Proceedings of 15th International Marine Design Conference (IMDC-2024)
June 2-6, 2024
Amsterdam, the Netherlands

Operational Data for Sea Margin Calculations in Early Ship Design

Sietske de Geus-Moussault [1a,2,*] Henk Seubers[3], Harry Linskens[3], Andrea Coraddu [1b] and Jeroen Pruyn[1c,4]

ABSTRACT

The current sea margin estimate applied in early ship design, commonly assumed 15-20% extra installed engine power, is not based on calculations, but has nonetheless become an industry standard. These sea margin estimations, applied in early ship design, are insufficiently accurate. This paper evaluates if a data driven approach is suitable to more accurately predict the sea margin in early ship design. Using operational data this method considers the whole operational profile of the vessel not limited to design or calm water conditions. A case study is performed where a data driven model is trained to make power predictions, subsequently this trained model is used to make calm water predictions. This proof of concept illustrates the potential of proposed method to be utilised for sea margin estimations in early ship design.

KEY WORDS

Sea Margin; Early Ship Design; Operational Data; Data Driven Model; Calm Water Predictions.

INTRODUCTION

In the early design stage of a vessel, many characteristics of the vessel are determined and fixed. When considering the propulsion system and power train, these choices have a large impact on the future emissions of the vessel, and care is required so as not to install too much power. A key parameter in this estimation is the sea margin, an addition of 15-20% to compensate for any unknowns and issues related to sailing on the oceans instead of in calm water (Esmailian et al. (2022), Islam and Soares (2022)). To the best of the authors' knowledge, there is no clear explanation for this value, and it has been maintained at this value despite improvements in estimation techniques and other scientific advances. As this directly impacts the installed power, reconsidering this value is warranted.

As explained, the sea margin is additionally installed engine power to attain added resistance due to external conditions, such as resistance from wind, waves, and/or fouling. To more accurately estimate the sea margin, the required extra power is needed to accurately estimate the added resistance. Based on the added resistance, a more accurate assumption for the required sea margin can be made, or at least a start can be made with understanding how this sea margin ensures safety and comfort during operations. To enhance the accuracy of models that predict the sea margin, it is essential to first assess the existing methods used to calculate added resistance. The following paragraphs will detail this evaluation.

[1] Department of Maritime and Transport Technology, Delft University of Technology, Delft, the Netherlands;
[a] ORCID: 0000-0002-5116-1491; [b] ORCID: 0000-0001-8891-4963; [c] ORCID: 0000-0002-4496-4544; [2] Maritime Institute Willem Barentsz, NHL Stenden University of Applied Sciences, Leeuwarden, the Netherlands; [3] DEKC Maritime, Groningen, the Netherlands [4] CoE HRTech, Maritime Innovation, Rotterdam University of Applied Sciences, Rotterdam, the Netherlands [*] Corresponding Author: s.r.a.degeus-moussault@tudelft.nl
Submitted: 29 March 2024, Revised: 1 May 2024, Accepted: 3 May 2024, Published: 23 May 2024

Conference paper, DOI: https://doi.org/10.59490/imdc.2024.882 e-ISSN: 3050-486

Literature Review

This section evaluates the current methods for calculating the added resistance: traditional methods, towing tank testing, and computational fluid dynamics.

Traditional Methods

A significant component of added resistance is wave resistance, which is defined as the resistance encountered by a ship as it moves through water. More precisely, *"the steady motion of a ship in initially smooth water assuming an ideal fluid"* (Bertram (2012)). Often considered to be the first breakthrough in wave resistance calculations is the 1898 paper of J.H. Michell, using a triple integral to calculate the wave resistance of a ship (Michell (1898)).

Another well-known method is Holtrop-Mennen, which can predict the resistance and propulsion data for different hull designs using regression analysis. First, in 1982, their method was used to develop a formula based on full-scale data and model experiments (Holtrop et al. (1982)). Subsequently, in 1984, the method was refined with more data and model tests to improve overall power predictions. The earlier method proved insufficient for more high-speed vessels at Froude numbers above 0.5 (Holtrop (1984)). In 1984, Holtrop and Mennen developed a formula to calculate the resistance, Equation 1, (Holtrop (1984)).

$$R_{Total} = R_F(1+k) + R_{APP} + R_W + R_B + R_{TR} + R_A \tag{1}$$

In which R_F is the frictional resistance; $(1 + K)$ is the form factor of hull; R_{APP} is the appendage resistance; R_W is the wave resistance; R_B is the additional pressure resistance of bulbous bow near the water pressure; R_{TR} is the additional pressure resistance due to transom immersion and R_A is the model ship correlation resistance. This is a parameter-set that describes the hull and is used to calculate the total resistance, the formula is made up out of the; frictional resistance, dependant on the form factor of the hull, the appendage resistance, wave resistance, bulbous bow pressure resistance, transom immersion pressure resistance and model-ship correlation resistance.

As this method relies on analysis of a finite combination of full-scale data and model tests, estimations and assumptions within the used parameters allow for less detail and robust results. When unorthodox combinations of parameters are used, it leads to inaccurate results (Holtrop et al. (1982)). This causes the regression method not to be the best fitting method analysing the vessel after construction (Petersen et al. (2012)).

As ship resistance calculations tend to focus on the resistance of ships through water, an under-exposed factor in overall resistance is wind loading. Not surprising as water is approximately 800 times more dense than air, however the surface area of ships above water can impact overall resistance (Blendermann (1994)). Particularly in the case of cargo ships, for instance, the surface area of the loaded vessel above water can be quite substantial. Blendermann (1994) identified parameters to calculate aerodynamic forces and moments to numerically simulate ship behaviour. This experiment was executed with different wind tunnel tests on scale models.

Resistance factors of hull coating are generally neglected, or general assumptions are made in traditional resistance calculations. Almost all ships have anti-fouling paint applied to their underwater hulls. The general state of this coating, combined with slime, shell, and weed growth, adds resistance as the ship moves through the water. Townsin (2003) describes a general roughness parameter to calculate the penalty of fouling, the ship's speed loss at constant power.

