

Power Increase due to Marine Biofouling: a Grey-box Model Approach

Matthew de Haas^{a,b}, Andrea Coraddu^a, Abdel-Ali El Mouhandiz^b, Nikoleta Dimitra Charisi^a, and
Austin A. Kana^{a*}

^aDepartment of Maritime and Transport Technology, Delft University of Technology, Delft, the Netherlands

^bFeadship, Hoofddorp, the Netherlands

*A.A.Kana@tudelft.nl

Abstract

This paper proposes a grey-box modelling approach to predict marine biofouling growth and its effects on ship performance. The approach combines empirical or experimental-based white-box models with data-driven black-box models. First, a white-box model is built to predict ship resistance considering a bare hull. This prediction is based on calm water resistance, wind, waves, and temperature differences. Subsequently, marine biofouling growth is predicted using an experimental model that estimates the level of roughness on the ship hull. Finally, a deep extreme learning machine is used as a black-box model, employing a feedforward neural network technique. To test the approach, a superyacht case study was selected as a category of vessel heavily exposed to fouling. The study used a 2-year dataset obtained through a collaboration with Feadship. Results showed that the black-box approach outperforms the white-box approach in predictive capabilities. However, when the knowledge encapsulated in the white-box model is included in the grey-box approach, the model shows the highest prediction accuracy achieved by leveraging less historical data. This study demonstrates the potential of the proposed grey-box approach to accurately predict marine biofouling growth and its effects on ship performance, which can benefit ship operators and designers in improving operational efficiency and reducing maintenance costs.

Keywords: Marine Biofouling; Power Increase Prediction; Yachts; Deep Extreme Learning Machine; Grey-box Models.

1 INTRODUCTION

Marine biofouling, a phenomenon of the accumulation of micro and macro-organisms on immersed surfaces, has a strong influence on the performance of vessels by increasing surface roughness and consequently increasing fuel consumption and emissions of greenhouse gases [1]. Biofouling creates roughness on the hull and propeller which leads to additional frictional resistance and loss of propeller efficiency, also known as additional sea margin. Research has shown that fuel costs can be increased up to 35% when the ship hull is heavily fouled [2]. Additionally, biofouling threatens ecological balance by transferring invasive aquatic species in waters where they have little to no natural enemies [3]. A trade-off is often made for ship owners and operators between the additional cost of maintenance to keep the ship clean compared to the increase in operational cost of sailing with a fouled hull. The current practice is that the hull and propeller are cleaned when other maintenance is scheduled, which does not guarantee optimal cleaning schedules [4].

Accurate prediction of biofouling may lead to significant benefits for ship design, maintenance and operations. For ship design, the added sea margin is the result of both added hull resistance and decreased propeller efficiency and can be taken into account either within hydrodynamic analysis or within the calculation of the powering of the vessel. Additionally, for ship's maintenance and operation, accurate marine biofouling prediction may lead to optimal maintenance and cleaning schedules.

Marine biofouling is a complex phenomenon because it depends on various variables such as sea-water surface temperature, salinity, acidity, speed of water flow, and light intensity [1]. There is currently no accurate and universal method for predicting biofouling and associated added sea margin [5]. The standard approach for estimating the speed loss is by applying ISO 19030 (ISO 19030-2, 2016), which prescribes methods for measuring changes in hull and propeller performance to give an indication for hull and propeller efficiency. However, this approach lacks a clear method on how to predict added sea margin due to fouling for ship design. As such, only low-fidelity analytical expressions ex-

ist recommended by the Propulsion Committee of the 28th ITTC, having an average error of around 20% [6].

State-of-the-art methods for predicting the additional sea margin and the effect of biofouling on a propeller's performance are based on first-principle experimental based models (e.g., [7]), computational fluid dynamics (CFD) (e.g., [8]–[10], and data-driven modelling [11]. The main strengths of first-principle experimental based models are their interpretability and low computational cost to perform evaluations. Such models can give a good insight into the relevant physics, such as those related to added frictional resistance of a flat plate [12]. However, conventional models cannot represent all fouling situations, and models are often limited to static growth of fouling [13]. On the other hand, CFD can provide accurate predictions of the additional sea margin. However, at a high computational cost, thus limiting results to the chosen ship hull. Authors of [12], [14] have shown promising results with only a few percentage deviation between predicted resistance and power compared to verified results.

Finally, data-driven black-box models are based on Machine Learning (ML) techniques and are able to address complex problems and improve the accuracy of the predictions. These data-driven models show great potential to predict added sea margin by giving new insights and accounting for all variables with limited simplifications [11]. However, data-driven models typically suffer from a lack of interpretability. Authors of [11] developed a data driven digital twin to estimate the speed loss due to marine biofouling, which showed significant improvement compared to ISO 19030.

Authors of [15] investigated and compared the applicability of both Artificial Neural Networks (ANNs) and Gaussian processes (GPs) to the prediction of fuel efficiency in ship propulsion by using sensors' data. White-box modelling trends were applied to account for fouling. The authors concluded that the ANN performed slightly better than the GPs. Additionally, [11] developed a data-driven digital twin using data from on-board sensors to estimate the speed loss due to marine biofouling. The method used was Deep Extreme Learning Machine (DELIM), which uses a feedforward neural network. This overcomes the problems resulting from the backward-propagation training algorithm with potentially low convergence rates, critical tuning of optimization parameters, and presence of local minima that call for multi-start and re-training

strategies [16].

To the best of the authors' knowledge, the applicability of a grey-box modelling approach, which combines a white- and a black-box model, for the prediction of marine biofouling has not yet been explored. The authors thus propose a grey-box approach to predict the power increase for ships due to marine biofouling. The white-box model is used for the estimation of the required power in a fouled situation, dependent on time and the ship environmental conditions. The collected sensors' data and the white-box power prediction are used as inputs for the black-box model based on DELIM.

To test the approach, a superyacht case study was selected because superyachts are particularly heavily exposed to marine biofouling due to their operational profile, which includes longer periods of being stationary compared to commercial vessels, often staying in ports or being anchored for long periods of time [17]. The operational profile of each superyacht can be very different and change over time. With an increasing and diversifying yacht fleet, understanding biofouling and its influence on yacht performance is important to minimizing their environmental footprint.

