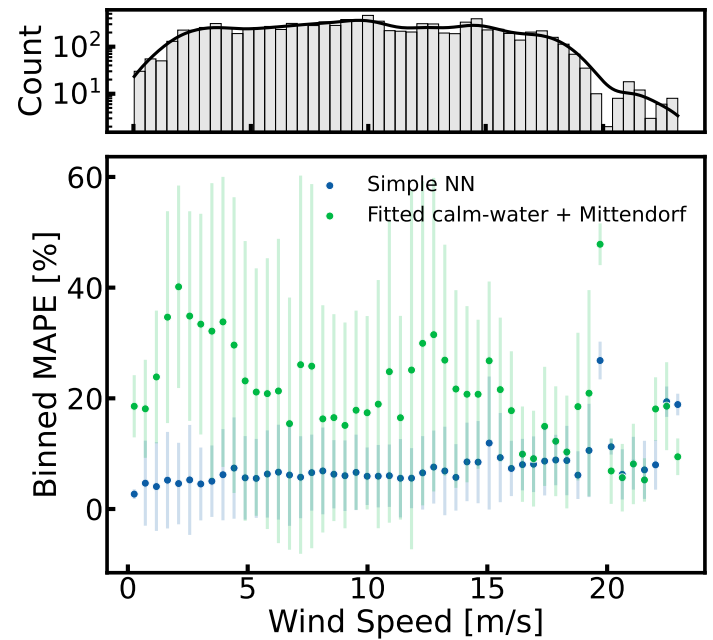


Fig. 9: MAPE binned w.r.t. the wind speed, evaluated on a chemical tanker. A comparison is made between the fitted calm-water model using Mittendorf’s wave resistance and a simple NN. The histogram shows the data spread of the wind speed.

the benefits against the limitations.

One of the main benefits of ML models is that they can be trained on existing data from a ship's performance, which means that they can be customized to a specific vessel and its operating conditions. This can provide more accurate and relevant results than using traditional methods.

Another advantage of ML is that it can automate some of the more tedious and time-consuming aspects of ship performance modeling, freeing engineers to focus on other tasks. Machine learning algorithms can also learn from data in real-time, which means that they can adapt to changes in a ship's operating conditions and provide updated predictions on an ongoing basis.

In contrast, one of the main constraints is the need for substantial amounts of high-quality data to train the models. This may be challenging for some ships or operating conditions where data may be limited or difficult to collect (e.g., only noon report data available or sensor drift).

Additionally, ML models can be difficult to interpret and explain, making it challenging for engineers to apply the results to their work. This issue can be addressed through the use of explainable artificial intelligence (XAI) techniques (Gunning et al., 2019) and physics-informed ML (PIML) (Karniadakis et al., 2021). XAI helps to make ML algorithms more transparent and interpretable, while PIML allows ML methods to follow physical relations.

Furthermore, ML algorithms can be subject to bias if the training data is not representative of the full range of conditions that a ship may encounter. This can lead to inaccurate or misleading results, which can be challenging to detect and correct. However, careful feature engineering and the use of PIML can help to ensure that the ML algorithm follows physical relationships beyond the range of the training data and is not biased by the data used to train it. This can help to improve the robustness and generalizability of the ML model (Vanackere et al., 2022).

Overall, while the use of ML in ship performance modeling offers many potential benefits, there are also significant limitations that one must consider. However, as the availability of data increases and the expertise in ML continues to grow, these restrictions are becoming easier to overcome.

6 CONCLUSION

We have outlined several (semi-)empirical approaches developed in the last decades of ship hydrodynamics. Most of these methods use theoretical approximations, after which they are fitted to experimental data. Researchers have been using increasingly advanced regression algorithms to enhance the experimental fits of their semi-empirical formulas.

Although their findings are crucial to obtaining a better understanding of ship performance in a research-based way, the developed methods might not be the best choice in practical applications as they require extensive ship particulars that are often unavailable.

Our benchmark of different added wave resistance theories, performed on in-service data from multiple vessels, supports this idea. Results showed that simple theories had similar accuracy to more complex methods, and that a data-driven calm-water resistance regression method proposed by Berthelsen and Nielsen (2021) was more useful than extrapolated sea trial curves for performance monitoring.

A natural next step to further improve accuracy and reduce the need for empirical methodologies and ship particulars, may be to leverage the data scalability of ML techniques. By applying ML techniques to ship-specific sensor data, the whole ship and its environment can be *learned* directly from real-world data.