Direct Model Testing in Towing Tanks

Experimental Fluid Dynamics, based on scaled ship models tested in towing tanks, provide an alternative to the traditional methods. The first internationally recognised towing tank experiments, are the results of the experiments executed by David Watson Taylor, published in his 1910 book ”The speed and power of ships”. In his book, Taylor describes the estimation of the flow resistance of 80 vessels that were model-tested in a towing tank. The sequential testing of variations in design characteristics of the models led to estimations of ship resistance (Taylor (2013)). The experiments by Taylor attributed a great deal to the knowledge of the impact of certain design choices that can be made in ship design. As experimental research relies on the repeatability of experiments in a fixed set of conditions, the smallest variation in conditions will lead to different results. Neglecting to conduct the experiments in identical temperature conditions by Taylor led the US Navy to carry out a re-analysis of the Taylor Standard Series in 1954 to correct for this variation in experimental conditions (Gertler (1954)).

The downsides to model tests are accuracy due to scaling and the overall costs of the experiments. Furthermore, the accuracy of towing tank can fluctuate as scaling problems tend to lead to a difference in wave behaviour between scale models and full size ships (Bertram (2022)). Testing in a towing tank is an expensive and time-consuming experiment that can cost tens of thousands of euros per test (Barczak (2020)).

Computational Fluid Dynamics

Computational fluid dynamics (CFD) is a form of fluid mechanics that uses numerical analysis and data structures to analyse fluid flows, and it can be used to simulate and calculate the flow around and, therewith, the added resistance of a hull-form. CFD was first developed in the 1950s with the emergence of the computer, as this opened up an more efficient method of the computation of complex partial differential equations like finite element methods (FEM) and finite difference methods (FDM) (Chung (2002)). Dynamic flows over intricate shapes can be calculated and analysed for both aerodynamics and hydrodynamics (Anderson and Wendt (1995)). In the 1960s simplified boundary layer equations were solved for ship hulls, those more basic elementary flow models led to less accurate results (Bertram (2022)). A Reynolds Averaged Navier-Stokes Solver (RANS) method was developed in the 1980s and was improved greatly with stern flow prediction in the 1990s and a numerical method for three dimensional flows was described (Chen et al. (1990)). However, because of insufficient knowledge about modeling of turbulence, propeller simulations were largely inaccurate (Zhang et al. (2006)).

To verify CFD model calculations a ship model was made open source available. The KVLCC (KRISO Very Large Crude Carrier model) was designed, analysed in both towing tank and CFD (Van et al. (2000)). Later, in 2005, a second model with a slight difference in hull shape was developed (Hino et al. (2021)). The hull lines, rudder, and propeller data are open source and available for validation and verification (Kvale (2014)). This model is used on a large scale in academics to increase the accuracy of CFD calculations. Sadat-Hosseini et al. (2013) verified a technique to calculate the added resistance of short and long waves using the KVLCC2 vessel.

Using numerical models for predicting ship performance in the design phase is becoming more common since the computational power of computers is far greater than it used to be, and RANS equations can now be more easily solved using CFD. The KVLCC2 vessel was used to predict and verify added resistance at constant forward velocity (Ozdemir and Barlas (2017)). In recent years, an effort has been made to make CFD more efficient. The CFD modelling process, which can last several weeks to months, is aided by the experience of ship designers. Because of this experience, the engineer can set design parameters to control the overall design process. The effectiveness of this process relies on the engineer’s skill.

Cui et al. (2012) adopted a machine learning approach to the early design phase in an effort to increase the robustness of the design. Weatheritt et al. (2017) used a machine learning approach on data in a CFD experiment to study linear and non-linear relationships using the least-squares regression technique. Zhao et al. (2020) build on this research using machine learning to develop a CFD training framework of RANS models.

CFD is often perceived to be time-consuming, complex, and expensive. With the computational power increase of computers and the user interface changes, this image problem might be fading, as the alternative, model testing in towing tanks, is time-consuming and expensive as well (Gatin (2019)).

Problem Definition

From the literature, it can be concluded that currently used methods to estimate the added resistance have their upsides and downsides. The traditional methods are more than 80 years old, based on ships from then and on only a selection of hull forms, making them not widely applicable to currently designed vessels. Towing tank tests lack accuracy due to scaling, are time-consuming, and expensive. Finally, CFD simulations are also expensive and considered time-consuming and complex. This illustrates that the current methods to calculate added resistance are not the most suitable to incorporate in a model to predict the sea margin in early ship design. Therefore, this research investigates whether we can utilise operational data to estimate the added resistance more accurately.

METHOD

A novel model has been developed and proposed to explore the possibility of using operational data to estimate added resistance. This model, which is currently based on operational data from a single cargo vessel, serves as a proof of concept. Successful validation of this model would allow for its expansion to include data from additional vessels, thereby enhancing its applicability and robustness In this section, the method is presented. First, the scope of the research is described. Then, in the second section, a description of the data is given. Finally, in the third section, the model description is presented.

Scope

This research covers the development of a model based on the data of one cargo vessel. The dataset consists of approximately 90,000 data points. The developed model is, in principle, suitable to make predictions based on data from this same vessel or sister vessels with similar dimensions. Nonetheless, when desired it is also possible to train the developed model with data from other vessels and therewith also make it able to make predictions about different vessels. However, this will not be investigated in this paper.

Data Description

In this research, real-world operational data from a single cargo vessel over a time span of two years is used. This data is combined with meteorological data from the Copernicus C3S Knowledge Base (Hersbach et al. (2023)). The data can be grouped into two categories: i) endogenous data, which describe the behaviour of the vessel, such as consumption of the main engine, the speed over ground, and the course over ground, and ii) exogenous data, which describe the metocean conditions, such as wind speed and direction, and wave height and direction. A summary of the operational data categorised by source is presented in Table 1.

Table 1: Summary of the Selected Operational Data categorised by Source

Source	Feature	Unit
Royal Wagenborg	Fuel Oil Consumption of ME	[t/day]
	Speed over Ground	[knots]
	Speed through Water	[knots]
	Draft Forward	[m]
	Draft Aft	[m]
	Shaft Generator Power	[kW]
	Deadweight	[t]
	Mean Engine Fuel Type	[MGO/HFO]
Copernicus Knowledge Base	Wind speed at 10 m (north-south)	[m/s]
	Wind speed at 10 m (west-east)	[m/s]
	Temperature at 2m	[K]
	Pressure at Mean Sea Level	[Pa]
	Mean Wave Direction	[deg]
	Significant Wave Height	[m]
	Wave Peak Period	[s]

Royal Wagenborg has provided data logs from a single cargo vessel over the period between April 13th, 2020, and April 13th, 2022. With a sampling interval of 5 minutes, various measurements, control settings, and manually entered voyage parameters were logged. Based on the GPS position and time, this data has been combined with hindcast weather data from the ERA5 dataset (Hersbach et al. (2023)). This weather data is available at one-hour intervals and a spatial scale of 0.5 degrees latitude and longitude (about 55 km). Relevant variables were selected, for the selection see Table 1, and interpolated in space along the ships track.

