

RETROFIT SOLUTIONS TO ACHIEVE 55% GHG REDUCTION BY 2030

Hull and propeller performance monitoring tool

WP 3 – Operational Synthesis and Optimization
Task 3.1 – Hull and propeller condition monitoring tool
D3.1 – Hull and propeller performance monitoring tool
Partners involved: SFWD , NTUA, CNR, AALTO, LASK, RINA, ARM, ATD
Authors: George Nikolaidis (SFWD), Vasilis Zagkas (SFWD), Reuben D'Souza (SFWD), Nikos Themelis (NTUA), Ageliki Kytariolou (NTUA), Mingyang Zhang (AALTO), Spiros Hirdaris (AALTO), Alessandro Maccari (RINA), Nikos Tsoulakos (LASK)



Project details

Project Title	RETROFIT SOLUTIONS TO ACHIEVE 55% GHG REDUCTION BY 2030
Project Type	Innovation Action
Project Acronym	RETROFIT55
Grant Agreement No.	101096068
Duration	36 M
Project Start Date	01/01/2023

Deliverable information

Status (F: final; D: draft; RD: revised draft)	F
Planned delivery date	31/12/2023
Actual delivery date	07/01/2023
<ul style="list-style-type: none"> • Dissemination level: PU – Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project’s page) • SEN – Sensitive, limited under the conditions of the Grant Agreement • Classified R-UE/EU-R – EU RESTRICTED under the Commission Decision No2015/444 • Classified C-UE/EU-C – EU CONFIDENTIAL under the Commission Decision No2015/444 • Classified S-UE/EU-S – EU SECRET under the Commission Decision No2015/444 	PU
Type: Report, Website, Other, Ethics	Report



Document history

Version	Date	Created/Amended by	Changes
01	04/12/2023	G. Nikolaidis, V. Zagkas, R. D'Souza (SFWD)	First draft
02	07/12/2023	A. Maccari (RINA)	Revised and expanded draft
03	15/12/2023	G. Nikolaidis, N. Themelis	New draft
04	19/12/2023	Emanuele Spinosa (CNR)	Format/Style review
05	27/12/2023	G. Nikolaidis, N. Themelis, Minyang Zhang	Review
06	28/12/2023	Emanuele Spinosa (CNR)	Minor corrections
Final	03/01/2024	Alessandro lafrati (CNR), N. Themelis (NTUA)	Final version

Quality check review

Reviewer (s)	Main changes / Actions
Manolis Angelou (NTUA)	Technical revision
Cecilia Leotardi (CNR)	Editorial revision
Cecilia Leotardi (CNR), Alessandro lafrati (CNR)	Final review of contents and submission to EC.



Table of contents

List of figures	6
List of tables.....	8
List of abbreviations	9
Executive Summary	10
1 Introduction	12
2 Literature review	15
2.1 Basics of the biofouling problem	15
2.2 Review of anti-fouling measures and technologies.....	17
2.3 Review of studies related with data-driven models for the performance monitoring	22
3 Collection of ship data: case study MV Kastor (Laskaridis Shipping).....	25
3.1 Criteria for selection of ship	25
3.2 Ship and equipment data	25
3.3 High-frequency data	26
3.4 Noon report data.....	26
3.5 Weather data	27
4 Exploratory data analysis of key parameters	28
4.1 Comparison between specific high-frequency data and noon reports	28
4.2 Loading conditions.....	32
4.3 Identification of maintenance actions using KPI.....	38
5 Data-driven models.....	41
5.1 Methodology	41
5.2 Feature engineering.....	41
5.3 Data preparation	43
5.4 Machine Learning (ML) model development.....	46
6 Optimization maintenance analysis.....	52
6.1 Method-1 for maintenance optimization.....	52
6.2 Method-2 for maintenance optimization.....	54
6.3 Estimation of added fuel cost.....	57
7 Hull condition-based monitoring and the regulatory framework	59
8 Concluding remarks	63
Appendix A: Decision Trees.....	67





Appendix B: Random Forests	69
Appendix C: Extremely Randomized Trees	70
Appendix D: Gradient Tree Boosting	71
Appendix E: Distance Correlation	72
Appendix F: List of high-frequency parameters selected for data analysis	73
Appendix G: An alternative approach based on fuel consumption prediction	80



List of figures

Figure 1: Macro-organisms that can adhere to submerged surfaces in the Mediterranean Sea (source: LABORATOIRE MAPIEM, [14]).....	15
Figure 2: (a) Global distribution of average annual sea-surface temperatures, (b) Distribution of salinity in Earth’s oceans and major seas, [17].....	16
Figure 3: Hull cleaning methods: (a) Manual cleaning (b) powered rotary brush systems, (c) innovative noncontact cleaning technologies.	19
Figure 4: Noon reports Mean Draft plotted against high-frequency Mean Draft.	28
Figure 5: Relative frequency histogram of difference between the Mean Draft from high frequency data and noon reports.	29
Figure 6: Sea height from noon reports plotted against significant height from the weather provider.	29
Figure 7: Relative frequency histogram of the difference between the significant wave height from the weather provider and sea height from noon reports	30
Figure 8: Relative frequency histogram of sea height from noon reports.....	30
Figure 9: Projection of the ship path on the world map during the recording period, utilizing high-frequency GPS signals.....	31
Figure 10: Overlaid histograms of the STW (Speed Through Water), SOG (Speed Over Ground) and of the vessel speed from noon reports.	32
Figure 11: Relative frequency histogram of sea current speed, calculated as the difference between the STW and the SOG	32
Figure 12: Relative frequency histogram of the Mean Draft from noon reports.....	33
Figure 13: Relative frequency histogram of the trim from noon reports.	34
Figure 14: Mean Draft plotted against trim using separate colours for the laden (red) and ballast condition (blue).	34
Figure 15: Relative frequency histogram of STW for loading condition 1 (laden).....	35
Figure 16: Relative frequency histogram of trim for loading condition 1 (laden).....	35
Figure 17: Relative frequency histogram of STW for loading condition 2 (laden).....	36
Figure 18: Relative frequency histogram of trim for loading condition 2 (laden).....	36
Figure 19: Relative frequency histogram of STW for loading condition 3 (ballast).	37
Figure 20: Relative frequency histogram of trim for loading condition 3 (ballast).	37
Figure 21: Relative frequency histogram of the STW for loading condition 4 (heavy ballast).	38
Figure 22: Relative frequency histogram of trim for loading condition 4 (heavy ballast).....	38
Figure 23: Timeline of the most important events during the vessel operation.	39
Figure 24: Utilization of KPI to monitor the status of the hull and the propeller performance, as well as to track the main maintenance events.	40
Figure 25: Steps of the development of the Machine Learning (ML) method to build the hull and propeller monitoring tool.....	41
Figure 26: Bar plot of the distance correlation coefficient between SHP and the considered features.	43
Figure 27: Comparison between the SHP signals over a one-day time interval, after smoothing with Simple Moving Average with different window sizes.	44
Figure 28: Comparison between the STW signals over a one-day time interval, and the signals obtained by using different values of k	45
Figure 29: Relative frequency histogram of SHP, then used to identify the tails of the statistical distribution of SHP.	45
Figure 30: Reduction in the number of datapoints at each stage of data preparation for the entire dataset.....	46

Figure 31: Flowchart illustrating the training and application process of Model-1 and Model-2 on separate datasets, within the context of hull and propeller condition monitoring.....	47
Figure 32: Plot of the predicted SHP values against the actual SHP values of the test set: (a) for Model-1, and (b) for Model-2.....	50
Figure 33: Plot of residual SHP values against the predicted SHP values on the test set: (a) for Model-1, and (b) for Model-2.....	51
Figure 34: Application of Model-1 to dataset B. (a): Scatter plot of the actual and predicted SHP over time, (b) scatter plot of the PI over time.....	53
Figure 35: Application of Model-1 to dataset D2. (a): Scatter plot of the actual and predicted SHP over time, (b) scatter plot of the PI over time.....	54
Figure 36: Predicted SHP values on the synthetic datasets plotted against STW, with respective least squares polynomial fits of 3rd degree for (a) laden condition and (b) ballast condition.....	56
Figure 37: Mean predicted SHP values on the synthetic laden and ballast datasets in terms of days, plotted with respective linear fits.	57
Figure 38: Flowchart of the multi-objective optimization method employed to achieve ship fuel consumption reduction based on the prediction model.....	80
Figure 39: Schematic of the deep learning method architecture for the ship fuel consumption prediction.	81
Figure 40: Visualization of the decision tree model with the best-performing hyperparameters obtained from the grid search. The full tree includes 90 levels (depths), for the reader's convenience only the first 4 levels are displayed here.	82
Figure 41: Correlation between the selected influencing factors on the ship fuel consumption, using the collected data streams.	83
Figure 42: Deep learning processing of the ship fuel consumption prediction for model training, testing and application.	84
Figure 43: Ship trajectories of a bulk carrier for training the ship fuel consumption prediction model from 01/2021 to 02/2023.	84
Figure 44: Training and testing loss over Epochs for the evaluation of model performance.	85
Figure 45: Ship trajectories of the bulk carrier for training the ship fuel consumption prediction model from 01/2023 to 06/2023.....	86
Figure 46: Error analysis of the fuel consumption prediction for the voyage 1.....	87
Figure 47: Comparison between the actual and the predicted ship fuel consumption for the eight voyages.	89

List of tables

Table 1: Anti-fouling coatings, based on [25] and expanded.	21
Table 2: Main characteristics of the ship under examination: MV Kastor (Laskaridis).....	25
Table 3: Main characteristics of the Main Engine (ME) of the ship under examination	25
Table 4: Main characteristics of the Auxiliary Engine (AE) of the of the ship under examination: MV Kastor (Laskaridis).....	26
Table 5: Basic information on the available datasets.....	27
Table 6: Definition of wave data parameters from weather providers.	27
Table 7: Descriptive statistics of high-frequency and noon report parameters.....	28
Table 8: Definition of loading conditions using Mean Draft ranges.	33
Table 9: Mean values and ranges of Mean Draft, trim and STW for each defined loading condition.	33
Table 10: Threshold values used for the filtering of the SHP, STW, Rate of Turn and Current speed.	39
Table 11: Definition of datasets based on the dates of the main maintenance events. Datasets D1 and D2, are subsets of dataset D.....	39
Table 12: Calculation of the distance correlation coefficient between SHP and the candidate features.	42
Table 13: List of the selected features with or without the Days feature.	43
Table 14: Threshold values used for filtering the ship navigation variables.	46
Table 15: Overview of the two ML models, their feature sets (see Table 13), and training datasets (see Table 11).....	46
Table 16: Overview of the hyperparameter values explored for each ML algorithm using randomized search combined with cross-validation.....	49
Table 17: Mean cross validated RMSE for each ML algorithm.	49
Table 18: Performance evaluation of Model-1 and Model-2 on the test set, using a variety of error metrics.....	49
Table 19: Application results of Model-1 to dataset B, showing the mean values of the actual and predicted SHP, as well as the mean PI value.....	53
Table 20: Application results of Model-1 to dataset D2, showing the mean values of the actual and predicted SHP and the mean PI value.	54
Table 21: Overview of the synthetic datasets designed separately for laden and ballast conditions.	55
Table 22: Mean predicted SHP and corresponding PI, when Model-2 is applied to the synthetic dataset representing the laden condition.....	55
Table 23: Mean predicted SHP and corresponding PI, when Model-2 is applied to the synthetic dataset representing the ballast condition.	55
Table 24: Estimated added fuel cost per day for the laden and ballast loading conditions, considering different values of the Days feature.....	58
Table 25: Model characteristics and the optimal hyperparameters.	85
Table 26: General information of the extra eight voyages used in the model validation.....	86
Table 27: Generalization ability evaluation of the trained ship fuel consumption prediction model for eight global voyages	87

List of abbreviations

Abbreviation	Definition
AE	Auxiliary Engine
AF	Antifouling
ANN	Artificial Neural network
CAP	Condition Assessment Program
CDP	Controlled Depletion Polymer
DAQ	Data Acquisition
DSC	Days Since Cleaning
EEDI	Energy Efficiency Design Index
FOC	Fuel Oil Consumption
FR	Foul Release
HSSC	Harmonized System of Survey and Certification
HFO	Heavy Fuel Oil
IACS	International Association of Classification Societies
IMO	International Maritime Organization
KPI	Key Performance Indicator
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ME	Main Engine
ML	Machine Learning
MLR	Multiple Linear Regression
NIS	Non-indigenous Species
PI	Power Increase
RIT	Remote Inspection Techniques
RMSE	Root Mean Square Error
RO	Recognized Organization
RoT	Rate of Turn
ROV	Remotely Operated Vehicles
RPM	Revolutions Per Minute
SD	Strategic Directions
SFOC	Specific Fuel Oil Consumption
SHP	Shaft Horsepower
SMA	Simple Moving Average
SOG	Speed Over Ground
SPC	Self-Polishing Copolymer
SSE	Sum of Squared Errors
STW	Speed Through Water
TBT	Tributyltin
TDNN	Tap-Delayed Neural Networks
UAV	Unmanned Autonomous Vehicles
UN	United Nations

Executive Summary

A hull and propeller biofouling condition monitoring framework is presented and demonstrated using as a case study a Kamsarmax size bulk-carrier examining an operation period of 2 and half years. The aim of the framework is to support the decision-making process for optimizing maintenance related to hull and propeller cleaning actions. For the monitoring of the performance an appropriate Key Performance Index (KPI) is defined, comparing the expected shaft power needed from the main engine with the measured one in the same operational conditions. For this calculation, a mathematical model for the prediction of the power is needed. The type of models developed in this study is data-driven and based on Machine Learning (ML) algorithms, due to the large volume of the available operational data from different sources and their improved accuracy.

A detailed review of the types of bio-fouling, the parameters affecting its growth and the methods and technologies of anti-fouling and surface cleaning is presented, to provide background knowledge. Moreover, the key works using ML algorithms for the hull and propeller performance prediction are reviewed.

Before developing the data-driven models, a thorough exploratory analysis of the key operational parameters is carried out. First, a statistical comparison of ship's speed, mean draft and wave height being key operational parameters and derived from different data sources is performed to evaluate their agreement and to justify data imputation choices for the creation of the dataset to be used. Secondly, four dominant loading conditions (two laden and two ballast), in terms of mean draft and trim, were identified and the respective ranges of ship's speed in each one is calculated. These realistic operational conditions will be utilized in the hydrodynamic optimization analysis that will be performed in WP2. Moreover, it is crucial for the development of the prediction models to know the maintenance actions during the monitoring period to define the appropriate datasets. The effect of hull and propeller cleaning events is also verified using as KPI the propeller loading coefficient.

The development of the data-driven models includes the process of feature engineering, which involves the creation of new parameters for a better representation of the physical problem and the selection of the parameters to be used in the models, using as criteria the distance correlation coefficient and domain knowledge. In addition, extensive filtering is applied to focus on meaningful time intervals to the problem, also involving a simple moving average to reduce noise. During the ML development stage, a comprehensive evaluation of various regression algorithms from the Decision Tree family is conducted. Simultaneously, an exploration of the hyperparameter search space takes place. As a result, the Extra-Trees algorithm emerges as the top performer, demonstrating remarkable effectiveness with a Mean Absolute Percentage Error (MAPE) of 1.1% on the test data.

The optimization maintenance analysis incorporates the KPI and two prediction models, corresponding to two different approaches to the problem. In the first, the model is employed as a reference performance baseline representing the clean vessel, to evaluate conditions with fouling. When applied to a dataset reflecting advanced stages of biofouling accumulation, the model predicts that if the vessel were in a clean state and operated under identical conditions, there would be an estimated average decrease of 11.3% in power requirements. In the second approach, the respective model includes the parameter "Days" and thus enables emulation of scenarios that refer to different points in time. To evaluate the fouling impact in time, synthetic datasets corresponding to a range of speeds, focused loadings and weather conditions are derived. By comparing the model predictions on synthetic data 300 days apart, while keeping other parameters constant, it is



determined that the fouled vessel requires an average 20.5% increase in power compared to a clean state.

Finally, to support in more practical terms the decision-making for the suitable time to perform a maintenance action, the analysis proceeds with the prediction of additional daily fuel cost due to hull and propeller fouling. Both approaches have estimated the power increase associated with the accumulation of biofouling over a specific time frame. This additional power requirement can be used to calculate the resulting fuel consumption excess, which directly translates into an added fuel cost per day. Consequently, the analysis reveals that a biofouling build-up over a 300-day period leads to an average daily increase of \$2,249 in fuel cost.



1 Introduction

To mitigate the effects of global climate change, there are increasing pressures on the maritime industry to reduce the environmental footprint of shipping. As early as 2011, the International Maritime Organization (IMO) adopted mandatory measures with the aim of optimizing the energy efficiency of ships. Since then, the IMO has been taking further action. Specifically, the IMO has already set a strategy to reduce greenhouse gases related to ships' operation based on short, medium, and long-term measures [1], while the most recent have been set out in the July 2023 revised strategy [2]. Under the revised strategy [2], the new targets include a 20% reduction in ship emissions by 2030, a 70% reduction by 2040 (compared to 2008 emission levels [3]), with the ultimate goal of near-zero emissions by 2050.

However, beyond the regulatory framework that steers shipping towards more sustainable solutions, a more general interest is observed from the various parties involved (e.g., ship owners and operators, charterers) to optimize the operational performance of ships, the reduction of fuel consumption and, ultimately, the reduction of emissions. As many studies show [4], the fuel cost constitutes by far the greatest part (up to 50- 70%) of the ships' operational cost and, therefore, the reduction of fuel consumption per voyage is a priority for all shipping companies. For example, reducing it just 1% can result in hundreds of thousands of dollars saved per year in the operation of large ships [4][4].

An overview of the reduction potential achieved by the candidate measures has been presented in [5] and [6]. According to [6], measures that control operational efficiency can lead to reductions of CO₂ emissions up to 48-60%. One option to do this is to increase the ship fuel economy, for example by upgrading the engines or retrofitting the ship design. Moreover, measures such as slow-streaming or weather routing are other options to reduce fuel consumption and subsequently Green-House Gas (GHG) emissions.

Therefore, the continuous monitoring of the ship operational performance has been on the spot and has now been established in the industry. This practice is evaluated through Key Performance Indicators (KPIs). In line with this trend, several practices have been implemented for the acquisition and analysis of operational data, to evaluate the ship performance that will assist the informed decision-making process. The data that can help to evaluate the energy efficiency of a ship include data related with the loading (e.g., drafts) and the operational parameters (e.g., speed, shaft power and RPM, ship heading), the weather conditions encountered (wind, waves and currents) as well as parameters related with the main and auxiliary machinery (consumptions and more detailed information, such as temperature, pressures etc.). Monitoring them often indicates clear differences between expected and actual performance throughout the life of the vessel. The collection and evaluation of these data may indicate specific trends in the behavior of the ship, mainly in terms of the power required for propulsion, the ability to reach specific speeds and the fuel consumption. These are factors that are directly linked to ship environmental impact and GHG emissions [4].

As mentioned previously and due to its significant contribution to the operating costs and profitability, the fuel consumption is a crucial aspect that vessel operators must consider while planning and carrying out voyages. The fuel consumption used for propulsion, according to [7], is strongly related to the ship drag forces and in particular to the frictional resistance, which is affected by the vessel hull fouling - biofouling. Marine fouling or biofouling is generated by the buildup of micro - and macro-organisms onto the ship hull, which increases the surface roughness and thus deteriorates the ship performance. The development of biofouling on the hull and propeller has a significant effect on the

overall hydrodynamic performance of a ship with relevant research showing that thin biofilm and algae growth (slime, algae biofilm) cause a reduction in power efficiency of more than 10%, while the growth of shellfish (barnacles) can lead to an 80% increase in power request [8]. The decrease in performance is a consequence of the increase in the roughness of the wetted surface of the hull, of the hull appendages (rudders, shaft brackets, bossings, pods, etc.) and of the propeller blades.

The increase in the roughness of the hull causes an increase in the thickness of the boundary layer and an increase in the frictional resistance to the motion of the ship in water. In addition, increasing the boundary layer thickness reduces the velocity of the water reaching the propeller, causing a reduction of the wake fraction, which further reduces the ship efficiency. Finally, the development of biofouling on the propeller blades modifies their original profile and roughness, causing an increase in the blade resistance and in the propeller torque, thus changing the propeller characteristics in free flow and resulting in an additional efficiency loss. According to [9], the generation of biofouling on ship hulls is influenced by several factors, such as the operational conditions (speed and time moored or at anchor), the trading routes and the sea-water characteristics, as well as the maintenance actions concerning the anti-fouling coating as well as the vessel hull cleaning frequency.

Monitoring the effect of biofouling development on ship performance is achieved by calculating appropriate KPIs. The KPIs act as a measure of the ship actual performance using operational data against the ideal reference performance typically obtained from sea trials. This is the standard approach of the ISO 19030 [10], which presents a default method addressing a combination of biofouling and mechanical types of damage that affect the hull and propeller performance (variable pitch propellers introduce additional complications that the current version of this standard is unable to accommodate). However, ISO 19030 [10], when compared with more advanced models based on data-driven methods, is less accurate and provides results with higher uncertainty (see for example [11]).

Therefore, the current study will exploit data-driven models based on Machine Learning algorithms for the development of a condition monitoring framework of the hull and propeller fouling, which will support the decision-making process to optimize maintenance actions. Specifically, the structure of the report is as follows. Section 2 presents the basic aspects of the biofouling problem, focusing on the different biofouling types and the parameters that affect its development. Moreover, it summarizes the maintenance methods related to the anti-fouling measures and the available technologies within the corresponding regulatory framework. In addition, the techniques for the hull and propeller cleaning are also described, to inform on the several maintenance actions that can be carried out by the ship operator. A brief literature review of the studies focusing on the hull and propeller performance monitoring using data-driven models is also presented. As the framework will be based on the analysis of operational data, Section 3 includes the presentation of the ship that will be used as case study and the reasons for its selection. Furthermore, three methods of data acquisition, e.g., high frequency data from automated logging, noon reports and information from a weather provider, will be analyzed, since the required information needs to be derived from a combination of all the available data sources. Section 4 presents an exploratory analysis of the key operational data that will help to develop the database to be used for the models in the next step. Initially, a comparison study of the key statistics of the basic parameters of the problem is presented to examine the level of agreement and the activities needed for the completion of the database. Next, the identification of the dominant loading and operational conditions is performed. This step is needed to verify the actual range of parameters that are significant to the problem, and it will also be

used in WP2 to account for realistic operational conditions in hydrodynamic optimization. Finally, a verification of the maintenance actions is performed through the use of a KPI and its capability to capture performance changes in specific times is assessed. Section 5 describes the steps for developing the data-driven models, the feature engineering, related with the creation of new parameters, the identification of the ones being more relevant to the problem, then the data preparation concerning the application of a proper filtering and smoothing method, which is crucial to ensure a high-quality database. The central focus of this section is dedicated to showing the development of the model through the application of multiple ML algorithms and the fine-tuning of hyperparameters using quantified metrics. Ultimately, two distinct models are presented, each one offering a unique approach to address the performance monitoring problem. Section 6 utilizes the models to perform the optimization of the maintenance actions. Particularly the KPI to be used is defined, while the methodology aims at calculating the additional daily fuel cost due to biofouling, considering however the time component of the biofouling growth. This piece of information is essential for an efficient decision-making process to select the time of maintenance actions. Section 7 presents an overview of the condition-based maintenance techniques and the relevant regulatory framework, which should progress in parallel to bridge the current gaps. Finally, the key conclusions of the study are reported. In addition, several appendices are included to support the study. Firstly, details about the ML algorithms used in the analysis are included, to support a more thorough knowledge of the models developed and the specification of the hyperparameters. Lastly, an additional study is included related to the prediction of the fuel consumption of the ship under examination, to support the operational optimization to be performed in the next tasks of the WP.

2 Literature review

2.1 Basics of the biofouling problem

Biofouling is the unwanted accumulation of marine life on a ship's submerged surface. Depending on the type of organisms that settle on the hull, hull appendages and propeller(s), biofouling is divided into micro-fouling and macro-fouling. The former involves the formation of a biofilm consisting of micro-organisms that adhere to the wetted surface and to each other through cell adhesion [12]. The latter refers to the attachment of macro-organisms such as weeds and shells, as shown in Figure 1. The biofilm, also termed slime, is the initial stage of fouling. As bacterial adhesion progresses the slime thickness increases, providing a nutrient-rich surface that facilitates weed and shell attachment [13]. The biofouling community is very diverse and can be classified into soft (non-calcareous) or hard (calcareous). Soft biofouling includes mostly slime, seaweed, algae, sponges and hydroids, whereas hard biofouling consists mainly of barnacles, molluscs, mussels and tubeworms.



Figure 1: Macro-organisms that can adhere to submerged surfaces in the Mediterranean Sea (source: LABORATOIRE MAPIEM, [14])

The degree and type of biofouling are heavily affected by the ship operational profile, i.e., speed, voyage duration and immobile periods. The formation of biofouling is more intense when the ship is stationary, commonly in port, as marine organisms can attach and reproduce more easily on a submerged surface at rest. Ships that operate at high speeds and remain stationary for short periods,

e.g., container ships, are typically fouled with slime and algae. In contrast, ships spending long periods immobile, such as naval vessels, experience macrofouling, which is more detrimental to their performance [15].

Moreover, biofouling growth is influenced by environmental conditions, such as water temperature, salinity, acidity, concentration of nutrients, flow velocity, and light intensity [16]. The nature of biofouling is also affected by aquatic biodiversity [13]. Organisms responsible for biofouling thrive in tropical waters. Thus, marine growth also depends on the season and geographic location. For completeness, the distributions of temperature and salinity in Earth's oceans and major seas are shown in Figure 2. It is also worth noting that, since separate areas of the ship hull provide varying environmental conditions, biofouling formation is not uniform across the hull surface. Different areas of the hull experience varying water flow, temperature, and exposure to sunlight, creating microhabitats that influence the settlement and growth of marine organisms. The areas of the hull that are sheltered from water flow, such as the flat bottom or areas near the waterline, are more prone to biofouling. These areas provide calmer conditions that are conducive to the settlement and growth of marine organisms. The lack of water flow allows organisms to attach more easily and form biofouling communities.

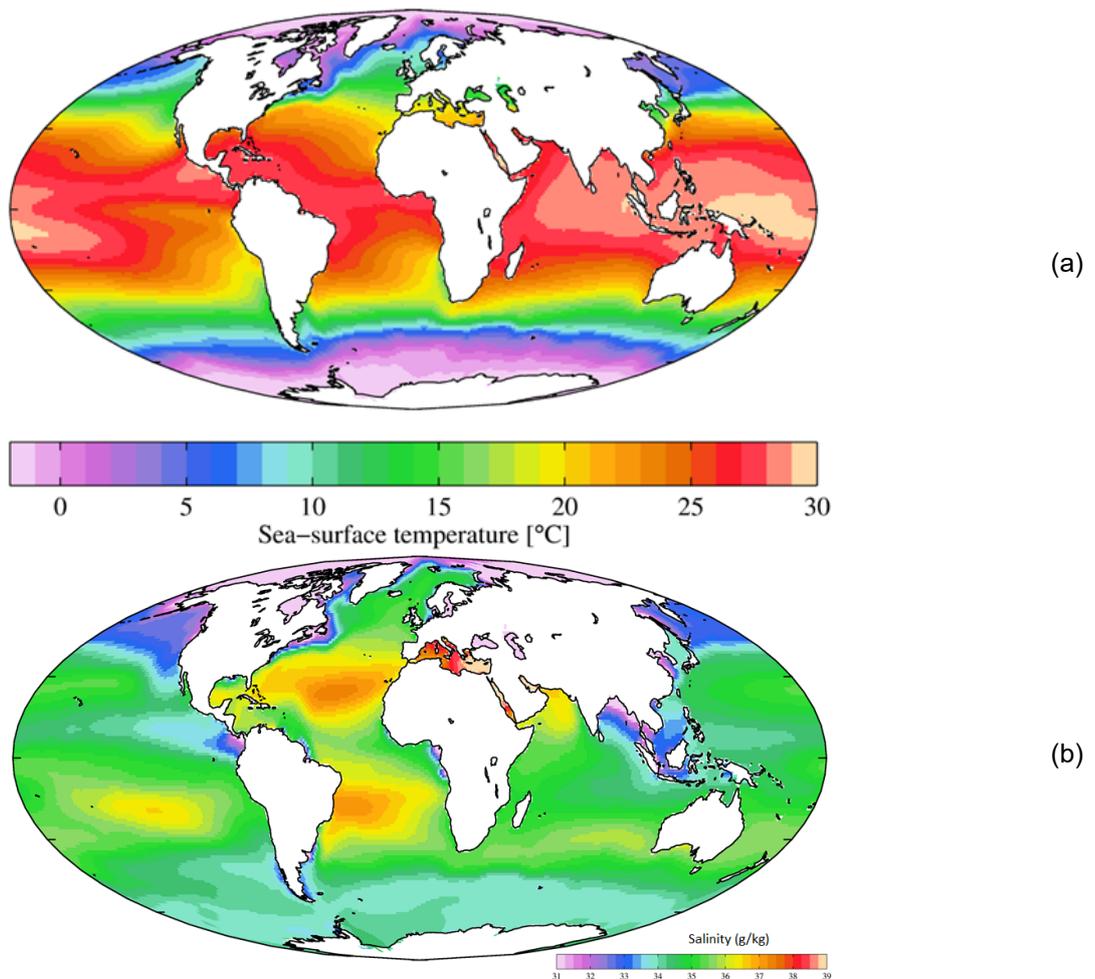


Figure 2: (a) Global distribution of average annual sea-surface temperatures, (b) Distribution of salinity in Earth's oceans and major seas, [17].

Biofouling poses a serious financial and environmental problem to the shipping industry, since it greatly diminishes the vessel hydrodynamic performance. Biofouling on a vessel leads to increased surface roughness, which in turn increases viscous drag and decreases the inflow speed at the propeller [8]. Such circumstances make the vessel less efficient when travelling through water, thus resulting in increased fuel consumption and increased emissions of greenhouse gases. The negative effect of biofouling on ship performance can be evaluated in terms of shaft power penalty, speed loss or increase in fuel oil consumption.