Our results showed that a simple NN outperforms all of the (semi-)empirical added wave resistance methods, which were used in combination with the data-driven calm-water model. This suggests that ML techniques may be a valuable tool for improving ship performance modeling in practical applications. Even a small increase in accuracy

could have significant economic and ecological benefits for the shipping industry. It requires a tiny change of perspective – transitioning from theoretical regression analysis to a more advanced approach – potentially yielding significant benefits.

REFERENCES

- Adland, R., Cariou, P., & Wolff, F.-C. (2020). Optimal ship speed and the cubic law revisited: Empirical evidence from an oil tanker fleet. *Transportation Research Part E: Logistics and Transportation Review*, 140, 101972.
- Alexiou, K., Pariotis, E. G., Leligou, H. C., & Zannis, T. C. (2022). Towards data-driven models in the prediction of ship performance (speed—power) in actual seas: A comparative study between modern approaches. *Energies*, 15(16), 6094.
- Berthelsen, F. H., & Nielsen, U. D. (2021). Prediction of ships' speed-power relationship at speed intervals below the design speed. *Transportation Research Part D: Transport and Environment*, 99, 102996.
- Breslin, J. P., & Andersen, P. (1995). Hydrodynamics of ship propellers. *Journal of Fluid Mechanics*, 293, 405–405.
- Coraddu, A., Oneto, L., Baldi, F., Cipollini, F., Atlar, M., & Savio, S. (2019). Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. *Ocean Engineering*, 186, 106063.
- Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., & Figari, M. (2016). Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 230(1), 136–153.
- Dalheim, Ø. Ø., & Steen, S. (2020). Added resistance and speed loss of a ship found using onboard monitoring data. *Journal of Ship Research*, 64(02), 99–117.
- Demirel, Y. K., Uzun, D., Zhang, Y., Fang, H.-C., Day, A. H., & Turan, O. (2017). Effect of barnacle fouling on ship resistance and powering. *Biofouling*, 33(10), 819–834.
- Fujiwara, T. (2006). A new estimation method of wind forces and moments acting on ships on the basis of physical components models. *Journal of the Japan society of naval architects and ocean engineers*, 2, 243–255.
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). Xai—explainable artificial intelligence. *Science robotics*, 4(37), eaay7120.
- Guo, B., Liang, Q., Tvette, H. A., Brinks, H., & Vanem, E. (2022). Combined machine learning and physics-based models for estimating fuel consumption of cargo ships. *Ocean Engineering*, 255, 111435.
- Gupta, P., Rasheed, A., & Steen, S. (2022). Ship performance monitoring using machine-learning. *Ocean Engineering*, 254, 111094.
- Hansen, S. V., Petersen, J., Jensen, J., & Lutzen, M. (2011). *Performance monitoring of ships*. DTU Mechanical Engineering.
- Holtrop, J., & Mennen, G. (1982). An approximate power prediction method. *International Shipbuilding Progress*, 29(335), 166–170.
- ISO. (2015). Ships and marine technology — guidelines for the assessment of speed and power performance by analysis of speed trial data. *INTERNATIONAL STANDARD, ISO15016*.
- ISO. (2019). Ships and marine technology — measurement of changes in hull and propeller performance. *INTERNATIONAL STANDARD, ISO19030*.
- ITTC. (2005). The specialist committee on stability in waves, final report and recommendations to the 24th ittc. *Proc. 24th ITTC*, 2, 369–407.
- ITTC. (2014). The specialist committee on stability in service, final report and recommendations to the 27th ittc. *Proc. 27th ITTC*.
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3(6), 422–440.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kitamura, F., Ueno, M., Fujiwara, T., & Sogihara, N. (2017). Estimation of above water structural parameters and wind loads on ships. *Ships and Offshore Structures*, 12(8), 1100–1108.
- Kristensen, H. O., & Lützen, M. (2012). Prediction of resistance and propulsion power of ships. *Clean Shipping Currents*, 1(6), 1–52.
- Liu, S., & Papanikolaou, A. (2016). Fast approach to the estimation of the added resistance of ships in head waves. *Ocean Engineering*, 112, 211–225.
- Liu, S., & Papanikolaou, A. (2020). Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. *Ocean Engineering*, 206, 107357.
- Liu, S., Papanikolaou, A., & Zaraphonitis, G. (2011). Prediction of added resistance of ships in waves. *Ocean Engineering*, 38(4), 641–650.
- MAN Energy Solutions. (2018). *Basic principles of ship propulsion*. Retrieved August 25, 2022, from https://www.man-es.com/docs/default-source/marine/tools/basic-principles-of-ship-propulsion_web_links.pdf?sfvrsn=12d1b862_10
- Mittendorf, M., Nielsen, U. D., & Bingham, H. B. (2022). Data-driven prediction of added-wave resistance on ships in oblique waves—a comparison between tree-based ensemble methods and artificial neural networks. *Applied Ocean Research*, 118, 102964.
- Mittendorf, M., Nielsen, U. D., Bingham, H. B., & Liu, S. (2022). Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading. *Ocean Engineering*, 251, 111040.
- Morobé, C. (2022). *3 examples of how improved ship performance modelling can save fuel*. Retrieved August 25, 2022, from <https://toqua.ai/whitepapers/3-examples-of-how-improved-ship-performance-modelling-can-save-fuel>
- Parlak, A., Islamoglu, Y., Yasar, H., & Egrisogut, A. (2006). Application of artificial neural network to predict specific fuel consumption and exhaust temperature for a diesel engine. *Applied Thermal Engineering*, 26(8-9), 824–828.
- Psaraftis, H. N., & Kontovas, C. A. (2014). Ship speed optimization: Concepts, models and combined speed-routing scenarios. *Transportation Research Part C: Emerging Technologies*, 44, 52–69.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation* (tech. rep.). California Univ San Diego La Jolla Inst for Cognitive Science.
- Shigunov, V. (2017). Added power in seaway. *Ship technology research*, 64(2), 65–75.
- Taskar, B., & Andersen, P. (2020). Benefit of speed reduction for ships in different weather conditions. *Transportation Research Part D: Transport and Environment*, 85, 102337.

