DTU Library

Data-driven Prediction of Added Resistance on Ships in Waves

Mittendorf, Malte

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Data-driven Prediction of Added Resistance on Ships in Waves

Malte Mittendorf

DTU
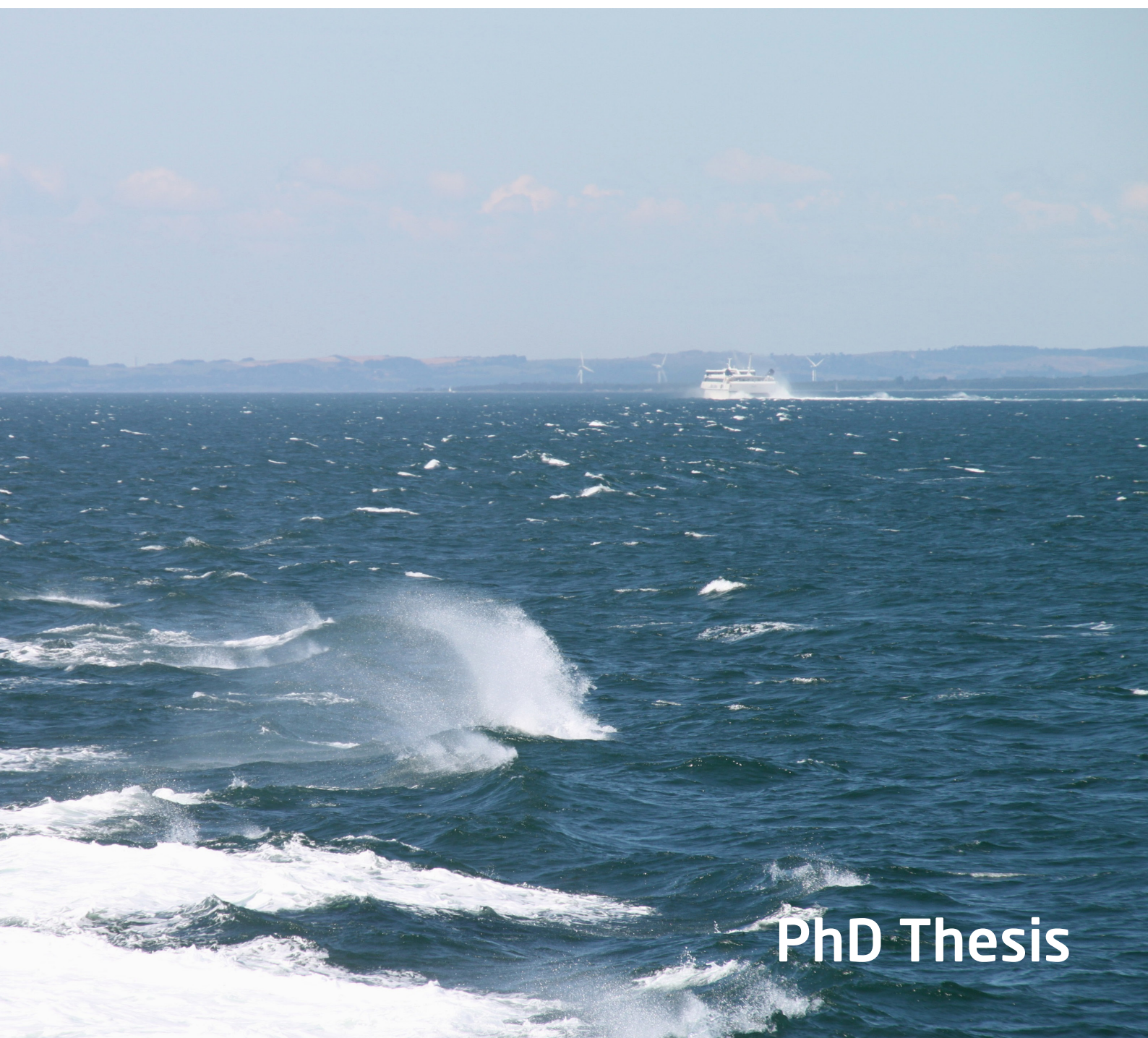
PhD Thesis

DATA-DRIVEN PREDICTION OF ADDED RESISTANCE ON SHIPS IN WAVES

PHD THESIS

Malte Mittendorf

Section for Fluid Mechanics, Coastal and Maritime Engineering
Department of Civil and Mechanical Engineering
Technical University of Denmark
February 2023

In physics the truth is rarely perfectly clear, and that is certainly universally the case in human affairs. Hence, what is not surrounded by uncertainty cannot be the truth.

– Richard P. Feynman, February 1976
Letter to the Editor of California Tech

This page is intentionally left blank.

Preface

The present thesis is submitted for partial fulfillment of the requirements for obtaining the Ph.D. degree at the Technical University of Denmark (DTU). The work has been carried out at the Section of Fluid Mechanics, Coastal and Maritime Engineering in the period between March 2020 and February 2023. Associate Professor Ulrik Dam Nielsen and Professor Harry B. Bingham supervised the study. The project was financially supported by Den Danske Maritime Fond (Projekt 2019-043), A/S D/S Orient's Fond, and the Department of Civil and Mechanical Engineering of DTU. This support is highly acknowledged.

Malte Mittendorf
Nærum, 28$^{\text{th}}$ February 2023

This page is intentionally left blank.

Executive Summary

In combination, international regulations and digitalization have led to increasing acquisition of in-service data from ships, including navigational and engine-related recordings. However, associated analyses are challenging due to the limited availability of reference data and operation in an environment governed by uncertainty, e.g. related to wind and waves. The added wave resistance can take up a relatively large proportion of a ship's required engine power. Still, its magnitude is notoriously difficult to determine – especially in short and oblique waves. Hence, the overarching goal of this thesis is to enhance the predictability and understanding of added resistance in real conditions using statistical and machine learning methods.

The first part of this thesis addresses the added resistance in regular waves and focuses on estimating the associated quadratic transfer function. Several machine learning methods are trained on the results of numerical methods for various hull shapes. The importance of data preprocessing and the generalization capability of neural networks stand out. A separate study investigates the uncertainty of the added resistance transfer function and its estimation via a semi-empirical formula. For this reason, parameter calibration of the underlying method is performed separately for both blunt-type and slender vessels. The methodology is based on experimental data, and a 90% prediction interval is implemented to improve the method's transparency in an adapted version.

An estimate of the corresponding wave energy density spectrum at the exact spatiotemporal point of operation is needed to calculate added resistance in actual seaways characterized by irregular waves. Hence, in the second part of the work, in-service sensor data of a container vessel and different neural networks are utilized for sea state identification. Overall, neural networks can produce a satisfactory mapping from measured vessel responses to sea state parameters.

The third part of the work is about correlating theoretical and empirical estimates of the mean added resistance in actual conditions using in-service data from a fleet of more than 200 container vessels. Theoretical estimates are calculated in the spectral domain by combining the adapted semi-empirical procedure and historical wave data. The empirical predictions are determined using the measured shaft power combined with a resistance decomposition. It is confirmed that the actual added resistance is highly complex to determine and subject to significant uncertainty.

In the fourth and final part of the work, synthetic monitoring data of a standard tanker (KVLCC2) is simulated for varying operating conditions using a semi-empirical framework. It is shown that the performance data of ships is subject to a distributional shift, and thus

neural networks are trained adaptively to pinpoint the added power due to biofouling. It turns out that methods for incremental learning are influenced by data quality and that the overall methodology may be applicable for determining the effect of energy-saving devices.

In the future, this work may serve the development of digital twins used to assess the safety and the energy efficiency of ship operations. Due to the sensitivity of machine learning methods to data quality and availability, it seems favorable to follow a hybrid approach in combination with established physical models. In a practical context, the findings of this work are applicable for voyage optimization or performance monitoring and may assist the maritime industry on its path to becoming sustainable.

Resumé

Internationale reguleringer og digitalisering har i kombination ført til stigende indsamling af in-service data fra skibe, herunder navigations- og motordata. Tilknyttede analyser er dog udfordrende på grund af begrænset tilgængelighed af referencedata og drift i et miljø, hvor usikkerheder er signifikante, f.eks. relateret til vind og bølger. Således kan den forøgede modstand et skib oplever under sejlads i bølger optage en relativt stor del af skibets motorkraft, men dens størrelse er vanskelig at bestemme – især i korte og skrå bølger. Derfor er det overordnede mål med denne afhandling at forbedre forudsigeligheden og forståelsen af ekstra modstand under virkelig sejlads ved hjælp af statistiske metoder og *machine learning* (maskinlæring).

Den første del af denne afhandling omhandler den modstands forøgelse i regulære bølger og fokuserer på estimeringen af den kvadratiske overføringsfunktion. Adskillige maskinlæringsmetoder trænes på resultater af numeriske metoder for en række forskellige skrogformer og vigtigheden af dataforbehandling samt generaliseringsevnen i neurale netværk viser sig. I en separat analyse undersøges usikkerheden af overføringsfunktionen af den forøgede modstand og dens estimering via en semi-empirisk formel. Af denne grund udføres parameterkalibrering af den underliggende metode separat for fyldige og slanke skibe. Metoden er baseret på eksperimentelle data og et 90% forudsigelsesinterval er implementeret for at forbedre metodens gennemsigtighed.

Til beregning af den forøgede modstand i virkelig sø, karakteriseret ved uregelmæssige bølger, er der behov for et estimat af det tilsvarende bølgespektrum ved skibets præcise geografiske placering i både tid og rum. Derfor anvendes i den anden del af afhandlingen sensordata fra et containerskib og forskellige neurale netværk til identifikation af søtilstanden. Samlet set viser det sig, at neurale netværk er i stand til at producere et tilfredsstillende forhold mellem målte skibsrespons og bølgeparametre.

Den tredje del af arbejdet handler om at korrelere teoretiske og empiriske estimater af den gennemsnitlige forøgede modstand under virkelige forhold ved hjælp af in-service data fra en flåde på mere end 200 containerskibe. Teoretiske estimater beregnes i det spektrale domæne ved at kombinere den tilpassede semi-empiriske procedure og historisk bølge data. De empiriske forudsigelser bestemmes ved hjælp af den målte motorkraft kombineret med modstandsdekomponering. Det bliver bekræftet, at analyser og beregninger af den forøgede modstand er komplekse, og estimater er generelt behæftet med stor usikkerhed.

I den fjerde og sidste del af afhandlingen simuleres syntetiske monitoreringsdata for et tankskib (KVLCC2) for varierende operationsprofiler ved hjælp af en semi-empirisk model. Det vises, at ydelses data fra skibe er genstand for et distributionsskifte, og derfor trænes

neurale netværk adaptivt for at lokalisere den krævede ekstra motorkraft som følge af begroning. Det viser sig, at metoder til inkrementel læring er påvirket af datakvalitet, og på samme tid at de udviklede metoder kan være anvendelige til at bestemme effekten af energibesparende tiltag gennem retrofitting.

I fremtiden kan dette projekts arbejde tjene til udvikling af digitale tvillinger, der bruges til at vurdere både sikkerhed og energieffektivitet ved skibsoperationer. På grund af maskinlæringsmetodernes følsomhed over for datakvalitet og tilgængelighed, virker det gunstigt at følge en hybrid metode, der kombinerer maskinlæring med etablerede fysiske modeller. I en praktisk sammenhæng er resultaterne af dette projekt anvendelige til ruteoptimering og ydelsesmonitorering og kan således hjælpe den maritime industri på dens vej mod at blive bæredygtig.

List of Publications

The present Ph.D. thesis comprises the following published or submitted journal papers:

[J_1] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B. (2022). *Data-driven prediction of added-wave resistance on ships in oblique waves - A comparison between tree-based ensemble methods and artificial neural networks.* Applied Ocean Research, Vol. 118, 102964

[J_2] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B.; LIU, S. (2022). *Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading.* Ocean Engineering, Vol. 251, 111040

[J_3] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B.; STORHAUG, G. (2022). *Sea state identification using machine learning - A comparative study based on in-service data from a container vessel.* Marine Structures, Vol. 85, 103274

[J_4] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B.; DIETZ, J. (2023). *Assessment of added resistance estimates based on monitoring data from a fleet of container vessels.* Ocean Engineering, Vol. 272, No. 113892

[J_5] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B. (2023). *Capturing the effect of marine growth on ships by adaptive machine learning.* (Under Review)

During the Ph.D. study period, the following conference papers have been presented and published as open access. It is noted that these are not included in this thesis.

[C_1] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B. (2022). *The Prediction of Sea State Parameters by Deep Learning Techniques using Ship Motion Data.* Proc. of the 7th World Maritime Technology Conference (WMTC), Copenhagen.

[C_2] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B.; GUNDERMANN, D.; SCHMODE, D.; DEYMIER, C. (2022). *Performance Analysis of a Gas Carrier using Continual Learning in a Data Stream Context.* Proc. of the 7th Hull Performance and Insight Conference (HullPIC), Tullamore.

[C_3] MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B.; STORHAUG, G. (2022). *On the Determination of the Relative Wave Direction based on Measured Ship Responses using Deep Multi-task Learning.* Proc. of the 14th Symposium on High-Performance Marine Vehicles (HIPER), Cortona.

This page is intentionally left blank.

CRediT Author Statements

[J_1] **Malte Mittendorf**: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Validation, Visualization, Writing – original draft. **Ulrik Dam Nielsen**: Writing – review & editing, Supervision, Project Administration, Resources, Funding Acquisition. **H. B. Bingham**: Writing – review & editing, Supervision, Resources, Project Administration.

[J_2] **Malte Mittendorf**: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Validation, Visualization, Writing – original draft. **Ulrik Dam Nielsen**: Writing – review & editing, Supervision, Project administration, Resources, Funding Acquisition. **Harry B. Bingham**: Writing – review & editing, Supervision, Resources, Project administration. **Shukui Liu**: Writing – review & editing, Software, Resources.

[J_3] **Malte Mittendorf**: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Validation, Visualization, Writing – original draft. **Ulrik Dam Nielsen**: Conceptualization, Methodology, Writing – review & editing, Supervision, Project Administration, Resources, Funding Acquisition. **Harry B. Bingham**: Conceptualization, Methodology, Writing – review & editing, Supervision, Resources, Project Administration. **Gaute Storhaug**: Conceptualization, Data Curation, Writing – review & editing, Resources.

[J_4] **Malte Mittendorf**: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Validation, Visualization, Writing - original draft. **Ulrik Dam Nielsen**: Conceptualization, Methodology, Writing - review & editing, Supervision, Project Administration, Resources, Funding Acquisition. **Harry B. Bingham**: Conceptualization, Methodology, Writing - review & editing, Supervision, Resources, Project Administration. **Jesper Dietz**: Conceptualization, Methodology, Data Curation, Writing - review & editing.

[J_5] **Malte Mittendorf**: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Validation, Visualization, Writing - original draft. **Ulrik Dam Nielsen**: Writing - review & editing, Supervision, Project Administration, Resources, Funding Acquisition. **Harry B. Bingham**: Writing - review & editing, Supervision, Resources, Project Administration.

This page is intentionally left blank.

Table of Contents

Chapter 1

Introduction

The shipping sector is standing at a crossroads. On one hand, around 80% of all traded goods are handled by seaborne transport and the demand is projected to grow in the coming decades, UNCTAD [127]. On the other hand, the world fleet contributes approximately 3% to anthropogenic carbon emissions, IMO [53]. Despite the industry's large share of global emissions, maritime transport is still considered the most efficient transport mode, Papanikolaou [98]. Nonetheless, the change in climate and its expected consequences call for a rapid and substantial reduction in emissions. In this light, the International Maritime Organization (IMO) proposed its Greenhouse Gas (GHG) strategy (cf. IMO [52]), which is depicted in Fig. 1.1 on the secondary axis in parallel with the historical and forecasted transport work of four shipping sectors on the primary axis according to DNV [26].
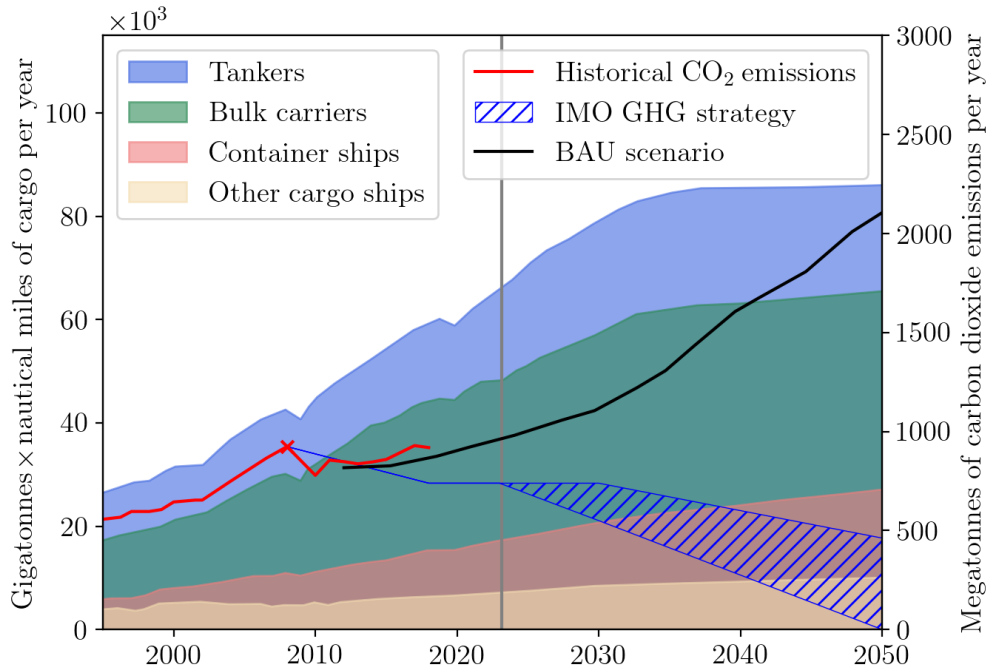

Figure 1.1: Transport work per ship type as well as historical and estimated CO_2 emissions of the global fleet in comparison to IMO's GHG strategy, cf. [19], [26], [52] and [53].

From Fig. 1.1, it is appreciated that maritime transport, as the spine of globalization, experienced sizable continuous growth in the last 30 years – excluding a few years with economic contraction. Expectedly, the depicted transport work of the international merchant fleet shows a profound correlation with historical CO_2 emissions, taken from Cames *et al.* [19] and IMO [53]. As a reaction, IMO ratified the first countermeasures curbing shipping-related emissions in 2011, and the initial GHG strategy entered into force in 2018. The respective goals of this legislation are in accordance with the Paris Agreement and are visualized in Fig. 1.1. It stands out that there is a range of possible decarbonization scenarios beyond 2023, which are, in fact, bounded by IMO's initial GHG strategy aiming at halving emissions from maritime transport by 2050, with 2008 as the base year. The lower and more ambitious strategy is preferred by several intergovernmental bodies and shipping companies that intend to reach carbon-neutral shipping by 2050. Subsequently, discussions regarding incorporating this progressive target in an amended GHG strategy will be part of the 80$^{\text{th}}$ session of the Marine Environment Protection Committee (MEPC) in summer 2023. Alongside IMO's ambitions, a business-as-usual (BAU) baseline projected with moderate economic growth (as of 2012) is depicted, taken from Cames *et al.* [19]. For the sake of simplicity, the associated prediction interval of the BAU projection is omitted herein. As can be inferred from Fig. 1.1, a minor disparity in magnitude between the BAU scenario and emissions manifests itself already in the years after 2012. A similar but even more significant emission and corresponding innovation gap stands out in the coming decades when considering IMO's GHG strategy and the projected BAU scenario. Hence, when we refer to the metaphor from the beginning, the case could be made that the shipping industry has chosen a path, i.e. it follows the BAU scenario – at least in the short term. However, it is found that the correlation between transport work and CO_2 emission is reduced after 2008, indicating an increase in overall energy efficiency in maritime transport.

Eventually closing the emission and innovation gap seen in Fig. 1.1 will be an enormous task in the years to come and includes all possible stakeholders in the maritime field. As a demarcation, the present thesis solely focuses on technical measures. It, therefore, disregards any logistical and marked-based efforts, even though these three domains may ultimately complement each other. Bouman *et al.* [15] conduct a thorough review of available literature focusing on reducing carbon emissions following novel approaches. Their results are reproduced in a condensed manner in Fig. 1.2, and three main disciplines are shown: (1) Ship design aspects, (2) alternative energy sources, and (3) ship operations. For the sake of clarity, the shown box plots indicate the interquartile range with their body, the median, as well as the maximum and minimum of the sample distribution. As an immediate finding, the most promising technical aspect for reducing carbon emissions seem to be novel fuel types, i.e. biofuels. Additionally, alternative and renewable fuels (or E-fuels) have received much attention recently since they show an even higher emission reduction potential. Indeed, these fuel types represent a cornerstone for decarbonizing shipping. Nevertheless, practical implementation of the corresponding fuels is, up until now, plagued by uncertainty since both the scalability of production and global availability, i.e. the supply infrastructure, are still out of reach, disqualifying them as an immediate measure. Furthermore, Lindstad *et al.* [78] state that E-fuels are per se not limited to newbuildings, but retrofitting existing vessels seems to be (as of now) not particularly economically viable. Hence, considering the current

world fleet composed of ca. 103,000 vessels (cf. [127]), it turns out that the implementation of energy-saving devices, such as air lubrication and wind propulsion equipment, will be a critical aspect for lowering carbon emissions in the short term. Subsequently, enormous research efforts have been invested in the respective areas, and technical solutions show an increasing level of maturity as well as applicability.
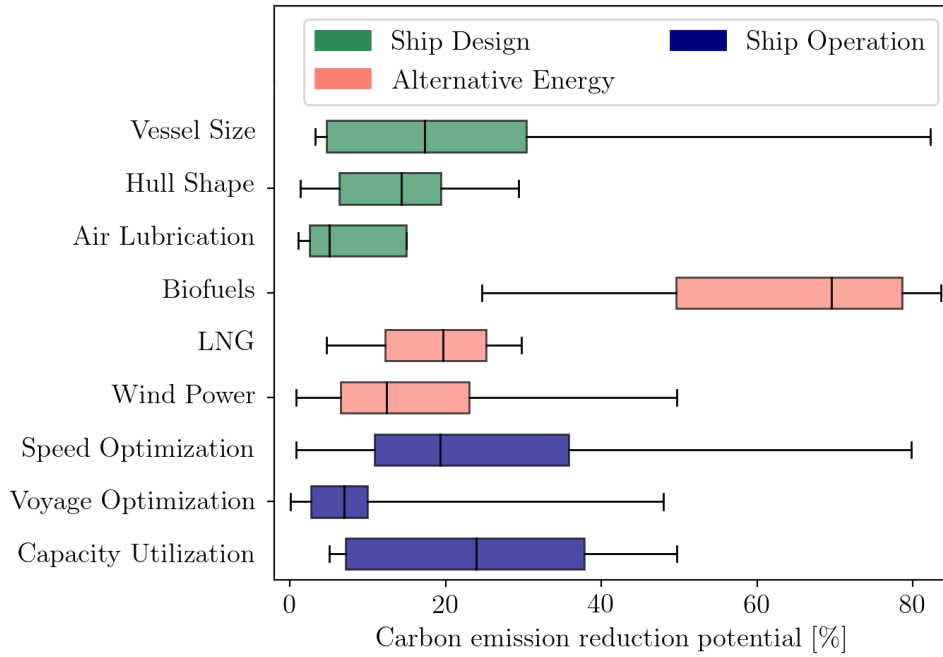
Figure 1.2: Emission reduction potential of several methods of three domains, according to Bouman *et al.* [15].

Another observation from Fig. 1.2 is that operational measures and design aspects show roughly similar magnitudes in CO_2 saving potential. It can be seen that *economy of scale*, i.e. vessel size, not only leads to a cost reduction from an economic perspective but also greater energy efficiency. Bouman *et al.* [15] emphasize that doubling a ship's cargo capacity leads, roughly, to an increase of 66% in fuel consumption. Thereby it is understood that a vessel's fuel consumption is directly linked to its carbon emissions and also makes up more than 50% of its operational cost, Perera and Guedes Soares [102]. Moreover, optimizing a ship's hull shape for the least hydrodynamic resistance is another important stepping stone for reducing emissions. In fact, ship design synthesis aims at optimizing vessels for their life cycle, i.e. tailoring the design in a multi-objective optimization for its actual service task and environment, Papanikolaou *et al.* [99]. In doing so, the two disciplines of ship design and operation, which were handled rather separately in the past, are brought closer together. For instance, Mittendorf and Papanikolaou [90] present the hull shape optimization of a fast passenger catamaran for its operational profile, i.e. the resistances at multiple speeds are considered as objective functions. However, the operational profile includes not only speed (and draft) variations but also environmental influences caused by the presence of wind and waves.

In the domain of ship operations, the voluntary reduction of the vessel's service speed, i.e. slow steaming, is a straightforward measure for increasing energy efficiency due to the proportionality of engine brake power to cubed ship speed. In fact, speed optimization shows a sizable emission savings potential in Fig. 1.2 with a higher mean than, e.g. optimizing vessel size (ship design). On a separate note, engine power limitations may be enforced in case of non-compliance to the newly adopted IMO regulations, leading to *involuntary* slow steaming. Still, Taskar and Andersen [121] found in a simulation-based study that the benefits of speed reduction depend highly on the encountered sea state and the extent of the added resistance due to waves. In fact, the latter is less affected by ship speed and thus increases relatively in magnitude as a fraction of the total resistance. It is noted that many savings estimates for speed optimization from the literature fall outside of the interquartile range, which indicates low agreement among the individual studies according to Bouman *et al.* [15]. Furthermore, voyage optimization is concerned with finding the optimal route considering economic constraints and maximizing energy efficiency as well as crew and cargo safety, Perera and Guedes Soares [102]. This is achievable through speed optimization and the circumvention of harsher weather conditions. In fact, D'Agostini *et al.* [24] show that storms may increase in their intensity due to consequences of climate change emphasizing the necessity of weather routing and accurate estimation methods of added resistance. It is believed that all three depicted operational measures are, in fact, intertwined and build on each other. The third operational aspect in Fig. 1.2 is capacity utilization which refers to exploiting network effects within a fleet of vessels. It shows a high impact, but it is stressed that only a handful of samples are available, underlining the uncertainty in the presented data. Another pivotal operational aspect is the utilization of in-service data of ships for the optimization of trim and maintenance schedules by performance monitoring, i.e. identifying the effect of marine growth. Overall, these additional operational measures are considered "low-hanging fruits", i.e. require little effort, but exhibit an approximate emission savings potential of 5-25%, according to IMO [55]. Referring to the existing world fleet, it shows that enhancing operational performance will be a driving factor for cutting fuel consumption and, thereby carbon emissions while maintaining the level of transport work.

Ultimately, Bouman *et al.* [15] show that all methods mentioned above in Fig. 1.2 have to be applied in conjunction to realize substantial emission savings for reaching IMO's intentions, as shown in Fig. 1.1. Broadly speaking, adopting innovative methods and technologies in the maritime industry has been regulation-driven for the last decade. As indicated by IMO [52], the present decade will be characterized by a predominant focus on reducing *carbon intensity*, which refers to a vessel's carbon emissions relative to its accomplished transport work. Following Lindstad *et al.* [78], GHG emissions per ton-mile of transport work must be reduced by at least 75-85% by 2050 to achieve the goal of the initial IMO strategy under the assumption of continuous growth within maritime transportation. Consequently, the Carbon Intensity Indicator (CII) came into action in 2023. The CII is an operational indicator determined annually and is subject to increasingly demanding limits, IMO [54]. The CII is not free of criticism since it employs the potential transport work by taking the maximum (or design) cargo capacity for calculation instead of considering the actual loading condition, cf. Wang *et al.* [131]. Besides volatile fuel prices and freight rates, the CII adds to the regulatory pressure and thus, to the overall uncertainty felt by ship

owners and operators. Still, the CII will act as an essential short-term measure in the IMO GHG strategy incentivizing shipping companies to streamline their fleet operations.

1.1 Motivation

Traditionally, fundamental ship characteristics, such as hull shape, propeller geometry, and main engine power, are determined and optimized in calm water conditions under design speed and draft. However, from the analysis of in-service data of ships, it turns out that ships have a more widespread operational profile, which is most often determined by economic circumstances. This hypothesis is solidified in Fig. 1.3, where approximately two years of ship telemetry data from eight container sister vessels (15,000 TEU) is depicted in two separate diagrams. In Fig. 1.3a, the mean draft of the vessels is shown for the non-dimensional advance speed, i.e. the Froude number Fn. In parallel, the initial design conditions are indicated and it stands out that these have never been experienced during the observation period. In fact, the vessels sailed predominantly with scantling draft and a lower service speed, thereby improving energy efficiency. For context, the data originates from the period between January 2019 and March 2021, i.e. includes the consequences of the severe disruptions in the global supply chains due to the worldwide COVID-19 pandemic. Moreover, it underlines that ships – and container ships in particular – do not have *one* point of operation, as often anticipated in design stages.
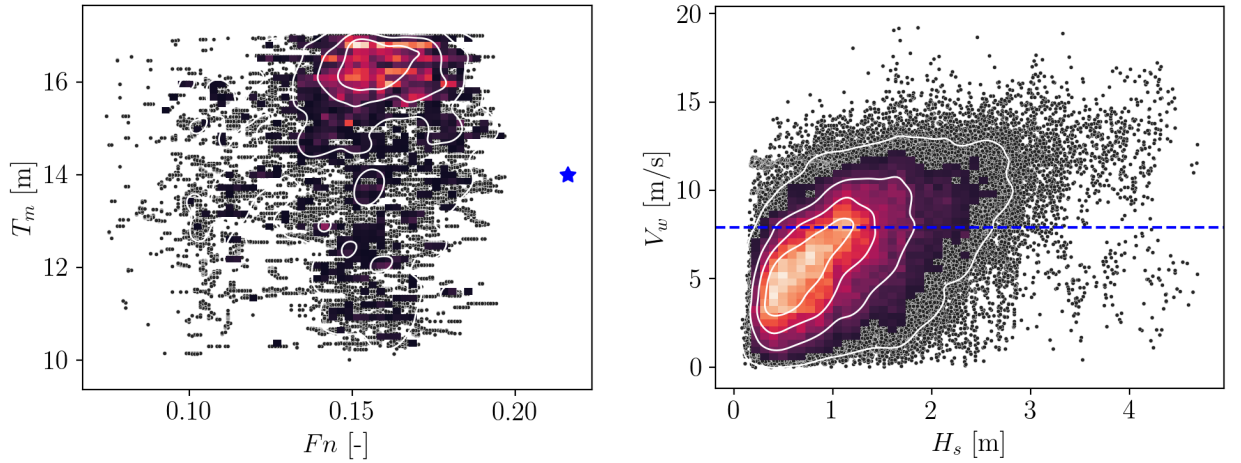
(a) Joint distribution of mean draft T_m and non-dimensional ship speed Fn. It is noted that the design conditions are indicated by a blue star.

(b) Encountered absolute wind speed V_w as a function of significant wave height H_s. The filtering threshold of ISO 19030 is shown.

Figure 1.3: Two joint distributions for eight 15,550 TEU sister vessels compiled from publicly available AIS and ERA5 hindcast data, as taken from the study presented in $[J_4]$. It is noted that the figures stack scatter, 2D histogram and kernel density estimate plots.

Interestingly, the re-emergence of vertical bow concepts in contemporary ship design leads not only to enhanced versatility regarding draft and speed variation, but also shows

beneficial behavior in case of the added resistance due to waves, as demonstrated by Yu *et al.* [132]. When considering environmental influences, such as waves, it stands out that the actual operational profile of the eight sister ships turns out to be even more complex as reflected by Fig. 1.3. Here, the encountered absolute (or true) wind speed V_w is displayed for the significant wave height H_s indicating the experience of relatively severe sea states (up to Beaufort 7) despite the use of routing optimization. However, for setting viable engine margins, only a lump-sum allowance (or sea margin) is used during early ship design taking the added resistance due the ambient environment into account, Strøm-Tejsen *et al.* [117]. In later design stages, the vessel may be optimized for least added resistance in waves using numerical methods, but it is rather uncommon to do so. Being a part of the IMO strategy [52], the Energy Efficiency Design Index (EEDI) and the EEXI (for existing ships) include the weather factor f_w, which expresses the speed loss due to waves in a Beaufort 6 sea state (corresponding to $H_s = 3.0$ m and $V_w = 12.6$ m/s) compared to calm water conditions. Taking Fig. 1.3b into account, the choice of the reference sea state can be considered as conservative. However, it underlines that accurate methods for the prediction of added resistance are required for a rational evaluation of f_w. With the introduction of the EEDI in 2011, the most straightforward measure for compliance was to decrease the installed main engine power (without any other design changes). In case of a few slower blunt type vessels, the lack of engine power led in harsh weather conditions to the loss of maneuverability and subsequently to the endorsement of the Minimum Propulsion Power (MPP) guidelines, IMO [51]. As stressed by Holt and Nielsen [48], reliable and practical methods for the prediction of added resistance are of high importance for the MPP assessment and for defining viable engine/propeller load diagrams.

It shows that the added resistance due to waves is ubiquitous in a ship's life cycle, i.e. throughout ship design and operation. As pointed out before, tremendous savings can be realized from optimizing fleet operations, but the ambient wind and wave conditions as well as their uncertainty are of great concern. Additionally, the assessment of ship operations is generally afflicted by the scarcity of reference data, i.e. reliable speed-power baselines. The use of sea trial or model test curves is often insufficient due to the mismatch between design and actual operating conditions, cf. Fig. 1.3a. Hence, Berthelsen and Nielsen [9] perform piece-wise regression on in-service data after imposing strict filtering criteria for approximating a quasi-calm water speed-power baseline for the assessment of ship performance. The satisfactory correction for environmental effects related to wind and waves has been a challenge in the field of vessel performance analysis ever since. In fact, the industry standard for in-service ship analyses ISO 19030 [58] imposes a draconian filtering threshold based solely on the measured absolute wind speed ($V_w < 7.9$ m/s), which is indicated in Fig. 1.3b. Following Bertram [11], the reasoning behind this threshold was that most methods were seen as inaccurate or expensive for the prediction of the added resistance in a practical context. By neglecting any swell waves, it was the aim to disregard data instances including adverse weather conditions above Beaufort 4. However, as can be seen in Fig. 1.3b, a large segment of the joint distribution, which could be deemed valid, is subsequently disregarded. In parallel, a sizeable sample size of severe swell-dominated sea states remains within the filtered dataset. These two inherent problems of the industry standard underline two key challenges, which will be addressed in the present thesis: (1) A

lack of practical methods for the reliable and accurate determination of the added resistance due to waves and (2) unavailability of sea state information for the exact vessel position. Physically, both aspects are interrelated when determining a ship's mean added resistance in a natural seaway, i.e. irregular waves.

1.2 Objectives and Scope

Stimulated by mandatory data acquisition onboard ships enforced by IMO and the EU (European Union), the availability of operational data of ships has been increasing and with it the tendency to utilize the data. Agarwala *et al.* [1] emphasize that digitalization will be a driving factor for lowering emissions from the shipping sector. In ship operations, the recent development in the field of Internet of Things (IoT), i.e. equipping ships (and other physical assets) with a sensor framework, and the emergence of data science – including machine learning – will be critical aspects going forward. Furthermore, digitalization enables the shipping sector to progress from the stages of reporting and monitoring to the third stage of predictive analysis, which subsequently facilitates the previously shown operational measures in Fig. 1.2.

The rationale of this thesis is to enhance predictability and understanding of added resistance on ships in all its facets employing data-driven methods. In the first part, regression models will be established to predict the added resistance in regular waves, i.e. the transfer function. The three overarching goals are: (1) The applicability without detailed hull shape information, (2) versatility in terms of operating conditions, and (3) the consideration of arbitrary wave heading angles. In fact, most numerical or experimental investigations consider exclusively head waves under design conditions and thus show a reduced practical relevance when referring to Fig. 1.3a.

The second part of this thesis is devoted to determining sea state information for the vessel's exact position, employing the wave buoy analogy, as described by Nielsen [95]. In other words, machine learning techniques will be utilized for mapping the measured ship responses to the encountered sea state parameters. Obviously, both added resistance transfer function and sea state data are equally important for estimating the added resistance in irregular waves. However, sea state information is commonly plagued by scarcity as well as uncertainty, which underlines the importance of using the ship itself for measuring the surrounding wave environment.

The third part is concerned with correlating theoretical estimates of added resistance in irregular waves to empirical predictions derived from in-service fleet performance data obtained on a large number of container vessels. In fact, investigations in irregular waves are most commonly conducted either numerically or experimentally under simplified assumptions. Herein, however, actual sensor data of ships in operation will be used and a data fusion concept is applied, i.e. hindcast wave data is used for modeling the ambient wave characteristics.

The fourth part proposes an adaptive machine learning methodology for ship performance monitoring. A synthetic dataset of in-service data for a standard tanker (KVLCC2) is derived for various conditions. The advantage of using simulation data is that a ground truth of the actual power increase due to fouling is available, allowing for sufficient model

assessment. In contrast to other studies, the model will be trained incrementally and the biofouling indicator will be determined implicitly.

Ultimately, it is the aim to consider the added resistance from different angles with a primary focus on the application in natural seaways. The present thesis has a strong focus on machine learning and statistical methods. In addition, the possible influence of uncertainty within the data will be an essential aspect, as an estimate is usually indispensable for decision-making.

1.3   Thesis Outline

The present thesis is divided into seven chapters – including this one. The following Chapter 2 covers the physical principles of ship resistance and the added resistance due to waves in particular. Naturally, not all theoretical details can be conveyed herein; therefore, basic knowledge of ship hydrodynamics and advanced mathematics are prerequisites. For more extensive studies of, e.g. individual expressions, reference will be made to relevant literature.

The added resistance of ships is a multi-faceted problem and the present thesis is organized accordingly. For visualizing the structure of the main part of the thesis, a Sierpiński triangle is utilized, as can be seen in Fig. 1.4. This thesis is paper-based, i.e. a collection of articles; hence, the reference of the corresponding journal paper will be indicated in parallel. As a guide for the reader, each chapter is composed of an introduction, a discussion part and an appended journal paper. It is recommended to read the journal paper directly after the introductory section.
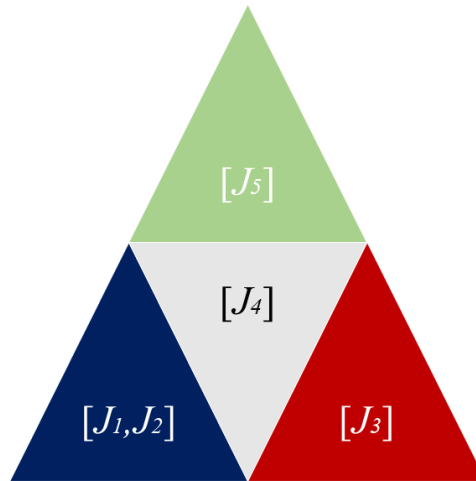
Figure 1.4: Composition of the present thesis.

The determination of the added resistance in a seaway involves both information about the ship's response and the conditions of the ambient wave system. As a consequence, the two pillars of the present thesis form the determination of the added resistance operator and an estimate of the sea state by means of a wave spectrum at the exact ship position: Chapter 3, which consists of journal paper $[J_1]$ and $[J_2]$, has its focus on the practical

determination of the added resistance transfer function. For this purpose, the concepts of machine learning and empirical risk minimization are introduced. Moreover, Chapter 4 is devoted to the identification of the sea state using the wave buoy analogy and includes [J_3]. In addition, the fundamentals of deep learning and the inherent uncertainty of sea state estimates are discussed. As indicated in Fig. 1.4, Chapter 5 and [J_4] deal with merging the two initial pillars (or parts) and compare theoretical estimates of the mean added resistance to empirical data obtained from actual in-service data from a fleet of container ships. In the present context, Chapter 5 is considered the scientific application, whereas Chapter 6 represents the industrial application of the previous findings. In the latter chapter and [J_5], the added *frictional* resistance caused by marine fouling is investigated and adaptive machine learning techniques estimate the corresponding power increase. The structure of the thesis and Fig. 1.4 also reflect the chronological order in which the papers have been prepared during the study period. Lastly, Chapter 7 reiterates the findings of the present thesis, outlines the accomplished contributions and provides recommendations for extending work in the future.

Chapter 2

Theoretical Background

An advancing ship experiences a drag force, i.e. the total ship resistance R_T, which can be broken down into several sub-components. The total resistance is crucial for determining the required engine power and propeller characteristics in both calm water and actual operating conditions. As can be seen in Eq. 2.1, the engine brake power P_B provides a thrust via a propulsor for balancing R_T.

$$P_B = \frac{R_T\,U}{\eta_D\,\eta_M} \tag{2.1}$$

From Eq. 2.1, it can be inferred that there is an equilibrium between brake power and effective power, i.e. the product of R_T and the ship's advance speed U, diminished by both propulsive η_D and transmission η_M efficiencies. In particular, η_D is a quantity of utmost importance for evaluating not only propeller properties but also the interaction between the propeller and hull flow. Both η_D and η_M may change considerably under sailing conditions in a seaway but are of minor interest in the present thesis. Extensive information regarding both propeller and engine behavior can be found in Carlton [20]. The propulsive performance in waves has been thoroughly studied in Saettone [107] and Nakamura and Naito [93]. As it stands out, the paramount quantity in Eq. 2.1 is the total resistance R_T, which is transmitted to the vessel by means of pressure p and shear stress distributions, where the latter is expressed by a matrix $\boldsymbol{\tau_w}$.

$$R_T = \int_S p\,n_x\,dS + \int_S \boldsymbol{\tau_w} \times \boldsymbol{n}\,dS \tag{2.2}$$

In Eq. 2.2, $\boldsymbol{n}$ indicates a normal vector and the subscript x stresses that only the longitudinal part of a vector is of interest. The individual hydrodynamic components are eventually determined by integrating both pressures as well as shear stresses over the wetted surface area of the ship S. The pressure resistance component, which corresponds to the first term, is mainly caused by the generation of surface waves and is affected by possible incident waves. The latter term in Eq. 2.2 corresponds to the frictional resistance due to the viscosity of the surrounding fluid and the development of a turbulent boundary layer between the ship hull surface and the far-field. It is noted that each possible sub-component of R_T can be split into both frictional and pressure parts. In addition, both contributions are linked through an interaction resistance - the viscous pressure resistance, which will be

described later. Bertram [10] states that resistance decomposition eases understanding and calculation but generally impedes capturing interaction effects sufficiently.

The reduction in ship advance speed U in waves can be both involuntary and voluntary, where the latter corresponds to the deliberate reduction in speed for the avoidance of, e.g. slamming or green water. Contrarily, the involuntary speed loss refers to the negative effect of the environment on the ship and can also be expressed in terms of a resistance surplus, i.e. an *added resistance*. The present work is limited to the latter perspective. A ship in actual operation is exposed to various additional external forces predominantly caused by wind and waves. Herein, we adopt a two-dimensional coordinate system, depicted in Fig. 2.1, with its origin at the vessel's center of gravity G. Hence, three degrees of freedom (DOF) are considered, i.e. x, y, and N, where the latter denotes rotations around the z-axis, which is pointing downwards. The vessel advances with forward speed U, and the possible influence of a pre-existing incident fluid velocity, i.e. ocean currents, is included in U.
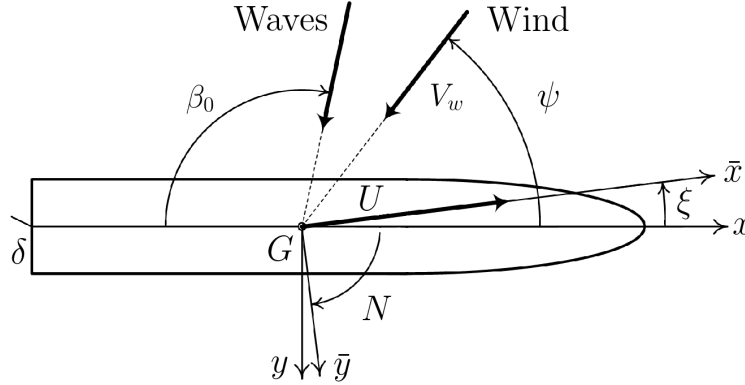
Figure 2.1: Adopted body-fixed coordinate system for a ship and external forces in a realistic environment under forward speed.

From Fig. 2.1, it is appreciated that the wind and wave directions follow a different heading convention. In the case of waves, the relative heading $\beta_0 = 180$ deg. refers to head waves and 0 deg. stands for following waves, whereas it is vice versa for wind directions. It is noted that the added resistance due to wind and waves is symmetric regarding the ship's center plane (given port/starboard symmetry). Additionally, the vessel may exhibit a drift (or leeway) angle ξ, and a rudder angle δ may be applied. In addition, drift forces and moments may act in all mentioned 3 DOF, but the present contribution has a sole focus on the longitudinal force. Shigunov [111] has shown that the added resistance is generally of highest importance under forward speed conditions compared to the other contributions. As another demarcation, the present thesis deals primarily with ship seakeeping and therefore disregards maneuvering aspects to a large degree. Nevertheless, knowledge about the added wave resistance is vital for assessing the maneuverability of ships in waves. For more information regarding ship maneuvering, see Brix [17].
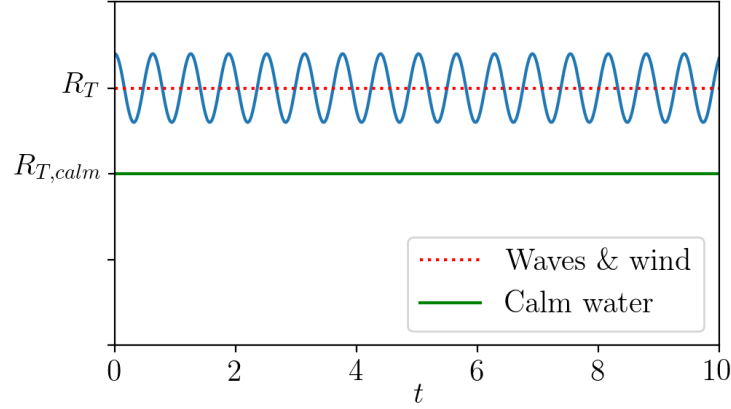
Figure 2.2: Time histories of the ship resistance in calm water and in a seaway considering steady conditions, adapted from Perez Arribas [103].

The decomposition of calm water and seaway-related ship resistance is shown in a simplified manner in Fig. 2.2, and it stated that both are treated independently. As mentioned by Blok [12], in case of added resistance, only the time-averaged value of the oscillations is of importance, as opposed to ship motion analysis. Under the assumption of deep water and steady conditions, i.e. the disregard of added resistance from steering, shallow water, and fouling, the total resistance in a seaway can be expressed through Eq. 2.3.

$$R_T = R_{T,calm} + R_{wind} + \bar{R}_{AW} \tag{2.3}$$

In the following, the three main contributions will be presented in greater detail. The calm water resistance is commonly decomposed into four components under the disregard of the appendage resistance, cf. Eq. 2.4. The fundamentals of resistance and propulsion in calm water can be found in Molland *et al.* [91].

$$R_{T,calm} = R_F + R_{VP} + R_W + R_{AA} \tag{2.4}$$

The frictional resistance R_F is Reynolds number Rn dependent, i.e. results from the viscous property of water and the shear stresses on the hull surface caused by the creation of a boundary layer. Following Froude, R_F can be determined by an equivalent plate, i.e. with the same length and wetted surface area S as the ship. In common practice, the ITTC (International Towing Tank Conference) 1957 correlation line is used for the empirical estimation of R_F. The overall procedure can be found in the 1978 power prediction method by ITTC [59]. In calm water conditions, ship-generated waves can be broken down into primary and secondary wave systems, Bertram [10]. The former has its origin in the pressure peaks at the stagnation points at bow and stern. The viscous pressure resistance R_{VP} links fluid viscosity and form effects of the hull shape. In fact, R_{VP} emerges from the primary wave system and is caused by the fact that the pressure peak in the aft part has a lower magnitude compared to the corresponding level of bow pressure. The resulting pressure difference stems from the loss of momentum within the flow along the ship hull contour caused by the presence of a boundary layer and possible separation bubbles in the stern region or at the transom. The viscous pressure resistance scales with both the Froude number and the Reynolds

number and is seen as an interaction term in Eq. 2.2. Following the ITTC'78 method, R_{VP} is calculated as a function of the frictional resistance R_F using the form factor k. The secondary wave system corresponds to the Kelvin wave pattern, which is transmitted downstream and consists of divergent longitudinal as well as transverse waves. The corresponding wave resistance R_W is connected to the lost energy required for maintaining the wave pattern. In fact, the ship induces multiple wave systems (predominantly in areas of higher pressure gradients) that interact with each other favorably or unfavorably, giving the wave resistance curve its characteristic humps and hollows. Thereby it is understood that R_W contains the wave *pattern* resistance, i.e. the potential (or inviscid) component, and the additional contributions due to wave breaking and spray, Molland *et al.* [91]. The wave resistance is generally determined in model tests and scales with Fn, i.e. it can be directly extrapolated from model to full scale. In early design stages, empirical procedures, by e.g. Guldhammer and Harvald [42], are commonly applied for an estimate of $R_{T,calm}$.

A ship is exposed to a two-phase flow. Even though air density is three orders of magnitudes lower than seawater density, the aerodynamic resistance (including wind) may take up a notable proportion of the total resistance in a seaway, which results from the higher air velocities, and a resistance force is generally proportional to the squared flow velocity. The air resistance R_{AA} of the ship is, in fact, part of the calm water resistance (cf. Eq. 2.4) and is generally of minor importance due to the difference in water and air densities. Typically, the air resistance can be obtained from dedicated wind tunnel tests or is extrapolated from experimental reference data. In contrast to the base flow of the air resistance, the flow profile of the separate wind resistance R_{wind} comprises an atmospheric boundary layer, Faltinsen [31]. According to the industry standard for conducting sea trials, ISO 15016 [57], the wind resistance R_{wind} can be calculated by Eq. 2.5, where the air density is $\rho_{air} = 1.22 \text{ kg/m}^3$.

$$R_{wind} = \frac{1}{2}\rho_{air}A_x(C_{AA}(\psi_r)V_{w,r}^2 - C_{AA}(0\text{deg.})U^2) \tag{2.5}$$

Given Eq. 2.5, it is appreciated that the overall aerodynamic resistance corresponding to the first term is subtracted by the calm water air resistance, which is defined in the second term, to obtain R_{wind}. It is noted that the absolute wind speed V_w and direction ψ, displayed in Fig. 2.1, have to be mapped from the absolute into the relative domain (with subscript r). The reference area A_x corresponds to the projected transverse area of the vessel above the waterline. The wind resistance coefficients C_{AA} are either experimentally determined or through empirical formulae. The industry standard ISO15016 [57], recommends the use of experimental data taken from the STA (Sea Trial Analysis) joint industry project or the regression method by Fujiwara *et al.* [34].

Similarly, the ISO15016 standard recommends various methods for wave correction in a natural seaway. The spectral method for the calculation of the mean added resistance $\bar{R}_{AW}$ builds upon linear superposition and is shown in Eq. 2.6. The Fourier integral has been proposed by Maruo [83] and is considered *the* core equation of this thesis.

$$\bar{R}_{AW} = 2\int_0^{2\pi}\int_0^\infty \frac{R_{AW}(\omega,\beta)}{\zeta_0^2}E(\omega,\beta)d\omega d\beta \tag{2.6}$$

In view of Eq. 2.6, it can be seen that the mean added resistance in short-crested waves $\bar{R}_{AW}$ is determined from the product of a quadratic transfer function, R_{AW}, and a directional wave spectrum, E, which is integrated over both wave frequency ω and wave direction β. It is noted that ζ_0 is the wave amplitude. The following Sec. 2.1 is devoted to the fundamentals of irregular ocean waves and in Sec. 2.2, the added resistance transfer function, and its calculation will be dealt with.