Model Description

In this paragraph, the setup and development of the model are explained. The explanation is divided into a few separate sections due to the nature of the work. First, the data preparation and combining of the two datasets is explained. Following, the model development and training are explained, and finally, the model implementation to predict the vessel's performance in calm water is elaborated upon.

Data Preparation

The first step in preparing the data is combining the two datasets to determine the weather conditions the ship is encountering. This is done based on location and time. For nondimensionalisation and scaling of derived variables some fixed reference values are applied, these are presented in Table 2.

Table 2: Fixed Reference Values

Variable	Symbol	Value	Unit
Length between Perpendiculars	LPP	130	[m]
Beam	B	15	[m]
Lightship Weight	LSW	3300	[t]
Height to top of Wheelhouse	H	25	[m]
Water Density	ρ_{water}	1025	[kg/m^3]
Acceleration of gravity	g	9.81	[m/s^2]

Following this first step, to evaluate the performance of the vessel, all measured conditions are expressed relative to the vessel, this means for example an apparent wind angle and apparent wind speed (Figure 1). In addition, derived variables are calculated that aid in interpreting the data, such as the wave encounter angle (WEA)

and rate of turn (ROT). All the derived variables calculated based on the measured variables are presented in Table 3.

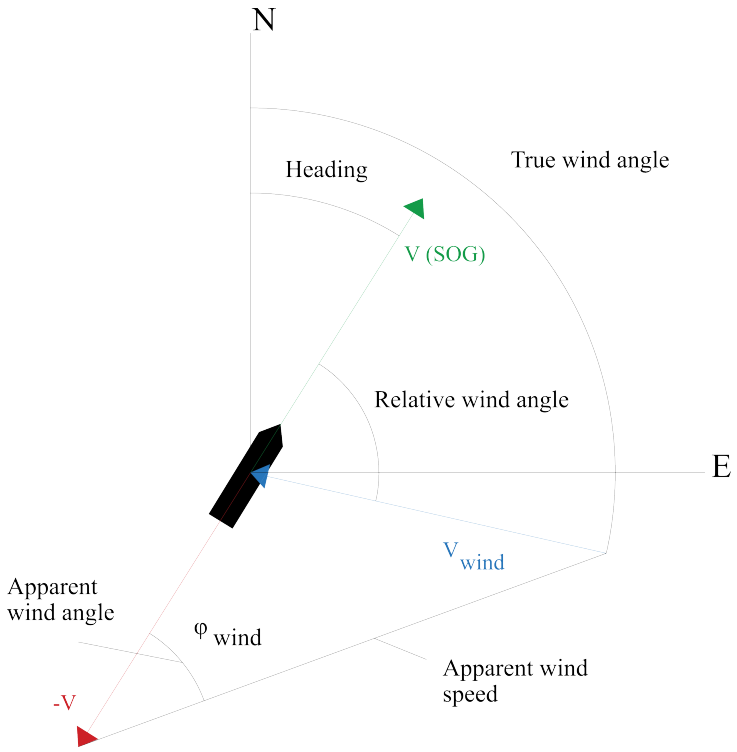

Figure 1: Graphical Representation of Relative Wind and Speed Calculation

Based on the length, the deadweight, the light ship weight, the forward and aft trim and textbook knowledge the *trim* (T), the *displacement* (D), and the *mean draft* (MD) are calculated. With the course over ground, ship speed, wind speed, ρ_{air}, and Equation 2 the *relative wind* (v), the *apparent wind angle* (AWA), and the *dynamic wind pressure* (Q_{wind}) are calculated, a graphical representation of this calculation is presented in Figure 1.

$$Q_{wind} = \frac{1}{2}\rho_{air}|v|^2 \tag{2}$$

With the course over ground, the mean wave direction, and Equation 3, the *WEA* is determined.

$$WEA = mean\ wave\ direction + 180 - course\ over\ ground \tag{3}$$

As a final step, also based on the graphical representation in Figure 1 and with the course over ground, the ships speed, and Equations 4 and 5, the *rate of turn* (ROT), *longitudinal acceleration* (A_x), and the *transverse acceleration* (A_y) are calculated. With that, all necessary derived parameters (Table 3) required as input for the model have been determined, and thus, the model can be developed.

$$ROT = \frac{dCOG}{dt} \tag{4}$$

$$A = \frac{dV}{dt} \tag{5}$$

Table 3: Derived Variables

Derived Variable	Symbol	Unit
Trim	T	[rad]
Displacement	D	[t]
Mean Draft	MD	[m]
Relative Wind Intensity	v	[m/s]
Apparent Wind Angle	AWA	[rad]
Dynamic Wind Pressure	Q_{wind}	[Pa]
Mean Wave Drift Force	MWDF	[N]
Wave Encounter Angle	WEA	[rad]
Rate of Turn	ROT	[deg/min]
Longitudinal Acceleration	A_x	[m/s^2]
Transverse Acceleration	A_y	[m/s^2]

Model Development

For the power prediction model a two step approach is required. First, the model needs to be developed and trained based on the parameters mentioned above. In the second step, the developed model is used to predict the required power in calm water conditions. The goal is to develop a model that can predict the required power given (average) speed, loading, and environmental conditions. Subsequently, the model will be implemented to predict the corrected required power, given speed, the design draft, and calm environmental conditions. Before constructing the model, some dependencies between the input variables have to be established.

In ship design, the power at the propeller shaft is of interest because this is measured during sea trials and separates the hydrodynamic performance from mechanical and thermal efficiencies. However, from the operational data, only fuel consumption is usually available to measure propulsive power when the shaft power is not measured. Therefore the model will aim to predict fuel consumption in ton per day.