- Tezdogan, T., Demirel, Y. K., Kellett, P., Khorasanchi, M., Incecik, A., & Turan, O. (2015). Full-scale unsteady rans cfd simulations of ship behaviour and performance in head seas due to slow steaming. *Ocean Engineering*, 97, 186–206.
- Tillig, F., Ringsberg, J. W., Mao, W., & Ramne, B. (2018). Analysis of uncertainties in the prediction of ships' fuel consumption—from early design to operation conditions. *Ships and Offshore Structures*, 13(sup1), 13–24.
- Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, 167, 107299.
- Valanto, P., & Hong, Y. (2015). Experimental investigation on ship wave added resistance in regular head, oblique, beam, and following waves. *The Twenty-fifth International Ocean and Polar Engineering Conference*.
- Vanackere, J., Gautama, S., & Morobé, C. (2022). *Robuustheid van fysisch-geinformeerde schip-prestatie modellen op basis van machinaal leren voor maritieme reisoptimalisatie*. Retrieved December 21, 2022, from <https://lib.ugent.be/catalog/rug01:003063868>

APPENDIX

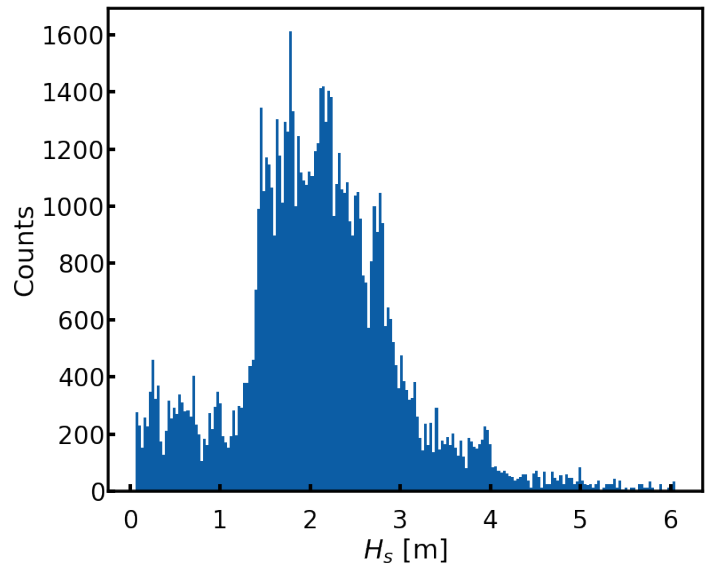


Fig. 10: Histogram of wave height.