2.1 Ocean Waves

It is said that Lord Rayleigh stated that "The basic law of the seaway is the apparent lack of any law.", Kinsman [69]. But mathematically, any time history of an irregular wave field ζ can be described by the superposition of an infinite number of regular harmonic waves with lengths λ and propagation directions β. The underlying assumptions of the wave elevation are stationarity, ergodicity, and Gaussianity throughout the observation time, emphasizing that merely statistical estimates can be made for a natural seaway. Moreover, it is a prerequisite that the considered water waves exhibit only a moderate steepness, i.e. follow linear theory. In order to determine the energy per square meter of sea surface area with angular wave frequency $\omega = 2\pi f$ and direction β, a directional wave energy density spectrum – or wave spectrum – is introduced. In Eq. 2.7, it can be seen how the directional wave spectrum $E(\omega, \beta)$ relates to the instantaneous wave elevation in the time domain following discrete Fourier analysis in real notation.

$$\zeta(x, y, t) = \sum_j^J \sum_l^L \sqrt{2E(\omega_j, \beta_l)\Delta\omega_j\Delta\beta_l} \times \cos\left[\omega_j t - k_j(x\cos\beta_l - y\sin\beta_l) + \varepsilon_{jl}\right] \qquad (2.7)$$

In Eq. 2.7, it can be seen that E relates to half the squared of the wave amplitude as well as that $\Delta\omega_j$ and $\Delta\beta_l$ denote frequency and direction increments, respectively. In addition, ε_{jl} corresponds to a random phase shift distributed in the range $[0, 2\pi]$ and the wave number k is defined by the dispersion relation in infinite water depth, i.e. $k_j = \omega_j^2/g$. Thereby it is understood that Eq. 2.7 corresponds to the general case of short-crested waves, which resembles the pattern of ocean waves more closely. In common practice, unidirectional or long-crested waves are assumed, and the expression in Eq. 2.7 can be simplified accordingly. The unidirectional wave spectrum F is given in Eq. 2.8.

$$F(\omega) = \int_0^{2\pi} E(\omega, \beta) \, d\beta \qquad (2.8)$$

In Eq. 2.8, it is appreciated that F is obtained by integrating E for all directions β. It is noted that both E and F reflect a short-term description of the sea state and may be obtained by analyzing finite sequences, e.g. 30 minutes, of measured wave elevation data. In typical engineering applications, the sea state can be approximated using idealized design wave spectra, e.g. by Pierson and Moskowitz [104], for obtaining F and a spreading function D is used for distributing the wave energy density around a mean vector, i.e. $E(\omega, \beta) = F(\omega)D(\beta)$. For doing so, integral sea state parameters are required as input

values, and in the following part, the three parameters of most importance for this thesis
will be introduced. The Beaufort scale classifies the sea state according to the prevalent
wind speed V_w and the significant wave height H_s. This sea state parameter is connected
to the energy content of the ambient wave system since it is determined by the zeroth order
spectral moment m_0, as can be seen in Eq. 2.9.

$$H_s = 4\sqrt{m_0} \text{ with } m_0 = \int_0^\infty F(\omega)\, d\omega \tag{2.9}$$

It is noted that the zeroth order spectral moments corresponds to the variance σ^2 of the
initial time series defined in Eq. 2.7. The peak period T_p is extracted from the integrated
wave spectrum $F(\omega)$ and corresponds to the inverse of the frequency where the wave energy
density is largest, see Eq. 2.10. In other words, T_p is an indication for the distribution of
wave energy within a given frequency range.

$$T_p = \frac{2\pi}{\omega_p} \text{ with } F(\omega_p) \equiv \max_\omega F(\omega) \tag{2.10}$$

The peak period T_p (alongside its peak value) and thus the shape of the wave spectrum
is determined mainly by fetch, i.e. the distance over which the wind generates ripples with
a constant velocity and the presence of possible swell waves, which may propagate with
an opposing direction to wind waves. In Eq. 2.11, the mean wave direction μ is obtained
following Longuet-Higgins *et al.* [82].

$$\mu = \arctan(d/c) \text{ with } \left\{ \begin{array}{c} d \\ c \end{array} \right\} = \int_0^{2\pi} \int_0^\infty E(\omega, \beta) \left\{ \begin{array}{c} \sin \\ \cos \end{array} \right\} \beta\, d\omega d\beta \tag{2.11}$$

It is stressed that μ corresponds to the mean direction in a global reference frame and
has a reduced physical relevance in multimodal sea states. For the determination of the
encountered mean wave direction, given in Fig. 2.1, $\beta_0 = \mu - \alpha$ has to be applied, where α
denotes the ship's heading. The intricacies of regular wave theory and the fundamentals of
obtaining E through field measurements can be found in Faltinsen [30], and Kinsman [69],
respectively.

2.2   Added Resistance due to Waves

The added wave resistance R_{AW} is considered the time-averaged longitudinal second-order
force an advancing ship experiences in the presence of incident waves apart from the calm
water resistance. Under non-forward speed conditions, this added resistance and the lon-
gitudinal wave drift force are identical. Following Strøm-Tejsen *et al.* [117], the added
wave resistance can be considered a non-viscous phenomenon and is most pronounced in
the proximity of head waves. Generally, R_{AW} results from the ship's generation of unsteady
waves and is considered proportional to the squared wave amplitude ζ_0, Faltinsen [31]. The
key driving factors of added resistance can be divided into (1) operational conditions, i.e.
wave length λ, amplitude ζ_0, and relative heading β_0 as well as ship speed U and (2) aspects
related to the vessel's hull itself, i.e. main particulars and hull shape. As a result of the ship
size dependency, the quadratic added resistance transfer function is commonly provided for

a non-dimensional frequency and by the coefficient C_{AW}, which is defined in Eq. 2.12. It is noted that L_{pp} is the length between perpendiculars, ρ denotes the water density, g is the gravitational acceleration, and B refers to the beam of the ship.

$$C_{AW} = \frac{R_{AW} L_{pp}}{\rho g \zeta_0^2 B^2} \tag{2.12}$$

According to Kashiwagi *et al.* [66], the added resistance can be decomposed into three sub-components: Diffraction, radiation, and their interaction. The diffraction component is caused by the ship's reflection of incident waves and is of the highest importance in short relative wave lengths λ/L_{pp}. It has been shown by Kashiwagi [65] both experimentally and numerically that the diffraction component of R_{AW} is nearly constant throughout the relevant regime of relative wave lengths considering waves with moderate steepness. However, with higher wave steepness, an increase in diffraction-related added resistance may occur in short waves due to non-linear effects, such as wave breaking. Kreitner [70] argued that the added resistance of ships is solely due to diffraction and derived a corresponding formula. However, Havelock [44] shows in a landmark paper, in which he pioneered calculating second-order wave forces for a vertical cylinder, that the force due to wave reflection may be surpassed by a component due to radiation. The contribution from radiation is linked to the first-order ship motion problem and originates from the lost energy caused by radiated waves due to damping. Strøm-Tejsen *et al.* [117] attribute this primarily to heave and pitch motions, even though other motion components become increasingly relevant in non-head wave conditions, Valanto and Hong [128]. The translational degrees of freedom (DOF) are surge, heave, and sway, whereas the rotational elements of the response vector are roll, yaw, and pitch. Fig. 2.3 illustrates the orientation of the mentioned DOF and the placement of the inertial coordinate system, which is moving with the ship but not experiencing its oscillatory motions.
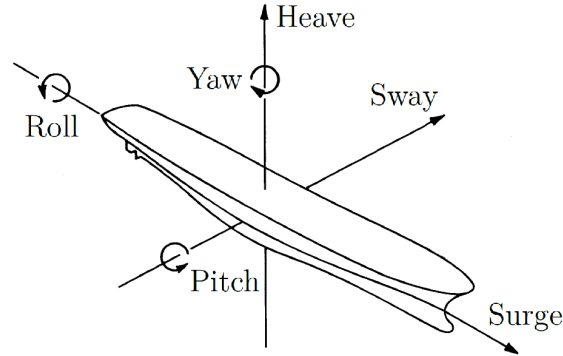
Figure 2.3: Adopted coordinate system with indicated degrees of freedom, adapted from Grim [41].

The radiation contribution is frequency dependent and reaches its maximum when the wave length is in the vicinity of the ship length, which corresponds to the peaks of heave and pitch transfer function. The added wave resistance R_{AW} originates in the lost energy due to waves transmitted from the ship to the far-field. Thereby, it is understood that incident

and ship-generated waves interfere with each other. The ship-generated waves are composed of radiated, diffracted, and Kelvin waves corresponding to the calm-water wave system. Due to these interaction effects, the radiation component may take negative values with larger λ/L_{pp} values, which then leads to an asymptotic behavior of added wave resistance towards zero in this regime, Kashiwagi *et al.* [66]. Hence, the ship generates no unsteady waves in longer relative wave lengths. Referring back to Eq. 2.2, added resistance due to waves contributes almost exclusively to the pressure component of R_T. The viscous added resistance component caused by an increase in wetted surface area and viscous damping – mainly caused by roll – is commonly disregarded in head waves but increases in magnitude in oblique waves.

The added wave resistance can be determined experimentally, numerically, or by regression analysis. Model tests towards R_{AW}, conducted under Fn similarity, were first initiated 60 years ago and aimed at increasing the understanding of ships in seafaring conditions. For this purpose, the resistance is measured in calm water and regular waves with different frequencies at a constant speed. The added resistance is the difference between the time-averaged longitudinal force in waves and the calm water resistance. Measuring the added resistance in model tests without interfering with the ship's motions is delicate. Grim [41] emphasizes that the inherent uncertainty of experimental data is mainly caused by the small magnitudes of added resistance in tank tests. Still, the measurement uncertainty of added resistance is rarely indicated, and the transparent uncertainty bars in the work of Joncquez *et al.* [61] are considered as rare in the existing literature. It can be seen that the reliability decreases drastically in shorter waves, which is due to small magnitudes of measured added resistance and small wave amplitudes. Additional sources of uncertainty of experimental studies towards added resistance are elaborated on in Park *et al.* [100]. Experimental Fluid Dynamics (EFD) results are seen as ground truth for validating numerical and deriving empirical methods.

Numerical approaches towards estimating R_{AW} split into viscid and inviscid methods. The latter pertain to potential flow theory, and the corresponding procedures are divided into far-field and near-field approaches. The former family of methods relies on energy or momentum conservation and considers their flux at infinity. The work that was performed by the ship transmitting unsteady waves to the far-field is determined. Contrarily, near-field methods build on pressure integration for calculating the longitudinal second-order force. Under the assumption of an ideal fluid, pressure integration is expressed in the first term in Eq. 2.2. The steady pressure distribution on the hull's wetted surface can be obtained using the first-order potentials. Mathematically, both approaches should converge to the same results when applied under consistent conditions; however, this is rarely achieved in practice due to different numerical implementations and simplifications. Perez Arribas [103] provides a comprehensive overview of both near- and far-field potential flow methods.

In a path-breaking work, Maruo [83] proposed the first far-field method for the calculation of added resistance for ships with forward speed by using the Kochin function, which allows the expression of the potentials of incident and radiated waves (caused by body motions and scattering) at infinity by integration over the ship's wetted surface. Maruo's theorem is founded on momentum conservation and is valid for all wave lengths and directions. Still, its practical implementation remained challenging and has been further elaborated on

in several succeeding publications, e.g. Maruo [84] and Newman [94]. In fact, Gerritsma and Beukelmann [36] build upon Maruo's theorem and derive a simplified radiated energy method using strip theory. However, Strøm-Tejsen *et al.* [117] report significant discrepancies between corresponding numerical results and model test data. Salvesen [108] presents increased accuracy by using a related far-field methodology and employing the STF strip theory by Salvesen *et al.* [109] for determining the first-order potentials underlining the importance of accurately solving the ship motion problem for the calculation of added resistance. Amini-Afshar [4] found that considering an initially disregarded term in Salvesen's formulation, i.e. the body disturbance, leads to better agreement with experimental data. But still, two-dimensional potential methods, i.e. strip theory, show reduced accuracy in the short wave regime, where the diffraction contribution is dominant. Hence, Kashiwagi [64] proposed his enhanced unified theory combining strip and slender body theory, which enables him to account better for three-dimensional and forward speed influences.

Boese [13] presented the first near-field method and considered a somewhat simplified pressure distribution on an average wetted surface area of the ship under forward speed. Faltinsen *et al.* [32] derived an analytical expression in short waves based on pressure integration only considering the diffraction component and its interaction with the base flow due to advance speed.

Nowadays, linear and (weakly) non-linear three-dimensional panel methods are frequently used for ship seakeeping computations and estimating added resistance. For instance, Joncquez *et al.* [61] employ a boundary element method in the time domain and use both near- and far-field methods for determining the second-order forces and moments acting on a ship – including R_{AW}. Moreover, Liu *et al.* [81] addressed the added resistance problem by a hybrid Rankine source Green's function method considering Maruo's theorem. In addition, Söding *et al.* [119] developed a frequency domain panel method with a non-linear free surface condition and pressure integration for the calculation of added resistance. For the sake of brevity, only three examples of panel methods for obtaining added resistance are presented herein.

Despite the efficiency of potential flow methods, Computational Fluid Dynamics (CFD) receives growing attention in ship hydrodynamics, which may result from the increasing availability of computational resources. These viscid field methods allow the numerical solution of RANS (Reynolds-averaged Navier Stokes) or Euler equations and enable taking non-linearities, such as wave breaking, into account. However, reliable and accurate results are primarily determined by the discretization quality in both time and space. Especially, the short wave application requires a fine discretization of the free surface and small time steps to prevent numerical diffusion and thereby increase computational cost, Sigmund [113]. Not only the added resistance in regular (and irregular) waves has been investigated in arbitrary wave headings using CFD, by e.g. Kim *et al.* [68], but also the behavior of propulsive coefficients in waves, by e.g. Mikkelsen *et al.* [85]. Even the maneuvering behavior of ships in waves can be modeled with sufficient accuracy, as shown by Uharek [126]. However, RANS methods are prohibitively expensive for systematic studies of added resistance in practice, according to Bertram [11].

The ship's hull geometry is a prerequisite when using the above-mentioned methods, which does not hold for empirical regression methods, primarily derived from experimental

data. Jinkine and Ferdinande [60] laid the groundwork with an empirical procedure for the radiation component of R_{AW} by examining cargo ships in medium to long wave lengths. In parallel, Fujii and Takahashi [33] derived a semi-empirical asymptotic formulation for added resistance due to bow wave reflection, which was later adapted and incorporated into the NMRI (National Maritime Research Insititute) formula, Tsujimoto *et al.* [123]. The NMRI formula applies this asymptotic formula as a correction term in short waves to results obtained from Maruo's theorem and is also part of the ISO 15016 [57] standard. Boom *et al.* [14] proposed the STAwave2 method, which employs an extension of the Jinkine and Ferdinande [60] method and a constant short wave correction. Liu and Papanikolaou [79] formulated a semi-empirical method by combining the work of Jinkine and Ferdinande [60] and a simplified version of Faltinsen's asymptotic formula for the prediction of added resistance in *arbitrary* wave heading angles. Moreover, Lang and Mao [74] combine the approach of Jinkine and Ferdinande [60] and the asymptotic formula of the Fujii and Takahashi [33]. It stands out that nearly all presented empirical methods rely on the exact or an adapted approach following Jinkine and Ferdinande [60].

In Fig. 2.4, several methods for the determination of added resistance will be shown, taking the KVLCC2 (KRISO Very Large Crude-oil Carrier 2) as a case ship, their characteristics will be discussed in the following. The main particulars of the KVLCC2 are listed in Tab. 2.1. The individual methods are validated against experimental data taken from Lee *et al.* [77] and Sadat-Hosseini *et al.* [106] in Fig. 2.4. Moreover, the examined conditions correspond to moderate forward speed ($Fn = 0.142$) and head waves under design draft conditions. In fact, a two-dimensional far-field method following Salvesen [108] is depicted (Amini-Afshar [4]), and a three-dimensional near-field method is shown (Söding and Shigunov [118]). Moreover, RANS results from Sadat-Hosseini *et al.* [106] are included as well as three of the above-mentioned semi-empirical methods. It can be seen that the maximum of the non-dimensional added resistance in head waves is located at a slightly larger wave length than ship length and reaches zero asymptotically in longer waves. The agreement for $\lambda/L_{pp} < 1.5$ is considered satisfactory for most methods, except the RANS method provides too high values compared to experimental data. In the proximity of resonance, all methods show an adequate magnitude in non-dimensional added resistance and also matching peak frequencies, except Boom *et al.* [14]. In short relative wave length, the methods exhibit less agreement. The potential flow methods show pronounced underprediction in the diffraction-dominated regime. It is noteworthy that the RANS method shows a similar tendency as the potential flow methods in short waves, and that the empirical approach by Boom *et al.* [14] yields too conservative estimates in the respective regime, cf. Fig. 2.4.

Table 2.1: Main particulars of the KVLCC2 in design conditions.

Definition	Symbol	Unit	Magnitude
Length between perpendiculars	L_{pp}	[m]	320.0
Beam	B	[m]	58.0
Draft	T	[m]	20.8
Block coefficient	C_B	[-]	0.808
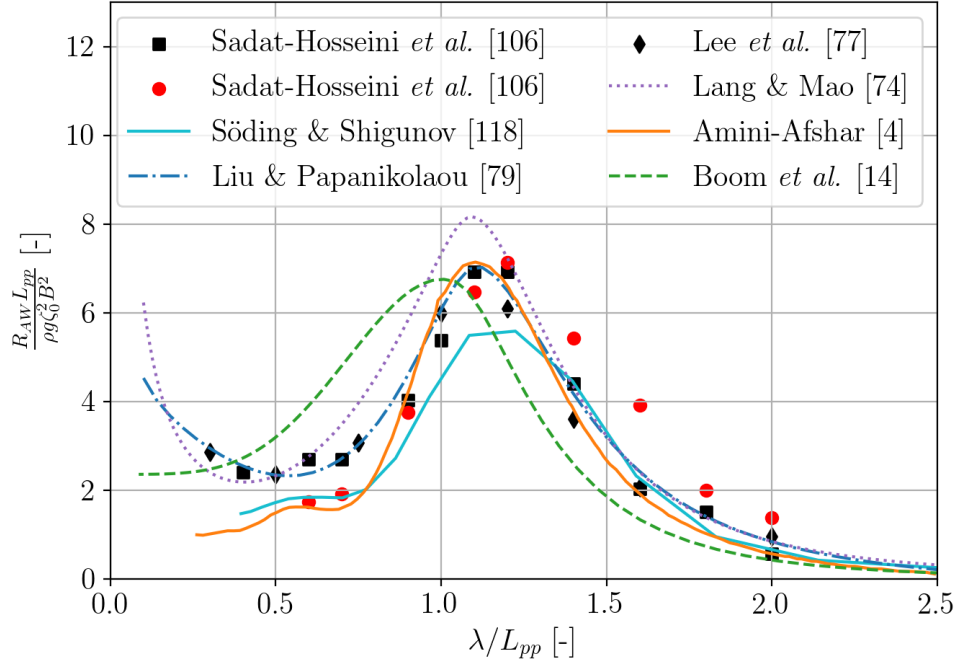
Figure 2.4: Comparison of numerical, experimental and empirical results of added resistance in case of the KVLCC2 for $Fn = 0.142$ and $\beta_0 = 180$ deg. Experimental and RANS data are indicated by black and red dots, respectively.

Kashiwagi [65] showed experimentally by comparing dynamometer and wave cut measurements (of unsteady waves) that a significant fraction of added resistance in the short wave region is caused by non-linear wave breaking effects. Hence, it is well-established that most potential flow theory methods show reduced accuracy in these respective wave lengths requiring empirical corrections. Söding and Shigunov [118] evaluate two asymptotic formulae for the application in short waves and found good performance of the asymptotic method by Faltinsen *et al.* [32] for application in a natural seaway. In contrast, the correction method for bow wave reflection by Tsujimoto *et al.* [123] turned out to be overly conservative in all cases. In Fig. 2.4, it can be seen that both regression methods by Liu and Papanikolaou [79] and Lang and Mao [74] are characterized by a good agreement with model test data in the shorter wave region, which may be a result of the incorporation of the mentioned asymptotic formulae. However, it is not possible to assess their estimates for $\lambda/L_{pp} < 0.3$ since no experimental data is available.

Overall, the agreement of most methods to experimental data is considered insufficient in shorter wave lengths. This has also been concluded in a numerical and experimental benchmark study, which was part of the SHOPERA (Energy Efficient Safe Ship Operation) EU project. In fact, Shigunov *et al.* [112] showed that all considered numerical models in this benchmark study exhibit great uncertainty in short waves, regardless of which methodology was chosen. In addition, it was found that an empirical method gave reasonable results for estimating the longitudinal wave drift force. Regarding experimental data, Sprenger *et al.* [115] indicate that generally good agreement between the individual experimental facilities was found during the SHOPERA benchmark study.

2.3 Concluding Remarks

This section is devoted to outlining the shortcomings of the state-of-the-art literature and emphasizing the goals of the present thesis. Referring back to Fig. 1.2, economy of scale has a tremendous effect on energy efficiency, which led to increasing vessel sizes in recent years. Following the thoughts of Söding and Shigunov [118], the regime of short waves is of utmost importance for larger ships as ocean waves generally exhibit lengths between 50 and 150m. When considering the ship length of the KVLCC2 L_{pp} = 320m, these conditions correspond to relative wave lengths λ/L_{pp} of 0.17 to 0.47. In other words, precisely the regime, where the agreement among all considered methods is lowest, cf. Fig. 2.4.

The added resistance in very short waves is a highly complex problem. For instance, Sprenger *et al.* [115] report based on experimental data that the second-order relationship between added resistance and wave amplitude was only confirmed in longer waves and seems questionable in very short waves (and shallow water). Moreover, for a container vessel in very short head waves ($\lambda/L_{pp} = 0.2$), Sigmund [113] shows that the viscous part of R_{AW} takes up 28.6% of the added resistance in full scale and 35.3% in model scale, which indirectly points at scaling issues in very short waves (assuming CFD results to be a ground truth). The uncertainty surrounding added resistance is a well-known problem but generally neither rigorously reported nor thoroughly investigated. Therefore, this work aims to enhance overall transparency for the prediction of added resistance.

The viscid contribution to added resistance increases not only in short waves but also in oblique waves due to the induced roll motions and the corresponding viscous damping due to vortex shedding (or oscillatory separation). It has been shown in multiple experimental studies that added resistance is largest in bow oblique waves ($\beta_0 = 150$ deg.), e.g. Valanto and Hong [128] or Fujii and Takahashi [33]. Still, most numerical considerations investigate head wave conditions exclusively. Moreover, when considering Eq. 2.6, it becomes evident that the added resistance transfer function must be computed for all heading angles under the assumption of short-crested waves. Hence, the present thesis considers added resistance for arbitrary wave heading.

Interestingly, one procedure for obtaining added resistance was omitted in the previous section: Full scale field measurement campaigns. These systematic studies had significant relevance before the advent of numerical methods, e.g. Kempf [67]. From the contemporary literature, however, it can be seen that there is an imbalance between numerical/theoretical and full scale studies even though data availability of in-service data has been increasing in recent years. For full scale studies, accurate information on the prevalent sea state is indispensable, but significant uncertainties plague associated data sources. Therefore, one chapter of this thesis will be directed toward ship-based sea state identification. In addition, another part is devoted to assessing full scale in-service data of ships for determining added resistance.

Empirical regression methods show not only satisfactory accuracy when predicting added resistance but also direct applicability. Regression is also a core discipline of machine learning and data-driven methods. Caused by increasing availability of computational power and data, machine learning is a quickly evolving research discipline with a wide variety of practical applications. Still, there appears to be a knowledge gap in the field of ship seakeeping

and added resistance in particular. For this reason, this thesis focuses on using data-driven methodologies to evaluate its benefits and possible imperfections when applied in the field of ship hydrodynamics.

Chapter 3

Determination of the Added Resistance Transfer Function

In-service analyses of vessel data suffer from a scarcity of reliable reference and sea state data. Hence, simplified formulae are commonly used in practice for considering the added resistance in seafaring conditions, e.g. Kreitner [70] and Kwon [73]. However, in combination with the neglect of swell waves, as suggested by ISO 19030, this may lead to underestimating added resistance severely. When considering Eq. 2.6, it can be seen that the added resistance is a function of the alignment of the peaks from a transfer function and a wave spectrum. In fact, Liu *et al.* [80] showed that the added resistance in swell waves may surpass the corresponding value in wind waves due to a wave period closer to the resonance region of the R_{AW} transfer function. This is certainly dependent on, e.g. ship size, loading conditions, and prevalent wave period, but most of these factors are not taken into account in simplified formulae. This in turn, underlines the necessity of the spectral formulation in Eq. 2.6 in industrial applications. As described in the previous Chapter 2, the practical and accurate determination of the added resistance transfer function remains challenging. When considering the assessment of ship operations, it becomes clear that exact hull shape information might be unavailable. Thus, the present chapter deals with the efficient and transparent estimation of the R_{AW} transfer function by different regression methods for a wide range of operational conditions.

3.1   Introduction

In 1959, Samuel [110] described machine learning as the "[...] field of study that gives computers the ability to learn without being explicitly programmed.". Mathematically, a machine learning model is well-described as a function $\mathcal{F}$ having an internal structure $\boldsymbol{\theta}$, which may be encoded in a parameter vector, e.g. for neural networks or a set of decision rules in the case of non-parametric tree-based methods.

$$\hat{y} = \mathcal{F}(\boldsymbol{x}|\boldsymbol{\theta}) \tag{3.1}$$

For finding a suitable internal model structure $\boldsymbol{\theta}$, both input vector $\boldsymbol{x} \in \mathbb{R}^n$ and output y must be known in advance. It is stressed that n denotes the dimensions of the feature

space and that a model only provides an approximation of the target variable $\hat{y}$. In the case of numerical tools (and other computer programs), $\boldsymbol{\theta}$ (and $\mathcal{F}$) have to be explicitly defined and provide $\hat{y}$ when given independent variables $\boldsymbol{x}$. In contrast, a machine learning model is found using an implicit approach without any strong assumptions or biases but under the condition of sufficient data availability. The actual model architectures are here of lower interest but will be elaborated on in $[J_1]$ and $[J_2]$.

In contrast to Samuel's relatively informal definition, Mitchell [86] postulates that a machine learning model is asked to "[...] learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E". Paraphrasing Mitchell's definition, a machine learning model builds on three pillars: (1) The model experiences (E) a representation, which corresponds to a dataset $\mathcal{D}$ characterized by an *empirical* joint distribution $\hat{p}_{Data}$. It is, however, vital to underscore that the training data is nothing but a representation of the actual underlying data-generating process with a joint distribution p_{Data}. (2) In the present thesis, the individual models learn from data in a supervised or inductive regression paradigm or task (T), i.e. $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}$. This learning paradigm requires a dataset $\mathcal{D}$ with sample length m composed of independent variables (or features) $\boldsymbol{x}$ and dependent variable (or target) y. For the sake of completeness, the two remaining learning paradigms are unsupervised learning, where algorithms process data without any given label, and reinforcement learning, where an agent, i.e. an algorithm, performs actions in an environment optimizing a reward function. (3) The third pillar of machine learning is the optimization of a performance measure P for obtaining an internal structure $\boldsymbol{\theta}$ maximizing generalization. In fact, P is defined using an unseen test or validation set and is thus optimized indirectly, which is a crucial difference to usual optimization procedures. In Eq. 3.2, it can be seen that the cost function J is defined as an average over the training data with $\mathbb{E}$ as expectation and $\mathcal{L}$ as an arbitrary per-sample loss function.

$$J(\boldsymbol{\theta}) = \mathbb{E}_{(\boldsymbol{x},y)\sim\hat{p}_{data}(\boldsymbol{x},y)}\mathcal{L}(\mathcal{F}(\boldsymbol{x}|\boldsymbol{\theta}), y) \tag{3.2}$$

In the field of machine learning, the term *risk* is of rather theoretical nature and expressed similarly as Eq. 3.2, but with respect to the data-generating distribution p_{Data}. However, in Eq 3.2, the empirical data distribution is utilized, and hence the empirical risk is minimized. In doing so, it is hoped that the actual risk is reduced accordingly. However, overly complex models can "memorize" certain training instances leading to a mismatch between risk and empirical risk, i.e. an increasing generalization error. This is known as overfitting and, hence in practical machine learning, mostly adapted (or regularized) versions of empirical risk minimization[1] are used for training. Further information about machine learning theory can be found in Mitchell [86].

In mathematical folklore, the No Free Lunch theorem indicates that, under certain constraints, it turns out that no optimization algorithm exists that performs universally better than any other, Goodfellow *et al.* [39]. Trivially, the same applies also to machine learning models, which consequently demands sensitivity studies for model type, complexity, and associated hyperparameters for reaching a satisfactory model capacity, i.e. optimal general-

---

[1]In $[J_2]$ on page 7, it must read *empirical* risk minimization.

ization. For this reason, $[J_1]$ has not only a focus on evaluating several different data-driven models but also on hyperparameter optimization. Additional features of this article are assessing the effect of different data preprocessing methods on model performance and comparing parametric and non-parametric machine learning models for the prediction of the added resistance transfer function. For this purpose, the database of the DTU design tool is utilized, which has undergone several iterations and is still in ongoing development. The data used builds upon results from three different potential flow theory methods. Hence, $[J_1]$ has a resemblance to the work of Alexanderson [3], where linear regression is applied to numerical results for the prediction of the radiation component in head waves. Contrarily, the entire quadratic transfer function is predicted in the head-to-beam wave regime in $[J_1]$ using both artificial neural networks and tree-based ensemble methods.

Cepowski [21] approximates experimental added resistance data in head waves employing neural networks. In the final step, an analytical formula is derived from the model parameters and activations. In doing so, it is shown how closely related empirical procedures and machine learning models are. Furthermore, both methodologies are inherently dependent on the training data and thus inherit any possible insufficiencies and uncertainties within the data. Therefore, it is attempted in $[J_2]$ to provide prediction intervals as a trust measure when estimating the R_{AW} transfer function. In fact, the well-established semi-empirical formula by Liu and Papanikolaou [79] is considered a model, and its parameters, i.e. the constants within the individual terms, are optimized (or calibrated) using Eq. 3.2 and various loss functions $\mathcal{L}$. The work has been disseminated at the DNV Nordic Maritime Universities workshop 2022 (Mittendorf *et al.* [89]) and the approach was implemented in the DTU in-house Ship Simulation Workbench (SSW), cf. Taskar and Andersen [120].
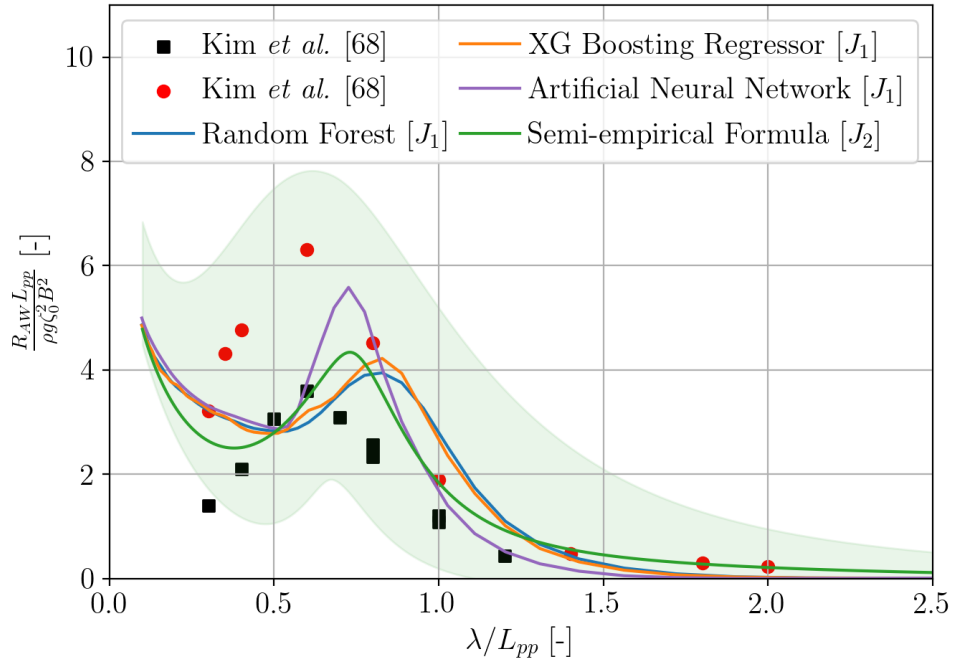

Figure 3.1: LNG carrier in bow oblique waves ($\beta_0 = 120$ deg.) and $Fn = 0.19$ with experimental data (black) and RANS results (red) taken from Kim *et al.* [68].

In the following, all developed methods from $[J_1]$ and $[J_2]$ are demonstrated. Their results are compared to experimental and numerical RANS data from Kim *et al.* [68] for an LNG carrier in bow oblique waves ($\beta_0 = 120$ deg.). The main particulars of the case ship are given in Tab. 3.1.

Table 3.1: Design main particulars of the LNG carrier, Kim *et al.* [68].

Symbol	Unit	Magnitude
L_{pp}	[m]	290.0
B	[m]	45.0
T	[m]	11.5
C_B	[-]	0.77

With attention to Fig. 3.1, a large discrepancy between experimental and RANS results can be observed. In addition, the herein-developed methods show reasonable agreement in longer waves but slightly deviant results in the proximity of resonant conditions, both in terms of peak value and frequency. In contrast, both RANS and EFD data seemingly agree in their peak frequencies, but a significant offset is observed for all considered λ/L_{pp}. In contrast to Fig. 2.4, the developed methods show profound agreement in short relative wave lengths. This is because all methods rely on the asymptotic method by Faltinsen *et al.* [32]. However, the validation in the region of very short waves is not possible due to a lack of experimental data. The prediction intervals improve transparency notably, but one experimental sample falls outside the lower uncertainty bound, indicating a non-conservative estimate in the short wave region. The tree-based ensemble methods, i.e. random forest and gradient boosting machine, occasionally show a lack of smoothness as opposed to the neural network, which exhibits good agreement with the mean of the semi-empirical formula.

3.2   Discussion

In essence, in $[J_1]$, it has been concluded that neural networks can possess a satisfactory generalization capability and robustness when applied to unseen data. Moreover, the significance of data preprocessing and hyperparameter optimization was underlined. Following Tsujimoto *et al.* [123], the added resistance in beam-to-following waves is assumed as zero in the work of $[J_1]$ due to typically small magnitudes, but large variance of the added resistance transfer function in this regime. In fact, similar observations can be made in $[J_2]$. Even though the underlying method of Liu and Papanikolaou [79] has a dedicated formulation for wave headings $\beta_0 < 90$ deg., it appeared that the residual variance was of similar size as compared to the head-to-beam wave region in spite of the smaller R_{AW} magnitudes. As mentioned in the previous chapter, the added resistance in following and stern quartering waves is problematic for all considered methods, and no clear trends can be established in this wave heading regime when analyzing the experimental database in $[J_2]$. Corresponding model tests are associated with large complexities and uncertainties – especially in short waves – due to small magnitudes of measured added resistance and wave amplitude. More-

over, contemporary ship designs exhibit pronounced transom sterns which may lead to flow separation under certain conditions, which leads to problems for simplified methods for estimating added resistance.

In Fig. 3.2, a number of empirical approaches are compared to numerical and experimental data of the Duisburg Test Case (DTC) in stern quartering waves ($\beta_0 = 30$ deg.) and low forward speed ($Fn = 0.052$) being part of the SHOPERA benchmark study, Shigunov *et al.* [112]. The main particulars of the DTC are given in Tab. 3.2, and it is noted that the methods of [J_1] are summarized as DTU Design Tool in Fig. 3.2 as the R_{AW} transfer function is assumed as zero is this scenario. Recently, the DTU design tool[2] and its database have been extended towards the prediction of added resistance in beam-to-following waves, Eftekhar [29].

Table 3.2: Design main particulars of the DTC, Shigunov *et al.* [112].

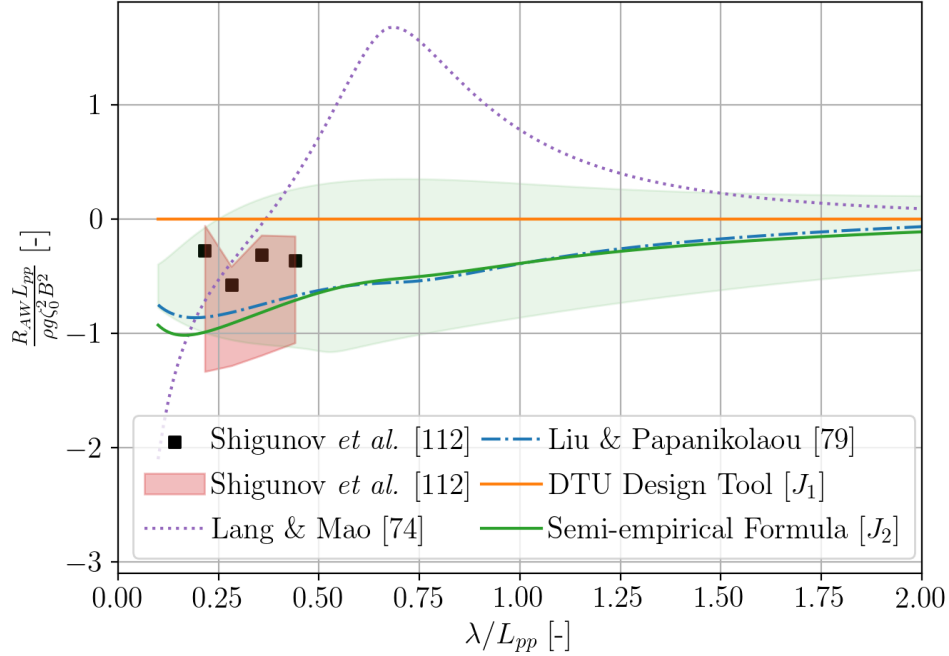| Symbol	Unit	Magnitude
L_{pp}	[m]	355.0
B	[m]	51.0
T	[m]	14.5
C_B	[-]	0.661

Figure 3.2: Model test and numerical data for the DTC in $Fn = 0.052$ and $\beta_0 = 30$ deg. according to Shigunov *et al.* [112]. It is noted that the numerical range is made of the extreme results taken from the numerical benchmark study.

---

[2]`https://gitlab.gbar.dtu.dk/oceanwave3d/DTU_DesignTool`

In view of Fig. 3.2, it can be seen that the herein proposed semi-empirical formula resembles a derivative of the method by Liu and Papanikolaou [79] but is characterized by enhanced transparency due to the prediction interval. Moreover, validating the individual methods is problematic because of the limited coverage of experimental data, which underlines the need for more systematic experimental and numerical studies for these wave headings. It is noteworthy that both the upper and lower bound of the numerical range result from different panel methods. The method by Lang and Mao [74] shows a resonant region in moderate wavelengths and asymptotic behavior in very short waves, as opposed to the other methods. Maruo's theorem shows that the added resistance in following waves depends on the encounter frequency. Simply speaking, the dispersion relation, i.e. the wave's phase velocity is inversely proportional to its frequency, leads in following waves to long waves overtaking the ship and to short waves being overtaken by the ship. Subsequently, the added resistance can take negative and positive values dependent on forward speed and wave frequency. This physical concept is incorporated in the method by Liu and Papanikolaou [79] in a simplified manner. Consequently, it also seems advantageous to introduce this principle into machine learning models for increased accuracy. It is believed that physics-informed machine learning shows strong potential in this respect. Karniadakis *et al.* [62] provide an extensive overview of merging ordinary or partial differential equations and machine learning models.

Overall, it is believed that the predictive performance of the methods presented in $[J_1]$ and $[J_2]$ can be improved by considering additional parameters (or features), such as wave steepness and flare angle above the waterline. In Blok [12], a pronounced correlation between flare angle and experimental added resistance is shown. Regarding wave steepness, Hengelmolen and Wellens [45] show an alternative way of normalizing the added resistance transfer function by wave steepness, which turned out to be beneficial in short waves (for fast semi-displacement vessels). In fact, Sigmund [113] shows numerically that the magnitude of added resistance in short waves has a dependency on wave steepness. Incorporating these additional parameters takes less effort when considering a machine learning method since entire terms of the semi-empirical formula have to be adjusted (or appended) manually.

The trade-off between accuracy and transparency is *the* key issue regarding machine learning. In this respect, it is believed that the generic approach for uncertainty quantification shown in $[J_2]$ is not only applicable to semi-empirical formulae but also in the field of machine learning – given that a custom loss function can be implemented and satisfactory availability of trustworthy reference data.

In the present context, numerical and experimental data have been treated separately, but the application of transfer learning shows considerable potential for further enhancing the accuracy and robustness of machine learning models. In other words, a neural network can be trained on numerical data and afterward retrained (or refined) using experimental data. Obviously, transfer learning is a discipline within the field of deep learning, i.e. a separate branch of machine learning focused on neural networks with multiple hidden layers. In the following chapter, the concept of deep learning will be presented and applied, although on a slightly different topic – on-site sea state identification.

[J_1] Data-driven prediction of added-wave resistance on ships in oblique waves – A comparison between tree-based ensemble methods and artificial neural networks

The paper entitled "Data-driven prediction of added-wave resistance on ships in oblique waves – A comparison between tree-based ensemble methods and artificial neural networks" is published in the Journal of Applied Ocean Research as:

Data-driven prediction of added-wave resistance on ships in oblique waves—A comparison between tree-based ensemble methods and artificial neural networks

Malte Mittendorf [a,*], Ulrik D. Nielsen [a,b], Harry B. Bingham [a]

[a] *DTU Mechanical Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*
[b] *NTNU Centre for Autonomous Marine Operations and Systems, 7052 Trondheim, Norway*

ABSTRACT

The present paper implements machine learning methods for the prediction of the added-wave resistance of ships in head to beam wave conditions. The study is focused on non-linear regression algorithms namely Random Forests, Extreme Gradient Boosting Machines and Multilayer Perceptrons. The employed dataset is derived from results of three different potential flow methods covering a wide range of operational conditions and 18 hull forms in total. The rational data preprocessing makes up the core part of the paper having its focal point on practical application. Moreover, a rigorous hyperparameter study based on Bayesian optimization is conducted, and the validation of the final models for three case studies against numerical and experimental data as well as two established prediction techniques shows satisfactory generalization in case of the neural network. The tree-based ensemble methods, on the other hand, are not able to generalize sufficiently from the given parameter discretization of the underlying dataset.

1. Introduction

The maritime industry is changing: The increased environmental awareness of society and policymakers has lead to the introduction of several rules and regulations incentivizing the enhancement of energy efficiency of ships, e.g. the mandatory compliance of new and existing ship designs with the Energy Efficiency Design Index (EEDI) baseline, International Maritime Organization (2011). Broadly speaking, the EEDI is the ratio of transport work and CO^2-emissions, i.e. it sets the environmental pollution in relation to the socioeconomic benefit, see Kristensen (2010) for greater detail. The legislation of the EEDI framework enforced a trend of reducing installed engine power and slow steaming. The latter (slow steaming) increases the added-wave resistance fraction of the total resistance which in combination with the former makes the risk of losing maneuverability in adverse weather conditions pivotal in ship design and operation. Moreover, the EEDI-equation includes the weather factor f_w incorporating the effect of the added resistance or, rather, the consequent involuntary speed loss in a seaway due to wind and waves. However, the reliable determination of the added-wave resistance and hence of the factor f_w is problematic both when using model test data and results of numerical calculations as has been shown by Shigunov et al. (2018). For this reason, the present work provides a machine learning (ML) approach for the

prediction of the added resistance in waves and compares the model estimates to traditional prediction methods. It is stressed that the ML task itself is considered as a supervised regression approach for an efficient estimate of the added-wave resistance without any detailed hull shape information.

1.1. Literature review

In early ship design stages as well as in the assessment of a vessels' operational performance an efficient prognosis of the added-wave resistance is essential; hence, it is compelling to pursue an empirical regression approach in this respect. Empirical formulas were and are still widely applied in naval architecture, but machine learning regression techniques are also well established in ship design – e.g. Bertram and Mesbahi (2000) – and ship hydrodynamics, especially as surrogate models that approximate results of computationally expensive simulation methods by regression models. For instance, Sclavounos and Ma (2018) estimate ship roll motions and their decay by a Support Vector Machine (SVM), which is a non-linear, instance-based ML-algorithm. Moreover, Mittendorf and Papanikolaou (2021) present a global surrogate-based multi-objective hull form optimization using a genetic algorithm applied to Multivariate Adaptive Regression Splines

---

(MARS) providing the calm water resistance of the case study. Oberhagemann et al. (2019) provide an approach for structural hull monitoring including extreme waves by combining Computational Fluid Dynamics (CFD) calculations with spectral methods and a tree-based ensemble approach by a Random Forest (RF). In contrast to the above-mentioned publications, artificial neural networks or rather multilayer perceptrons (MLP) are predominantly used in state-of-the-art publications for the prediction of added-wave resistance on ships. Cepowski (2020), for instance, trains an MLP on benchmark model test data in head wave conditions and extracts an empirical formula from the resulting model parameters and activations. Moreover, Herradon de Grado and Bertram (2016) train a MLP based on CFD results of the added-wave resistance of one container vessel for headings ranging from head to following waves. In addition, they consider the added-wind resistance by training another MLP on data derived from an empirical formula.

1.2. Motivation and objective

As can be inferred from the presented literature, the individual approaches are either limited to one case ship or to one wave direction. Therefore it is the aim of this paper to provide a more widely applicable surrogate model; both in terms of geometric variation of the hull and with regards to the relative wave direction. The latter is of high relevance as model test data suggests that the added-wave resistance, R_{AW}, will be most prominent in head to bow oblique waves, Valanto and Hong (2015). Another novelty will be the focus on the data preprocessing (feature engineering) and hyperparameter study as these steps are highly influential on the model performance and are not considered in the above-mentioned publications. Moreover, it is envisioned to investigate a variety of machine learning models, and not MLP exclusively, since these are characterized by numerous inherent caveats, such as no scale invariance, general opaqueness and sensitivity to multicollinearity,[1] cf. Hastie et al. (2009). In contrast, non-parametric tree-based ensemble models, such as random forests and (extreme) gradient boosting methods are not affected by the mentioned drawbacks. Hence, this contribution is focused on a comparison between MLP and tree-based ensemble methods applied to the prediction of the added-wave resistance of ships based on a database generated using three different numerical calculation methods in head and bow oblique waves, i.e. in head to beam seas.

Finally, the goal of this contribution is to illuminate the following research question: Do tree-based ensemble methods yield higher accuracy and transparency compared to multilayer perceptrons for the prediction of the added-wave resistance on ships? In this respect, the developed statistical models will be compared to the semi-empirical formula of Liu and Papanikolaou (2020), model test data and RANS (Reynolds averaged Navier Stokes) CFD results for several validation cases. It is stressed that this study is limited to long-crested and regular waves; however, the extension to irregular waves is trivial given a suitable wave energy density spectrum.

1.3. Composition

In Section 2, a brief sketch of the theory behind the added-wave resistance and the applied machine learning models will be provided. Section 3 introduces the underlying database and contains the applied methodology, i.e. data preprocessing, model selection and the overall training configuration. The main findings and results are displayed in Section 4. Moreover, the final models will be applied to three case studies and the models' performances are discussed. Finally, in Section 5, conclusions are drawn and several suggestions for future work are presented.

---

[1] Multicollinearity expresses when two or more predictors are highly related to each other, i.e. do not provide unique or independent information to the regression approach and thus negatively impact model performance.

2. Theoretical background

2.1. Added-wave resistance

The added-wave resistance R_{AW} is the unsteady longitudinal force a ship experiences apart from the calm-water and wind resistances in a realistic seaway. The so-called drift force is of second order and constitutes three components: Wave diffraction, radiation and their interaction. Following Faltinsen (2005), the added resistance in waves results from the ship's ability of generating unsteady waves. A ship in resonant heave and pitch conditions generates the most pronounced waves per unit of wave amplitude due to the large relative vertical motions. For more comprehensive theoretical considerations regarding the added resistance in waves reference is made to the literature, e.g. Strøm-Tejsen et al. (1973).

The added-wave resistance is expressed in non-dimensional form in Eq. (1).

$$C_{AW} = \frac{R_{AW} L_{pp}}{\rho g \zeta^2 B_m^2} \tag{1}$$

In view of Eq. (1), the added-wave resistance is dependent on the ship's length and the beam denoted as L_{pp} and B_m respectively, as well as the wave amplitude ζ, the gravity acceleration is g, and the seawater density ρ. In addition, the ship's draft T_D, the longitudinal radius of gyration k_{yy}, the bow shape as well as the fullness of the hull, i.e. the block coefficient C_B, have a pronounced impact on the magnitude of R_{AW}. Furthermore, operational conditions, such as loading conditions (i.e. GM_T), Froude number $Fn = \frac{U}{\sqrt{gL_{pp}}}$, relative wave direction β, intrinsic wave frequency ω_0 and wave steepness H/λ are also of significant importance.

There are two main approaches towards the calculation of the added-wave resistance using potential flow theory: (1) The far-field methods and (2) the near-field methods. The far-field methods are based on energy considerations for the reflected and radiated waves, the momentum flux at infinity and the work performed by the floating structure. These calculation methods are split further into two sub-methods for the determination of added resistance in waves: Maruo (1957) and Salvesen (1978) proposed approaches based on momentum-conservation, whereas Gerritsma and Beukelman (1971) formulated a far-field method based on the radiated energy method. Both approaches were implemented using the Salvesen, Tuck and Faltinsen (STF) strip theory developed by Salvesen et al. (1970) and are still widely used today. Conversely, the near-field methods determine the steady second order force by integration of the hydrodynamic, steady pressure acting on the wetted surface of the ship's hull. This pressure is be calculated exactly from the first order potential and its derivatives (cf. Boese (1970)). Faltinsen et al. (1980) proposed a near-field approach by an asymptotic formula for the added resistance in short waves which is based on slender body theory. The herein employed dataset consists of results of three different methods in their respective preferred definition range: (1) A Green's Function Method is used for the calculation of the drift force in non-forward speed conditions, (2) results of Salvesen's formulation are employed for the determination of the added-wave resistance in large to medium wavelengths and (3) Faltinsen's asymptotic formula is applied in the region of high wave frequencies. Since details about the dataset are given in Section 3, it is sufficient to mention that higher wave frequencies have the most practical relevance since ships sail the majority of the time in low to moderate sea states, (Faltinsen, 2005). Liu and Papanikolaou (2017) show that relatively short waves $\lambda/L_{pp} < 0.5$ contribute significantly to the added resistance of large ships in seafaring conditions, but are often neglected.

Ultimately, the inherent non-linear characteristics of the added-wave resistance lead to major complexities for theoretical calculations. In fact, the SHOPERA (energy efficient Safe SHip OPERAtion) benchmark study revealed remarkable deviations of numerical, theoretical simulation results and model test data, cf. Shigunov et al. (2018).

Even though RANS methods are, in fact, able to reproduce non-linear characteristics such as wave breaking, the accuracy of these methods is not superior in the above mentioned SHOPERA benchmark study. CFD simulations require a very fine spatial and temporal discretization in order to prevent numerical diffusion of the wave elevation in the free surface region, resulting in a high computational cost. The majority of studies pertaining the added resistance in waves is focused on head wave conditions (e.g. Sigmund and el Moctar (2018) or Sadat-Hosseini et al. (2013)), but recent studies such as Uharek (2019) or Kim et al. (2021) consider relative wave headings $\beta \in [180, 0]$ deg. discretized by 30 deg. increments.