The impact on performance varies with the type and stage of biofouling. According to [18], slime alone on a vessel can increase power requirement up to 26%, compared to the clean state, and heavy (calcareous) biofouling can lead to an 86% power penalty at 15 knots. Reviewed studies [19] on the drag resistance associated with several biofouling states and concluded that a lightly fouled vessel requires 10% more fuel on average, whereas a heavily fouled vessel requires 35% more fuel. Taking into consideration that the fuel cost of a large ship may account for more than 75% of its operating costs [20], the importance of mitigating the negative effects of biofouling is crucial. As a counter-measure, ships are dry-docked periodically, and anti-fouling paints are applied on the hull to impede the adherence of marine organisms. Additionally, hull cleaning and propeller polishing is performed either while the ship is berthed (underwater hull cleaning) or in dry dock, using various devices (e.g., rotary brushes, water jet technology, ultrasonic technology, laser technology) to remove biofilm and calcareous formations [21].

2.2 Review of anti-fouling measures and technologies

The maritime industry is at a turning point, needing to balance effective antifouling methods with environmental safety. The need for innovative, sustainable, and biocide-free antifouling paints is not just a trend, but also a necessity. The evolution of marine anti-fouling paints was explored, highlighting their crucial role in the ship maintenance and the environmental challenges they pose [22]. While Anticorrosive Protection Painting (APP) is often side-lined in scientific discussions, it is vital to prevent marine organism growth on ship hulls. However, the paints must now also meet stricter environmental laws, reducing harmful emissions and avoiding biocides. Recent trends show that shipyards are employing automated systems for painting, aligning with laws that protect the environment by lowering harmful chemicals in marine paints.

Due to the environmental damage caused by the tributyltin (TBT) present in paints, several countries, IMO and the European Union have banned their use after 2003 and their presence on ships after 2008 [3]. IMO AFS convention bans harmful organotin in anti-fouling paints for ships and establishes safeguards against other potentially harmful substances. Historically, various chemicals, including the harmful tributyltin (TBT), were used to prevent marine life from attaching to ship hulls, impacting marine ecosystems and potentially entering the food chain. Recent studies revealed environmental damages, leading to the AFS Convention regulations on ship coatings, applicable to all member-flagged ships and those entering members' ports.

Paint manufacturers substituted TBT with other chemicals like copper, zinc and added artificial biocides. However, these substances also proved environmentally damaging, or their effects are still unclear. As a result, there is a growing emphasis on developing paints without biocides that prevent marine organisms from sticking to ships, focusing on how well organisms can adhere to painted surfaces or if they can be easily washed off. Despite this, many organisms, like algae and oysters, strongly attach to surfaces and are hard to remove, even when the ships operate at higher speeds.

The prohibition of tributyltin TBT-based paints, due to their environmental impacts, has set off the substantial transformation within the antifouling paint industry, leading to the emergence of several tin-free alternatives on the commercial market, which provide effective marine biofouling prevention with minimal environmental harm. Historically, TBT self-polishing copolymer paints (TBT-SPC) have been dominant in combating biofouling, used extensively across the global fleet and resulting in substantial economic gains. However, the adverse environmental effects of these systems cannot be overlooked. This acknowledgment has led to the establishment of stringent regulations worldwide, limiting the use and release of TBT compounds and ultimately prompting the industry to innovate TBT-free alternatives (Table 1) that maintain economic viability while reducing environmental harm [23]. In the same study an in-depth examination of the evolution, current status, and future prospects of the Anti-Fouling (AF) technology have also been presented, with a particular focus on the development of efficient and environmentally sustainable anti-fouling coatings. There is a notable research gap concerning the impact of varying seawater parameters on the performance of active chemical AF paints. Recent findings indicate that sea conditions significantly influence the chemical reactions and diffusion processes critical to the functionality of biocide-based AF paints. Back in the mid-1800s, various paints were formulated using toxicants, such as copper oxide, arsenic, and mercury oxide, dispersed in polymers and solvents like turpentine oil, naphtha, and benzene. After the World War II, the industry underwent significant changes with the introduction of new synthetic resins, the phasing out of harmful substances and technological advancements like airless spraying. The organotin compounds, introduced during this era, notably enhanced the AF paint performance.

The mechanics behind insoluble and soluble matrix paints and their evolution, have highlighted the role of components like rosin and triorganotin derivatives. The specifics of TBT self-polishing paints and the urgent turn to TBT-free systems considering environmental concerns have also been explored.

An extensive review and critical analysis of various ship hull cleaning technologies, emphasizing their evolution and significance in maintaining ship performance and preventing environmental pollution were presented in [21]. Marine biofouling potentially enables the spread of invasive species, particularly when vessels remain inactive for extended periods, such as while mooring.

Traditional cleaning techniques have heavily relied on manual labour, with workers manually scrubbing or wiping to remove biofouling, a method still widespread in hull maintenance. However, technological progress has introduced more efficient tools and methods that significantly reduce manual labour and enhance the cleaning operations, such as powered rotary brush systems and innovative noncontact cleaning technologies, see Figure 3.

Common strategies for biofouling management include dry-docking cleaning, application of antifouling paints and regular underwater cleaning. Noncontact techniques, such as Cavi-Jet pistols used by divers for intricate surfaces, are preferred for minimizing damage to hull coatings compared to rotary brushes. Heating methods have also proven effective in exterminating a range of marine organisms, especially in scenarios of light and moderate biofouling.

Furthermore, cleaning devices ranging from handheld cleaners to advanced robotic systems can also be utilized. Underwater hull cleaning robots are spotlighted as optimal solutions, due to the issues associated with diver-led cleaning, such as high physical demand, low efficiency, restricted operation time and safety concerns. These robots can navigate close to the hull without causing damage, achieving precise movement control with six degrees of freedom (DOF) and high positional accuracy. However, challenges persist in controlling the attachment and release mechanisms,

particularly when using permanent magnets for the robot's adherence to the hull, with electromagnets presented as a viable alternative.

It is recalled that BugWright2 “Autonomous Robotic Inspection and Maintenance on Ship Hulls and Storage Tanks” - a collaborative R&D project co-funded by the European Union's Horizon 2020 Research and Innovation programme, participated by RINA Services - is currently addressing state-of-the-art Remote Inspection Techniques (RIT), namely Unmanned Autonomous Vehicles (UAVs), Remotely Operated Vehicles (ROVs), and magnetic crawlers as part of a cascade of technological innovation for underwater hull monitoring and cleaning. For more details, the reader can refer to the BugWright2 webpage (<https://www.bugwright2.eu/>). Further considerations on these issues and on the regulatory aspects of hull condition-based monitoring and the relevant regulatory framework are reported in the Section 7 of this deliverable.

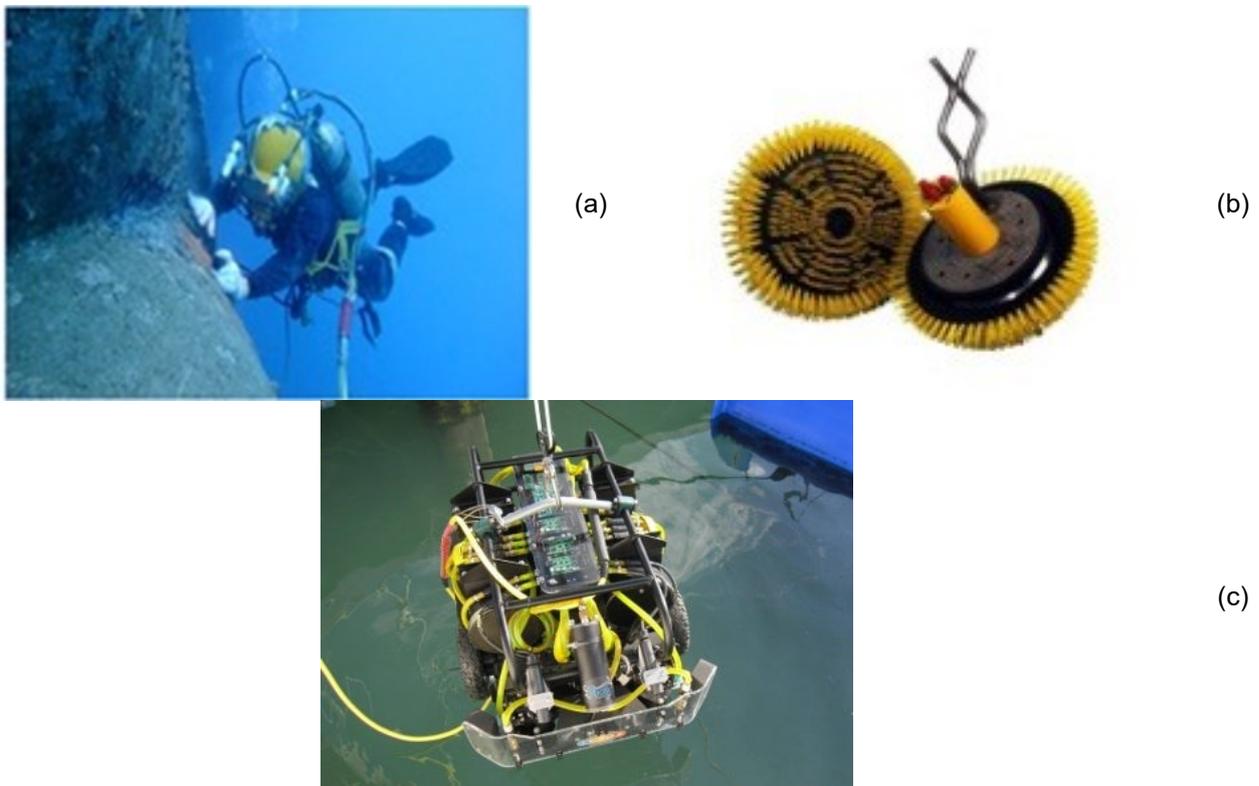


Figure 3: Hull cleaning methods: (a) Manual cleaning¹ (b) powered rotary brush systems², (c) innovative noncontact cleaning technologies³.

The balance between effective hull in-water cleaning practices and the preservation of fouling-control coatings, emphasizing the need to avoid coating damage for extended lifespan and minimized antifoulant release were investigated in [24]. The research, conducted over a year, employed a novel immersed waterjet, to identify the minimal cleaning forces sufficient for early-stage fouling removal without inflicting harm on biocidal antifouling or biocide-free Foul-Release (FR) coatings. Findings indicate that bi-monthly or monthly cleanings, utilizing a maximum wall shear stress of ~ 1.3 kPa and a jet stagnation pressure of ~ 0.17 MPa, effectively maintain the integrity of both AF and FR coatings.

¹ <https://www.nereussubsea.com/blog/underwater-hull-cleaning/>

² <https://rts.as/wp-content/uploads/2020/06/RQ-021-Rotary-Cleaning-Brush.pdf>

³ <https://www.naval-technology.com/news/newsus-navy-test-searobotics-hullbug-autonomous-cleaning-system/>



Notably, the AF coating requires bi-monthly cleanings to restrict fouling, while FR coatings achieve similar results even without cleaning. The study also addresses the environmental implications of prevalent fouling prevention methods, including chemical pollution from antifoulant release, potentially intensified by in-water cleaning, and the danger of non-indigenous species dispersion due to compromised hull coatings or uncontained waste during cleanings. The research concludes that while no significant wear is observed on the coatings during minimal force cleanings, further precision is necessary to estimate the biocide release for future environmental risk assessments. Additionally, while regular cleaning reduces fouling levels, particularly for biocidal coatings, active vessels may require less frequent cleanings, except from certain hull areas less exposed to hydrodynamic forces.

In [25] the historical, environmental, and regulatory aspects of antifouling strategies are addressed, emphasizing on the ecological impact of biocidal AF technologies and the urgent need for environmentally alternatives. Biofouling, a significant concern for marine vessels, not only increases fuel consumption due to increased drag, but also poses an ecosystem risk through the transport of Non-Indigenous Species (NIS). Traditional biocidal solutions, particularly copper and tributyltin (TBT), have introduced extensive environmental contamination, prompting a global TBT ban in 2008 and stricter copper regulations, leading to the investigation of non-toxic coatings (Table 1).

Synthesis of the ecological impacts of past and present biocides underscores critical research gaps and offers guidance for future antifouling management, with focus on the transition from toxic substances like copper and TBT to potentially less harmful alternatives. These include foul-release coatings, leveraging silicone elastomers, waxes, or oils, and natural coatings derived from marine organisms.

While effective alternative paints exist, their suitability varies across different vessels, and the ecological advantages of eliminating toxic biocides must be balanced against potential increases in fuel consumption and NIS dissemination. Regulatory decisions around AF paints require a holistic consideration of both economic and ecological trade-offs. Although the TBT ban was economically motivated, reflecting its detrimental effects on the oyster industry, the ban also accounted for its ecological consequences, bioaccumulation risks, and human health threats. However, the shift to copper and booster biocides may prolong environmental toxicity and promote NIS resilience.

The ideal elimination of toxic biocides from AF paints confronts the challenge of identifying viable replacements, particularly for copper. Current alternatives like foul-release coatings and natural compounds show potential but lack widespread applicability (Table 1).

In [26] a comprehensive overview of the relationship between hull fouling and its impact on ship performance is presented. Selecting the appropriate coating depends on factors like salinity, vessel service speed, national regulations, compatibility and cost. The complex process of biological fouling depends on numerous factors, including the ship loading condition, operational zones, anti-fouling paint effectiveness and environmental conditions. Ships in constant motion are less prone to marine growth accumulation than those in prolonged mooring or anchoring. Hull cathodic protection is more effective when the ship is in motion.

During the ship operation, surface roughness may increase due to coating cracking, damage, and corrosion, all of which can attract marine growth. Various types of coatings are used to combat fouling, each with different characteristics. Controlled Depletion Polymer (CDP) coatings release biocides through rosin matrix dissolution, Self-Polishing Copolymer (SPC) coatings release biocides through hydrolysis and Foul-Release Coatings utilize non-stick properties to prevent fouling



adhesion (Table 1). Most anti-fouling coatings currently in use are self-polishing copper and tin-based paints, although some countries consider prohibiting these substances.

Table 1: Anti-fouling coatings, based on [25] and expanded.

Anti-fouling system	Action – Main components
TBT Self-polishing copolymer (SPC) coatings	Zinc oxide and insoluble pigments or copper oxide, tri-organo-tin and co-biocides. TBT biocide is chemically bonded within a copolymer resin, and its gradual release occurs through hydrolysis when in contact with seawater, resulting in a slow and consistent discharge of the biocide.
Tin-free SPC coatings	A Cu/Zn/Silyl copolymer resin, containing Cu particles and booster biocides uniformly dispersed within the paint matrix, undergoes hydrolysis upon exposure to seawater, ensuring a steady and controlled release of the biocide.
Tin-free conventional coatings	Cu particles and booster biocides are dispersed within a soluble or insoluble paint binder, and their release into seawater occurs through dissolution, resulting in a gradual and diminishing discharge of the biocide.
Booster biocides	Tin-free Conventional and SPC AF paints frequently incorporate herbicides/pesticides to enhance their effectiveness against Cu-tolerant algae, often achieving improved results.
Foul-release coatings	Surfaces with low energy and minimal adhesiveness typically feature silicone elastomers and frequently incorporate silicone oils.
Biomimetics	Natural anti-fouling (AF) compounds produced by marine organisms, such as secondary metabolites, or surfaces designed based on natural microtopography.
Controlled depletion paints (CDPs) (increase of soluble matrix technology with use of new resins)	CDP coatings utilize hydration as part of their mechanism.
Hybrid systems (CDPs/SPCs)	Potentially involving biocides like Zineb or Cu pyrithione for enhanced anti-fouling properties.
Organic coatings	Superhydrophobic polyurea/TiO ₂ composite coating with rapid self-healing ability by simple sprinkling or spraying of modified TiO ₂ nanoparticles on the brushed polyurea coating [27].
Graphene-based anti-fouling coatings	Graphene oxide and reduced graphene oxide. Zinc-rich graphene coatings are effective in case of defects in the coating [28].

2.3 Review of studies related with data-driven models for the performance monitoring

Several data-driven models for the prediction of fuel consumption were examined in [29]. Operational data from a containership were utilised to develop the models, specifically Artificial Neural Networks (ANN) and Multiple Linear Regression (MLR) models. For the development of these models, unlike other studies that use fuel consumption per hour as the dependent variable, fuel consumption per distance travelled was used, to account more emphatically on the operational conditions, such as the loading condition and the weather. Then, for the selection of the dependent variables to be used for the model development, the "Domain Knowledge" and "Statistical method based on Lasso Regularization" methods were examined. The evaluation of these methods was based on the calculation of the mean absolute error, considering several combinations of variable feature selections and model training methods. It was derived that the more accurate results obtained by the usage of Artificial Neural Networks in combination with either the "Domain Knowledge" method or the "Lasso Regularization" method. Finally, a sensitivity analysis of the model was performed regarding the ship draft. For this analysis, the combination of ANN - Domain Knowledge was used, aiming at the calculation of the optimal draft for the reduction of fuel consumption.

A comparison study of various machine learning and statistical methods to predict the propulsion power needed to achieve a specific speed was presented in [30]. The methods include XGBoost, ANN, Support Vector Regression as well as statistical methods such as Linear and Polynomial Regression and Generalized Additive Model. The parameters used as input were ship speed, mean draft, trim, heading, significant wave height, mean wave period, mean wave heading and wind speed. The root means square error and the complexity degree are used to compare the examined methods. After collecting and processing data from a tanker and a RoRo vessel, it was found that machine learning algorithms offer greater reliability and accuracy compared to statistical methods. However, the statistical methods are superior in the time required for "training".

The fuel consumption of a container using various machine learning algorithms was predicted in [31]. Specifically, different predictive models such as Multiple Linear Regression, Ridge and LASSO Regression, Support Vector Regression, Tree-Based Algorithms, and Boosting Algorithms were examined using noon reports and data from engine log-books. For the validation of the predictive models, the K-fold cross-validation method was used and a correlation analysis to identify the relationship among the different variables was carried out. Additionally, for the evaluation of the accuracy of each method, a root mean square error analysis was performed, reporting that it is more advantageous for larger datasets. After gathering all the information, it was concluded that the best-performing and most accurate predictive model is the Gradient Boosting Regression.

A publicly available dataset collected from onboard sensors of a ferry over a period of two months, to train two types of neural network models: instantaneous and predictive was used in [32]. The instantaneous is a feed-forward neural network that estimates an output Y_n (e.g., the main engine's fuel consumption) based on the current input vector X_n and ignores the variation in time of the propulsion parameters. To estimate the difference in the target variable at the following time step X_{n+1} , the predictive model uses a Tap-Delayed Neural Networks (TDNN), a type of recurrent neural network. The residual error of the predictive mode in the training set is also used to fit a probability distribution that will be added to the model's final predictions as noise. With the instantaneous model a mean relative error of 1.50% was obtained on the estimation of Fuel Oil Consumption (FOC), whereas the TDNN model had no comparable metric estimated in the study.

In [33], an artificial neural network (ANN) is combined with a seventh-degree polynomial regression model to predict the brake power and fuel consumption of the main engine. It is emphasized that the available dataset contains information from two sea voyages of a Very Large Crude Carrier (VLCC) under the same loading condition, which limits the generalization ability to a range of conditions. One voyage is used to train the ANN for predicting the brake power, which is then fed into the seventh-degree polynomial regression model to predict the specific fuel consumption. The other voyage is used to evaluate the performance of the two models, resulting in the ANN achieving a coefficient of determination $R^2=0.984$ and the polynomial regression achieving $R^2=0.981$. Furthermore, the two models are utilized to estimate fuel savings in a Just-In-Time arrival scenario for the case where the ship was waiting one day for berthing. Therefore, assuming that the ship follows the same route and encounters the same weather conditions, selecting a lower navigation speed to arrive at the port without the need for waiting results in fuel savings of 24.24%.

In [34], a weather-based route planning algorithm is presented, in combination with an artificial neural network (ANN) that predicts the fuel consumption of the main engine. In this framework, data from automated logging and a meteorological service provider are utilized. During the data pre-processing stage, measurements of each parameter that deviate more than three standard deviations from the mean are discarded. It is also reported that the generated ANN achieves a coefficient of determination $R^2=0.894$ on the validation set. To determine the optimal weather-based route, the modified Dijkstra algorithm is iteratively applied, considering the dynamic variation of weather conditions. This methodology is applied in the case of a voyage from the Gulf of Guinea to Marseille, and two scenarios are examined corresponding to different navigation speeds and loading conditions. The criterion for selecting the optimal route is the minimization of the fuel consumption.

In [35] the quantification of the impact of ship design measures on its energy efficiency is attempted using data-driven models predicting the Main Engine (ME) FOC. To this end, data from automated recording and meteorological service providers are utilized, which have been collected before and after the retrofit measures, which concern bulbous bow modification and propeller exchange. Specifically, data cleaning is performed by rejecting points based on threshold values and by discarding outliers through a suitable statistical procedure involving median filtering. To monitor fouling, the Days feature is included, which measures the time elapsed since the beginning of the data set without any cleaning or retrofitting events. Then, 10-fold cross validation is utilized to explore the hyperparameter search space of feed-forward ANNs. At this stage, the contribution of various input variables is confirmed by comparing the results of models with different sets of features. In the test set, the final model achieves an R^2 score of 0.980 and an MAE of 1.245 tn/d. Next, the estimation of fuel savings as a result of the retrofitting is performed using ANN models in three different ways: a) the ANN model created with pre-retrofitting data is applied to post-retrofitting data; b) two ANN models - one trained on pre-retrofitting data and the other on post-retrofitting data - are applied to the same set of synthetic data; c) these two models are applied to scenarios with different sailing speeds. These approaches yield an average reduction of approximately 15% in the ME FOC. Furthermore, the analysis highlights the varying enhancements in hydrodynamic efficiency across different navigation speeds.

In [36] ML techniques are employed to predict power and quantify the performance degradation due to fouling, using high-frequency data from a containership spanning 12 months. In their study, the Days Since Cleaning (DSC) feature is introduced to monitor fouling. To determine the best performing model, four ML algorithms are evaluated: K-Nearest Neighbours, Decision Tree, Random Forest and ANN. The hyperparameters of each model are tuned using grid search and cross



validation. The best performance is achieved by Random Forest, with a MAPE of 1.17% on the test set. The gain in prediction accuracy is examined when DSC and significant wave height features are included. Then, predictions are made with the model on a synthetic data set, where STW varies from 14 to 22 kn, DSC ranges from 0 to 360 days, and the other parameters are held constant. Two power-speed curves are presented, corresponding to the DSC values of 60 and 360, obtained by fitting a third order curve to the prediction points at a given DSC. Thus, an average power increase of 5.2% is reported between the DSC values, which is attributed to fouling. The additional power requirement leads to an excess fuel cost of £2,500 per day. This method can be utilized by shipping companies to determine the optimal timing for conducting hull cleaning.



3 Collection of ship data: case study MV Kastor (Laskaridis Shipping)

3.1 Criteria for selection of ship

The selection of the examined ship was based on criteria related not only with this specific task. Specifically, the ship shall be equipped with sensors and systems providing the necessary high-frequency signals for the analysis of the hydrodynamic performance of hull and propeller. In addition, the collection of these signals shall be supported by an automated data acquisition system, combined with the daily reports (noon reports), which are filled in manually by the ship crew, and supplemented by weather data from a reliable provider. The previous data shall be available for a time duration of more than 1.5 years, to adequately identify any performance degradation.

Moreover, bearing in mind the wider needs of the project, from the available fleet of bulk-carrier vessels of LASKARIDIS Shipping, for the selection of the ship type a size criterion was set related with examination of retrofit measures and specifically that of a wind-assisted propulsion system. Specifically, the maximum size was set to that of Kamsarmax ships, to maximise the effects of the Wind Assisted Ship Propulsion (WASP) on the ship energy efficiency. Furthermore, it would be beneficial for the wider scope of the project if a main engine power limitation is applied to the ship to satisfy the EEXI requirements. Therefore, the implementation of suitable retrofit measures could be examined as a solution to avoid the ME power limitation to satisfy the EEXI requirements. Based on the previous criteria, the characteristics of the selected ship are presented in the next section.

3.2 Ship and equipment data

Table 2: Main characteristics of the ship under examination: MV Kastor (Laskaridis)

Length overall	229.00 [m]
Length between perpendiculars	225.50 [m]
Breadth, moulded	32.26 [m]
Depth, moulded	20.05 [m]
Summer load line draught, moulded	14.45 [m]
Deadweight at summer load draught	80996.1 [t]

Table 3: Main characteristics of the Main Engine (ME) of the ship under examination

Manufacturer	HYUNDAI-MAN B&W
Type	6S60ME-C8.5
Maximum continuous rating (MCRME)	9930 kW x 90.4 rpm
Limited maximum continuous rating with engine power limitation (MCRME,lim)	8230 kW
SFC at 75% of MCRME or 83% of MCRME lim	166.81 g/kWh
Number of engines	1
Fuel type	HFO

Table 4: Main characteristics of the Auxiliary Engine (AE) of the of the ship under examination: MV Kastor (Laskaridis)

Manufacturer	YANMAR CO., LTD
Type	6EY22LW
Maximum continuous rating (MCRAE)	800 kW x 720 rpm
SFC at 50% of MCRAE	215 g/kWh
Number of sets	3
Fuel type	HFO

3.3 High-frequency data

A list of all parameters collected by the available high-frequency data acquisition system has been provided by LASKARIDIS. The available dataset corresponds to two and a half years (from February 2021 to July 2023), with a one-minute sampling period, so it consists of 1,311,000 data points. The list contains 559 parameters, some of which correspond to direct measurements from onboard sensors, some come from weather data providers, some are for the Data Acquisition (DAQ) provider internal use only, and some are built and processed. LASKARIDIS specifies that the names of parameters coming from sensors end with the name of the source (e.g., “_AMS”), the names of weather provider parameters start with “DTN_”, the names of parameters acquired via the DAQ for internal use contain joined words (e.g., “mainEnginePower”), and the names of processed and built parameters have “-” between words (e.g., “Propeller-Shaft-Power”). A file that would contain all these parameters for a two-and-a-half-year period would be quite large and difficult to handle. Moreover, some of these parameters would not add any value to data analysis, while increasing the size of the dataset. Thus, a choice had to be made about the parameters that LASKARIDIS should provide.

In particular, the signals coming from onboard sensors are essential as well as reliable, so all of them have been included in the dataset. Weather data provider parameters are also necessary and have to be included. However, parameters built by the DAQ providers aiming to provide their own reports to the shipping company, have been omitted because their exact definition is unknown. Similarly, parameters developed to treat outliers have been left out, since each data analyst uses his/her own methods for this operation. As a result, a dataset with 268 parameters has been obtained and shared with the partners as a CSV file. For easier handling, the same dataset has also been provided split into four CSV files.

When examining the data, it appeared that some parameters were not available throughout the recording period. For instance, draft measurements are valid until November 2021, since after that date the sensors failed and produced faulty values. Available weather provider data starts from April 2022. Moreover, the Main Engine fuel oil supply volumetric flow meter malfunctioned between November 2021 and July 2022, as well as between September 2022 and January 2023, due to a mechanical failure.

3.4 Noon report data

In addition to the high-frequency data, LASKARIDIS has also provided a noon report dataset for the same two-and-a-half-year period, consisting of 930 datapoints. Noon reports are prepared on a daily basis, including fewer parameters than the automated logging system of the high-frequency data, and the values they provide are less accurate. Nevertheless, it is important to include noon report

data in the analysis because they are easier to handle due to their small size, providing a quick assessment of the vessel performance. Noon reports can also be used to fill in missing values from high-frequency data in cases where similar parameters exist. As an example, draft, loading condition and weather parameters from noon reports have been filled in per minute based on their daily values, to have the same time resolution as the high-frequency data. By merging these filled in noon report parameters with the original high-frequency data, a new dataset has been created and then shared with the partners.

Table 5: Basic information on the available datasets

	High Frequency Data	Noon Reports
Period	02-2021 to 07-2023	02-2021 to 07-2023
Sampling interval	1 minute	1 day
Number of points	1,311,000	930

3.5 Weather data

Weather data from third-party providers have an actual time resolution of one hour. To align this data with the rest of the dataset, the DAQ provider modifies the time resolution, converting it to one minute. Available parameters include sea temperature, air pressure and temperature, wind speed and direction, total current speed and direction. More importantly, parameters for wave, wind-wave and swell characteristics have been provided, together with explanations listed in Table 6.

Table 6: Definition of wave data parameters from weather providers.

Significant wave height H_s	Spectral approximation of the observed mean wave height of the largest 33% of the waves within a 20-minute window.
Max wave height	Mean of 1% highest waves or $1.67H_s$ according to deep water Rayleigh distribution.
Mean wave direction	Mean direction where waves are arriving from. Mind the combination of multiple wave trains.
Mean wavelength	On deep water (>500m) the mean wavelength can be computed as $1.56T^2$. Account for the dispersion relation on shallow water.

4 Exploratory data analysis of key parameters

4.1 Comparison between specific high-frequency data and noon reports

Automated logging is certainly more reliable due to the higher frequency and accuracy of the measurements, but noon report information is available throughout the examined period and should therefore be considered in the cases where high frequency recording is missing. Having created a dataset containing ship operation parameters from both high-frequency recording and noon reports, the first step of the analysis is to compare the available parameters from both these sources.

Mean draft is calculated as the average of the fore and aft draft measurements. High frequency draft measurements are valid from February to November 2021. Mean draft values from noon reports during this period are plotted against the respective mean draft obtained from the high frequency sensors, as shown in Figure 4. The clustering of the points around the 45-degree line confirms the agreement between the two recording methods. The horizontal lines appearing in this plot show the increased variability of the continuous monitoring sensor values compared to the noon report draft values, which appear somewhat discrete; this is attributed to the low frequency of noon reports. Figure 5 shows a histogram of the difference of the mean drafts by the two sources; a mean difference of -0.54 m is calculated, which accounts for bias, as well as a standard deviation of 1.41 m, because of the high frequency recording variance.