For this research, the collected data provided by Feadship came from both on-board sensors and company's databases. The available data contains ship design specifications, various captain logs, maintenance, engine, motion, voyage report, and auxiliary power data. Here the voyage report data contains onboard feedback monitoring of the following parameters: ship speed and heading, wave conditions (height, period, and directions), wind conditions (speed and direction), and corresponding measured operational profiles from Feadship fleet.

2 PHYSICS-BASED MODEL FOR BIO-FOULING ESTIMATION

This section explains the physical model for predicting the fouled ship power. This contains both a short discussion on the predicted smooth ship resistance in Section 2.1, together with further elaboration on the prediction of fouling roughness in Section 2.2, while the fouled ship power prediction is presented in Section 2.3.

2.1 Smooth Ship Resistance

The ship resistance is predicted based on calm water, wind, and wave resistance together with differences due to temperature. The method proposed is in line with recommendations given by the [18].

As a basis for the resistance prediction in any given condition, the calm water resistance is computed based on the ship's speed. The speed over water is used, rather than the speed over ground as measured with AIS. The relative wind speed is based on the ship's speed, wind speed, and wind direction. The wind resistance coefficient, area of maximum transverse section exposed to wind, air density, and relative wind speed have been used to determine the added resistance due to wind in accordance with the methodology reported in [18].

To determine the added wave resistance, first the added thrust in waves is found based on ship speed, heading, length, displacement, and water-plane coefficient of the foreship [19]. With the added thrust in waves known, the added wave resistance can be found. An actual sea state is normally described by a wave spectrum such as the one proposed by Pierson-Moskowitz [20]. To allow for flexible spectrum shapes, the spectrum is multiplied with the peak enhancement factor, using the JONSWAP spectrum [21]. Next, the directional wave spectrum is found by multiplying the JONSWAP spectrum with the angular distribution function. Last, a correction accounts for the difference in resistance due to change in water temperature and difference in ship draft due to salinity. With this approach, a change of frictional resistance coefficient and change in resistance due to ship displacement can be found [18], [22].

2.2 Fouling roughness

For the prediction of biofouling growth, the model proposed by Uzun, Demirel, Coraddu, *et al.* was used [13]. The model makes use of two main principles: i) a fouling rating and ii) the fouling surface coverage for calcareous fouling. The fouling rating forms a basis for the model, combining slime, non-shells organisms, and calcareous fouling into one overall fouling rating. Together, this gives a good indication of the level of fouling present on the ship and its resulting roughness. However, when calcareous fouling is present on the ship, its level of surface coverage can be a dominant factor. Due to this, the authors introduced the calcareous surface coverage as an additional parameter. For this reason, a different function is used for the biofouling growth when the calcareous surface coverage increases above 5%.

With analyses limited to the given regions, the authors suggest interpolating and extrapolating found patterns based on sea surface temperature as the dominant fouling parameter. With the help of

the proposed functions, biofouling growth trends for the Equator and Mediterranean can be inter- and extrapolated for all locations, to obtain the roughness thickness present on the ship. In this case, the roughness is modelled by using the equivalent sand roughness height (k_s). For a full explanation of the model, see [13].

2.3 Fouled Ship Power Prediction

The effects of the obtained roughness on the hull and propeller surface are determined next. First, the added frictional resistance coefficient (ΔC_F) as a result of this roughness can be calculated. This is done based on the equivalent sand roughness height, the ship waterline length (L_{WL}), and Reynolds number (Re) with the function of [23] (Equation 1).

$$\Delta C_F = 0.044 \left[\left(\frac{k_s}{L_{WL}} \right)^{\frac{1}{3}} - 10 \cdot Re^{-\frac{1}{3}} \right] + 0.000125 \quad (1)$$

Next, the added frictional resistance due to biofouling is found via Equation 2.

$$\Delta R_F = \frac{1}{2} \rho S \Delta C_F V_S^2 \quad (2)$$

The approach by [7] was employed to simulate the impacts of biofouling on the ship propeller. With this model, the change in thrust and torque coefficient due to biofouling is computed, and a new open water efficiency is found. This is done by finding the change in drag and lift for both coefficients (see [7] for the full method). However, it can be seen that the changes in the coefficients are a function of the propeller characteristics: propeller pitch, diameter, number of blades, chord length, and maximum thickness, together with fouling roughness. It is important to mention that both chord length and maximum thickness are taken at a radius of 0.75. Next, the drag and lift coefficient can be determined for the propeller in smooth and rough condition. Here, the smooth frictional coefficient can be found by either using Schroeinherr's friction line (in Equation 3) or with the ITTC-1957 skin friction line (Equation 4), based on the method of the [24]. For the rough condition, the added frictional resistance coefficient is found using Equation 1, where the plate length (L_{WL}) is taken as the chord length (c) at radius 0.75R.

$$\frac{0.242}{\sqrt{C_{FS}}} = \log(Re \cdot C_{FS}) \quad (3)$$

$$C_{FS} = \frac{0.075}{(\log(Re) - 2)^2} \quad (4)$$

With changes found, the open water efficiency for a fouled propeller can be found with the propeller thrust and torque coefficient for rough conditions.

Last, it can be noted that while marine biofouling mainly has an influence on the frictional resistance of the ship, it also has some effects on the wave resistance. One of the key findings by [25] is the decreasing wave resistance with an increasing surface roughness. This trend was later also found and confirmed by others [10], [26], [27]. It is important to mention that these findings go against the traditional view that wave-making resistance is not affected by hull-roughness [28]. The author proposes to use changes in wave coefficients based on [26], and interpolate and extrapolate these between researched speeds and equivalent sand roughness heights.

With a resistance prediction for smooth ship outlined and fouling growth and effects predicted, next a fouled ship power prediction can be made. First, the total resistance (R_T) can be found based on the calm water resistance (R_{calm}), the air drag resistance (R_{AA}), the wave resistance including changes due to biofouling ($R_W + \Delta R_W$), friction changes due to temperature ($\Delta R_{\Delta T}$), changes due to displacement ($R_{\Delta D}$), and added frictional resistance due to biofouling (ΔR_F), as shown in Equation 5

$$R_T = R_{calm} + R_{AA} + R_W + \Delta R_W + \Delta R_{\Delta T} + \Delta R_{\Delta D} + \Delta R_F \quad (5)$$

Next, the fouled brake power (P_{BR}) can be predicted, with the help of the found total resistance for the fouled situation together with ship speed, hull efficiency (η_H), rough open water efficiency (η_{OR}), relative rotative efficiency (η_R), propulsive efficiency (η_D), gearbox efficiency (η_{GB}) and shaft efficiency (η_S) as shown in Equations 6 and 7.

$$P_E = R_T V_S \quad (6)$$

$$P_{BR} = \frac{P_E}{\eta_H \cdot \eta_{OR} \cdot \eta_R \cdot \eta_{GB} \cdot \eta_S} \quad (7)$$

3 GREY-BOX MODEL APPROACH

In this section, a comprehensive description of the grey-box model applied for predicting marine biofouling growth and its effects on ship performance is provided. First, the underlying principles of DELM are detailed in Section 3.1, while the characteristics and parameters of the model's input are described in Section 3.2.