2.2. Machine learning

Mitchell (1997) defines a machine learning task as the following: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E". In the herein presented work, the task T is considered as a multivariate regression problem, i.e. the learning algorithm is asked to approximate the function $f : \mathbb{R}_n \to \mathbb{R}$. The supervised machine learning algorithm experiences E, a dataset $\mathcal{D}\{(\mathbf{x_i}, y_i)\}(|\mathcal{D}| = n, \mathbf{x_i} \in \mathbb{R}^d, y_i \in \mathbb{R})$, where a feature vector $\mathbf{x_i}$ is associated with a known scalar target y_i and the feature matrix is of shape $n \times d$, where n corresponds to the samplesize and d reflects the dimensions of the feature space. Moreover, the approximation $\hat{y}_i$ is obtained by training the model, i.e. minimizing the performance measure P which is specific to the task T. In the following, the fundamentals of the individual learning algorithms are provided. The models are selected such that they are able to handle high-dimensional, non-linear and continuous data.

2.2.1. Artificial neural networks

Neural networks are widely applied as information processing systems and common learning algorithms which are inspired by the governing principles of mammalian brains and their dendrites. These models are directed acyclic graphs and are used as universal function estimators, since the combination of affine functions, i.e. matrix operations, and non-linearities introduced via activation functions, lead to the *Universal Approximation Theorem*, cf. Goodfellow et al. (2016). The most widespread artificial neural network is the multilayer perceptron (MLP), and in its basic configuration it comprises three layers: An input layer, a hidden layer and an output layer; hence, it can be considered as a composite function.

$$\hat{y}_i = f_L[W_L, f_{L-1}(W_{L-1}, \dots, f_1\{W_1, \mathbf{x_i}\})] \tag{2}$$

In Eq. (2), f_i corresponds to the non-linearity or activation function applied in layer L on the weight matrix W_i. It is noted that the usual bias term b is omitted. The parameters or weight matrices W_i are initialized randomly and optimized during training. In general, the training process of neural networks is formulated as a gradient-based optimization problem: The loss function P is obtained by backpropagation and minimized by, e.g. a stochastic gradient descent algorithm. Recently, deep architectures achieved outstanding results in a variety of different disciplines ranging from image recognition to natural language processing. In the herein presented work, relatively shallow MLP architectures are applied for the non-linear regression approach. Further and more elaborate details of neural networks can be found in Goodfellow et al. (2016) and Hastie et al. (2009).

2.2.2. Tree-based ensemble methods

Ensemble methods combine predictions from multiple so-called *weak* machine learning algorithms or base learners for more accurate predictions. The building block of the herein employed ensemble models is the non-parametric Decision Tree. This greedy[2] learning algorithm infers decision rules through features in a supervised environment by splitting the feature space into cuboid regions maximizing the information gain and minimizing impurity in a recursive and progressive fashion. Impurity – or variance in case of regression – refers to the homogeneity of the labels at each node and the information gain is the difference between the parent node impurity and the weighted sum of the child node impurities, i.e. the variance reduction. In this study, the CART (Classification and Regression Tree) implementation of Breiman et al. (1984) is utilized and it constructs binary trees using the feature and threshold that yields the largest information gain at each node. In a regression context it is vital to note that predictions of decision trees are not smooth, but *constant and piecewise approximations*. For this reason, TEM (Tree-Based Ensemble Methods) are known – like most other ML methods – for their insufficient extrapolation capabilities. For more detailed information about regression trees, cf. Hastie et al. (2009) or Bishop (2006).

In this study, two different ensemble methods are considered: (1) In a *bagging* (bootstrap aggregating) approach, a multitude of weak learners are trained in *parallel* on random subsets of the entire training data with replacement. (2) *Boosting*, on the other hand, is a sequential approach and each individual, base learner is influenced by the imperfections of the previous model. Due to the *sequential* manner it allows the definition and minimization of an arbitrary loss function. However, boosting does obviously not reach the computational efficiency of bagging. In the following, additional details and comments about the considered ensemble methods (Random Forests and Extreme Gradient Boosting Trees) are given.

Random forests

Random Forest (RF) Regressors were proposed by Breiman (2001) and grow decision trees in a probabilistic scheme. The algorithm employs decorrelated decision trees in its CART implementation as weak, underfitting learners circumventing the pitfall of overfitting. The output is calculated as the average prediction of the individual decision trees.

$$\hat{y}_i = \frac{1}{M} \sum_{m=1}^{M} \phi_m(\mathbf{x_i}) \tag{3}$$

As can be inferred in Eq. (3), M corresponds to the number of base learners and ϕ_m is considered as the output of the decision tree trained on a bootstrap of the data with a random subsample of the features, i.e. random variable selection at each node. These two sources of randomness decrease the variance of the random forest estimator. In fact, the weak base learners typically show high variance, but the enforced randomness yields decision trees with decoupled or rather decorrelated prediction errors. Hence, by averaging (cf. Eq. (3)) the individual predictions, some errors cancel each other. The variance reduction leads to increased model generalization and robustness to outliers, but might in turn slightly increase the bias.

Extreme gradient boosting trees

Extreme Gradient Boosting Regressors (XGBR) are based on the boosting meta-algorithm, which yields an ensemble model using decision trees as sequential base learners. XGBRs construct an additive, stage-wise model and it based on the optimization of arbitrary, differentiable and regularized loss functions. Gradient descent algorithms are utilized for the creation of new trees based on the residual errors of the previous one, i.e. with the addition of each tree, the model becomes more expressive. In contrast to random forests, each decision

---

[2] Greedy algorithms build a solution to an optimization problem in a heuristic and recursive multi-step manner by making locally optimal choices at each level.

tree contains weights w_j assigned to each leaf j and its value as well as the tree structure itself are determined by minimizing the objective function. The loss function $\mathcal{L}$ comprises of either the $L1$ or $L2$ error and a regularization term penalizing model complexity, cf. Eq. (4).

$$\mathcal{L} = \sum_i l(y_i, \hat{y}_i) + \sum_k \gamma T + \frac{1}{2} \lambda \|w\|^2 \qquad (4)$$

It is noted that l is a differentiable loss function, T corresponds to the number of leaf nodes, whereas γ and λ are used as hyperparameters for regularization. The hyperparameter λ decreases the risk of overfitting, whereas γ penalizes each tree k for growing additional leaves. The optimal w_j is obtained by using a second-order Taylor expansion of the formulated loss function (cf. Eq. (4)). The determination of the optimal tree structure is performed by a greedy method enabling the selection of the best split candidate for adding new leaves. Lastly, the sequential, additive prediction process is evident in the following equation due to the iteration index denoted as t.

$$\hat{y}_i^{(t)} = \sum_{m=1}^{M} \phi_m(\mathbf{x_i}) = \hat{y}_{i,t-1} + \phi_t(\mathbf{x_i}) \qquad (5)$$

The scores w_j and respective tree structure are summarized in ϕ_t in Eq. (5). Only the fundamentals of XGBR were conveyed herein, detailed information, such as split finding algorithms, weighted quantile sketch and system design are presented more elaborately in Chen and Guestrin (2016). Lastly, the XGBR algorithm is versatile and works well on structured data; however, its use is problematic for sparse and dispersed data. In comparison to the random forest, the XGBR algorithm has more adjustable hyperparameters requiring more computational effort during the parameter study. Ultimately, this algorithm does not only reduce variance – similarly to the random forest – due to the ensemble approach, but in most cases also the bias due to its sequential approach as every decision tree is affected by the imperfections or errors of the previous one.

3. Data and methods

3.1. Dataset

The underlying data of the DTU Design Tool, cf. Nielsen (2015) and Martinsen (2016), was used for the training setup of the machine learning models. The DTU design tool aggregates results of three different calculation methods for the determination of the added-wave resistance R_{AW} in irregular waves and employs linear interpolation on a set of quadratic transfer functions. As mentioned in Section 1, the momentum conservation method by Salvesen (1978) was used for low to medium wave frequencies, and the pressure integration method by an asymptotic formula proposed by Faltinsen et al. (1980) was utilized for the high wave frequencies. The required quantities, such as added mass and damping matrices as well as first order potential were computed by the strip theory solver I-Ship Petersen (2000), which is based on the STF method (Salvesen et al., 1970). Lastly, the calculations for zero-forward speed were carried out using the Greens Function panel method WAMIT (Lee and Newman, 2013). The main particulars of the target ship and the parameters of a parametric wave energy spectrum are then specified as an input to the MATLAB tool which performs the interpolations.

The program constitutes two sets of hull variants: One slender, relatively fast container vessel type ($C_B = 0.58$) and a blunt, relatively slow bulk carrier or tanker type ($C_B = 0.84$). It is stressed that the results for the two parent hulls are considered as separate datasets in the presented work. In addition, both datasets are identical considering their shape and derivation by numerical tools. One database is set up for nine scaling variants totaling 18 different hulls and their non-dimensional main particulars are presented in Table 1. The ratios are selected for minimizing the epistemic uncertainty, i.e. the entire design space is taken into account. The hull #9 may appear to be unrealistic,

Table 1
The scaling variants of the parent hulls from Martinsen (2016).

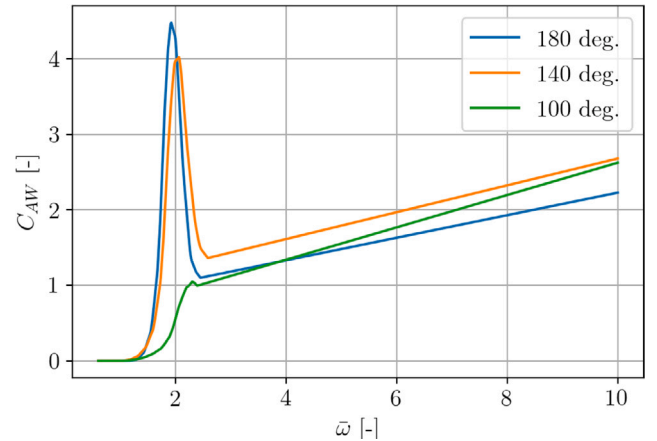|	L_{pp}/B_m	B_m/T_D	L_{pp}/T_D
#1	4.0	2.0	8.0
#2	4.0	3.5	14.0
#3	4.0	5.0	20.0
#4	6.0	2.0	12.0
#5	6.0	3.5	21.0
#6	6.0	5.0	30.0
#7	8.0	2.0	16.0
#8	8.0	3.5	28.0
#9	8.0	5.0	40.0

Fig. 1. Sample plot of hull #1 ($C_B = 0.84$) at $Fn = 0.1$ of the non-dimensional added-wave resistance C_{AW} and for various relative wave headings β.

however, the ratios of the main particulars of contemporary cruise ships closely match the selected values.

On a sidenote, all nine blunt-type scaling variants share the same longitudinal radius of gyration $k_{yy}/L_{pp} = 0.25$, block coefficient $C_B = 0.84$ and waterplane coefficient $C_{WP} = 0.92$. The same holds for the slender container ship hulls with the following characteristics: $k_{yy}/L_{pp} = 0.25$, $C_B = 0.58$ and $C_{WP} = 0.79$. It is noteworthy that the scaling combinations are chosen to reflect the variety of the current world fleet, Nielsen (2015). The database consists of a large set of quadratic transfer functions given for the non-dimensional wave frequency $\bar{\omega} = \omega_0/\sqrt{g/L_{pp}}$ discretized into 941 values and $\bar{\omega} \in [0.6, 10.0]$.

In view of Fig. 1, it becomes clear that assembling the database using three different calculation method leads to discontinuities in the slope of the curves and somewhat arbitrary transitions posing a challenge to the present regression approach. Especially, since the intersection points of the individual quadratic transfer functions between Faltinsen's formula and Salvesen's method lead to inconsistencies, as in some cases the two curves do not intersect at all. In these rare cases, the Salvesen estimate is used instead for C_{AW} in short waves, (Martinsen, 2016).

Six different ship speeds have been taken into account considering Froude numbers in the interval $Fn \in [0.0, 0.25]$ with increments of 0.05. The wave direction $\beta \in [90, 180]°$ is discretized into 10 directions covering head to beam seas. It is stressed that the added resistance for beam to following seas is assumed to be zero. Moreover, the slenderness ratio $L/\nabla^{1/3}$, where ∇ corresponds to the ship's volume displacement, and the following ratios L_{pp}/B_m, B_m/T_D and L_{pp}/T_D are included in the dataset. Ultimately, each database has a sample size of $n = 508,140$ and one exemplary and properly non-dimensionalized sample in a ML setup has the following shape $\mathbf{x_i} = [L_{pp}/B_m, B_m/T_D, L_{pp}/T_D, L_{pp}/\nabla^{1/3}, \beta, Fn, \bar{\omega}]$ and $y_i = C_{AW}$.
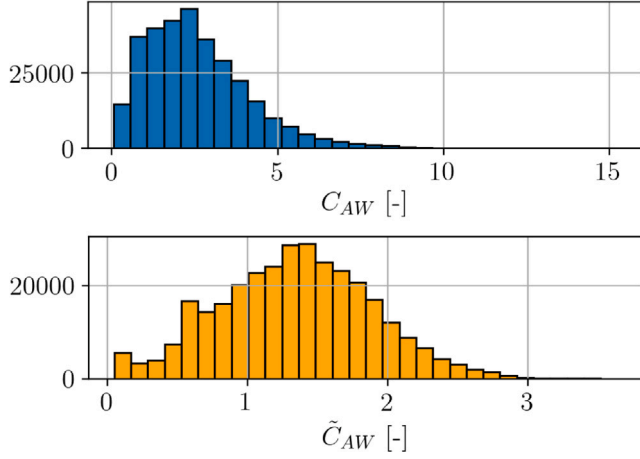
Fig. 2. The application of the Power Transform to the target distribution in black leads to the transformed one in orange. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Methodology

The applied methodology of the presented work consists of three parts: The initial data preprocessing and the model selection part in which the hyperparameter optimization is conducted. The final model assessment is carried out with respect to accuracy, generalization capability, and transparency. From a statistical point of view, the dataset has three distinct characteristics: (1) The target distribution is highly skewed, (2) the behavior of the target is non-linear and (3) the dataset is - taking both databases into account - with a samplesize of $n > 10^6$ relatively rich. Hence, the learning algorithms were selected such that they are capable of handling non-linear and high-dimensional, tabular data. It is noted, that the data preprocessing and model selection part is only carried out for the $C_B = 0.84$ dataset and the findings are adopted for the second database ($C_B = 0.58$).

A transformation of the dependent variable y_i is required as the original target distribution is biased and unbalanced towards lower C_{AW} values (cf. Fig. 2). The Gaussianity or normality of the target distribution is a presupposition of linear regression techniques in statistical modeling and has also a beneficial effect on the performance of non-linear machine learning methods. Logarithmic transforms are monotonic and do not change the locations of optima and also lead to far more stable numerical algorithms. Thus, prior to the training process, the target C_{AW} was transformed according to the Yeo and Johnson (2000) power transformation to $\tilde{C}_{AW}$. The Yeo-Johnson transformation was chosen over the Box–Cox transformation, since it does not impose any restrictions on y_i. The Yeo-Johnson transformation maximizes the marginal likelihood function by the modification of the hyperparameter λ and is defined in the following Eq. (6).

$$y_i^{(\lambda)} = \begin{cases} ((y_i + 1)^{\lambda} - 1)/\lambda & \text{if } \lambda \neq 0, y \geq 0 \\ \log(y_i + 1) & \text{if } \lambda = 0, y \geq 0 \\ -[(-y_i + 1)^{(2-\lambda)} - 1]/(2 - \lambda) & \text{if } \lambda \neq 2, y < 0 \\ -\log(-y_i + 1) & \text{if } \lambda = 2, y < 0 \end{cases} \quad (6)$$

It is stressed that the used transformation is more versatile than the well-known Box-Cox transformation due to the treatment of negative values. After the first calculation, λ was set to $\lambda = 0.1814$ in order to maintain reproducibility. Such a low λ value indicates that the optimal transformation is quite close to the logarithm. The effect of the Yeo–Johnson transformation on the specific target distribution can clearly be seen in Fig. 2.

Fig. 2 shows that the initial target distribution (black) has been transformed into a symmetric and asymptotic distribution (orange)

which resembles a standard normal distribution. It is stressed that the transformation does not impact values equal to zero and thus these are disregarded in Fig. 2 for the sake of clarity. Neural networks are not scale invariant, hence a transformation is necessary for their application. On the other hand, tree-based ensemble methods are scale invariant, but their performance is negatively affected by dispersed data. For this reason and in order to maintain a consistent training methodology across all models, transformed values were fed into both model types. Due to the evenly discretized dataset only normalization, i.e. scaling according to the extreme values, is applied to the feature space.

Furthermore, the polynomial feature expansion up to second order was applied, in order to capture the non-linear behavior of the target by extending the initial set of predictors. It is common practice to generate additional feature incorporating non-linear interactions between the independent variables $\mathbf{x_i}$. Hence, the dimensions of the feature space is increased to $d = 35$ due to the linear combination and squaring of all features. In general, the polynomial expansion is a well-established approach when applying linear regression to non-linear data, and it does also have a positive impact on the performance of non-linear regression algorithms.

Feature importance measures are examined for the examination of the models' transparency and the effect of the feature expansion/ elimination. The TEM provide an intrinsic estimate by the mean decrease impurity (MDI) method which was proposed by Breiman (2001). MDI determines each feature importance as the decrease in node impurity weighted by the probability of reaching that node. The probability can be obtained by dividing the number of samples reaching this node by the total samplesize. Averaging the estimates of the predictive ability over several decision trees leads to a variance decrease of the importance measure and thus it can be used for feature selection. Since neural networks are in general considered as opaque and do not provide a built-in feature importance estimate, the permutation feature importance method is applied. Permutation or mean decrease accuracy (MDA) feature importance is an iterative, model-agnostic algorithm and indicates the impact of one particular feature on the model score by disregarding it during training. The method was proposed by Breiman (2001) for random forests and it is considered trivial that the computational effort scales with the dimensions of the feature space. However, the estimates are not affected by multicollinearity as in the case of the MDI approach. Ultimately, it is stressed that both methods are not directly comparable, but complement each other.

The increase in dimensions of the feature space leads to sparsity and machine learning algorithms are negatively affected by the so-called *curse of dimensionality* (for mathematical details cf. Hastie et al. (2009)). Hence, the number of features will be reduced to the most important subset in three iterations: (1) Initially, the most unrelated features are dropped based on a chosen threshold in view of the individual feature importances. (2) The tree-based methods 'vote' via recursive feature elimination by their cross validation score on the discard of features, noticing that Recursive Feature Elimination (RFE) was proposed by Guyon et al. (2002). It is a selection method that fits a model and removes the weakest feature until the loss function converges. (3) The previous method is only based on the TEM and thus on the MDI feature importance estimate which is affected by collinear features, which will inevitably result from the applied feature expansion technique. The performance of MLP are negatively influenced by the occurrence of multicollinearity. For this problem, the iterative variance inflation factor (VIF) method is used for eliminating highly correlated predictors. The VIF is defined in Eq. (7).

$$VIF = \frac{1}{1 - R_i^2} \text{ with } R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y}_i)^2} \quad (7)$$

The VIF approach was proposed by Lin et al. (2011) and is a measure for the increase in variance of ordinary least squares (OLS) parameter estimates if an additional variable i is added to the linear

regression model. In Eq. (7), R^2 corresponds to the coefficient of determination and the overbar indicates the mean. The VIF is a measure of multicollinearity and the threshold for the VIF of each feature is set to $VIF < 10$.

The dataset does not require any further preprocessing such as filtering or imputation as the data shows only minor variance and no missing values. Prior to the training phase, the modified dataset is resampled and split into a training and a test set, where the test set makes up 20% of the initial data. The hyperparameter study as well as the model selection is only carried out on the training set and the considered metrics are the root mean squared error (RMSE) (cf. Eq. (8)) and the mean absolute error (MAE) which is defined in Eq. (9). It is stressed that both metrics have the same unit and smaller values indicate more accurate predictions.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{8}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{9}$$

The RMSE puts more weight on outliers and is sensitive to variance, whereas the MAE expresses the magnitude of the error regardless of the direction and thus is more robust to noise.

The *No Free Lunch Theorem* (cf. Goodfellow et al. (2016)) states that one cannot assume, that one particular architecture or parameter combination performs better on a variety of tasks; hence, it is of paramount importance to conduct a hyperparameter optimization. Instead of a uninformed method, such as exhaustive grid search, an informed and thus more elegant optimization approach by Bayesian optimization was utilized. Bayesian optimization is a sequential, semi-supervised method employing Bayes theorem for adaptively approximating the objective function and obtaining the optimum without considering any gradient information, but a surrogate of the function of interest by a Gaussian process. The algorithm requires an exploration as well as an exploitation phase and initially a prior is forced onto the loss landscape. After the function evaluation, or update, the posterior is used for adjusting the acquisition function. The next candidate is queried based on the value of the acquisition function. Bayesian optimization is advantageous for objective function that are expensive to evaluate, like the model score during hyperparameter optimization. The herein applied optimization procedure was set up as a minimization problem of the RMSE of a holdout set (20% of training data) with ten initial samples (exploration), 20 exploitative iterations maximum and the model type was a Gaussian Process with Expected Improvement as acquisition function. The overall objective was to find the optimal model capacity, i.e. the tradeoff between variance and bias. For more elaborate details about global Bayesian optimization cf. Mockus (1989).

All calculations were performed on a virtual machine on a server equipped with two Intel Xeon processors (2.2 GHz), 13 GB RAM and an NVidia Tesla K80 GPU. The utilized machine learning frameworks were scikit-learn (Pedregosa et al., 2011), XGBoost (Chen and Guestrin, 2016), GPyOpt (The GPyOpt authors, 2016) for the hyperparameter tuning and statsmodels (Seabold and Perktold, 2020) for the VIF procedure. It is stressed that both scikit-learn and XGBoost employ an optimized version of the CART algorithm for the TEM.

4. Results and discussion

4.1. Feature selection

In the present study, the *Pearson* correlation coefficient is neither suited for data exploration nor for feature selection, as it is only capable of expressing linear interdependencies. To pinpoint the *non-linear* driving factors of R_{AW}, according to the chosen models, the features are presented according to their relevance obtained from the already described mean decrease impurity (MDI) and mean decrease accuracy
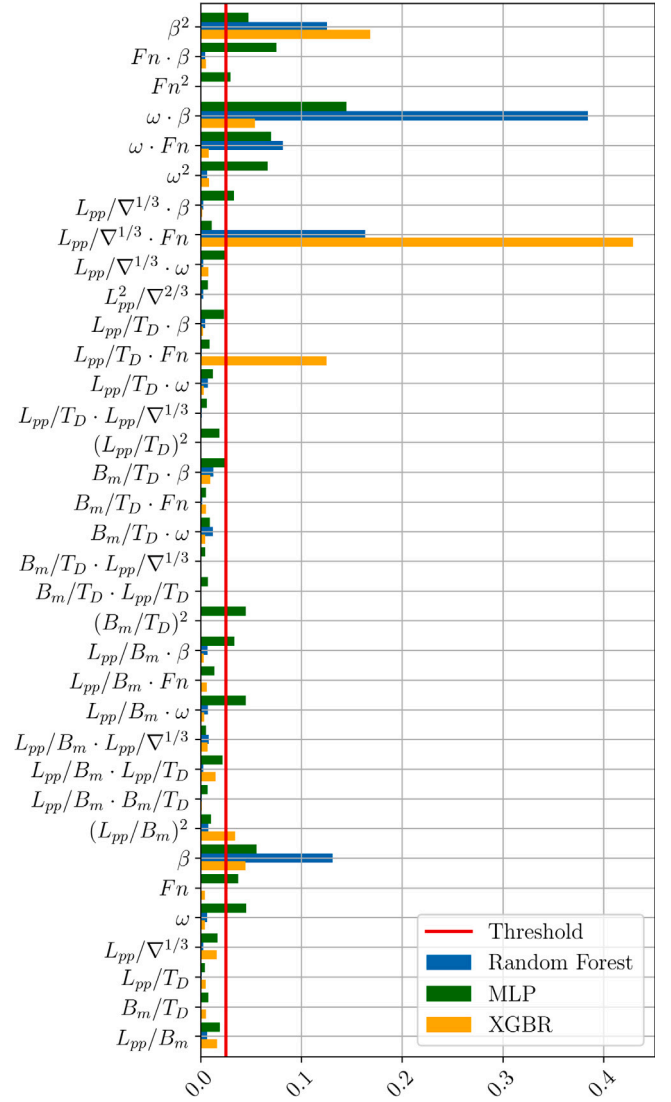

Fig. 3. Individual feature importance estimates for the expanded dataset including a threshold (red) for feature selection. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(MDA) method. The threshold for the *sequential* feature selection is set to 2.5% in the first iteration and is given in Fig. 3 alongside the individual feature influences on the model performance. It is noted that all values of one model sum up to one and that the results are computed using the libraries scikit-learn (Pedregosa et al., 2011) and XGBoost (Chen and Guestrin, 2016).

In view of Fig. 3 it is appreciated that the models' estimates agree mostly on the relevance of the predictors. Moreover, exclusively investigating the initial set of features (i.e. the original database) it is clear that the operational conditions, such as β, Fn and $\bar{\omega}$ are most influential for the models' predictions. Furthermore, Fig. 3 shows that the polynomial expansion leads on one hand to highly expressive variables as the linear combinations (e.g. $\bar{\omega}\beta$), but also the squared versions (e.g. β^2) are of great relevance for the models. Hence, the overall tradeoff between model accuracy and transparency is underlined as it can be seen that the machine learning models choose unphysical, highly abstract features over physical and human-readable ones. But on the other hand, the feature expansion also leads to multicollinearity. An indication for that are the discrepancies between the MDA and MDI estimates (cf. Fig. 3), where the latter
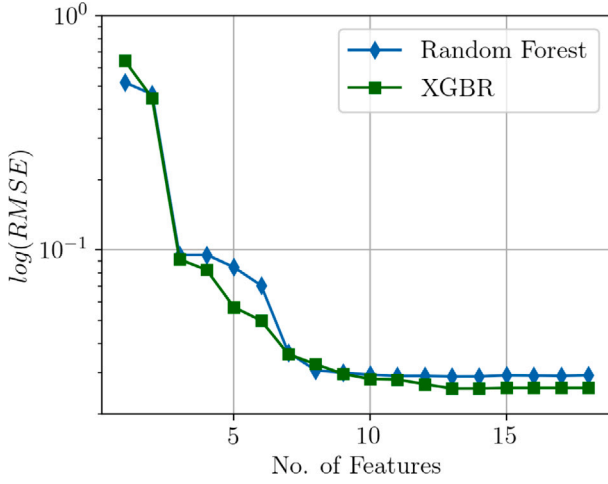
Fig. 4. Convergence study of the $RMSE$ of the validation set applying Recursive Feature Elimination (RFE) for the tree-based ensemble methods.

Table 2
Results of VIF obtained after iterative multicollinearity analysis.

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
β^2	19.16	18.49	11.83	11.5	7.12	2.04
$\bar{\omega}\beta$	7.09	6.98	6.97	6.66	6.65	6.64
$\bar{\omega}Fn$	1.35	1.35	1.35	1.35	1.35	1.31
$\bar{\omega}^2$	3.4	3.37	3.36	3.22	3.22	3.05
$L_{pp}/\nabla^{1/3}Fn$	8.81	8.20	8.2	8.06	8.04	5.4
L_{pp}/TFn	6.36	5.69	5.69	5.57	5.55	3.53
$(B_m/T_D)^2$	7.26	6.33	6.28	6.01	0.85	0.51
$L_{pp}/B_m\bar{\omega}$	21.22	19.34	19.32	5.17	5.17	5.14
$(L_{pp}/B_m)^2$	9.69	9.61	2.12	2.06	2.07	1.85
β	11.42	11.12	11.05	11.05	10.97	elim.
$B_m/T_D\beta$	50.47	15.42	15.33	14.98	elim.	elim.
$L_{pp}/\nabla^{1/3}\bar{\omega}$	27.6	23.98	23.98	elim.	elim.	elim.
$L_{pp}/B_m\beta$	137.36	25.24	elim.	elim.	elim.	elim.
$L_{pp}/\nabla^{1/3}\beta$	222.38	elim.	elim.	elim.	elim.	elim.

is affected by multicollinearity. These differences are visible for instance for $L_{pp}/\nabla^{1/3}Fn$ or $\bar{\omega}\beta$. Ultimately, the MDI estimates show general agreement with the MDA feature importance of the multilayer perceptron (MLP), underlining the sufficient transparency of the tree-based ensemble methods (TEM) at much lower computational cost. The positive effect of the feature expansion was proven and the dimensions of the feature space were reduced according to the chosen threshold of 2.5% from $d = 35$ to $d = 18$. Considering Fig. 3, the feature vector $\mathbf{x_i}$ has the following shape after the first iteration: $\mathbf{x_i} = [\beta^2, Fn\beta, Fn^2, \bar{\omega}\beta, \bar{\omega}Fn, \bar{\omega}^2, L_{pp}/\nabla^{1/3}\beta, L_{pp}/\nabla^{1/3}Fn, L_{pp}/\nabla^{1/3}\bar{\omega}, L_{pp}/T_DFn, B_m/T_D\beta, (B_m/T_D)^2, L_{pp}/B_m\beta, L_{pp}/B_m\bar{\omega}, (L_{pp}/B_m)^2, \beta, Fn, \bar{\omega}]$. It is stressed that the described first iteration facilitates the second iteration and saves computational effort considerably.

The second feature selection iteration employs recursive feature elimination (RFE), i.e. it is based on the individual model MDI estimates of the TEM. For this purpose, the scikit-learn (Pedregosa et al., 2011) function including cross validation was employed. The convergence of $RMSE$ of both models applied on a cross validation set is presented in Fig. 4 for the increasing number of variable dimensions d.

Considering Fig. 4, it shows that there is good overall agreement between random forest (RF) and extreme gradient boosting regressor (XGBR) in their convergence behavior and the sensitivity to the individual features. According to both models $d = 14$ is optimal and Fn, Fn^2, $Fn\beta$ as well as $\bar{\omega}$ were excluded from the feature space based on the ranking that results from the feature elimination procedure. It is interesting that operational features, such as Fn and β, are disregarded in this step, even though they appear as highly relevant in Fig. 3.

Subsequently, the application of the variance inflation factor (VIF) procedure is carried out, in order to eliminate highly correlated and thus redundant features. The feature with the highest VIF is dropped and the procedure is conducted again with the updated set of features until all features exhibit $VIF < 10$. The respective VIF values for all predictors are presented for all 6 necessary iterations in Table 2. It is noted that the statsmodels library (Seabold and Perktold, 2020) was used for computing the variance inflation factor.

Lastly, the final feature vector has the following shape: $\mathbf{x_i} = [\beta^2, \bar{\omega}\beta, \bar{\omega}Fn, \bar{\omega}^2, L_{pp}/\nabla^{1/3}Fn, L_{pp}/T_DFn, (B_m/T_D)^2, L_{pp}/B_m\bar{\omega}, (L_{pp}/B_m)^2]$. As only features resulting from the polynomial expansion are part of the final feature space the relevance of this step in the data preprocessing step is emphasized. Carrying out the feature selection methodology, the expanded feature space is reduced from $d = 35$ to $d = 9$ (cf. Table 2) positively affecting computational efficiency. The possible beneficial effect on model performance of the presented feature selection is studied in the following part.

4.2. Model selection

In this section, the final model characteristics are determined using Bayesian optimization. Moreover, the individual performances of the methods will be assessed in use of the test set, but initially the baseline architectures of the considered models are presented. It is stressed that the naming convention of the employed software packages are used herein.

The baseline random forest is characterized by 100 estimators, whereas the hyperparameter `min_samples_split`, which represents the minimum number of samples required to split an internal node, is set to 2. Increasing this parameter, the tree becomes more constrained as it has to consider more samples at each node. The parameter `min_samples_leafs` is defined as 4 and it is similar to `min_samples_splits`; however, the latter describes the minimum number of samples at the leafs. The maximum depth of the tree `max_depth` is equal to 10 indicating how deep the tree can be built. The deeper the tree, the more splits and thus the more information can be captured. Furthermore, the number of features `max_features` to be considered while searching for the best split is specified as d - the number of features. In addition, either the $L1$ or $L2$ error at the leafs is minimized.

The initial XGBR model is characterized by 100 n_estimators, a linear loss function and a `learning_rate` of 0.1. Moreover, `max_depth` is set to 10 and the regularization term $\gamma = 0.0$. The parameter `colsample_by_tree` denotes the subsample ratio of features for each split at each level and is thus comparable to `max_features`. Also, `min_child_weight` defines the minimum sum of weights w_j of all samples required in a child node, i.e. related to `min_child_leaf` and capable to prevent overfitting.

The baseline MLP architecture comprises one hidden layer with 100 neurons including $L2$ regularization as well as a ReLU (Rectifier Linear Unit, $x = max(0, x)$) activation following input and hidden layer. The output neuron is activated by a linear function and the Adam (Adaptive Moment Estimation) optimizer (cf. Kingma and Ba (2015)) minimizes the $L2$ error w.r.t. the learning rate.

Throughout the parameter study, the `hidden_layer_sizes`, the `activation_function` as well as the $L2$ penalty or regularization term α are optimized for the MLP. All mentioned hyperparameters of RF and XGBR are variables during optimization. The baseline and final model hyperparameters as well as their definition range are provided in Table 3. The default values are adopted from the chosen software packages scikit-learn (Pedregosa et al., 2011) and XGBoost (Chen and Guestrin, 2016), whereas the ranges are located around them.

The Bayesian optimization was fed with 10 initial explorative samples and ran for 20 iterations during the exploitation optimizing the $RMSE$ of a cross validation set. In view of Table 3, it is evident that the resulting parameters mostly correspond to either the upper parameter bound or the baseline value. For instance, n_estimators

Table 3

Hyperparameters of the three examined models including baseline and bounding values.

|	Hyperparameter	Baseline	Range	Result
RF	n_estimators	100	[50, 150]	150
	max_depth	10	[1, 30]	30
	min_sample_leafs	4	[1, 5]	1
	min_sample_split	2	[2, 10]	2
	max_features	d	$\{d, \sqrt{d}, \log_2(d)\}$	d
	criterion	MSE	$\{MSE, MAE\}$	MSE
XGBR	n_estimators	100	[50, 150]	150
	learning_rate	0.1	[0.05, 0.1]	0.1
	max_depth	10	[1, 30]	20
	min_child_weight	1	[1, 10]	1
	γ	0	[0.05, 5]	0.1
	colsample_by_tree	1	[0.5, 0.8]	0.8
MLP	hidden_layer_sizes	(100)	$\{(100), (50, 50, 50), (50, 100, 50), (100, 100)\}$	(50, 100, 50)
	activation	ReLU	{ReLU, sigmoid, tanh}	ReLU
	α	10^{-4}	$[10^{-4}, 0.05]$	10^{-4}
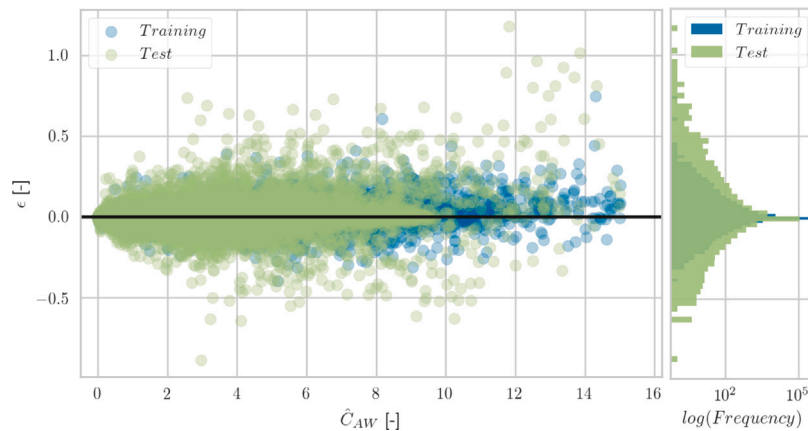
Fig. 5. Residuals plot for the RF on the $C_B = 0.84$ database. Note that the histogram (right-side plot) of the residuals presents the occurrence frequency in logarithmic scale.

in case of the TEM and `hidden_layer_sizes` reach their upper constraint after the optimization for the largest model complexity. On first sight this fact may increase the risk of overfitting (insufficient model capacity), but the cross validation score is taken into account in the parameter study in order to prevent overfitting. The possible occurrence of overfitting and the overall model performance will be assessed in the following.

In Figs. 5–7 the residuals ϵ of the models' predictions are presented for both the training and test datasets, accompanied by the histogram of the residuals in logarithmic scale. It is stressed that only the results for the $C_B = 0.84$ dataset are presented herein and the corresponding results for the $C_B = 0.58$ dataset are placed in the appendix. It is evident that all three models are biased towards underpredicting the non-dimensional added resistance C_{AW} as the histograms are shifted towards positive ϵ values. Furthermore, considerable heteroskedasticity is noticeable for all three models, i.e. an increase in variance with increasing magnitude of the target variable. This phenomenon is also observed in the residual analysis of the semi-empirical formula of Liu and Papanikolaou (2020) and results from the non-linear behavior of the added-wave resistance. However, the heteroskedasticity is most pronounced for the MLP and decreases for the TEM. Moreover, we see the largest bias and variance for the MLP. The RF on the other hand, exhibits robust estimates with little noise; however, it is conspicuous that the test predictions are more biased towards underpredicting the target than the training predictions. The discrepancies in the residual distributions are not as pronounced for the XGBR and the MLP. Whether this is an indication of overfitting in case of the random forest will be assessed in the following analysis of the metrics of the test predictions. It is noted that a minor discrepancy between training and test performance is expected as it is unseen data.

The models are assessed based the metrics $RMSE$ and MAE using the out-of-sample (test) dataset in Table 4. It is stressed that (a) corresponds to the application of the Yeo–Johnson-Transformation and the described data preprocessing method. In addition, the performance on a dataset (b) without the described preprocessing methodology, i.e. only normalizing the target according to the extreme values and the use of the initial set of features is presented. Hence, the effect of the preprocessing methodology is examined. In addition, the performance of the chosen baseline models is given in order to underline the increase in accuracy resulting from the conducted parameter study. It is noted that Table 4 presents the model performance for both datasets and that the metrics for the in-sample set for both datasets are located in the appendix of this contribution.

The observations from Table 4 are two fold: (1) The performance increase resulting from the hyperparameter optimization and (2) the effect of the data preprocessing described in the previous section. Regarding the first point, the MLP model experiences the largest performance gain of 72.7%, whereas the RF and XGBR achieve an increase of 40.05% and 11.25%, respectively. It is stressed that the presented percentage increase is an average for both metrics and databases. Regarding the second point, it shows that the described data preprocessing method leads to superior model accuracy in comparison to the dataset without the applied methodology (b). The performance of the XGBR is affected the most with 170.5% higher accuracy relating (a) to (b). The MLP as well as the RF yield a performance increase of 35.51% and 36.87%, respectively. For the sake of clarity the results of Table 4 are presented in Fig. 8 in histograms.

Ultimately, the RF achieves the highest accuracy according to both metrics (bold in Table 4). Whether this is an indication of the superior generalization capability will be assessed in the model validation
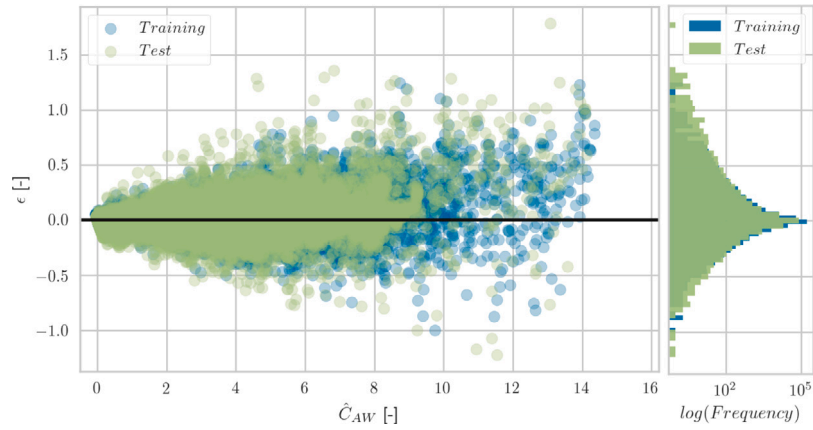
Fig. 6. Residuals plot for the XGBR on the $C_B = 0.84$ database. Note that the histogram (right-side plot) of the residuals presents the occurrence frequency in logarithmic scale.
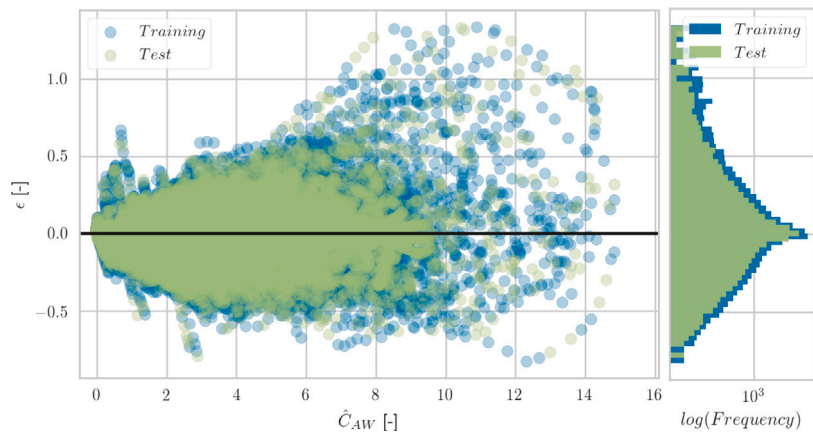

Fig. 7. Residuals plot for the MLP on the $C_B = 0.84$ database. Note that the histogram (right-side plot) of the residuals presents the occurrence frequency in logarithmic scale.

Table 4
Metrics of predictions using the test dataset for both datasets.

		$C_B = 0.84$			$C_B = 0.58$		
		RF	XGBR	MLP	RF	XGBR	MLP
Baseline	$RMSE$	0.0502	0.0789	0.2680	0.0497	0.0727	0.2808
	MAE	0.0130	0.0343	0.1304	0.0130	0.0328	0.1392
(a)	$RMSE$	**0.0306**	0.0645	0.0730	0.0310	0.0647	0.0611
		−38.98%	−18.27%	−72.74%	−37.61%	−11.04%	−78.25%
	MAE	**0.0075**	0.0310	0.0393	0.0076	0.0308	0.0416
		−41.83%	−9.65%	−69.84%	−41.78%	−6.07%	−70.11%
(b)	$RMSE$	0.0407	0.1597	0.0878	0.0369	0.1573	0.0848
		−18.85%	102.57%	−67.23%	−25.70%	116.27%	−69.80%
	MAE	0.0101	0.0921	0.0578	0.0122	0.0906	0.0566
		−22.39%	168.60%	−55.65%	−5.69%	175.96%	−59.38%

section. However, it proves that the RF did not overfit on the in-sample data despite the discrepancies of the residual histograms found in Fig. 5. In addition, it can be observed that the RF shows the smallest variance in model performance, cf. Fig. 8. Hence, it is concluded that the RF is in this case not only the most accurate regressor, but also the most robust one. The MLP on the other hand shows a considerable increase in performance after the parameter study and also the data preprocessing method affects the results noticeably. For the XGBR, the parameter study has only a small influence on performance and the poor performance on the (b) dataset is conspicuous. It is assumed that the dispersed target distribution (cf. upper plot in Fig. 2) is the reason for this performance drop. Outliers have a negative impact on boosting as each new tree builds on previous tree residuals. Thus boosting ensemble methods have a disproportionate bias towards outliers as they have much larger residuals than other samples. Still, the insufficient

performance on (b) is higher than expected. Lastly, one can conclude that it is justifiable to only optimize the models using one dataset ($C_B = 0.84$) as both the performance increase due to the hyperparameter optimization and the data preprocessing are similar.

4.3. Model validation

In the following, the models are validated on three case studies against publicly available model test and computational fluid dynamics results — RANS in particular. Additionally, the estimates of the DTU Design Tool and a semi-empirical formula proposed by Liu and Papanikolaou (2020) are considered. The following hull forms were selected as benchmarks from the literature: The full-type KVLCC2 (KRISO Very Large Crude oil Carrier) is investigated in head waves.
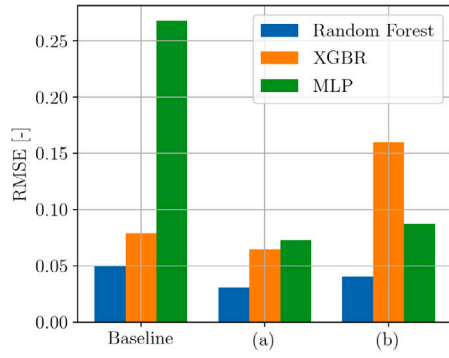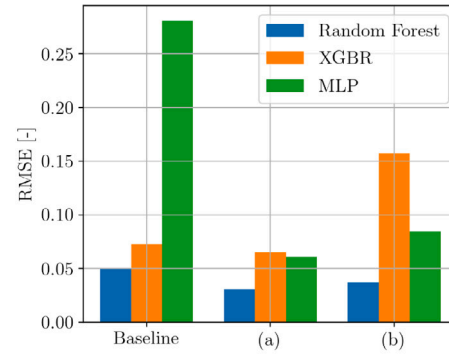
(a) Results for the $C_B = 0.84$ database.

(b) Results for the $C_B = 0.58$ database.

Fig. 8. Histogram of the $RMSE$ of the model predictions applied on the test set.

Table 5
Main particulars and operational conditions of the three case studies adopted from Sadat-Hosseini et al. (2013), Fujii and Takahashi (1975) and Kim et al. (2021).

	Unit	KVLCC2	S175	LNG-C
L_{pp}	[m]	320.0	175.0	290.0
B_m	[m]	58.0	25.4	45.0
T_D	[m]	20.8	9.5	11.5
C_B	[–]	0.808	0.561	0.79
k_{yy}/L_{pp}	[–]	0.25	0.24	0.25
Fn	[–]	0.142	0.15	0.19
β	[deg.]	180	150	120

On the other hand, the S175 container ship Fujii and Takahashi (1975) and the LNG carrier of Kim et al. (2021) are considered in bow oblique waves. The main particulars of the ships as well as their forward speed and relative wave direction are provided in Table 5.

As can be seen in Table 5, one slender and two full-type vessels were selected for model validation. All geometries remain within the definition range of the database (cf. Table 1), but do not exactly match the scaling ratios of the initial hull forms; thus, the generalization capability of the data-driven models can be assessed in a proper fashion. Besides the agreement in the magnitude of C_{AW}, the adequate representation of the resonant conditions is of high interest as the mean added resistance $\bar{R}_{AW}$ in irregular waves is a function of the overlap between the individual peaks of the prevalent wave energy density spectrum and the quadratic transfer function.

Initially, the blunt KVLCC2 tanker is considered in head waves $\beta = 180°$ at $Fn = 0.142$. In Fig. 9, the estimated quadratic transfer functions of C_{AW} are presented vs. the relative wavelength λ/L_{pp}. The experimental fluid dynamics (EFD) samples of Lee et al. (2013) and Sadat-Hosseini et al. (2013) as well as the RANS results of the latter publication are considered for validation.

In view of Fig. 9, it is stated that all estimates of the ML models are in accordance with the DTU Design Tool which is expected as the underlying database is the same. Moreover, the models show quantitative agreement with the CFD as well as EFD samples in long waves and qualitative agreement with their trend in the short wave regime. However, both the resonance frequency and the corresponding magnitude are not very well captured. Lastly, the semi-empirical formula provides the most appropriate estimate in comparison to the model test and CFD results. On a sidenote, the two separate experimental datasets show minor variability or rather uncertainty - especially near resonance. In addition, the RANS results differ notably from the experimental data in long and short waves.

The slender S175 container ship is examined in bow oblique waves ($\beta = 150°$) at $Fn = 0.15$. In Fig. 10, the data-driven models are validated against the RANS results of Uharek (2019) and the experimental data of Fujii and Takahashi (1975).

As can be inferred from Fig. 10, none of the ML methods captures the resonance region correctly as the individual peaks deviate in both wavelength and in magnitude. Furthermore, the ML models show a large variation in their accuracy. The MLP exhibits good agreement in long and short waves and it is evident that the tree-based ensemble methods (TEM) provide lower C_{AW} estimates in the proximity of resonance compared to the corresponding MLP values. The DTU Design Tool and the semi-empirical formula by Liu and Papanikolaou (2020) are in general in agreement with the numerical and experimental results. Once again, the discrepancy between RANS and EFD data is noted and the TEM are not able to reflect the resonant conditions correctly.

The individual results of the twinscrew LNG carrier, which is examined in oblique waves $\beta = 150°$ and moderate forward speed conditions ($Fn = 0.19$), are provided in Fig. 11. The benchmark data is established by numerical and experimental data adopted from Kim et al. (2021).

The immediate observation from Fig. 11 is the discrepancy between RANS and EFD data. In addition, the MLP, the DTU Design Tool as well as the semi-empirical formula by Liu and Papanikolaou (2020) show matching estimates — in resonant conditions in particular. However, the results of the TEM deviate both quantitatively and qualitatively in this regime. The data of the tree-based ensemble methods shows minor irregularities due to the fact that TEM only provide piecewise approximations. The satisfactory agreement of the ML models and DTU Design Tool is evident in short waves, where the semi-empirical formula of Liu and Papanikolaou (2020) yields higher values.