Table 7: Descriptive statistics of high-frequency and noon report parameters.

	Mean draft difference [m]	Wave height difference [m]
Mean	-0.54	-0.02
Standard deviation	1.41	0.84

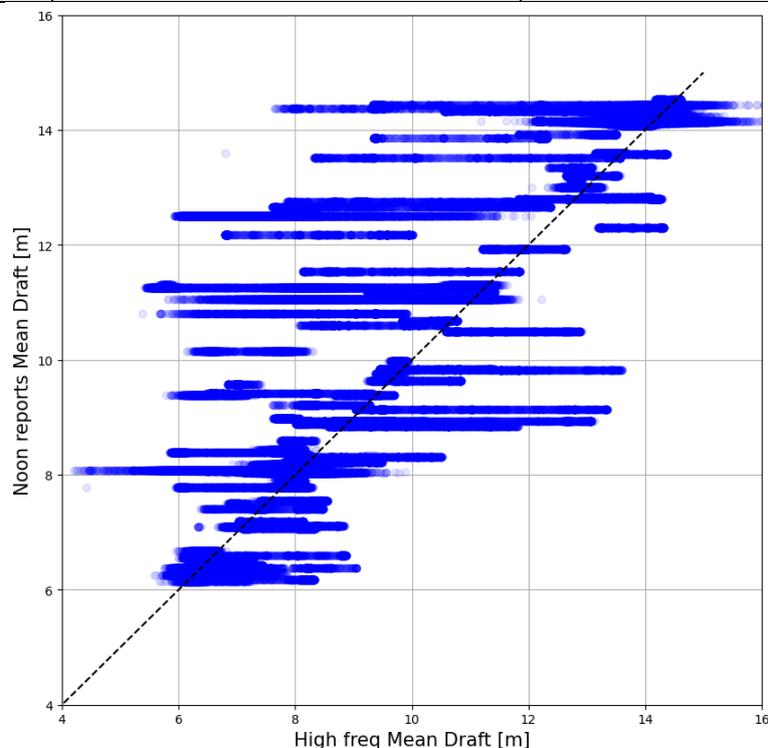


Figure 4: Noon reports Mean Draft plotted against high-frequency Mean Draft.

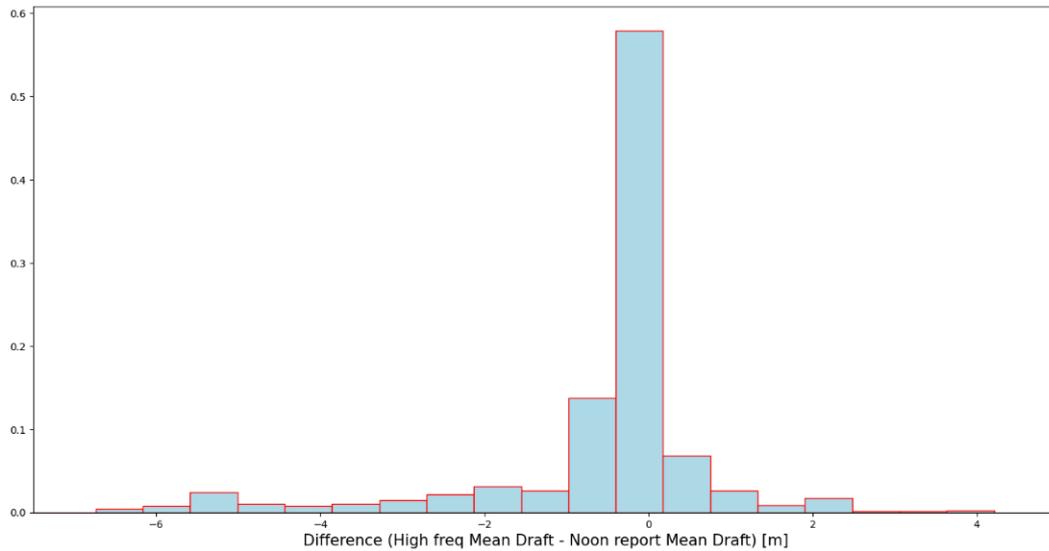


Figure 5: Relative frequency histogram of difference between the Mean Draft from high frequency data and noon reports.

Data from the weather provider are available from April 2022 to February 2023. Sea height values from noon reports during this period are plotted against the respective significant wave height obtained from the weather provider, as shown in Figure 6. Both parameters represent the combined effect of swell and wind waves. Again, the clustering of the points around the 45-degree line indicates agreement of the two reporting methods, while the horizontal lines are due to the higher resolution in time of the weather provider data, compared to the noon reports.

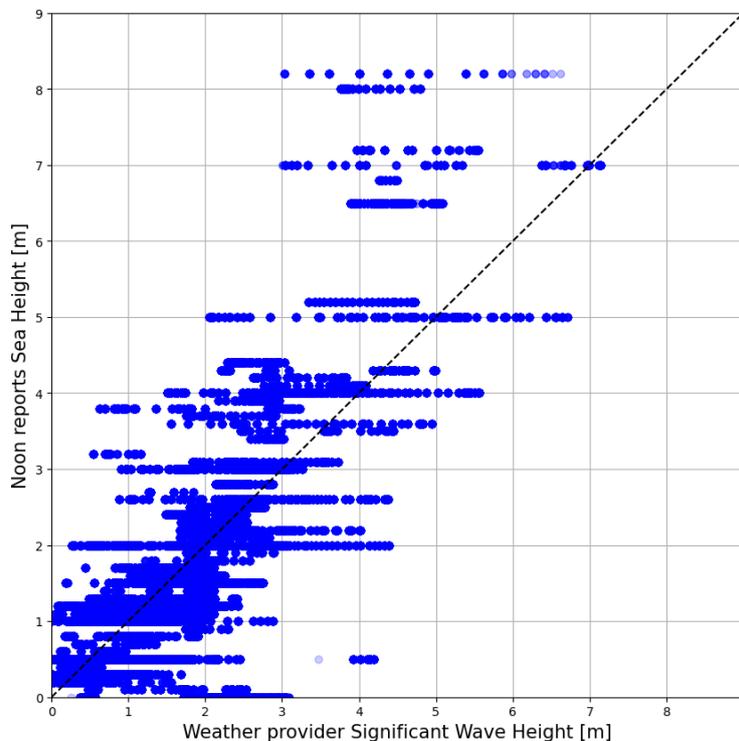


Figure 6: Sea height from noon reports plotted against significant height from the weather provider.

Figure 7 shows a histogram of the difference of the two parameters mentioned above: a mean of -0.02 m and a standard deviation of 0.84 m are calculated. Furthermore, Figure 8 shows a histogram of sea height from noon reports; a mean of 1.73 m and a standard deviation of 1.39 m are calculated. It should be noted that Figure 6, Figure 7, and Figure 8 correspond to STW greater than 6 kn, which is assumed as representative of the navigational conditions in open sea.

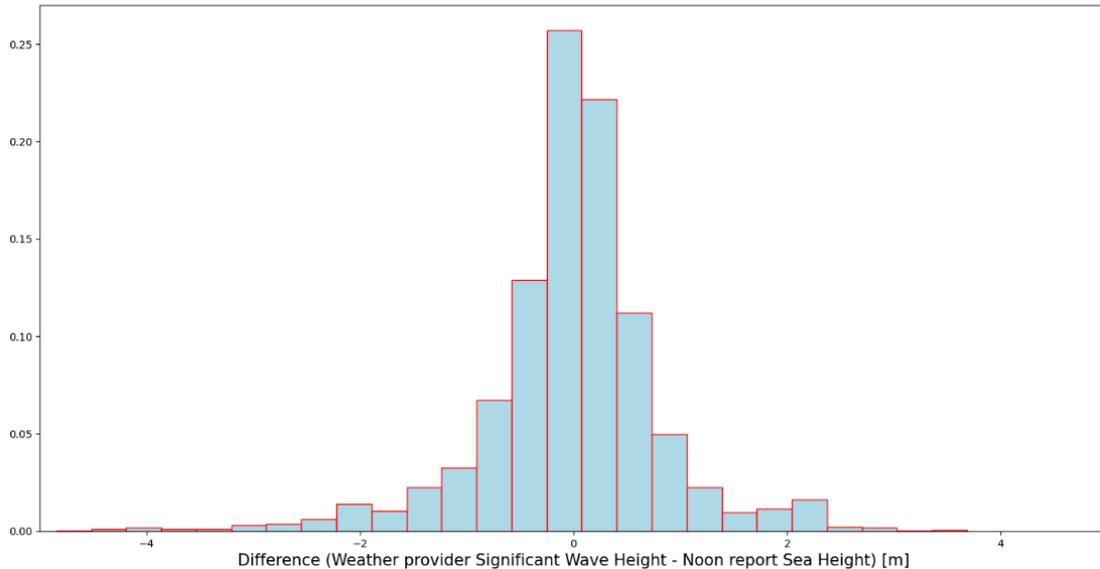


Figure 7: Relative frequency histogram of the difference between the significant wave height from the weather provider and sea height from noon reports

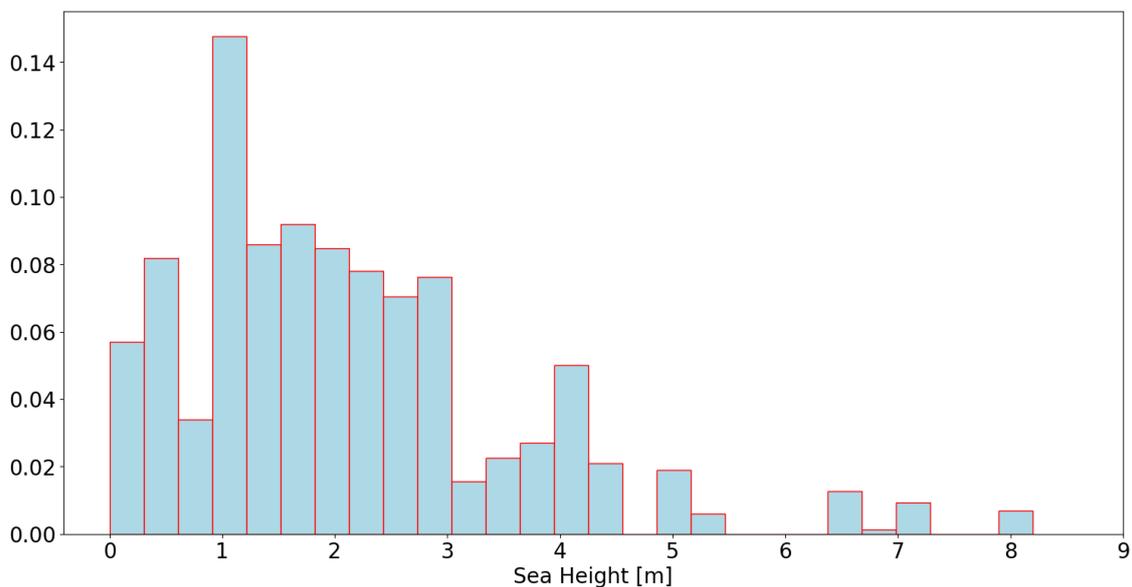


Figure 8: Relative frequency histogram of sea height from noon reports

Based on the previous observations, the noon reports draft and sea height parameters are then used, instead of their high-frequency counterparts. Another option would be to use high-frequency values where available, and noon reports where high-frequency values are not available, but this would lead to an uneven analysis.

Another useful tool for exploring the ship operation is to project its path on the world map by simply superimposing the GPS longitude and latitude signals on the map, as shown in Figure 9. This helps to identify the routes followed by the ship during the two-and-a-half-year recording period, information essential for implementing physics based or data driven weather routing. According to LASKARIDIS, a total of 60 voyages were performed within this period.

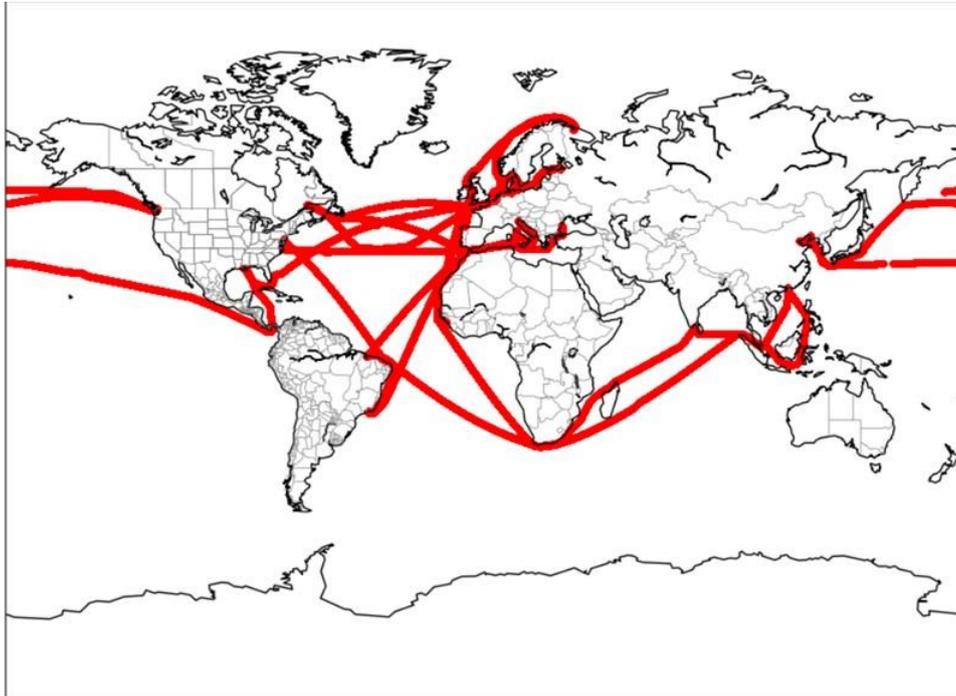


Figure 9: Projection of the ship path on the world map during the recording period, utilizing high-frequency GPS signals.

A comparison is then made between three different vessel speed parameters: the first two are the Speed Through Water (STW) and the Speed Over Ground (SOG) from high-frequency recording and the third is from noon reports. In Figure 10 the distributions of the three speed parameters are superimposed, creating an overlaid histogram. Here, it is worth observing that the noon reports values lie close to SOG, while there is a noticeable deviation from STW, which is higher overall, because of sea currents. Hence, the difference between SOG and STW is used to determine the sea current speed, which has a mean of -0.74 kn (STW>SOG) and a standard deviation of 1.10 kn, and its histogram is shown in Figure 11.

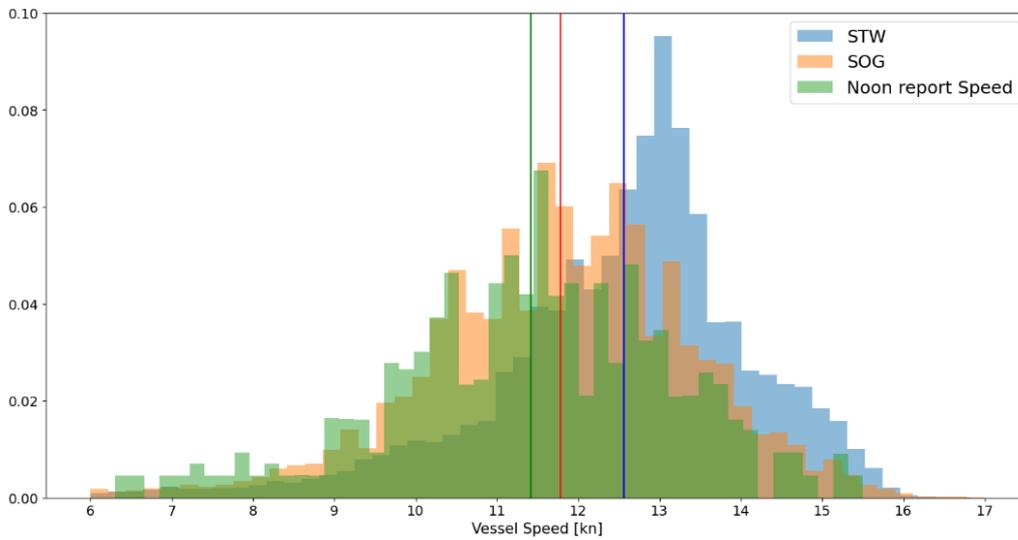


Figure 10: Overlaid histograms of the STW (Speed Through Water), SOG (Speed Over Ground) and of the vessel speed from noon reports.

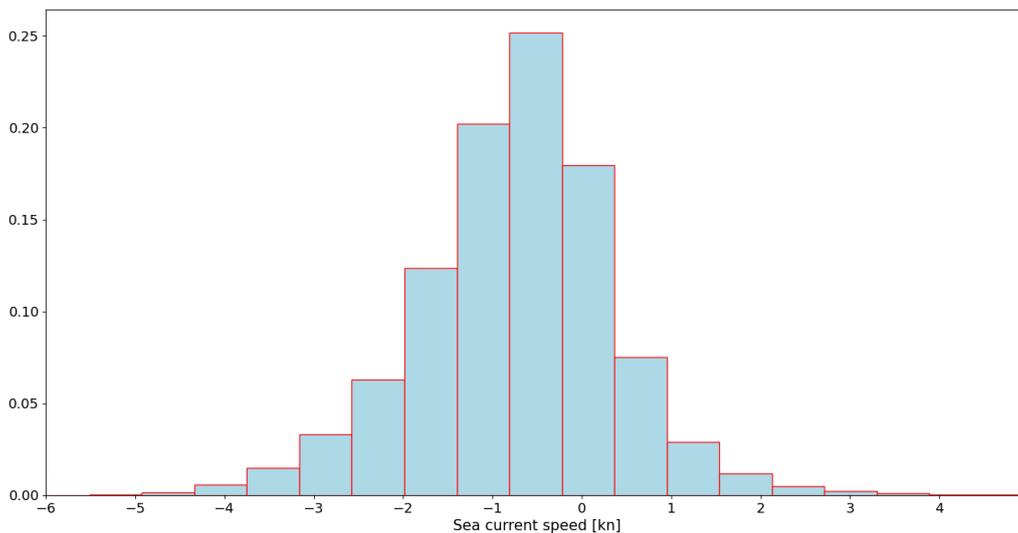


Figure 11: Relative frequency histogram of sea current speed, calculated as the difference between the STW and the SOG

4.2 Loading conditions

Upon examining the draft distribution, several loading conditions that the vessel experienced during the reporting period are revealed. To filter out port operations, the criterion $STW > 6$ kn is applied. The histogram in

Figure 12 provides a visualization of the distribution of the mean draft, which appears to be multimodal. Specifically, four peaks are prominent near the values of 6.3 m, 8.1 m, 13.2 m and 14.3 m, corresponding to four distinct loading conditions, two of which are laden and the others ballast. The draft range of each loading condition is listed in Table 8, as derived from the spread of the distribution around these peaks, also considering the trim and stability booklet of the examined ship. In addition, the trim is calculated as the difference of the fore and aft draft measurements. In the histogram of

Figure 13 it is indicated that the distribution of trim is bimodal: the first peak is close to -0.2 m, corresponding to the two laden conditions, and the second is close to -2.6 m, corresponding to the two ballast conditions, as shown in Figure 14.

For better understanding, each loading condition is reviewed separately below. Figure 15, Figure 17, Figure 19, and Figure 21 show the histograms of the STW, while Figure 16, Figure 18, Figure 20, and Figure 22 show the histograms of the trim for the above-mentioned loading conditions. More details about each loading condition are provided in Table 9, where mean values and ranges for the mean draft, trim and STW are reported. The aim of this table is to help define the respective ranges in the design of experiments for the surrogate models used in WP2.

Table 8: Definition of loading conditions using Mean Draft ranges.

Loading condition	Type	Draft range [m]
1	Laden	$TM > 13.5$
2	Laden	$12 < TM < 13.5$
3	Ballast	$6 < TM < 6.5$
4	Heavy ballast	$7.75 < TM < 8.5$

Table 9: Mean values and ranges of Mean Draft, trim and STW for each defined loading condition.

Loading condition	TM [m] Mean	TM [m] Range	Trim [m] Mean	Trim [m] Range	STW [kn] Mean	STW [kn] Range
1	14.30	[13.5, 14.5]	-0.16	[-0.5, 0]	12.40	[9, 15]
2	13.13	[12, 13.5]	-0.16	[-0.5, 0]	12.45	[11, 15]
3	6.31	[6, 7]	-2.89	[-3.5, -1.8]	12.94	[9, 16]
4	8.11	[7.75, 8.5]	-2.01	[-3, -1.5]	12.52	[9, 16]

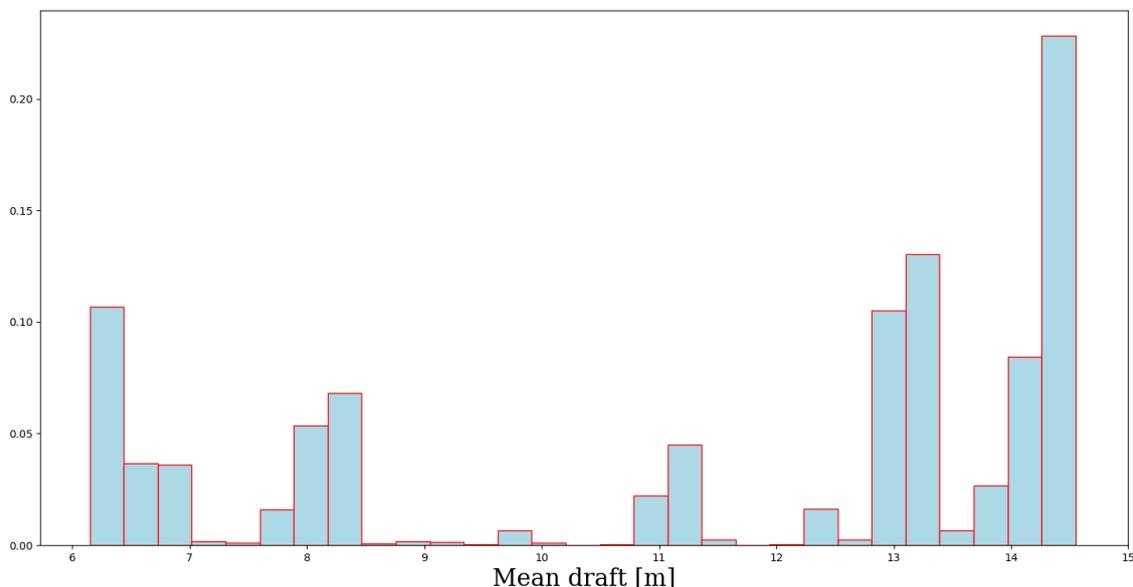


Figure 12: Relative frequency histogram of the Mean Draft from noon reports.

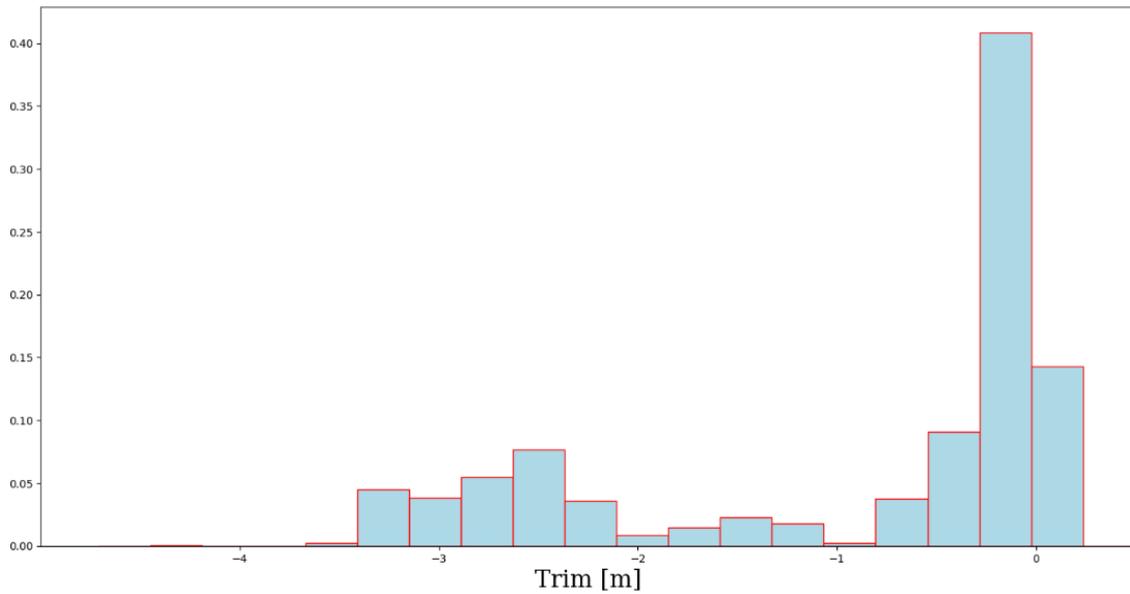


Figure 13: Relative frequency histogram of the trim from noon reports.

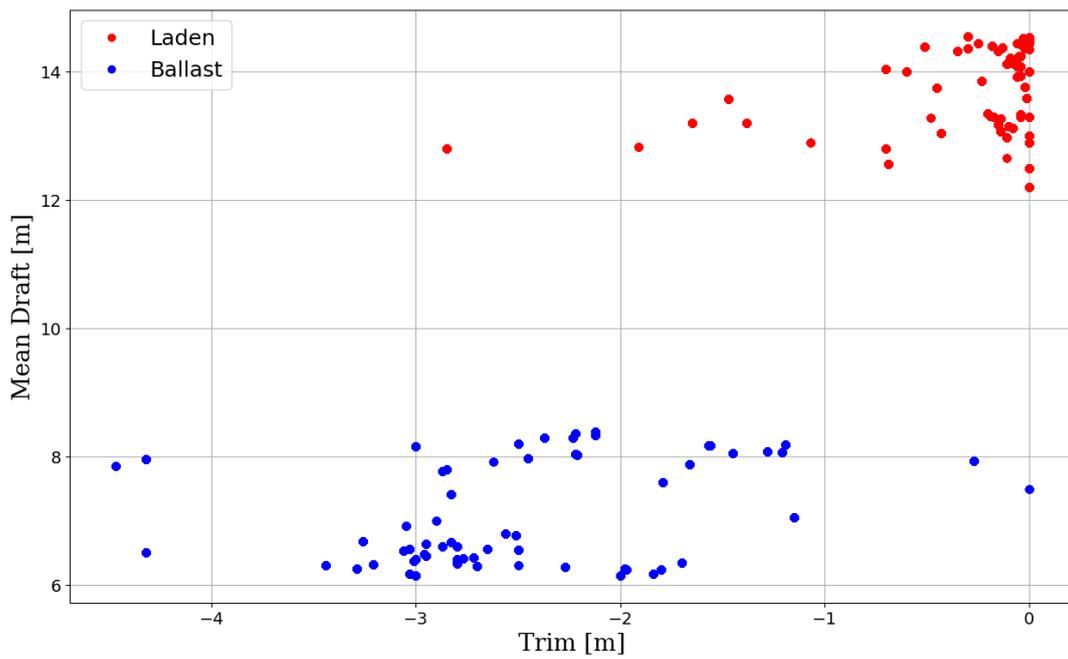


Figure 14: Mean Draft plotted against trim using separate colours for the laden (red) and ballast condition (blue).

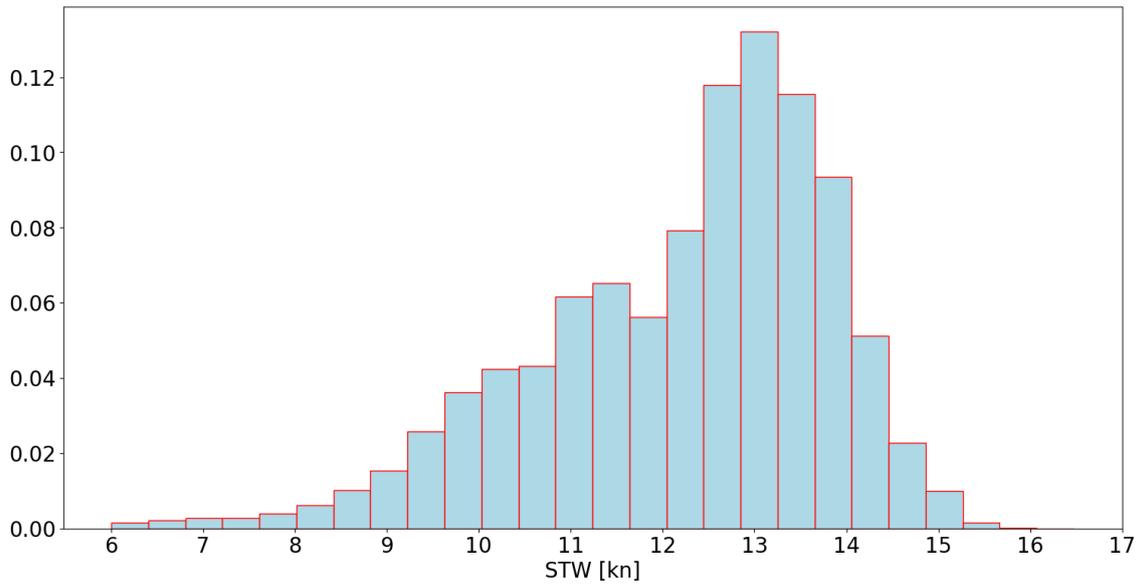


Figure 15: Relative frequency histogram of STW for loading condition 1 (laden).