3.1 Deep Extreme Learning Machine

The task of predicting marine biofouling growth and its subsequent impact on ship performance, based on the data delineated in Section 1, can be mapped into the classical ML regression problem [29].

To better comprehend the aforementioned problem, let's recall the fundamental concepts of the ML regression problem. Let us define $\mathcal{X} \subseteq \mathbb{R}^d$ as the input space composed of d distinct features, and $\mathcal{Y} \subseteq \mathbb{R}$ as the corresponding output space.

Consider a sequence of $n \in \mathbb{N}^*$ distinct samples, symbolized as $\mathcal{D}_n = \{(x_1, y_1), \dots, (x_n, y_n)\}$, where each $x_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$ for all $i \in 1, \dots, n$. These samples are independently drawn from an undefined probability distribution μ encompassing $\mathcal{X} \times \mathcal{Y}$. Within this scenario, we opt for a function (or model) $f : \mathcal{X} \rightarrow \mathcal{Y}$ from a set \mathcal{F} of potential models. An algorithm, characterized by its hyperparameters \mathcal{H} and denoted as $\mathcal{A}_{\mathcal{H}} : \mathcal{D}_n \times \mathcal{F} \rightarrow f$, is employed to choose a model from the suite of possible choices, guided by the available dataset.

The efficacy of the function f in modeling the unobserved system \mathfrak{S} is evaluated by employing a predetermined loss function, $\ell : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, \infty)$. Given that the issue at hand is one of regression, the most fitting choice for the loss function is the squared loss, expressed as $\ell(f(\mathbf{x}), y) = [f(\mathbf{x}) - y]^2$ [30]. Consequently, we can define the true error, or the generalization error, of f as

$$L(f) = \mathbb{E}_{(\mathbf{x}, y)} \ell(f(\mathbf{x}), y). \quad (8)$$

Since $L(f)$ cannot be computed, its empirical estimator (the empirical error) can be derived as follow

$$\hat{L}(f) = \frac{1}{n} \sum_{i=1}^n \ell(f(\mathbf{x}_i), y_i). \quad (9)$$

When it comes to the selection of an algorithm, this paper capitalizes on the DELM. Various algorithms for tackling regression problems abound in the existing literature [29]. In particular, three principal categories of methods have demonstrated practical effectiveness [29], [31], [32]: kernel methods, ensemble methods, and neural networks. In our study, we leverage insights from [11] to adopt a specific subset of neural networks, namely, the DELM [33]. DELM represents an evolution from the Shallow Extreme Learning Machine (SELM), developed for single-hidden-layer feedforward neural networks, with the aim of creating an algorithm capable of not only learning new features from available raw variables but also building a robust regression model.

SELM were originally developed for the single-hidden-layer feedforward neural networks

$$f(\mathbf{x}) = \sum_{i=1}^h w_i g_i(\mathbf{x}). \quad (10)$$

Here, $g_i : \mathbb{R}^d \rightarrow \mathbb{R}$, $i \in 1, \dots, h$ denotes the output from the hidden layer corresponding to the input sample $\mathbf{x} \in \mathbb{R}^d$, while $\mathbf{w} \in \mathbb{R}^h$ represents the output weight vector linking the hidden layer to the output layer.

The input layer, equipped with d neurons, communicates with the hidden layer (which has h neurons) via a set of weights $W \in \mathbb{R}^{h \times d}$ and a nonlinear activation function. For this study, we selected the tanh function as the activation function, as suggested in the seminal work of [33]. It is worth noting, however, that the choice of other activation functions, such as the sigmoid function, doesn't significantly impact the final performance, $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. Consequently, the response of the i -th hidden neuron to an input stimulus \mathbf{x} is given by:

$$g_i(\mathbf{x}) = \varphi \left(W_{i,0} + \sum_{j=1}^d W_{i,j} x_j \right). \quad (11)$$

In SELM, the parameters W are randomly assigned. A weight vector, $\mathbf{w} \in \mathbb{R}^h$, devoid of any bias, bridges the hidden neurons to the output neuron. The comprehensive output function of the network is given by:

$$f(\mathbf{x}) = \sum_{i=1}^h w_i \varphi \left(W_{i,0} + \sum_{j=1}^d W_{i,j} x_j \right). \quad (12)$$

For practicality, we define an activation matrix, $A \in \mathbb{R}^{n \times h}$, in which the element $A_{i,j}$ signifies the activation value of the j -th hidden neuron for the i -th input pattern. Consequently, the A matrix takes the form:

$$A = \begin{bmatrix} \varphi_1(\mathbf{x}_1) & \dots & \varphi_h(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{x}_n) & \dots & \varphi_h(\mathbf{x}_n) \end{bmatrix}. \quad (13)$$

In SELM models, the weights W are set randomly and remain unmodified, leaving the quantity \mathbf{w} in Eq.(12) as the sole degree of freedom. This circumstance simplifies the training to a direct Regularized Least Squares (RLS) problem [34]:

$$\mathbf{w}^* = \arg \min_{\mathbf{w}} \|A\mathbf{w} - \mathbf{y}\|^2 + \lambda \|\mathbf{w}\|^2, \quad (14)$$

where, $\lambda \in [0, \infty)$ signifies a hyperparameter that requires tuning during the Model Selection (MS) phase [35]. This tuning process establishes a balance between model complexity and accuracy, measured by the square loss and the $L2$ regularizer respectively. As a result, the optimal weight vector, denoted as \mathbf{w}^* , can be determined as follows:

$$\mathbf{w}^* = (A^T A + \lambda I)^+ A^T \mathbf{y}, \quad (15)$$

where $I \in \mathbb{R}^{h \times h}$ denotes an identity matrix, and $(\cdot)^+$ refers to the Moore-Penrose matrix pseudoinverse. It's crucial to note that h , the count of hidden neurons, is another hyperparameter requiring fine-tuning, based on the specific problem under consideration. Additionally, other regularizers, such as sparse regularizers, can be employed [36].