4.4. General discussion

The herein examined data-driven models were able to represent the frequency coherent character of the added-wave resistance. The validation shows sufficient performance for the MLP and highlights its applicability as a design aid and surrogate model. The TEM on the other hand show deviant results and high bias in resonant conditions when applied on unseen data. In long and even short waves, the agreement is considered as reasonable in most cases. The MLP shows overall sufficient resemblance to the two established prediction methods, i.e. the DTU Design Tool (Nielsen, 2015; Martinsen, 2016) and the semi-empirical formula by Liu and Papanikolaou (2020). As the added-wave resistance is generally considered as a non-viscous phenomenon, the application of potential theory methods is appropriate. However, considering further perspectives and the application to real data, all of the underlying methods are subject to the limitations of potential theory, such as the assumption of linearity including only small perturbations, small amplitudes, moderate wave steepness and the absence of viscous effects. Salvesen's method disregards the hull geometry above the waterline and includes only the Neumann–Kelvin linearized interaction between the steady and the unsteady flows. In
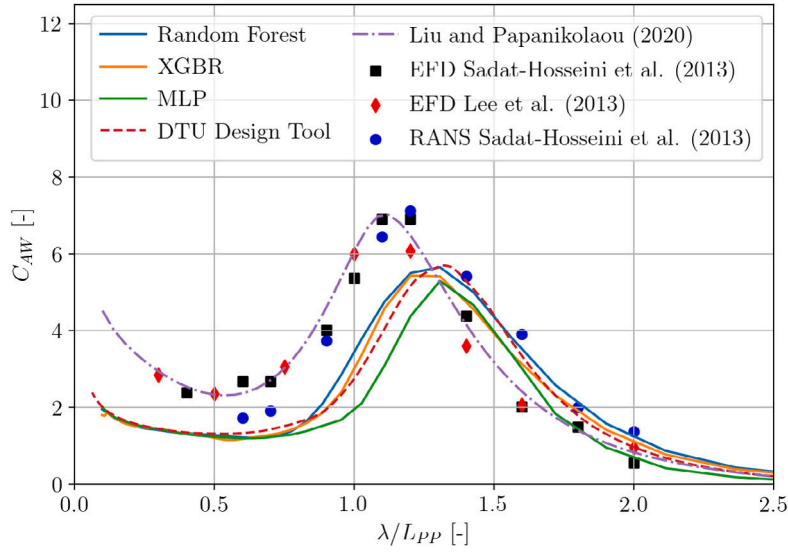
Fig. 9. Non-dimensional added-resistance C_{AW} of the KVLCC2 tanker for $\beta = 180°$ and $Fn = 0.142$.
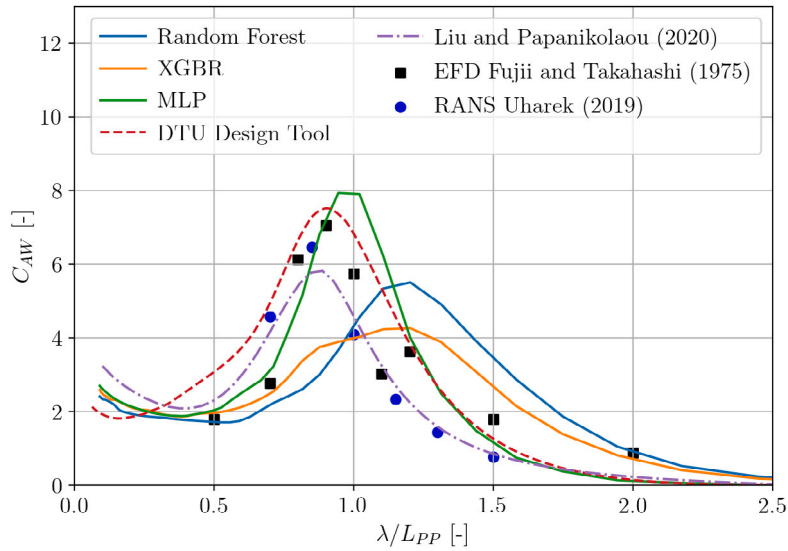

Fig. 10. Non-dimensional added-resistance C_{AW} of the S175 containership for $\beta = 150°$ and $Fn = 0.15$.

addition it is based on strip theory and inherits the corresponding assumptions and approximations. Similarly, Faltinsen's formula relies on the assumption that the ship has vertical walls, i.e. no flare is considered. Ultimately, the TEM match the abilities of the MLP in numerous characteristics, for instance capturing non-linear interactions between the predictors or handling the transition from Salvesen's to Faltinsen's estimate in the underlying database, but essentially the fundamental tradeoff in machine learning between transparency and accuracy, or rather in generalization capability, manifests itself in the present study through the transparent TEM showing inferior results compared to the opaque MLP. The MLP shows higher model generalization due to its multilayered architecture, but as the number of hidden layers increases the estimates and the mathematics behind them get more complex and opaque. The TEM on the other hand, have a clear and transparent structure and provide not only a decision path for every prediction, but also a direct estimate of each feature's predictive power. Still, the ensemble methods are affected by the fundamental limitations of decision trees and the CART implementation. On one hand, the CART algorithm is known for its transparency and simplicity as it grows binary tree graphs. In turn, this comes at the cost of instabilities and

inaccuracies as soon as the estimator is applied on unseen data. This instability is due to the hierarchical structure because the effect of an imperfection in the top split is propagated down to all child nodes. The MARS algorithm (cf. (Mittendorf and Papanikolaou, 2021)) has a mathematical resemblance to the CART algorithm and alleviates its lack of smoothness, but is not as scalable and hence not considered in the present work. Lastly, considerable variance in the estimates of the added-wave resistance was noticeable, not only for data-driven models, but also in the numerical results and even in the experimental data uncertainty was present, cf. Fig. 9. This reminds us that both the experimental and numerical determination of added-wave resistance, in particular near resonance and in oblique waves, is challenging. It could be even argued that a deterministic prediction is not *yet* suited for practical applications. Park et al. (2015) provide several sources of uncertainty regarding experiments pertaining to the added-wave resistance. This motivates dedicated studies for the uncertainty quantification in the direction of predicting the added-wave resistance using both numerical and empirical methods. From the literature it is evident that the uncertainty of added resistance is, indeed, well-known, but *not* well characterized.
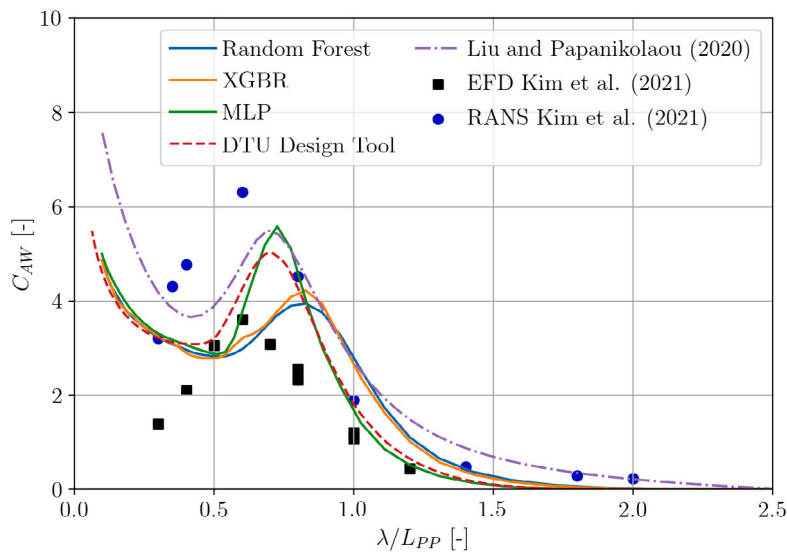
Fig. 11. Non-dimensional added-resistance C_{AW} of the LNG carrier for $\beta = 120°$ and $Fn = 0.19$.

5. Conclusions

The prediction of the second order steady wave forces on ships is a classical hydrodynamics problem and has been addressed in this contribution using different non-linear machine learning techniques. This study proposed a data-driven methodology for the estimation of the non-dimensional added-wave resistance C_{AW} based on results obtained from three different numerical calculation methods based on potential flow theory. The multivariate regression approach was focused on the comparison of tree-based ensemble methods and artificial neural networks in the context of prediction accuracy, robustness and transparency. Moreover, the data preprocessing and hyperparameter optimization were key features of the suggested methodology. As a result, all applied methods were able to produce a non-linear and consistent mapping from abstract features to the target C_{AW} with sufficient accuracy on in- and out-of-sample data. The application to real world case studies revealed insufficient results of the tree-based ensemble methods compared to the neural network. In comparison to well established design aids, the study proves that the multilayer perceptron is able to generate C_{AW} estimates of similar quality. Lastly, the present work highlights that the reliable and accurate prediction of the added-wave resistance is still a challenging task for all examined methods and will remain a competitive and demanding research topic in the coming future.

Addressing the research question from the paper's introduction, it appears that tree-based ensemble methods show comparable or even higher precision than the multilayer perceptron on in-sample data; however, the final validation study revealed that tree-based ensemble methods are rigid in their application and become unstable applied on new data differing from the initial discretization or distribution. This is a result of the inherent drawbacks of the Classification and Regression Trees (CART) algorithm as its simplicity comes at the cost of instability and inaccuracies as the weak learners tend to become overly complex. In fact, this leads to decreased generalization capability and overfitting. Moreover, tree-based ensemble methods show high bias in the sense that they underestimate the magnitude of the added-wave resistance under resonant conditions for most cases. As already noted, tree-based ensemble methods do not perform sufficiently well on sparse data, therefore the minimization of epistemic uncertainty is pivotal for the improvement of the model performance. Thus, the generation of an updated database with additional parameter combinations and parameter variation is an important aspect for future work. In addition,

the consideration of different longitudinal radii of gyration k_{yy} and block coefficients C_B values is required for improving the prediction of the added-wave resistance and the corresponding resonance frequency. Lastly, the calibration on experimental data of the models using transfer learning is a promising aspect of future work.

The implicit feature processing capabilities of deep neural networks make their application appealing as the herein labor intensive preprocessing is not required using these advanced architectures. Considering deep learning methods and their ability handling high-dimensional tensors, it is possible to utilize the hull shape itself as input as has been proposed in the study by Taniguchi and Ichinose (2020). They employ a Generative Adversarial Network for the generation of novel hull shapes and a convolutional neural network as a discriminator trained on CFD results of the wake profile. Such sophisticated layer types and architectures have immense potential for future work regarding the development of initial design aids and also hull shape optimization approaches.

CRediT authorship contribution statement

Malte Mittendorf: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualization, Writing – original draft & editing. **Ulrik D. Nielsen:** Writing – review & editing, Supervision, Project administration, Resources, Funding acquisition. **Harry B. Bingham:** Writing – review & editing, Supervision, Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
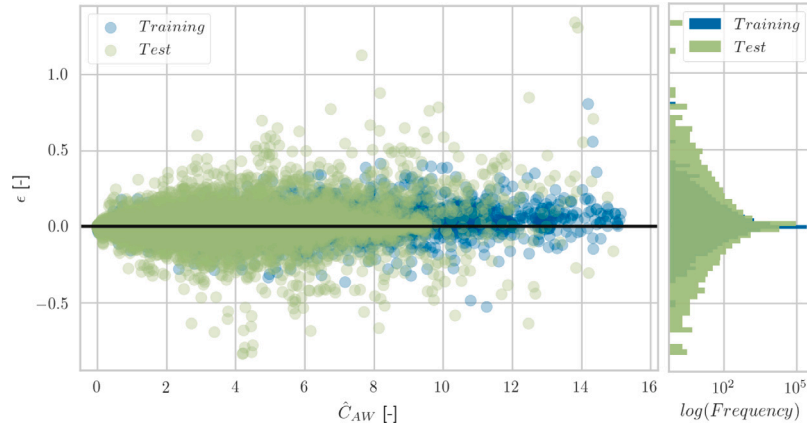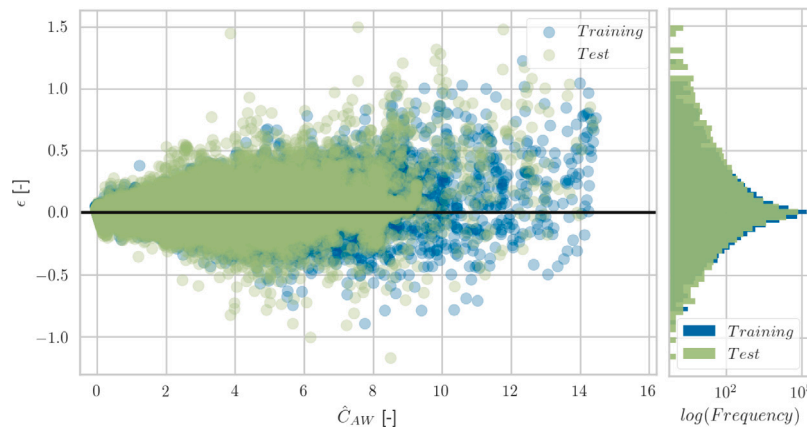
Acknowledgments

Fig. 12. Residuals plot for the RF on the $C_B = 0.58$ database.

Fig. 13. Residuals plot for the XGBR on the $C_B = 0.58$ database.

Table 6
Metrics for the prediction of the train dataset for both databases.

		$C_B = 0.84$			$C_B = 0.58$		
		RF	XGBR	MLP	RF	XGBR	MLP
Baseline	$RMSE$	0.0322	0.0634	0.2674	0.0318	0.0596	0.2844
	MAE	0.0083	0.0295	0.1299	0.0082	0.0285	0.1393
(a)	$RMSE$	0.0121	0.0492	0.0726	0.0121	0.0484	0.0605
	MAE	0.0029	0.0258	0.0392	0.0029	0.0256	0.0414
(b)	$RMSE$	0.0154	0.1439	0.0877	0.0150	0.1423	0.0840
	MAE	0.0043	0.0870	0.0577	0.0054	0.0859	0.0563

Table 7
Main particulars of the additional four case studies taken from Valanto and Hong (2015), Shigunov et al. (2018), Yasukawa et al. (2019) and Park et al. (2019).

	Unit	HSVA	DTC	S-Cb84	SNU
L_{pp}	[m]	220.32	355.0	178.0	323.0
B_m	[m]	32.04	51.0	32.26	60.0
T_m	[m]	7.2	14.5	11.57	21.0
C_B	[-]	0.654	0.661	0.84	0.83
k_{yy}/L_{pp}	[-]	0.263	0.27	0.25	0.258
Fn	[-]	0.233	0.052	0.0	0.137
β	[deg.]	120	180	150	150

Appendix

A.1. Performance on training dataset

See Table 6.

A.2. Residuals plots for the $C_B = 0.58$ database

See Figs. 12–14.

A.3. Additional validation cases

In addition to the three validation cases in Section 4, the models are validated for four other case studies. The ships are selected such that they cover different operational conditions and hull shapes. The main particulars of the hulls are presented in Table 7 and it is evident that two slender and two full-type ships are chosen.

Initially, the slender HSVA (Hamburg Ship Model Basin) cruise ship is considered in oblique waves $\beta = 120°$ and relatively high forward speed $Fn = 0.233$. In Fig. 15, the transfer function of C_{AW} is presented for the normalized wavelength λ/L_{pp} alongside the experimental samples of Valanto and Hong (Valanto and Hong, 2015).

As can be inferred from Fig. 15, all three ML models are consistent with both model test results and the semi-empirical formula. However, it is conspicuous that none of the methods capture the resonance region correctly as the individual peaks deviate in both wave length and in magnitude. It is evident that the TEM provide lower C_{AW} estimates in the proximity of resonance compared to the MLP. Lastly, the MLP provides the most accurate ML prediction considering both model test data and the formula of Liu and Papanikolaou (2020). One possible explanation for the deviations in case of the HSVA cruise ship could be that the k_{yy}/L_{pp} is kept constant in the database at 0.25 and the
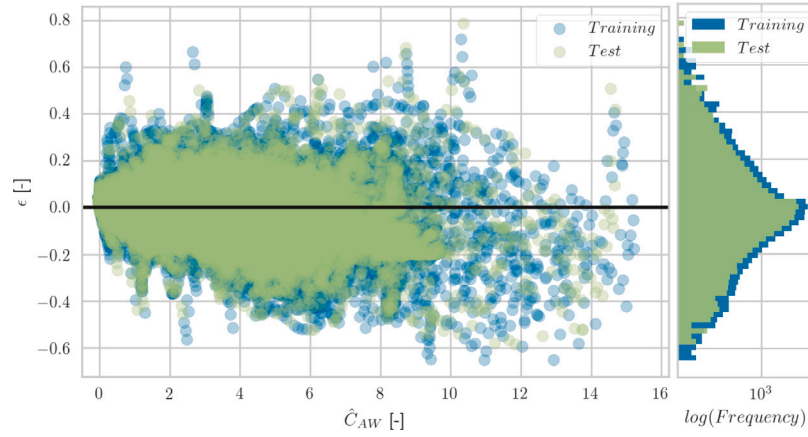
Fig. 14. Residuals plot for the MLP on the $C_B = 0.58$ database.
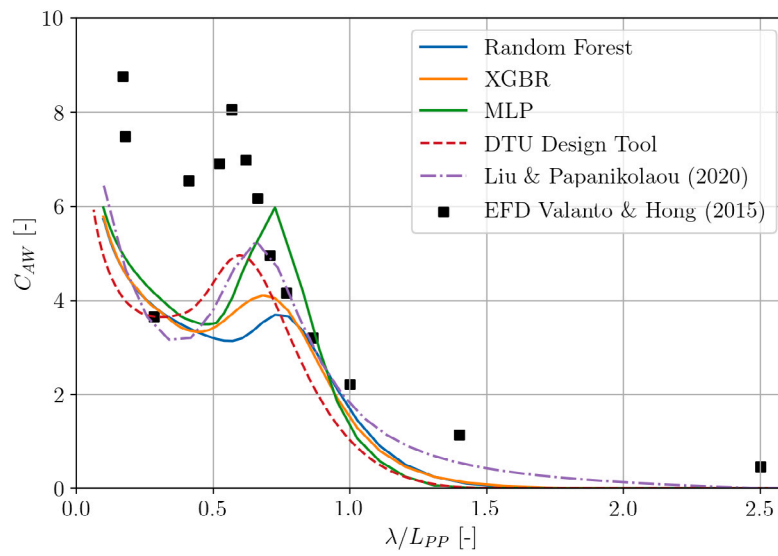
Fig. 15. C_{AW} for the HSVA cruise ship for $\beta = 120°$ and $Fn = 0.233$.
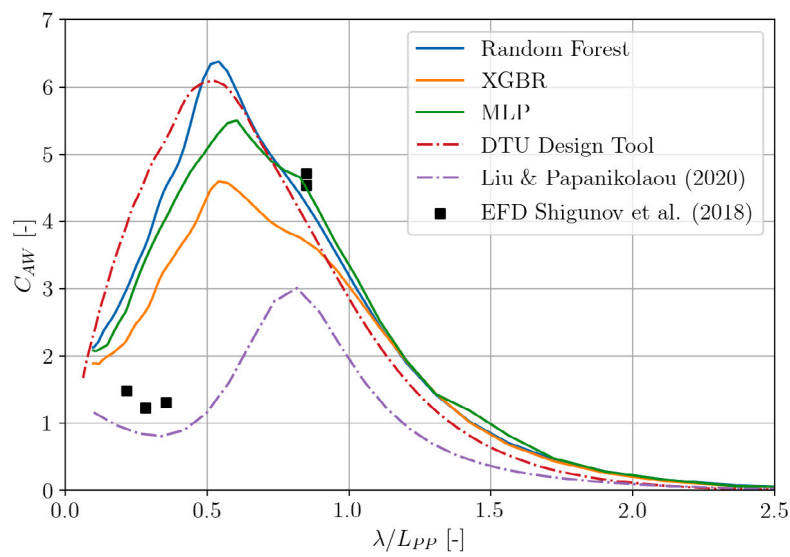
Fig. 16. C_{AW} for the DTC container ship for $\beta = 180°$ and $Fn = 0.052$.

cruise ship has a $k_{yy}/_{Lpp} = 0.263$, (Valanto and Hong, 2015). In addition, the loading condition (i.e. GM) is neither variable in the semi-empirical formula of Liu and Papanikolaou (2020) nor in the used dataset. The weight distribution in general is a considerable source of
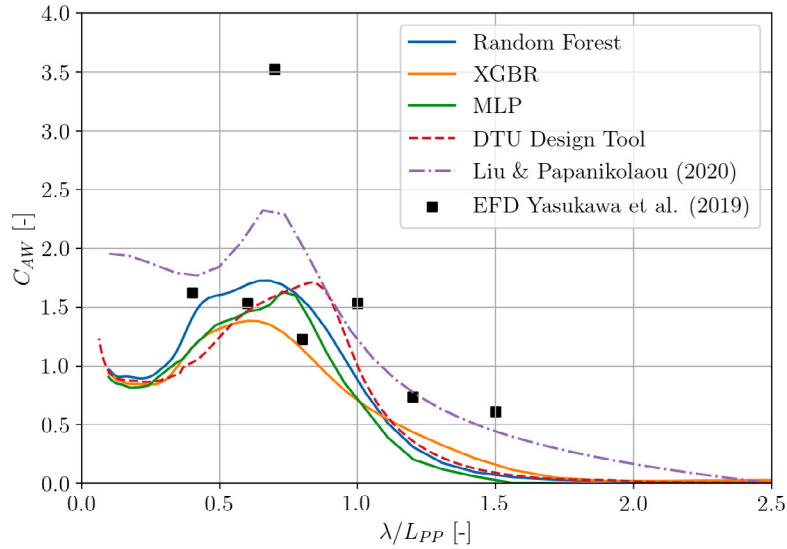
Fig. 17. C_{AW} for the S-Cb84 bulk carrier for $\beta = 150°$ and $Fn = 0.0$.

Fig. 18. C_{AW} for the SNU tanker for $\beta = 150°$ and $Fn = 0.137$.

uncertainty during model tests as Park et al. (2015) mention. Moreover, the occurrence of breaking waves as reported by Valanto and Hong (2015), the relatively high forward speed and potential wall reflections for oblique waves could be additional driving factors of experimental uncertainty.

The slim-shaped Duisburg Test Case (DTC) was studied extensively for its manoeuvering characteristics in waves during the SHOPERA EU project. The case with head waves ($\beta = 180°$) and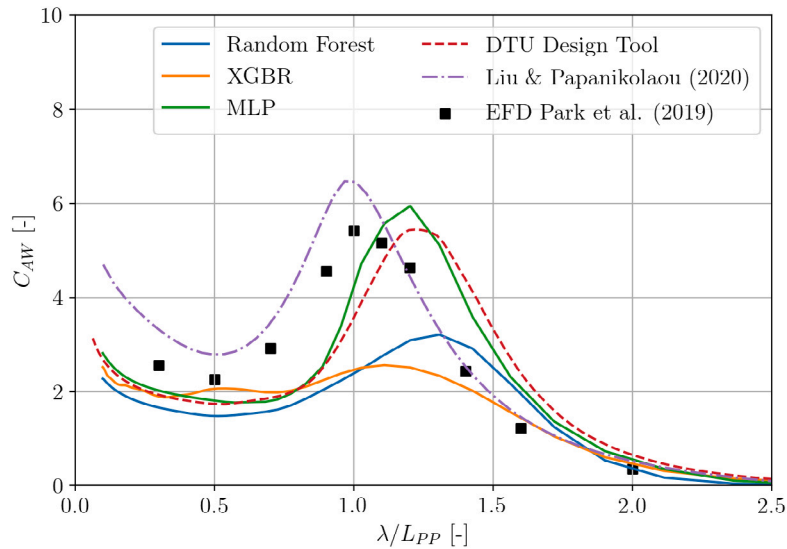 relatively low forward speed with $Fn = 0.052$ was examined. The results for this case are presented in Fig. 16 with the samples from Shigunov et al. (2018) as benchmark.

In view of Fig. 16, it is assumed that the ML models provide reasonable estimates in the regime of long waves due to the asymptotic behavior, but in short wavelengths drastic deviations are noticeable in comparison to the benchmark results. Conversely, the semi-empirical formula captures the added-wave resistance in short waves sufficiently, but underestimates the magnitude for the resonant frequencies. Also, the estimates of the TEM – the XGBR in particular – show considerable variance due the fact that TEM only yield piecewise approximations.

In the following, a case study without forward speed ($Fn = 0.0$) is investigated with the full-type S-Cb84 bulk carrier in oblique waves

$\beta = 150°$ and the results are presented in Fig. 17 together with the model test data of Yasukawa et al. (2019).

As can be inferred in Fig. 17, the estimates of the ML models are qualitatively in accordance with the benchmark data and the Liu Papanikolaou formula. Moreover, the semi-empirical method shows superior accuracy in long waves, whereas the ML methods exhibit good agreement in shorter waves.

The last test case is the SNU (Seoul National University) tanker, which is characterized by high fullness, in oblique waves $\beta = 150°$ and medium forward speed conditions ($Fn = 0.137$). The experimental data of Park et al. (2019) is considered for validation.

The ML models show a large variance in their accuracy, viz. Fig. 18. In particular, the results of the TEM deviate dramatically both quantitatively and qualitatively. The predictions of the MLP match the magnitude of the experimental data, but the peak frequencies differ slightly. Once again, the data of the TEM shows irregularities in short waves.

References

Bertram, V., Mesbahi, E., 2000. Applications of neural nets in ship design. In: Jahrbuch der Schiffbautechnischen Gesellschaft. Springer.

Bishop, C., 2006. Pattern Recognition and Machine Learning. Springer Verlag, New York, http://dx.doi.org/10.1117/1.2819119.

Boese, P., 1970. Eine Einfache Methode Zur Berechnung Der Widerstandserhoehung eines Schiffes im Seegang, Vol. 258. Schriftenreihe Schiffbau, pp. 1–9. http://dx.doi.org/10.15480/882.648.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. http://dx.doi.org/10.1023/A:1010933404324.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and regression trees. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, ISBN: 9780412048418.

Cepowski, T., 2020. The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network. Ocean Eng. 195, http://dx.doi.org/10.1016/j.oceaneng.2019.106657.

Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, http://dx.doi.org/10.1145/2939672.2939785.

Faltinsen, O.M., 2005. Hydrodynamics of High-Speed Marine Vehicles. Cambridge University Press, http://dx.doi.org/10.1017/CBO9780511546068.

Faltinsen, O.M., Minsaas, K.J., Liapis, N., Skjordal, S.O., 1980. Prediction of resistance and propulsion of a ship in a seaway. In: Inui, T. (Ed.), Proceeding of 13th Symposium on Naval Hydrodynamics. The Shipbuilding Research Association of Japan, Tokyo (Japan), pp. 505–529.

Fujii, H., Takahashi, T., 1975. Experimental study on the resistance increase of a large full ship in regular oblique waves. J. Soc. Naval Archit. Japan 1975 (137), 132–137.

Gerritsma, J., Beukelman, W., 1971. Analysis of the Resistance Increase in Waves of a Fast Cargo Ship.. Technical Report 334, Shipbuilding laboratory, Technological University Delft, Delft, The Netherlands.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. The MIT Press, ISBN: 0262035618.

Herradon de Grado, E., Bertram, V., 2016. Predicting added resistance in wind and waves employing artificial neural nets. In: In Proc. of the 1st Hull Performance & Insight Conference, HullPIC'16. Pavone.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Mach. Learn. 46 (1), 389–422. http://dx.doi.org/10.1023/A:1012487302797.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer Science+Business Media, New York, http://dx.doi.org/10.1007/978-0-387-84858-7.

International Maritime Organization, 2011. Amendments To the Annex of the Protocol of 1997 to Amend the International Convention for the Prevention of Pollution from Ships, 1973, As Modified By the Protocol of 1978 Relating Thereto, Vol. 203 no. 62. Resolution MEPC.

Kim, T., Yoo, S., Kim, H.J., 2021. Estimation of added resistance of an LNG carrier in oblique waves. Ocean Eng. 231, 109068. http://dx.doi.org/10.1016/j.oceaneng.2021.109068.

Kingma, D.P., Ba, J., 2015. Adam: a method for stochastic optimization. In: Proc. of 3rd International Conference on Learning Representations, ICLR '15, San Diego.

Kristensen, H.O.H., 2010. Model for environmental assessment of container ship transport. Soc. Nav. Archit. Mar. Eng. Trans. 118, 122–139. http://dx.doi.org/10.5957/SMC-2010-T41.

Lee, C.H., Newman, J.N., 2013. Wamit user manual. http://www.wamit.com/manual.html version 7.0.

Lee, J.H., Seo, M.G., Park, D.M., Yang, K.K., Kim, K.H., Kim, Y., 2013. Study on the effects of hull form on added resistance. In: Proc. of the 12th International Symposium on Practical Design of Ships and Other Floating Structures, Changwon, Korea, pp. 329–337.

Lin, D., Foster, D.P., Ungar, L.H., 2011. VIF regression: A fast regression algorithm for large data. J. Amer. Statist. Assoc. 106 (493), 232–247. http://dx.doi.org/10.1198/jasa.2011.tm10113.

Liu, S., Papanikolaou, A., 2017. On the prediction of the added resistance of large ships in representative seaways. Ships Offshore Struct. 12 (5), 690–696. http://dx.doi.org/10.1080/17445302.2016.1200452.

Liu, S., Papanikolaou, A., 2020. Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. Ocean Eng. 206, http://dx.doi.org/10.1016/j.oceaneng.2020.107357.

Martinsen, M.A., 2016. A design tool for estimating wave added resistance of container ships (Master's Thesis). Technical University of Denmark.

Maruo, H., 1957. The excess resistance of a ship in rough seas. Int. Shipbuild. Prog. 4 (35), 337–345. http://dx.doi.org/10.3233/ISP-1957-43501.

Mitchell, T.M., 1997. Machine Learning. McGraw-Hill, New York.

Mittendorf, M., Papanikolaou, A.D., 2021. Hydrodynamic hull form optimization of fast catamarans using surrogate models. Ship Tech. Res. 68 (1), 14–26. http://dx.doi.org/10.1080/09377255.2020.1802165.

Mockus, J., 1989. Bayesian Approach to Global Optimization. Kluwer Academic Publishers, Dordrecht.

Nielsen, C.S., 2015. A ship design tool for estimating added resistance in waves (Master's Thesis). Technical University of Denmark.

Oberhagemann, J., Kaufmann, J., Ervik, A., Gramstad, O., Helmers, J., Sireta, F.X., 2019. Wave Load and Response Predictions Combining HOSM, CFD and Machine Learning. In: Proc. of the 38th Int. Conf. on Ocean, Offshore and Arctic Engineering OMAE 2019, Glasgow, Scotland, http://dx.doi.org/10.1115/OMAE2019-95352.

Park, D.-M., Lee, J.-H., Jung, Y.-W., Lee, J., Kim, Y., Gerhardt, F., 2019. Experimental and numerical studies on added resistance of ship in oblique sea conditions. Ocean Eng. 186, 106070. http://dx.doi.org/10.1016/j.oceaneng.2019.05.052.

Park, D.-M., Lee, J., Kim, Y., 2015. Uncertainty analysis for added resistance experiment of KVLCC2 ship. Ocean Eng. 95, 143–156. http://dx.doi.org/10.1016/j.oceaneng.2014.12.007.

Pedregosa, F., et al., 2011. Scikit-learn: Machine learning in python. Mach. Learn. Res. 12, 2825–2830.

Petersen, T., 2000. Wave Load Prediction-a Design Tool. Department of Naval Architecture and Ocean Engineering, Technical University of Denmark.

Sadat-Hosseini, H., Wu, P., Carrica, P., Kim, H., Toda, Y., Stern, F., 2013. CFD verification and validation of added resistance and motions of KVLCC2 with fixed and free surge in short and long head waves. Ocean Eng. 59, 240–273. http://dx.doi.org/10.1016/j.oceaneng.2012.12.016.

Salvesen, N., 1978. Added resistance of ships in waves. J. Hydronaut. 12 (1), 21–34. http://dx.doi.org/10.2514/3.63110.

Salvesen, N., Tuck, E., Faltinsen, O., 1970. Ship motions and sea loads. In: Annual Meeting of the Society of Naval Architects and Marine Engineers, SNAME, New York.

Sclavounos, P.D., Ma, Y., 2018. Artificial Intelligence Machine Learning in Marine Hydrodynamics. In: Proc. 37th Int. Conf. on Ocean, Offshore & Arctic Engineering OMAE 2018, Madrid, Spain.

Seabold, S., Perktold, J., 2010. statsmodels: Econometric and statistical modeling with python. In: Proc. of 9th Python in Science Conference, http://dx.doi.org/10.25080/Majora-92bf1922-011.

Shigunov, V., el Moctar, O., Papanikolaou, A., Potthoff, R., Liu, S., 2018. International benchmark study on numerical simulation methods for prediction of maneuverability of ships in waves. Ocean Eng. 165, 365–385. http://dx.doi.org/10.1016/j.oceaneng.2018.07.031.

Sigmund, S., el Moctar, O., 2018. Numerical and experimental investigation of added resistance of different ship types in short and long waves. Ocean Eng. 147, 51–67. http://dx.doi.org/10.1016/j.oceaneng.2017.10.010.

Strøm-Tejsen, J., Yeh, H.Y.H., Moran, D.D., 1973. Added resistance in waves. Soc. Nav. Archit. Mar. Eng. Trans. 81, 109–143.

Taniguchi, T., Ichinose, Y., 2020. Hull form design support tool based on machine learning. In: Proc. of the 19th Conference on Computer and IT Applications in the Maritime Industries, COMPIT'20, Pontignano.

The GPyOpt authors, 2016. GPyOpt: A Bayesian optimization framework in python. http://github.com/SheffieldML/GPyOpt.

Uharek, S.A., 2019. Numerical Prediction of Ship Manoeuvring Performance in Waves (Doctoral Thesis). Technical University of Berlin.

Valanto, P., Hong, Y., 2015. Experimental investigation on ship wave added resistance in regular head, oblique, beam, and following waves. In: Proc. of the 25th International Society of Offshore and Polar Engineers.

Yasukawa, H., Hirata, N., Matsumoto, A., Kuroiwa, R., Mizokami, S., 2019. Evaluations of wave-induced steady forces and turning motion of a full hull ship in waves. J. Mar. Sci. Technol 24 (1), 1–15. http://dx.doi.org/10.1007/s00773-018-0537-3.

Yeo, I.-K., Johnson, R.A., 2000. A new family of power transformations to improve normality or symmetry. Biometrika 87 (4), 954–959. http://dx.doi.org/10.1093/biomet/87.4.954.

[J_2] Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading

The paper entitled "Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading" is published in the Journal of Ocean Engineering as:

Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading

Malte Mittendorf [a,*], Ulrik D. Nielsen [a,c], Harry B. Bingham [a], Shukui Liu [b]

[a] *DTU Mechanical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*
[b] *School of Mechanical and Aerospace Engineering, Nanyang Technological University, SG-639798 Singapore, Singapore*
[c] *Centre for Autonomous Marine Operations and Systems, NTNU AMOS, NO-7052 Trondheim, Norway*

ARTICLE INFO

ABSTRACT

The present paper examines a semi-empirical framework for the estimation of added resistance in arbitrary wave heading under consideration of uncertainty quantification. In this respect, the calibration of the formula's parameter vector is conducted based on particle swarm optimization as well as a database of model test results comprising 25 different ships and around 1100 samples. In the first iteration, the minimization of reducible systematic uncertainty is of interest and the effect of four objective functions on prediction accuracy is evaluated. Moreover, two different parameter combinations were obtained for blunt ($C_B \geq 0.70$) and slender-type ships. Conversely, the irreducible statistical uncertainty, i.e. the inherent noise of the experimental data, is taken into account by a quantile regression procedure. Applying this approach, a 90% prediction interval for the formula's estimates is implemented using the skewed version of the superior loss function in the previous iteration. The practical relevance of an uncertainty estimate for the prediction of the added resistance is emphasized in the final part, in which the proposed approach is validated in regular waves against model test data and other well-established prediction methods. In general, the validation studies suggest satisfactory performance and reliability of the adapted semi-empirical formulation.

1. Introduction

The increasing pressure at societal and legislative levels for the reduction of carbon emissions in the maritime sector makes the accurate and *transparent* prediction of the actual required engine power in realistic seaways an imperative throughout the ship's life cycle. The added resistance is of high relevance in ship design and operation due to its impact on safety and energy efficiency. The decrease in service speed of merchant ships (slow steaming) leads to a larger contribution of the added resistance to the total resistance in relative terms. Thus, the added resistance is mainly a concern for energy efficiency considerations in moderate sea states. In severe sea states, however, the risk of losing maneuverability – for tankers and bulk carriers in particular – is pivotal due to the prevalent trend of reducing installed engine power, in order to comply with the enforced energy efficiency regulations such as the Energy Efficiency Design Index (EEDI), Papanikolaou (2018). This safety concern motivated the legislation of the Minimum Propulsion Power (MPP) regulations by the International Maritime Organization (IMO) for blunt-type ships, in which several semi-empirical formulas are recommended for the practical determination of the added resistance and the associated speed loss in a seaway, cf. IMO (2013, 2016) and IMO (2017).

The numerical calculation of added resistance has attracted much attention over last years using not only potential flow theory methods, such as strip theory (e.g. Amini-Afshar and Bingham (2021)) and panel codes (e.g. Söding et al. (2014)), but also using RANS (Reynolds-Averaged Navier Stokes) methods for calculating the added resistance (e.g. Hizir et al. (2019)), the related propulsive coefficients (e.g. Mikkelsen et al. (2022)) or the maneuvering coefficients (e.g. Uharek (2019)). However, the estimation of hydrodynamic characteristics in waves, such as the added resistance, is subject to large uncertainties both in terms of theoretical/numerical and experimental methods. The SHOPERA (Energy efficient safe SHip OPERAtion) benchmark study, which was part of the eponymous EU-project, revealed remarkable deviations in relatively short wave conditions between numerical results and model test data, cf. Shigunov et al. (2018). As a matter of fact, precisely these conditions (i.e. $\lambda/L_{pp} < 0.5$) are nowadays of high importance according to Kwon (1981) and Minsaas et al. (1983) as ships predominantly sail in these low to moderate sea states. This is a consequence of increasing ship size (economy of scale) and also due to the use of routing optimization/weather routing. Still, short waves impose tremendous challenges on contemporary numerical

---

* Corresponding author.
  E-mail address: mamit@mek.dtu.dk (M. Mittendorf).

methods and push their applicability towards and beyond the limits. On one hand, the importance of non-linear effects, such as wave breaking in short waves impedes the use of potential flow methods, such as strip theory, for the prediction of added resistance. Additionally, strip theory codes show a lack of accuracy in following and quartering waves, Park et al. (2019). On the other hand, short waves lead to a extremely fine spatial discretization of the free surface requiring immense amounts of computational effort for RANS methods. The computational cost is amplified by the required small timesteps for the temporal discretization, in order to prevent numerical diffusion of the wave elevation ($CFL = \frac{u\Delta t}{\Delta x} < 0.5$ in free surface region), Sigmund (2018).

Despite the increase in the availability of computational resources and the accuracy of high-fidelity Computational Fluid Dynamics (CFD) calculation tools, semi-empirical formulas are still widely applied for the transparent and efficient assessment of a ship's hydrodynamic characteristics in ship design and operation. The development of (semi-)empirical methods for the estimation of the added resistance in waves is a competitive research area and sparked several different studies over the past years. The application cases for such semi-empirical formulas are manifold, e.g. the wave correction during sea trials (Tsujimoto et al., 2021), overall hull and propeller performance monitoring (Taskar and Andersen, 2021) and for bow shape optimization under realistic conditions, Bolbot and Papanikolaou (2016). The bow shape and the stem profile in particular are highly influential on the magnitude of the added resistance. For instance, Yasukawa and Tsujimoto (2020) conducted systematic model tests and showed a significant decrease in added resistance throughout the total range of relative wavelengths for a vertical stem with no protruding bulb compared to a conventional stem geometry with a bulbous bow. In the field of ship operation, the reliable prediction of the added resistance has also considerable influence on routing optimization (weather routing) as well as monitoring the ship technical performance over time and scheduling hull/engine maintenance.

1.1. Literature review

Havelock (1940) pioneered the scientific work of theoretically approximating second order wave forces by deriving a formulation for the steady force acting on vertical circular cylinders. Fujii and Takahashi (1975) proposed their well-renowned formula for the prediction of added resistance caused by diffraction, i.e. wave reflection, by introducing two correction coefficients to Havelock's approach. These two coefficients take forward speed as well as "finite draft" into account and their derivation was based on theoretical results from Ursell and Dean (1947). Takahashi (1988) modified the original formula using model test data. Later, Tsujimoto et al. (2008) and Kuroda et al. (2008) fine-tuned the empirical corrections of the two additional parameters based on experimental data and proposed the so-called NMRI (National Maritime Research Institute Japan) formula. However, Kwon (1981) pointed out that the finite draft influences both the partial reflection effect and energy transmission below the hull. Liu (2020) as well as Mourkogiannis and Liu (2021) proposed a physically more solid semi-empirical formula taking the effects of forward speed, partial reflection and the finite water depth as well as the energy transmission below the hull into account.

In parallel, Jinkine and Ferdinande (1974) proposed an empirical method for the calculation of the added resistance of fast cargo ships due to radiation, i.e. in long waves. The extension of this method for short waves using experimental data resulted in the STAwaveI and STAwaveII formula for the estimation of the added resistance, where the latter was recommended for the use in sea trial corrections by the ITTC (2014). Lang and Mao (2021) proposed a combined approach in their CTH formula (Chalmers Tekniska Högskola) by taking the NMRI formula for short waves and the Jinkine and Ferdinande (1974) approach for long waves. Furthermore, Lang and Mao implemented

several correction coefficients in short waves based on experimental data. Faltinsen et al. (1980) developed an analytical asymptotic formula for the calculation of the added resistance in short waves. Kashiwagi et al. (2010) as well as Liu and Papanikolaou (2013) introduced practical corrections for advance speed and short waves, respectively. Just recently, Yang et al. (2018) also modified Faltinsen's asymptotic formula in order to reflect the three following aspects: finite draft, local steady velocity, and the hull shape above the waterline.

Liu et al. (2015), Liu and Papanikolaou (2016) and Liu et al. (2019) established an efficient semi-empirical formula by combining an enhanced version of the work of Jinkine and Ferdinande (1974) in long waves and an adapted version of Faltinsen's asymptotic formula (Faltinsen et al., 1980) in short waves simplified for practical use. Valanto and Hong (2015) show that the added resistance becomes most prominent in bow oblique waves and thus the consideration of heading angles $\beta \neq 180$ deg. is vital for practical application. Hence, the approach for head waves was lately extended to arbitrary heading angles and the proposed approach was validated against a comprehensive database of model test results, Liu and Papanikolaou (2020). Upon successful validation studies, it was recently adopted by ITTC recommended procedure for analyzing sea trial results, ITTC (2021) and Wang et al. (2021). Moreover, the MEPC (Marine Environment Protection Committee) 76 circular (IMO, 2016, 2017, 2021) also endorsed this method for the assessment of the minimum propulsion power of blunt-type ships. Furthermore, Nielsen (2015) and Martinsen (2016) developed the DTU design tool, which is a linear interpolation method aggregating results of three different potential flow theory methods for a wide range of wave headings and hull shapes. Machine Learning (ML) techniques are increasingly employed in the realm of ship hydrodynamics. Thus, Mittendorf et al. (2022) compare Artificial Neural Networks (ANN) and ensemble tree-based methods for the prediction of the added resistance in head to beam waves using the underlying database of the DTU design tool. In addition, Cepowski (2020) proposed another solely data-driven formula by training an ANN on model test data in head wave conditions and extracted the corresponding equation from the obtained model parameters and activations.

One can conclude that (semi-)empirical methods for the estimation of the added resistance in waves have evolved over time. The mathematical frameworks underwent multiple iterations and experienced an increase in accuracy as more model test data became available and as data analysis methods became more advanced. Holt and Nielsen (2021) assessed the recently passed MPP regulations using empirical methods for the estimation of added resistance. Their study sparked the idea of the present contribution in adapting the longitudinal radius of gyration k_{yy} for the increase in prediction accuracy of the STAwaveII formula. Just recently, Gatchell (2018) presented a parameter calibration of the Hollenbach (1998) calm water resistance estimation method for contemporary RoPax hull shapes using CFD results and Particle Swarm Optimization (PSO). Nikolopoulos and Boulougouris (2019) on the other hand employed the multi-objective genetic algorithm NSGA-II (Non-dominated Sorting Genetic Algorithm) to adapt the Holtrop and Mennen (1982) framework statistically for blunt ships using model test data. Moreover, they expressed the uncertainty for each prediction *quantitatively* by an analytical equation. Lastly, the rigorous parameter calibration as well as uncertainty quantification are still unexplored in the field of added resistance, and hence represent the novelties of the following contribution.

1.2. Motivation and objective

The majority of estimation methods for the added resistance in waves only consider head waves and provide deterministic predictions showing severe practical limitations for these established formulas. In this work, however, we make use of the formulation of Liu and Papanikolaou (2020) for arbitrary wave heading and attempt to minimize the reducible uncertainty pertaining the estimation of the added

resistance and express the irreducible component *quantitatively*. Quoting the famous aphorism of George Box,[1] it is the overall aim of this paper to enhance the usefulness of the semi-empirical formula of Liu and Papanikolaou (2020) and accompany the estimates with uncertainty bounds. In essence, the presented approach combines the transparency of a simple semi-empirical formula with the accuracy and predictive power of a supervised machine learning method in a multivariate regression context. This will be addressed in two stages: (1) In the *holistic* parameter calibration or rather training procedure a metaheuristic optimization algorithm and a dataset of benchmark model test results are employed. The parameter vector will be adjusted minimizing, and comparing several different loss functions. (2) In the second iteration, new parameters are derived using a quantile loss function for lower and upper bounds establishing a 90% prediction interval. The resulting framework is comparable to the method of Hollenbach (1998) which provides lower and upper bounds, along with the actual prediction, for the estimation of the calm water resistance. Furthermore, it is envisioned to provide a pseudo-stochastic prediction of the added resistance for arbitrary heading instead of a deterministic one as a trustworthy representation of uncertainty is an imperative in safety-critical domains such as wave loads on ships.

1.3. Composition

In the following Section 2, the analysis of the used database and the residual analysis of the underlying (Liu and Papanikolaou, 2020) formula will be presented. Section 3 is devoted to elucidate the applied calibration methodology and presents the adapted and parameterized semi-empirical formulation as well as the particle swarm optimizer. Section 4 provides the obtained results and discussion, which includes the validation in regular waves investigating a number of case studies. In the final Section 5, conclusions are drawn and suggestions for extending the work are presented.

2. Data analysis

Experiments on added resistance started in the 1960s and were exploited to gain understanding of the behavior and performance of ships in a realistic seaway as well as for the validation of numerical and empirical tools. Conducting this kind of experiments requires highly accurate measurement equipment, as the added resistance is a higher order force and in most cases only of small magnitude compared to the calm water resistance. In general, the added resistance is seen as a *non-viscous* phenomenon and can be obtained from the square of ship-generated waves, consisting of a linear superposition of diffraction and radiation waves. Kashiwagi (2013) shows that the diffraction component is not frequency dependent and nearly constant throughout the entire regime of wavelengths. Initially, this may seem contradictory, however, Kashiwagi et al. (2010) conclude from experimental data in head wave conditions that the radiation component takes negative values in longer waves due to the interaction terms. Thus, the sum of both contributions yield small, but positive values approaching zero with increasing relative wavelength. The radiation component is, in fact, frequency dependent and dominant in long to medium wavelengths. On the other hand, the diffraction part is relatively more pronounced in shorter waves. Blok (1993) splits the contributions to the added resistance into (1) operational conditions, i.e. wavelength λ, amplitude ζ_a and relative heading β as well as ship speed Fn and (2) driving factors related to the ship, i.e. main particulars and hull shape.

Table 1 presents an array of experimental data for head, oblique and following waves and represents the employed database of this work. The main particulars of the examined ships are indicated and also the references are given in Table 1. It is stressed that only realistic

ship designs are taken into account and mathematical hull definitions, such as the Wigley hull, are disregarded due to the missing similarity to contemporary hull forms. The database is characterized by $L_{pp} \in [90, 355]$m, $C_B \in [0.503, 0.858]$ and $Fn \in [0.0, 0.3]$. According to Papanikolaou (2014), the L/B and B/T ratios resemble the current world fleet only to some extent: The L/B ratios show small variation $L/B \in [5.05, 7.5]$ compared to $L/B \in [4.5, 8.5]$ for the world fleet. (2) The B/T ratios are slightly higher than the values of the current world fleet with $B/T \in [2.5, 4.7]$ compared to $B/T \in [2.1, 4.0]$. The mentioned deviations are the first indications for systematic or rather *epistemic* uncertainty of the available data. The decomposition of the database regarding ship type is presented in Fig. 1(a) and reveals only minor deviations to the current (2019) composition of the world fleet according to gross tonnage, cf. Equasis (2019). However, general cargo ships are over-represented in the present database comparing 24% to only 4.5% in reality. In addition, the corresponding number of bulk carriers is too small comparing 16% to 34% for the world fleet as of 2019.

The uncertainty of the added resistance in seafaring conditions is a well-known problem, but is despite its relevance, not well studied. Park et al. (2015) provide uncertainty sources of respective model tests and Papanikolaou and Liu (2021) address this problem qualitatively. Initially, two uncertainty categories are introduced: (A) *Epistemic* or systematic uncertainty results from limitations of the available data, i.e. missing observations. For instance, in short waves ($\lambda/L_{pp} < 0.5$) and non-head wave conditions considerable epistemic uncertainty is present. The database consists of 1127 samples of 25 ships in total, but for $\beta < 180$ deg. only 470 samples are included corresponding to 42% of the total number of samples, whereas the sample size for $\lambda/L_{pp} < 0.5$ is even smaller at 231, cf. Fig. 1(b). (B) *Aleatoric* or statistical uncertainty on the other hand, is due to the stochastic (non-deterministic) character of the observations themselves and is thus an inherent part of the data. Hence, the aleatoric uncertainty does not decrease by increasing the sample size of the model test results. Soize (2017) and Hüllermeier and Waegeman (2021) provide comprehensive overviews of uncertainty quantification and the distinction of uncertainty components.

Park et al. (2015) list five sources of aleatoric uncertainty regarding experiments for the added resistance: (1) Instrument uncertainty, (2) mass distribution uncertainty, (3) calibration uncertainty, (4) measurement uncertainty and (5) data reduction uncertainty. Papanikolaou and Liu (2021) also highlight the uncertain influence of wave steepness on data quality in short waves as well as limitations of towing tanks for measuring the added resistance in short and especially oblique as well as following waves due to, for example, wall effects, such as reflection. Finally, experiments pertaining the added resistance are carried out under partial dynamic similarity, i.e. constant Froude number. Thus, viscous forces are, in fact, relatively more pronounced than in full scale, which might also introduce additional uncertainty in short waves.

Conducting residual analysis of the original approach of Liu and Papanikolaou (2020), it is observed that the residual variance not only increases with the magnitude of the added resistance coefficient, but also with increasing forward speed (Fn). This phenomenon is known as heteroskedasticity, i.e. the variance growth with an increase of a target variable. The well-established p-test of Breusch and Pagan (1979) identifies, whether regression models exhibit heteroskedastic residuals and violate the normality of the residuals and thus the assumption of linear regression. If the test statistics have a *p*-value below a sensible threshold (e.g. $p < 0.05$), the null hypothesis H_0 of homoskedasticity is rejected. The residuals of the Liu and Papanikolaou formula resulting from application to the present dataset have a *p*-value of $p \ll 0.05$, which proves heteroskedasticity. In the following parameter calibration we attempt to reduce the degree of heteroskedasticity. Furthermore, the residual analysis revealed a structural break at $C_B = 0.70$, which is reflected in the experimental data of the bluntness coefficient B_f of the original NMRI formula, cf. Fig. 6 in Liu and Papanikolaou (2016).

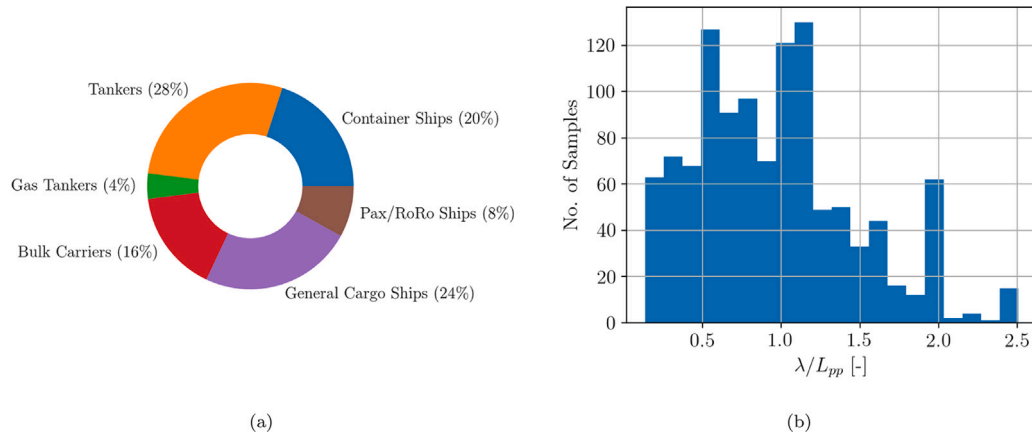[1] "All models are wrong, but some are useful".

Fig. 1. Breakdown per ship type (a) and relative wavelength (b).

Table 1
Database of benchmark model test data.