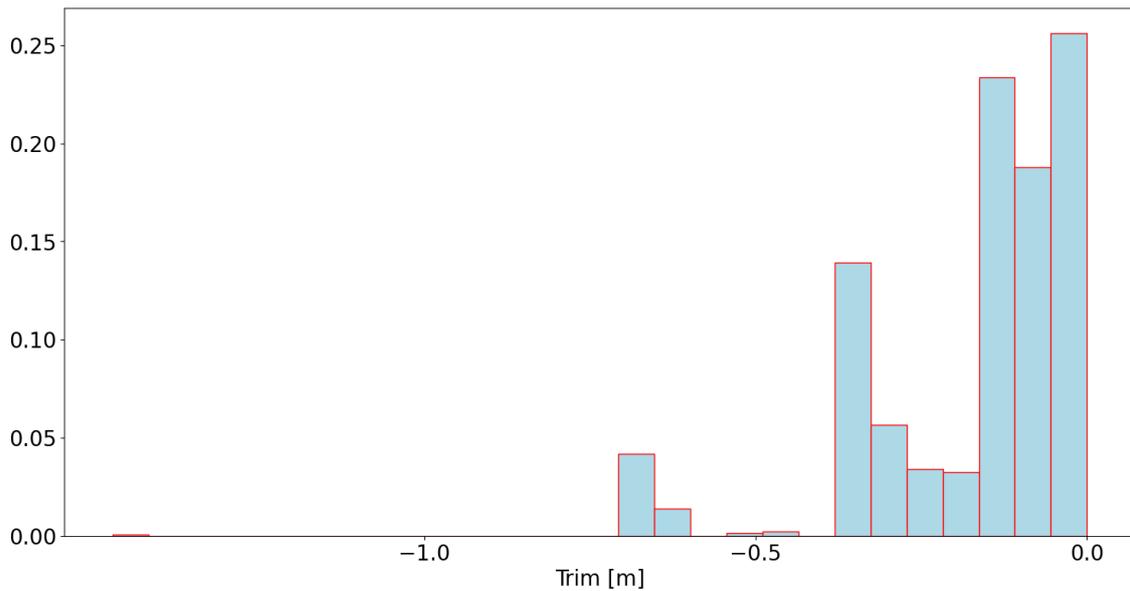


Figure 16: Relative frequency histogram of trim for loading condition 1 (laden).

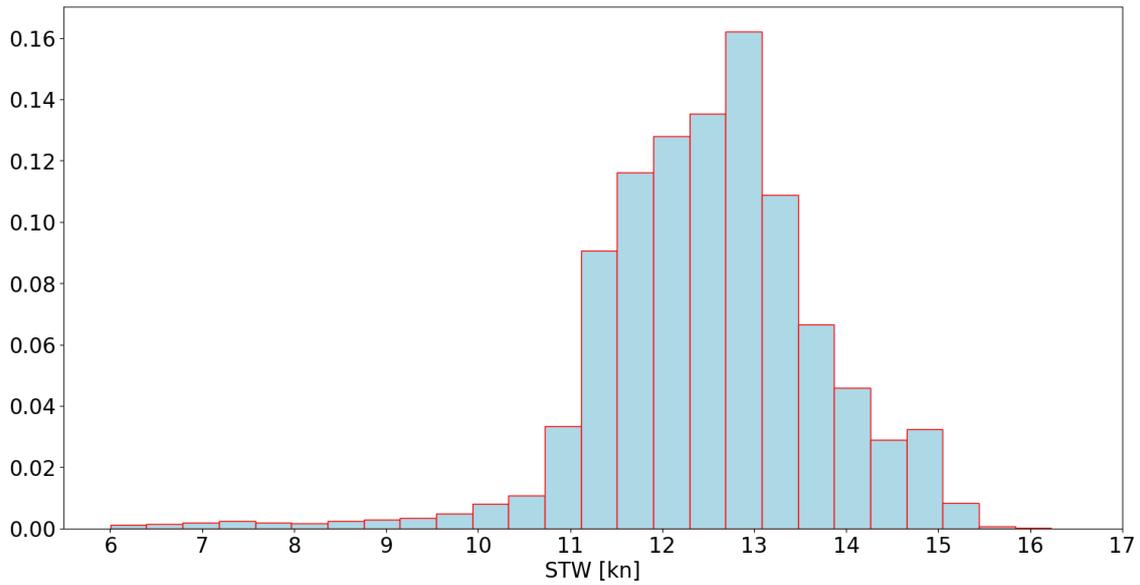


Figure 17: Relative frequency histogram of STW for loading condition 2 (laden).

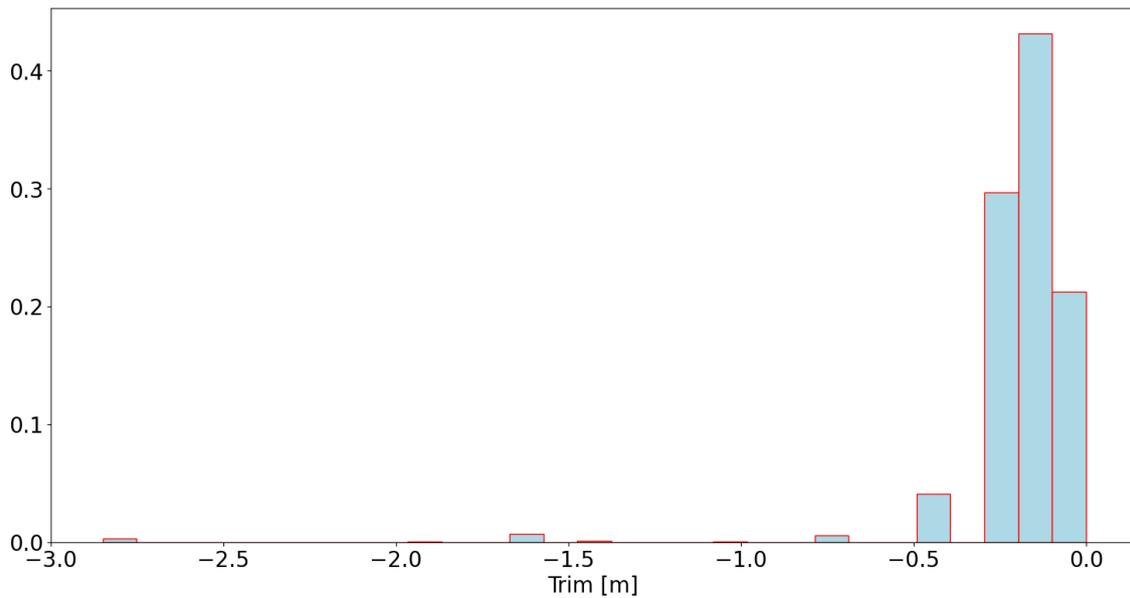


Figure 18: Relative frequency histogram of trim for loading condition 2 (laden).

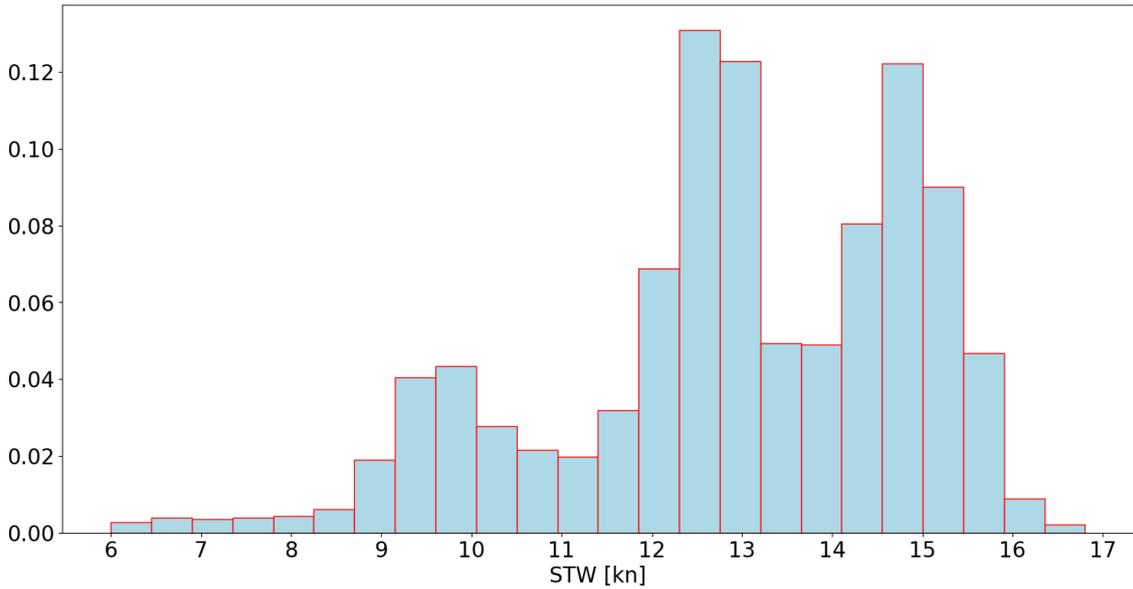


Figure 19: Relative frequency histogram of STW for loading condition 3 (ballast).

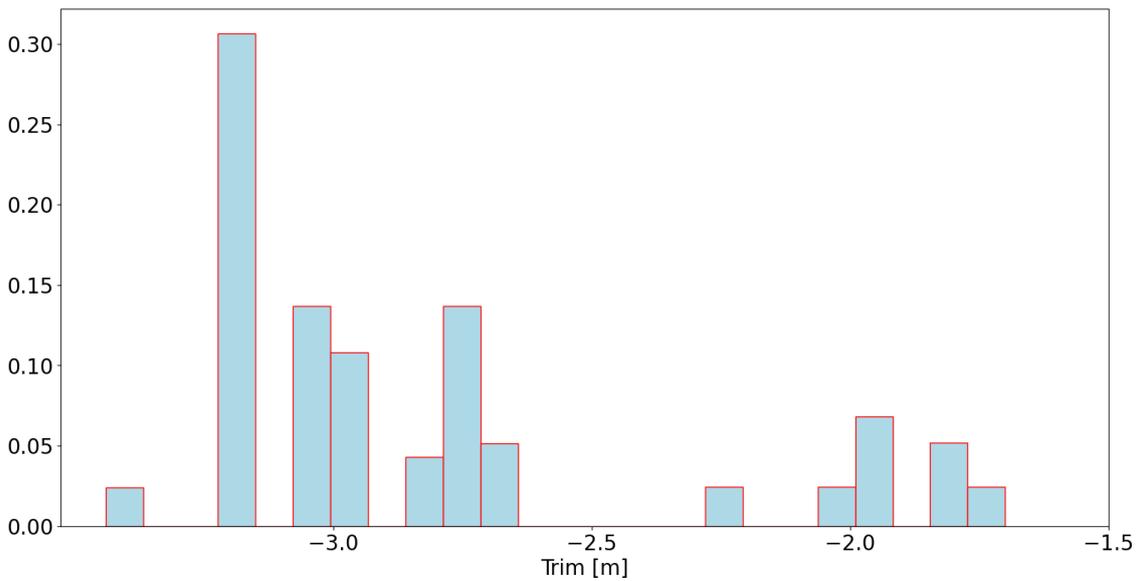


Figure 20: Relative frequency histogram of trim for loading condition 3 (ballast).

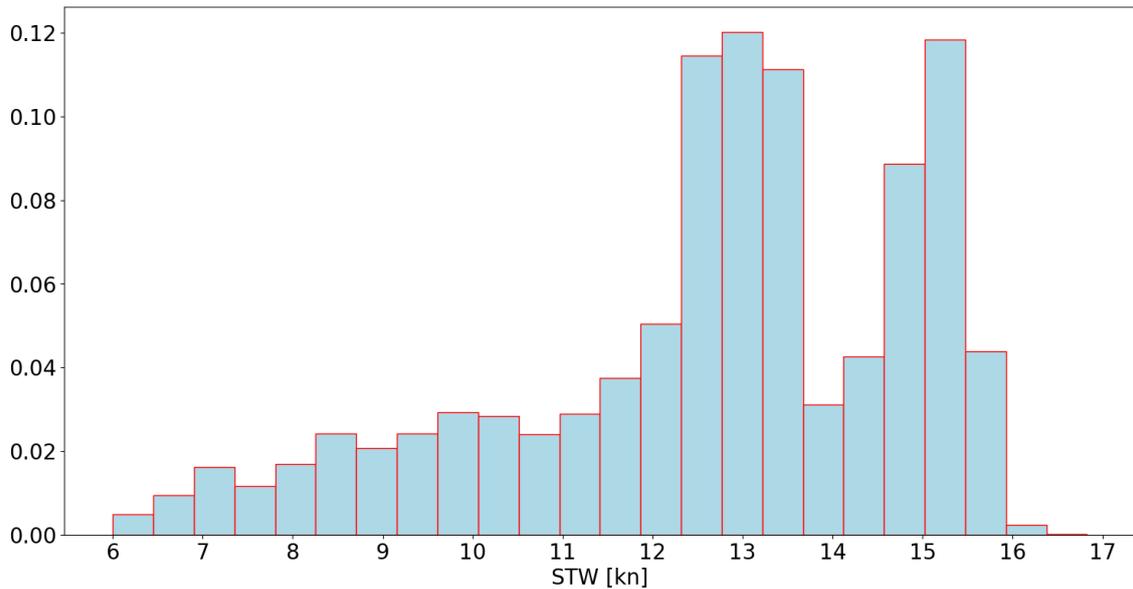


Figure 21: Relative frequency histogram of the STW for loading condition 4 (heavy ballast).

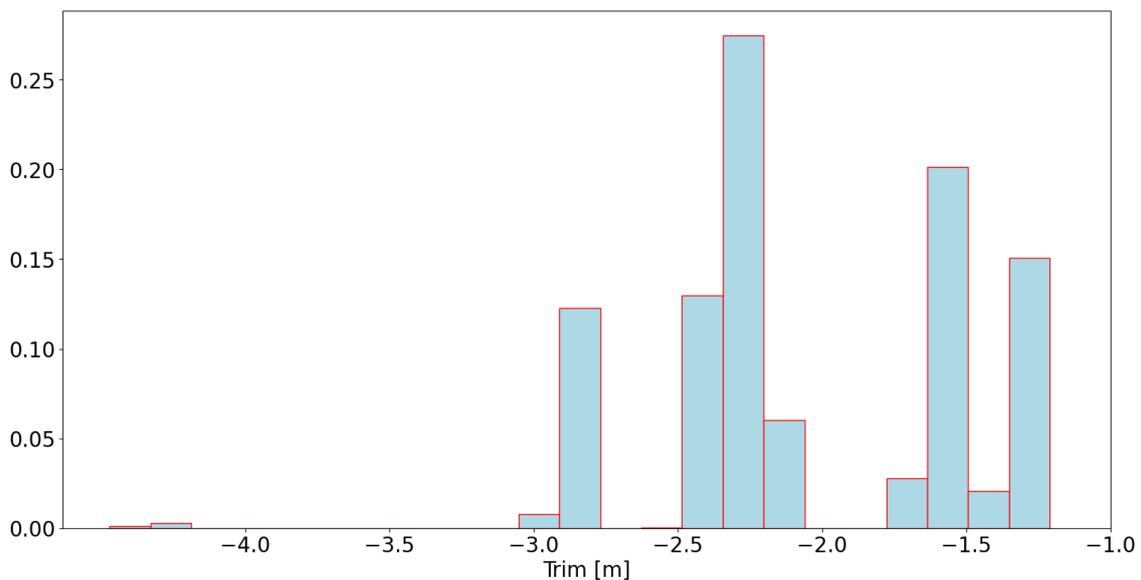


Figure 22: Relative frequency histogram of trim for loading condition 4 (heavy ballast).

4.3 Identification of maintenance actions using KPI

An assessment of the vessel hydrodynamic performance makes sense only in sailing conditions. Thus, it is important to identify into the data records the time intervals during which the ship was traveling. For this purpose, a list of the start and end dates-times of voyages during the reporting period is derived from noon reports, and then a dataset containing only voyages is extracted to exclude port calls. To ensure that the data analysis is focused on specific conditions related with open-sea operation, certain criteria are subsequently applied to the dataset, i.e., data points are rejected based on threshold values. A list of the lower and upper limits used in this study is provided in Table 10.

Table 10: Threshold values used for the filtering of the SHP, STW, Rate of Turn and Current speed.

Variable	Unit	Lower Limit	Upper Limit
SHP	kW	3000	-
STW	kn	6	-
Rate of turn	deg/min	-6	6
Current speed	kn	-3	2

Furthermore, due to fouling accumulation, the vessel hydrodynamic performance may differ drastically at different points in the operation timeline. To assess the changes in vessel performance over time, it is necessary to know the dates of the various maintenance events that took place during the reporting period. In Figure 23 the vessel operation timeline is shown, containing the date of each important event. The dataset is split, based on maintenance events, into four distinct datasets (A, B, C, D), described in Table 11. During dataset A the propeller was damaged, so the performance was bad and non-representative of the vessel normal condition. Dataset B represents a period during which marine growth has accumulated. Due to the short duration between cleaning events, Dataset C captures a limited period where fouling build-up is not substantial. Dataset D, spanning over a year, contains the most appropriate data for studying fouling accumulation, due to its extended timeframe. Datasets D1 and D2 are subsets of D: the former spans the first 2.5 months of D, the latter the last 3 months.

Table 11: Definition of datasets based on the dates of the main maintenance events. Datasets D1 and D2, are subsets of dataset D.

Dataset ID	from	to
Dataset A	05-02-2021	19-07-2021
Dataset B	21-07-2021	08-01-2022
Dataset C	05-02-2022	21-05-2022
Dataset D	22-05-2022	29-06-2023
Dataset D1	22-05-2022	08-08-2022
Dataset D2	23-03-2023	29-06-2023

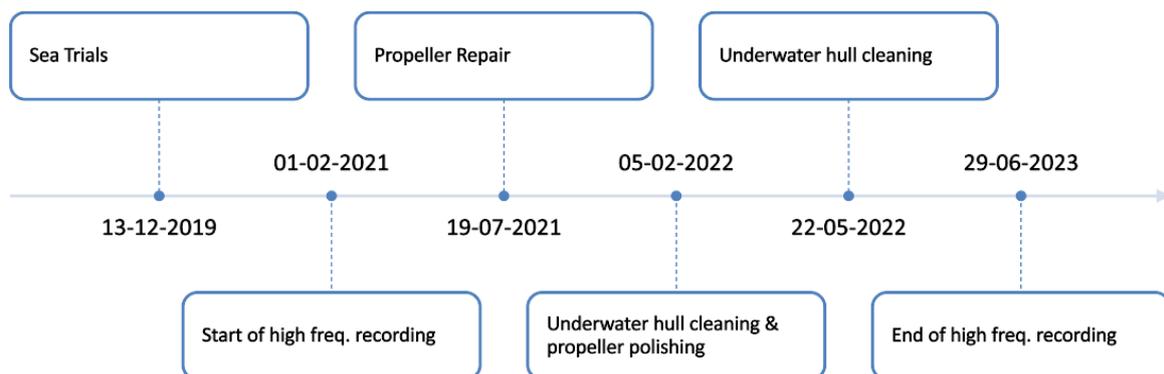


Figure 23: Timeline of the most important events during the vessel operation.

To assess the impact of hull and propeller fouling on ship performance, the change over time of an appropriate KPI is observed. The KPI employed in this study is a customary propeller loading coefficient defined as follows:

$$KPI = \frac{P}{N^3} \quad (1)$$

where P is the ME power in kW and N is the propeller rate of revolution in RMP. In Figure 24 the KPI is plotted throughout the reporting period using separate colours for laden (green) and ballast (blue) conditions to demonstrate that as the draft increases, the KPI increases as well. The maintenance events dates are also shown in the plot, based on which the dataset is split into individual periods. Due to high variance caused by transients and various loading and weather conditions, it is difficult to directly interpret the variation of KPI. Therefore, a first order trendline is fit over the data available in each period in-between two consecutive maintenance events. The uptrend observed in each period trendline accounts for the hull and propeller deterioration due to fouling. In addition, the drop of the KPI noticed after each maintenance event represents the improvement in the vessel performance achieved as a result of the maintenance itself.

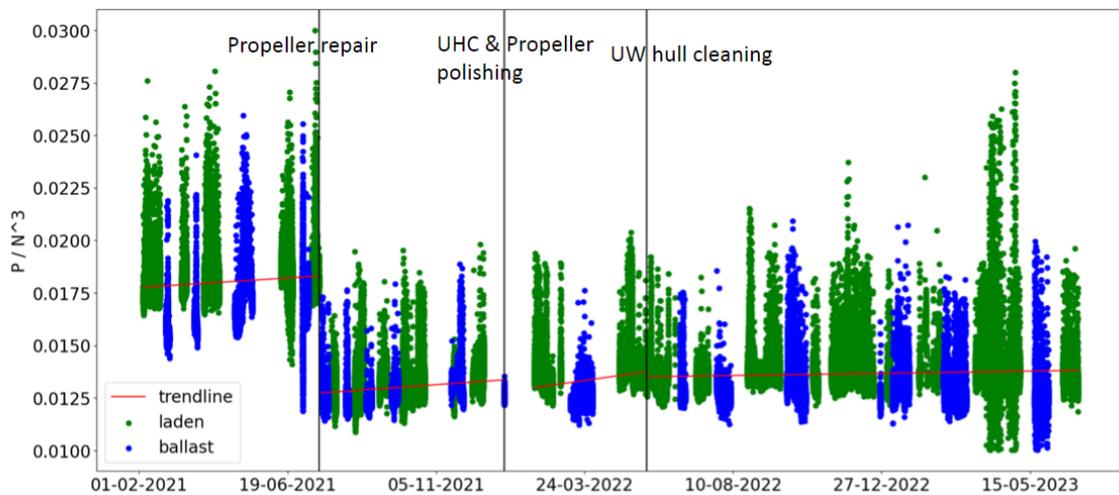


Figure 24: Utilization of KPI to monitor the status of the hull and the propeller performance, as well as to track the main maintenance events.

Despite being intuitive and easy to implement, this approach presents a few shortcomings. First, it fails to answer the crucial question on how much power is saved as a result of a maintenance event, and, second, it does not account for the various factors affecting the ME power consumption. Power normalization methods for environmental factors exist [10], but they serve as a coarse approximation, since they rely on assumptions that introduce uncertainty. For a more accurate and holistic approach to the problem, sophisticated data-driven methods ought to be employed. Using a well-tuned, highly accurate black-box model, the vessel's behaviour in various conditions can be predicted. Such a model constitutes a digital twin of the vessel, meaning that it can be used to simulate scenarios [35].

5 Data-driven models

5.1 Methodology

Data-driven models which are developed using the available operational data and suitable Machine Learning (ML) algorithms are henceforth referred to as ML models. The workflow of an ML process aimed at assessing ship performance and supporting the maintenance optimization is shown in Figure 25. Given the compromised state of the FOC measurements, the Shaft Horse-Power (SHP) is selected as the output variable of the ML model. The input variables that feed the model to predict ME SHP are termed features. The first step of the process regards feature engineering, which consists of feature construction and feature selection. The data preparation step deals with the elimination of erroneous data points, to derive a dataset free from anomalies focusing on the desired operational conditions. Subsequently, various ML algorithms are evaluated, and different hyperparameter combinations are explored, using state-of-the-art techniques, to develop a model with high predictive accuracy. Hull and propeller condition monitoring is achieved by employing the ML model to evaluate the expected ship performance.



Figure 25: Steps of the development of the Machine Learning (ML) method to build the hull and propeller monitoring tool.

5.2 Feature engineering

Feature engineering requires domain knowledge and comprises feature construction and feature selection. In feature construction, new variables are built using existing ones for a more accurate physical description of the problem. Specifically, aft and fore draft measurements are combined to calculate the mean draft and trim of the vessel, according to Eq. 2 and 3. In addition, the sea current speed is calculated as the difference between SOG and STW, as stated in Eq. 4. The relative wind speed and direction measurements are combined to determine the longitudinal and transverse components of the relative wind speed, according to Eq. 5 and 6. Similarly, in Eq. 7 the relative wave angle is calculated to determine the longitudinal and transverse components of the significant wave height, according to Eq. 8 and 9. More in general, the installation of a dedicated wave radar could be recommended to provide a detailed dataset, useful to determine the actual sea conditions and to support additional ship functions (e.g. seakeeping monitoring, hull-induced stresses, hydroelasticity, Decision Support System for navigation) To account for the strong time dependence of fouling growth, the Days feature is added, which is equal to the number of days passed since the last hull and propeller maintenance event. For each data point recorded on a given date, the Days feature is calculated according to Eq. 10. Thus, for each dataset mentioned in Table 11, the Days feature is created separately, based on the respective starting date.

$$\text{Mean draft} = \frac{\text{Fore Draft} + \text{Aft Draft}}{2} \quad (2)$$

$$\text{Trim} = \text{Fore Draft} - \text{Aft Draft} \quad (3)$$

$$\text{Current} = \text{SOG} - \text{STW} \quad (4)$$

$$\text{Wind } x = \text{Relative Wind Speed} \cdot \cos(\text{Relative Wind Angle}) \quad (5)$$

$$\text{Wind } y = \text{Relative Wind Speed} \cdot \sin(\text{Relative Wind Angle}) \quad (6)$$

$$\text{Relative Wave Angle} = \text{abs}(\text{Vessel Heading} - \text{WaveDirection}) \quad (7)$$

$$\text{Wave } x = \text{Wave Height} \cdot \cos(\text{Relative Wave Angle}) \quad (8)$$

$$\text{Wave } y = \text{Wave Height} \cdot \sin(\text{Relative Wave Angle}) \quad (9)$$

$$\text{Days} = \text{Date} - \text{Date of last maintenance event} \quad (10)$$

In feature selection, the most important input variables are determined with the help of appropriate statistical tools, whereas irrelevant or redundant input variables are eliminated, to simplify the model and prevent overfitting. To determine the necessary features, two criteria are used: comprehension of the physical problem and correlation analysis. Based on domain knowledge, among all available variables a few are shortlisted as feature candidates for SHP prediction. Then, correlation analysis is performed on those candidates using the distance correlation method, and the results are presented in Table 12. Distance correlation was presented by [37] and constitutes a measure of statistical dependence between two random variables that captures both linear and nonlinear relationships. More details about this are presented in Appendix E: Distance Correlation.

Table 12: Calculation of the distance correlation coefficient between SHP and the candidate features.

Feature	Unit	Distance Correlation Coefficient
RPM	rpm	0.8620
STW	kn	0.4509
Mean Draft	m	0.3014
Trim	m	0.3724
RoT	deg/min	0.0778
Wave	m	0.1434
Wind	m/s	0.1816
Wave x	m	0.2007
Wave y	m	0.1115
Wind x	m/s	0.2252
Wind y	m/s	0.0939
Current	kn	0.1737
Days	days	0.2895

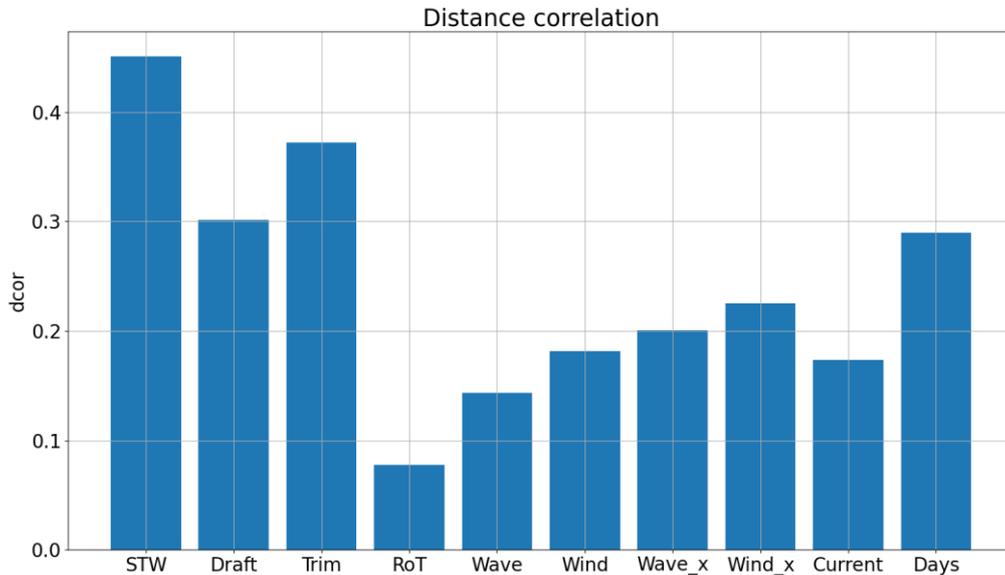


Figure 26: Bar plot of the distance correlation coefficient between SHP and the considered features.

Correlation analysis confirms what is expected: RPM and STW are the candidate features with the highest correlation values with the SHP variable, which highlights their importance as predictors. However, the RPM variable is not included as a feature, because it would limit the applicability of the model as an emulator, even though it would increase its predictive accuracy. It is worth mentioning that the correlation coefficients values of the longitudinal components of the relative wind speed and the significant wave height are higher than the respective values of both the transverse components and the original variables. In addition, it is worth noting that although the correlation coefficient of the Rate of Turn (RoT) variable has a small value, it is useful to include a variable related to manoeuvring.

In Figure 26, a bar plot is shown, displaying the distance correlation coefficient between the SHP and the considered features. Eventually, two sets of features are selected, one including the Days feature and the other not including it, as listed in Table 13. The models derived by each feature set will be used in a different way, to monitor the hull and propeller condition.

Table 13: List of the selected features with or without the Days feature.

Feature set w/o Days	Feature set w/ Days
STW	STW
Mean Draft	Mean Draft
Trim	Trim
RoT	RoT
Wind x	Wind x
Wave x	Wave x
Current	Current
-	Days

5.3 Data preparation

Before feeding the data to an ML algorithm, a suitable preparation is required to ensure their quality. In the present study, data preparation is implemented in three steps:

- data smoothing
- voyage extraction
- threshold value filtering

These processes are described hereafter.

Data smoothing aims to remove signal noise, while preserving the main trend. To do this, first, a Simple Moving Average (SMA) filter is applied to each signal (features and output variables). The signal is denoted as $X(i)$, where i represents the data sample at a certain time. The equation for calculating the SMA at the time i , denoted as $SMA(i)$, with a window size of k can be expressed as:

$$SMA(i) = \frac{X(i) + X(i - 1) + X(i - 2) + \dots + X(i - k + 1)}{k} \quad (11)$$

Then, each SMA-smoothed signal is decimated by a factor of k , i.e., only one sample out of k samples is retained, where k is equal to the window size used for the moving average. In Figure 27 and Figure 28, window sizes of 5, 10 and 15 for two variables are compared to show the effect of k . Clearly, the wider is the time window, the more intense is the effect of the smoothing. The value of $k = 5$ minutes is commonly used in the literature, see for example [35] and [11], since it reduces the sharpest undesired fluctuations.