The resistance is split into individual contributions, which are modelled independently. Some modelling assumptions have already been made here to avoid infinite values in the model. Towing power is defined by speed through water, not speed over ground. The reason is that the towing speed-power curve (corrected to calm-water) is a property of the ship, while the speed-over-ground-power curve is a property of the ship and the local current. In other words, the corrected power data will not coincide with a single curve when plotted against speed-over-ground. The resistance contributions are modeled linearly using coefficients such as, calm water, wind, and wave influence.

$$P_{tow} = V_w(R_{calm} + R_{wind} + R_{wave} + R_{inertia}) \tag{6}$$

The friction, pressure, and wave-making resistances are all modeled to scale with the dynamic pressures.

$$Q_w = \frac{1}{2}\rho_w V_w^2\,, \quad Q_a = \frac{1}{2}\rho_a V_a^2 \tag{7}$$

No Froude-number effect is modelled, so the model should only be applied to displacement vessels well below the critical Froude number. The reference area for calm water resistance is an indication of the wetted area based on the main dimensions and the draft. The reference area for wind resistance is an approximate frontal area above the waterline, including the accommodation.

$$R_{calm} = LPP * (C_{calm}B + C_{draft}T) * Q_w \tag{8}$$

$$R_{wind} = C_{wind}(\alpha_{wind}) * (H - T)B * Q_a \tag{9}$$

The wave-induced resistance is based on the significant wave height H_s, the wave encounter angle α, and the wave length λ_p at the peak period, resulting in Equation 10. This is, in essence, the mean wave drift force (Journée et al. (2015)) with a coefficient consisting of a spline (in the frequency/wavelength) (Press et al. (2011)) and a Fourier series (in the wave encounter angle).

$$R_{wave} = C_{wave}\left(\alpha, \frac{\lambda_p}{LPP}\right)\frac{1}{16}\rho_w g B H_s^2 \tag{10}$$

We aim to develop a model that defines the speed-power curve under specific conditions: calm water, no wind, and constant speed and direction. This model will consider a vessel at its design draft, maintaining an even keel and consistent power output. To extract this knowledge from the model, we first need to develop and train it. Earlier sections have already established the model's input parameters and their interdependencies. The model's input features and target feature are summarised in Table 4.

Table 4: Input Parameters of the Model

Input and Target Features	**Symbol**
LPP	LPP
Beam	B
Speed	V
Trim	T
Displacement	D
Relative Wind Intensity	v
Apparent Wind Angle	AWA
Dynamic Wind Pressure	Q_w
Wave Encounter Angle	WEA
Rate of Turn	ROT
Longitudinal Acceleration	A_x
Transverse Acceleration	A_y
Measured Fuel Consumption	FOC

One of the objectives of this study is to develop a model for predicting the FOC based on the input parameters outlined in Table 4. This learning problem can be formulated as a supervised Machine Learning (ML) problem, specifically a regression problem (Shalev-Shwartz and Ben-David (2014)). In regression analysis, an input space $\mathcal{X} \subseteq \mathbb{R}^d$ is comprised of d features (in this case, the parameters in Table 4) and the output space, $\mathcal{Y} \subseteq \mathbb{R}$, corresponds to FOC. A dataset of n examples, denoted as $\mathcal{D}_n = (\boldsymbol{x}_1, y_1), \cdots, (\boldsymbol{x}_n, y_n)$, represents input/output relationships where $\boldsymbol{x}_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}\ \forall i \in 1, \cdots, n$. The aim is to learn the unknown input/output function $\mu : \mathcal{X} \to \mathcal{Y}$ based solely on $\mathcal{D}n$. An ML regression algorithm A, characterized by its hyperparameters $\mathcal{H}$, selects a model f from a set of potential models $\mathcal{F}$ based on available data $A_{\mathcal{H}} : \mathcal{D}_n \times \mathcal{F} \to f$. The set $\mathcal{F}$ is typically unknown and depends on the choices of A and $\mathcal{H}$.

Many different ML algorithms exist in the literature (Shalev-Shwartz and Ben-David (2014); Chandrashekar and Sahin (2014)) but, as the no-free-lunch theorem states (Wolpert (2002)), there is no way to determine a-priori the best ML algorithms to use for a specific application. Since our goal is to establish a proof-of-concept, we have opted to use a Regularized Least Squares (RLS) model as our benchmark. This decision is supported by the fact that linear regression models are frequently employed in similar engineering contexts. Additionally, preliminary exploratory data analysis revealed a linear correlation between several input parameters and the target feature FOC. Thus, initiating our investigation with a linear model as a benchmark provides a logical and strategic starting point. RLS is a regression method that introduces a regularization term to the traditional least squares problem to control the complexity of the model and prevent overfitting. The objective of RLS is to minimize the sum of squared residuals, similar to ordinary least squares, but with an additional penalty term that discourages

large values of the model parameters. The regularization term is typically a function of the model parameters, such as the L2 norm (also known as Ridge regression) or the L1 norm (also known as Lasso regression). The L2 norm encourages small parameter values, leading to a more stable model with lower variance, while the L1 norm can lead to sparse solutions, where some parameters are exactly zero, effectively performing feature selection. The balance between the fit to the data and the regularization is controlled by the only hyperparameter of this algorithm λ. A larger λ increases the impact of the regularization term, leading to a simpler model, while a smaller λ allows the model to fit more closely to the data, potentially at the risk of overfitting. The accuracy of model f in approximating μ is evaluated using a prescribed metric $M : f \rightarrow \mathbb{R}$. Multiple metrics are available for regression analysis in ML (Aggarwal (2015)). However, due to the physical significance of FOC, this study will focus on three primary metrics: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R^2). A statistically consistent model selection and error estimation process was conducted to identify the optimal hyperparameters and evaluate the performance of the final model based on the desired metrics.

We divide all available data into two subsets: training and test sets. Specifically, 80% of the data is allocated for training the model, which uses measured power as its input (as shown in Model in Figure 2). Following the training phase, we utilize the remaining 20% of the data as a test set to validate the model's performance. This approach ensures that we can assess the model's accuracy and generalizability to unseen data.

The results from this validation of the model are described in the model implementation paragraph.