Ship	L_{pp} [m]	B [m]	T [m]	C_B [–]	Fn [–]	β [deg.]	Reference
170BC	279	45	16.5	0.858	0.128	{180,150,120, 90,60,30,0}	Kashiwagi et al. (2004)
DTC	355	51	14.5	0.661	{0.0,0.052,0.139}	{180,150,120, 90,60,30,0}	Shigunov et al. (2018) and Sprenger et al. (2017)
HSVA cruise ship	220.27	32.04	7.2	0.654	0.233	{180,150,120, 90,60,30,0}	Valanto and Hong (2015)
KCS	230	32.2	10.8	0.651	0.26	{180,90,0}	Simonsen et al. (2014) and Stocker (2016)
KVLCC2	320	58	20.8	0.808	{0.0,0.055,0.11,0.142}	{180,150,120, 90,60,30,0}	Shigunov et al. (2018), Guo and Steen (2010) and Sadat-Hosseini et al. (2013)
S175	175	25.4	9.5	0.572	{0.15,0.2,0.25,0.3}	{180,150,120, 90,60,30,0}	Fujii and Takahashi (1975) and Nakamura and Naito (1977)
S-Cb84	178	32.26	11.57	0.84	{0.0,0.048,0.099,0.166}	{180,150, 90,30}	Yasukawa et al. (2019)
S-LNGC	290	45	11.5	0.79	0.19	{180,150,120, 60,30,0}	Kim et al. (2021)
SNU tanker	323	60.0	21.0	0.83	0.137	{180,150,120, 90,60,30,0}	Park et al. (2019)
Aframax tanker	239	44	13.6	0.835	0.154	180	Oh et al. (2015)
Bulk carrier	285	50	18.5	0.829	{0.0,0.05,0.1,0.15}	180	Kadomatsu et al. (1989)
Container ship	300	40	14	0.66	0.247	180	Tsujimoto et al. (2008)
Product carrier	145.4	23.4	8	0.757	0.177	180	Li et al. (2016)
S.A. van der Stel	152.5	22.8	9.14	0.563	{0.15,0.2,0.25,0.3}	180	Gerritsma and Beukelman (1972)
			5.2	0.503	{0.15,0.2,0.25,0.3}	180	Gerritsma and Beukelman (1972)
RIOS bulker	2.0	0.333	0.107	0.8	0.18	180	Kashiwagi et al. (2019)
RoPax	90	17.82	4.2	0.549	0.087	180	Sprenger et al. (2015)
Series 60 ($C_B = 0.6$)	121.96	16.254	6.492	0.6	{0.266,0.283}	180	Strøm-Tejsen et al. (1973)
Series 60 ($C_B = 0.65$)	121.96	16.816	6.73	0.65	{0.237, 0.254}	180	Strøm-Tejsen et al. (1973)
Series 60 ($C_B = 0.7$)	121.96	17.42	6.79	0.7	{0.207,0.222}	180	Strøm-Tejsen et al. (1973)
Series 60 ($C_B = 0.75$)	121.96	18.062	7.22	0.75	{0.177,0.195}	180	Strøm-Tejsen et al. (1973)
Series 60 ($C_B = 0.8$)	121.96	18.757	7.495	0.8	{0.147,0.165}	180	Strøm-Tejsen et al. (1973)
SR221C tanker	320	58	19.3	0.803	0.15	180	Kashiwagi et al. (2004)
Supramax tanker	192	36	11.2	0.84	0.17	180	Yu et al. (2017)
ULYSSES tanker	187.3	32	12	0.82	{0.06,0.12,0.17}	180	SSPA (2012)
WILS II container ship	324	48.4	15	0.602	0.183	180	Söding et al. (2014)

This structural break is highlighted in Fig. 2, in which the residuals are presented for the block coefficient C_B and the Froude number Fn.

It is assumed that the variance increase with higher forward speed is the same as for slender ships (i.e. $C_B < 0.70$) since they usually operate under higher speeds compared to more full-type vessels. In fact, they are subsequently also experimentally examined at these relatively high speeds. Furthermore, the residuals are reasonably balanced for $C_B \geq 0.70$, whereas the formula provides in general too small values for slender ships. The biased residuals for the fast and fine ships strongly motivates splitting the database at $C_B = 0.70$ and obtaining two separate updated parameter vectors for the Liu and Papanikolaou (2020) method. The resulting two datasets are presented in Fig. 3 and comprise 598 for $C_B \geq 0.70$ and 529 samples for $C_B < 0.70$.

It is appreciated that not only the residuals, but also the slender ships' added resistance coefficients are characterized by larger variance compared to the corresponding values of the blunt-type ships. Moreover, it is obvious that the database is biased (unbalanced) towards head wave conditions and relative wavelengths $\lambda/L_{pp} \leq 1.0$, cf. Fig. 1(b). In addition, the dataset for the full-type ships is slightly larger, but in general the database is considered as sufficiently balanced. Furthermore, it is apparent in Fig. 3 (mainly in (a)) that the maximum of the non-dimensional added resistance in head waves appears at a slightly larger wavelength than ship length and reaches zero asymptotically for wavelengths larger than twice the ship's length. Two possible reasons leading to the notable variance in head wave conditions may be attributed to inaccuracies in the experimental setup and the use of
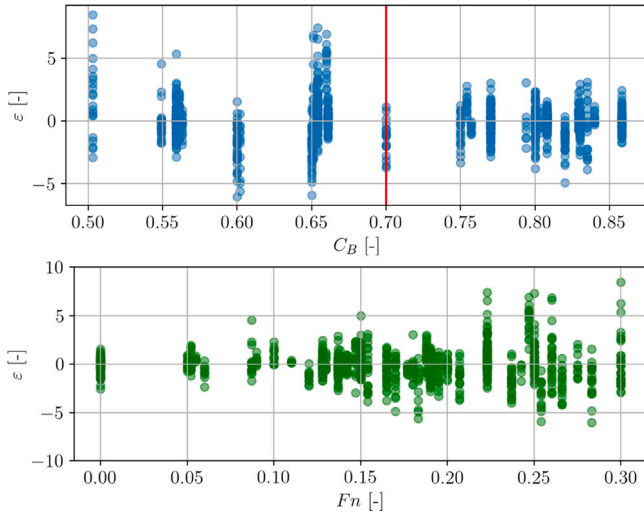
Fig. 2. Residual analysis ε for block coefficient C_B and Froude number Fn. It is noted that the residuals are defined as $\varepsilon_i = y_i - \hat{y}_i$, where y_i is considered as the ground truth.

different wave amplitudes and thus wave steepnesses. Additionally, the incident wave amplitude ζ_a is neither spatially nor temporally constant in a towing tank (not even in one run of experiments). Since, the added resistance is proportional to ζ_a^2, spatio-temporal uncertainty of ζ_a leads to large scatter or rather aleatoric uncertainty. The attributed scatter vanishes when normalizing the experimental data with the measured ζ_a^2, which is also indicated by Maruo (1957). Another reason for uncertainty might be different natural frequencies in heave as well as pitch due to different advance speeds. Despite the theory of Faltinsen's asymptotic formula, the non-dimensional added resistance is trivially zero for $\lambda/L_{pp} = 0.0$. In short waves, the non-dimensional added resistance reaches nearly a constant value, but in some cases tends towards large magnitude and variance, which is partly owed to wave breaking effects at the bow. The latter phenomenon is presumably attributed to relatively high wave steepnesses H/λ, as Park et al. (2015) pointed out. Sigmund (2018) comes to the same conclusion by comparing three different wave steepnesses and their influence on the non-dimensional added resistance by conducting RANS simulations. It was found that increasing the wave steepness in turn slightly decreases the magnitude of C_{AW} in long waves and increases C_{AW} in non-linear fashion in short waves. It is thought that all experiments were carried out with different wave steepnesses, but more importantly under relatively large (and varying) wave amplitudes ζ_a magnifying small added resistance values in order to be measured sufficiently. Unfortunately, it is not possible to study the dependency of the residuals on the wave steepness or wave amplitude as most publications do not specify these values. Moreover, the influence of forward speed $Fn \neq 0$ and non-linear effects, such as wave breaking, contribute to the increase of C_{AW} in relatively short waves. Ultimately, the added resistance is largest in head to beam waves and only of comparably small magnitude in beam to following seas. Hence, the DTU design tool (cf. Nielsen (2015) and Martinsen (2016)) and the NMRI formula (cf. Tsujimoto et al. (2008)) assume the added resistance in waves to be equal to zero for the latter range of wave headings as it is thought that the degree of uncertainty exceeds the actual magnitude of the wave-added resistance. However, the adoption of semi-empirical formulas for the prediction of added resistance depends on the appropriate representation of physical phenomena including the possibility of estimating a negative value like the wave drift force in following waves. Therefore, the presented approach imposes no restrictions of the considered relative wave heading angle β.

3. Applied methodology

The present parameter calibration procedure requires the minimization of one objective, or rather loss, function $\mathfrak{L}$ which is subject to the parameter vector x_i and independent variables θ_m. This methodology can be expressed mathematically as: $\min \mathfrak{L}(x_i | \theta_m)$. Herein, the meta-heuristic Particle Swarm Optimization (PSO) algorithm is applied for finding the global minimum due to the highly discontinuous and scattered loss landscape, i.e. without a smooth gradient. The randomized, population-based PSO method was proposed by Kennedy and Eberhart (1995) and mimics the social behavior of bird or fish swarms. Each particle is characterized by a position vector $x_i = [x_1, x_2, \ldots, x_n]^T$ in the n-dimensional search space $\Omega \subseteq \mathbb{R}^n$ and a corresponding velocity vector v_i.

$$v_j^{k+1} = \Phi v_j^k + c_1 r_1 (P_j^k - x_j^k) + c_2 r_2 (G_j^k - x_j^k) \tag{1}$$

In Eq. (1), j denotes the jth particle and k is the iteration index. Moreover, Φ is the inertia weight or the momentum term, whereas c_1 and c_2 are two strictly positive weights for balancing the exploitation of P_j^k and G_j^k – the particle's best and the global best parameter combination, respectively. The social behavior is implemented using the globally or locally best solutions and c_1 and c_2 are considered as cognitive and social impact, respectively. Hence, the choice of the c_1 and c_2 coefficients is of high importance for the PSO algorithm performing well throughout the optimization procedure. On the other hand, r_1 and r_2 are two random variables drawn from a uniform distribution between 0 and 1. The velocity update of the parameter vector x_j^{k+1} is defined in Eq. (2).

$$x_j^{k+1} = x_j^k + v_j^{k+1}. \tag{2}$$

As has been described, balancing the cognitive and social impact coefficients is decisive for the convergence of the optimization procedure. Therefore, the *adaptive* PSO algorithm proposed by Zhan et al. (2009) is applied in the present work. Ultimately, it suffices to specify initial values for Φ, c_1 and c_2, and after each iteration the corresponding values are changed dynamically. The updated values are determined by the spread from the global optimum. In the present paper, the implementation of the Python software package *Pymoo* developed by Blank and Deb (2020) is employed and we choose the following initial values: $\Phi = 0.9$, $c_1 = 2.0$ and $c_2 = 2.0$. Moreover, an upper velocity bound is enforced in order to prevent the rapid movement of particles in the search space Ω, and the corresponding value is defined as $v_{max} = 0.025$. The population size is 25 and termination criteria are the following: (1) The threshold of the change of objective function is set to 10^{-6} and (2) the maximum number of generations is 10^5. Moreover, the initial solutions are not drawn randomly, but by Latin Hypercube Sampling which is a stratified sampling technique proposed by McKay et al. (1979). The computations were performed on an Intel© Core™ i7-8565U CPU, 1.80 GHz with 16 GB physical memory (RAM).

3.1. Adapted semi-empirical framework

The mathematical framework for the prediction of the added resistance R_{AW} combines a variation of the empirical approach of Jinkine and Ferdinande (1974) for long waves and a simplified version of Faltinsen's asymptotic formula in the short wavelength regime . Further simplifications and empirical regression coefficients were introduced, Liu et al. (2015). The original approach originated in the SHOPERA EU-project which addressed the rational calculation of the hydrodynamic behavior of ships in waves. In the present paper, we make use of the latest version of this method which imposes no restrictions on the wave heading, Liu and Papanikolaou (2020). The input vector to the formula is the following $\theta_m = [\lambda, L_{pp}, B, T_f, T_a, C_B, L_E, L_R, \beta, k_{yy}, Fn]^T$. It is noted that T_a and T_f denote the draft at in aft and forward position, respectively. Moreover, L_E and L_R are the lengths of entrance and describe the length between the fore- and aft-most position to the start

(a) $C_B \geq 0.70$ dataset with 598 samples.

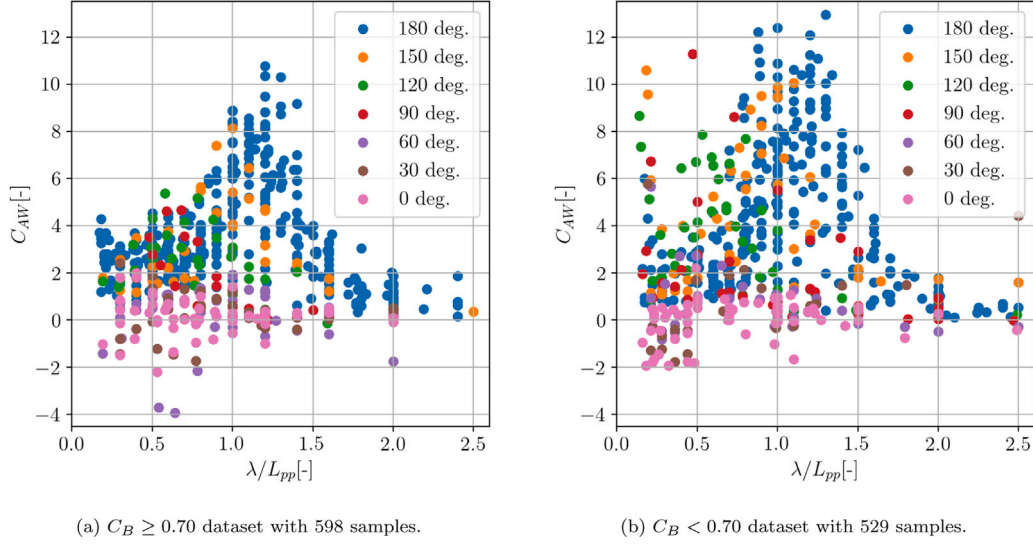(b) $C_B < 0.70$ dataset with 529 samples.

Fig. 3. Samples of the C_{AW} database. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and end of the parallel midship body. The parameter vector x_i comprises 29 constants of the initial formulation by Liu and Papanikolaou (2020), which are presented in Table 2 alongside the corresponding updated values. It is stressed that the signs of the parameters are not changed and only strictly positive values are considered in the parameter (or position) vector x_i. For detailed information and guiding sketches, cf. Liu and Papanikolaou (2020). In contrast to the original paper, we present the non-dimensional added resistance coefficient C_{AW}, cf. Eq. (3).

$$C_{AW} = \frac{R_{AW} L_{pp}}{\rho g \zeta_a^2 B^2} \tag{3}$$

It is noted that g denotes the gravity acceleration and ρ reflects the sea water density. In this work, the non-dimensional added resistance coefficient C_{AW} is decomposed into (1) a base term $C_{AW,M}$ resulting from ship motion and (2) a residual term for bow reflection $C_{AW,R}$ due to diffraction in short waves. The latter contribution is considered as a practical correction as the method by Jinkine and Ferdinande (1974) shows only satisfactory accuracy in medium to long waves. A similar correction approach is described by Tsujimoto et al. (2008). Following the approach of Liu and Papanikolaou (2020), the reflection contribution is calculated using a modified version of Faltinsen's formula and takes the following form.

$$C_{AW,R} = \sum_{i=1}^{4} C_{AW,R}^i. \tag{4}$$

The reflection component is expressed in Eq. (4) separately for each of the four segments S_i. Note that S_1 and S_2 are the port and starboard parts of the bow, whereas S_3 and S_4 are the port and starboard parts of the stern part, respectively.

$$C_{AW,R}^1 = x_2 \frac{L_{pp}}{B} \alpha_{T*} \left\{ \sin^2(E_1 - \beta) + \frac{2\omega_0 U}{g} \right.$$
$$\left. \times \left[\cos E_1 \cos(E_1 - \beta) - \cos \beta \right] \right\} \left(\frac{0.87}{C_B} \right)^{(1+4\sqrt{Fn})f(\beta)} \tag{5}$$

$$C_{AW,R}^2 = x_2 \frac{L_{pp}}{B} \alpha_{T*} \left\{ \sin^2(E_1 + \beta) + \frac{2\omega_0 U}{g} \right.$$
$$\left. \times \left[\cos E_1 \cos(E_1 + \beta) - \cos \beta \right] \right\} \left(\frac{0.87}{C_B} \right)^{(1+4\sqrt{Fn})f(\beta)} \tag{6}$$

We note that ω_0 is the intrinsic wave frequency, the ship's advance speed is denoted with U in [m/s] and α_{T*} is the draft coefficient which

will be defined at a later stage. The form factor $\left(\frac{0.87}{C_B} \right)^{(1+4\sqrt{Fn})f(\beta)}$, which accounts for non-linear effects induced by the fore shoulder, is neglected (i.e. equal to 1) in following wave conditions (cf. Eqs. (7) & (8)) and not considered in the parameter calibration as its modification affected convergence behavior negatively.

$$C_{AW,R}^3 = -x_2 \frac{L_{pp}}{B} \alpha_{T*}$$
$$\times \left\{ \sin^2(E_2 + \beta) + \frac{2\omega_0 U}{g} [\cos E_2 \cos(E_2 + \beta) - \cos \beta] \right\} \tag{7}$$

$$C_{AW,R}^4 = -x_2 \frac{L_{pp}}{B} \alpha_{T*}$$
$$\times \left\{ \sin^2(E_2 - \beta) + \frac{2\omega_0 U}{g} [\cos E_2 \cos(E_2 - \beta) - \cos \beta] \right\} \tag{8}$$

It is stressed that C_{AWR}^1 is valid for $E_1 \leq \beta \leq \pi$, C_{AWR}^2 in $\pi - E_1 \leq \beta \leq \pi$, $C_{AW,R}^3$ is valid for angles in $0 \leq \beta \leq \pi - E_2$ and C_{AWR}^4 is defined for $0 \leq \beta \leq E_2$. Both E_1 and E_2 are considered as entrance angles of the aft and bow part of the ship of interest.

$$E_1 = \arctan \left(\frac{0.99 \cdot B/2}{L_E} \right) \tag{9}$$

$$E_2 = \arctan \left(\frac{0.99 \cdot B/2}{L_R} \right) \tag{10}$$

The non-dimensional added resistance due to motion is calculated in the following way using several regression coefficients:

$$C_{AW,M} = x_1 a_1 a_2 a_3 \bar{\omega}^{b_1} e^{\frac{b_1}{d_1}(1 - \bar{\omega}_1^d)}. \tag{11}$$

In the following, the heading dependent regression coefficient a_1 for the calculation of the motion component $C_{AW,M}$ is presented.

a_1

$$= \begin{cases} x_3 C_B^{x_4} (x_5 k_{yy})^2 \left(\frac{x_6}{C_B} \right)^{-(1+Fn)\cos\beta} \left(\ln \frac{B}{T_{max}} \right)^{-1} \frac{(1-2\cos\beta)}{x_7} & \text{for } \beta \in [\frac{\pi}{2}, \pi] \\ \text{linear interpolation between } \beta = \pi/2 \text{ and } \beta = 0 & \text{for } \beta \in]0, \frac{\pi}{2}[\\ f(U, V_g) & \text{for } \beta = 0 \end{cases} \tag{12}$$

The coefficient a_1 requires special attention in beam to following waves and its determination is described briefly in the next paragraph. For elaborate explanations and the theoretical derivation consult the original paper of Liu and Papanikolaou (2020).

The calculation of a_1 in stern oblique waves $\beta \in]0, 90[$ deg. is performed by linear interpolation between the beam and following
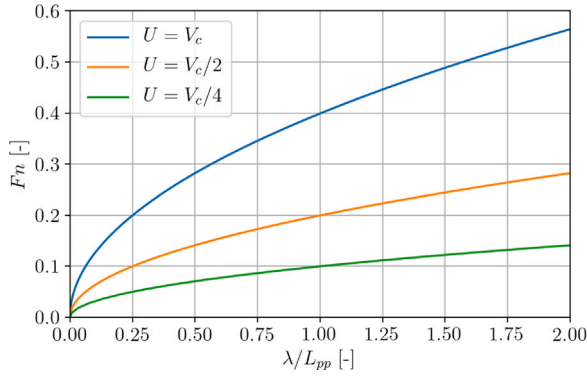
Fig. 4. Ship forward-speed for which the maximum added resistance changes its sign dependent on the relative wavelength, according to Liu and Papanikolaou (2020). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

sea case. The coefficient a_1 in following waves is highly dependent on the wave's phase velocity V_c as well as the ship advance speed U. The calculation is based on the investigation of the energy exchange between ship and incident wave, cf. 4. Following the thoughts of Liu and Papanikolaou (2020): (1) For $Fn = 0$, the added resistance and thus a_1 are negative, (2) on the green curve ($U = V_c/4$) it is assumed that the radiation component vanishes ($a_1 = 0$) and (3) on the orange curve ($U = V_c/2$) it is thought that C_{AW} is equal to the magnitude as in the head wave case, but for non-forward speed, i.e. a_1 is positive, but calculated for $Fn = 0$. In addition, (4) the value at any condition in between will be obtained by linear interpolation and (5) conditions above the orange curve will be approximated using the relative Froude number $Fn_r = \frac{U - V_c/2}{\sqrt{gL_{pp}}}$. The remaining regression coefficients are defined as follows:

$$a_2 = \begin{cases} 0.0072 + x_8 Fn & \text{for } Fn < 0.12 \\ Fn^{x_9} e^{(-x_{10} Fn)} & \text{for } Fn \geq 0.12 \end{cases} \tag{13}$$

$$a_3 = 1.0 + x_{11} \text{atan} \frac{|T_a - T_f|}{L_{pp}} \tag{14}$$

$$b_1 = \begin{cases} x_{12} & \text{for } \bar{\omega} < 1 \\ -x_{13} & \text{elsewhere} \end{cases} \tag{15}$$

$$d_1 = \begin{cases} x_{14} \left(\frac{L_{pp} C_B}{B} \right)^{-x_{15}} & \text{for } \bar{\omega} < 1 \\ -x_{14} \left(\frac{L_{pp} C_B}{B} \right)^{-x_{15}} \times \left(4 - \frac{x_{16} \text{atan}|T_a - T_f|}{L_{pp}} \right) & \text{elsewhere} \end{cases} \tag{16}$$

$$\bar{\omega} = x_{17} \sqrt[3]{k_{yy}} \sqrt{\frac{L_{pp}}{\lambda}} \left[1 - \frac{x_{18}}{C_B} \left(\ln \frac{B_m}{T_{max}} - \ln x_{19} \right) \right] \left(\frac{C_B}{x_{20}} \right)^{x_{21}}$$
$$\times \left[(-x_{22} Fn^2 + x_{23} Fn) |\cos \beta| + \frac{x_{24}(13 + \cos 2\beta)}{x_{25}} \right]. \tag{17}$$

It is noted that T_{max} is defined as $T_{max} = \max(T_a, T_f)$. The remaining definitions are needed for the calculation of the reflection part, but are presented in the last part as modifications are introduced in these equations. Initially, the constraint enforcing the added resistance to be zero for $\lambda/L_{pp} > 2.5$ is removed in the draft coefficient α_{T_*} as it led to discontinuous results after initial calibration attempts. Instead, the database was enriched with multiple artificial samples in the respective range of wavelengths with $C_{AW} = 0.0$ introducing domain knowledge into the otherwise solely data-driven methodology. Conversely, Cepowski (2020) did not include this kind of boundary condition in his empirical machine learning approach and in some cases the added resistance showed an unphysical increase in long waves. Referring to the initially defined uncertainty categories, the epistemic uncertainty

in very long waves is reduced by the definition of artificial samples.

$$f(\beta) = \begin{cases} -\cos \beta & \text{for } \beta \in [\pi - E_1, 0] \\ 0 & \text{for } \beta < \pi - E_1 \end{cases} \tag{18}$$

$$\alpha_{T_*} = 1 - e^{-x_{26} \pi \left(\frac{T^*}{\lambda} - \frac{T^*}{x_{27} L_{pp}} \right)}. \tag{19}$$

For the bow segments S_1 and S_2 the expression $T^* = T_{max}$ is valid. In the aft part, however, the following Eq. (20) is used. In the original paper, a condition was imposed at $C_B = 0.75$, but in the present approach, two separate parameter vectors are obtained with the transition at $C_B = 0.70$. Hence, the equation of T^* for the stern simplifies to:

$$T^* = \frac{T_{max} \left(x_{28} + \sqrt{|\cos \beta|} \right)}{x_{29}}. \tag{20}$$

The presented framework acts as a model in the following calibration or training procedure. Adopting the statistical notation, we consider the estimate of the semi-empirical framework C_{AW} as the approximation $\hat{y}_i$ and the database presented in Section 2 reflects the ground truth y_i. Ultimately, the presented formulation as well as the derived parameters are made publicly available on GitLab.[2]

3.2. Loss functions

As the performance of the semi-empirical method is not only subject to the uncertainty of the data, e.g. experimental uncertainty, but also to model form uncertainty, a parameter calibration approach is conducted. Essentially, we attempt to deduce an optimized parameter vector x_i using an objective or rather loss function and labeled data, i.e. the database. Generally speaking, this methodology is defined – in machine-learning terms – as Structural Risk Minimization (SRM) and its concrete form is defined in Eq. (21), cf. Wang et al. (2020).

$$\min \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}_{x_i}(\hat{y}_i) \quad \text{where } \hat{y}_i = f(\theta_m) \tag{21}$$

In the above-mentioned equation, N corresponds to the number of samples and f denotes the previously presented semi-empirical formulation. The choice of the loss function $\mathcal{L}$ is highly influential on the quality of the results and is affected by the presence of outliers and noise in the data. For gradient-based optimization, e.g. stochastic gradient descent algorithms in deep learning, the differentiability of the loss function is of high importance. In this work, however, we use a gradient-free optimizer, i.e. particle swarm optimization. In Section 2 it was stated that the employed datasets are – the one for slender ships in particular – characterized by large variance and heteroskedasticity. For this reason, we investigate the effect of four different loss functions: (1) the squared loss, (2) the absolute loss, (3) the Huber loss and (4) the log-cosh loss. The latter two are, in fact, hybrid versions of the former two.

1. **Squared Loss**

$$\mathcal{L}_{L2} = (y_i - \hat{y}_i)^2. \tag{22}$$

The squared loss or $L2$ error is the most commonly used loss function and puts strong weight on outliers and variance.

2. **Absolute Loss**

$$\mathcal{L}_{L1} = |y_i - \hat{y}_i|. \tag{23}$$

The $L1$ or absolute error is robust and expresses the magnitude of errors without considering their direction.

Comparing the absolute to the squared loss, it proves that the $L2$ error is more stable and provides a closed form solution due to its continuous

---

[2] https://gitlab.gbar.dtu.dk/mamit/RAW_Formula.

(a) Comparison of four different loss functions.

(b) Quantile loss functions for three different γ.
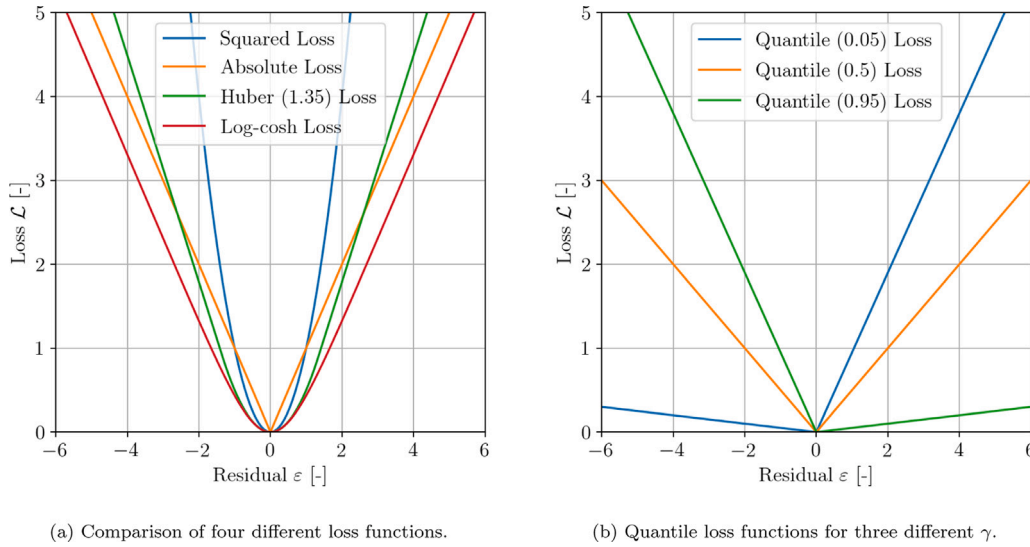
Fig. 5. Illustration of the behavior of the individual loss functions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

derivative, Wang et al. (2020). Essentially, the Huber and the log-cosh losses are considered as elegant combinations of the $L1$ and $L2$ losses by restricting the influence of outliers and by also having a continuous derivative. Hence, they combine the advantages of both $L1$ and $L2$ loss.

3. **Huber Loss**

$$\mathcal{L}_{Huber} = \begin{cases} \frac{1}{2}(\hat{y}_i - y_i)^2 & \text{for } |\hat{y}_i - y_i| \leq \delta \\ \delta|\hat{y}_i - y_i| - \frac{1}{2}\delta^2 & \text{otherwise} \end{cases} \qquad (24)$$

The Huber loss function is a piecewise function of squared and absolute loss. The transition is controlled by the hyperparameter δ and samples within this boundary use the squared loss, whereas outside the boundary it is vice versa. In doing so, the drawback of the squared loss can be diminished, i.e. putting less weight on outliers and thus avoiding overfitting. The use of the Huber loss is more complex due to the hyperparameter δ, but Huber himself suggested the robust value of $\delta = 1.35$, cf. Huber (1981).

4. **Log-cosh Loss**

$$\mathcal{L}_{log-cosh} = \log(\cosh(\hat{y}_i - y_i)). \qquad (25)$$

The Log-cosh loss function is the logarithm of the hyperbolic cosine of the residual, i.e. when the error is small, the loss is close to $\frac{1}{2}\varepsilon^2$ and when the error is large, the loss is approximately $|\varepsilon| - \log 2$. Hence, its behavior is very similar to the Huber loss function, i.e. has all of its advantages, but it does not require the definition (and optimization) of a hyperparameter.

The individual loss functions are depicted in Fig. 5(a) and it is visible that the squared loss punishes large errors the most, whereas the absolute loss is the only discontinuous function. For more information about the described loss functions, cf. Wang et al. (2020).

The objective or loss function $\mathcal{L}$ of the parameter calibration is not subject to any equality or inequality constraints, but the search space Ω itself is constrained by lower and upper bounds of the variables $x_i^l \leq x_i \leq x_i^u, i = 1, \ldots, n$. The constraints of the design space are highly influential on the quality of the results and were found in an iterative process. Ultimately, the described methodology has a resemblance to a supervised regression approach in machine learning.

3.3. Quantile regression

Quantile Regression is a non-linear regression approach proposed by Koenker and Hallock (2001). Its application is considered herein as

it is robust to the presence of outliers and considered as an extension to linear regression, i.e. in case the corresponding assumptions are not valid. This includes linearity, homoskedasticity and normality. As pointed out it Section 2, none of these prerequisites are met in the present work. Thus, for the determination of the uncertainty bounds, SRM is employed again (cf. Eq. (21)) and the quantile loss function is minimized. In general, this particular loss function is in its original form an extended version of the absolute loss, which qualitatively matches Eq. (23) when $\gamma = 0.5$ (i.e. the 50th percentile). The original quantile loss function as proposed by Koenker and Hallock (2001) is given in Eq. (26).

$$\mathcal{L}_\gamma(y, \hat{y}) = \begin{cases} (\gamma - 1)(y_i - \hat{y}_i) & \text{for } y_i < \hat{y}_i \\ \gamma(y_i - \hat{y}_i) & \text{otherwise} \end{cases} \text{ where the quantile is } \gamma \in [0, 1]$$

$$(26)$$

Using a tilted or *pinball* loss function in SRM enables the formula to reflect a specified quantile instead of the mean. Quantile loss functions adjust the weight of the individual sample's error according the given quantile γ, i.e. a smaller γ increases the magnitude of the loss of those samples with a negative residual ε and introduces more punishment for overestimation. For large γ it is vice versa. An illustration of several quantile loss functions are presented in Fig. 5(b) and for more elaborate details regarding quantile regression consult Koenker and Hallock (2001).

In this paper, a 90% prediction interval is desired and thus we choose two objective functions with $\gamma = 0.05$ and $\gamma = 0.95$. However, the actual definition of the used quantile loss function is dependent on the performance of the other four losses in the first stage, i.e. the parameter calibration. In addition, to the 29 parameters, an offset parameter x_{30} is defined and initialized with $x_{30} = \pm 0.1$, and is bounded by $x_{30} \in [0.0, 0.2]$. The updated and newly obtained parameters as well as the corresponding constraints will be presented in the following section.

4. Results and discussion

4.1. Parameter calibration

The following part presents the obtained results and their interpretation. It is noted that all samples of the S-Cb84 (cf. Table 1) are left out of the training dataset as it acts as an out-of-sample test case. Moreover, the training dataset comprises artificial samples in longer waves $\lambda/L_{pp} > 2.5$ as an enforced boundary condition. In Table 2,

Table 2

Initial, bounding and final values of the parameter vector. Note, that the values in the parentheses reflect the parameters of lower and upper bound, respectively.

	Baseline	Lower constraint	Upper constraint	Final parameters	
				$C_B < 0.70$	$C_B \geq 0.70$
x_1	4.0	3.0	8.0	4.06 {3.03,4.27}	6.99 {3.0,4.59}
x_2	0.5625	0.42	0.705	0.705 {0.42,0.705}	0.425 {0.420,0.565}
x_3	60.3	50	100	56.23 {49.97,55.88}	70.92 {55.84,54.28}
x_4	1.34	1	2	1.185 {1.66,1.06}	2.0 {1.329,1.236}
x_5	4.0	2.5	5	4.057 {3.466,5.0}	3.013 {4.312,5.0}
x_6	0.87	0.65	1.09	0.65 {1.09,0.65}	0.9 {1.09,0.65}
x_7	3.0	2.25	5	2.25 {3.53,2.538}	3.726 {3.75,2.25}
x_8	0.1676	0.1	0.3	0.134 {0.12,0.21}	0.281 {0.146,0.21}
x_9	1.5	1.2	1.9	1.626 {1.483,1.502}	1.388 {1.793,1.549}
x_{10}	3.5	2.5	4.5	3.585 {2.62,3.583}	3.869 {3.566,3.954}
x_{11}	28.7	22	36	29.107 {30.924,27.079}	17.963 {27.628,33.777}
x_{12}	11	8	14	10.868 {12.974,8.2}	11.5 {12.298,14.0}
x_{13}	8.5	5.5	10.7	9.878 {10.7,6.37}	5.589 {10.7,6.37}
x_{14}	566	425	800	480.961 {707.94,425.0}	473.743 {679.27,425.0}
x_{15}	2.66	1.95	3.5	2.62 {1.95,3.5}	2.214 {2.309,3.5}
x_{16}	125	100	200	123.163 {93.535,121.801}	153.646 {130.0,118.02}
x_{17}	2.142	1.6	5.0	1.745 {2.166,2.041}	2.433 {2.231,2.043}
x_{18}	0.111	0.08	0.3	0.107 {0.14,0.0836}	0.0557 {0.083,0.105}
x_{19}	2.75	1.0	3.5	3.5 {2.896,3.5}	2.410{2.638,2.268}
x_{20}	0.65	0.4	0.825	0.825 {0.825,0.639}	0.813 {0.784,0.702}
x_{21}	0.17	0.125	0.3	0.125 {0.131,0.125}	0.3 {0.22,0.125}
x_{22}	1.377	1.03	2.0	1.03 {1.03,1.03}	1.008 {1.351,1.03}
x_{23}	1.157	1.0	1.45	1.201 {1.422,1.134}	1.076 {1.318,1.45}
x_{24}	0.618	0.4	0.9	0.723 {0.632,0.603}	0.778 {0.627,0.647}
x_{25}	14.0	10	20	13.630 {17.286,12.364}	19.241 {15.26,14.423}
x_{26}	4	3.0	5.0	5.0 {3.0,5.0}	3.029 {3.0,3.0}
x_{27}	2.5	1.0	4.0	3.15 {3.114,3.15}	3.845 {2.408,3.15}
x_{28}	{4.0,2.0}	1.0	6.0	2.897 {6.0,1.0}	3.123 {6.0,1.0}
x_{29}	{5.0,3.0}	2.0	7.5	5.552 {2.0,7.5}	4.527 {2.0,7.5}
x_{30}	±0.1	0.0	±0.2	±0.2	±0.2

the derived parameters as well as the corresponding initial values and constraints are presented.

The determination of sensible constraints is the most time-consuming part, as their values affect the balance between over- and under-fitting, i.e. the *bias–variance-tradeoff*, cf. Hastie et al. (2009) for detailed information. Essentially, the presented parameter bounds were obtained in an iterative manner and the final values are, in fact, considered as relatively narrow. This is a result of the fact that the lower bound does not capture the resonance region sufficiently with broader constraints. Other than that, it seems that the constraints are chosen well since only a few final parameters reach the lower or upper bound. Moreover, noticeable differences between initial values and final parameters are observable underlining the degree of reduced uncertainty. Also the parameter vectors for $C_B \geq 0.70$ and $C_B < 0.70$ – both of mean and the bounds – differ notably in their magnitudes.

In this contribution, we consider the coefficient of determination R^2, the *Pearson* correlation coefficient r and the variance σ^2 as metrics for the assessment of the prediction accuracy. It is noted that the corresponding equations are presented in Eqs. (27)–(29) and that overbar indicates then mean.

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y}_i)^2} \qquad (27)$$

$$r = \frac{\Sigma(y_i - \bar{y})\Sigma(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\Sigma(y_i - \bar{y})^2 \Sigma(\hat{y}_i - \bar{\hat{y}})^2}} \qquad (28)$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2. \qquad (29)$$

The performance of the individual loss functions is examined for the whole database (cf. Section 2) and not the separate sets. It is noted that the database corresponds to the initial one and not to the training set. The metrics are presented in Table 3 and the results of the squared loss are highlighted as the derived parameters are most accurate and subsequently adopted.

Table 3

Metrics for the performance of individual parameter combinations on the whole database.

	Baseline	Squared loss	Absolute loss	Huber(1.35) loss	Log-cosh loss
R^2	0.672	**0.731**	0.711	0.722	0.721
r	0.835	**0.856**	0.848	0.852	0.852
σ^2	2.540	**2.074**	2.174	2.121	2.124

As can be inferred in Table 3, a performance increase of approximately 8.8% considering R^2 is achieved by the proposed methodology using parameters of the $L2$ procedure and compared to the metrics of the initial formulation. Moreover, the Pearson correlation coefficient improved and the variance decreased significantly across the board. The $L2$ loss performs superior compared to the $L1$ loss and the metrics of Huber and Log-cosh losses are of similar character as expected due to their comparable behavior. Lastly, the Breusch Pagan test is conducted for the residuals of the updated formula, but they are still heteroskedastic, since we have to reject H_0 again as the *p*-value is $p \ll 0.05$. The heteroskedasticity is also visible in the residuals plot in Fig. 6(a), in which the residuals of the original and updated formulation are displayed in parallel.

In general, the distinction between epistemic and aleatoric uncertainty is imprecise in the context of model assessment, Soize (2017). Hence, we introduce the term model form uncertainty dividing into *parameter* and *structural* uncertainty, where the latter corresponds to the model's bias, e.g. inadequate assumptions. The parameter calibration leads to a decrease in variance of the residuals, but, as indicated by the Breusch Pagan p-test, to no reduction in heteroskedasticity, cf. Fig. 6. Moreover, for low to medium magnitudes of the added resistance coefficient it is perceptible that the method tends to provide higher values compared to the database (overprediction). Whereas in the regime of large C_{AW} magnitude, the approach is characterized by underprediction, cf. Fig. 6(a). This shows strongly biased residuals and
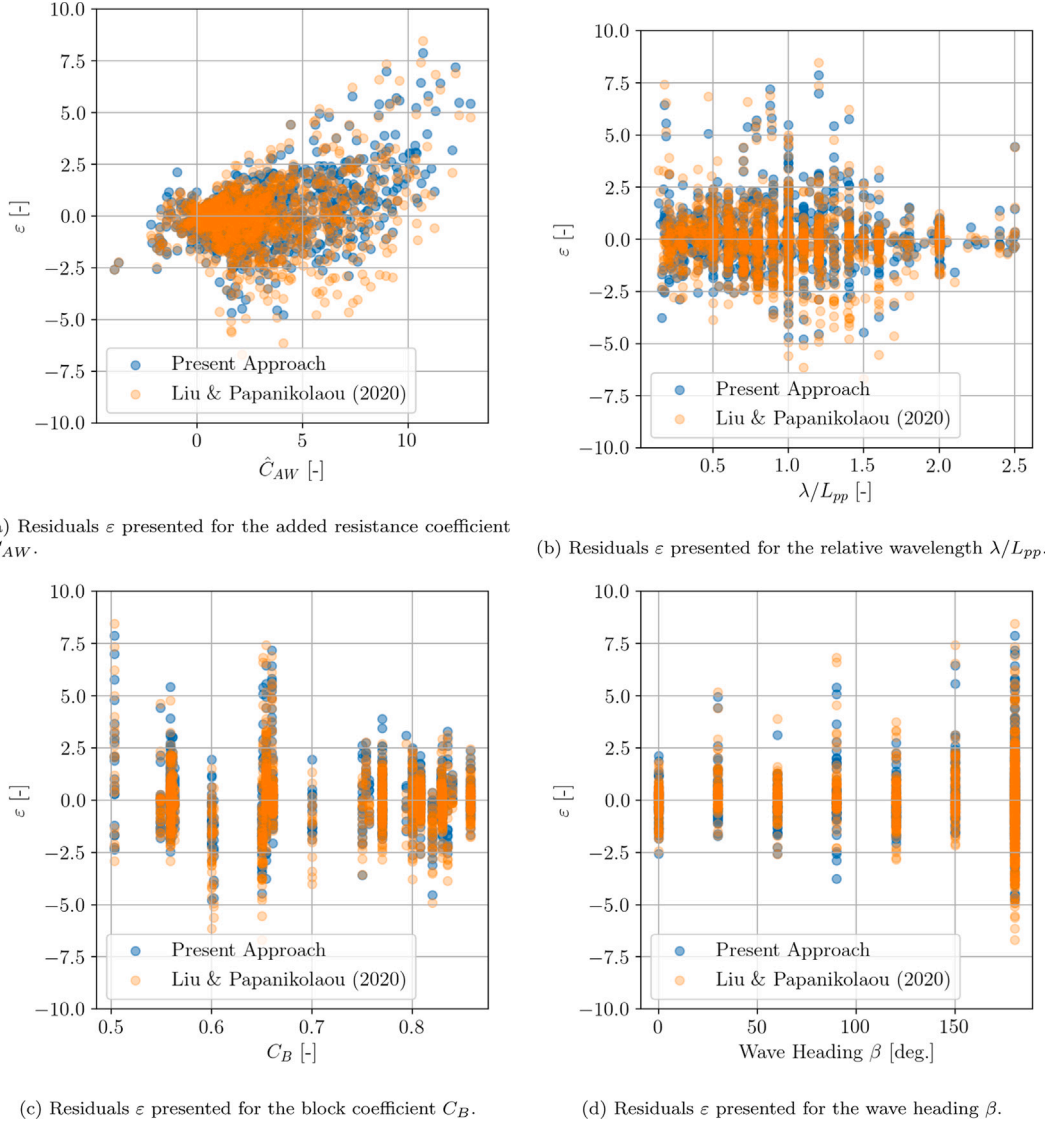
(a) Residuals ε presented for the added resistance coefficient C_{AW}.

(b) Residuals ε presented for the relative wavelength λ/L_{pp}.

(c) Residuals ε presented for the block coefficient C_B.

(d) Residuals ε presented for the wave heading β.

Fig. 6. Residual analysis comparing the presented approach to the original method. It is noted that the residuals are defined as $\varepsilon_i = y_i - \hat{y}_i$, where y_i is considered as the ground truth. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

therefore considerable structural uncertainty. In view of Fig. 6(b) it is concluded that the largest residual variance is present in short waves and in the region close to resonant conditions (i.e. $\lambda/L_{pp} \in [0.5, 1.5]$). Ultimately, only a minor variance decrease is noticeable considering the blockcoefficient C_B and the wave heading β, cf. Figs. 6(c) and 6(d). However, the residuals are still biased towards lower estimates for slender hull shapes and also in head oblique wave conditions (i.e. $\beta \in [180, 150]$ deg.) Moreover, we see heavily skewed residuals for $\beta = 30$ deg. and a shift in the bias from the original to the present method in beam waves ($\beta = 90$ deg.). Ultimately, a substantial variance reduction (parameter uncertainty) was achieved, but the bias (structural uncertainty) remained stable to a large extend.

Since the parameters derived by minimizing the $L2$ loss are adopted, the squared version of the quantile loss function is used in the second task: The outlier robust quantile regression. The modified quantile loss function is defined in Eq. (30).

$$\mathcal{L}_{\gamma, L2}(y, \hat{y}) = \begin{cases} (\gamma - 1)(y_i - \hat{y}_i)^2 & \text{for } y_i < \hat{y}_i \\ \gamma(y_i - \hat{y}_i)^2 & \text{otherwise.} \end{cases} \tag{30}$$

The convergence behavior of the determination of the four necessary uncertainty bounds is presented in Fig. 7, and it is stated that obtaining suitable parameters for the 90% prediction interval required more

iterations, i.e. more computational effort, for the slender-type ships. Moreover, the corresponding values for the blunt-type ships are not only characterized by fewer iterations, but also by lower magnitude of the respective loss functions. This is a result of the sensitivity to large variance (outliers) of squared loss functions and the larger inherent variance in the case of the slender-type ship database. Ultimately, all four optimization runs converged successfully and the resulting parameters are presented in Table 2 in the parentheses.

4.2. Validation in regular waves

In the following, the validation of the developed method against model test results and other established (semi-)empirical methods for the determination of the added resistance is carried out. In this respect, we investigate four ships and illuminate the influence of forward-speed on the uncertainty, the in- and out-of-sample prediction in various wave headings as well as the comparison to established methods, e.g. the DTU design tool. It is noted that the corresponding main particulars of the individual test cases are listed in Table 1.

As in the original paper we validate the proposed method considering the S-Cb84 bulk carrier as an out-of-sample test case for the assessment of the overall generalization capability. The case ship is
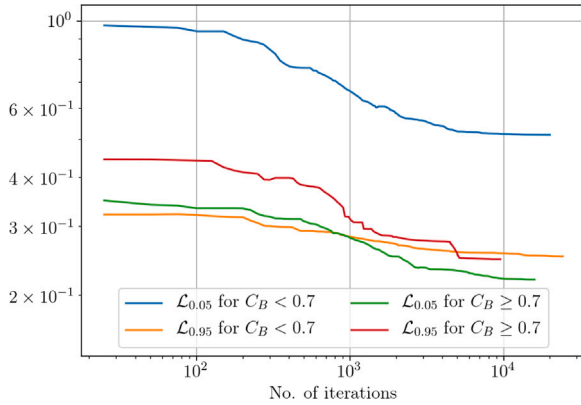
Fig. 7. Convergence study of the quantile loss functions for the individual bounds. Note that both axes are in logarithmic scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
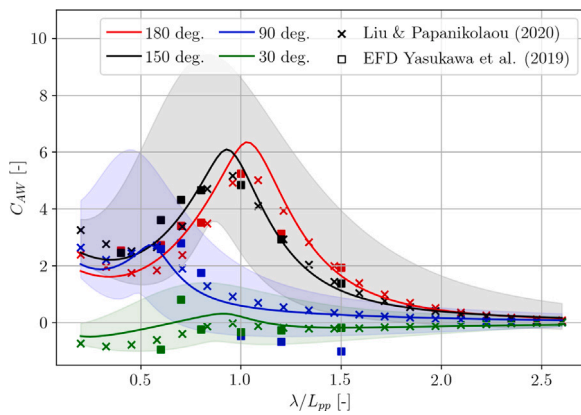

Fig. 9. Effect of forward speed investigating the S60 ship ($C_B = 0.8$) in head waves. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 8. Assessment of the performance for the out-of-sample test case S-Cb84 at $Fn = 0.1$. The uncertainty estimate for the head wave case is excluded, as it almost maps on top of the one for 150 deg. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 10. S175 container ship at $Fn = 0.15$ for a variety of wave headings. As a sidenote, all experimental samples are taken from Fujii and Takahashi (1975). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

evaluated in several heading angles at $Fn = 0.1$ and the model test data is taken from Yasukawa et al. (2019).

As can be inferred from Fig. 8, the accuracy of the proposed method appears as sufficient in comparison to the original method. Moreover, the prediction interval is remarkably accurate and insightful, since almost all EFD (Experimental Fluid Dynamics) samples stay within the uncertainty bounds and the interval's width has a distinct dependency on the relative wavelength. The largest uncertainty is present in resonant conditions and the width narrows for short and especially for long waves. It is noted that the uncertainty estimate for head waves is not included as it intersects with the corresponding one of $\beta = 150$ deg. to a large extent. Even though the S-Cb84 was not part of the training dataset, the mean predictions are of comparable quality to the original method of Liu and Papanikolaou (2020).

The second case study is the Series 60 ship ($C_B = 0.8$) and the behavior of the uncertainty estimate in head waves with increasing forward speed is studied. The proposed method is validated against the EFD data from Strøm-Tejsen et al. (1973) in Fig. 9.

In view of Fig. 9, it is stated that all model test results remain inside the 90% prediction interval and also that the dependency of the uncertainty on the advance speed is evident. It is appreciated that the mean estimates of the proposed method provide lower predictions for the entire relative wavelength regime. Moreover, the peak frequency according to the semi-empirical formula deviates minimally
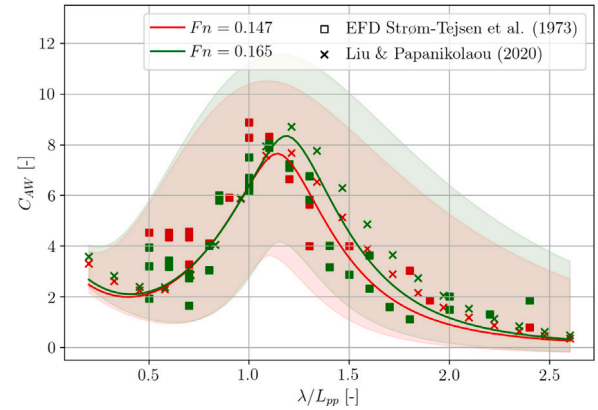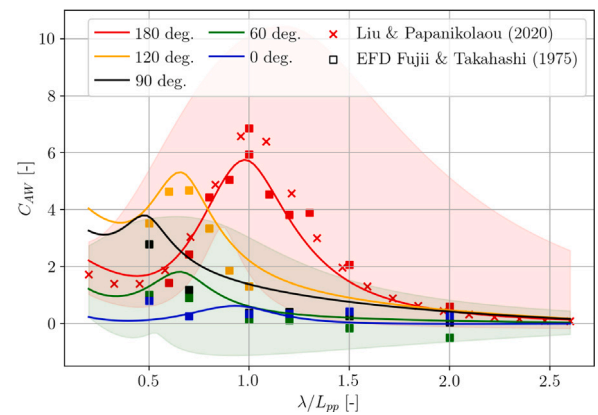
from the experimental resonance frequency. Lastly, the variance increase with larger advance speed is not only visible for the 90% prediction intervals, but also in the EFD data itself.

The slender S175 container ship is the third test case and will be examined for five different heading angles in Fig. 10 at $Fn = 0.15$. The experimental results of Fujii and Takahashi (1975) are considered in the following.