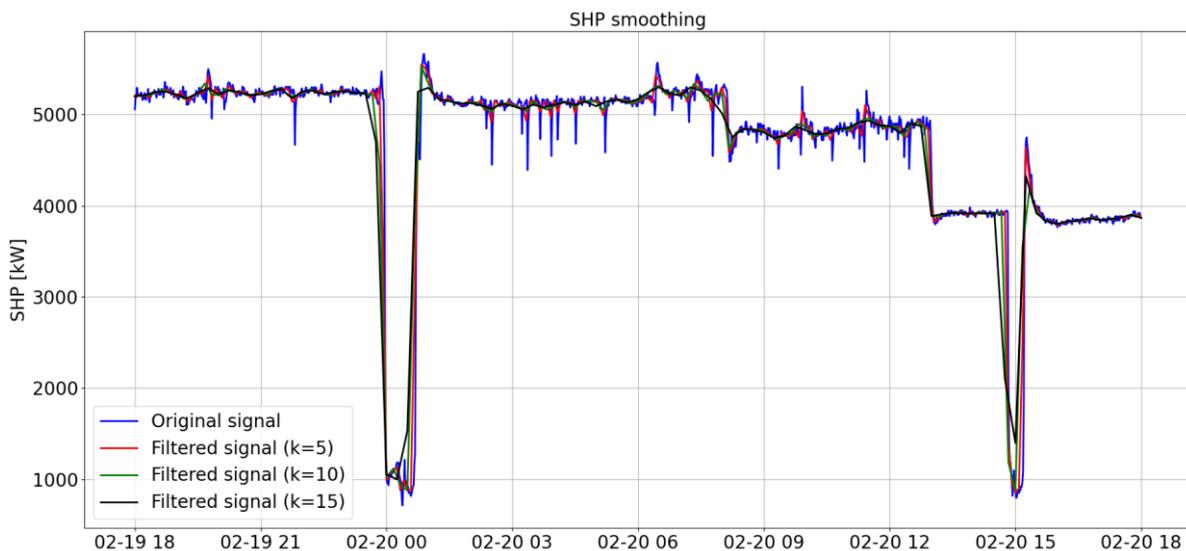


Figure 27: Comparison between the SHP signals over a one-day time interval, after smoothing with Simple Moving Average with different window sizes.

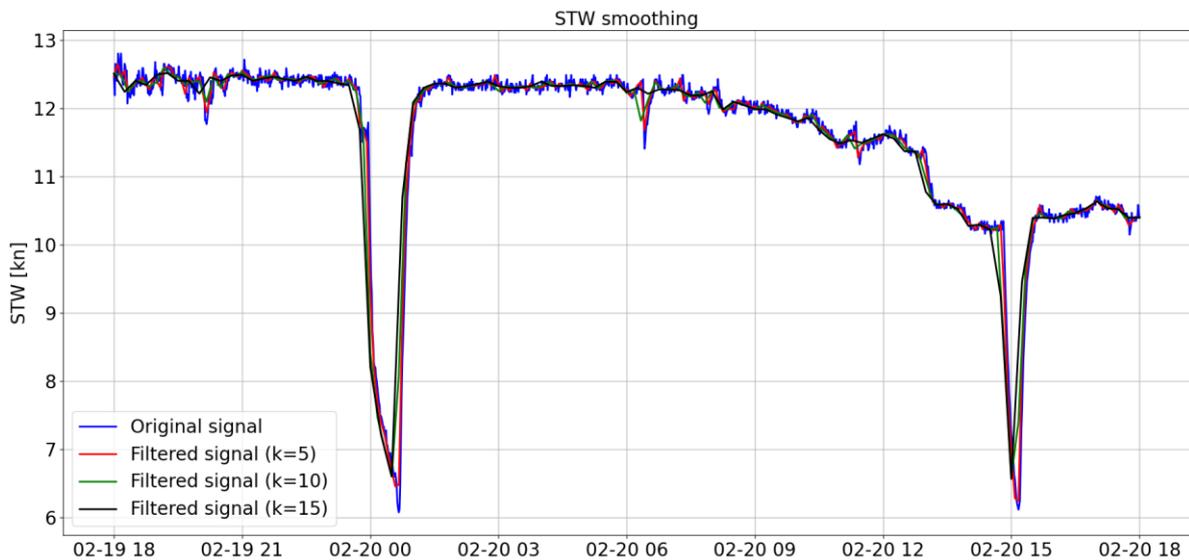


Figure 28: Comparison between the STW signals over a one-day time interval, and the signals obtained by using different values of k .

Voyage extraction is the process of identifying the intervals within the dataset that correspond to actual ship voyages, therefore excluding port calls. A list of the start and end date/time of voyages during the reporting period is derived from noon reports. Subsequently, a dataset containing only voyages is extracted.

Threshold value filtering involves the rejection of data points based on threshold values. In such a way, only the values inside the operational range of choice are kept. The histogram of each variable is examined, see for example in Figure 29, to identify the tails of the distributions. Table 14 lists the lower and upper limits of each variable.

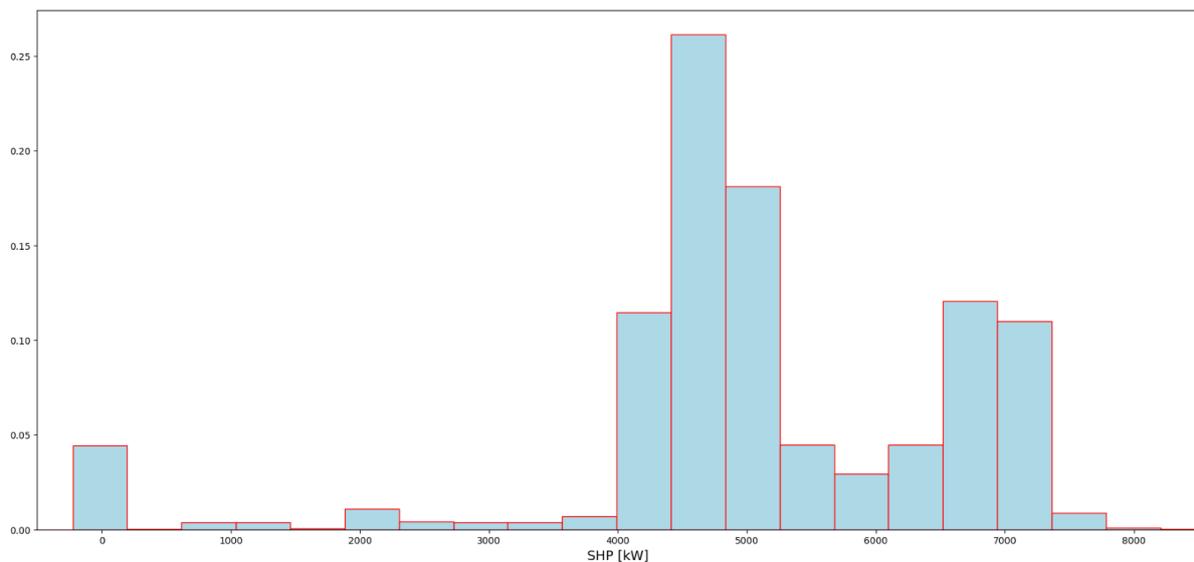


Figure 29: Relative frequency histogram of SHP, then used to identify the tails of the statistical distribution of SHP.

Table 14: Threshold values used for filtering the ship navigation variables.

Variable	Unit	Lower limit	Upper limit
SHP	kW	3000	-
RPM	rpm	55	-
STW	kn	6	-
RoT	deg/min	-6	6
Current	kn	-3	2

These data preparation steps cause a progressive reduction of the number of data points with respect of the raw data, as shown in Figure 30, where the number of points after each step is indicated, showing that the greatest reduction occurs during the smoothing process. This reduction is a side-effect, however deemed necessary to eliminate spikes and reduce the noise present in the raw data.

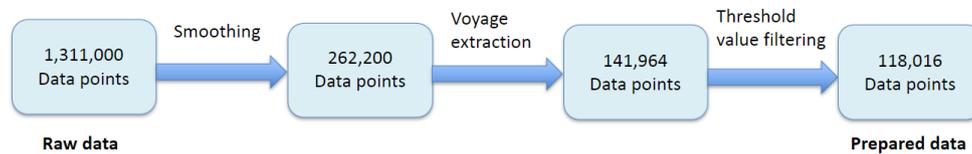


Figure 30: Reduction in the number of datapoints at each stage of data preparation for the entire dataset.

The last step of data preparation involves splitting the dataset based on maintenance events into four distinct datasets (A, B, C, D) described in Table 11. In addition, dataset D1 corresponding to 2.5-months of clean propeller and hull condition is extracted, as well as dataset D2 corresponding to the last 3 months of recording.

5.4 Machine Learning (ML) model development

Based on the two feature sets described in Table 13 and the datasets in Table 11, two separate ML models are developed. The feature set without Days is employed in Model-1, which is trained on dataset D1, while the feature set with Days is employed in Model-2, which is trained on dataset D2. Details of these models are given in Table 15.

Table 15: Overview of the two ML models, their feature sets (see Table 13), and training datasets (see Table 11).

	Model-1	Model-2
Feature set	Without Days (7 features)	With Days (8 features)
Training set	Dataset D1 (22-05-2022 to 08-08-2022)	Dataset D (22-05-2022 to 29-06-2023)

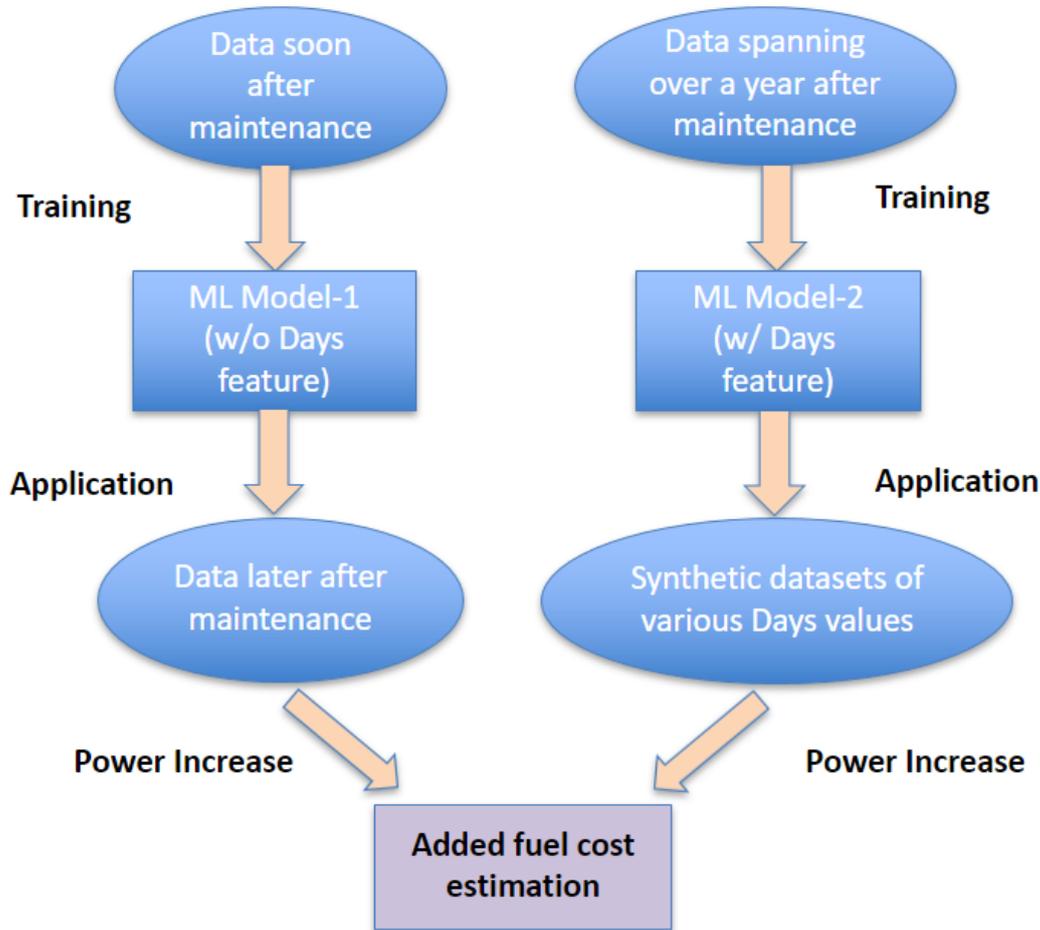


Figure 31: Flowchart illustrating the training and application process of Model-1 and Model-2 on separate datasets, within the context of hull and propeller condition monitoring.

The procedure for the development of both models is the following one. First, the dataset is randomly shuffled and split into training sets and test sets using a ratio of 80% to 20% respectively. Then various ML algorithms, suitable for regression tasks, are evaluated on the training set using 5-fold cross-validation. The ML algorithms considered are Decision Trees and ensemble methods based on Decision Trees, i.e., Random Forests, Extremely Randomized Trees (Extra-Trees), and Gradient Tree Boosting. Theoretical descriptions of the examined machine learning algorithms can be found in Appendixes A, B, C and D. These appendices provide explanations of the underlying mechanisms of each algorithm, as well as insights into the specific hyperparameters that were tuned during this study.

In summary, Decision Trees can capture complex interactions and patterns, thus they can handle a wide range of regression problems, including those with nonlinear or non-monotonic relationships. Ensemble methods combine multiple Decision Trees, each one trained on different subsets of the data, providing multiple predictions. The combination of the information coming from these multiple predictions improves the overall accuracy and allows a generalization of the model. Decision Trees are prone to overfitting, which occurs when the model becomes too specific to the training data and performs poorly on unseen data, i.e., instances that were not part of the training dataset. Ensemble methods alleviate this issue by reducing overfitting through techniques like random subspace sampling (Random Forests) or gradient-based optimization (Gradient Boosting). These methods

introduce randomness and regularization, leading to better generalization and improved performance on unseen data.

Moreover, the averaging employed in Ensemble's methods helps to mitigate the influence of individual noisy predictions and leads to more robust overall predictions. In ensemble methods, multiple models are combined, and their predictions are averaged to produce a final prediction. By doing so, the impact of individual predictions that may contain noise or errors is reduced, resulting in more reliable and stable predictions across the ensemble.

Being every problem different, the most suitable algorithm cannot be known a priori, thus a comparative analysis is required. In addition, the performance of the algorithm is highly dependent on the configuration of the hyperparameters. To determine a set of well performing hyperparameters, the hyperparameter search space should be explored. An exhaustive exploration of the search space is impossible, so only a few selected combinations are tested, either with grid search or a randomized search. Grid search evaluates all the possible combinations of specified hyperparameter values, whereas randomized search selects a random value for each hyperparameter in each iteration, so that a set number of random combinations are evaluated. For a more extensive exploration of the search space, the randomized search is performed over a wide range of hyperparameter values, coupled with 5-fold cross validation, to account for intrinsic randomness in the training. The metric used to evaluate the combination of hyperparameter values is the mean cross validated Root Mean Squared Error (RMSE), which provides a measure of model generalization capacity. Thus, for each algorithm the best combination of hyperparameters values is found (see Table 16), and the respective mean cross validated RMSE is reported in Table 17. By comparing the cross validated error results, the ML algorithm with the lowest error is selected, i.e., the Extra-Trees ensemble method.

The model development process is completed by evaluating the selected ML algorithm on the test set, with the objective to assess the model generalization capacity. The test set provides a reliable final measure of how well the model performs when applied to unseen data, because it has not been involved in either training or validation. Thus, error metrics suitable for regression tasks are used, and the results are reported in Table 18. Moreover, in Figure 32 the SHP values predicted by the model in the test set are plotted against the actual SHP values, and the points are clustered around the 45-degree line. Similarly, in Figure 33 the residual SHP values (actual minus predicted) in the test set are plotted against the predicted SHP values, and the points are clustered around the horizontal line through zero. A comparison between the results of Model-1 and Model-2 in both the test set and cross validation reveals a slightly better performance of the model with Days, which is due to the additional information provided by the temporal feature.

The performance of the model is evaluated by using error metrics on the test set, which consists of data absent from the training and validation process. Let \hat{y}_i be the value predicted by the model for i -th sample, y_i the respective actual value and $\epsilon_i = y_i - \hat{y}_i$ the corresponding error. Thus, the following error metrics are defined, which are suitable for regression analysis [35]:

- Coefficient of Determination, R^2 :

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (12)$$

where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ and $\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n \epsilon_i^2$

- Mean Absolute Error, MAE

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| = \frac{1}{n} \sum_{i=1}^n |\epsilon_i| \quad (13)$$

- Root Mean Squared Error, *RMSE*

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} = \sqrt{\sum_{i=1}^n \epsilon_i^2} \quad (14)$$

- Mean Absolute Percentage Error, *MAPE*:

$$MAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| / \max(\epsilon, |y_i|) \cdot 100\% \quad (15)$$

where ϵ is an arbitrarily small positive number, to avoid zeroing the denominator.

Table 16: Overview of the hyperparameter values explored for each ML algorithm using randomized search combined with cross-validation.

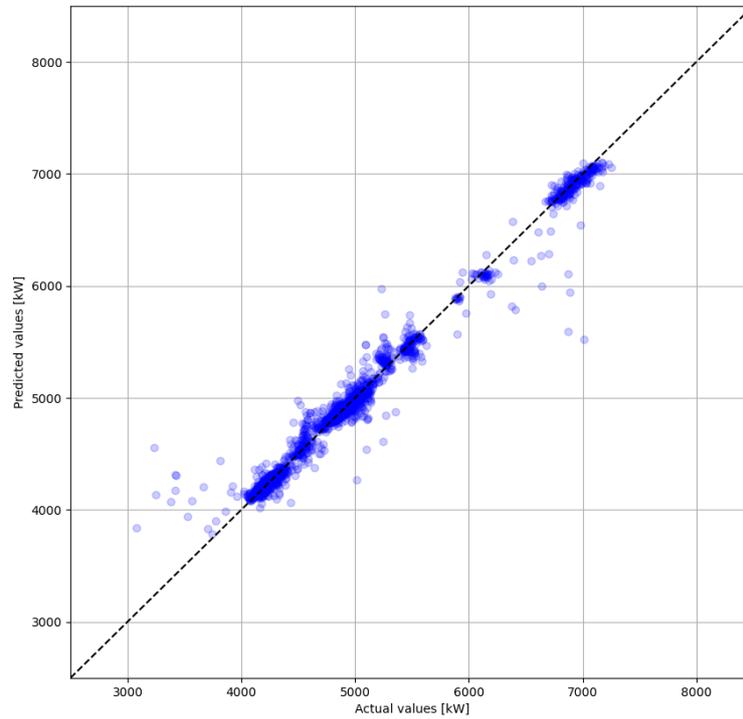
	Hyperparameter	min	max	step	best
Decision Tree	max_depth	5	30	1	23
	min_samples_leaf	1	20	1	3
Random Forest	max_depth	5	30	1	20
	min_samples_leaf	1	20	1	4
	n_estimators	20	200	5	110
Extra Trees	max_depth	5	30	1	20
	min_samples_leaf	1	20	1	4
	n_estimators	20	200	5	120
Gradient Tree Boosting	max_depth	5	30	1	14
	n_estimators	20	200	5	135
	learning_rate	0.05	1	0.01	0.3

Table 17: Mean cross validated RMSE for each ML algorithm.

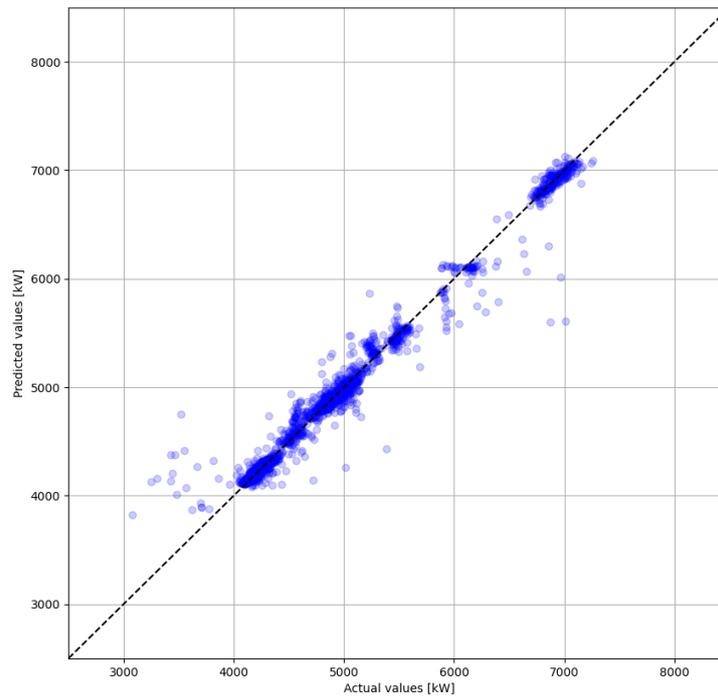
ML algorithm	Model-1 mean cross validated RMSE [kW]	Model-2 mean cross validated RMSE [kW]
Decision Tree	185.8	151.2
Random Forest	158.2	130.7
Extra-Trees	158.9	127.2
Gradient Tree Boosting	159.3	128.6

Table 18: Performance evaluation of Model-1 and Model-2 on the test set, using a variety of error metrics.

	Model-1	Model-2
R ²	0.979	0.981
MAE [kW]	56.6	54.1
MAPE	1.12%	1.11%
RMSE [kW]	115.1	109.9



(a)



(b)

Figure 32: Plot of the predicted SHP values against the actual SHP values of the test set: (a) for Model-1, and (b) for Model-2.

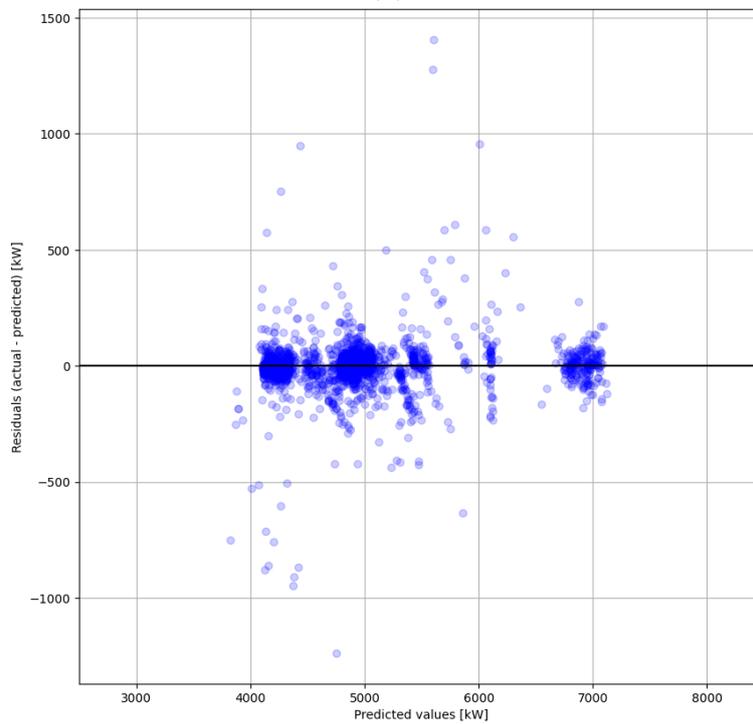
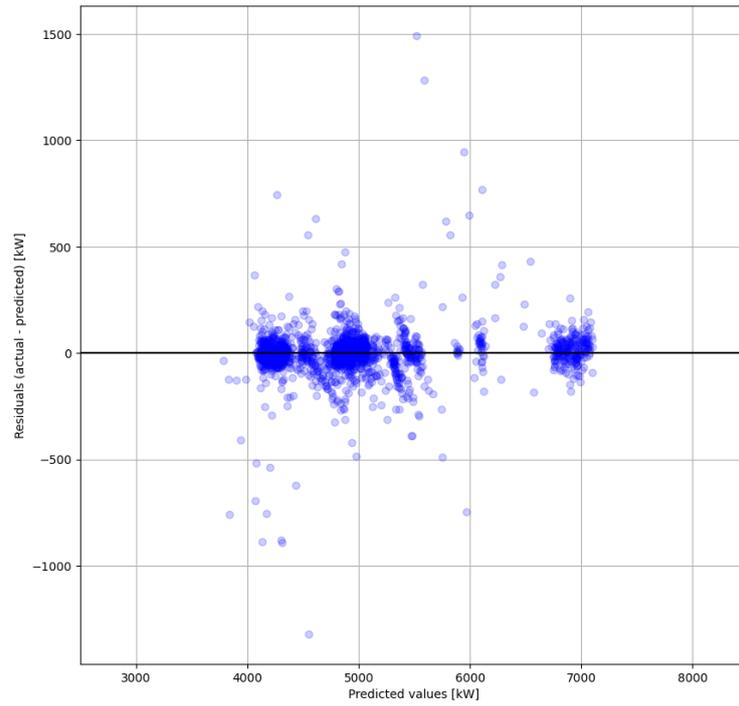


Figure 33: Plot of residual SHP values against the predicted SHP values on the test set: (a) for Model-1, and (b) for Model-2.

6 Optimization maintenance analysis

The purpose of this task is the development of a hull and propeller condition monitoring system for the MV Kastor ship to support maintenance optimization by making use of operational data, estimating the expected behaviour and using appropriate performance indicators. As time elapses, marine growth on hull and propeller progresses, and thus the vessel viscous drag increases. This results in an increase in the thrust that must be delivered by the propulsion system to maintain a given ship speed, which leads to increased shaft power demand. The main assumption behind this reasoning is that, apart from the fouling condition of the vessel, the other conditions affecting the power demand remain the same. The goal of the data-driven modelling approach followed in the present study is to satisfy this assumption. Thus, two methods are proposed to assess the fouling condition using the ML models developed in the previous chapter. The power increase indicator (PI) is used, which is defined as:

$$PI = 100 \frac{P_f - P_c}{P_c} \quad (16)$$

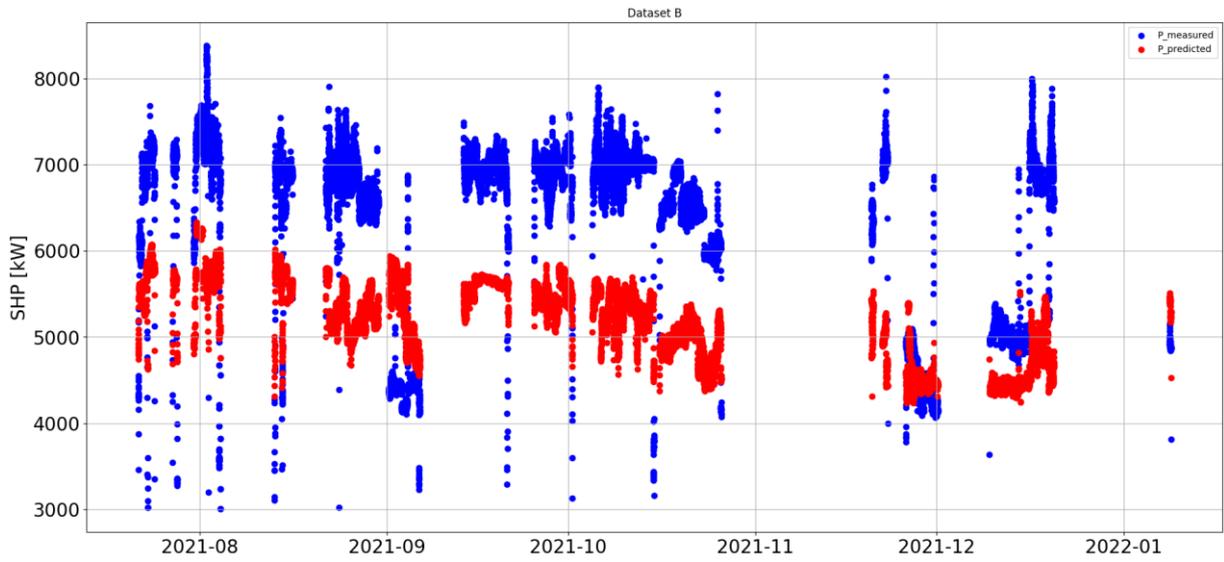
where P_c is the SHP in clean vessel condition and P_f is the SHP in fouled vessel condition, both values corresponding to the same operational conditions.

6.1 Method-1 for maintenance optimization

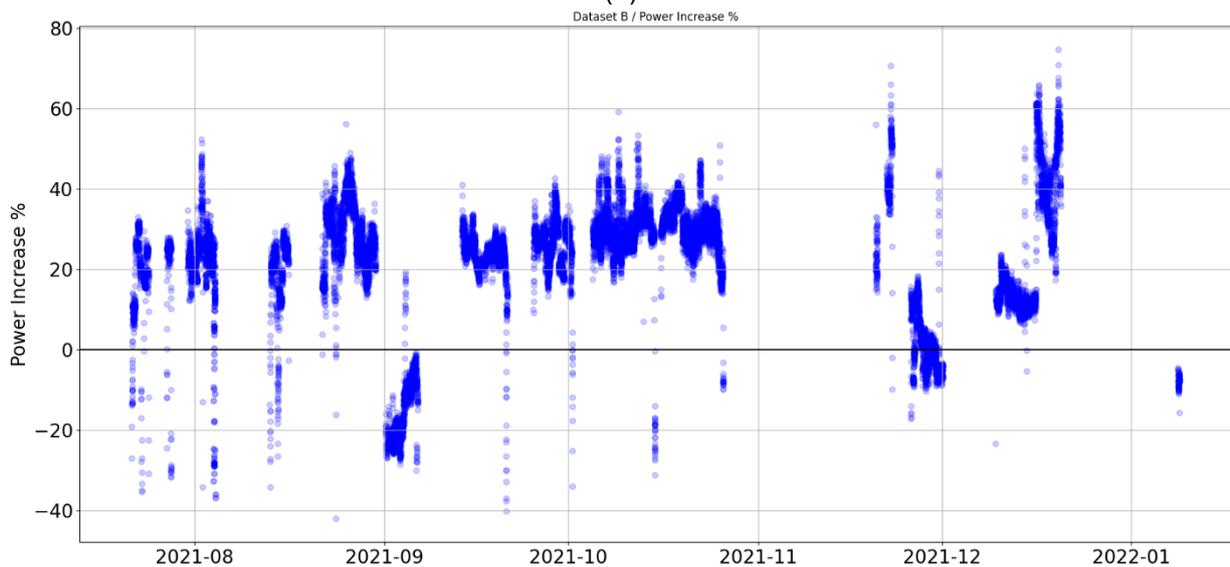
The main idea here is to employ Model-1 as a reference performance baseline, representing the vessel clean condition, to evaluate conditions with fouling. Model-1 is trained on data corresponding to a few months after hull cleaning and propeller polishing, so it emulates the behaviour of the vessel when it is clean. The approach of Method-1 is to apply Model-1 on data corresponding to a period much later in time after the ship maintenance, to capture the power penalty caused by biofouling. According to Table 11, the two datasets corresponding to advanced biofouling are B and D1. The difference observed between the actual power and the values predicted by Model-1, when applied to these datasets, serves as an estimate of the hull and propeller deterioration due to biofouling.