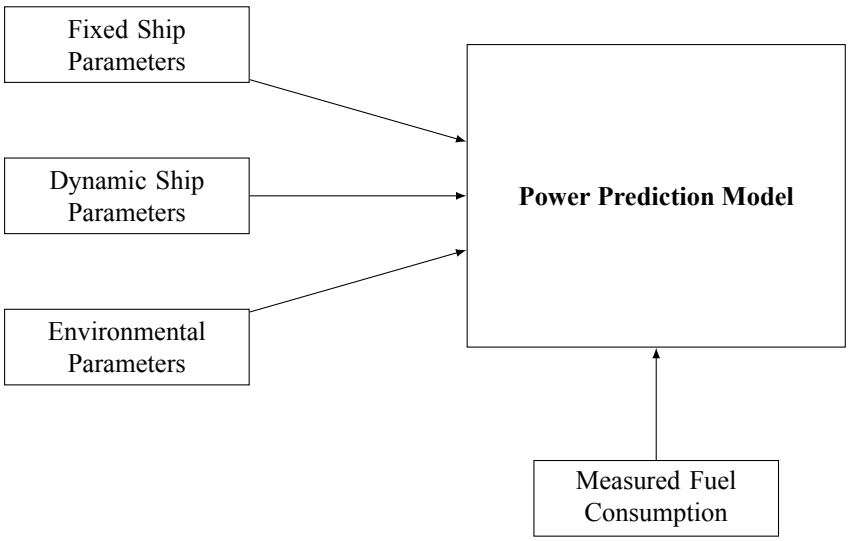

Figure 2: Training and Development of Power Prediction Model

Model Implementation

The first step is to employ the model developed in the previous section to predict the calm water speed-power curve, which assumes no wind, constant speed and direction, design draft with even keel, and constant power take-off. The performance of the model is assessed through error estimation and model validation, as depicted in Figure 3 and summarized in Table 5. These visual and tabular representations indicate that the model achieves a satisfactory level of accuracy. However, there remains potential for enhancement. Consequently, pursuing further research with alternative (non-linear) modeling approaches represents a viable direction for future studies.

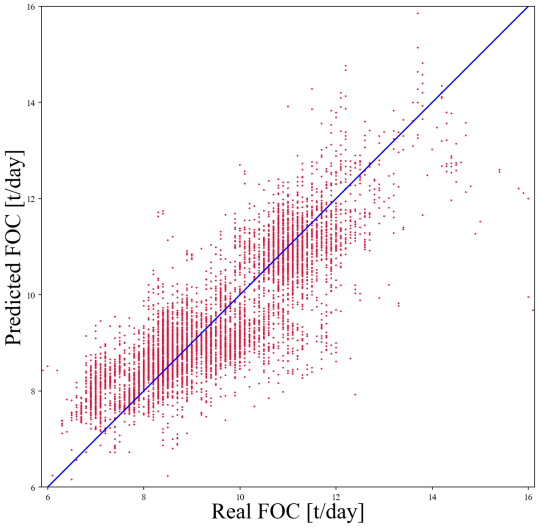

Figure 3: Scatterplot Predicted versus Measured

Table 5: Error Estimates of the Model

Metric	**Accuracy**
MAE	0.64 ± 0.0070
MAPE	17.79 ± 0.1281
R^2	0.706 ± 0.0163

The second step involves using the model to generate power predictions under calm water conditions. This step requires modifications to some of the input parameters previously utilized for model development, as detailed in Table 4 and Figure 2. These parameters are adjusted to reflect calm water conditions and are displayed in blue in Table 6 and Figure 4. Additionally, it is important to note that the power measurements, which were initially used as inputs in the model, are no longer included.

Table 6: Input Features of the Model

Input Feature	**Alteration**
LPP	as before
Beam	as before
Speed	as before
Trim	Design Draft
Displacement	Design Draft
Relative Wind	No wind
Apparent Wind Angle	No wind
Q_{wind}	No wind
Wave Encounter Angle	No waves
Rate of Turn	No Rate of Turn
Longitudinal Acceleration	No Acceleration
Transverse Acceleration	No Acceleration
Measured Fuel Consumption	No input

With these altered input features we will test our hypothesis, *that it is possible to make calm water predictions based on real-world operational data*. In the next section, the results are presented and discussed.

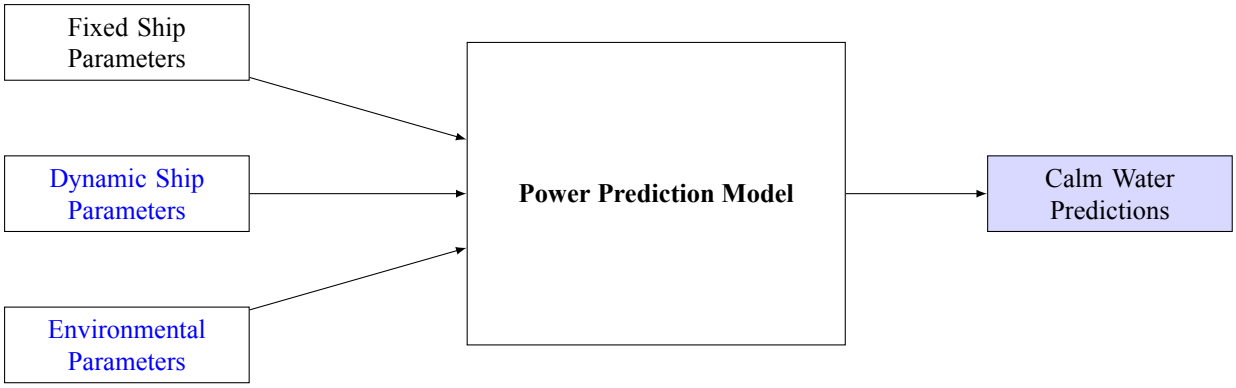

Figure 4: Implementation Power Prediction Model

RESULTS & DISCUSSION

In this section, the results of the calm water predictions are presented and discussed. In the scatter plots, the fuel consumption (t/day) is on the y-axis, and the speed over ground (knots) is on the x-axis. The original measured fuel consumption of the main engine, denoted by y, is depicted in blue, while the predictions of fuel consumption, corrected for factors such as draft, maneuvering, and weather conditions, are illustrated in orange.

Figure 5 shows a scatter of the measured fuel consumption, and in Figure 6, the measured and the corrected are reported.