All of the EFD samples stay well within the uncertainty bounds and the present method's mean predictions show sufficient agreement to the model test results in all considered heading angles, but tend to underestimate the non-dimensional added resistance in long head waves. In addition, it evident that the proposed approach provides higher mean estimates in short waves, but lower ones in resonant conditions compared to the original method. This is the exact opposite to the observations made for the S-Cb84 bulk carrier. The final case study is a full-type LNG carrier in oblique waves with $\beta = 150$ deg. and moderate forward speed at $Fn = 0.19$. The proposed method will be compared to several well-established approaches, such as the STAwaveII (ITTC, 2017), the DTU design tool, the CTH formula proposed by Lang and Mao (2021) and the original approach by Liu and Papanikolaou (2020). All of these formulations were introduced in Section 1.1. The model test results of Kim et al. (2021) are considered as the ground truth in Fig. 11.

As one can appreciate in Fig. 11, the uncertainty bounds enclose the majority of the EFD results. Moreover, the mean estimate of the present
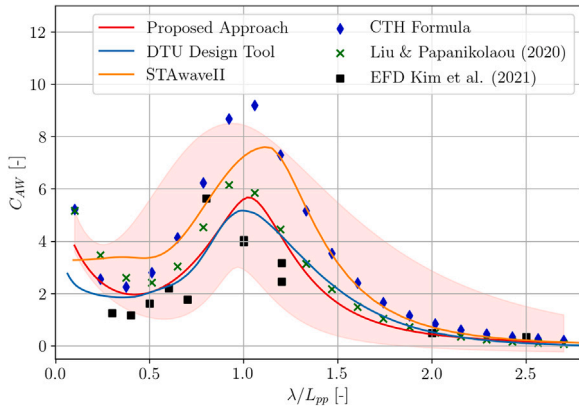
Fig. 11. Comparison of several prediction methods for the S-LNGC at $Fn = 0.19$ in $\beta = 150$ deg. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

method tends towards the lower bound in very short waves, which might indicate that the parameter constraints in the above methodology were selected too narrow as two samples fall outside the prediction interval in the short wave regime. Other than that, the present method approximates the non-dimensional added resistance well and is in satisfactory agreement with the DTU design tool. The STAwaveII and the CTH formula show the tendency to over predict the added resistance in the resonance region as well as for long waves and the estimates of the CTH formula even fall outside the 90% prediction interval. It is assumed that the deviation in case of the CTH formula estimates results from the two opposing approaches for the calculation of the reflection contribution as both the CTH and Liu and Papanikolaou (2020) formulas apply the (Jinkine and Ferdinande, 1974) approach for the motion component. Furthermore, Lang and Mao (2021) derive a different, simplified methodology for the calculation of added resistance in oblique and following seas from the analysis of heave, pitch and roll transfer functions and introduce two additional correction coefficients. Comparing the original formula of Liu and Papanikolaou (2020) to the calibrated one, we conclude that the same observations as in the case of the S60 ship are perceptible, i.e. lower estimates for the whole range of relative wavelengths. Lastly, it is not possible to properly assess the performance in very short waves (i.e. $\lambda/L_{pp} < 0.3$) as there are no available model test results.

4.3. General remarks and discussion

The mean estimate of the proposed method is of comparable accuracy to the original formula; however, the prediction interval proved to be highly insightful and shows that the reliable and accurate estimation of the second order wave drift force on ships is still a complex research area. The *quantitative* prediction of uncertainty around the added resistance is the main novelty of this contribution. In general, the shape of the 90% prediction interval matches expectations to a large extend, but especially in very short waves the uncertainties were assumed to be larger. In case of the S-LNGC this fact was most striking as the mean estimate converged towards the lower uncertainty bound and the EFD data fell out of the prediction interval, cf. Fig. 11. The narrow bandwidth in the short wave regime results from the epistemic uncertainty and the limited availability of samples in this range of relative wavelengths, cf. Fig. 1(b).

Moreover, one can conclude in view of Fig. 10, that the reflection component $C_{AW,R}$ has increased for slender ships relative to the formula of Liu and Papanikolaou (2020), whereas for the motion contribution $C_{AW,M}$ it is vice versa. For blunt-type ships, it is visible that the reflection component has decreased considerably, cf. Fig. 9.

In Section 2, it was assumed that the large variance in the slender ship ($C_B < 0.7$) database is attributed to high forward speeds. Additionally, the variance might result from more pronounced non-linear behavior, such as wave breaking in short waves, which is mainly due to the hull shape (and wave steepness): (1) The significant flare angles in the bow region, (2) the relatively long bulbous bow and (3) extreme B/T ratios (in case of passenger ships) lead to a considerable reflection contribution. Hizir et al. (2019) discuss the *geometric non-linearity* of slender hulls and Mourkogiannis and Liu (2021) analyze the bow reflection of ships with high B/T ratios both analytically and numerically. The accuracy for blunt and slender ships as well as the smoothness of the method is investigated in Fig. 12 considering the HSVA cruise ship and the SNU tanker. It is noted that both ships are in-sample testcases and that the proper assessment and validation of the formula's accuracy is not possible using a tilted three-dimensional figure.

It is thought that the presented method provides satisfactory and smooth estimates, but fine-tuning is needed especially in short waves as considerable deviations are evident in both Figs. 12(a) and 12(b). Moreover, conservative estimates for the full-type tanker and underprediction for the fast and slender cruise ship are perceptible. The latter problem has already been observed in Fig. 2 and similar deviations for slender ships were also observed by Tsujimoto et al. (2008) using their NMRI formula. Lastly, the reduction of the residuals' variance was achieved and thus it was possible to diminish the parameter uncertainty. However, even after the parameter calibration, the residuals revealed significant bias in Fig. 6 indicating strong structural uncertainty. Although it was attempted to lower this uncertainty component by splitting the method at $C_B = 0.70$ and deriving two separate parameter sets, the impact was negligible. The extension of the underlying semi-empirical framework of Liu and Papanikolaou (2020) is thus an important aspect for the reduction of structural uncertainty. As a sidenote, during extensive validation studies it appeared in a few cases that in following seas and relatively low forward speed ($Fn < 0.1$) the uncertainty bounds intersect. It is thought that this results from the small magnitude of the added resistance for low Fn and the conditional calculation of a_1. Ultimately, an exception is included in the developed code to circumvent this drawback and it also requires further studies in the future in order to overcome this issue.

The calculation of the added resistance in following and quartering waves utilizing the Liu and Papanikolaou formula is based on linear interpolation and dependent on encounter frequency. The formulation of Maruo (1957), on the other hand, is exact under the assumptions of an ideal fluid and shows that the sign of the added resistance in following waves can be obtained by the balance of the magnitude of three integrals dependent on the encounter frequency. Still, the drawback of this approach is the required knowledge of detailed hull form information as well as the increased computational effort for the calculation of the Kochin function. However, the focus of the presented formula is on the efficient and practical determination of the added resistance in waves with rudimentary hull form information, e.g. in early ship design stages. Therefore, a modified and simplified reasoning using Maruo's theory is chosen for the calculation of the sign of the added resistance in following waves (cf. Eq. (12) and Fig. 4). In addition, the development of a correction approach dependent on wave steepness would lead to improved results of the semi-empirical framework especially in short waves and for slender ships, however, there is often a lack of corresponding H/λ data in publicly available experimental data.

The previously mentioned assumption of the DTU design tool and the NMRI formula, i.e. considering the added resistance to be zero in stern oblique waves due to the surrounding uncertainty is addressed. From a practical perspective the assumption appears as legitimate in view of Figs. 8 and 10 as the degree of uncertainty engulfs the actual magnitude of the predicted added resistance. However, from a scientific standpoint this observation motivates further rigorous experimental
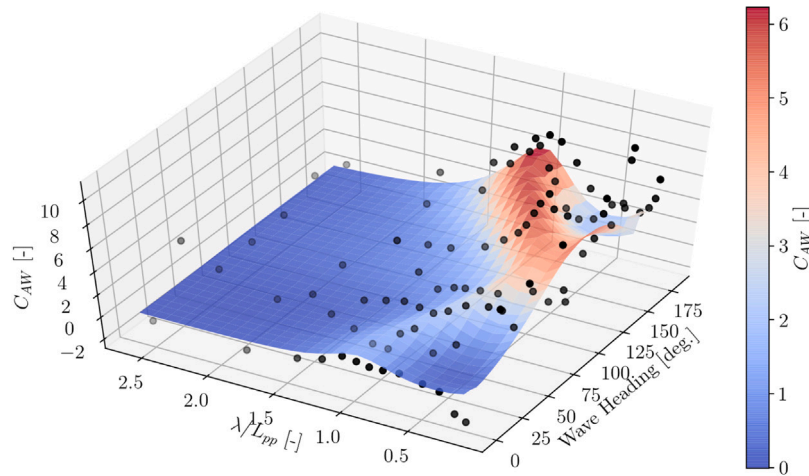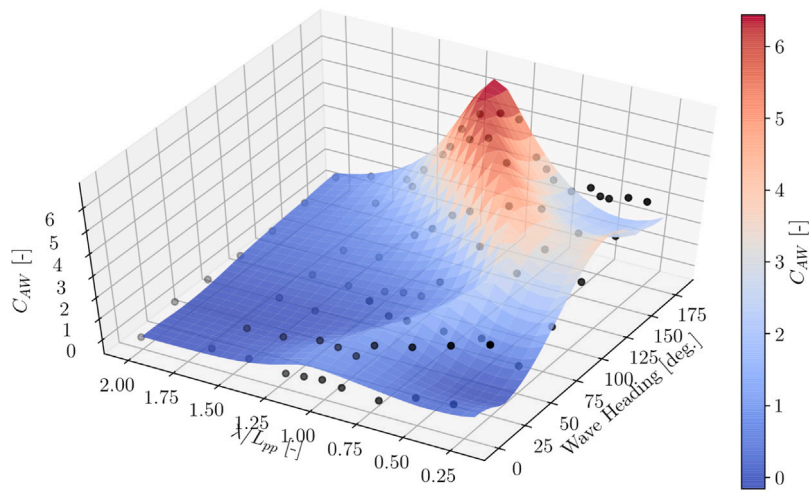
(a) Slender HSVA cruise ship ($Fn = 0.233$), cf. Valanto and Hong (2015).

(b) Full-type SNU tanker ($Fn = 0.137$), cf. Park et al. (2019).

Fig. 12. Illustration of the present method compared to model test data for two case studies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and numerical investigation of C_{AW} in wave headings of $\beta \in [0, 90]$ deg., in order to further reduce the epistemic uncertainty in this regime of wave headings. In addition, more experimental data for slender ships, e.g. container and naval ships, are required. It is well known in statistical modeling that the model itself can only be as good as the underlying dataset. Hence, the extension of the database with model test samples showing high data quality in oblique waves is essential for the enhancement of the Liu and Papanikolaou formula applying the present approach.

The prediction of uncertainty bounds of the added resistance enables the possible application of the proposed method for the determination of outliers and erroneous data, i.e. *outlier detection*, in either experimental or numerical data. The definition of a corresponding decision boundary is non-trivial in this case as the added resistance is inherently non-linear (and heteroskedastic). Hence, a constant (i.e. linear) criterion for anomaly detection for the entire C_{AW} regime derived, e.g. from either the variance or the interquartile range of the underlying dataset is considered as not appropriate. In this respect, the application of unsupervised machine learning methods for classification show potential for accurate non-linear outlier detection for C_{AW}. However, this kind of methods is in general considered as opaque and thus the proposed approach could show better applicability due to its transparency

resulting from the used simple algebraic expressions. Ultimately, an investigation in this direction is an interesting aspect for further work.

5. Summary and conclusion

The present study proposed a data-driven methodology for parameter calibration of a semi-empirical approach for estimation of the added resistance. The calibration was performed for the reduction of the inherent uncertainties of the added resistance in a deterministic manner. In this contribution, the mathematical formulation of Liu and Papanikolaou (2020) was adjusted and the parameter vector was optimized with respect to two datasets for full and slender ships consisting of several hundred of data points from publicly available model tests resembling the current world fleet sufficiently. In doing so, the method's transparency is maintained and the advantages of supervised regression models are leveraged. The minimization was defined as a single objective optimization problem and due to the complex and scattered loss landscape, the stochastic and metaheuristic particle swarm optimizer was used. It was found that out of four different loss functions, the squared loss function performed superior according to the chosen metrics. The calibration of the semi-empirical definition led to an increase of around 9% in performance and thereby the parameter uncertainty was diminished significantly. The 90% prediction interval was established by deriving additional parameter sets which were

determined by optimizing the tilted quantile version of the squared loss function. The uncertainty bounds proved to be highly valuable in a practical context for expressing the inherent irreducible uncertainty of the added resistance. Finally, the proposed method was validated against experimental data and several well-established formulas for the prediction of the added resistance. The validation studies suggest satisfactory accuracy of the mean estimate and the prediction interval on in- and out-of-sample data.

The implementation of the uncertainty bounds increases the relevance of the used formula for practical application. In addition, the described calibration methodology is considered as "method-agnostic" or generic for uncertainty quantification of (semi-)empirical formulas in general for e.g. the calculation of motion transfer functions using the closed form expressions of Jensen et al. (2004) as long as enough experimental data is available. The present study was limited to regular long-crested regular waves, however, the calculation of the mean added resistance in irregular waves is trivial given a suitable (directional) wave energy density spectrum and under the assumption of linear superposition and time invariance. However, the mentioned prerequisites are invalid under certain circumstances, e.g. in severe sea states or under changing weather conditions. For this reason, higher order spectral analysis may be required for an accurate estimate of the mean added resistance in a realistic seaway as shown by Hasselmann (1966). Furthermore, a corresponding uncertainty estimate of the prevalent wave spectrum is necessary for a consistent calculation. Lastly, the present paper advocates the implementation of a trustworthy representation of uncertainty when predicting safety-related quantities in ship hydrodynamics, such as the added resistance or speed loss in a realistic seaway.

Even though the method has been compared to multiple established approaches, additional and more rigorous validation studies are necessary for universal applicable recommendations and limitations for application. Moreover, sensitivity studies for the variation in draft and trim are important subjects of future work, in order to prove the applicability of the proposed method in ship operation for, e.g. performance monitoring or weather routing. Moreover, it seems appealing to enrich the presented database with high-fidelity CFD results – in short and oblique waves in particular – for further reduction of epistemic uncertainty as model test results from the public domain are scarce in these specific conditions. In Section 2, it was decided to split the approach for the block coefficient; however, dividing the method for different forward speed regimes and deriving updated parameter sets might have serious potential for further reducing structural uncertainty. The application of a parameter calibration combined with a bias correction introduced by an additive discrepancy expression in a probabilistic or Bayesian methodology could also have a tremendous effect on reducing and *explaining* the model form uncertainty, cf. Wu et al. (2021).

Reproducibility

In order to facilitate adoption and adaptation of the presented approach, the Python code as well as an Excel implementation is publicly available on https://gitlab.gbar.dtu.dk/mamit/RAW_Formula.

CRediT authorship contribution statement

Malte Mittendorf: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualization, Writing – original draft, Writing – review & editing. **Ulrik D. Nielsen:** Writing – review & editing, Supervision, Project administration, Resources, Funding acquisition. **Harry B. Bingham:** Writing – review & editing, Supervision, Resources, Project administration. **Shukui Liu:** Writing – review & editing, Software, Resources.

References

Amini-Afshar, M., Bingham, H.B., 2021. Added resistance using salvesen – tuck – faltinsen strip theory and the kochin function. Appl. Ocean Res. 106, 102481.

Blank, J., Deb, K., 2020. Pymoo: Multi-objective optimization in python. IEEE Access 8, 89497–89509.

Blok, J.J., 1993. The Resistance Increase of a Ship in Waves (Ph.D. thesis). Technical University of Delft.

Bolbot, V., Papanikolaou, A., 2016. Parametric, multi-objective optimisation of ships bow for the added resistance in waves. Ship Technol. Res. 63 (3), 171–180.

Breusch, T.S., Pagan, A.R., 1979. A simple test for heteroskedasticity and random coefficient variation. Econometrica 47 (5), 1287–1294.

Cepowski, T., 2020. The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network. J. Ocean Eng. 195.

Equasis, 2019. Statistics from Equasis - Electronic Quality Ship Information System. The 2019 World Fleet Report.

Faltinsen, O.M., Minsaas, K.J., Liapis, N., Skjordal, S.O., 1980. Prediction of resistance and propulsion of a ship in a seaway. In: Proc. of 13th Symposium on Naval Hydrodynamics. Tokyo (Japan): The Shipbuilding Research Association of Japan. pp. 505–529.

Fujii, H., Takahashi, T., 1975. Experimental study on the resistance increase of a large full ship in regular oblique waves. J. Soc. Nav. Archit. Japan 137, 132–137.

Gatchell, S., 2018. Stability and Hydrodynamic Tools – Level 1. HOLISHIP Internal Report D3.2.

Gerritsma, J., Beukelman, W., 1972. Analysis of the resistance increase in waves of a fast cargo-ship. Int. Shipbuild. Prog. 18 (217).

Guo, B., Steen, S., 2010. Added resistance of a VLCC in short waves. In: Proc. 29th Int. Conf. on Ocean, Offshore & Arctic Engineering. OMAE 2010.

Hasselmann, K., 1966. On non-linear ship motions in irregular waves. J. Ship Res. 10, 64–68.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer Science+Business Media, New York.

Havelock, T.H., 1940. The pressure of water waves upon a fixed obstacle. Proc. R. Soc. Lond. Ser. A 175 (963), 409–421.

Hizir, O., Kim, M., Turan, O., Day, A., Incecik, A., Lee, Y., 2019. Numerical studies on non-linearity of added resistance and ship motions of KVLCC2 in short and long waves. Int. J. Nav. Archit. Ocean Eng. 11 (1), 143–153.

Hollenbach, K.U., 1998. Estimating resistance and propulsion for single-screw and twin-screw ships. J. Ship Tech. Res. 45 (2), 72–76.

Holt, P., Nielsen, U.D., 2021. Preliminary assessment of increased main engine load as a consequence of added wave resistance in the light of minimum propulsion power. J. Appl. Ocean Res. 108, 102543.

Holtrop, J., Mennen, G., 1982. An approximate power prediction method. Int. Shipbuild. Prog. 29 (335), 166–170.

Huber, P.J., 1981. Robust Statistics. John Wiley and Sons, New York.

Hüllermeier, E., Waegeman, W., 2021. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. Mach. Learn. 110, 457–506.

IMO, 2013. Interim guidelines for determining minimum propulsion power to maintain the maneuverability in adverse conditions. London, international maritime organization. In: MEPC 232(65), London, UK.

IMO, 2016. Supplementary information on the draft revised guidelines for determining minimum propulsion power to maintain the maneuverability of ships in adverse conditions. In: MEPC 70/INF.30, London, UK.

IMO, 2017. Draft revised guidelines for determining minimum propulsion power to maintain the manoeuvrability of ships in adverse conditions. In: MEPC 71/INF.28. London, UK.

IMO, 2021. Report of the correspondence group on air pollution and energy efficiency, submitted by Japan. In: MEPC 76/5/1, London, UK.

ITTC, 2014. Procedures and guidelines: Prediction of power increase in irregular waves from model test. In: International Towing Tank Conference.

ITTC, 2017. Recommended procedures and guidelines - preparation, conduct and analysis of speed/power trials. In: International Towing Tank Conference.

ITTC, 2021. The specialist committee on ships in operation at sea (SOS): Final report and recommendations to the 29th ITTC. In: International Towing Tank Conference.

Jensen, J.J., Mansour, A.E., Olsen, A.S., 2004. Estimation of ship motions using closed-form expressions. Ocean Eng. 31 (1), 61–85.

Jinkine, V., Ferdinande, V., 1974. A method for predicting the added resistance of fast cargo ships in head waves. Int. Shipbuild. Prog. 21 (238), 149–167.

Kadomatsu, K., Inoue, Y., Takarada, N., 1989. On the added resistance and speed loss in low forward speed range. J. Soc. Nav. Archit. Japan 166, 217–223.

Kashiwagi, M., 2013. Hydrodynamic study on added resistance using unsteady wave analysis. J. Ship Res. 57 (4), 220–240.

Kashiwagi, M., Ikeda, T., Sasakawa, T., 2010. Effects of forward speed of a ship on added resistance in waves. Int. J. Offshore Polar Eng. 20 (3), 196–203.

Kashiwagi, M., Iwashita, H., Miura, S., Hinatsu, M., 2019. Study on added resistance with measured unsteady pressure distribution on ship-hull surface. In: Proc. of 34th IWWWFB (Noah's on the Beach, Newcastle, NSW, Australia). pp. 81–84.

Kashiwagi, M., Sugimoto, K., Ueda, T., Yamasaki, K., Arihama, K., Kimura, K., Yamashita, R., Itoh, A., Mizokami, S., 2004. An analysis system for propulsive performance in waves. J. Kansai Soc. Nav. Archit. 241, 67–82.

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proc. of IEEE Int. Conf. on Neural Networks, Vol. IV. pp. 1942–1948.

Kim, T., Yoo, S., Kim, H.J., 2021. Estimation of added resistance of an LNG carrier in oblique waves. J. Ocean Eng. 231, 109068.

Koenker, R., Hallock, K.F., 2001. Quantile regression. J. Econ. Perspect. 15 (4), 143–156.

Kuroda, M., Tsujimoto, M., Fujiwara, T., 2008. Investigation on components of added resistance in short waves. J. Soc. Nav. Archit. Ocean Eng. 8.

Kwon, Y.J., 1981. The Effect of Weather, Particularly Short Sea Waves, on Ship Speed Performance (Ph.D. thesis). Department of Naval Architecture and Shipbuilding, University of Newcastle upon Tyne.

Lang, X., Mao, W., 2021. A practical speed loss prediction model at arbitrary wave heading for ship voyage optimization. J. Mar. Sci. Appl. 20, 410–425.

Li, C., Ma, X., Chen, W., Li, J., Dong, G., 2016. Experimental investigation of self propulsion factor for a ship in regular waves. J. Shipbuild. China 57 (1), 1–8.

Liu, S., 2020. Revisiting the influence of a ship's draft on the drift force due to diffraction effect. J. Ship Tech. Res. 67 (3), 175–180.

Liu, S., Papanikolaou, A., 2013. Added resistance of ships in quartering seas. In: Proc. V Int. Conf. on Computational Methods for Coupled Problems in Science and Engineering Coupled Problems 2013, Spain.

Liu, S., Papanikolaou, A., 2016. Fast approach to the estimation of the added resistance of ships in head waves. J. Ocean Eng. 112, 211–225.

Liu, S., Papanikolaou, A., 2020. Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. J. Ocean Eng. 206.

Liu, S., Papanikolaou, A., Zaraphonitis, G., 2015. Practical approach to the added resistance of a ship in short waves. In: Proc. of the 25th Int. Offshore and Polar Engineering Conference, Kona, Hawaii, ISOPE.

Liu, S., Shang, B., Papanikolaou, A., 2019. On the resistance and speed loss of full type ships in a seaway. J. Ship Tech. Res. 66 (3), 161–179.

Martinsen, M.A., 2016. A Design Tool for Estimating Wave Added Resistance of Container Ships (Master's Thesis). Technical University of Denmark.

Maruo, H., 1957. The excess resistance of a ship in rough seas. Int. Shipbuild. Prog. 4 (35).

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (2), 239–245.

Mikkelsen, H., Shao, Y., Walther, J.H., 2022. Numerical study of nominal wake fields of a container ship in oblique regular waves. Appl. Ocean Res. 119, 102968.

Minsaas, K., Faltinsen, O.M., Persson, B., 1983. On the importance of added resistance, propeller immersion and propeller ventilation for large ships in a seaway. In: Proc. of 2nd Int. Symp. on Practical Design in Shipbuilding (PRADS'83), Tokyo and Seoul.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., 2022. Data-driven prediction of added-wave resistance on ships in oblique waves - a comparison between tree-based ensemble methods and artificial neural networks. J. Appl. Ocean Res. 118 (102964).

Mourkogiannis, D., Liu, S., 2021. Investigation of the influence of the main dimensional ratios of a ship on the added resistance and drift force in short waves. In: Proc. of 31st Int. Ocean and Polar Engineering Conference, Rhodes, Greece (Virtual).

Nakamura, S., Naito, S., 1977. Propulsive performance of a containership in waves. J. Soc. Nav. Archit. Japan 15, 24–48.

Nielsen, C.S., 2015. A Ship Design Tool for Estimating Added Resistance in Waves (Master's Thesis). Technical University of Denmark.

Nikolopoulos, L., Boulougouris, E., 2019. A study on the statistical calibration of the holtrop and mennen approximate power prediction method for full hull form, low froude number vessels. J. Ship. Prod. Des. 35, 41–68.

Oh, S., Yang, J., Park, S.H., 2015. Computational and experimental studies on added resistance of aframax-class tankers in head seas. J. Soc. Nav. Archit. Korea 52 (6), 471–477.

Papanikolaou, A., 2014. Ship Design – Methodologies of Preliminary Design. Springer.

Papanikolaou, A.D., 2018. Hydrodynamics & ship design: Optimization of ships for minimum power and safe navigation in adverse weather conditions. In: 41st Georg Weinblum Memorial Lecture, Hamburg.

Papanikolaou, A., Liu, S., 2021. On the uncertainties in the estimation of the added resistance of ships in waves in view of practical applications. In: Proc. of 5th Joint ISSC and ITTC International Workshop on Uncertainty Modelling in Wave Description and Wave Induced Responses.

Park, D.M., Lee, J., Jung, Y., Lee, J., Kim, Y., Gerhardt, F., 2019. Experimental and numerical studies on added resistance of ship in oblique sea conditions. J. Ocean Eng. 186, 106070.

Park, D.M., Lee, J., Kim, Y., 2015. Uncertainty analysis for added resistance experiment of KVLCC2 ship. J. Ocean Eng. 95, 143–156.

Sadat-Hosseini, H., Wu, P., Carrica, P., Kim, H., Toda, Y., Stern, F., 2013. CFD verification and validation of added resistance and motions of KVLCC2 with fixed and free surge in short and long head waves. J. Ocean Eng. 59, 240–273.

Shigunov, V., el Moctar, O., Papanikolaou, A., Potthoff, R., Liu, S., 2018. International benchmark study on numerical simulation methods for prediction of maneuverability of ships in waves. J. Ocean Eng. 165, 365–385.

Sigmund, S., 2018. Performance of Ships in Waves (Ph.D. thesis). University of Duisburg-Essen.

Simonsen, C.D., Otzen, J.F., Nielsen, C., Stern, F., 2014. CFD prediction of added resistance of the KCS in regular head and oblique waves. In: Proc. of 30th Symp. on Naval Hydrodynamics, Hobart, Tasmania, Australia.

Söding, H., Shigunov, V., Schellin, T., el Moctar, O., 2014. A rankine panel method for added resistance of ships in waves. J. Offshore Mech. Arct. Eng. 136.

Soize, C., 2017. Uncertainty Quantification: An Accelerated Course with Advanced Applications in Computational Engineering. In: Interdisciplinary Applied Mathematics, vol. 47, Springer.

Sprenger, F., Maron, A., Delefortrie, G., Hochbaum, A., Fathi, D., 2015. Mid-term review of tank test results. In: SHOPERA Deliverable D3 2.

Sprenger, F., Maron, A., Delefortrie, G., Van Zwijnsvoorde, T., Cura-Hochbaum, A., Lengwinat, A., Papanikolaou, A., 2017. Experimental studies on seakeeping and manoeuvrability in adverse conditions.. J. Ship Res. 61 (3), 131–152.

SSPA, 2012. Ulysses project phase II: Seakeeping test for 2020 ship in sspa's maritime dynamic laboratory. In: ULYSSES Seakeeping and Manoeuvring Model Tests.

Stocker, M.R., 2016. Surge Free Added Resistance Tests in Oblique Wave Headings for the KRISO Container Ship Model (M.Sc. Thesis). University of Iowa.

Strøm-Tejsen, J., Yeh, H.Y.H., Moran, D.D., 1973. Added resistance in waves. Soc. Nav. Archit. Mar. Eng. Trans. 81, 109–143.

Takahashi, T., 1988. A practical prediction method of added resistance of a ship in waves and the direction of its application to hull form design. Trans. West Japan Soc. Nav. Archit. 75, 75–95.

Taskar, B., Andersen, P., 2021. Comparison of added resistance methods using digital twin and full-scale data. J. Ocean Eng. 229, 108710.

Tsujimoto, M., Shibata, K., Kuroda, M., Takagi, K., 2008. A practical correction method for added resistance in waves. J. Japan Soc. Nav. Archit. Ocean Eng. 8, 177–184.

Tsujimoto, M., Yasukawa, H., Yamamoto, K., Il Lee, T., 2021. Validation of added resistance in waves by tank tests and sea trial data. J. Ship Tech. Res..

Uharek, S., 2019. Numerical Prediction of Ship Manoeuvering Performance in Waves (Doctoral Thesis). Technical University of Berlin.

Ursell, F., Dean, W.R., 1947. The effect of a fixed vertical barrier on surface waves in deep water. Math. Proc. Cambridge Philos. Soc. 43 (3), 374–382.

Valanto, P., Hong, Y., 2015. Experimental investigation on ship wave added resistance in regular head, oblique, beam, and following waves. In: Proc. of the 25th Int. Society of Offshore and Polar Engineers.

Wang, J., Bielicki, S., Kluwe, F., Xin, G., Orihara, H., Kume, K., Oh, S., Liu, S., Feng, P., 2021. Validation study on a new semi-empirical method for the prediction of added resistance in waves of arbitrary heading in analyzing ship speed trial results. J. Ocean Eng. 240, 109959.

Wang, Q., Ma, Y., Zhao, K., Tian, Y., 2020. A comprehensive survey of loss functions in machine learning. Ann. Data Sci..

Wu, X., Xie, Z., Alsafadi, F., Kozlowski, T., 2021. A comprehensive survey of inverse uncertainty quantification of physical model parameters in nuclear system thermal-hydraulics codes. arXiv:2104.12919v1.

Yang, K., Kim, Y., Jung, Y., 2018. Enhancement of asymptotic formula for added resistance of ships in short waves. J. Ocean Eng. 148, 211–222.

Yasukawa, H., Hirata, N., Matsumoto, A., Kuroiwa, R., Mizokami, S., 2019. Evaluations of wave-induced steady forces and turning motion of a full hull ship in waves. J. Mar. Sci. Technol. 24 (1), 1–15.

Yasukawa, H., Tsujimoto, M., 2020. Impact of bow shape on added resistance of a full hull ship in head waves. J. Ship Tech. Res. 67 (3), 136–147.

Yu, J.W., Lee, C.M., Lee, I., Choi, J.E., 2017. Bow hull-form optimization in waves of a 66,000 DWT bulk carrier. Int. J. Naval. Archit. Ocean. Eng. 9 (5), 499–508.

Zhan, Z., Zhang, J., Li, Y., Chung, H.S., 2009. Adaptive particle swarm optimization. IEEE Trans. Syst. Man Cybern. B 39 (6), 1362–1381.

Chapter 4

Ship-based Prediction of Sea State Information

Mathematically, estimating the mean added wave resistance in irregular waves is straightforward using Eq. 2.6, given a suitable quadratic transfer function and a wave energy density spectrum. In practice, however, sea state information is sparse and can, thus, be considered an Achilles heel for the determination of $\bar{R}_{AW}$ in actual conditions, according to Bertram [11]. Therefore, the present chapter considers the ship as a sailing wave buoy, i.e. the wave-induced ship response is used for estimating the governing sea state conditions in an inverse problem. The first studies on the *wave buoy analogy* started in the 1970s and Nielsen [95] provides an extensive overview of model-based methods for sea state estimation, which utilize the ship motion transfer functions. Research pertaining to on-site sea state estimation evolved in several iterations, where considering ship forward speed and eventually beam-to-following waves were considered as separate crucial milestones, Iseki and Ohtsu [56]. Recently, the application of deep neural networks showed promising results using the wave buoy analogy concept, e.g. Düz *et al.* [28].

4.1   Introduction

In his 1950 seminal paper, Alan Turing asked, "Can machines think?", Turing [125]. Additionally, he stated "[...] that these machines are intended to carry out any operations which could be done by a human computer". Indeed, this is a definition of artificial *general* intelligence, i.e. a human-like performance, which is considered the ultimate goal in the research field of machine learning. Exactly 70 years later, in 2020, Brown *et al.* [18] proposed an autoregressive language model called GPT3 (Generative Pre-trained Transformer 3), which has an enormous number of 175 billion parameters, and the later ChatGPT application built upon this particular model. Brown *et al.* [18] state that human evaluators struggled to distinguish between human-written and news articles generated by GPT3, indicating that GPT3 may exhibit human-like behavior in specific disciplines. However, several insufficiencies and erroneous behavior of GPT3 were reported as well.

Deep learning is the underlying methodology on which GPT3 is founded, and it is considered a sub-field of machine learning. Formally, deep learning is concerned with neural

networks with more than one hidden layer and follows *connectionism*, which attempts to emulate cognitive processes by an interconnected network of nodes (or simplified neurons). Thereby it is understood that these connections can be adjusted by weights. In its most generic form, deep neural networks can be considered composite functions (or acyclic graphs) of several different units, which can be seen in Eq. 4.1.

$$\hat{y} = \phi_l[\boldsymbol{\theta}_l, \phi_{l-1}(\boldsymbol{\theta}_{l-1}, \ldots, \phi_1\{\boldsymbol{\theta}_1, \boldsymbol{x}\})] \tag{4.1}$$

As can be inferred from Eq. 4.1, a neural network produces (in this case) a feedforward mapping, i.e. the information flows in one direction from input to output through hidden layers. Moreover, each affine layer with index l comprises parameters $\boldsymbol{\Theta}$ and a non-linear activation function ϕ. As a side note, the bias terms are disregarded in Eq. 4.1. Theoretically, according to the universal approximation theorem, any neural network with a linear output layer and at least one hidden layer with a sufficient number of neurons can represent any possible function, Goodfellow *et al.* [39]. Even though this network may *represent* any function, it is uncertain whether the training algorithm is capable of learning this function either due to overfitting or caused by the optimization algorithm, which may be unable to find a parameter combination fitting the function correctly. The optimal set of parameters is obtained by gradient-based optimization, and the gradients are obtained through backpropagation, which feeds errors from output to input nodes, as proposed by Rumelhart *et al.* [105]. Quintessential, the multilayer perceptron in $[J_1]$ can already be considered a deep neural network, but it treats its inputs as spatially invariant. In contrast, convolutional and recurrent neural networks are able to process higher-dimensional spatially coherent or even sequential data. Deep neural networks generally exhibit a higher scalability compared to traditional machine learning techniques in terms of input dimensions and sample size. However, due to the larger number of parameters, both computational effort and the risk of overfitting (on small datasets) are typically higher in comparison to other machine learning methods. It is important to stress that the field of deep learning is still an active research discipline and that certain aspects still lack a thorough theoretical foundation. A concise overview of the deep learning development is given by LeCun *et al.* [76], whereas a theoretical background can be found in Goodfellow *et al.* [39].

One of the most significant advantages of deep learning techniques is the implicit feature construction, i.e. the ability to generate meaningful data without any manual preprocessing, which is also known as end-to-end learning. In contrast, traditional machine learning approaches require time-consuming feature extraction before modeling, as shown by Tu *et al.* [124], where they utilize time/frequency domain features for classifying the sea state according to the Beaufort scale. Following the work of Tu *et al.* [124], the overall methodology is named sea state *identification*, to differentiate from sea state *estimation*, where the entire directional wave spectrum is predicted. Herein, the models are trained for providing the mean integral sea state parameters. It is noted that these may have a reduced physical relevance in multimodal sea states. The overarching methodology of this chapter and $[J_3]$ is depicted in Fig. 4.1. Following linear and time-invariant filter theory, the ship acts as a filter on the unknown instantaneous wave elevation, and the motion/structural ship response can be measured and acts as input for the neural network. The green arrows indicate a physical process, whereas the lower opposing red arrows show the data flow in the inverse procedure
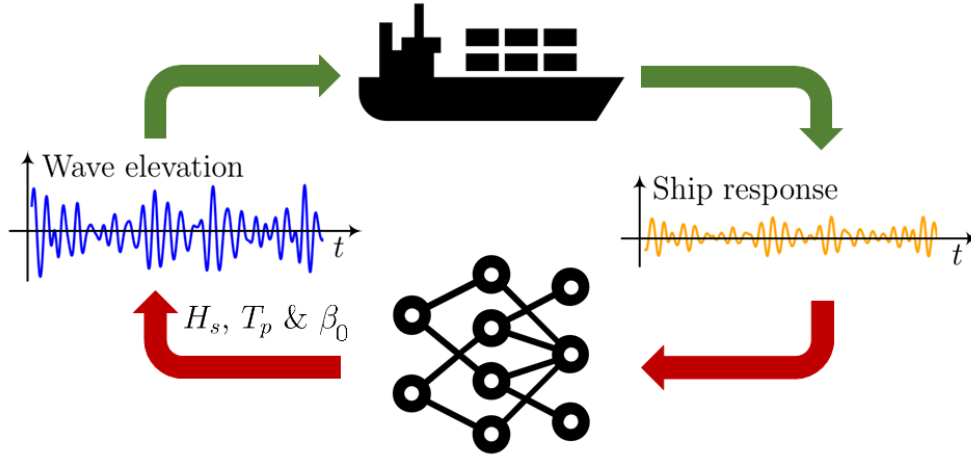
performed by a neural network.

Figure 4.1: Simplified concept of the wave buoy analogy, where green arrows indicate the filtering effect of a ship and red arrows resemble the inverse mapping problem. It is noted that the wave elevation is not reconstructed, but the underlying sea state parameters.

Given Fig. 4.1, it can be seen that the model uses the ship response as an input and maps it back to the governing sea state parameters $\{H_s, T_p, \beta_0\}$. Obtaining these quantities beforehand by, e.g. a wave radar, is seen as most critical for this supervised learning study. Another key aspect of $[J_3]$ is to compare time and frequency domain approaches in terms of accuracy and computational efficiency. Initially, a study on simulated motion data for a container vessel in unimodal and unidirectional seas was conducted and presented in $[C_1]$. This study provided evidence that frequency domain models may show a substantially lower computational cost, allow considering longer time series samples without increasing CPU time, and are likely of a more robust nature. In agreement with other publications, it was confirmed that predicting the mean encounter wave direction is delicate and requires information about the phase relationships of the individual responses. Estimating both energy content and its distribution within the wave energy density spectrum, i.e. H_s and T_p, appears relatively straightforward – even under changing forward speed. The enclosed $[J_3]$ builds upon these findings and presents a rigorous comparison of time and frequency domain models based on actual in-situ data of a container ship operating in the Northern Atlantic. Contrary to most state-of-the-art publications, it has been found in $[J_3]$ that considering frequency domain features, i.e. cross-spectral analysis, for data preprocessing shows superior characteristics. The application of multi-task learning, i.e. the simultaneous consideration of several learning tasks, in the sea state identification domain is considered the main novelty of this article. The work has been disseminated at the second Marine AI WISE seminar organized by TUMSAT (Tokyo University of Marine Science and Technology) in Mittendorf and Nielsen [87]. In addition, an extending study aimed at different preprocessing methodologies of the mean encounter wave direction is presented in $[C_3]$. Lastly, DTU has been part of the VALID III JIP (Joint Industry Project), and the work of $[J_3]$ has been applied to two coastguard cutters. The dataset consisted of motion recordings and

wave hindcast data, and the frequency domain approach showed satisfactory robustness and accuracy – even in coastal waters with multimodal sea states. The corresponding results are presented in the final deliverable of the JIP, Drummen and Hageman [27]. During the individual studies, the versatility and modularity of deep learning stood out very positively, but the increased computational effort required the use of resources provided by DTU's computing cluster.

4.2   Discussion

Knowledge about the ambient wave environment is indispensable for decision support for seafarers to maintain ship safety and enhance energy efficiency. Generally, the determination of sea state parameters and especially the estimation of the directional wave spectrum are of great complexity. Moreover, *all* possible data sources and their estimates are characterized by inherent uncertainty. For instance, the used X-band wave radar provides an estimate of the directional wave energy density spectrum at the exact position of the vessel in time and space but needs empirical corrections for calibrating the energy content, i.e. the significant wave height, and may show reduced accuracy due to rain clutter. In fact, Ardhuin *et al.* [7] provide a review of different data sources of sea state data with an emphasis on the associated uncertainties. They postulate that narrower uncertainty tolerances compared to worldwide standards are needed for satisfactory application in, e.g. coastal areas. Regarding the three sea state parameters, the presently required accuracy regarding H_s is in the range of 5–10%, 0.1 to 1.0 s for T_p, and about 10 deg. for β_0. With that in mind, the sea state data from the wave radar is compared to reanalysis data from the public domain ERA5 database, as proposed by Hersbach *et al.* [46]. This *hindcast* service builds upon data assimilation by using the third-generation numerical ECWAM wave model and calibration by altimeter or wave buoy data. The sea state parameters along the ship's route are interpolated using a bi-linear scheme, as shown by Nielsen [96]. The direct comparison of wave radar and hindcast data is shown in Fig. 4.2 using separate correlation plots.
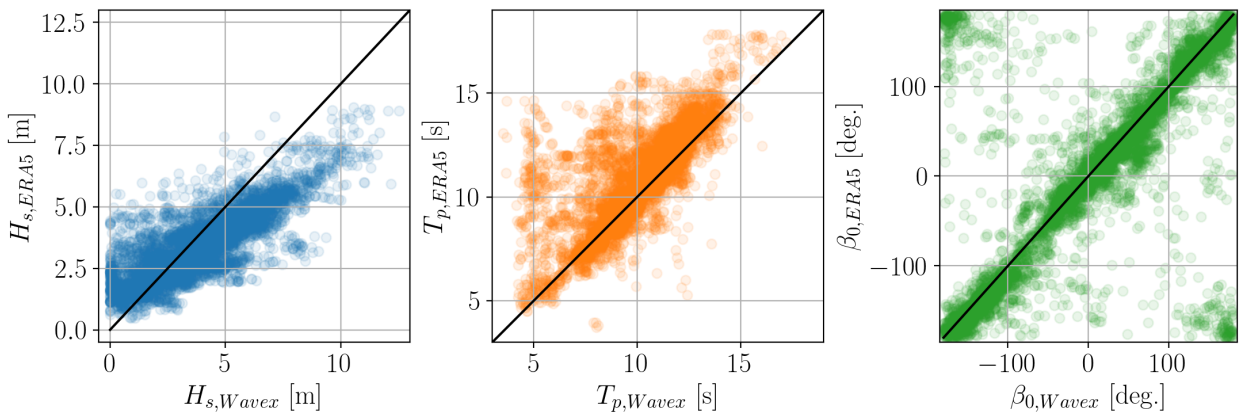

Figure 4.2: Comparison of integral sea state parameters H_s, T_p and β_0 according to ERA5 and the X-band wave radar mounted on the case ship.

In the case of peak period and mean encounter wave heading, a low bias but significant variance can be seen in Fig. 4.2. For H_s, however, a pronounced systematic bias stands out when comparing hindcast and wave radar data. When considering the ERA5 data as a ground truth, it is stated that the wave radar data is non-conservative in higher sea states and vice versa in lower sea states. As mentioned before, wave radars require empirical procedures for providing H_s correctly, and it was stated in Storhaug *et al.* [116] that the installed wave radar was not calibrated for the specific case ship. In fact, this could be a cause for the matching empirical and long-term joint distribution of H_s and T_p in $[J_3]$ since the use of weather routing would suggest lower values of encountered H_s. Gangeskar [35] shows that the technology of X-band wave radars improved significantly in comparison to the used wave radar during the measurement campaign, which was conducted 15 years ago. But still, the observed differences in H_s and, even more so, the large variance regarding all sea state parameters in Fig. 4.2 motivate a data fusion approach, i.e. combining several methods for minimizing uncertainty. For instance, Thornhill and Stredulinsky [122] show a robust and efficient hybrid methodology for calibrating the estimated significant wave height by an X-band wave radar using the experienced ship motions from a sailing vessel. More-over, an ensemble approach using several hindcast vendors shows potential for obtaining an uncertainty estimate, following Vettor and Guedes Soares [129]. In fact, when calculating $\bar{R}_{AW}$ with an uncertainty-aware transfer function estimate via the method proposed in $[J_2]$, a consistent calculation must consider uncertainty ranges for the prevailing sea state parameters as well. It will be an interesting aspect of future work to study deep learning-based methods for uncertainty quantification parallel to sea state identification. However, these uncertainty bands would reflect a combination of uncertainty related to data and actual modeling. Another essential element for further work is predicting the full directional wave spectrum for more accurate information about sea states comprising wind and swell, which is a more demanding task when using deep learning. Overall, it is concluded that deriving reliable sea state data remains challenging. Still, it is believed that the symbiosis between deep learning and established physical models will lead to an enhanced accuracy and robustness of corresponding estimates.

[J_3] Sea state identification using machine learning - A comparative study based on in-service data from a container vessel

The paper entitled "Sea state identification using machine learning - A comparative study based on in-service data from a container vessel" is published in the Journal of Marine Structures as:

MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B.; STORHAUG, G. (2022). *Sea State Identification using Machine Learning - A Comparative Study based on In-service Data from a Container Vessel.* J. Marine Structures Vol. 85, No. 103274 `https://doi.org/10.1016/j.marstruc.2022.103274`
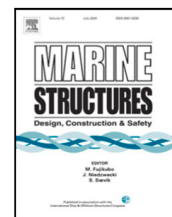
Sea state identification using machine learning—A comparative study based on in-service data from a container vessel

Malte Mittendorf [a,*], Ulrik Dam Nielsen [a,c], Harry B. Bingham [a], Gaute Storhaug [b]

[a] *DTU Civil and Mechanical Engineering, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark*
[b] *Maritime Advisory, DNV AS, NO-1363, Høvik, Norway*
[c] *Centre for Autonomous Marine Operations and Systems, NTNU AMOS, NO-7052, Trondheim, Norway*

ARTICLE INFO

ABSTRACT

This paper is concerned with a machine learning-based approach for sea state estimation using the wave buoy analogy. In-situ sensor data of an advancing medium-size container vessel has been utilized for the prediction of integral sea state parameters. The main novelty of this contribution is the rigorous comparison of time and frequency domain models in terms of accuracy, robustness and computational cost. The frequency domain model is trained on sequences of spectral ordinates derived from cross response spectra, while the time domain model is applied to 5-minute time series of ship responses. Multiple deep neural networks were trained and the sensitivity of individual sensor recordings, sample length, and frequency discretization on estimation accuracy was analysed. An Inception Architecture adapted for sequential data yields the highest out of sample performance in both considered domains. Additionally, multi-task learning was employed, as it is known for increased generalization capability and diminished uncertainty. Overall, it was found that the frequency domain method provides both superior performance and significantly less computational effort for training.

1. Introduction

1.1. Motivation

Ship safety and operational efficiency depend to a large degree on the prevalent sea state. As such, on-board identification of the ambient wave system may assist the crew in the decision making processes for minimizing risks related to critical wave encounters. For instance, large roll amplitudes are presently an immense concern due to the increasing number of lost containers at sea over the last couple of years resulting from wave-related impact, Meister et al. [1]. Furthermore, estimates of the sea state experienced throughout the vessel's operational profile are of significant importance for shore-based vessel performance analysis for the assessment of general energy efficiency and scheduling maintenance. Therefore, an accurate and reliable estimate of the prevailing wave energy density spectrum is the key aspect of on-board decision support systems. In addition, with an ever-increasing interest in autonomous ships, the significance of real time estimates of the wave environment and the corresponding ship response grows further as the expertise of seafarers may not necessarily be available, Jalonen et al. [2]. For all of this, it is thought that ship response-based sea state estimation (SSE) could be an essential building block. Hereby, it is understood that data in terms of wave-induced responses from the ship is processed, thus facilitating a real-time identification of the sea state at the ship's exact position. In fact, this is the underlying idea of the wave buoy analogy, as presented by Nielsen [3]. Broadly speaking, the wave buoy

---

* Corresponding author.
E-mail address: mamit@mek.dtu.dk (M. Mittendorf).

analogy considers the ship as a wave rider buoy and establishes an inverse mathematical relationship between measured responses and the encountered directional wave spectrum or the corresponding integral parameters.

1.2. Literature review

The methods for ship-based sea state estimation are manifold and an overview is given in Nielsen [4]. Initial studies addressing the wave buoy analogy were carried out in the 1970s by e.g. Takekuma and Takahashi [5], but without considering forward speed, i.e. the Doppler shift. Iseki and Ohtsu [6] and Nielsen [7] present methodologies based on Bayes theorem for the calculation of directional wave spectra using both complex-valued transfer functions and cross response spectra under forward speed conditions. Following the convention of Nielsen [3], the techniques for ship-board SSE are split into non-parametric and parametric approaches: The former provides the directional or 2D wave spectrum, while parametric techniques by e.g. [8–10] yield input values to parameterized wave energy density spectra with the possible addition of spreading functions. The applicability and accuracy of the aforementioned estimation methods inherently depend on the availability of reliable transfer functions, also referred to by Response Amplitude Operators (RAO). Hence, Nielsen et al. [11] propose a correction or rather *calibration* methodology for the pitch RAO using met-ocean hindcast data and in-service ship motion recordings of a container vessel. Mounet et al. [12] extend this work and merge the correction technique into a sea state estimation approach using a network of ships as observation platforms. Nevertheless, in individual cases, transfer functions may simply be unknown to the ship operator due to a lack of detailed hull shape information. Moreover, the presented procedures are premised on the assumptions of linearity for which transfer functions can be applied. These potential disadvantages of techniques relying on RAOs subsequently motivate data-driven machine learning studies for the estimation of sea state conditions.

Nowadays, machine learning techniques are universally applied as so-called surrogate models in general ship hydrodynamics, i.e. the model approximates data of computationally expensive methods in a regression task, e.g. Mittendorf and Papanikolaou [13]. In the field of ship-based SSE, machine learning is increasingly applied as well: Åvist and Pyörre [14], for instance, apply traditional machine learning – non-parametric regression techniques in particular – for the prediction of the significant wave height and encountered wave direction based on frequency domain features. Furthermore, Han et al. [15] predict the sea state parametrically using a research vessel as a case study and compare three different machine learning methods. In addition, the data preprocessing part is another key aspect and the feature space comprises elements from multiple domains. More importantly, deep learning methods, i.e. artificial neural networks with more than one hidden layer, achieved significant results in a variety of tasks ranging from image to speech recognition. The increased accuracy, versatility and scalability compared to traditional machine learning methods is credited to special layer-types, such as convolutional or recurrent layers. For the theoretical intricacies of deep learning consult Goodfellow et al. [16]. Its rapid development over the last decade subsequently sparked several studies in the ship-based SSE domain. Cheng et al. [17] develop an end-to-end classification approach via a multi-channel convolutional network for sea state estimation using the Beaufort sea scale for dynamic positioning, i.e. without forward speed. Instead of using raw time series data, Cheng et al. [18] convert the ship motions to spectrograms and train a neural network in an image recognition task predicting the sea state scale. The work is extended for forward speed cases in Cheng et al. [19] and an advanced architecture is employed. Their so-called SSENET features attention mechanisms and residual skip connections for enhanced performance. Moreover, Düz et al. [20] present a real-time multivariate time series *regression* approach for integral sea state parameters applying multiple deep architectures on 2.5-min motion samples of a frigate-type ship. One distinct aspect of their work is the procedure of transfer learning: Initially, the model is trained on simulated data obtained from time domain potential theory calculations and then re-trained on in-situ measurement data. The displayed results show sufficient accuracy for the prediction of the sea state parameters under forward speed conditions. Kawai et al. [21] present a simulation-based study of a container carrier in the frequency domain by extracting sequences of spectral values from a set of cross response spectra. The convolutional neural network predicts the parameters for an Ochi–Hubble type spectrum, i.e. for both wind and swell waves. Scholcz and Mak [22] extend the work of Düz et al. [20] and present a deep learning methodology for the non-parametric estimation of the directional wave spectrum based on wave radar data using a convolutional encoder–decoder network applied to in-service time series data. Lastly, Han et al. [23] provide an investigation for non-parametric SSE and establish an approach based on a generative adversarial network, in which the generator predicts the 2D spectrum relying on cross response spectra, and the discriminator classifies the validity of the prediction. This iterative approach is then compared to a traditional model-based approach and shows satisfactory accuracy in non-forward speed scenarios; emphasizing that simulated time series data has been considered exclusively.