Dataset B is about two years away from sea trials and no hull cleaning or propeller polishing has been performed in the meantime. Thus, it represents a period in which severe marine growth has occurred, and, as a result, the hydrodynamic performance has deteriorated. In Figure 34, a scatter plot of the actual power and that predicted by Model-1 is shown, as well as a scatter plot of the power increase indicator. The estimated mean power increase is 21.7%, as reported in Table 19.

Dataset D2 is extracted about a year away from the previous hull cleaning event, three months prior to when the propeller was polished. In this period, biofouling is expected to have caused a considerable power penalty, even if not as significant as than in dataset B. Indeed, when Model-1 is applied to dataset D2, the actual power values are greater than the predicted values, which translates into a significant power increase, as shown in Figure 35. Specifically, the mean power increase is estimated to be 10.8%, as reported in Table 20.



(a)



(b)

Figure 34: Application of Model-1 to dataset B. (a): Scatter plot of the actual and predicted SHP over time, (b) scatter plot of the PI over time.

Table 19: Application results of Model-1 to dataset B, showing the mean values of the actual and predicted SHP, as well as the mean PI value.

Mean actual SHP	kW	6,342
Mean predicted SHP	kW	5,209
Difference in means	kW	1,133
Mean PI	-	21.7%

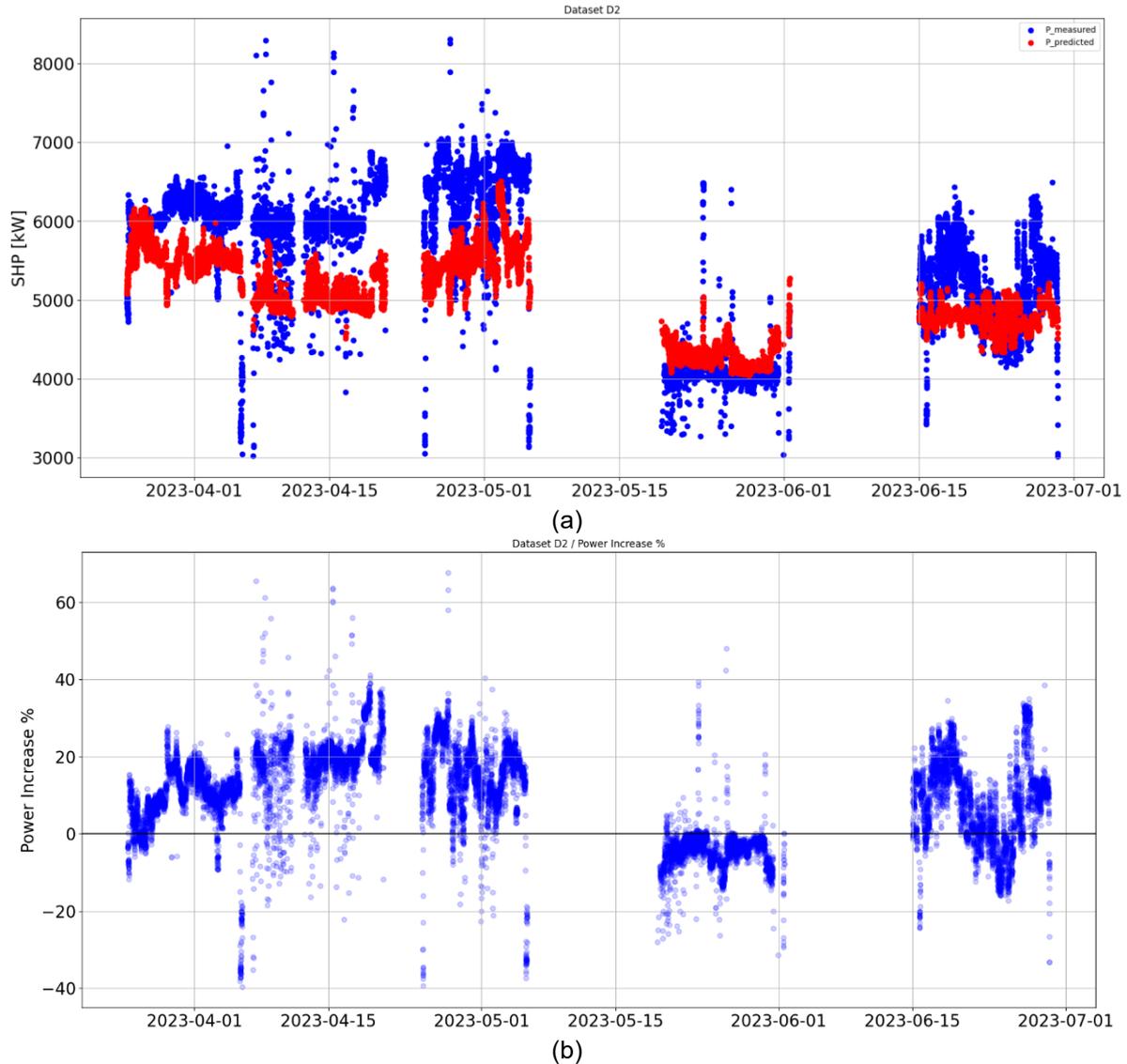


Figure 35: Application of Model-1 to dataset D2. (a): Scatter plot of the actual and predicted SHP over time, (b) scatter plot of the PI over time.

Table 20: Application results of Model-1 to dataset D2, showing the mean values of the actual and predicted SHP and the mean PI value.

Mean actual SHP	kW	5,546
Mean predicted SHP	kW	4,984
Difference in means	kW	562
Mean PI	-	11.3%

6.2 Method-2 for maintenance optimization

In this approach Model-2 is employed to monitor the advance of biofouling in the period corresponding to dataset D. The distinctive feature of Model-2 is Days, which enables emulation of scenarios referring to different points in time. To achieve this, synthetic datasets are created by assigning specific values to the model features. Although the goal here is to evaluate the effect of

Days, there are other input variables playing a crucial role in the output, namely the STW, the Draft and the Trim. Thus, two loading conditions, one laden and one ballast, are considered, setting Draft and Trim equal to the respective mode values of dataset D. Then, for each condition, values evenly spaced from 12 to 14 kn in steps of 0.5 kn are assigned to the STW, since most values of dataset D are in this range. Furthermore, a calm sea condition with no current effect is considered, so the wave, wind and current variables are set equal to zero. Additionally, the RoT variable is assumed to be zero, indicating that the vessel is traveling on a steady course. Lastly, for each loading condition, three values for the Days feature, those most frequently found in dataset D, are selected over the entire period of dataset D. The feature values listed above are summarized in Table 21 which describes the synthetic datasets of laden and ballast conditions.

Then, Model-2 is applied to the synthetic datasets and predictions are grouped by loading condition and day. In Figure 36, the predicted power is plotted against the STW, and a least squares polynomial fit of 3rd degree is applied to each group of points. This fitting process yields an estimate of the propeller curves for laden and ballast conditions, respectively. The mean predicted power is calculated in each group of points, as listed in Table 22 and Table 23. In Figure 37, for each loading condition the values of the mean predicted power are plotted against the corresponding values of Days, along with the corresponding best-fit straight lines.

Table 21: Overview of the synthetic datasets designed separately for laden and ballast conditions.

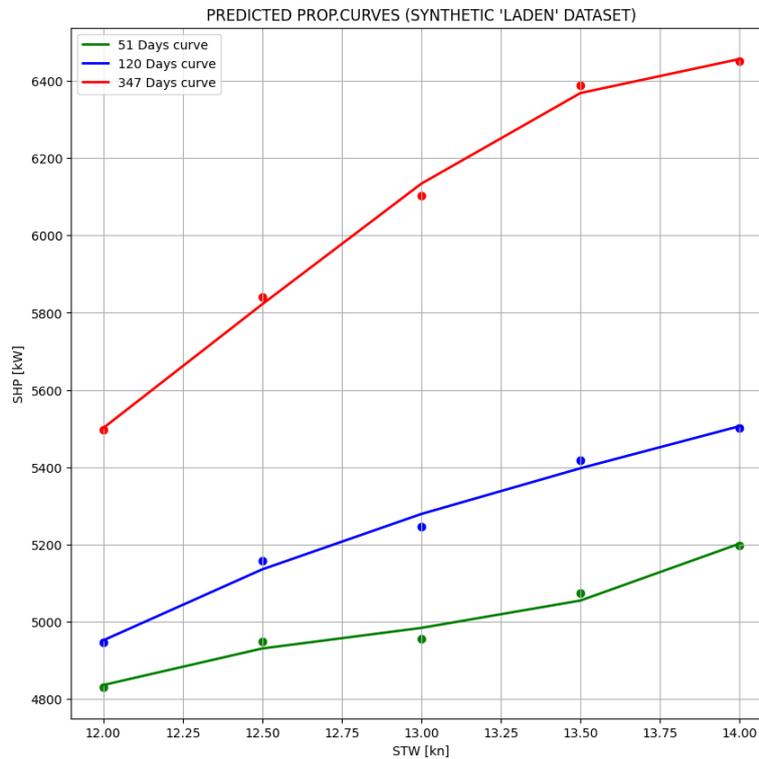
		Laden	Ballast
STW	kn	12.0, 12.5, 13.0, 13.5, 14.0	12.0, 12.5, 13.0, 13.5, 14.0
Draft	m	14.4	6.81
Trim	m	-0.13	-2.56
RoT	deg/min	0	0
Wind x	m/s	0	0
Wave x	m	0	0
Current	kn	0	0
Days	days	51, 120, 354	31, 229, 297

Table 22: Mean predicted SHP and corresponding PI, when Model-2 is applied to the synthetic dataset representing the laden condition.

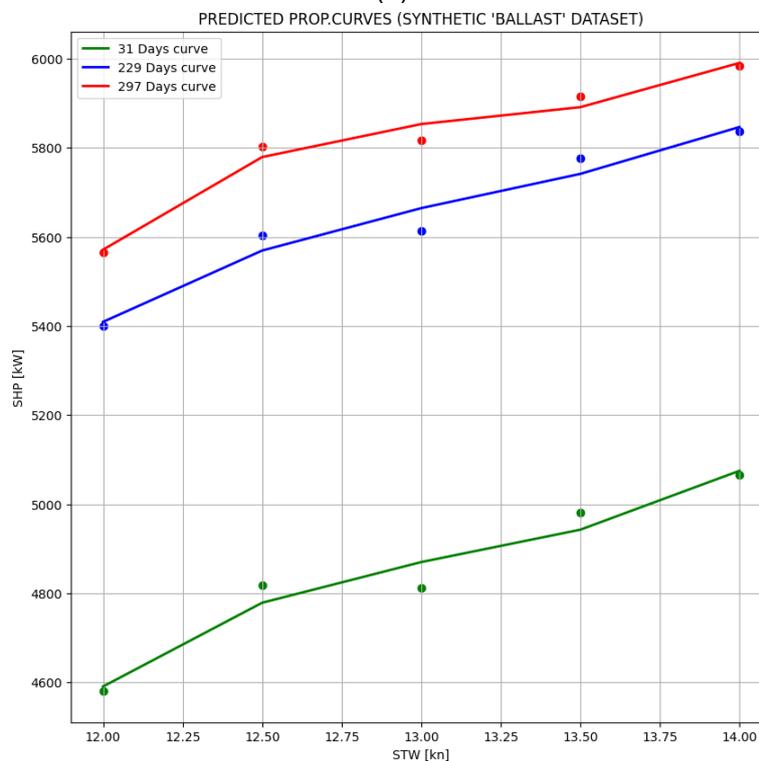
	Mean predicted SHP on synthetic 'Laden' dataset [kW]	Mean PI
Days = 51	5,006	-
Days = 120	5,283	5.5%
Days = 347	6,062	21.1%

Table 23: Mean predicted SHP and corresponding PI, when Model-2 is applied to the synthetic dataset representing the ballast condition.

	Mean predicted power on synthetic 'Ballast' dataset [kW]	Mean PI
Days = 31	4,851	-
Days = 229	5,640	16.3%
Days = 297	5,817	19.9%



(a)



(b)

Figure 36: Predicted SHP values on the synthetic datasets plotted against STW, with respective least squares polynomial fits of 3rd degree for (a) laden condition and (b) ballast condition.

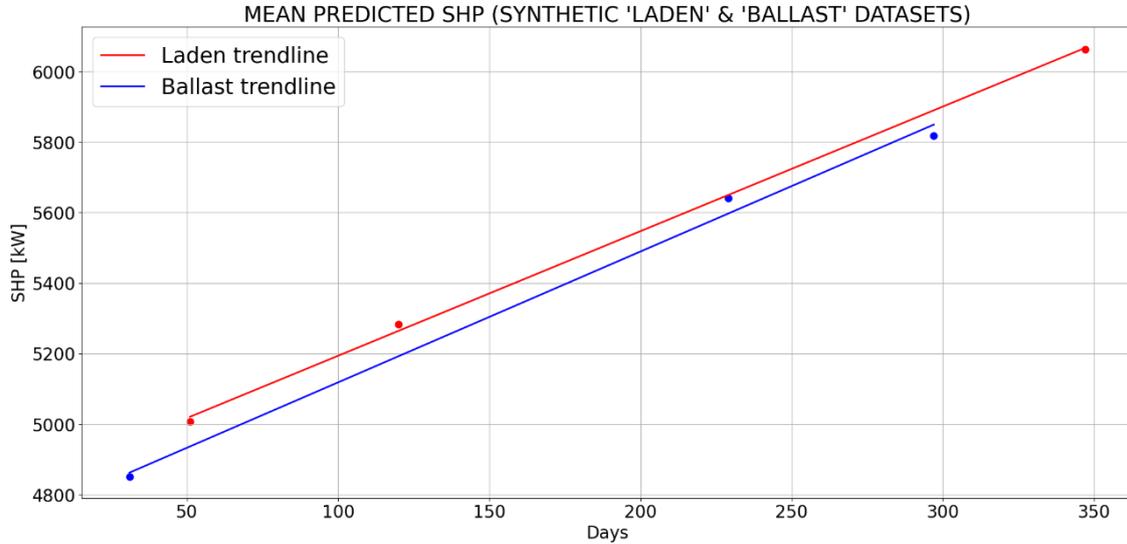


Figure 37: Mean predicted SHP values on the synthetic laden and ballast datasets in terms of days, plotted with respective linear fits.

6.3 Estimation of added fuel cost

For the maintenance optimization, the added fuel cost due to biofouling must be evaluated. The methods described above provide an estimate of the power penalty caused by biofouling growth over a certain period after hull cleaning and propeller polishing. To calculate the consequent increase in FOC (dFOC), the approximate values of specific fuel oil consumption (SFOC) at different loads are used based on ME's shop tests. Since the power values are known, the dFOC is calculated as follows:

$$dFOC = P_f \cdot SFOC(P_f) - P_i \cdot SFOC(P_i) \quad (17)$$

where P_i and P_f are the initial and final SHP values respectively, and $SFOC(P_i)$ and $SFOC(P_f)$ are the corresponding SFOC values. According to LASKARIDIS, the average fuel cost per ton during the examined period is 556 \$/ton, so the added fuel cost per day (dCost) is derived from the following equation:

$$dCost = dFOC \cdot FuelCost \quad (18)$$

For instance, by applying this reasoning to the power values predicted with Method-2 for each loading condition and for different ship speed scenarios, the added fuel cost per day due to biofouling which has accumulated during the corresponding period is estimated and reported in Table 24. Hence, the added fuel cost when the vessel operates under the conditions assumed in Method-2 is estimated to be on average equal to \$2,249 per day. Considering that hull cleaning and propeller polishing for a Kamsarmax vessel ranges between \$15,000 and \$18,000, where the higher one corresponds to the laden condition, it is concluded that on average in eight days of fouled operation the accumulated additional fuel cost equals the maintenance cost. Furthermore, the mean predicted power at each loading condition is linearly extrapolated using the straight-line fit to the points in Table 22 and Table 23 to predict the power for an assumed value of Days equal to 400. Thus, based on the behaviour of the vessel in past, on Day 400 an average added fuel cost of \$2,743 per day is predicted. This information can be used by the shipping company to decide the right time to perform



the next maintenance action, although other factors may influence this decision, such as the ship operation schedule.

Table 24: Estimated added fuel cost per day for the laden and ballast loading conditions, considering different values of the Days feature.

	From Day	To Day	dCost at STW=14kn	Mean dCost at STW=12÷14kn	Mean dCost at STW=12÷14kn and Day 400
Laden	51	347	2,791	2,362	2,769
Ballast	31	297	1,807	2,136	2,716



7 Hull condition-based monitoring and the regulatory framework

Over the last few years, Remote Inspection Techniques (RIT) have been introduced to perform general inspections of the steel structures on ships and on floating offshore units. RIT are based on machine learning, offering time-efficient and conceivably cost-effective alternatives to existing manual-driven survey and maintenance operations.

In specific conditions, manual inspections could be totally or partially replaced by Unmanned Autonomous Vehicles (UAVs), Remotely Operated Vehicles (ROVs), magnetic crawlers and any other technological / functional solutions approved by Classification societies.

Noticeably, RITs have already been approved by several flag State Administrations on a case-by-case basis. National flag State authorities, Classification societies and ship owners are slowly but steadily adapting to RIT-based alternatives.

UAVs can perform general visual inspections, ultrasonic thickness measurements and close-up surveys on ships as part of statutory and/or Classification surveys. ROVs are tethered, maneuverable underwater robots that can perform tasks underwater, without the need for divers. Magnetic crawlers can also conduct underwater measurements, scanning hull plates. Some crawlers are also capable of hull cleaning.

The EU landscape of robotics for vessel structure and hull inspection and maintenance requires further insights into key elements towards a harmonized regulatory blueprint, which could serve as a foundation for the anticipated international stand-alone *Guideline for end-users*, bridging all potential gaps through cooperation-based strategic techno-regulatory governance founded on critical safety, security, quality, performance, reliability and efficiency standards.

The United Nations (UN) Climate Change Conference of the Parties (COP26) in Glasgow (2021) also stressed the need to mitigate biofouling build-up, to mitigate GHG emissions,

Anyway, any maritime regulatory development is to be tied to International Maritime Organization (IMO) Strategic Directions (SDs):

- (SD 1) aiming at the efficient and consistent implementation and enforcement of the provisions of the IMO instruments
- (SD 2) aiming at integrating and advancing technologies in the regulatory framework
- (SD 3) intending to respond to climate change by reducing greenhouse emissions
- (SD6) addressing human element related issues including consideration of new technologies and human-centred design
- (SD7) ensuring regulatory effectiveness in the development of advancing technologies, IMO 2022, Resolution A.1149 (32).

The IMO International Conventions are subjected to continuous amendments. The introduction of risk assessment techniques, as part of formal safety assessments and goal-based standards methodologies, represents a new regime. This might even result in a decision to carry out surveys depending on risk profiles as part of an extended Condition Assessment Program (CAP), aimed at assigning ratings based on the condition of a ship, independently of its classification.

In this context, it is acknowledged that novel data detection methods, machine learning modelling techniques and new technologies to diagnose hull and propeller fouling would enable a better asset



management, giving the ship owners and operators methods to predict the hull condition and to suggest the best time for hull maintenance.

The IMO Resolution MEPC.365(79) 2022 “*Guidelines on survey and certification of the Energy Efficiency Design Index (EEDI)*” contains the provisions for the estimation and calculation of resistance of the full-scale ship and propulsion power (speed and power curves) as part of the procedure for the survey and certification of the attained EEDI, in representative sea, wind and weather conditions. According to the paragraph 4.2.6, the preliminary EEDI verification “...*should focus on the calculation process of the attained EEDI to ensure that it is technically sound and reasonable and follows regulation 22 of MARPOL Annex VI and the EEDI Calculation Guidelines*”.

The following two Notes apply:

“Note 1: A possible way forward for more robust verification is to establish a standard methodology of deriving the ship speed from the outcome of tank tests, by setting standard values for experience-based correction factors such as roughness coefficient and wake scaling coefficient. In this way, ship-by-ship performance comparisons could be made more objectively by excluding the possibility of arbitrary setting of experience-based parameters. If such standardization is sought, this would have an implication on how the ship speed adjustment based on sea trial results should be conducted, in accordance with paragraph 4.3.8 of these guidelines.

Note 2: A joint industry standard to support the method and role of the verifier is expected to be developed”.

The shortcoming of this IMO Resolution is the lack of prescriptive indications on the hull conditions, which are implicitly addressed by the vague reference to the “*roughness coefficient*” in the correlation between model tests and sea trials, which in turn is affected by the actual hull roughness, effects of butt-weld seams and joints, coating and cleaning. Note 2 opens the way to a more comprehensive definition of these aspects, encompassed into a new industry standard.

From a Classification society perspective, RINA introduced in its Rules for the Classification of Ships, Part F, Chapter 12, the provisions for Planned Maintenance Systems and Condition Based Maintenance. While these provisions are intended to define the conditions to issue a corresponding Additional Class Notation (PMS/CBM) and in its current edition they are limited to propulsion systems, they may evolve to include new guidelines for ship lifecycle hull and propeller cleaning.

Current underwater inspections, carried out with the support of Unmanned Autonomous Vehicles (UAVs) and Remotely Operated Vehicles (ROVs), would allow the Class Surveyor to achieve a judgement/view on the hull inspection equivalent to the same inspection attended on board (RINA MNO 147, 2019). The advantages of this service are:

- Management of simple, but critical, situations caused by logistic problems linked to the position/port where the ship is
- Shorter response times for the request and execution of the surveys with more efficient schedule
- Reduction of travel or other local costs related to the survey
- Lower operational downtime
- Faster accessibility and examination of the initial condition of assessment
- Data record tracking and condition comparison with past maintenance records
- Data sharing with multiple recipients/legal entities in real time





- Development of archives to maintain the data, which can be used for research purposes (by shipyards, classification societies' technical teams, etc.)
- Flag state acceptance" in case of statutory Surveys.

RITs for underwater inspection/in-water surveys are broadly covered by Class additional notations, which usually include:

- Detailed plans of the hull and hull appendages (all shell openings, stem, rudder and fittings, sternpost, fins, propeller with identification of each blade, anodes and securing arrangements, bilge keels, welded seams, and butts)
- Rudder arrangement
- Tail-shaft arrangement
- Identification marks on the hull to facilitate the in-water survey (in particular, the positions of transverse watertight bulkheads)
- Full photographic documentations
- Thickness readings, close-up and non-destructive testing.

Underwater surveys require extreme care in case of corroded structures or with coating in poor condition. The practice of taking thickness readings in conjunction with close-up and hull inspection is often delegated to service providers authorized either by the Flag Administration or the Classification society, in compliance with the International Association of Classification Societies' (IACS) unified requirement (UR) Z17.

In general, the outside of the ship's hull and related items are to be examined on two occasions in the five-year period of the certificate of Class, with a maximum of 36 months between surveys.

Some minimum standards of RIT have been specified in section 1.1 of IACS Recommendation 42 "Remote inspection techniques may include the use of: Divers; Unmanned robot arm; Remote Operated Vehicles (ROV); Climbers; Drones; Other means acceptable to the Society".

To ensure uniform guidance by any Classification societies, IACS also developed UR Z29 "Remote Classification Surveys" (entered into force on 1 January 2023), to move forward towards class certification and full Administration acceptance.

High-definition cameras, artificial lighting, high precision tracking sensors and 3D scene reconstruction models are essential elements for underwater hull condition monitoring. High definition/high quality data play an important role for detecting the hull conditions as well as any structural defect. Digital data such as photos, live-stream and recorded videos are the predominant outcomes of such inspections. In this process, "metadata" could also be generated, which include time/date stamps, GPS location, camera orientation, focal length, shutter speed, aperture setting, ISO level, camera type, lens type, etc.

It is useful to remind that it is the owner's responsibility to properly maintain the ship in the period between surveys (remote or in-person), and it is the duty of the owner, or their representative, to inform the Class Society of any events or circumstances that may affect the continued conformance of the ship with Class Rules.

IMO has recently embarked on the development of guidance for assessments and applications of remote surveys, ISM Code audits and ISPS Code verifications, with 2024 as the target completion year. This may likely result in amendments to current instruments such as Survey Guidelines under the Harmonized System of Survey and Certification (HSSC), 2019, Resolution A.1140(31), or





guidelines to other security-related instruments, where appropriate, with reference to IACS rules and requirements ([38], [39]) to streamline the usage of remote inspection techniques and establish a strong foundation for this kind of surveys. In parallel, policymakers could consider developing and harmonizing existing flag state-initiated practices, given that all IMO rules and requirements concerning survey/inspection are aimed at flag States that can then delegate responsibilities to Classification societies as Recognized Organization (RO).

Fragmentation in methodologies must be carefully avoided. Uniformity contributes to certainty, that in turn is an acknowledgment that technology-policy interface developments can keep pace with innovation.



8 Concluding remarks

In conclusion, the study successfully developed a hull and propeller condition monitoring tool for the MV Kastor ship using operational data collected over a period of two-and-a-half years. The analysis utilized high frequency data and noon reports to study the distributions of key operational parameters and determine the vessel loading conditions and operating profile. This information formed the basis for detailed simulations aimed at a hydrodynamic optimization. A preliminary assessment of the vessel hydrodynamic condition was conducted using a suitable KPI, which allowed for the identification of performance degradation due to fouling and the impact of maintenance events.

Furthermore, the study implemented a data-driven approach to predict the SHP in realistic operational conditions. Following feature engineering and data preparation, various Machine Learning algorithms were evaluated, and an extensive exploration of hyperparameters was performed to identify the model with the best generalization capability. As a result, a MAPE of 1.1% was achieved on previously unseen test data. A model trained on data collected shortly after maintenance was applied to data collected at a later stage after maintenance. The PI, defined as the normalized difference between the actual and predicted SHP, was utilized to describe the deterioration of the vessel condition. In a separate approach, the feature Days was incorporated to indicate the elapsed time since the previous cleaning event. A model incorporating this feature was trained using a dataset spanning over a year, and it was then applied to synthetic datasets representing both laden and ballast conditions across a range of vessel speeds, with various values of Days. By comparing the predictions made for different Days values, while keeping the other features constant, the PI resulting from biofouling was derived.

These data-driven approaches provide an estimation of the PI caused by marine biofouling growth within a specific timeframe. By incorporating SFOC and fuel cost information, the subsequent daily increase in fuel cost can be determined. As a result, it is estimated that over a span of approximately 10 months, biofouling leads to an average daily fuel cost increase of \$2,249. Shipping companies can compare the cumulative additional fuel cost with the cost of maintenance, considering the vessel schedule, to strategically determine the optimal timing to perform a vessel cleaning. Hence, by providing valuable insights and practical implications, the findings of this study enable cost-effective decision-making within the shipping industry. These insights support the development of more informed and efficient maintenance strategies, ultimately leading to improved operational efficiency.

It is also essential that the current regulatory framework is improved to bridge the existing gaps, foster innovation, and take advantage of remote inspections and digital technologies to obtain high quality data from underwater hull monitoring activities.