Given its shallow architecture, SELM might not offer efficient feature learning, even when h is large. As feature learning frequently enhances the final model's accuracy, multi-layer (deep) solutions are often required. In this context, [37] develops multi-layer learning architectures using ELM-based autoencoder (AE) as the fundamental building block, leading to the creation of DELM. In a DELM, each layer i out of the l layers — each composed of $h_{i \in 1, \dots, l}$ neurons — strives to reconstruct the input data. The outputs from the previous layer are then utilized as inputs for the next. Consequently, instead of yielding a single output, a sequence of outputs \hat{x}_j is obtained, with $j \in 1, \dots, d$, such that:

$$\hat{x}_j = \sum_{i=1}^h w_{i,j} \varphi \left(W_{i,0} + \sum_{j=1}^d W_{i,j} x_j \right) \quad (16)$$

$$\hat{x}_j = \sum_{i=1}^h w_{i,j} \varphi_i(\mathbf{x}), \quad (17)$$

where $w_{i,j}$ with $i \in \{1, \dots, h\}$ are found with the same approach of SELM.

In the DELM model, before the supervised regularized least mean square optimization occurs, the encoded outputs are directly channeled to the last layer for decision-making, bypassing any random feature mapping. Unlike SELM, DELM doesn't necessitate fine-tuning for the entire system, enabling much faster training speed than traditional backpropagation-based Deep Learning. Training a DELM is essentially equivalent to training multiple SELMs. Therefore, the advantages of a deep architecture can be harnessed using only the optimization tools designed for the SELM. It's worth noting that the DELM model encompasses numerous hyperparameters: the number of layers, the number of nodes per layer, and the regularization coefficient,

expressed as $\mathcal{H} = \{l, h_1, \dots, h_l, \lambda\}$. These parameters must be carefully fine-tuned to minimize the final model's generalization error. Accordingly, a model selection phase consistent with [38] has been conducted in this study.

We utilize the nonparametric Bootstrap approach for model selection, a frequently implemented method within the resampling method family. The original dataset \mathcal{D}_n has been resampled once or several times (n_r), either with or without replacement, to create two independent datasets — the training set $\mathcal{L}_{n_l}^r$ and the validation set $\mathcal{V}_{n_v}^r$, respectively. Here, $r \in \{1, \dots, n_r\}$ and the two sets are mutually exclusive and collectively exhaustive: $\mathcal{L}_{n_l}^r \cap \mathcal{V}_{n_v}^r = \emptyset$, $\mathcal{L}_{n_l}^r \cup \mathcal{V}_{n_v}^r = \mathcal{D}_n$. Following this, to perform the model selection phase and identify the optimal combination of hyperparameters \mathcal{H} from a set of possible ones $\mathcal{S}_{\mathcal{H}} = \{\mathcal{H}_1, \mathcal{H}_2, \dots\}$ for the algorithm $\mathcal{A}_{\mathcal{H}}$, we apply the subsequent procedure:

$$\mathcal{H}^* : \min_{\mathcal{H} \in \mathcal{S}_{\mathcal{H}}} \frac{1}{n_r} \sum_{r=1}^{n_r} \frac{1}{n_v} \sum_{(\mathbf{x}, y) \in \mathcal{V}_{n_v}^r} \ell(\mathcal{A}_{\mathcal{H}, \mathcal{L}_{n_l}^r}(\mathbf{x}), y), \quad (18)$$

where $\mathcal{A}_{\mathcal{H}, \mathcal{L}_{n_l}^r}$ is a model built with the algorithm $\mathcal{A}_{\mathcal{H}}$ trained with $\mathcal{L}_{n_l}^r$. Since the data in $\mathcal{L}_{n_l}^r$ is independent of that in $\mathcal{V}_{n_v}^r$, the optimized hyperparameters \mathcal{H}^* should achieve low error rates on a dataset distinct from the one used for training. It's important to note that the nonparametric Bootstrap approach differs from other resampling methods in two key aspects: firstly, $n_l = n$ and secondly, $\mathcal{L}_{n_l}^r$ is sampled with replacement from \mathcal{D}_n . We also note that $\mathcal{V}_{n_v}^r$ is the complement of $\mathcal{L}_{n_l}^r$ within \mathcal{D}_n , that is, $\mathcal{V}_{n_v}^r = \mathcal{D}_n \setminus \mathcal{L}_{n_l}^r$.

3.2 Model Input

Using the selected grey-box approach, all available data, as previously described, has been incorporated into the grey-box model, along with the estimate found from the white-box prediction, following the serial grey-box configuration of [39]. In the white-box biofouling growth approach, the fouling at each anchorage was predicted in conjunction with its corresponding average sea surface temperature, therefore, for the grey-box model, the sea surface temperature is used (see Fig 1). Note that sensor data for the sea surface temperature was not available for the full investigated period, when this was the case an additional dataset with the ships location and time was employed for an estimate of this parameter. The white-box approach contains both the total anchorage days since cleaning that

are processed per anchorage and the sailing days since cleaning, and the biofouling growth model does not have a prediction of how this changes during sailing, as the model is based only on static tests. Nonetheless, this parameter is entered into the grey-box model so that possible fouling changes during sailing can be found. Last, the delivered brake power by the ship is taken as the model output at different speeds (Fig 2). An overview of the input and output parameters is shown in Table 1.

Figure 1: Sea surface temperature over time.

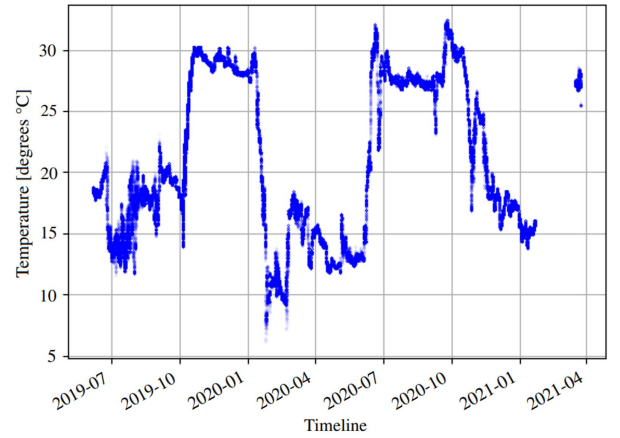


Figure 2: Ship speed vs power.