1.3. Objective

This paper focuses on deep learning methods exclusively, whereas model-based approaches, dependent on availability of RAOs, are not considered. In view of state of the art literature, it is concluded that the majority of deep learning studies use time domain data. Thus, the main novelty of this paper is the parallel application of deep neural networks in both the time and frequency domains. In addition, a large part of the literature's deep learning studies are based on simulated ship motion data. In the present paper, however, in-service sensor recordings from a container vessel sailing in the Northern Atlantic are studied. A regression approach is proposed using sequential data from the time and frequency domain. The inverse mapping from several ship responses to prevalent sea conditions, i.e. significant wave height, peak period and mean encounter wave direction, is achieved by advanced deep neural networks, such as residual networks. Additionally, a novel approach for obtaining the optimal combination of ship responses for sea state identification is demonstrated in a feature importance study. Moreover, the trade-off between sample length

Fig. 1. Case study in service taken from [24].

Table 1
Main particulars of the Panamax container vessel, cf. [24].

Dimension	Unit	Magnitude
L_{pp}	[m]	232.0
B	[m]	32.2
T	[m]	10.78
C_B	[–]	0.685
dwt	[t]	40,900
P_B (@ 93 rpm)	[kW]	25,786

and frequency discretization will be determined in the frequency domain in another sensitivity study. Initially, four multi-output-regressor architectures are compared in terms of prediction accuracy, and two different multi-task configurations will be applied to the best-performing model in the first iteration. The application of multi-task learning for sea state estimation is another novelty of this investigation, as the recent literature is only focused on multi-output estimators. Lastly, a contribution of this work is the development of data-driven methods for the mapping of in-situ ship motion recordings to sea state parameters derived from data of a directional wave radar.

1.4. Composition

The remaining sections of the article are organized in the following way: In the upcoming Section 2, the case study, its sensor infrastructure and the data filtering methodology will be presented. Section 3 has its focus on the proposed methodology and conveys the applied neural network architectures as well as the concept of multi-task learning. Furthermore, the obtained results are shown and discussed in Section 4. In the final Section 5, the described work is summarized and suggestions for future work are presented.

2. Data analysis

The employed case ship of the present study is a 2800 TEU Panamax container vessel built in 1998. The vessel is displayed in Fig. 1 and its main particulars are listed in Table 1. During a four-year period between August 2007 and July 2011, the vessel was equipped with an extensive sensor framework and sailed in cross-Atlantic service. Even though the actual data acquisition period was conducted over 4 years, the Miros wave radar Wavex was only installed until March 2009. Hence, the span of the data used herein is 1.5 years and the GPS position of the vessel during that time is depicted in Fig. 2. Additional details of the measurement campaign including a study pertaining to structural fatigue due to wave excitations, are described in Storhaug et al. [24].

It is noteworthy that draft and trim were not measured. This sort of epistemic (i.e. systematic) uncertainty within the data may be influential on the following machine learning approach. In the paper of Storhaug et al. [24], it is stated that the typical transit draft is approximately 9.5 m and the transit trim lies between 0.5 to 1.0 m. In addition, the loading condition, and thus the transverse metacentric height GM_T, is not included in the dataset adding further uncertainty due to the widespread operational profile of container vessels, in general.

In Fig. 2, the GPS position history of the case ship is visible and the effect of seamanship as well as weather routing stands out since several routes deviate notably from the shortest distance. In one particular case, the container vessel even sails around the British islands in order to circumnavigate possible adverse weather conditions. The case ship trades between Western Europe (France, Belgium, Great Britain and Germany) and Quebec in Canada. Moreover, the ship obviously operates in coastal and restricted waters, such as the St. Lawrence river; however, the focus of the present study is on deep water conditions. For this reason, samples possibly influenced by shallow water or refraction from the coastline are disregarded by enforcing spatial boundaries at −55 and −5 degrees of Longitude as indicated in Fig. 2.
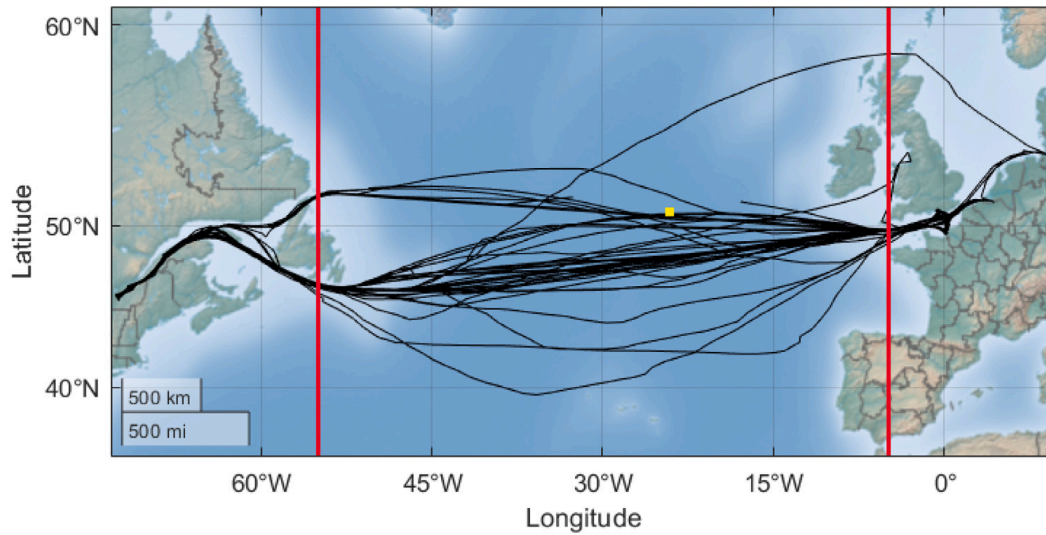
Fig. 2. Overview of the vessel's GPS position history (black) and the applied spatial boundaries (red). The yellow dot indicates one particular measurement sample.

The on-board measurement system comprises: (1) A Miros Wavex directional wave radar providing an averaged spectrum in a 30 min interval based on 1-min sample periods. The directional wave spectrum $E(f, \mu)$ is discretized into 31 frequencies f [Hz] and 36 directions μ [rad]. X-band marine radars are nowadays frequently applied as ocean remote sensors onboard of vessels due to their high spatio-temporal resolution. However, the accurate estimation of the significant wave height is a matter of empirical corrections (cf. Gangeskar [25]) and also the effect of rain clutter on data quality is noteworthy (cf. Chen et al. [26]). The accuracy of the Wavex radar is in the range of ±10% in case of the three sea state parameters, cf. Miros [27]. The working principles of wave radars are conveyed in Barstow et al. [28]. (2) The Motion Response Unit (MRU) was located 78.5 m forward of the aft perpendicular (AP), 11.7 m above base line (BL), and at the centre line (CL). The MRU was installed in a socket filtering high frequency noise in the range of 50–100 Hz caused by e.g. thrusters or pumps. (3) The bow accelerometer is installed on a vertical pillar in the bosun store at the forward perpendicular and measures the vertical bow acceleration. (4) Moreover, four strain gauges are attached to longitudinal stiffeners at port and starboard in the aft section (50.3 m from AP) and amidships (118.7 m from AP). As described in [24], the port and starboard strain sensors are aggregated into two artificial or virtual sensors indicating vertical bending and axial stress in the aft and midship section. Lastly, (5) the propeller revolutions (rpm), (6) the rudder angle δ, (7) the GPS position, as well as (8) the Speed Over Ground (SOG) were obtained during data acquisition period.

The present study is focused on the prediction of the three integral sea state parameters: The significant wave height H_s, peak period T_p and mean encounter wave direction β. The parameters were obtained from the directional wave spectrum provided by the Wavex wave radar measurements. The significant wave height H_s relates to the zeroth order spectral moment m_n (with $n = 0$) and is defined using Eqs. (1)–(3). It is noted that the angular wave frequency is $\omega = 2\pi f$ and the directional wave energy density spectrum is denoted as $E(\omega, \mu)$.

$$H_s = 4\sqrt{m_0} \tag{1}$$

$$m_n = \int_0^\infty \omega^n F(\omega)\, d\omega \tag{2}$$

$$F(\omega) = \int_{-\pi}^{\pi} E(\omega, \mu)\, d\mu \tag{3}$$

The peak period T_p is extracted from the integrated wave spectrum $F(\omega)$ and corresponds to the period at which the wave energy density is highest, cf. Eq. (4).

$$T_p = \frac{2\pi}{\omega_p} \;,\; F(\omega_p) \equiv \max_\omega F(\omega) \tag{4}$$

The mean encounter wave direction β is calculated as the circular mean according to Longuet-Higgins et al. [29] in Eqs. (5)–(7). It is noted that the direction of the measured wave spectrum is relative to the ship's heading and, thus, no transformation is necessary.

$$\beta = \arctan(d/c) \tag{5}$$

$$d = \int_{-\pi}^{\pi} \int_0^\infty E(\omega, \mu) \sin(\mu)\, d\omega d\mu \tag{6}$$

$$c = \int_{-\pi}^{\pi} \int_0^\infty E(\omega, \mu) \cos(\mu)\, d\omega d\mu \tag{7}$$
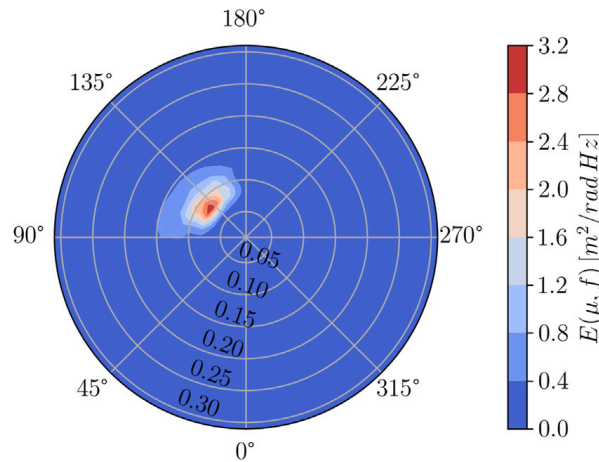
Fig. 3. Approaching directional wave spectrum E obtained at the yellow dot in Fig. 2. It is stressed that $\beta = 180$deg. indicates head waves and $\beta = 0$deg. stands for following waves; that is, in the polar diagram, the ship can be imagined to have its centre line aligned with the chord (diameter) from 0 to 180 deg.

In Fig. 3, a sample directional wave spectrum is depicted in its approaching form and it was obtained at the yellow dot in Fig. 2 on the 9th April 2008 (01:30 UTC). It is noted that $\beta = 180$deg. refers to head wave conditions. As can be seen, the ship encountered bow oblique waves and the calculated parameters using the aforementioned equations are: $H_s = 4.0$ m, $T_p = 12.1$ s and $\beta = 124.8$deg.

The available data is preprocessed and filtered in four individual steps: (1) Initially, the wave radar data was merged with the available GPS data and filtered for deep sea conditions using the enforced geographic boundaries (cf. Fig. 2). This reduced the wave radar samples from 10 790 to 7051. (2) In addition, the wave radar data was synchronized with the entirety of sensor readings, but samples with missing or corrupted recordings were disregarded. The majority of the readings had different sample frequencies ranging from 100 to 1 Hz and hence 25 min time series samples sampled with a consistent frequency of 5 Hz were extracted for each timestamp, i.e. in a 30 min interval. It is noted that 25 min are chosen as the maximum sample length since recordings close to the 30 min threshold were frequently missing. In addition, the focus of the herein presented work is on using smaller time frames, noting that other studies use longer durations, say, from 30 min in [23] to 60 min in [21]. The synchronization step decreased the size of the dataset further to 5182 samples. (3) Lastly, the data was cleaned manually from erroneous time series and (4) samples with significant wave heights $H_s < 0.5$m were excluded, as the vessel's response is negligible in this case. Altogether, the final dataset comprises 4779 samples, and the distributions of the resulting sea state parameters as well as the ship advance speed U_s are presented in Fig. 4.

As can be inferred from Fig. 4(a), the probability density of the significant wave height follows an exponential distribution. The theoretical probability density functions (PDF) Weibull and Gumbel are fitted to the measured data distribution using the Python library scipy.[1] The conclusions drawn in Nielsen and Ikonomakis [30] are also observed in this case: The Gumbel distribution matches the actual data with higher accuracy than the Weibull PDF. In addition, it is visible that the ship is experiencing harsh conditions despite seamanship and the utilization of weather routing. The distribution of the peak period, which is depicted in Fig. 4(b), is symmetrical and centred around 10 s. The histogram of the relative wave direction as shown in Fig. 4(c) conveys that predominantly head and following wave conditions are encountered which results from the fact that the Northern Atlantic is known for dominant westerly wind conditions. Furthermore, the speed variation including involuntary and voluntary speed loss due to waves can be seen in Fig. 4(d), where the data distribution arranges around an mean advance speed of 19 knots. In addition, the measured joint distribution of H_s and T_p is compared to long term wave climate statistics in the Northern Atlantic provided by Söding [31] in Fig. 5. It is noted that Fig. 5(a) is a combined scatter, kernel density and two-dimensional histogram plot for demonstrating the overall data distribution as well as possible outliers in parallel.

As can be inferred from the comparison of Figs. 5(a) and 5(b), the joint distributions are in good agreement, in general. However, the covariance in case of the measurements is slightly larger compared to the modelled long term statistics. It is noted that the latter is based on hindcast wave spectra determined from wind fields during 10 years, [31]. Hence, different sources and magnitudes of epistemic (i.e. systematic) measurement uncertainty may be responsible for the minor deviation of both joint distributions. On the other hand, it is stated in [24] that the calibration of the Wavex wave radar is based on earlier studies and has not been performed for this particular vessel. Even in case of a slightly biased wave radar, it has no direct impact on the machine learning methodology itself, as the offset remained the same throughout the measurement campaign. Furthermore, transfer learning enables neural networks to be retrained and adjusted on newly obtained sea state data — possibly even from different vessels.
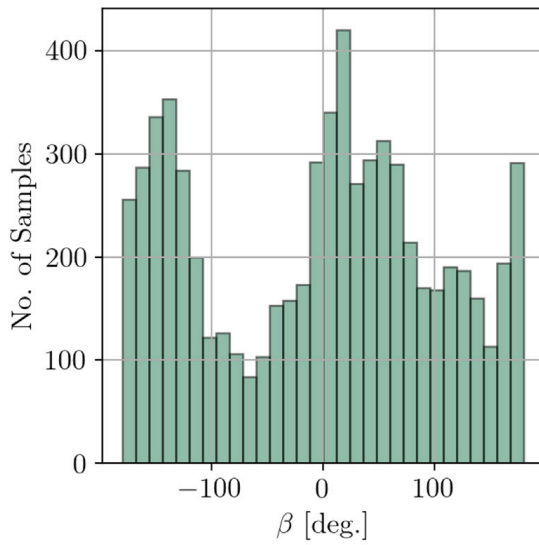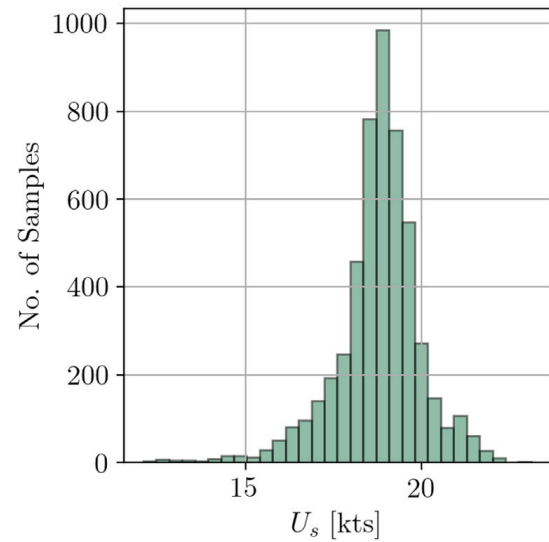
---

[1] https://scipy.org/.
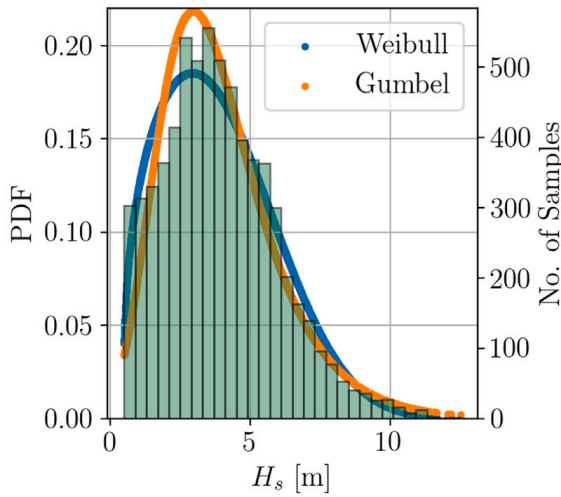
(a) Histogram of the significant wave height H_s including two parameterized PDF.
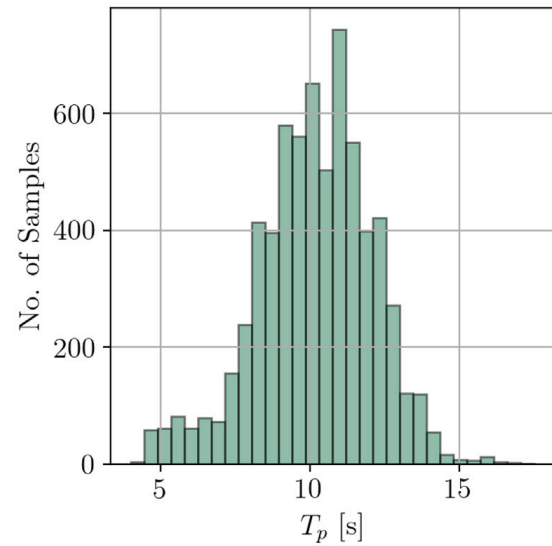
(b) Histogram of the peak period T_p.

(c) Histogram of the relative wave direction β, where |180| deg. indicates head waves and 0 deg denotes following waves.

(d) Histogram of the ship advance speed U_s according to GPS data.

Fig. 4. Data distributions of integral sea state parameters and ship advance speed.

3. Methodology

The surface wave elevation in realistic seaways is often assumed to be stationary, ergodic, and a Gaussian random process throughout the observation time. Following St. Denis and Pierson [32], a ship can be seen as a linear and time-invariant filter in the frequency domain under these assumptions. Specifically, this leads to the underlying principle of the wave buoy analogy, emphasizing a number of associated characteristics and concerns: (1) The filtering characteristics of ships are governed by their hull geometries and the relative size compared to waves. (2) Generally, a ship acts as a low-pass filter which means that it is irresponsive to higher wave frequencies. Furthermore, (3) when sailing with non-zero forward speed the Doppler shift is introduced and (4) the assumption of the ship as a linear and time-invariant filter is subject to uncertainty under severe sea states or changing wave conditions — especially in case of forward speed. All of these characteristics have an impact on a machine learning-based sea state identification methodology.

3.1. Data processing

Initially, Mittendorf et al. [33] suggest that transformations to non-linear and skewed target variables positively affect prediction accuracy even for non-linear regression algorithms, such as neural networks. For this reason, the logarithm is applied to the significant wave height H_s as it leads to a more symmetrical shape of the resulting distribution. It is noted that the application
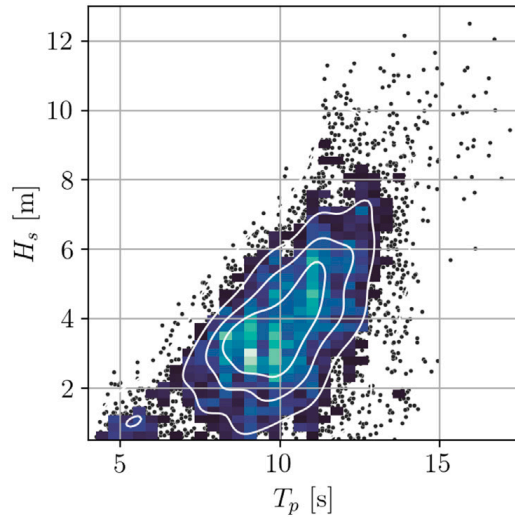
(a) Joint distribution of H_s and T_p in deep sea conditions for 1.5 years.
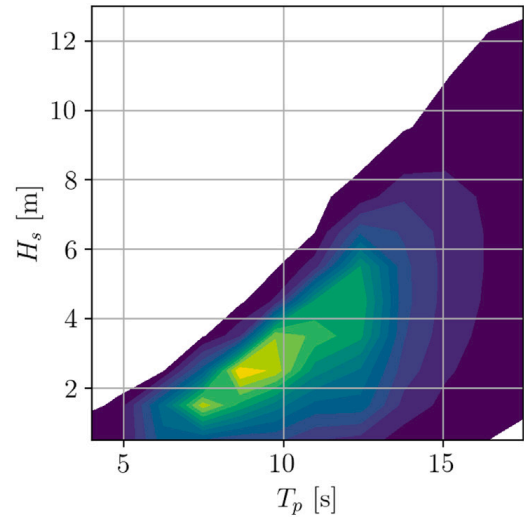
(b) Northern Atlantic wave climate statistics from Söding [31].

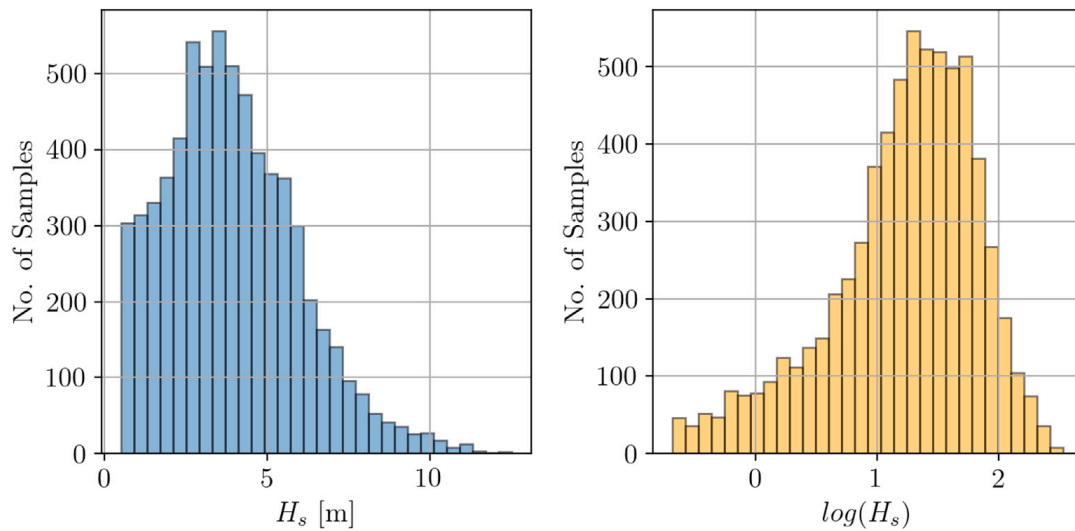**Fig. 5.** Comparison of the wave statistics in the Northern Atlantic.

Fig. 6. Effect of the logarithmic transformation on the H_s distribution.

of the log-transform is considered as standard practice in statistics and time series analysis for forcing a more symmetrical shape onto an exponential distribution.

As shown in Fig. 6, applying the logarithm to the H_s distribution leads, in fact, to a more symmetrical shape, even though it still shows minor asymmetry in this case. Moreover, the relative mean wave direction β is decomposed into corresponding sine and cosine values, in order to circumvent the circular ambivalence in the machine learning approach. The peak period remains unchanged. Ultimately, the output vector in both the time and frequency domain methodologies has the shape of $\{\log(H_s), T_p, \sin(\beta), \cos(\beta)\}$ and is passed through a linear activation function. It is noted that input and output data are normalized beforehand, which will be described at a later stage.

In an earlier study of Mittendorf et al. [34], a time and frequency domain approach were applied based on simulated data of a vessel in unimodal and unidirectional sea states. In that study, a frequency domain model was trained on spectral moments of the auto cross response spectra and the peak spectral ordinates, as well as the corresponding peak frequencies obtained from the off-diagonal cross response spectra. The temporal method, on the other hand, was based on raw heave, pitch and roll time series. In the present work, raw 5 min acceleration time series with a sample frequency of 5 Hz are fed into the time domain models based on the findings of [34]. This leads to an input matrix of $1500 \times M$, where M is the number of features. A sample plot of 5 min acceleration time series resulting from the encountered sea state in Fig. 3 is depicted in Fig. 7. It is noted that a coordinate system is adopted herein, where the x_1-axis coincides with the ship's centreline (positive forward) and the x_2-coordinate points upwards.
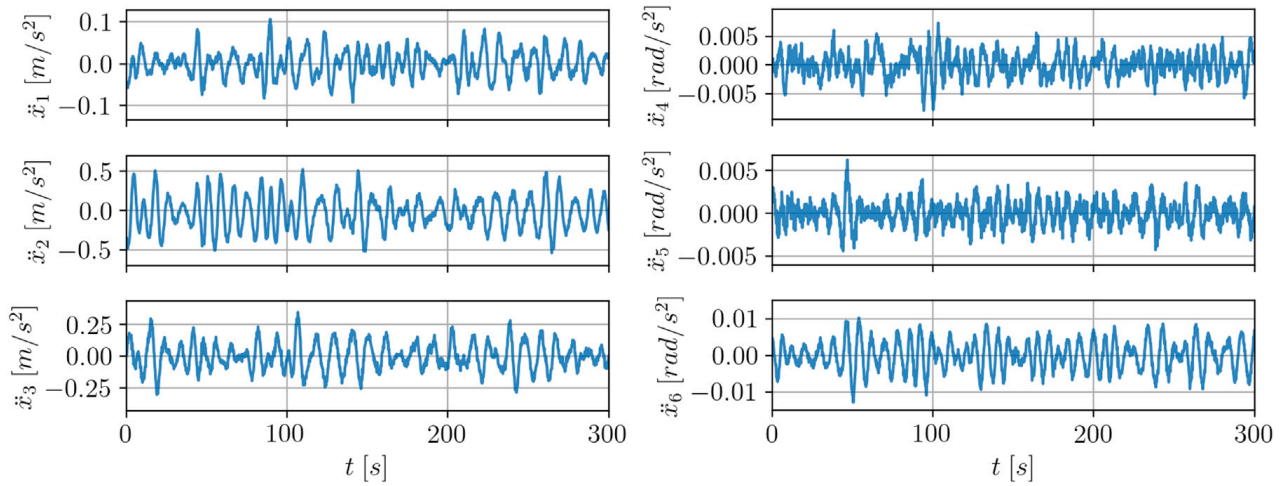
Fig. 7. Measured accelerations corresponding to Fig. 3 in 6 degrees of freedom, namely surge, heave, sway, roll, yaw and pitch (starting at the upper left plot).
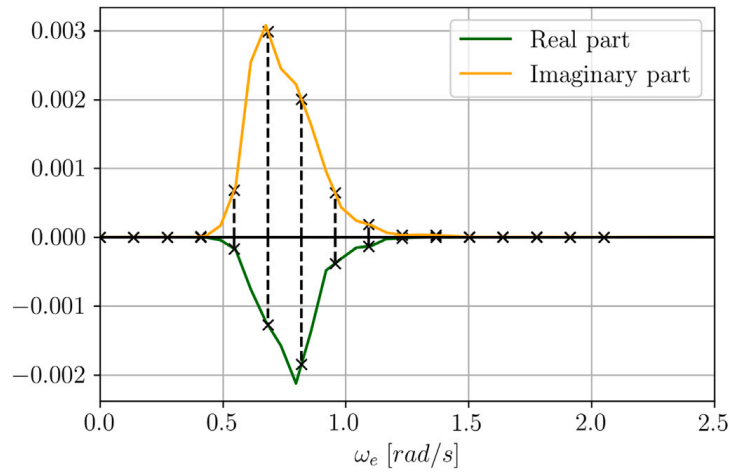

Fig. 8. Pitch-heave acceleration cross response spectrum demonstrating the frequency discretization with 16 spectral ordinates (black) resulting from the conditions in Fig. 3.

It is stressed that not only the 6 Degree of Freedom (DOF) accelerations – shown in Fig. 7 – are considered in the feature space, but also the bow acceleration and the strain measurements at the aft and midship sections. The underlying reasons are that both a second reference point on the ship and an indication of the hull flexure might add value to the sea state identification approach. The optimal number of features M will be determined in sensitivity studies in both the temporal and spectral domain.

The prediction of the relative wave direction is a function of the individual phase differences of accelerations in several degrees of freedom, which are obviously directly available in the time domain, cf. Fig. 7. In the frequency domain, however, cross spectral analysis of two individual time series, is used as predictor for β. In contrast to [34], sequences of spectral ordinates extracted from cross response spectra are fed into the neural network. This leads to an $N \times N$ complex-valued matrix, where N is the number of the considered DOF. It is noted that only one side of the off-diagonals are taken into account, as the matrix is complex conjugate symmetric. It is noted that the sensor recordings of the bow acceleration and the strain measurements are also examined in the sensitivity study and may be part of the final feature space. The matrix of cross-spectral densities is determined using the Welch algorithm (cf. Welch [35]) with a Hanning window as well as a segment length of 512. The employed software package for the digital signal processing procedures is scipy. In Fig. 8, the pitch-heave acceleration cross spectrum and the chosen frequency discretization are visible for the same conditions as in Fig. 3.

In view of Fig. 8, it is appreciated that the discrete spectral ordinates are extracted in the range of $\omega_e \in [0, 2.05]$ rad/s with a number of discrete frequencies denoted as Ω. The cut off frequency is selected for filtering the measurement noise and for responses due to excitations which are not related to waves, e.g. hull-structural vibrations. From Fig. 8, it appears that the peak value is well captured in case of the imaginary part of the cross response spectrum, but not in case of the real part. Thus, the number of discrete frequencies Ω is part of a second sensitivity study, in which also the effect of the sample length on prediction accuracy will be investigated. For reference, 71 and 100 discrete spectral ordinates were extracted from individual cross response spectra in the work of [21,23], respectively. Lastly, the shape of the input matrix for the machine learning model in the spectral domain has a shape of $\Omega \times M$, where M takes a maximum value of 43. Firstly, 30 real and imaginary sequences from the off diagonals of the

cross response spectra and 6 sequences from the auto cross response spectra, i.e. the diagonal elements, will be part of the largest possible feature space (with $N = 6$). Moreover, 2 response spectra from the strain measurements in the aft and midship sections may be included in the input matrix. In addition, the real-valued response spectrum of the bow acceleration as well as 4 real and imaginary sequences from the cross spectra of bow and heave/pitch accelerations may be also considered as features. It is stressed that only real valued numbers are permitted as input to the models. The optimal number of features M and thus the shape of the input matrices in both domains will be obtained in sensitivity studies presented in the further course of this contribution.

Before feeding the data into the neural networks, the dataset is split into training and validation sets, where the latter is exclusively used for the model assessment. The validation set makes up 20% of the initial dataset and comprises 956 samples. The training set, on the other hand, has a sample size of 3823. Neural networks are not scale invariant, thus, the individual sequences as well as the elements of the output vector were normalized according to the extreme values of the training set.

3.2. Model architectures

The machine learning task itself is considered as a supervised regression approach and the models are trained on sequences from the temporal and spectral domain. As has been stated in Section 1.3, the focus of this contribution is on deep learning methods. Artificial neural networks are utilized as universal function estimators and are considered as composite functions $\mathcal{F}$ of two essential building blocks: The weight matrix and the non-linear activation function. Concise theoretical considerations of artificial neural networks are presented in Mittendorf et al. [33] and more detailed information is delivered by Goodfellow et al. [16].

Broadly speaking, deep learning architectures are capable of coping with high dimensional input tensors and have the ability of implicit feature construction, i.e. derive meaningful data *themselves*. Additional advantages of deep neural networks over traditional machine learning algorithms include higher scalability and increased generalization capability. The caveats of deep neural networks are on one hand the relatively high computational cost of their training process, and on the other hand the tendency of overfitting on small datasets due to their relatively large amount of trainable parameters. While, one particular problem has been most influential on the development of deep learning methods: As the name suggests, deep learning achieves its superior performance by stacking individual layers. However, very deep models suffer from the *vanishing gradient problem*. The parameters of neural networks are updated under backpropagation in a gradient-based optimization problem of an arbitrary loss function $\mathcal{L}$. However, with an increasing number of hidden layers, the magnitude of the gradient, i.e. the impact on parameter adjustment, reduces. Batch normalization as proposed by Ioffe and Szegedy [36] standardizes the output of hidden layers before applying the activation function mitigating vanishing gradients by resetting the parameter distribution. In addition, batch normalization leads to robust neural networks reducing the need for hyperparameter optimization. The simplified activation function Rectified Linear Unit ($ReLU = max(0, x)$) not only decreases the chance of vanishing gradients, but also speeds up the training process, as it is faster to differentiate in comparison to traditional activation functions, such as the hyperbolic tangent function. Moreover, numerous advanced architectures, such as the LSTM (Long short-term memory) by Hochreiter and Schmidhuber [37] or residual networks by He et al. [38], were proposed for alleviating the occurrence of vanishing gradients. In this paper, the residual network as well as the Inception model proposed by Szegedy et al. [39] are applied.[2] The herein employed models feature one-dimensional convolutional layers, since they are best-suited to sequential data. Convolutional layers utilize spatially shared weights and are often followed by a pooling or subsampling procedure. In theory, the convolution procedure has a multidimensional tensor as input, which is modified by a kernel whose parameters are trainable. For more elaborate details consult the work of Krizhevsky et al. [40]. The concepts and structures of the herein applied models are described in the following paragraphs. Moreover, the model architectures and hyperparameters are the same in both the time and frequency domain for consistency.

3.2.1. Residual network

The Residual Network (ResNet) features a block-wise architecture and the blocks are bypassed by identity mappings or gates. The ResNet was proposed by He et al. [38] and its concept draws inspiration from the pyramidal cells of the cerebral cortex and its skip connections. The feedforward residual network is made out of multiple residual blocks as shown in Fig. 9(a) and the skip connections allow the development of very deep models without vanishing gradients. The internal shape of the proposed residual block comprises three convolutional layers of constant filter size, but with 8, 4 and 2 kernels, respectively. Filter size refers to the number of output dimensions, whereas the kernel size denotes the length of the receptive field, i.e. of the convolutional window. Lastly, the output of the block is the sum of the last layer's output and the input matrix gated through the skip connection activated by ReLU, as can be seen in Fig. 9(a). It is noted that the identity mapping is multiplied by a linear projection for expanding the channels of the skip gate, in order to match the shape of the residual block's output. The 4 blocks of the chosen model have filters with the size of 32, 64, 64 and 64, respectively. The tail part of the ResNet consists of an average pooling layer and a fully connected layer with a width of 50 neurons. It is noted that batch normalization is applied after each convolutional layer.

---

[2] An early version of the Inception Network, namely the GoogLeNet, and a residual network were the winners of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014 and 2015, respectively.
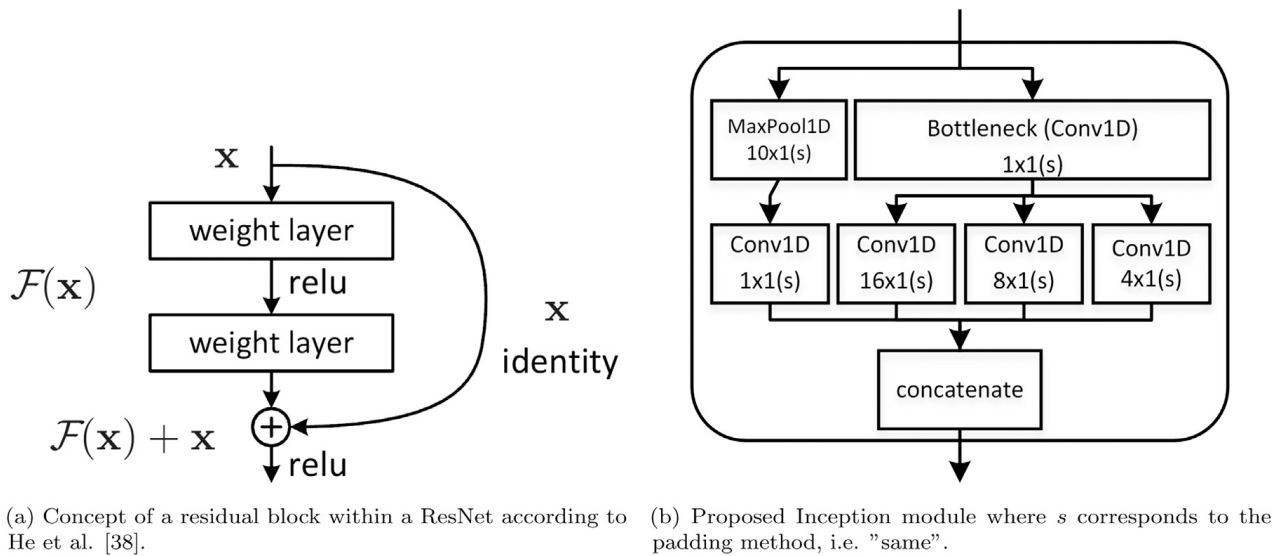
(a) Concept of a residual block within a ResNet according to He et al. [38].

(b) Proposed Inception module where s corresponds to the padding method, i.e. "same".

Fig. 9. Underlying concepts of the ResNet (a) and Inception (b).

3.2.2. Inception network

The Inception architecture was initially proposed by Szegedy et al. [39] in an end-to-end image classification procedure. The model follows a similar block-wise methodology as the ResNet employing the novel Inception module as its building block. The fundamental idea is that several convolutional layers of constant filter size of 16, but with different kernel sizes, i.e. different receptive fields, are employed simultaneously at the same level. In doing so, the individual receptive fields capture different patterns in the feature maps at varying scales. This concept is inspired by principles of the visual cortex in the human brain. The Inception architecture is named after the eponymous movie[3] indicating the common premise of embedding either networks within networks in case of the model or dreams within dreams as in the movie. The architecture evolved in multiple iterations to Inception-v4, where the model was extended with residual connections for increased performance, Szegedy et al. [41]. The proposed Inception module has been adapted for sequential data following Ismail Fawaz et al. [42], i.e. one-dimensional convolutional layers are employed. As can be inferred from Fig. 9(b), the module splits into four elements: (1) The bottleneck or dimensionality reduction element is introduced in front of the computationally intensive parallel layers and takes cross-channel patterns into account. The bottleneck layer reduces computational cost as well as the number of trainable parameters. (2) The second element is made out of the three parallel convolutional layers with different kernel sizes, i.e. the core part of the Inception module. In the present work, kernel sizes $l \in \{16, 8, 4\}$ are applied. (3) The third part of the module is the skip connection, which is inspired by the ResNet, and consists of a pooling layer and another dimensionality reduction layer. (4) The two gates are concatenated along the depth dimension in the last element. It is stressed that the Inception architecture has additional outer residual skip connections bypassing entire Inception modules. The general structure of the proposed Inception network can be inferred from Fig. 9(a) when replacing the residual block with the Inception module shown in Fig. 9(b). The herein proposed Inception network is defined with filter size of 16 and 8 inception modules in total with individual residual connections. In front of the output layer, the activations are fed through an average pooling layer and batch normalization is also applied in the entire model prior to all activation functions.

3.2.3. Benchmark models

For comparison of ResNet and Inception to state of the art literature, two models are adopted as benchmarks: The convolutional neural network (CNN) from Kawai et al. [21] and the multichannel convolutional LSTM (MLSTM) network as proposed by Düz et al. [20]. The former network was applied in the frequency domain on cross spectral sequences, whereas the latter one used raw time series as input for the prediction of sea state parameters.

3.3. Multi-task learning

Multi-Task Learning (MTL) has been pioneered by Caruana [43] and is considered as *parallel* transfer learning. As opposed to a multioutput-regression task, each output is now considered as a separate task and has, thus, its own dedicated branch of fully connected hidden layers and corresponding output layers. In the present paper, it is thought that considering sea state estimation as a multi-task learning approach may be beneficial as there are obviously physical interdependencies between H_s and T_p. According to Ruder [44], a hard parameter sharing configuration is selected herein, as can be seen in Fig. 10, which mitigates overfitting and improves generalization immensely.

---

[3] Inception is a science fiction movie from 2010 written and directed by Christopher Nolan.
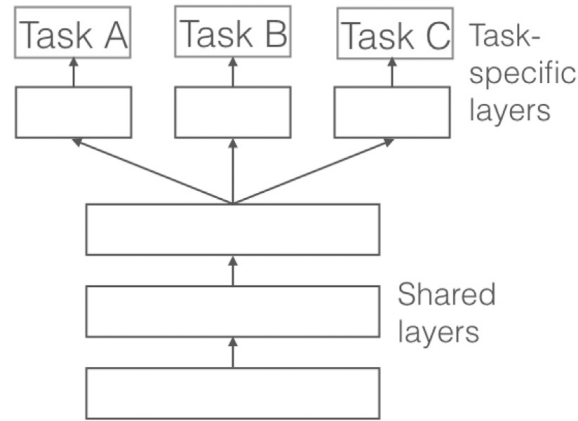
Fig. 10. Hard parameter sharing concept from Ruder [44].

In view of Fig. 10, it is appreciated that the model is said to learn several tasks at once using a shared latent data representation. Following [44], MTL improves the model's generalization capability based on eavesdropping, i.e. the network learns complex features and the relationships between other dependent variables, instead of receiving them as inputs. In addition, MTL leads to increased regularization for improving the model's robustness by reducing Rademacher complexity, i.e. the model's tendency to fit to stochastic noise. In this work, each task-related branch consists of two hidden layers with 100 and 50 neurons and an output layer. Two different MTL architectures are examined in this contribution: (1) The *MTL* architecture has three branches for each of the sea state parameters, however, the output layer of the third branch has a length of two, as both the sine and cosine of β are considered. (2) The *MTL+* architecture, however, takes the advance speed U_s as a fourth task into account. It has been demonstrated by Caruana and de Sa [45] that potential input features with low variance may be better predicted as an additional task instead, when the overall input data is subject to large variance — as in the present case. In addition, the inclusion of the mean advance speed in the input space is impractical due to the mismatching dimensions. It is thought that the model is able to derive an understanding of the Doppler shift and thus the forward speed for the correct prediction of the sea state. Hence, it will be studied herein, whether the prediction of the forward speed enhances the prediction of the sea state parameters and could act simultaneously as a reliability indicator. Lastly, the tasks of the two investigated architectures are related or even complementary, however, adversarial training as in [23] is also feasible using multi-task learning.

3.4. Training setup

In the present work, the ADAM optimizer – proposed by Kingma and Ba [46] – is employed as the learning algorithm. The ADAptive Momentum estimation (ADAM) optimizer uses the exponential moving average of the gradient and scales the learning rate, i.e. the step size, according to the squared gradient. Moreover, the applied loss function is the mean squared error, the chosen batchsize is 64, and the initial learning rate is set to 10^{-3}. Mak and Düz [47] suggest that k-fold cross validation with shuffled samples leads to increased performance in a sea state estimation methodology. Hence, we apply shuffled 5-fold cross validation in 300 epochs. The employed model-checkpoint-callback stores the model's parameters with the smallest cross validation loss, in order to save the model with the highest generalization capability before overfitting occurs with an increasing number of epochs. The computations of the training procedures were carried out using a GPU node of the DTU computing center equipped with two Nvidia© Volta-100 GPUs, each with 16 GB of memory (RAM) and multiple Intel© Xeon© Gold 6126 CPU with 2.60 GHz were used. Moreover, the utilized programming language is Python 3.6 and the deep learning framework is TensorFlow 2.6 as proposed by Abadi et al. [48]. For GPU parallelization of the computations CUDA (Compute Unified Device Architecture) was employed.

4. Results and discussion

In the following, the obtained results are presented and discussed. The used metrics are the root mean squared error ($RMSE$) and the mean absolute error (MAE), and their definitions are given in Eqs. (8) and (9). It is noted that $\hat{y}_i$ indicates the model's prediction, whereas y_i refers to the wave radar data.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2} \tag{8}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i| \tag{9}$$

The $RMSE$ attributes more weight to outliers or variance and the MAE indicates the magnitude of the error without considering the sign. Due to the circular ambiguity in case of the mean encounter wave direction, a similar approach as in Nielsen [49] is
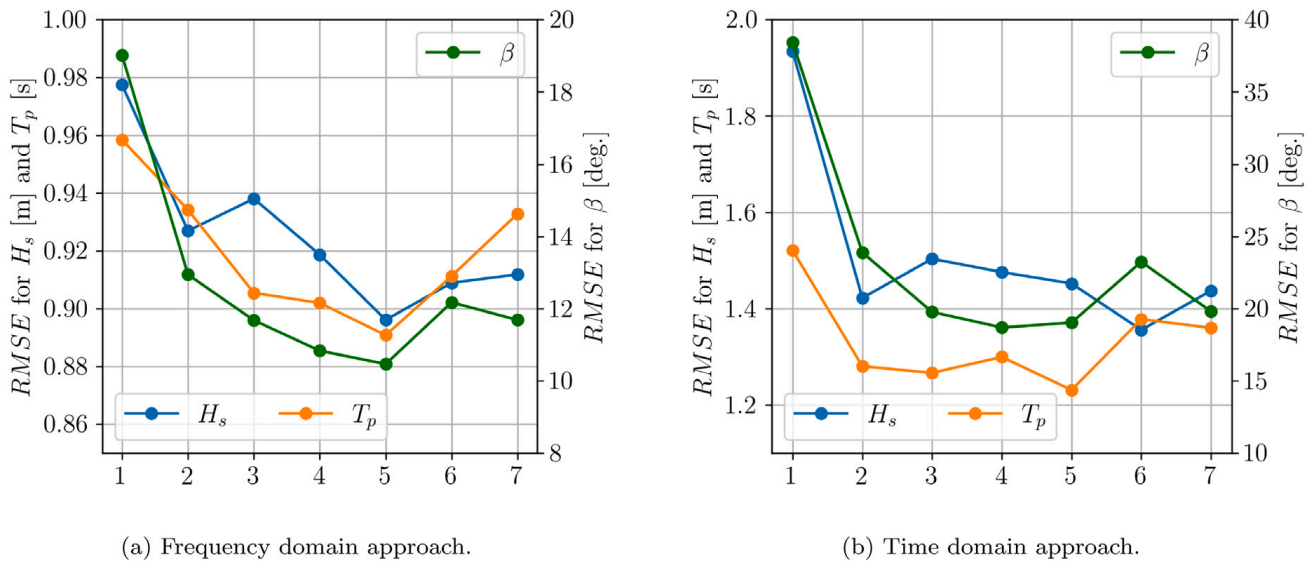
(a) Frequency domain approach.

(b) Time domain approach.

Fig. 11. Optimal selection of the considered ship responses according to the performance on the validation set. Note that the individual scales on the ordinates are different.

adopted for error calculation, i.e. absolute values of β are considered when calculating the performance metrics. It is stressed that the accuracy on the *validation* dataset is assessed exclusively in this section and that the model performance on the *training* set is presented in the Appendix.

4.1. Sensitivity studies

Model-based approaches of the wave buoy analogy, e.g. Nielsen [7], traditionally consider heave and pitch as well as one additional asymmetric motion, such as roll or sway for satisfactory results with a distinguishing between waves approaching from port or starboard side. The use of surge and yaw, however, is generally not pursued, since the RAOs provided by, say, strip theory often turn out to be unreliable due to the missing restoring term. However, as the present study is not reliant on RAOs, all 6 DOF accelerations are considered in the sensitivity study. Mittendorf et al. [33] show in multiple iterations that there is an optimal number of features M to be found in a machine learning approach. In this work, the trade-off between accuracy and complexity will be determined in a convergence study, i.e. the number of features and indirectly the number of measurement devices are increased incrementally. Therefore, multiple models are trained on varying, expanding input matrices, and their $RMSE$ on the validation dataset is taken as the underlying score for all three sea state parameters. Based on the findings of Mittendorf et al. [34], a ResNet was chosen as the baseline model using a frequency discretization with $\Omega = 42$ components. In Fig. 11, the following combinations are considered in individual scenarios: **(1)** heave and pitch acceleration, **(2)** + roll acceleration, **(3)** + sway acceleration, **(4)** + surge acceleration, **(5)** + yaw acceleration, **(6)** + vertical bow acceleration and **(7)** + deck strain measurements in the aft and mid-ship positions.

When comparing Figs. 11(a) and 11(b), it is appreciated that the shape of the individual curves of the scores for H_s, T_p and β is generally in qualitative agreement — even across domains. Moreover, it stands out that the consideration of all 6 DOF accelerations yields the lowest scores, i.e. the highest out-of-sample accuracy. Thus, it is concluded from a practical point of view that, in the present case, only the motion recordings from the MRU are needed; in the sole event of sea state identification. In fact, bow acceleration and strain measurements do not lead to increased accuracy, but add complexity and noise into the machine learning model. This may be attributed to the fact that the contained information in the mentioned features is redundant, but a larger input matrix leads to a larger number of model parameters (i.e. complexity). Interestingly, the magnitude of the $RMSE$ of the targets in the time domain are roughly twice the ones in the frequency domain model, which is likely to be caused by the shorter considered sample length. As expected, the inclusion of roll acceleration in the second scenario leads to the most significant increase in accuracy for all sea state parameters. In theory, H_s and T_p, i.e. the energy content of the spectrum and its distribution, may be determined by just considering heave and pitch. However, the consideration of an asymmetric acceleration is essential for the sufficient estimation of β. Lastly, an aggregated sensitivity indicator is used, i.e. irrespective of different conditions, in this study. However, Montazeri et al. [50] as well as Andersen and Storhaug [51] indicate that the sensitivity of individual responses is not constant, but variable under different sea state conditions. For instance, the relative importance of considering the vertical bending moment or its axial stress in the hull girder increases in relatively low period waves, due to increased bandwidth of these RAOs compared to the motion response RAOs, [50]. Therefore, [51] suggest an adaptive selection of the optimal responses in case of a model-based methodology. Yet, it is stressed that a dynamic methodology is not feasible in this case, as the underlying computation graph of neural networks is static. Ultimately, the optimal number of features is identified as $M = N = 6$ in both domains.
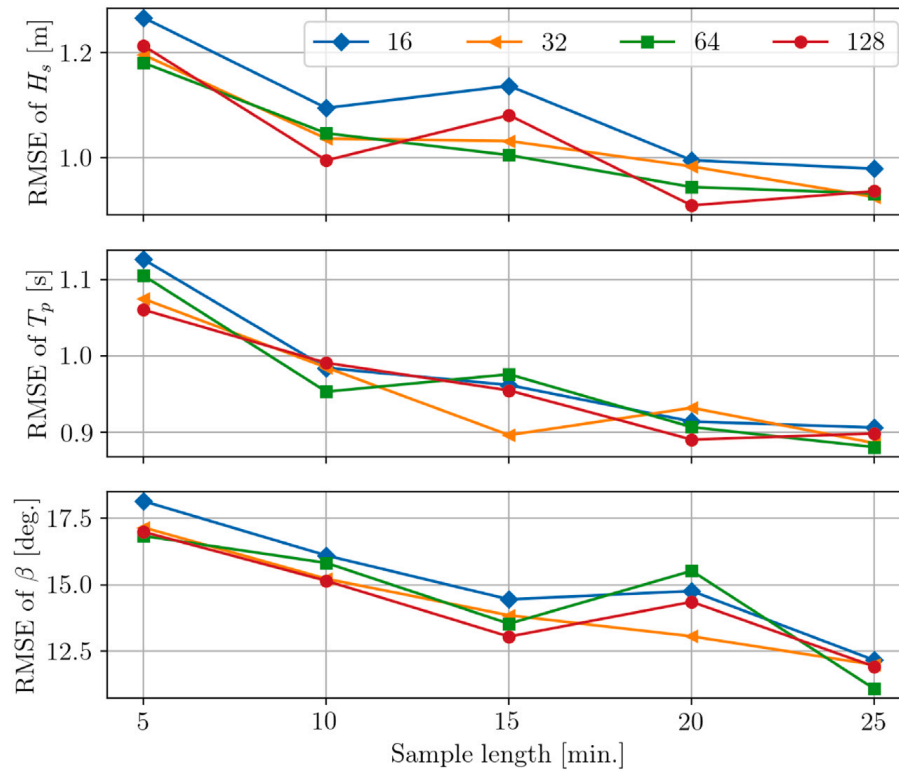
Fig. 12. Sensitivity study for sample length L and frequency discretization with Ω indicated in the legend.