References

- [1] IMO (2018), Resolution MEPC.304 (72), Initial IMO Strategy on Reduction of GHG Emissions from Ships, International Maritime Organisation, London, UK.
- [2] IMO (2023), Resolution MEPC.377(80), 2023 IMO Strategy on Reduction of GHG Emissions from Ships, International Maritime Organization, London, UK.
- [3] IMO (2008), International Convention on the Control of Harmful Anti-fouling Systems on Ships, International Maritime Organization, London, UK.
- [4] Barreiro, J., Zaragoza, S., & Diaz-Casas, V. (2022), Review of ship energy efficiency, *Ocean Engineering*, 257, Vol. 257, 111594, <https://doi.org/10.1016/j.oceaneng.2022.111594>.
- [5] Brynolf, S., Baldi, Fr., Johnson, H. (2016), Energy efficiency and fuel changes to reduce environmental impacts, In: Andersson, K., Brynolf, S., et al., (Eds.). *Shipping and the environment: improving environmental performance in marine transportation*, Springer, Berlin, Heidelberg, 295–339. http://dx.doi.org/10.1007/978-3-662-49045-7_10.
- [6] Bouman, E. A., Lindstad, E., Rialland, A. I., Stramman, A. H. (2017), State-of-the-art technologies, measures, and potential for reducing GHG emissions from shipping – A review, *Transportation Research Part D: Transport and Environment*, 52, 408-421, <http://dx.doi.org/10.1016/j.trd.2017.03.022>.
- [7] Yusim, A. K., Utama, I. K. A. P. (2017), An Investigation Into The Drag Increase on Roughen Surface due to Marine Fouling Growth, *Journal for Technology and Science*, Vol. 28 (3).
- [8] Carhen, Al., Altar, M. (2020), Four KPIs for the assessment of biofouling effect on ship performance, *Ocean Engineering*, Vol. 217, <https://doi.org/10.1016/j.oceaneng.2020.107971>.
- [9] Arndt, E., Robinson, A., Hester, S., Woodham, B., Wilkinson, P., Gorgula, S., & Brooks, B. (2021), Factors that influence vessel biofouling and its prevention and management Final report for CEBRA Project 190803.
- [10] ISO (2016), ISO 19030: Ships and Marine Technology- Measurement of Changes in Hull and Propeller Performance, International Organization of Standardization, Geneva.
- [11] Karagiannidis, P., Themelis, N. (2021), Data-driven modelling of ship propulsion and the effect of data pre-processing on the prediction of ship fuel consumption and speed loss, *Ocean Engineering*, 222:108616, <https://doi.org/10.1016/j.oceaneng.2021.108616>.
- [12] López, D., Vlamakis, H., Kolter, R. (2010). "Biofilms". *Cold Spring Harbor Perspectives in Biology*. 2 (7).
- [13] Candries, M., Altar, M., Anderson, C.D. (2003), Estimating the impact of new-generation antifouling on ship performance: the presence of slime, *Journal of Marine Engineering & Technology*, Vol.2.
- [14] Bressy, C., Lejars, M. (2014), Marine Fouling: An Overview, *Journal of Ocean Technology*, Vol.9.
- [15] Valchev, I., Coraddu, A., Kalikatzarakis, M., Geertsma, R., Oneto, L. (2022), Numerical methods for monitoring and evaluating the biofouling state and effects on vessels' hull and propeller performance: A review, *Ocean Engineering*, Vol. 251, 110883, <https://doi.org/10.1016/j.oceaneng.2022.110883>.
- [16] Uzun, D., Demirel, Y.K., Coraddu, A., Turan, O. (2019), Time-dependent biofouling growth model for predicting the effects of biofouling on ship resistance and powering, *Ocean Engineering*, Vol. 191, 106432, <https://doi.org/10.1016/j.oceaneng.2019.106432>.
- [17] Earle, S. (2019). *Physical Geology – 2nd Edition*. Victoria, B.C.: BCcampus. Retrieved from <https://opentextbc.ca/physicalgeology2ed/>.
- [18] Schultz ,M.,P. (2007), Effects of coating roughness and biofouling on ship resistance and powering, *Biofouling* 23 (5), 331–341, <https://doi.org/10.1080/08927010701461974>.
- [19] Munk, Y., Kane, D., Yebra, D.M. (2009), The effects of corrosion and fouling on the performance of ocean-going vessels: a naval architectural perspective, Editor(s): Claire

- Hellio, Diego Yebra, In Woodhead Publishing Series in Metals and Surface Engineering, Advances in Marine Antifouling Coatings and Technologies, Woodhead Publishing, 2009, Pages 148-176.
- [20] Ronen, D. (2011) The Effect of Oil Price on Containership Speed and Fleet Size. *Journal of Operational Research Society*, 62, 211-216.
- [21] Song, C.; Cui, W. (2020), Review of Underwater ship hull cleaning technologies, *Journal of Marine Science and Application*, 19, 415-429.
- [22] Almeida, E., Diamantino, T. C., Sousa, O. (2007), Marine paints: The particular case of antifouling paints, *Progress in Organic Coatings*, 59, 2-20.
- [23] Yebra, D. M., Kiil, S., Dam-Johansen, K. (2003), Antifouling technology – past, present and future steps towards efficient and environmentally friendly antifouling coatings, *Progress in Organic Coatings*, 50, 75-104.
- [24] Oliveira, D. R.; Granhag, L. (2020), Ship hull in-water cleaning and its effects on fouling-control coatings, *Biofouling*, 36:3, 332-350.
- [25] Dafforn, K. A., Lewis, J. A., Johnston, E. L. (2011), Antifouling strategies: History and regulation, ecological impacts and mitigation, *Marine pollution Bulletin*, 62, 453-465.
- [26] IMO (2016), Module 4 – Ship Board Energy Management, International Maritime Organization London, UK.
- [27] Gao, J., Zhang, K., Li, H., Lang, C., Zhang, L. (2023), Exo-friendly intrinsic self-healing superhydrophobic polyurea/TiO₂ composite coatings for underwater drag reduction and antifouling, *Progress in Organic Coatings*, 183.
- [28] Sharma, A.; Sharma, S. (2023), Graphene-based polymer coatings for preventing marine corrosion: a review, *Journal of Coatings Technology and Research*, 20, 413-432.
- [29] Kim, Y. R., Jung, M., & Park, J. B. (2021), Development of a fuel consumption prediction model based on machine learning using ship in-service data, *Journal of Marine Science and Engineering*, 9(2), 1–25. <https://doi.org/10.3390/jmse9020137>.
- [30] Lang, X., Wu, D., & Mao, W. (2022), Comparison of supervised machine learning methods to predict ship propulsion power at sea, *Ocean Engineering*, 245, <https://doi.org/10.1016/j.oceaneng.2021.110387>.
- [31] Uyanık, T., Karatuğ, Ç., Arslanoğlu, Y. (2020), Machine learning approach to ship fuel consumption: A case of container vessel. *Transportation Research Part D: Transport and Environment*, <https://doi.org/10.1016/j.trd.2020.102389>.
- [32] Petersen, J.P., Winther, O., Jacobsen, D.J. (2012), A Machine-Learning Approach to Predict Main Energy Consumption under Realistic Operational Conditions, *Ship Technology Research*, 59 (1).
- [33] Farag, Y., B., Ölçer Al. (2020), The development of a ship performance model in varying operating conditions based on ANN and regression techniques. *Ocean Engineering*, 198:106972, <https://doi.org/10.1016/j.oceaneng.2020.106972>.
- [34] Gkerekos, C., Lazakis, I. (2020), A novel, data-driven heuristic framework for vessel weather routing, *Ocean Engineering*, 197:106887, <https://doi.org/10.1016/j.oceaneng.2019.106887>.
- [35] Nikolaidis, G., Themelis, N. (2022), Examining the performance of retrofit measures in real ship operation using data-driven models, *Ship Technology Research*, 69:3, 170-180, <https://doi.org/10.1080/09377255.2022.2109327>.
- [36] Laurie A, Anderlini E, Dietz J, Thomas G. (2021). Machine learning for shaft power prediction and analysis of fouling related performance deterioration. *Ocean Engineering*. 234:108886.
- [37] Székely, G.J., Rizzo, L.M., Bakirov, N.K. (2007), Measuring and testing dependence by correlation of distances., *The Annals of Statistics*, vol. 35, no. 6, 2007, pp. 2769-2794.
- [38] IACS, (2016), Recommendation No. 42: Guidelines for use of remote inspection techniques for surveys, International Association of Classification Societies.



- [39] IACS, (2022), Unified Requirement on Remote Classification Surveys UR Z29, International Association of Classification Societies.
- [40] Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A. (1984), Classification and Regression Trees, Chapman and Hall.
- [41] Breiman, L. (2001), Random Forests. Machine Learning 45, 5–32.
- [42] Geurts, P., Ernst, D., Wehenkel, L. (2006). Extremely randomized trees. Machine learning, 63(1), 3-42.
- [43] Friedman, J.H. (2001), Greedy function approximation: A gradient boosting machine, Ann. Statist. 29 (5) 1189 – 1232.
- [44] Staudemeyer, R. C., Morris, E. R. (2019), Understanding LSTM--a tutorial into long short-term memory recurrent neural networks, arXiv preprint arXiv:1909.09586.
- [45] Zhang, M., Kujala, P., & Hirdaris, D. (2023), A deep learning method for the prediction of ship fuel consumption in real operational conditions, Engineering applications of artificial intelligence. (Under review).



Appendix A: Decision Trees

Decision Trees are powerful and interpretable models used for both classification and regression tasks. In this report, we focus on Decision Trees for regression, specifically those that employ the CART (Classification and Regression Trees) algorithm [40].

A Decision Tree for regression is a predictive model that uses a tree-like structure to make predictions on continuous numerical target variables. The CART algorithm constructs the Decision Tree by recursively partitioning the feature space into smaller regions, with the goal of minimizing the sum of squared errors (SSE) within each partition. The CART algorithm starts with the entire dataset and selects a feature v to split the data. It evaluates different split points for the selected feature and chooses the one that results in the lowest SSE. The splitting process divides the dataset into two subsets A_1 and A_2 based on the feature value f_v , creating left and right child nodes. The cost function minimized with the split is the following:

$$J(v, f_v) = \frac{k_1}{k} SSE_1 + \frac{k_2}{k} SSE_2 \quad (\text{A.1})$$

where k , k_1 , and k_2 are the numbers of instances in the original dataset and in subsets A_1 and A_2 respectively. Thus, the SSE is given as:

$$SSE_j = \sum_{i \in A_j} (y_i - \bar{y}_j)^2, \text{ for } j = 1, 2 \quad (\text{A.2})$$

$$\bar{y}_j = \frac{1}{k_j} \sum_{i \in A_j} y_i, \text{ for } j = 1, 2 \quad (\text{A.3})$$

The splitting process is applied recursively to each child node until a stopping criterion is met. This criterion can be a maximum tree depth, a minimum number of samples per leaf node, or a minimum reduction in SSE. The recursive nature of the algorithm allows to capture complex relationships between features and the target variable. When the recursion stops, the resulting tree contains terminal nodes called leaf nodes. Each leaf node represents a prediction for the target variable in the corresponding region of the feature space. In case of regression, the prediction is typically the mean or median value of the target variable within that region. The predictions from the leaf nodes form the final output of the Decision Tree model.

Decision Trees offer several advantages when used for regression analysis. First and foremost, they provide interpretability, allowing analysts to gain a clear and intuitive understanding of the decision-making process. By following the path from the root node to a leaf node, one can discern how each input feature contributes to the predictions, enabling valuable insights. Furthermore, Decision Trees excel at capturing nonlinear relationships between features and the target variable. This flexibility makes Decision Trees particularly suitable for datasets with intricate and nonlinear patterns.

However, Decision Trees have certain disadvantages that should be addressed. Intricate trees are often produced that are overly complex and struggle to generalize well to new data. This phenomenon is known as overfitting. To avoid overfitting, it is crucial to employ various techniques, such as specifying a minimum number of samples required at a leaf node or setting a maximum depth for the tree. Another challenge with Decision Trees is their susceptibility to instability caused by minor variations in the input data. Even slight changes can lead to the creation of entirely different



trees. To tackle this issue, Decision Trees are frequently used as part of an ensemble, where multiple trees are combined to make more robust predictions. By leveraging the collective wisdom of multiple trees, the instability problem can be mitigated, resulting in more reliable and accurate predictions.

Typically, the “max_depth” and “min_samples_leaf” hyperparameters are tuned when creating a Decision Tree model. Specifically, the “max_depth” hyperparameter determines the maximum depth or levels of the Decision Tree. It controls how many splits or branches the tree can have from the root node to the leaf nodes. A higher value can result in a more complex tree that can potentially overfit the training data, while a lower value can lead to underfitting. On the other hand, the “min_samples_leaf” hyperparameter sets the minimum number of samples required to be in a leaf node. If the number of samples in a potential leaf node is less than the specified value, further splitting is halted, and the node becomes a leaf. A higher value can prevent the tree from creating very specific rules for a small subset of data, reducing overfitting. These hyperparameters play a crucial role in controlling the complexity and generalization ability of a Decision Tree model for regression tasks.



Appendix B: Random Forests

Random Forests for regression combine the strength of Decision Trees with ensemble techniques [41]. Instead of relying on a single Decision Tree, random Forests aggregate predictions from multiple trees, providing more robust and accurate regression results. The CART algorithm is employed to construct individual Decision Trees within the Random Forest.

Random Forests are constructed through the following steps. First, the algorithm begins by bootstrapping, where subsets of the original training data are randomly selected with replacement. These subsets, called bootstrap samples, have the same size as the original dataset but can contain duplicate instances. Next, for each bootstrap sample, a Decision Tree is constructed using the CART algorithm. To introduce diversity among the trees in the ensemble, feature randomness is incorporated. At each split during the Decision Tree construction, a random subset of features is considered. Finally, once all the trees are built, predictions are made by aggregating the individual tree predictions. This is achieved by calculating the mean of the predicted values from each tree. The "n_estimators" hyperparameter refers to the number of individual Decision Trees that are created and combined in the ensemble.

Random Forests provide several advantages for regression tasks. Firstly, they offer improved accuracy by aggregating predictions from multiple trees. This aggregation process helps to reduce overfitting and enhance the ability of the model to generalize well to unseen data. Moreover, the ensemble nature of Random Forests enables them to capture complex relationships between features and the target variable, leading to more accurate predictions. Secondly, Random Forests exhibit robustness to outliers. Due to the averaging of predictions from multiple trees, outliers have a limited impact on the final predictions. This property makes Random Forests more reliable when dealing with noisy or skewed data, since the presence of outliers does not heavily distort the overall predictions.

Appendix C: Extremely Randomized Trees

Extremely Randomized Trees build upon the idea of Random Forests by introducing additional randomness during the tree construction process [42]. This randomness leads to further diversity among the trees in the ensemble, which can enhance their predictive performance. The CART algorithm is utilized to construct individual Decision Trees within Extremely Randomized Trees.

The construction of Extremely Randomized Trees involves several steps. Firstly, bootstrapping is employed, similar to Random Forests, to create multiple bootstrap samples from the original training data. Each bootstrap sample is generated by randomly selecting instances with replacement. Secondly, feature randomness is introduced in Extremely Randomized Trees. At each split point during the tree construction process, thresholds for each feature are randomly selected without evaluating the optimal split points. This random selection of thresholds adds more randomness and reduces correlation between the trees in the ensemble. Thirdly, individual Decision Trees are constructed using the CART algorithm. The feature space is recursively partitioned, and each tree is grown by selecting random thresholds for each feature and choosing the split that minimizes the SSE within each leaf node. Finally, after constructing all the trees, predictions are made by taking the mean of the predicted values from each tree. The hyperparameter "n_estimators" specifies the number of individual decision trees that are generated and aggregated as an ensemble.

Extremely Randomized Trees offer numerous benefits when applied to regression tasks. Firstly, they offer improved diversity among the ensemble of trees. The additional randomness introduced during the tree construction process enhances the diversity among the trees, which can help reduce overfitting and improve the generalization performance of the model. Secondly, Extremely Randomized Trees tend to have lower variance compared to traditional random forests. By randomly selecting split points without evaluating optimal splits, they are less influenced by noise or outliers in the data, resulting in more robust predictions. Lastly, Extremely Randomized Trees are efficient in terms of computational cost. The random selection of thresholds reduces the need for exhaustive search for optimal splits, making the training process faster compared to other ensemble methods.

Appendix D: Gradient Tree Boosting

Gradient Tree Boosting, or Gradient Boosted Regression Trees (GBRT), is an ensemble learning method that combines the predictive power of Decision Trees with the concept of boosting [43]. It builds an ensemble of weak prediction models, typically Decision Trees, in a stage-wise manner. Based on the principle of gradient boosting, each subsequent model tries to correct the residuals (errors) of the previous models.

The construction of GBRT entails a series of distinct steps. Firstly, the predictions are initialized by using a selected value, such as the mean of the target variable. This value acts as the starting point for subsequent iterations. An iterative process then begins: the difference between the actual target values and the prediction from the previous iteration is calculated, resulting in residuals that represent the errors to be rectified by the current model. Following that, for each iteration, a new Decision Tree is built using the CART algorithm. After constructing the tree, the predictions from the tree are multiplied by a learning rate, which governs the contribution of each tree to the final prediction. These predictions are then aggregated to update the overall prediction of the GBRT. This iterative process continues with each iteration aiming to reduce the remaining residuals. Multiple Decision Trees are added to the ensemble, and their predictions are combined with the previous predictions to progressively refine the overall model.

GBRT is renowned for its high predictive accuracy. Through iterative improvement based on residuals, GBRT can effectively capture complex relationships and deliver superior performance compared to individual Decision Trees. However, GBRT can be prone to overfitting, especially when the dataset is small or noisy. The iterative nature of GBRT may lead to the model memorizing the training data, resulting in poor generalization to new, unseen data. To prevent overfitting, it is important to find an optimal value for the 'n_estimators' hyperparameter, which determines the number of Trees in the ensemble.

Appendix E: Distance Correlation

Distance correlation, as defined by [37], is a measure that captures both linear and nonlinear associations between variables in regression analysis. While traditional correlation measures such as Pearson's correlation coefficient are effective for capturing linear relationships, they often fail to detect complex dependencies that may exist in the data. Distance correlation overcomes this limitation by considering the distances between observations rather than assuming a specific functional form. It provides a robust measure of dependence that can reveal intricate relationships that may be missed by traditional methods.

The concept of distance correlation is based on the notions of distance covariance and distance variance. The distance covariance quantifies the similarity of paired observations in terms of their distances to other observations, while the distance variance measures the dispersion of the distance between paired observations. By normalizing the distance covariance with the square root of the product of the distance variances, distance correlation is obtained. The resulting measure ranges between 0 and 1, with 0 indicating independence and 1 indicating perfect dependence. This makes distance correlation a valuable tool for feature selection, as it captures a wide range of relationships and provides a comprehensive understanding of the data.

Specifically, the distance covariance between two random variables X and Y is defined as the square root of their distance variance:

$$dCov(X, Y) = \sqrt{dVar(X, Y)} \quad (E.1)$$

The distance variance between X and Y measures the dispersion of the distances between paired observations. It is calculated as follows:

$$dVar(X, Y) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n S(X_i, X_j) S(Y_i, Y_j) \quad (E.2)$$

where $S(X_i, X_j)$ and $S(Y_i, Y_j)$ represent the Euclidean distances between observations X_i and X_j , and between Y_i and Y_j , respectively, and n represents the sample size.

The distance correlation between X and Y is obtained by normalizing the distance covariance with the square root of the product of the distance variances:

$$dCor(X, Y) = \frac{dCov(X, Y)}{\sqrt{dVar(X) dVar(Y)}} \quad (E.3)$$

where $dVar(X)$ and $dVar(Y)$ represent the distance variances of X and Y , respectively.

Appendix F: List of high-frequency parameters selected for data analysis

	Parameter name
1	TIME
2	TIME_NUM
3	Longitudinal Water Speed_BRG_SLOG
4	Total Cumulative Water Distance_BRG_SLOG
5	Water Distance Since Reset_BRG_SLOG
6	Water Depth Relative to the Transducer_BRG_ECHO
7	Offset from Transducer_BRG_ECHO
8	Maximun Range Scale in Use_BRG_ECHO
9	Wind Angle_BRG_WIND
10	Wind Speed_BRG_WIND
11	UTC_BRG_GPS_
12	Latitude_BRG_GPS_
13	Longitude_BRG_GPS_
14	Speed Over Ground_BRG_GPS_
15	True Course Over Ground_BRG_GPS_
16	Date_BRG_GPS
17	Magnetic Variation_degrees E/W_BRG_GPS
18	Latitude_BRG_GPS
19	Longitude_BRG_GPS
20	UTC_BRG_GPS
21	True Course Over Ground_BRG_GPS
22	Magnetic Course Over Ground_BRG_GPS
23	Speed Over Ground_knots_BRG_GPS
24	Water Depth_meters_BRG_ECHO
25	True Heading_BRG_GYRO
26	Rate of Turn_BRG_GYRO
27	Time_BRG_GPS
28	Day_BRG_GPS
29	Month_BRG_GPS
30	Year_BRG_GPS
31	DG1 Power Output_DGP
32	DG2 Power Output_DGP
33	DG3 Power Output_DGP
34	DG1 Power Output_mA_DGP
35	DG2 Power Output_mA_DGP
36	DG3 Power Output_mA_DGP

37	M/E T/C RPM_IND1
38	M/E Scavenge Air Pressure_IND1
39	M/E T/C RPM_mA_IND1
40	M/E Scavenge Air Pressure_mA_IND1
41	Fuel Oil Volumetric Flow (ME supply) TRQM
42	Fuel Oil Temperature (ME supply) TRQM
43	Fuel Oil Volumetric Flow (ME return) TRQM
44	Fuel Oil Temperature (ME return) TRQM
45	Fuel Oil Volumetric Flow (GE supply) TRQM
46	Fuel Oil Temperature (GE supply) TRQM
47	Fuel Oil Volumetric Flow (GE return) TRQM
48	Fuel Oil Temperature (GE return) TRQM
49	M/E Shaft RPM_TRQM
50	Shaft Torque_TRQM
51	Shaft Thrust_TRQM
52	Shaft Power_TRQM
53	Speed Over Ground_TRQM
54	Speed Through Water_TRQM
55	ME Consumption_TRQM
56	DG Consumption_TRQM
57	No 1 ballast pump running_AMS
58	No 2 ballast pump running_AMS
59	Fore draft_AMS
60	Aft draft_AMS
61	Middle draft(P)_AMS
62	Middle draft(S)_AMS
63	No 1 HFO stor tk(P) temp_AMS
64	No 1 HFO stor tk(S) temp_AMS
65	No 2 HFO stor tk(P) temp_AMS
66	No 2 HFO stor tk(S) temp_AMS
67	No 3 LS HFO stor tk(S) temp_AMS
68	No 1 LS HFO sett tk temp_AMS
69	No 2 HFO sett tk temp_AMS
70	No 1 LS HFO serv tk temp_AMS
71	No 2 HFO serv tk temp_AMS
72	Trim_AMS
73	List_AMS
74	FP tk volume_AMS
75	AP tk volume_AMS

76	No 1 WB tk(P) volume_AMS
77	No 1 WB tk(S) volume_AMS
78	No 2 WB tk(P) volume_AMS
79	No 2 WB tk(S) volume_AMS
80	No 3&4 WB tk(P) volume_AMS
81	No 3&4 WB tk(S) volume_AMS
82	No 5&6 WB tk(P) volume_AMS
83	No 5&6 WB tk(S) volume_AMS
84	No 7 WB tk(P) volume_AMS
85	No 7 WB tk(S) volume_AMS
86	No 1 HFO stor tk(P) volume_AMS
87	No 1 HFO stor tk(S) volume_AMS
88	No 2 HFO stor tk(P) volume_AMS
89	No 2 HFO stor tk(S) volume_AMS
90	No 3 LS HFO stor tk(S) volume_AMS
91	No 1 LS HFO sett tk volume_AMS
92	No 2 HFO sett tk volume_AMS
93	No 1 LS HFO serv tk volume_AMS
94	No 2 HFO serv tk volume_AMS
95	MDO stor tk volume_AMS
96	LS MGO stor tk volume_AMS
97	MDO serv tk volume_AMS
98	LS MGO serv tk volume_AMS
99	Fresh water tank(P) volume_AMS
100	Fresh water tank(S) volume_AMS
101	Cyl 01 fore main bearing temp_AMS
102	Cyl 01 AFT main bearing temp_AMS
103	Cyl 02 AFT main bearing temp_AMS
104	Cyl 03 AFT main bearing temp_AMS
105	Cyl 04 AFT main bearing temp_AMS
106	Cyl 05 AFT main bearing temp_AMS
107	Cyl 06 AFT main bearing temp_AMS
108	Cyl 01 crank pin bearing temp_AMS
109	Cyl 02 crank pin bearing temp_AMS
110	Cyl 03 crank pin bearing temp_AMS
111	Cyl 04 crank pin bearing temp_AMS
112	Cyl 05 crank pin bearing temp_AMS
113	Cyl 06 crank pin bearing temp_AMS
114	Thrust main bearing temp_AMS

115	ME main LO inlet press_AMS
116	ME FO inlet press_AMS
117	ME JCW inlet press_AMS
118	ME scav air receiver inlet pres_AMS
119	ME fuel index_AMS
120	ME TC speed_AMS
121	GE1 FO inlet press_AMS
122	GE1 start air press_AMS
123	GE1 HT FW inlet press_AMS
124	GE1 LO inlet press_AMS
125	GE1 LT FW inlet press_AMS
126	GE1 HT FW outlet temp_AMS
127	GE1 LO inlet temp_AMS
128	GE1 cyl 1_3 exh gas TC inlet tem_AMS
129	GE1 cyl 4_6 exh gas TC inlet tem_AMS
130	GE1 bearing temp_AMS
131	GE2 FO inlet press_AMS
132	GE2 start air press_AMS
133	GE2 HT FW inlet press_AMS
134	GE2 LO inlet press_AMS
135	GE2 LT FW inlet press_AMS
136	GE2 HT FW outlet temp_AMS
137	GE2 LO inlet temp_AMS
138	GE2 cyl 1_3 exh gas TC inlet tem_AMS
139	GE2 cyl 4_6 exh gas TC inlet tem_AMS
140	GE2 bearing temp_AMS
141	GE3 bearing temp_AMS
142	GE3 FO inlet press_AMS
143	GE3 start air press_AMS
144	GE3 HT FW inlet press_AMS
145	GE3 LO inlet press_AMS
146	GE3 LT FW inlet press_AMS
147	GE3 HT FW outlet temp_AMS
148	GE3 LO inlet temp_AMS
149	GE3 cyl 1_3 exh gas TC inlet tem_AMS
150	GE3 cyl 4_6 exh gas TC inlet tem_AMS
151	ME TC LO outlet temp_AMS
152	ME main LO inlet temp_AMS
153	ME cyl lub oil temp_AMS

154	ME FO inlet temp_AMS
155	ME JCW inlet temp_AMS
156	ME cyl 1 exh gas outlet temp_AMS
157	ME cyl 2 exh gas outlet temp_AMS
158	ME cyl 3 exh gas outlet temp_AMS
159	ME cyl 4 exh gas outlet temp_AMS
160	ME cyl 5 exh gas outlet temp_AMS
161	ME cyl 6 exh gas outlet temp_AMS
162	ME TC exh gas inlet temp_AMS
163	ME TC exh gas outlet temp_AMS
164	ME cyl 1 JCW outlet temp_AMS
165	ME cyl 2 JCW outlet temp_AMS
166	ME cyl 3 JCW outlet temp_AMS
167	ME cyl 4 JCW outlet temp_AMS
168	ME cyl 5 JCW outlet temp_AMS
169	ME cyl 6 JCW outlet temp_AMS
170	ME JCW outlet press_AMS
171	ME FO inlet press_RED_AMS
172	ME JCW inlet press_RED_AMS
173	ME scav air receiver inlet pres_RED_AMS
174	ME control air inlet press_AMS
175	ME cyl 1 scav air temp_AMS
176	ME cyl 2 scav air temp_AMS
177	ME cyl 3 scav air temp_AMS
178	ME cyl 4 scav air temp_AMS
179	ME cyl 5 scav air temp_AMS
180	ME cyl 6 scav air temp_AMS
181	ME scav air receiver temp_AMS
182	ME air cooler cool w inlet tem_AMS
183	ME air cooler cool w outlet te_AMS
184	ME air cooler cool w inlet pre_AMS
185	ME TC LO inlet press_AMS
186	ME main LO inlet press_RED_AMS
187	ME axial vibration_AMS
188	ME cyl 1 PCO outlet temp_AMS
189	ME cyl 2 PCO outlet temp_AMS
190	ME cyl 3 PCO outlet temp_AMS
191	ME cyl 4 PCO outlet temp_AMS
192	ME cyl 5 PCO outlet temp_AMS

193	ME cyl 6 PCO outlet temp_AMS
194	ME thrust bearing pad temp_AMS
195	Intermediate shaft bearing temp_AMS
196	Stern tube bearing temp_AMS
197	ME RPM_AMS
198	Speed-Over-Ground
199	Speed-Through-Water
200	TC-rpm
201	ME-Scavenge-Air-Press
202	Fuel-Index-Position
203	Propeller-Shaft-Power
204	Exhaust-temp-bef-TC
205	Exhaust-temp-after-TC
206	Rel-Wind-Speed
207	Rel-Wind-Direction
208	Vessel-Heading
209	Propeller-Shaft-RPM
210	DG-FM-Inlet-temp
211	DG-FM-Outlet-temp
212	ME-FM-Inlet-temp
213	DG1-Power
214	DG2-Power
215	DG3-Power
216	DG1-LO-inlet-press
217	DG2-LO-inlet-press
218	DG3-LO-inlet-press
219	DG-Inlet-FO-vol-flow-cons
220	DG-Outlet-FO-vol-flow-cons
221	ME-Inlet-FO-vol-flow-cons
222	ME-Cyl1-Exh-Gas-Out-temp
223	ME-Cyl2-Exh-Gas-Out-temp
224	ME-Cyl3-Exh-Gas-Out-temp
225	ME-Cyl4-Exh-Gas-Out-temp
226	ME-Cyl5-Exh-Gas-Out-temp
227	ME-Cyl6-Exh-Gas-Out-temp
228	Density-15oC-FO
229	Draft-Aft
230	Draft-Fwd
231	Density-15oC-HFO

232	Density-15oC-MGO
233	Draft-Mean
234	Fuel-LCV
235	DG-FO-Cons
236	ME-FO-Cons
237	ME-Cyl1-CFW-Out-temp
238	ME-Cyl2-CFW-Out-temp
239	ME-Cyl3-CFW-Out-temp
240	ME-Cyl4-CFW-Out-temp
241	ME-Cyl5-CFW-Out-temp
242	ME-Cyl6-CFW-Out-temp
243	ME-CFW-Out-temp-avg
244	DTN_AIR_PRESSURE_MEAN_SEA_LEVEL
245	DTN_WINDWAVE_SIGN_WAVE_HEIGHT
246	DTN_WINDWAVE_MAX_WAVE_HEIGHT
247	DTN_WINDWAVE_MEAN_WAVE_DIRECTION
248	DTN_WINDWAVE_MEAN_WAVE_LENGTH
249	DTN_AIR_TEMPERATURE_10M
250	DTN_SEA_TEMPERATURE_0M
251	DTN_WIND_SPEED_10M
252	DTN_WIND_SPEED_GUST_10M
253	DTN_WIND_DIRECTION_10M
254	DTN_SIGNIFICANT_WAVE_HEIGHT
255	DTN_MAXIMUM_WAVE_HEIGHT
256	DTN_MEAN_WAVE_DIRECTION
257	DTN_MEAN_WAVE_LENGTH
258	DTN_NORTHWARD_CURRENT
259	DTN_EASTWARD_CURRENT
260	DTN_SWELL_SIGNIFICANT_WAVE_HEIGHT
261	DTN_SWELL_MAXIMUM_WAVE_HEIGHT
262	DTN_SWELL_MEAN_WAVE_DIRECTION
263	DTN_SWELL_MEAN_WAVE_LENGTH
264	DTN_CURRENT_DIRECTION
265	DTN_CURRENT_SPEED
266	DTN_AIR_PRESSURE_MEAN_SEA_LEVEL_ACTUAL
267	DTN_AIR_TEMPERATURE_10M_ACTUAL
268	DTN_SEA_TEMPERATURE_0M_ACTUAL

Appendix G: An alternative approach based on fuel consumption prediction

The study presented in this section has been carried out by AALTO, to support the operational optimization actions to be carried in the next tasks of the project. In particular, it presents a Deep Learning Method for ship fuel consumption prediction that utilizes big data from sensors, voyage reporting and hydrometeorological data, comprising the 266 variables available for the ship examined in the previous sections. Decision Tree (DT) is used to evaluate feature importance and to select key influencing factor on ship fuel consumption. The ship fuel consumption prediction model is developed by incorporating attention mechanisms into Bidirectional Long Short-Term Memory (Bi-LSTM) networks. Bi-LSTM networks are a type of recurrent neural network (RNN) that can capture long-term dependencies in time series data. They consist of two LSTM layers that process the data in both forward and backward directions, allowing the network to have more context and a better understanding of the data stream. Attention mechanisms enable neural networks to focus on specific parts of the input data when making predictions. The proposed model can be used to capture ship energy system and to predict ship fuel consumption in real operational condition.