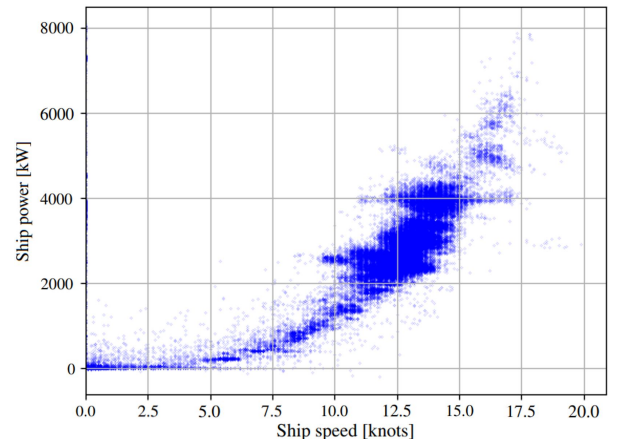


Table 1: Input and output for grey-box model.

Input	Unit
White-box prediction	[kW]
Ship speed	[knots]
Ship direction	[°]
Wave height	[m]
Wave direction	[°]
Wave period	[s]
Wind speed	[m/s]
Wind direction	[°]
Average sea surface temperature	[°C]
Anchorage days since clean ship	[days]
Sailing days since clean ship	[days]
Average ship speed	[knots]
Mean roll	[°]
Mean pitch	[°]
Roll deviation	[°/s]
Pitch deviation	[°/s]
Output	Unit
Fouled ship power	[kW]

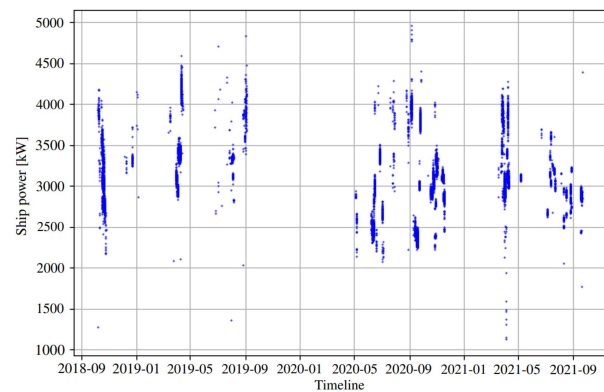
Once input and output have been defined, the data still requires preparation for use in the grey-box model. To achieve this, the data was filtered using Chauvenet’s criterion [40]. Moreover, a filtering technique that incorporates engineering knowledge to curate the data has been applied. The data selection for these predictions was determined by specific criteria related to ship speed and wind conditions. For ship speed, we incorporated all data exceeding 10 knots. Meanwhile, for wind speed, we restricted the dataset to conditions below 8 m/s to exclude instances of severe weather, which render ship performance difficult to predict. In addition to these factors, changes in ship speed were also considered. Significant speed fluctuations in a short period of time typically indicate that the ship is either accelerating or decelerating. As the current methodology does not account for these motions, they could potentially lead to inaccurate power predictions and adversely affect the quality of analysis data. Therefore, with data sampling every three minutes, we chose to analyze only those ship speeds where the preceding speed was within a one-knot range. This approach led to a total number of 34,448 samples. This approach was adopted based on the study cited in [11].

4 RESULTS AND DISCUSSION

In this section, we will evaluate the performance of the proposed grey-box approach utilizing the validation techniques outlined in Section 3. First, we examine the accuracy of the white-box model,

detailed in Section 2, that forms the basis of the grey-box construction. We present its performance through both quantitative and qualitative metrics. The dataset was filtered to focus on a fixed speed of 13.5 knots, one of the most frequently attained speeds, using for the overall prediction accuracy the mean absolute percentage error (MAPE). To give better insight into which extent this number is true, a 95% confidence level interval has been applied. Upon inspection of this filtered data, a noticeable drift in power increase during operation can be discerned, as illustrated in Fig 3. Moreover, after cleaning periods in early 2020, we observe a decrement in power, reverting back to lower values. This observation substantiates the presumption that the visible power increase is attributable to marine biofouling, rather than, for instance, the loss of performances of other propulsion system components.

Figure 3: Power usage over time for 13.5 knots.



The white-box model’s predictions showcased an accuracy of 85%, as shown in Figure 4. This model, grounded in established physical laws and principles, possesses the advantage of being interpretable and reliable under conditions akin to the ones it was formulated for. However, our analysis revealed a consistent trend of underestimation. Even in scenarios shortly following cleaning procedures, the white-box model’s predictions fell short of the actual measurements. This systematic bias suggests that certain aspects of the resistance are not being fully accounted for within this model. The degree of this bias is expected to vary across different vessels, underscoring the inherent challenge in formulating a general model suitable for all ship types. Therefore, a pertinent line of inquiry is to determine the extent to which the white-box method can capture the intricacies of ship power prediction. In addition, it’s worth investigating whether the introduction of additional parameters or refinement of existing ones

can address this model’s underestimation bias.

Moving onto the black-box model, this approach exhibited an accuracy of 90%, as shown in Fig 5. This model has a similar input as the grey-box model as illustrated in Table 1, except for the white-box prediction as an input parameter. Black-box models, being purely data-driven, tend to be highly flexible and can potentially model complex, nonlinear relationships that the white-box models may not adequately capture. However, they are susceptible to high variance error, a problem often associated with overfitting. Thus, although it delivers a marginally improved accuracy compared to the white-box model, it’s crucial to evaluate its performance over a wide range of operational conditions and for different ship types to ensure robustness.

Lastly, the grey-box model, a fusion of the principles-based approach of the white-box model and the empirical learning of the black-box model, accomplished the highest prediction accuracy of almost 92%, as shown in Figure 6. By integrating known physical relationships and data-driven elements, the grey-box model can effectively strike a balance between bias and variance, consequently leading to a superior predictive performance. Notably, another crucial advantage of the grey-box model lies in its efficiency with respect to data requirements. While data-driven models, like the black-box model, typically necessitate large quantities of data to attain high accuracy, the incorporation of known physical laws in the grey-box model allows it to achieve comparable, if not superior, performance levels with significantly less data, as it has prior knowledge/assumptions about the problem domain. This trait makes the grey-box model particularly attractive in scenarios where data collection may be expensive, time-consuming, or otherwise challenging. However, caution is advised to avoid over-dependence on the data-driven component in the grey-box model, as this could veer the model towards overfitting. Thus, while leveraging the flexibility offered by the data-driven component, it is critical to continually reference and respect the governing physical principles to maintain the robustness and generalizability of the model.