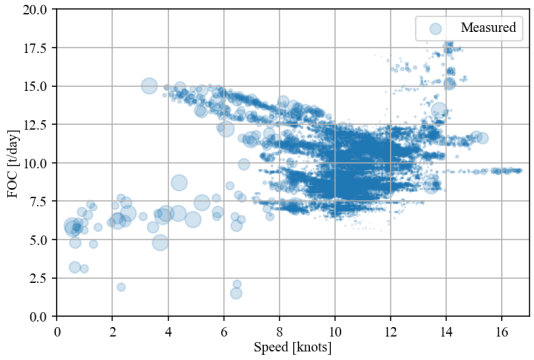

Figure 5: Speed-Power Plot based on the Measured Operational Data

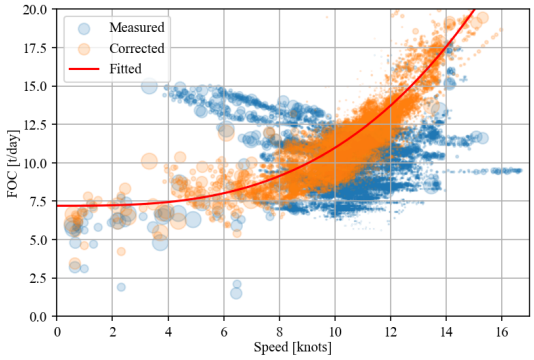

Figure 6: Speed-Power Plot with Measured and Corrected

In Figure 6, it is clear that the corrected data points (orange) are less scattered than the original measured operational data (blue). Furthermore, it is observed that the orange scatters approach the fitted red line. This fitted line is an ideal representation of the developed model, meaning the calm water situation. The orange scatters approaching the fitted line illustrate that the proposed methodology is capable of correcting for weather conditions, but also that there is still room for improvement. Moreover, it can be observed that the two cloudy protrusions in the blue scatter (coming from the center of the plot pointing towards the top left of the plot, between 4 and 8 knots, and 12.5 and 15 t/day) are no longer visible in the calm water predictions. This is a positive confirmation of correcting for weather conditions, as these two protrusions represent two storms (on 27 December 2020 and 11 March 2021). Also, these two protrusions make it clear that the weather conditions (in this case, the two storms) result in a significant involuntary speed reduction, leading to a speed between 4 and 8 knots with a fuel consumption between 12.5 and 15 t/day.

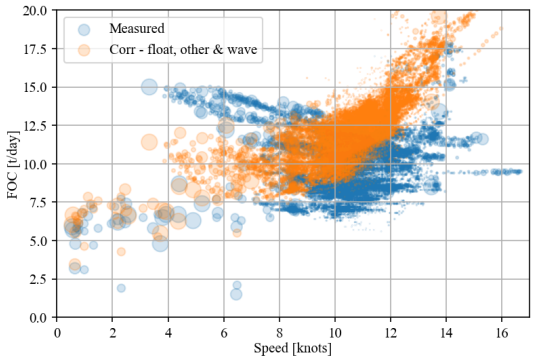

Figure 7: Speed-Power Plot with Corrections for All Parameters Except Wind

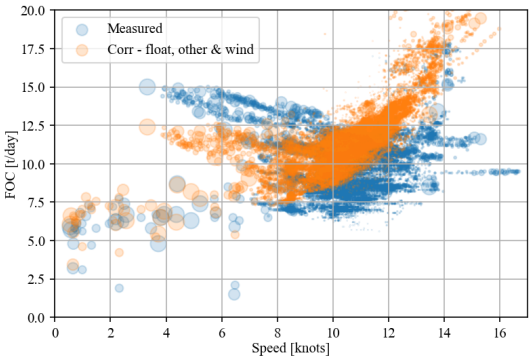

Figure 8: Speed-Power Plot with Corrections for All Parameters Except Waves

When making the calm water predictions, several intermediate steps are taken between Figure 5 and Figure 6. The intermediate steps most interesting to discuss are the situations without the correction for wind (Figure 7) and without the correction for waves (Figure 8). When making the corrections for all parameters except for wind, we get the results plotted in Figure 7. Alternatively, when making the corrections for all parameters except for waves we get the results plotted in Figure 8. When evaluating the difference between the two situations, Figure 7 and Figure 8, it can be seen that for the conditions the vessel encountered between April 2020 and April 2022, the waves appear to have a larger impact on the fuel consumption. This can be best seen when looking at the orange scatter clouds between 4 and 8 knots and comparing the change in these clouds between the measured data (Figure 5), the corrections except wind (Figure 7), and the corrections except waves (Figure 8). When comparing these two predictions with the measured data and each other, it is evident that the effect of the waves is larger than the effect of the wind. Furthermore, when looking at Figure 6, it can be observed that the model still has some difficulty completely filtering out the effect of these two storms, as the orange cloud between 4 and 8 knots remains more scattered than at higher speeds.

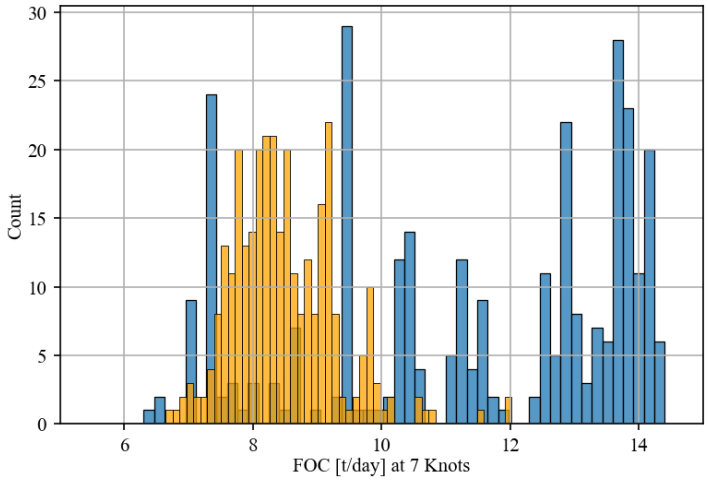

Figure 9: Histogram of the Distribution of the Fuel Consumption at 7 knots

To properly evaluate and compare the distribution, three histograms are plotted in Figures 9, 10, and 11. These histograms are cross-sections of the distribution of the fuel consumption at low, design, and high speed. Furthermore, a summative table presenting the mean and variance at all speeds between 7 and 14 knots is given in Table 7. When evaluating the distribution of the fuel consumption at low speed, 7 knots (Figure 9) it is clear that the spread has decreased between the measured (blue) and the predicted calm water (orange) fuel consumption. The measured fuel consumption is spread between 6.5 and 14.5 tonne per day, and the calm water predictions are spread between 6.75 and 12. This reduced the range from 8 for the measured data to 5.25 for the predicted

calm water. Also, when comparing the distribution of the measured and predicted, it can be observed that the orange graph is far more like a bell curve than the blue distribution.