The objective of the ship fuel consumption prediction model is to create an intelligent decision support tool to reduce fuel consumption during actual ship operations, as shown in Figure 38. The motivations for this study are:

- ship fuel consumption is influenced by numerous interconnected factors, such as operational conditions, navigation data, engine performance, and weather conditions. The model aims to unravel the complex relationship between these factors and fuel consumption, forming the basis for a multi-objective optimization;
- the model used is data-driven, generative, and regression-based, designed to mimic ship energy systems in real operations. This model is specifically tailored for the ship under examination, providing rapid and precise estimates of fuel consumption via an optimization process;
- the model has the potential to determine optimal ship operation commands and routes that achieve a fuel consumption reduction.

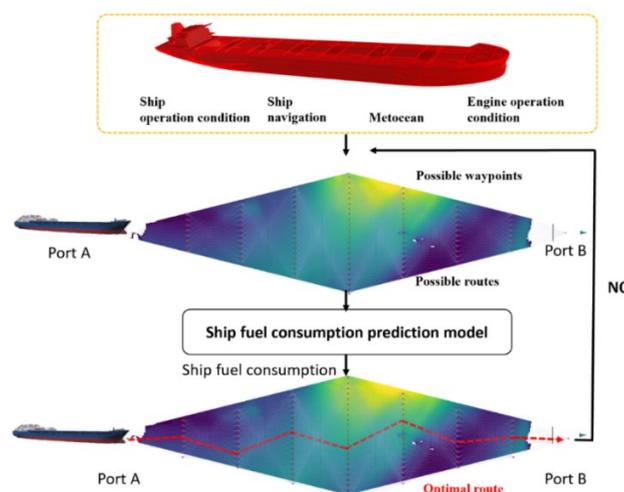


Figure 38: Flowchart of the multi-objective optimization method employed to achieve ship fuel consumption reduction based on the prediction model.

To predict ship fuel consumption in real operations, this section introduces a deep learning model based on Bi-LSTM with attention mechanisms ([44], [45]), as shown in Figure 39. Since the fuel consumption prediction model combines the strengths of Bi-LSTM, which can capture both past and future dependencies, and attention mechanisms, which allow the model to focus on relevant parts of the input data streams, it demonstrates a good performance in terms of effectively capturing and utilizing historical information ([45]).

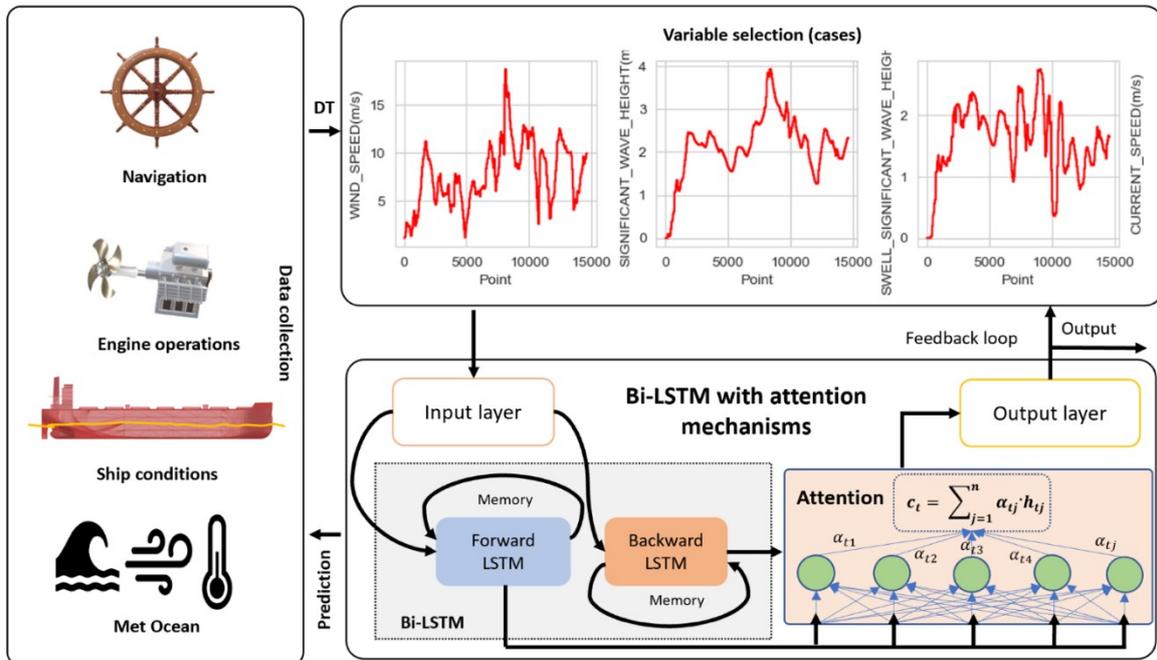


Figure 39: Schematic of the deep learning method architecture for the ship fuel consumption prediction.

To identify the key influencing factors while disregarding irrelevant data, the study employed a DT regression method. This approach facilitated the evaluation of the importance of a comprehensive set of 266 variables (comprising of 265 parameters and ship fuel consumption in the time domain). By utilizing this method, the study aimed to extract meaningful insights and discern the significant predictors that have a substantial impact on ship fuel consumption. Additionally, a DT regression model was employed to analyse the importance of the 265 variables in relation to ship fuel consumption. The 265 variables were designated as the X database, while ship fuel consumption was designated as the Y database. The X and Y database were split into training and testing sets, with the training set accounting for 80% of the available data streams and the testing set for the remaining 20%. Finally, these data were normalized to facilitate the training of the decision tree. The best-performing DT model was determined through a grid-search cross-validation method, representing the decision tree regressor with the hyperparameters that generated the lowest negative Mean Absolute Error (MAE). Consequently, the best decision tree was obtained with $d=90$, $m=4$, and $r=1$, and the decision tree is as shown in **Errore. L'origine riferimento non è stata trovata.**

This study employed an optimized decision tree model alongside a collected dataset to assess the different variable importance. The aim was to identify the key factors affecting ship fuel consumption.

In addition, Figure 41 demonstrates the correlation among these factors and fuel consumption. Notably, strong positive correlations were observed between the propeller shaft RPM, the main engine shaft RPM, the main engine shaft torque, and fuel consumption. Moderate correlations exist between speed, engine temperature, trim, and fuel consumption. Interestingly, ship course, heading, and hydrometeorological factors displayed little relevance or even negative correlation with fuel consumption. This discrepancy is due to the use of different coordinate systems. The DT emphasizes the importance of combining hydrometeorological values with their directions to determine fuel consumption accurately in operational scenarios.

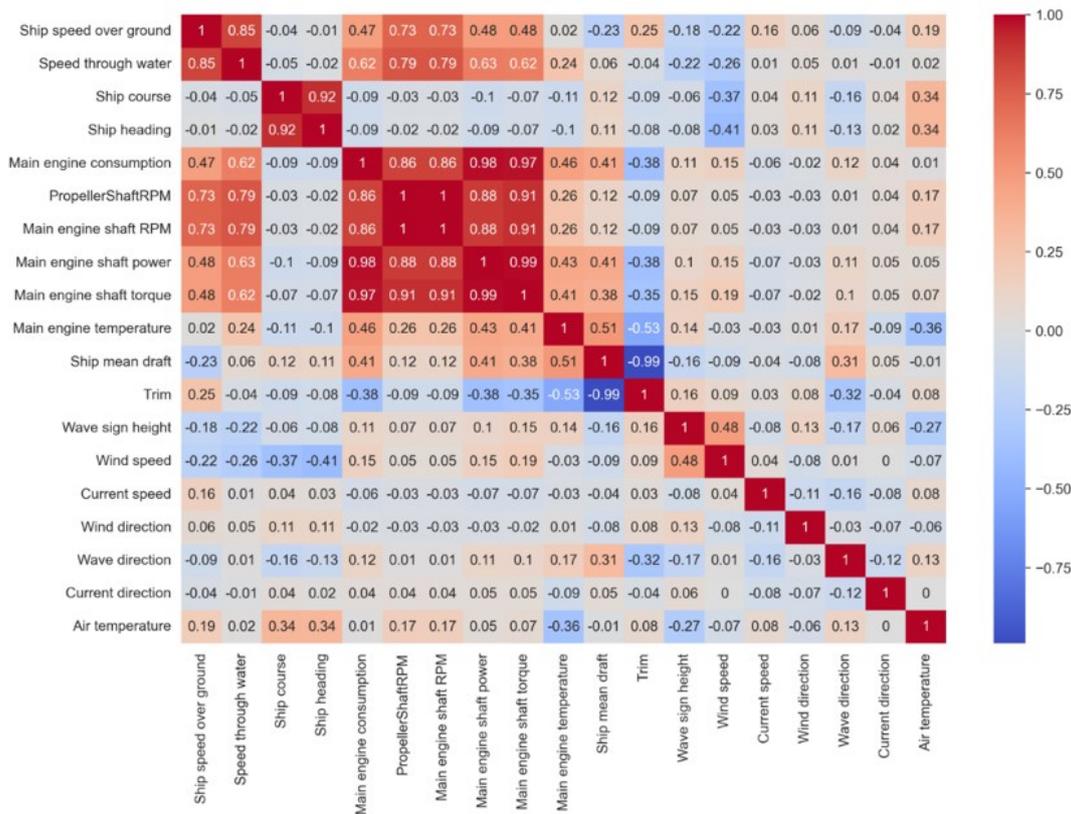


Figure 41: Correlation between the selected influencing factors on the ship fuel consumption, using the collected data streams.

The information on ship navigation conditions (speed, course, heading), ship operation (draft, trim), engine operation (propeller, main engine shaft RPM and torque) and external operational conditions serve as primary control parameters (or influencing factors) for managing the ship energy system and navigation system. These factors are interconnected, making it difficult to optimize ship operations for reduction of fuel consumption when considering only individual or a few factors (see more in Figure 41). The knowledge of these factors forms the basis for a multi-objective optimization algorithm (see more in Figure 38). It is worth noting that the propeller/main engine shaft RPM and torque are closely related and exhibit interdependencies, see Figure 45. Hence, they are not included as input parameters (see Figure 42).

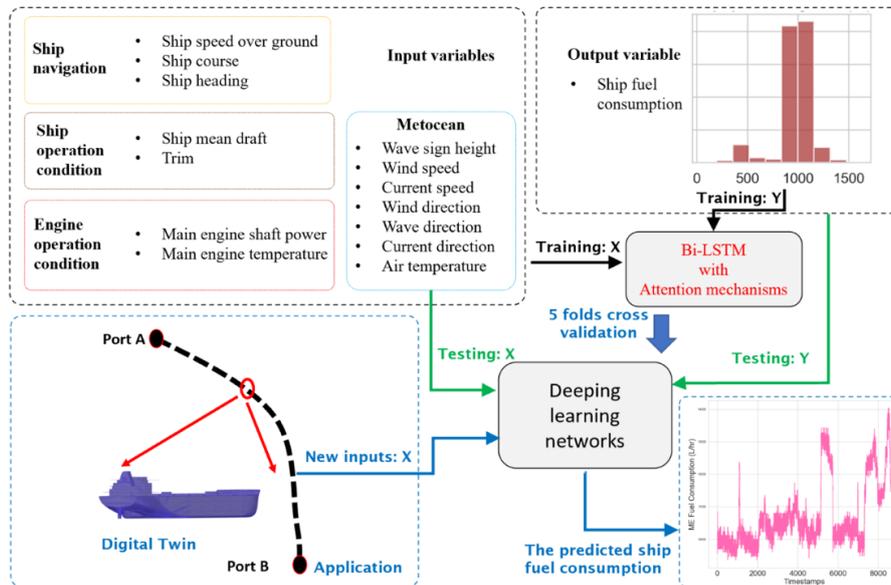


Figure 42: Deep learning processing of the ship fuel consumption prediction for model training, testing and application.

The deep learning model architecture used for training is composed of an input layer, three bi-LSTM layers, three attention layers, and an output layer. Each bi-LSTM layer featured 128 hidden units. The optimization of hyperparameters, including the regularization parameter, dropout rate and learning rate for the bi-LSTM with attention mechanism model, was conducted through a grid search employing a five-fold cross-validation approach.

The best hyperparameters were selected based on the lowest Mean Squared Error (MSE) observed during the validation process (Table 25). The dataset was divided into five subsets for cross-validation. Training and validation losses were computed for the respective 80% and 20% splits of the dataset as depicted in Figure 43.

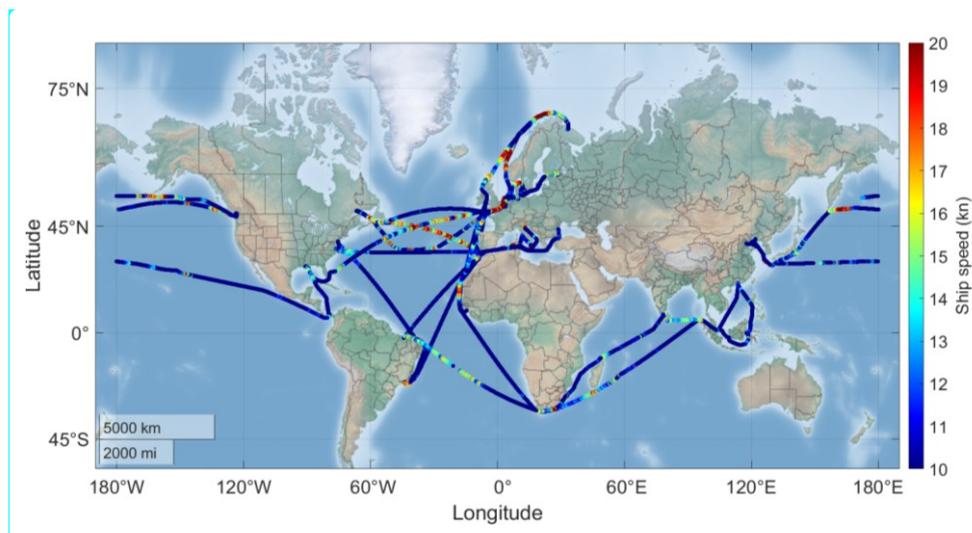


Figure 43: Ship trajectories of a bulk carrier for training the ship fuel consumption prediction model from 01/2021 to 02/2023.

Table 25: Model characteristics and the optimal hyperparameters.

Model	Input variables	Output variable	Layers
Bi-LSTM with attention mechanisms	14	1	7
Hidden units per layer	Optimizer	Batch Size	Early stopping
128	Adam	48	Patience=10
Dropout rate	Learning rate	Epochs	Regularization param
0.2	5e-05	178	0.1

As depicted in Figure 44, both training and validation losses steadily decreased and stabilized around the 178th epoch. At this point, if no discernible improvement in validation performance occurred, the training process was prematurely terminated to prevent overfitting. The model optimal fitting state was achieved after 178 epochs, indicating a good balance between the model performance and the training data. Such a balance ensures that the model neither suffers from overfitting nor from underfitting, resulting in an effective generalization to new data while capturing training data complexities.

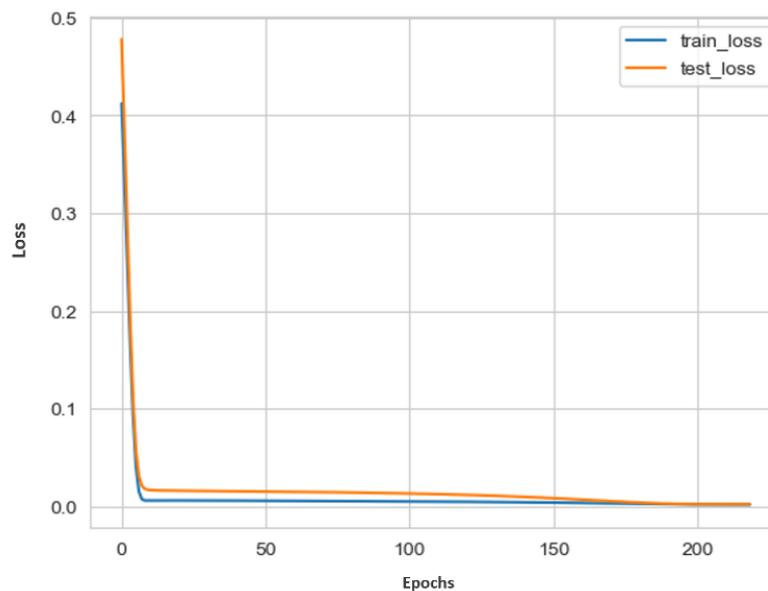


Figure 44: Training and testing loss over Epochs for the evaluation of model performance.

To expand the applicability of the trained model to a next voyage of the same ship, the new data stream was collected from January 2023 to June 2023. This new data stream encompasses eight entire voyages worldwide, as illustrated in Figure 45.

The longest trajectory among these voyages is 7223.9 nautical miles, while the shortest one is 980 nautical miles. Voyage details are provided in Table 26. Utilizing the accumulated navigation data, ship operating conditions, engine operating conditions, and Metocean observations, the predictive capabilities of the trained model to estimate fuel consumption for this specific itinerary were tested.

Table 26: General information of the extra eight voyages used in the model validation.

	Condition	Departure/destination	Voyage lengths	Number of data
Voyage 1	Laden condition	Vancouver/Yantai	7223.90 nm	28,479
Voyage 2	Ballast condition	Yantai/Hong Kong	980.22 nm	6,081
Voyage 3	Laden condition	Hong Kong/Sarina	3035.639 nm	17,347
Voyage 4	Ballast condition	Sarina/Singapore	3764.36 nm	19,545
Voyage 5	Laden condition	Singapore/Pouad	4724.23 nm	19,749
Voyage 6	Laden condition	Pouad/Rotterdam	3954.41 nm	8,299
Voyage 7	Ballast condition	Rotterdam/Belem	4230.67 nm	19,930
Voyage 8	Laden condition	Belem/Rotterdam	4197.65 nm	21,206

Save model: `model.save('Ship fuel consumption model.h5')`

Load model: `loaded_model = load_model('Ship fuel consumption model.h5')`

Use the model: `Predictions = loaded_model.predict(New_inputs)`

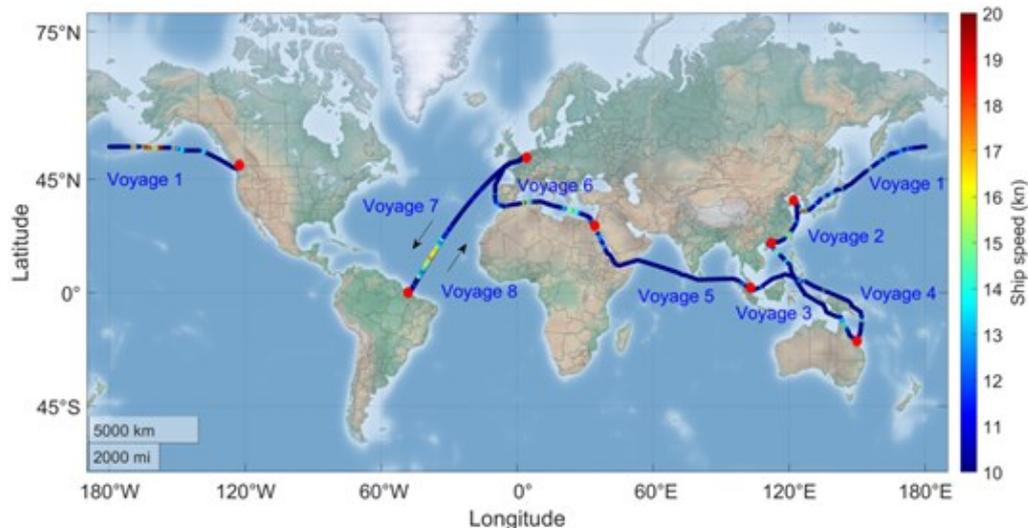


Figure 45: Ship trajectories of the bulk carrier for training the ship fuel consumption prediction model from 01/2023 to 06/2023

As shown in Figure 45, the trained ship fuel consumption model was loaded and then ship fuel consumptions of eight entire voyages were predicted based on the trained model in the time domain. For example, in Figure 46, it is shown that more than 90% of the prediction errors are below 10% and that the average error of the ship fuel consumption on the voyage 1 is estimated with an error of -0.51 %. Furthermore, the evaluation metrics (R², RMSE, MAE and average error rate) are used to present the discrepancy between the real and the predicted fuel consumptions. The evaluation of the generalization ability of the model are presented in Table 27. The results, as shown in Figure 47, reveal the outcomes of fuel consumption predictions. These results indicate that the R² values range from 0.71 to 0.94. Voyage 2 exhibits the smallest Mean Absolute Error (MAE) at 28.19 L/h, and the Root Mean Square Error (RMSE) for voyage 2 is also the lowest at 38.99 L/h. Furthermore, the average error rate values range from -0.51% to 5.56% and, the smallest average error rate of -0.51% is observed in voyage 1.

These findings show that the trained model is an efficient tool for forecasting fuel consumption during comparable voyages. Consequently, the application of the trained model can facilitate an efficient fuel management and optimization in real operational conditions.

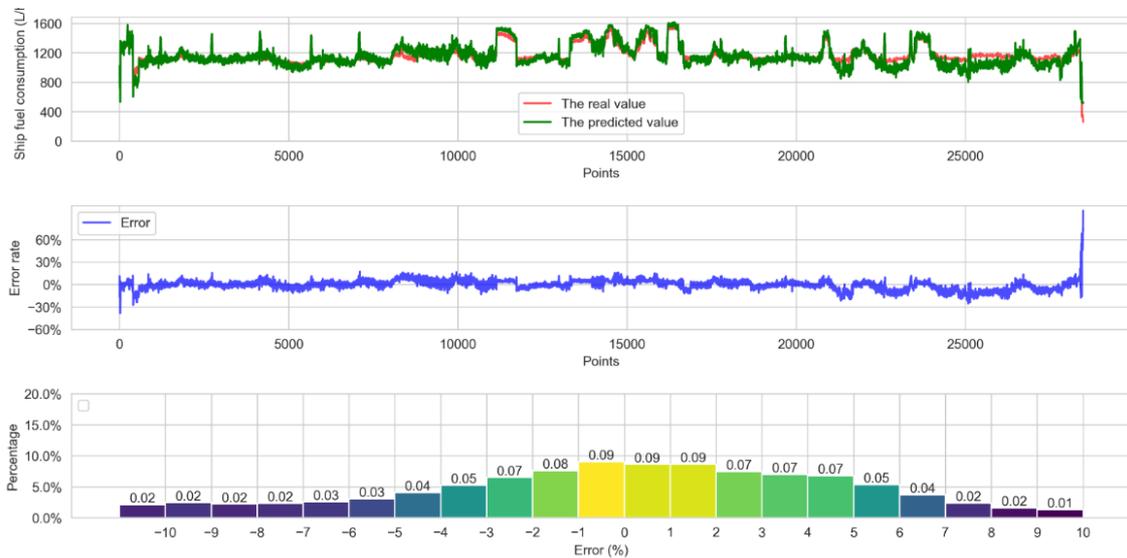


Figure 46: Error analysis of the fuel consumption prediction for the voyage 14

Table 27: Generalization ability evaluation of the trained ship fuel consumption prediction model for eight global voyages⁵

	Voyage 1	Voyage 2	Voyage 3	Voyage 4	Voyage 5	Voyage 6	Voyage 7	Voyage 8
R ²	0.84	0.73	0.88	0.71	0.94	0.94	0.92	0.80
RMSE (L/h)	53.24	38.99	61.02	72.72	29.09	57.33	42.14	86.14
MAE (L/h)	39.99	28.19	42.54	58.05	19.89	41.50	28.47	59.67
Average error rate	-0.51%	2.64%	-2.68%	5.56%	2.53%	1.36%	2.48%	-3.59%

Note: The used eight voyages are presented in Table 26 and Figure 45.

As a conclusion, this appendix has introduced a deep learning approach for precise ship fuel consumption prediction under real operational conditions, focusing on a Kamsarmax bulk carrier. The study encompasses three key aspects:

- evaluation of the significance of the factors affecting fuel consumption,
- training of deep learning neural networks to model the operation of the ship energy systems,
- the real-time prediction of the ship fuel consumption for long voyages.

⁴ Top panel: the red line refers to the actual values of fuel consumption, the green line refers to the predicted values. Center panel: the blue line refers to the error in the time domain ($e_n = (y_n - \hat{y}_n)/y_n$), y_n is the actual value, \hat{y}_n denotes the predicted value). Bottom panel: presents the prediction error distributions.

⁵ RMSE, MSE, and the R2 value are used to measure the diffidence between the real value and the predicted value.



Data from sea trials spanning two years were utilized for the model development and validation. The Bi-LSTM model with attention mechanisms emerges as the most effective choice for real-time fuel consumption prediction during a voyage. This approach appears very promising for the ship route optimization, lowering the level of emissions and reducing operational costs, with benefits both for the environment and for the ship operators.



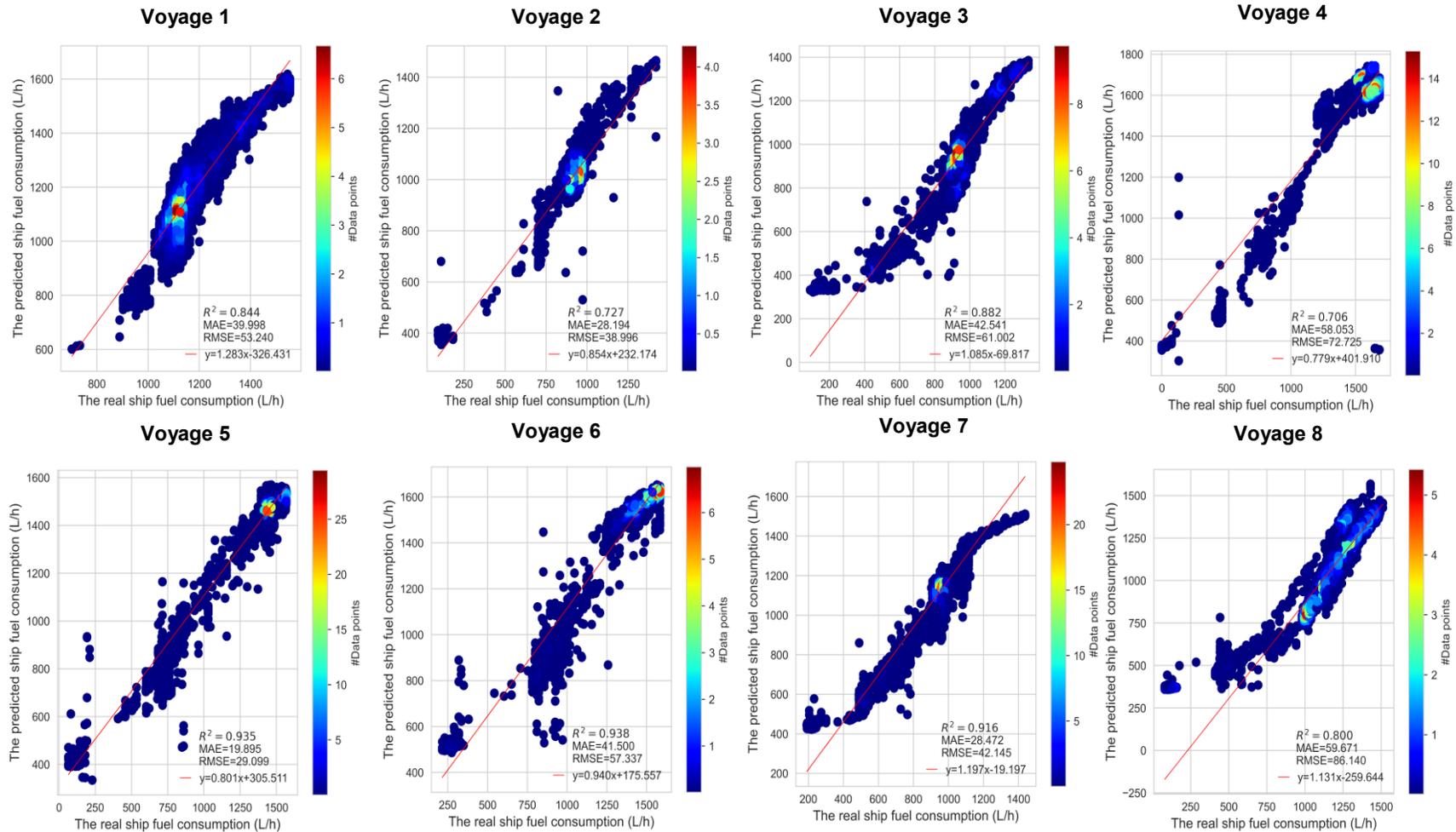


Figure 47: Comparison between the actual and the predicted ship fuel consumption for the eight voyages.