A comparative analysis of the models, as summarized in Table 2, clearly indicates a hierarchy in performance over the test dataset: Grey-Box > Black-Box > White-Box. Yet, it’s important to note that each model’s performance should not be evaluated merely on accuracy, but also on bias-variance tradeoff, interpretability, and applicability to diverse ship types and operational conditions.

Table 2: Performance comparison

Model	MAPE
White-Box	14.7% ± 0.3%
Black-Box	10.2% ± 0.2%
Grey-Box	8.6% ± 0.2%

Figure 4: Physics-based model predicted and measured power.

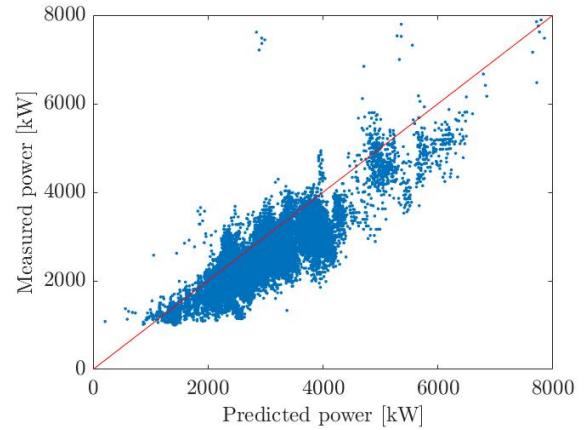


Figure 5: Black-box predicted and measured power.

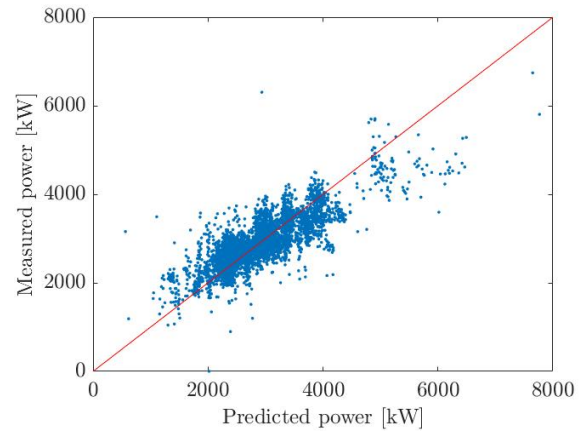
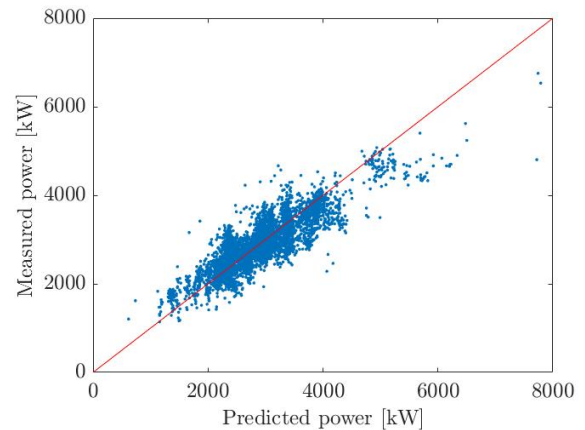


Figure 6: Grey-box predicted and measured power.



In order to assess the models' predictive performance beyond the confines of the data used for training, validation, and test, a preliminary evaluation was conducted. Fig 7 presents an application of the model over an extended period. To facilitate this, all parameters, with the exception of anchorage days, sailing days, and the predictions from the white-box model, were assumed to be at their mean value. The yacht's activity, characterized by sailing 16% of the time, was used to establish the correlation between anchorage and sailing days. Fig 7 illustrates the mean predictions for the black-box and grey-box models with a line, while the prediction area represents the interval of confidence of the results over 30 repetitions. These repetitions are performed, as deep learning models are stochastic models, making use of randomness while being fit on the data. Employing this approach provides insights into the areas where the predictions are mostly congruent and the regions that exhibit significant inter-model variance in predictions. It highlights models ability to capture the inherent uncertainties associated with real-world conditions, thereby lending confidence to their utility in practical scenarios. Last, the trained anchorage days range is also highlighted in the figure, as it shows a more clear division between the interpolation and extrapolation capacities of the models employed.

Figure 7: Comparison between white-, black-, and grey-box prediction for longer period of time.

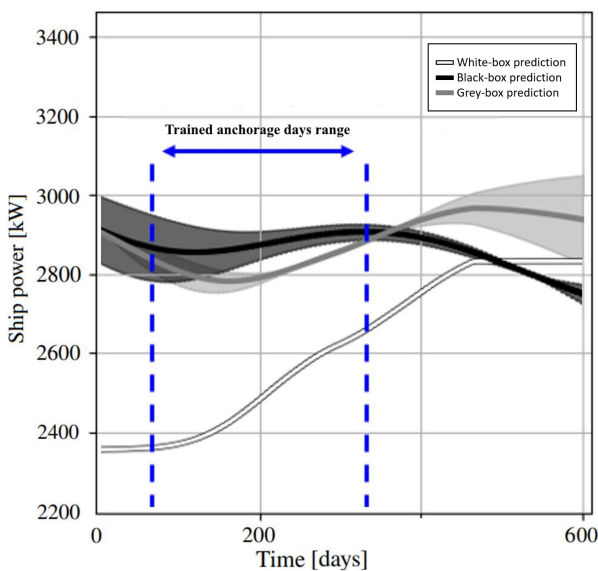


Fig 7 illustrates that the black-box and grey-box models display comparable trends within the range of trained data, whereas the white-box model consistently generates lower predictions. However, upon extrapolating the data-driven models, a sig-

nificant degree of variance and potential inaccuracy in their results becomes evident. Such observations align with the common understanding that data-driven models often grapple with challenges in the context of extrapolation. The exhibited large variance and outcomes underscore the limitations of such models when applied beyond the bounds of their training data. This highlights the limitations for application of the proposed models, particularly in scenarios when extreme extrapolation is necessary.