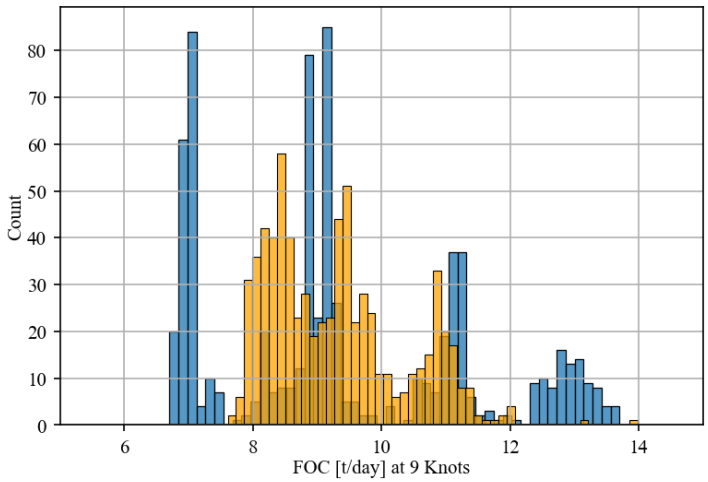

Figure 10: Histogram of the Distribution of the Fuel Consumption at 9 knots

When evaluating the distribution of the fuel consumption at the design speed, 9 knots (Figure 10), it is clear that also at this speed, the spread has decreased between the measured (blue) and the calm water (orange) fuel consumption. The measured fuel consumption is spread between 7 and 14 tonne per day, and the calm water predictions are spread between 8 and 12. This results in a reduction from 7 for the measured data to 4 for the predicted calm water. Furthermore, when comparing the distributions, it can be observed that, in this case, the orange distribution is more in the shape of a bell curve than the original blue distribution.

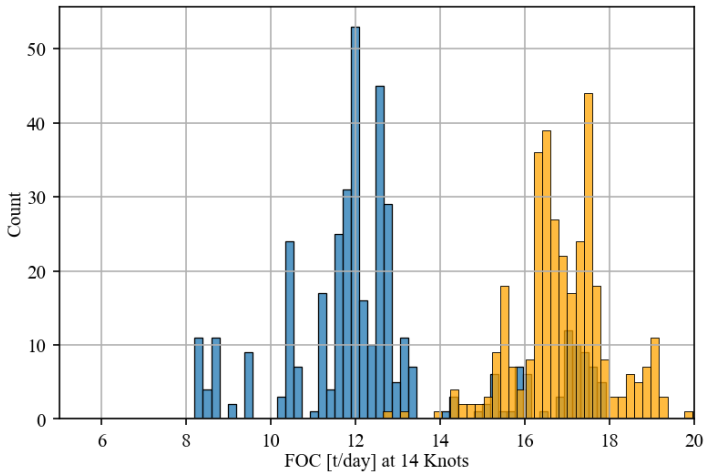

Figure 11: Histogram of the Distribution of the Fuel Consumption at 14 knots

When finally also evaluating the distribution of the fuel consumption at high speed, in this case, 14 knots (Figure 11), it is clear that also at this speed, the spread has decreased between the measured (blue) and the calm water (orange) fuel consumption. The measured fuel consumption is spread between 8 and 18 tonne per day, and the calm water predictions are spread between 14 and 20. Meaning that also, for this situation, there is a significant decrease in the range, from 10 for the measured to 6 for the calm water predictions. Furthermore, when comparing the distributions in Figure 11, it can be observed that the orange graph is more in the shape of a bell curve than the blue distribution. However, for this speed it is also noted that the bell curve has shifted from the original center around 12 tonne per day to a new center around 17.

Table 7: FOC Mean and Variance at Different Speeds

Speed	FOC Mean [t/day]		FOC Variance [t/day]	
[knots]	Measured	Calm Water	Measured	Calm Water
7	11.41	8.48	5.62	0.65
8	9.96	9.29	4.86	0.99
9	9.37	9.29	3.90	1.08
10	8.89	10.19	2.05	0.92
11	9.38	11.42	2.03	0.99
12	9.85	12.51	1.77	1.19
13	10.73	14.80	2.64	2.16
14	12.54	17.65	5.32	4.10

When comparing the mean and variance of the measured and the calm water prediction in Table 7, it can be seen that the variance has decreased at all speeds. Furthermore, it becomes clear that at the lower speeds, the mean in the calm water predictions is lower than the corresponding mean of the measured data, while at the higher speeds, the calm water mean is higher than the corresponding mean of the measured data. These reductions in range and shifts of the mean to lower fuel consumption at reduced speeds not only validate the concept of the proposed method but are also highly explicable. In Figure 9, it can be seen that the corrected predicted fuel consumption (orange) has shifted to the left, a lower fuel consumption. When evaluating Figure 10, it can be seen that for this speed, the predicted fuel consumption is more centered compared to the measured fuel consumption. Finally, when looking at Figure 11, the fuel consumption at 14 knots, it can be seen that the predictions have moved more to the right. This seems very logical because the lower and higher speeds are outside the normal operating envelope of the vessel. The lower speeds with high fuel consumption are involuntary, in this case mostly due to storms. The higher speeds with low fuel consumption are only possible with positive help from a tailwind and/or current. This means that when we remove the meteorological effects, low speeds with high fuel consumption and high speeds with low fuel consumption no longer occur.

Overall, it is observed that the predicted calm water fuel consumption has a decreased variance relative to the variance of the original measured fuel consumption. Furthermore, the histograms in Figures 9, 10, and 11 confirm what is also observed in Figure 6 that the developed method for making calm water predictions works as it results in less scattered data points, all more centered around the fitted line. Implying the developed method works.

CONCLUSIONS

This paper presents a method to make calm water power predictions. This is achieved through a data driven model based on operational data. A case study proved the concept and showed the potential of this method. It illustrated that operational data can be used to make calm water predictions, correcting for meteorological effects. When comparing the calm water power predictions with the actual delivered power the sea margin can be calculated. The proposed methodology will enable a more accurate sea margin calculation in early ship design and will thus allow for a better understanding of the required power. This will lead to a better-informed decision-making process regarding the propulsion system and/or the possibilities of alternative fuels, thus having a positive effect on the footprint of a vessel.

This novel approach for sea margin estimation offers a quick and substantiated estimation in the early-design stage of a ship, making it a promising starting point for further research. Recommendations for further research comprise further developing the model by testing different algorithms, including more data from different vessels to train and test the model, and finally evaluating whether a pattern in sea margin for certain passages can be recognised.

CONTRIBUTION STATEMENT

Sietske de Geus-Moussault: Conceptualization; data curation, methodology; writing – original draft. **Henk Seubers:** conceptualization; data curation, methodology; writing – review and editing. **Harry Linskens:** conceptualization; data curation, methodology; writing – review and editing. **Andrea Coraddu:** conceptualization; supervision; writing – review and editing. **Jeroen Pruyn:** conceptualization; supervision; writing – review and editing.

ACKNOWLEDGEMENTS

This research was performed based on data from a cargo vessel of Royal Wagenborg the authors would like to acknowledge and thank Royal Wagenborg for their support of this research.

This research has been performed within the TODDIS project, partially funded by the Dutch Research Council (NWO) under grant agreement Raak-Pro program 2018, n° 03.023.

This research contains Copernicus Climate Change Service Information (2024), neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

REFERENCES

Aggarwal, C. C. (2015). *Data mining: the textbook*. Springer.

Anderson, J. D. and Wendt, J. (1995). *Computational fluid dynamics*, volume 206. Springer.

Barczak, N. (2020). *The ship towing tank*. DMS.

Bertram, V. (2012). *Practical ship hydrodynamics*. Elsevier.

Bertram, V. (2022). Realistic assessment of saving potential for energy saving options. In *Sustainable Energy Systems on Ships*, pages 451–467. Elsevier.

Blendermann, W. (1994). Parameter identification of wind loads on ships. *Journal of Wind Engineering and Industrial Aerodynamics*, 51(3):339–351.

Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28.

Chen, H. C., Patel, V. C., and Ju, S. (1990). Solutions of reynolds-averaged navier-stokes equations for three-dimensional incompressible flows. *Journal of Computational Physics*, 88(2):305–336.

Chung, T. J. (2002). *Computational fluid dynamics*. Cambridge university press.

Cui, H., Turan, O., and Sayer, P. (2012). Learning-based ship design optimization approach. *Computer-Aided Design*, 44(3):186–195.

Esmailian, E., Steen, S., and Koushan, K. (2022). Ship design for real sea states under uncertainty. *Ocean Engineering*, 266(5). doi.org/10.1016/j.oceaneng.2022.113127.

Gatin, I. (2019). Cfd in the marine industry: Today and tomorrow.

Gertler, M. (1954). *A reanalysis of the original test data for the Taylor Standard series*. Department of the Navy, Washington.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N. (2023). Era5 hourly data on single levels from 1940 to present. copernicus climate change service (c3s) climate data store (cds), doi: 10.24381/cds.adbb2d47 (accessed on 01/02/2023).

Hino, T., Hirata, N., Stern, F., Larsson, L., Visonneau, M., and Kim, J. (2021). Introduction, conclusions and recommendations. In *Numerical Ship Hydrodynamics: An Assessment of the Tokyo 2015 Workshop*, pages 1–21. Springer.

Holtrop, J. (1984). A statistical re-analysis of resistance and propulsion data. *Published in International Shipbuilding Progress, ISP, Volume 31, Number 363*.

Holtrop, J., Mennen, G. G. J., et al. (1982). An approximate power prediction method. *International Shipbuilding Progress*, 29(335):166–170.

Islam, H. and Soares, C. (2022). Head wave simulation of a kriso container ship model using openfoam for the assessment of sea margin. *ASME. J. Offshore Mech. Arct. Eng.*, 144(3). doi.org/10.1115/1.4053538.

Journée, J. M. J., Massie, W. W., and Huijsmans, R. H. M. (2015). Offshore hydromechanics. *3rd edition*. Delft.

Kvale, J. M. (2014). Revised simulation model for a very large crude carrier (vlcc). Master's thesis, Institutt for marin teknikk.

Michell, J. H. (1898). The wave-resistance of a ship. *Philosophical Magazine*, 45(5):106–123.

Ozdemir, Y. H. and Barlas, B. (2017). Numerical study of ship motions and added resistance in regular incident waves of kvlcc2 model. *International Journal of Naval Architecture and Ocean Engineering*, 9(2):149–159.

Petersen, J. P., Jacobsen, D. J., and Winther, O. (2012). Statistical modelling for ship propulsion efficiency. *Journal of marine science and technology*, 17:30–39.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2011). *Numerical Recipes: The Art of Scientific Computing*. Third Edition. ISBN 978-0-521-88068-8.

Sadat-Hosseini, H., Wu, P.-C., Carrica, P. M., Kim, H., Toda, Y., and Stern, F. (2013). Cfd verification and validation of added resistance and motions of kvlcc2 with fixed and free surge in short and long head waves. *Ocean Engineering*, 59:240–273.

Shalev-Shwartz, S. and Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.

Taylor, D. W. (2013). *The speed and power of ships*. BoD–Books on Demand.

Townsin, R. L. (2003). The ship hull fouling penalty. *Biofouling*, 19(S1):9–15.

Van, S.-H., Kim, W.-J., and Kim, D.-H. (2000). Experimental investigation of local flow around kriso 3600teu container ship model in towing tank. *Journal of the Society of Naval Architects of Korea*, 37(3):1–10.

Weatheritt, J., Pichler, R., Sandberg, R. D., Laskowski, G., and Michelassi, V. (2017). Machine learning for turbulence model development using a high-fidelity hpt cascade simulation. In *Turbo Expo: Power for Land, Sea, and Air*, volume 50794, page V02BT41A015. American Society of Mechanical Engineers.

Wolpert, D. H. (2002). The supervised learning no-free-lunch theorems. In *Soft computing and industry*.

Zhang, Z.-r., Hui, L., Zhu, S.-p., and Feng, Z. (2006). Application of cfd in ship engineering design practice and ship hydrodynamics. *Journal of Hydrodynamics, Ser. B*, 18(3):315–322.

Zhao, Y., Akolekar, H. D., Weatheritt, J., Michelassi, V., and Sandberg, R. D. (2020). Rans turbulence model development using cfd-driven machine learning. *Journal of Computational Physics*, 411:109413.