Marine fouling growth encompasses various stages, such as the formation of slime, the growth of non-shell organisms, and the occurrence of calcareous fouling. However, given that our model was solely trained on data from the initial year of fouling growth, it remains unfamiliar with the settlement of barnacles and the onset of calcareous fouling. This unfamiliarity becomes evident when the white-box model indicates a significant power increase at more advanced stages of fouling. While the grey-box model attempts to harness the strengths of both white- and black-box models, its effectiveness seems to primarily lie in improving upon the black-box model predictions. Conversely, the white-box model still appears to have the upper hand in extrapolation and in providing transparent and interpretable insights into the fouling process. This underscores the need for a more comprehensive training dataset that spans across the various stages of marine fouling, thereby improving the predictive capacity of the model for long-term and advanced fouling situations.

The white-box model, as employed in this study, is indeed a practical solution that assimilates some of the most relevant research pertinent to this field of application. However, it's critical to underline that this model, despite its theoretical robustness, manifests certain limitations in terms of accuracy. Its predictive performance is low, and the model also demonstrates a higher level of variance, introducing a degree of uncertainty into the predictions. This model also exhibits substantial bias error, which may significantly skew the predictions and result in systematic deviations from the actual data. These aspects, coupled with the limitations in addressing complex, multi-staged fouling processes, call for refinement and enhancement of the white-box model to ensure it's aptly equipped to handle the intricacies of this application.

This observed scatter can be attributed primarily to the inconsistencies between the ship's power, which tends to remain stable, and the ship's speed,

which fluctuates. This disparity introduces inaccuracies across the dataset, despite the overall trend displaying promising prediction potential. Another contributing factor to the data scatter is the derivation process for the ship's speed. Initially extracted from the Automatic Identification System (AIS), the speed was then computed over water using predicted currents. However, it's important to acknowledge that directly measuring the speed over water could potentially improve prediction accuracy. Even slight variations in speed can significantly impact the predicted calm water resistance and the power required, underscoring the need for precise speed measurements in enhancing the predictive accuracy of these models.

5 CONCLUSIONS

The research outlined in this paper introduces a grey-box modeling strategy aimed at predicting marine biofouling growth and its subsequent impact on ship performance. This approach synergistically integrates the empirical insights of white-box models and the data-intensive capabilities of black-box models. Initially, a white-box model is presented to predict ship resistance, considering variables such as calm water resistance, wind, waves, and temperature discrepancies. The prediction of marine biofouling growth is subsequently managed via an experimental model designed to estimate the roughness level of the ship hull. Lastly, a Deep Extreme Learning Machine is utilized as a black-box model, which incorporates a feedforward neural network technique.

With the proposed approach certain limitations inherent in the white-box model can be partially mitigated through the application of trained data-driven models. These models possess the capacity to understand how available input can be utilized to enhance the predictive accuracy of the white-box models.

Moreover, in situations where a fleet of similar ships is available, the correlations discerned by the grey-box model can be leveraged to refine the predictions of full white-box models for other ships. This can be achieved even in the absence of trainable data. This highlights the potential of the grey-box model as a powerful predictive tool that can efficiently use available information to enhance prediction accuracy, thus providing a robust method that can be applied across multiple ships, irrespective of the available data for each vessel.

To validate the proposed approach, a superyacht

case study was chosen as the testing ground due to the category's high susceptibility to fouling. This involved the analysis of a 2-year dataset acquired through a partnership with Feadship.

The results presented highlighted that while the black-box model displayed superior predictive capabilities compared to its white-box counterpart, it was the grey-box model that exhibited the best performance. By incorporating the knowledge encapsulated within the white-box model, the grey-box model displayed the highest prediction accuracy whilst requiring less historical data. These findings underline the potential of the grey-box model as a tool for accurately predicting marine biofouling growth and its impact on ship performance. The practical implications of this research extend to ship operators and designers, who could leverage these insights to enhance operational efficiency and minimize maintenance costs.

The results also suggest for a more nuanced understanding of model selection and design in predictive tasks. While black-box models may often deliver higher prediction accuracy in situations where ample historical data is available, the incorporation of domain knowledge via grey-box models can help achieve similar, if not better, performance with less data. This highlights the importance of continued research in hybrid modelling techniques, especially in domains where data may be costly or difficult to acquire. While the data-driven models, and especially the grey-box model showed the highest potential for the capture of the biofouling modeling and ship power prediction within the researched period, its lacking extrapolation capacity was also identified. Enabling the dataset with more vessels, a wider range of operational profiles, and later biofouling growth stages such as barnacle growth, would improve model performance.

From the perspective of ship design, the current approaches towards predicting fouling are either absent or rely heavily on rudimentary approximations. The primary assessments for powering calculations and ship speed are generally conducted for clean hulls, given that the biofouling issue is predominantly addressed during maintenance and operational phases rather than the design stage. Nonetheless, pertinent information that could allow for preliminary predictions on biofouling development is often available even at the early stages of ship design. The present research was initiated with the intention to incorporate and utilize this information effectively. Utilizing the proposed model, it is possible to craft ship-specific predictions grounded on

the prospective operational profile and forecasted environmental conditions. This, in turn, facilitates more informed decisions regarding the selection of antifouling systems and the determination of necessary engine margins or propulsion layouts. When these predictions are coupled with the operational profile of a yacht, it is possible to calculate the fuel penalty in conjunction with the costs of antifouling, docking, and cleaning.

ACKNOWLEDGEMENTS

The authors are grateful for the support of Feadship / De Voogt Naval Architects as the research presented in this paper was generated as part of their sea margin project, where yacht sensor data was shared together with in-house methods.

REFERENCES

- [1] C. Bressy and M. Lejars, "Marine fouling : An overview," *Journal of Ocean Technology*, vol. 9, pp. 19–28, Dec. 2014.
- [2] Y. Demirel, O. Turan, and A. Incecik, "Predicting the effect of biofouling on ship resistance using CFD," *Applied Ocean Research*, vol. 62, pp. 100–118, 2017.
- [3] S. A. Alghamdi and R. A. Quijada Cordova, "The impact of biofouling on marine environment: A qualitative review of the current antifouling technologies," M.S. thesis, World Maritime University, Malmö, Sweden, 2019.
- [4] L. Kjær, D. Pigosso, T. McAloone, and M. Birkved, "Guidelines for evaluating the environmental performance of product/service-systems through life cycle assessment," *Journal of Cleaner Production*, vol. 190, Apr. 2018.
- [5] ITTC Technical Committee on Resistance and Propulsion. "Review of state of the art." (2021), [Online]. Available: <https://ittc.info/media/9825/resistance-and-propulsion-committee.pdf>.
- [6] Propulsion Committee of the 28th ITTC, "1978 ITTC Performance Prediction Method," *ITTC - Recommended Procedures and Guidelines*, 2017.
- [7] A. Farkas, N. Degiuli, I. Martić, and I. Ančić, "Performance prediction method for fouled surfaces," *Applied Ocean Research*, vol. 99, Apr. 2020.
- [8] A. Farkas, N. Degiuli, and I. Martić, "The impact of biofouling on the propeller performance," *Ocean Engineering*, vol. 219, p. 108376, Nov. 2020.
- [9] D. Owen, Y. Demirel, E. Oguz, T. Tezdogan, and A. Incecik, "Investigating the effect of biofouling on propeller characteristics using cfd," *Ocean Engineering*, vol. 159, Feb. 2018.
- [10] S. Song, Y. Demirel, C. De Marco, T. Tezdogan, and M. Atlar, "Fouling effect on the resistance of different ship types," *Ocean Engineering*, vol. 216, p. 107736, 2020.
- [11] A. Coraddu, L. Oneto, F. Baldi, F. Cipollini, M. Atlar, and s. Savio, "Data-driven ship digital twin for estimating the speed loss caused by the marine fouling," *Ocean Engineering*, vol. 186, p. 106063, 2019.
- [12] Y. K. Demirel, S. Song, O. Turan, and A. Incecik, "Practical added resistance diagrams to predict fouling impact on ship performance," *Ocean Engineering*, vol. 186, p. 106112, 2019, ISSN: 0029-8018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0029801819303087>.
- [13] D. Uzun, Y. Demirel, A. Coraddu, and O. Turan, "Time-dependent biofouling growth model for predicting the effects of biofouling on ship resistance and powering," *Ocean Engineering*, vol. 191, 2019.
- [14] Y. Demirel, M. Khorasanchi, O. Turan, A. Incecik, and M. Schultz, "A CFD model for the frictional resistance prediction of antifouling coatings," *Ocean Engineering*, vol. 89, pp. 21–31, 2014.
- [15] J. P. Petersen, D. J. Jacobsen, and O. Winther, "Statistical modelling for ship propulsion efficiency," *Journal of marine science and technology*, vol. 17, no. 1, pp. 30–39, 2012.
- [16] G.-B. Huang, "An insight into extreme learning machines: Random neurons, random features and kernels," *Cognitive Computation*, vol. 6, no. 3, pp. 376–390, 2014.
- [17] L. Chambers, K. Stokes, F. Walsh, and R. Wood, "Modern approaches to marine antifouling coatings," *Surface and Coatings Technology*, vol. 201, pp. 3642–3652, 2006.
- [18] ITTC, "Analysis of Speed/Power Trial Data," *ITTC - Recommended Procedures and Guidelines*, 2014.
- [19] R. Grin, "On the prediction of wave-added resistance with empirical methods," *Journal of Ship Production and Design*, vol. 31, no. 03, pp. 181–191, 2015.
- [20] W. J. Pierson Jr and L. Moskowitz, "A proposed spectral form for fully developed wind seas based on the similarity theory of sa kitaigorodskii," *Journal of geophysical research*, vol. 69, no. 24, pp. 5181–5190, 1964.
- [21] K. Hasselmann, T. P. Barnett, E. Bouws, *et al.*, "Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap).," *Ergaenzungsheft zur Deutschen Hydrographischen Zeitschrift, Reihe A*, 1973.
- [22] 26th ITTC Specialist Committee on Uncertainty Analysis, "Fresh Water and Seawater Properties," *ITTC - Recommended Procedures*, 2011.
- [23] R. Townsin, "The ship hull fouling penalty," *Biofouling*, vol. 19, no. S1, pp. 9–15, 2003.

- [24] ITTC, "Recommended Procedures and Guideline Practical Guidelines for Ship CFD Applications," *ITTC - Recommended Procedures and Guidelines*, 2011.
- [25] Y. Demirel, M. Khorasanchi, O. Turan, and A. Incecik, "On the importance of antifouling coatings regarding ship resistance and powering," in *International Conference on Technologies, Operations, Logistics and Modelling for Low Carbon Shipping*, 2013.
- [26] D. Oliveira, A. I. Larsson, and L. Granhag, "Effect of ship hull form on the resistance penalty from biofouling," *Biofouling*, vol. 34, no. 3, pp. 262–272, 2018.
- [27] P. Kanninen, P. Peltonen, and V. Vuorinen, "Full-scale ship stern wave with the modelled and resolved turbulence including the hull roughness effect," *Ocean Engineering*, vol. 245, p. 110434, 2022.
- [28] W. H. O. Institution and U. S. N. D. B. of Ships, *Marine fouling and its prevention*. United States Naval Institute, 1952.
- [29] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.
- [30] L. Rosasco, E. De Vito, A. Caponnetto, M. Piana, and A. Verri, "Are loss functions all the same?" *Neural Computation*, vol. 16, no. 5, pp. 1063–1076, 2004.
- [31] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?" *The journal of machine learning research*, vol. 15, no. 1, pp. 3133–3181, 2014.
- [32] M. Wainberg, B. Alipanahi, and B. J. Frey, "Are random forests truly the best classifiers?" *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 3837–3841, 2016.
- [33] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *IEEE International Joint Conference on Neural Networks*, 2004.
- [34] A. Caponnetto and E. De Vito, "Optimal rates for the regularized least-squares algorithm," *Foundations of Computational Mathematics*, vol. 7, no. 3, pp. 331–368, 2007.
- [35] L. Oneto, "Model selection and error estimation without the agonizing pain," *WIREs Data Mining and Knowledge Discovery*, 2018.
- [36] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
- [37] L. L. C. Kasun, H. Zhou, G.-B. Huang, and C. M. Vong, "Representational learning with elms for big data," *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 31–34, 2013.
- [38] L. Oneto, "Model selection and error estimation without the agonizing pain," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, e1252, 2018.
- [39] L. Leifsson, H. Sævarsdóttir, S. Sigurdsson, and A. Vésteinsson, "Grey-box modeling of an ocean vessel for operational optimization," *Simulation Modelling Practice and Theory*, vol. 16, pp. 923–932, Sep. 2008.
- [40] A. F. R. F. Rochim, *Chauvenet's criterion, peirce's criterion, and thompson's criterion (literatures review)*, 2021.