Linear spectral analysis is based on the assumption of stationarity, however, it has been shown by Brodtkorb et al. [52] that the time-invariance of sea state parameters in 30 min samples is subject to uncertainty — even in a simulation-based study under zero-forward speed conditions. It is thought that forward speed and in-situ measurement data result in additional uncertainty regarding stationarity. In their work [52], it is deduced that the relative wave direction is most susceptible to instationarity, whereas H_s and T_p are seen as more stable. For minimizing this kind of uncertainty, it is the objective of this work to use shorter sample lengths as compared to 30 min in [23] and 60 min in [21]. However, the uncertainty regarding stable spectral analysis is a counteracting contribution, which will increase with shorter time series samples. Hence, the upcoming sensitivity study is about obtaining the compromise of the smallest possible sample length with the most accurate frequency resolution at the same time. The sample length L is discretized as $L \in \{5, 10, 15, 20, 25\}$min. Obviously, the frequency discretization is interrelated and, thus, the number of frequencies Ω is considered as the second discrete variable in the study, i.e. $\Omega \in \{16, 32, 64, 128\}$. For reference, [21] use 71 and [23] employ 100 discrete frequencies. It is noted that the $RMSE$ of the model applied to the validation dataset is taken as the score.

In view of Fig. 12, the conclusions drawn in Iseki and Nielsen [53] are confirmed, i.e. the longer the time frame, the clearer the frequency resolution, and thus the higher the model score. For H_s and T_p, monotone curves with minor variance are apparent, which is expected as the accuracy obviously will be higher with less variance in the frequency representation. However, the same is visible in case of β, even though a local minimum is located at $L = 15$min. This is in contradiction to [52] and results, on one hand, from the considered conditions, i.e. in deep water and without any major heading changes due to steering (mostly stationary conditions). On the other hand, the presented sensitivity score is highly biased by the model and its learning algorithm. Thus, the model not only reflects physical interdependencies, but also model uncertainty, since the loss function, which is minimized during training, is the sum of the errors of all three target variables. Moreover, it is trivially concluded that, in general, a finer frequency discretization results in lower error values, i.e. higher model performance or accuracy. However, in Fig. 12 it is visible that $\Omega = 64$ leads to the smallest validation error, which may be explained by the increased complexity of using 128 discrete frequencies. In the latter case, no additional information is provided to the model, but the large feature space leads to more trainable parameters, i.e. complexity. Ultimately, 25 min are identified as the best suited sample length and 64 discrete frequencies are chosen as the dimension of the feature space for the frequency domain model. Subsequently, it can be seen that the uncertainty pertaining to the spectral analysis is of greater relevance than the uncertainty of the time invariance assumption. The X-band wave radar provides samples in a 30-min interval at time t_{radar}. Thus, for consistency, we choose the starting point $t_1 = t_{radar} - 25$ min in the frequency domain approach, and in the time domain approach $t_2 = t_{radar} - 5$ min. On a side note, it can be inferred that the higher validation error in the time domain, cf. Fig. 11(b), results from the smaller sample length as the errors using $L = 5$ min are of similar magnitude in both time and frequency domains.

Table 2

Metrics of model performance on the validation set in the frequency and time domains. The lowest error values are indicated in bold font.

		Frequency domain				Time domain			
		ResNet	Inception	MLSTM	CNN	ResNet	Inception	MLSTM	CNN
H_s	$RMSE$	0.932	**0.927**	1.089	1.273	1.702	**1.345**	2.702	1.358
	MAE	**0.620**	0.626	0.767	0.868	**0.909**	0.916	1.979	0.991
T_p	$RMSE$	0.881	**0.878**	0.925	1.051	**1.253**	1.266	2.379	1.300
	MAE	0.539	**0.529**	0.596	0.714	0.935	**0.847**	2.122	0.975
β	$RMSE$	11.09	**10.50**	12.12	14.87	19.03	**16.67**	95.34	26.00
	MAE	12.34	**11.29**	77.60	16.67	12.33	**11.29**	77.60	16.67

(a) Frequency domain approach.

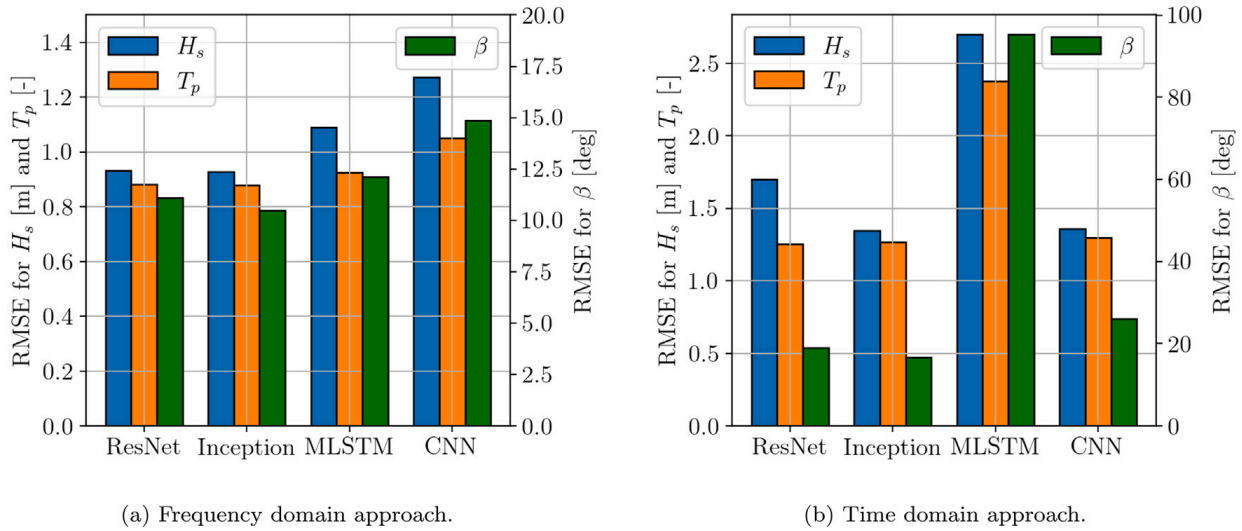(b) Time domain approach.

Fig. 13. Histograms of the model performance on the validation set using the $RMSE$. Note that β is presented on the secondary axis.

4.2. Model assessment

The performance of frequency and time domain models is presented in the following in parallel. In Table 2, the metrics of the prediction error for the validation dataset are presented for all four models in both domains. Moreover, the lowest error values, i.e. the highest accuracy, are indicated in bold font. For the sake of clarity, corresponding histograms of the $RMSE$ taken from Table 2 are provided for all models in Figs. 13(a) and 13(b).

In view of Table 2, it is stated that the Inception network yields the lowest errors, i.e. the highest accuracy, on the validation dataset in either domain. However, it turns out that the error values of the ResNet are comparable and even lower in certain cases indicating similar model performance compared to the Inception architecture. The performance of MLSTM and CNN in the frequency domain are considered as sufficient. However, the MLSTM shows a considerable drop-off in accuracy in the time domain. It is thought that this results from the parallel architecture, i.e. splitting the model into a convolutional and a recurrent branch, leading to decreased robustness. In contrast, convolutional layers are frequently used as feature construction layers prior to LSTM cells in a sequential fashion. Moreover, it is obvious in the spectral approach (cf. Fig. 13(a)) that out of sample accuracy is dependent on model complexity, as the Inception network comprises the most trainable parameters and the CNN the least. Conversely, in the time domain, the overall picture is scattered, as the approach is subject to more variance and stochastic noise. The variance results from the short considered sample lengths and is amplified by the larger number of parameters due to the increased input matrix compared to the frequency domain approach. Thus, the temporal model is more susceptible to perturbations because of its complexity. Ultimately, the Inception network is used as the underlying model for the multi-task learning methodology and the results in both domains are conveyed numerically in Table 3 and visually using histograms in Fig. 14.

In view of Table 3, it is stated that the MTL-Inception models show satisfactory accuracy on the unseen validation dataset in both configurations. In general, they achieve higher performances than the corresponding multi-output approach as presented in Table 2. It can be seen in Fig. 14 that multi-task learning, indeed, may have a beneficial impact on model performance in a sea state identification approach, however, this applies predominantly to the time domain. In Section 3.3 it has been described that MTL reduces the model's tendency to fit to stochastic noise, i.e. Rademacher complexity, (cf. Ruder [44]) and that the MTL+ architecture performs well under noisy conditions, [45]. As can be seen in Table 3, the prediction of advance speed U_s in the MTL+ architecture is considered as satisfactory and, in fact, facilitates the prediction of sea state parameters despite the inherent uncertainty in the time domain approach. The $RMSE$ is a measure of the prediction's variance and decreased consistently applying both MTL approaches. In turn, the effect of MTL in the frequency domain is not as pronounced, since there is not as large variance as in the temporal

Table 3

Metrics of model performance on the validation set in the frequency and time domains using multi-task learning. The lowest error values are indicated in bold font.

|		Frequency domain			Time domain		
		Inception	MTL	MTL+	Inception	MTL	MTL+
H_s	$RMSE$	0.927	0.883	**0.871**	1.345	**1.228**	1.248
	MAE	0.626	0.582	**0.580**	0.953	0.885	**0.858**
T_p	$RMSE$	0.878	**0.855**	0.903	1.266	1.124	**1.087**
	MAE	0.529	**0.472**	0.501	0.916	0.752	**0.744**
β	$RMSE$	**10.50**	12.82	11.57	**16.67**	18.75	16.84
	MAE	**6.446**	6.916	6.767	11.29	11.56	**10.33**
U_s	$RMSE$			0.322			0.441
	MAE			0.227			0.331
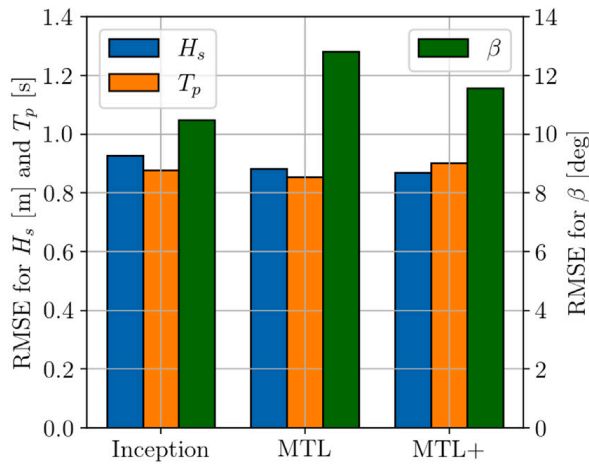
(a) Frequency domain approach.
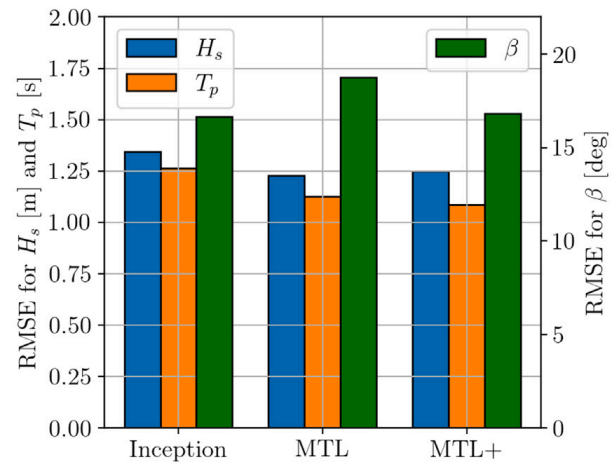
(b) Time domain approach.

**Fig. 14.** $RMSE$ histogram indicating the performance on the validation dataset of different multi-task architectures. Note that β is presented on the secondary axis.

approach. Hence, the spectral approach is considered as more robust and it is concluded that the MTL methodology increases robustness as well as accuracy for the temporal approach notably. Altogether, the MTL version is taken in the frequency domain as the final model and the MTL+ architecture is adopted in the temporal approach. In Fig. 15, the predictions on the validation set of the spectral MTL-Inception model are presented for ground truth values. Additionally, the loss curves of training and cross validation set are presented on a logarithmic scale. It is stressed that Figs. 15(a)–15(c) are of similar type to Fig. 5(a). Moreover, corresponding correlation and loss plots of the temporal MTL+-Inception model are depicted in Fig. 16 for comparison.

Generally, it is seen from Figs. 15 and 16 that both model types are capable of providing accurate predictions of the governing sea state. Moreover, it proves again that the temporal approach is characterized by higher variance compared to the spectral approach. This applies not only to the prediction of the three target variables, but also to the learning procedure, i.e. the behaviour of the loss functions. The loss curves are already volatile in the frequency domain due to the application of 5-fold cross validation, however, the variance is even larger in the time domain. From Fig. 16(d), it can be inferred that the time domain model was trained for only 200 epochs in order to save computational resources, as it was evident in the spectral approach (cf. Fig. 15(d)) that the model generalization converged before reaching 100 epochs anyhow. Therefore, implementing an early stopping callback is a key aspect of extending the proposed training procedure. In case of the peak period T_p, minor heteroskedastic behaviour of the predictions is visible, i.e. the variance is not constant for the entire definition range, cf. Figs. 15(b) and 16(b). Instead, the variance increases, as the magnitude of T_p decreases. It is assumed that this is due to the low pass filtering effect of the vessel, i.e. there is no vessel response in low-period waves introducing uncertainty. However, in case of higher H_s and T_p values, both spectral and temporal models exhibit reduced accuracy, which is due to the unbalanced training data and amplified by possible non-linear behaviour in more severe sea states. Furthermore, when comparing both H_s and T_p for both models, it shows that the time domain model is subject to larger uncertainty due to the more widespread joint distributions arranged around the identity line in the figures. In turn, the variance of β reveals only a small increase in the time domain model, when investigating Figs. 15(c) and 16(c). However, several outliers are visible in Figs. 15(c) and 16(c) in beam-sea conditions, i.e. $\beta = \pm 90$deg. It is noted that the scattered squares in Fig. 16 are part of the 2D histogram and their variance stems from the uncertainty of the predictions in the time domain. Obviously, the sea state identification approach is subject to uncertainty resulting from e.g. the unknown loading conditions, as has been mentioned in Section 2. Moreover, possible unfiltered measurement errors in both wave radar and MRU, i.e. corrupted samples, increase the
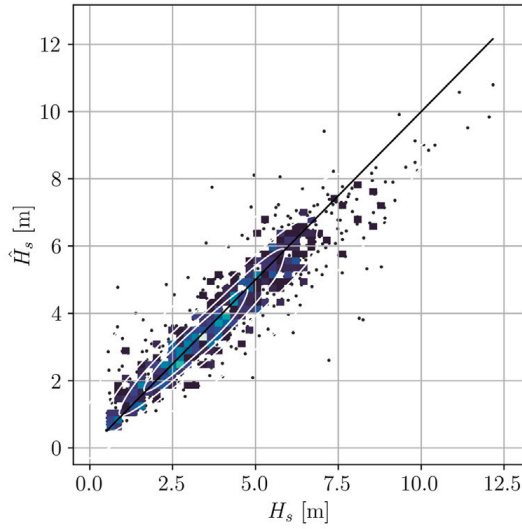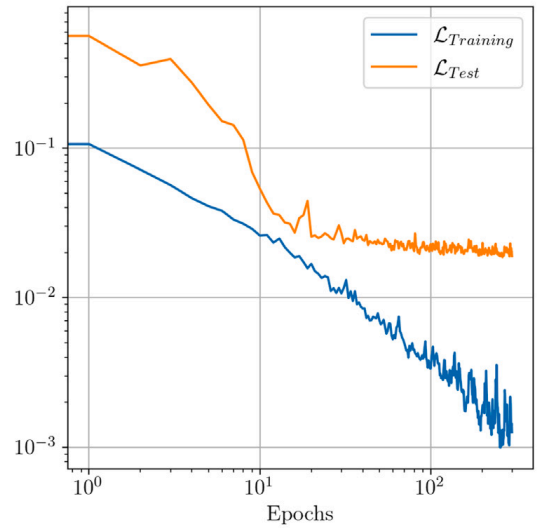
(a) Predictions of the significant wave height H_s based on the validation set.
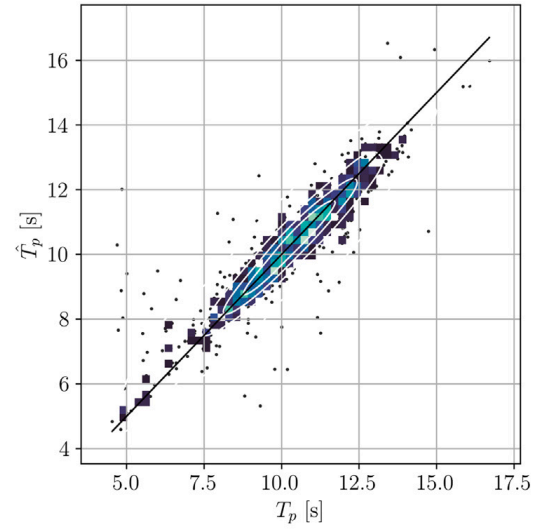
(b) Predictions of the peak period T_p based on the validation set.
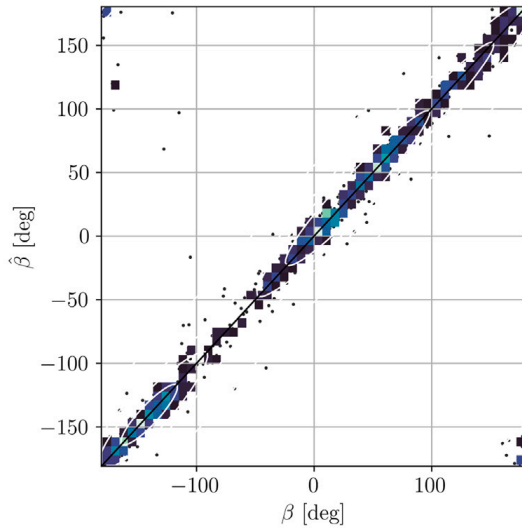
(c) Predictions of the relative wave direction β based on the validation set.
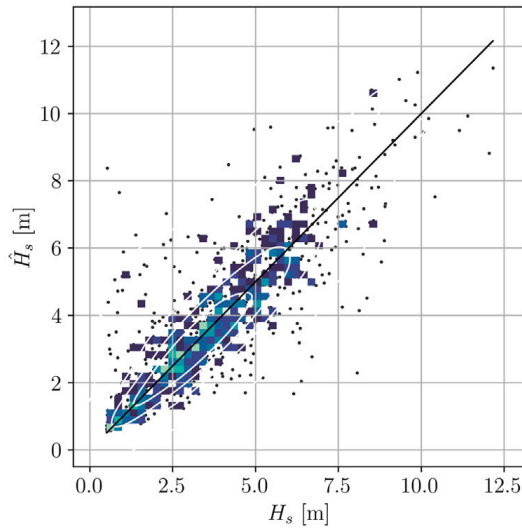
(d) Loss curves for both training and cross-validation set. Note that both axes are in logarithmic scale.

Fig. 15. Correlation plots of the target variables and the loss curves in the frequency domain. Note that overhat denotes the model predictions.

aleatoric or statistical uncertainty. As discussed in Chen et al. [26], wave radars exhibit significant inaccuracies under precipitation due to rain clutter.

4.3. General discussions

In the present work, it has been shown that the frequency domain models achieve higher out-of-sample accuracy, i.e. generalization capability, which is mainly due to the consideration of longer sample lengths as compared to the time domain. It is stressed that both the computational effort and model complexity in the time domain are proportional to the considered sample length. Conversely, both of these remain constant in the frequency domain. The computational time for one epoch – using the exact same hardware and model architecture (ResNet) – was 56 s in the time domain and in the frequency domain one epoch took 13 s. Specifically, the frequency domain requires approximately 77% less computing time, which is generally in accordance with the ratio of the numbers of elements of the two input feature spaces, i.e. 1500×6 in the time domain and 64×36 in the frequency domain, having obviously a direct impact on the trainable parameters of the model. Ultimately, it is concluded that the spectral approach is characterized by better accuracy, robustness and computational efficiency. This finding is somewhat consistent with other recent research on machine learning-driven sea state estimation that focuses on frameworks formulated in the frequency
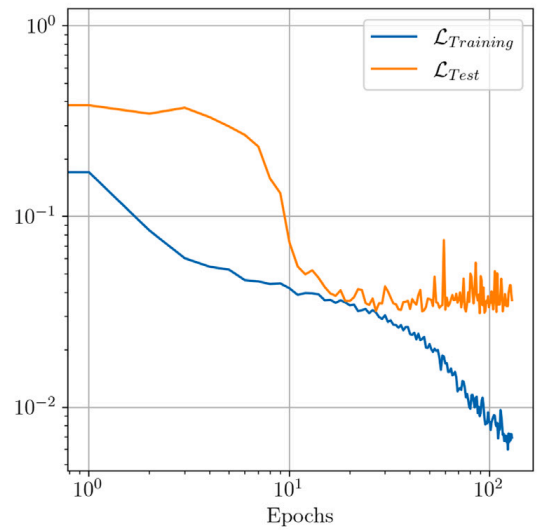
(a) Predictions of the significant wave height H_s for the validation set.
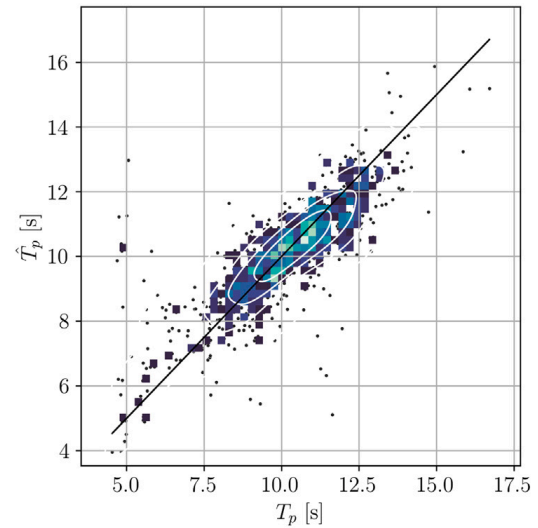
(b) Predictions of the peak period T_p for the validation set.
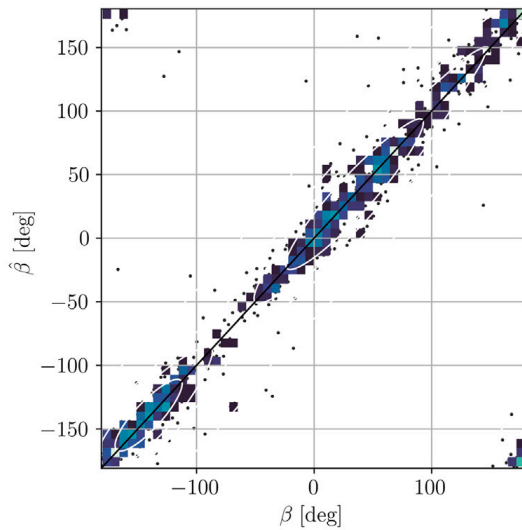
(c) Predictions of the mean encounter wave direction β for the validation set.

(d) Loss curves for both training and cross-validation set. Note that both axes are in logarithmic scale.

Fig. 16. Correlation plots of the target variables and the loss curves in the time domain. Note that overhat denotes the model predictions.

domain, e.g., [21,23]. In this context, it is also noteworthy that the generalization and extension towards methods for estimating the full directional wave spectrum is an important aspect of future work, such as in [22,23].

The main drawback of frequency domain approaches is considered to be the dependency on relatively long time windows, up to 25-30 min as found in this study, for a reasonable frequency resolution. Inherently, this can compromise results due to problems related to nonstationary conditions, emphasizing that even the sea state itself, as encountered from a sailing ship going maybe +20 knots, can vary because of spatial and temporal progression [30,49]. On the other hand, the literature provides promising studies that could be considered for mitigation of the drawback. For instance, Takami et al. [54] present the use of the prolate spheroidal wave functions for spectral analysis. Its main advantage over, say, Fast Fourier Transformation (FFT) is to provide a higher frequency resolution when applied on shorter time series, and with no need for manual choices with respect to smoothing. Alternatively, Cheng et al. [18] employ a hybrid approach, based on time–frequency representations of ship motions using spectrograms also indicating promising results.

Generally, sea state estimation is subject to substantial inherent uncertainty of both aleatoric and epistemic type. As such, the extension of this work towards uncertainty quantification and minimization of sea state estimation seems promising. On one hand, Mounet et al. [12] provide a methodology using a model-based sea state estimation approach while minimizing epistemic uncertainty by both reducing the uncertainty of RAOs by tuning, and spatial uncertainty by considering a network of ships at the same time. On the other hand, Han et al. [55] present an approach of visualizing and expressing the surrounding aleatoric or statistical

uncertainty in a hybrid approach with a model-based approach and Gaussian process regression. It is thought that the herein presented methodology may be extended towards both directions, however, the focus will be on the latter as aleatoric uncertainty was more prominent in the presented results. Thus, two individual approaches show potential for the transparent representation of uncertainty: (1) Following Mittendorf et al. [56], a quantile regression approach, i.e. training the neural network on quantile loss functions, seems worthwhile in a deterministic attempt. (2) Conversely, establishing a Bayesian or probabilistic model using the Monte Carlo dropout method as proposed by Gal and Ghahramani [57] shows great potential for future work, but may also be limited by increased computational effort. Lastly, Bitner-Gregersen et al. [58] conclude that the identification of the sea state with its surrounding uncertainty is and will be a vital research area in the maritime sector; and the changing climate contributes additional relevance to the topic.

5. Conclusions

Real time monitoring of the governing wave environment during a vessel's passage is directly linked to aspects of safety and economy. For this reason, the present contribution established a non-linear mapping from in-situ ship responses to prevalent sea state parameters using machine learning. The case ship was a midsize container vessel trading in Northern Atlantic. The ship was equipped with a wave radar from which data was collected during 1.5 years. Machine Learning frameworks were formulated in both the time and frequency domains using four different deep neural network architectures. The Inception model gave the highest performance and was applied in a multi-task learning setting. The model assessment suggested satisfactory performance of the multi-task frequency domain model on unseen validation data. The time domain models, on the other hand, exhibited substantial aleatoric (or statistical) uncertainty due to the considered short sample length. Lastly, the frequency domain model also showed superior characteristics in terms of computational effort and robustness. From a practical perspective, it was demonstrated that a machine learning methodology for sea state identification is applicable under realistic conditions, exclusively using response measurements from a cost effective MRU equipment.

Initially, it was thought that additional sensor recordings, such as relative wind speed and direction, could act as well-suited predictors for the sea state and for reducing uncertainty. Moreover, the consideration of propeller revolutions and rudder angle (or rather their variances) could be beneficial in a machine learning-based sea state estimation approach. These sensor readings were disregarded herein, since they are mostly instationary and filtering these samples reduces the sample size even more. Additionally, it stood out that a data acquisition period of 1.5 years only leads to approximately 5000 valid samples, i.e. complying to the enforced boundary conditions defined in Section 2. The major drawback of deep learning methods is definitely the required large amount of training data as well as limited applicability for extreme events. In turn, this motivates a parallel hybrid approach using a model-based method relying on RAOs and a machine learning model simultaneously, as investigated by Han et al. [55]. Model-based methods are not dependent on data availability, except for making the actual estimate, and may be considered as *self-supervised* in machine learning terms as they generate their directional wave spectrum estimate from an RAO database, but without any ground truth information. Therefore, it seems appealing to train a baseline model on results of model-based techniques and re-train an advanced model on e.g. wave radar data in a transfer learning approach, similar to the work of Düz et al. [20]. Hence, model-based and machine learning approaches are not seen as competitors, but as complementary — particularly when facing scarcity of sea state data. Moreover, the methodology is not dependent on wave radar data, but may also be applied to hindcast metocean data, e.g. from ERA 5 [59] being part of the EU Copernicus programme. Lastly, it is thought that the herein developed model may be considered as a baseline and could be adapted to other vessels using transfer learning in case of limited data availability. Possibly, the transfer learning procedure may be carried out in an incremental learning approach, i.e. the model is adapted in an online fashion, as more data becomes available. All of the above are interesting aspects of extending the present work.

CRediT authorship contribution statement

Malte Mittendorf: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualization, Writing – original draft. **Ulrik Dam Nielsen:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Resources, Funding acquisition. **Harry B. Bingham:** Conceptualization, Methodology, Writing – review & editing, Supervision, Resources, Project administration. **Gaute Storhaug:** Conceptualization, Data curation, Writing – review & editing, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Table 4

Metrics of predictions using the training dataset in both domains. The lowest error values are indicated in bold font.

		Frequency domain				Time domain			
		ResNet	Inception	MLSTM	CNN	ResNet	Inception	MLSTM	CNN
H_s	*RMSE*	**0.413**	0.425	0.649	1.033	1.353	**0.883**	2.568	1.196
	MAE	**0.255**	0.268	0.492	0.666	1.062	**0.631**	2.116	0.890
T_p	*RMSE*	0.393	**0.381**	0.506	0.893	1.022	**1.002**	1.653	1.276
	MAE	0.225	**0.216**	0.329	0.633	0.768	**0.755**	1.374	0.956
β	*RMSE*	5.106	**4.910**	5.486	10.84	13.49	**10.20**	31.99	21.38
	MAE	2.641	**2.372**	3.024	6.920	9.261	**7.130**	20.17	14.16

Table 5

Metrics of predictions using the training dataset in both domains using multi-task learning. The lowest error values are indicated in bold font.

		Frequency domain			Time domain		
		Inception	MTL	MTL+	Inception	MTL	MTL+
H_s	*RMSE*	0.425	0.341	**0.324**	0.883	0.945	**0.672**
	MAE	0.268	0.179	**0.171**	0.631	0.699	**0.461**
T_p	*RMSE*	0.381	**0.356**	0.371	1.002	0.845	**0.639**
	MAE	0.216	**0.170**	0.192	0.755	0.599	**0.463**
β	*RMSE*	**4.910**	5.031	5.082	10.20	12.75	**9.022**
	MAE	**2.372**	2.548	2.569	7.130	8.864	**5.888**
U_s	*RMSE*			0.167			0.378
	MAE			0.103			0.282

Acknowledgements

Appendix

Model performance on training data

See Table 4 for the results of the four initial models and Table 5 for data pertaining to the multi-task learning approach.

References

[1] Meister S, Taylordean S, Singer D. Predicting extreme parametric roll in container ships. In: Proc. of the 14th int. symp. on practical design of ships and other floating structures (PRADS'19), Yokohama, Japan. Springer; 2019, p. 703–18.

[2] Jalonen R, Tuominen R, Wahlström M. Remote and autonomous ships—the next steps: Safety and security in autonomous shipping—challenges for research and development. In: Rolls-Royce, Buckingham Gate, London: The advanced autonomous waterborne applications (AAWA). 2016, p. 56–73.

[3] Nielsen UD. Sea state estimation based on measurements of wave-induced ship responses (Dr. Techn. thesis), DTU Mechanical Engineering, Kgs. Lyngby, Denmark; 2018.

[4] Nielsen UD. A concise account of techniques available for shipboard sea state estimation. Ocean Eng 2017;129:352–62.

[5] Takekuma K, Takahashi T. On the evaluation of sea spectra based on measured ship motions. Trans West-Jpn Soc Nav Archit 1973;45:51–7.

[6] Iseki T, Ohtsu K. Bayesian estimation of directional wave spectra based on ship motions. Control Eng Pract 2000;8:215–9.

[7] Nielsen UD. Estimation of on-site directional wave spectra from measured ship responses. Mar Struct 2006;19(1):33–69.

[8] Pascoal R, Perera LP, Soares CG. Estimation of directional sea spectra from ship motions in sea trials. Ocean Eng 2017;132:126–37.

[9] Tannuri EA, Sparano JV, Simos AN, Cruz JJD. Estimating directional wave spectrum based on stationary ship motion measurements. Appl Ocean Res 2003;25:243–61.

[10] Nielsen UD, Stredulinsky DC. Sea state estimation from an advancing ship – A comparative study using sea trial data. Appl Ocean Res 2012;34:33–44.

[11] Nielsen UD, Mounet REG, Brodtkorb AH. Tuning of transfer functions for analysis of wave-ship interactions. Mar Struct 2021;79:103029.

[12] Mounet REG, Nielsen UD, Brodtkorb AH. Simultaneous sea state estimation and transfer function tuning using a network of dynamically positioned ships. Appl Ocean Res 2022. (Under review).

[13] Mittendorf M, Papanikolaou AD. Hydrodynamic hull form optimization of fast catamarans using surrogate models. Ship Technol Res 2021;68(1):14–26. http://dx.doi.org/10.1080/09377255.2020.1802165.

[14] Åvist P, Pyörre J. Modeling the impact of significant wave height and wave vector using an on-board attitude sensor network. In: Proc. of the 12th int. conf. on computer and IT applications in the maritime industries, Cortona; 2013.

[15] Han P, Li G, Skjong S, Wu B, Zhang H. Data-driven sea state estimation for vessels using multi-domain features from motion responses. In: 2021 IEEE international conference on robotics and automation (ICRA). 2021, p. 2120–6. http://dx.doi.org/10.1109/ICRA48506.2021.9561261.

[16] Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge: The MIT Press; 2016.

[17] Cheng X, Li G, Skulstad R, Chen S, Hildre HP, Zhang H. Modeling and analysis of motion data from dynamically positioned vessels for sea state estimation. In: Proc. of 2019 int. conf. on robotics and automation (ICRA). IEEE; 2019, p. 6644–50.

[18] Cheng X, Li G, Skulstad R, Zhang H, Chen S. SpectralSeaNet: Spectrogram and convolutional network-based sea state estimation. In: Proc. of IECON 2020 the 46th annual conference of the IEEE industrial electronics society; 2020.

[19] Cheng X, Li G, Li G, Chen S, Hildre HP, Zhang H. A novel densely connected convolutional neural network for sea-state estimation using ship motion data. IEEE Trans Instrum Meas 2020;69(9):5984–93.

[20] Düz B, Mak B, Hageman R, Grasso N. Real time estimation of local wave characteristics from ship motions using artificial neural networks. In: Proc. 14th int. symp. practical design of ships and other floating structures (PRADS'19), Yokohama, Japan; 2019.

[21] Kawai T, Kawamura Y, Okada T, Mitsuyuki T, Chen X. Sea state estimation using monitoring data by convolutional neural network. J Mar Sci Technol 2021;26:947–62.

[22] Scholcz TP, Mak B. Ship as a wave buoy – Estimating full directional wave spectra from in-service ship motion measurements using deep learning. In: Proc. of 39th int. conf. on ocean, offshore and arctic engineering (OMAE 2020), Fort Lauderdale, USA; 2020.

[23] Han P, Li G, Skjong S, Zhang H. Directional wave spectrum estimation with ship motion responses using adversarial networks. Mar Struct 2022;83:103159.

[24] Storhaug G, Moe E, Piedras Lopes TA. Whipping measurements onboard a midsize container vessel operating in the north atlantic. In: Marintec China proceedings (RINA, CMP, and SNAME). 2007.

[25] Gangeskar R. An algorithm for estimation of wave height from shadowing in X-Band radar sea surface images. IEEE Trans Geosci Remote Sens 2014;52(6):3373–81.

[26] Chen X, Huang W, Zhao C, Tian Y. Rain detection from X-Band marine radar images: A support vector machine-based approach. IEEE Trans Geosci Remote Sens 2020;58(3):2115–23.

[27] Miros. Wavex® v5.6 - the x-band radar-based wave and current monitoring system. 2017.

[28] Barstow S, Bidlot J, Caires S, Donelan M, Drennan W, Dupuis H, Graber H, Green J, Grønlie Ø, Guérin C, Gurgel K-W, Günther H, Hauser D, Hayes K, Hessner K, Hoja D, Icard D, Kahma K, Keller W, Wyatt L. Measuring and analysing the directional spectrum of ocean waves. 2005, COST 714.

[29] Longuet-Higgins MS, Cartwright DE, Smith ND. Observations of the directional spectrum of sea waves using motions of a floating buoy. In: Ocean wave spectra. New York: Prentice Hall; 1963, p. 111–36.

[30] Nielsen UD, Ikonomakis A. Wave conditions encountered by ships—A report from a larger shipping company based on ERA5. Ocean Eng 2021;237:109584.

[31] Söding H. Global seaway statistics. Schriftenreihe Schiffbau, report no. 610, Hamburg: Technical University Hamburg-Harburg; 2001.

[32] St. Denis M, Pierson WJ. On the motion of ships in confused seas. SNAME Transactions; 1953.

[33] Mittendorf M, Nielsen UD, Bingham HB. Data-driven prediction of added-wave resistance on ships in oblique waves - a comparison between tree-based ensemble methods and artificial neural networks. J Appl Ocean Res 2022;118(102964).

[34] Mittendorf M, Nielsen UD, Bingham HB. The prediction of sea state parameters by deep learning techniques using ship motion data. In: Proc. of 7th world maritime technology conference (WMTC'22), Copenhagen, Denmark; 2022.

[35] Welch P. The use of fast Fourier transform for the estimation of power spectra: a method based ontime averaging over short, modified periodograms. IEEE Trans Audio Electroacoust 1967;15:70–3.

[36] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: Bach F, Blei D, editors. Proc. of the 32nd int. conf. on machine learning. Proceedings of machine learning research, vol. 37, Lille, France: PMLR; 2015, p. 448–56.

[37] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80.

[38] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proc. 2016 IEEE conference on computer vision and pattern recognition (CVPR); 2016, p. 770-8.

[39] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan V, Rabinovich A. Going deeper with convolutions. In: Proc. of the IEEE conference on computer vision and pattern recognition; 2015, p. 1–9.

[40] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. 2012, NIPS 2012: 1106-1114.

[41] Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proc. of the 31st AAAI conference on artificial intelligence; 2017, p. 4278–84.

[42] Ismail Fawaz H, Lucas B, Forestier G, Pelletier C, Schmidt DF, Weber J, Webb GI, Idoumghar L, Muller P-A, cois Petitjean F. InceptionTime: Finding AlexNet for time series classification. Data Min Knowl Disc 2020;34:1936–62.

[43] Caruana R. Multi-task learning. Mach Learn 1997;28:41–75.

[44] Ruder S. An overview of multi-task learning in deep neural networks. 2017, arXiv:1706.05098v1.

[45] Caruana R, de Sa V. Promoting poor features to supervisors: Some inputs work better as outputs. In: Advances in neural information processing systems 9: proceedings of the 1996 conference, Vol. 9. 1997.

[46] Kingma DP, Ba J. Adam: a method for stochastic optimization. In: Proc. of 3rd int. conf. on learning representations (ICLR '15), San Diego; 2015.

[47] Mak B, Düz B. Ship as a wave buoy – estimating relative wave direction from in-service ship motion measurements using machine learning. In: Proc. 38th int. conf. on ocean, offshore & arctic engineering OMAE 2019, Glasgow, Scotland; 2019.

[48] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X. TensorFlow: Large-scale machine learning on heterogeneous systems. 2015, URL: https://www.tensorflow.org/, software available from tensorflow.org.

[49] Nielsen UD. Spatio-temporal variation in sea state parameters along virtual ship route paths. J Oper Oceanogr 2021.

[50] Montazeri N, Nielsen UD, Jensen JJ. Selection of the optimum combination of responses for wave buoy analogy - an approach based on local sensitivity analysis. In: Proc. of 13th PRADS, Copenhagen, Denmark; 2016.

[51] Andersen IMV, Storhaug G. Dynamic selection of ship responses for estimation of on-site directional wave spectra. In: Proc. 31st int. conf. on ocean, offshore and arctic engineering (OMAE'12), Rio de Janeiro, Brazil; 2012.

[52] Brodtkorb AH, Nielsen UD, Sørensen AJ. Online wave estimation using vessel motion measurements. In: Proc. of 11th IFAC conf. CAMS, Opatija, Croatia; 2018.

[53] Iseki T, Nielsen UD. Study on short-term variability of ship responses in waves. J Jpn Inst Navig 2015;132:51–7.

[54] Takami T, Nielsen UD, Jensen JJ. Estimation of autocorrelation function and spectrum density of wave-induced responses using prolate spheroidal wave functions. J Mar Sci Technol 2020;26.

[55] Han P, Li G, Cheng X, Skjong S, Zhang H. An uncertainty-aware hybrid approach for sea state estimation using ship motion responses. IEEE Trans Ind Inf 2022;18(2):891–900.

[56] Mittendorf M, Nielsen UD, Bingham HB, Liu S. Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading. Ocean Eng 2022;251(111040).

[57] Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In: Balcan MF, Weinberger KQ, editors. Proc. of the 33rd int. conf. on machine learning. Proceedings of machine learning research, vol. 48, New York, New York, USA: PMLR; 2016, p. 1050–9.

[58] Bitner-Gregersen EM, Bhattacharya SK, Chatjigeorgiou IK, Eames I, Ellermann K, Ewans K, Hermanski G, Johnson MC, Ma N, Maisondieu C, Nilva A, Rychlik I, Waseda T. Recent developments of ocean environmental description with focus on uncertainties. Ocean Eng 2014;86:26–46.

[59] Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut J-N. Era5 hourly data on single levels from 1979 to present. 2018, Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 24-Feb-2022).

Chapter 5

Estimation of Added Resistance in Irregular Waves using In-service Data

Digitalization and imposed regulations are decisive driving factors for increasing data acquisition onboard ships. On the one hand, collecting operational data became mandatory due to several rules stipulated by IMO and the EU. On the other hand, utilizing performance data for overall transparency and optimization of fleet operations is another motivation for shipping companies to equip their fleet with an array of sensors. Aldous *et al.* [2] show empirically that the inherent uncertainty of in-situ sensor data is significantly lower than manually collected noon reports. However, other complexities arise when dealing with high-frequency data, such as handling and filtering large amounts of data. In this respect, the work of Dalheim and Steen [25] provides a very detailed preprocessing methodology for, e.g. filtering for steady conditions. In addition, several sensors are known for their proneness to error and sensor drift, such as the Doppler velocity log for measuring Speed Through Water (STW), Ikonomakis *et al.* [50]. Still, the installation of sensor frameworks on ships allows for more detailed full scale analyses in comparison to sea trials, which are rather limited in temporal coverage and do not consider the vessel in actual operation with, e.g. the experience of harsh weather conditions. Therefore, it is the goal of this chapter to utilize fleet performance data for an investigation towards the mean added resistance in a seaway.

5.1   Introduction

For reliable performance monitoring and routing optimization, an accurate estimate of the mean added resistance $\bar{R}_{AW}$ is indispensable. For instance, when optimizing sailing routes for wind-assisted ships, a sufficient method for estimating added wave resistance is essential since favorable aerodynamic conditions may correspond to unfavorable hydrodynamic conditions due to the coupling between prevailing wind and wave conditions. The assessment of the minimum propulsion power requirement for blunt-type ships and the wave correction as part of sea trial analyses are additional application fields that require accurate estimates of R_{AW} transfer function *and* sea state. Still, most numerical studies disregard the application in irregular waves entirely, and a majority of studies towards $\bar{R}_{AW}$ employ unimodal and unidirectional wave spectra, e.g. Nabergoj and Prpić-Oršić [92] or Perez Arribas [103].

When in fact, it is well-known that a real *short-crested* seaway comprises several partitions from wind and swell. In [J4], it is found that 78% of all observed sea states were, in fact, bimodal assuming hindcast wave data as ground truth. Liu *et al.* [80] list three requirements for estimation methods for practical application in irregular waves: (1) Applicability to arbitrary wave headings, (2) coverage of a wide range of operating conditions including advance speed, draft, and trim, and (3) a sufficient accuracy in shorter waves.

The calculation of $\bar{R}_{AW}$ relies fundamentally on Maruo's formulation given in Eq. 2.6, which is based on linear spectral analysis. For demonstration, several methods presented in Fig. 2.4 are applied in an irregular wave scenario for the same case ship – the KVLCC2. In Fig. 5.1, $\bar{R}_{AW}$ is normalized by the calm water resistance $R_{T,calm}$, as taken from experiments conducted by Yu *et al.* [133], and shown for all mean wave heading angles β_0. The considered sea state is characterized by only wind waves (i.e. unimodal) and approximated by a Pierson and Moskowitz [104] spectrum with respect to $H_s = 4.0$ m and $T_p = 11.0$ s. The Kreitner [70] and Boom *et al.* [14] methods are included due to their frequent use in practice for, e.g. performance monitoring or sea trial analysis. The method by Boom *et al.* [14] is calculated for long-crested waves, and the Kreitner [70] method may be outside of its definition range since $H_s > 2$ m. However, the specified peak wave period reflects a relative wave length of $\lambda/L_{pp} = 0.59$ for the KVLCC2, and hence the regime can still be considered diffraction-dominated, for which the Kreitner [70] is deemed valid. The remaining four methods are valid for arbitrary wave heading angles and are subsequently applied in short-crested waves, which is facilitated by overlaying the unidirectional spectrum with a $\cos^{2s}$ spreading function.
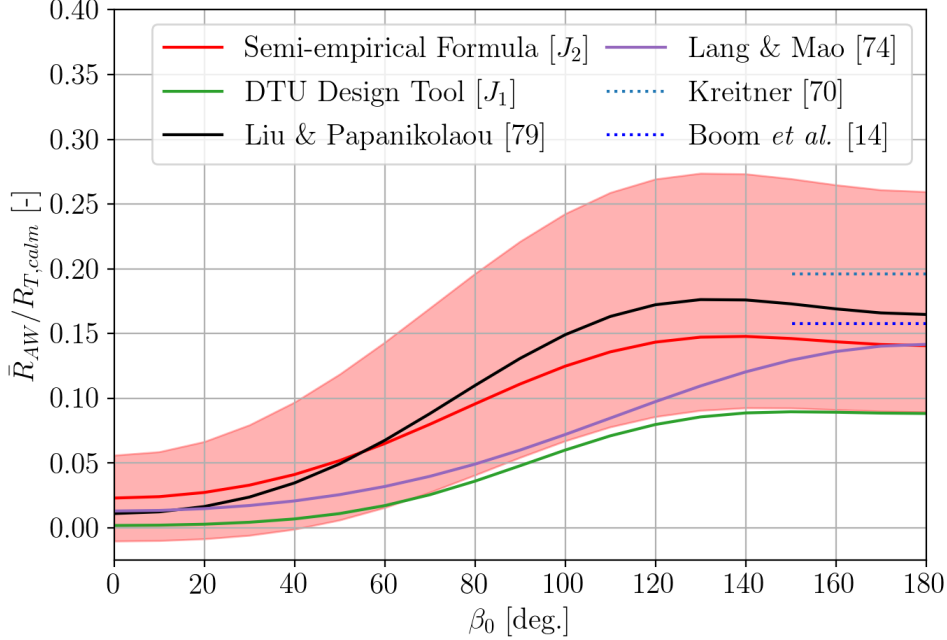

Figure 5.1: The relative mean added resistance in irregular short-crested waves for the KVLCC2 at $Fn = 0.142$ as a function of the mean wave heading β_0 using a unimodal Pierson and Moskowitz [104] spectrum for $H_s = 4.0$ m and $T_p = 11.0$ s, i.e. Beaufort 6.

Given Fig. 5.1, it is observed that $\bar{R}_{AW}$ can take up almost 20% of the calm water resis-

tance in a Beaufort 6 sea state at design speed, and it is emphasized that this fraction may increase with a decrease in forward speed, e.g. in a slow steaming regime. As anticipated, the individual methods show a clear dependency on β_0; however, only the two related methods by $[J_2]$ and Liu and Papanikolaou [79] show that $\bar{R}_{AW}$ is largest in bow oblique waves, as opposed to the DTU design tool and Lang and Mao [74]. Moreover, the DTU design tool shows good qualitative agreement with other methods but generally non-conservative estimates since it nearly coincides with the lower uncertainty bound of the $[J_2]$ method. In contrast, both Kreitner [70] and Boom *et al.* methods exhibit relatively more conservative predictions of $\bar{R}_{AW}$, but still fall into the prediction interval of $[J_2]$. As a side note, a minor increase in added resistance in the proximity of stern quartering waves has been reported in several studies, e.g. Blok [12]. Valanto and Hong [128] assume that the increase of the yaw and roll in these wave headings may be causal for this. However, based on the experimental database in $[J_2]$, it is not possible to draw any firm conclusions in this regard, indicating the need for more in-depth studies – also for possible extensions of semi-empirical methods.

Unlike most other studies and Fig. 5.1, the present chapter focuses on estimating the in-situ added resistance using a data fusion approach of in-service ship monitoring and historical wave data. In 1936, Kempf [67] conducted thorough experiments for correlating both full and model scale results. The study considers one case ship with extensive measurement equipment recording, e.g. ship motions, propeller revolutions, and the sea state (visually). Overall, reasonable accuracy was found between model and full scale, and a speed loss of 40% in a Beaufort 8 sea state was reported. Approximately 80 years later, Vitali *et al.* [130] present a study combining AIS (Automatic Identification System) and wave hindcast data for determining the speed loss of a large fleet of container ships with respect their original design speeds. Moreover, they compare their estimates to results obtained from the formula of Kwon [73] and find that this approach is generally non-conservative. A possible explanation could be the missing differentiation between involuntary and voluntary speed loss the this study. Nielsen *et al.* [97] examine the added resistance in an indirect approach using high-frequency data from a container vessel. The method builds on the measured shaft power and empirical resistance decomposition for calculating the added resistance as the surplus to the theoretical shaft power in calm weather. Generally, satisfactory results have been obtained, but significant variance and presumably corrupt data in some cases in following waves stood out.

The overarching goal of this chapter and $[J_4]$ is to apply the indirect approach of Nielsen *et al.* [97] in parallel to the spectral formulation given in Eq. 2.6. For this purpose, a large-scale fleet performance dataset from a major shipping company will be utilized, similar to one used in Ikonomakis *et al.* [50]. Moreover, wave hindcast data from the ERA5 database will be employed. In fact, the modeling of the prevailing wave environment receives a large focus and surpasses current state-of-the-art publications. For instance, Lang and Mao [74] consider only a unimodal directional wave spectrum and Liu *et al.* [80] consider only wind waves as short-crested, but swell waves as unidirectional. Herein, both partitions will be regarded as short-crested, and the peak enhancement factor (as part of the JONSWAP spectrum) will be variable for a detailed representation of the wave environment in regions with limited fetch, i.e. coastal areas. Moreover, the method developed in $[J_2]$ will be used for calculating the transfer function, and the empirical Hollenbach [47] method is used for

estimating the calm water resistance. By using in-service data, it is the goal to investigate the mean added wave resistance in actual conditions without any simplifications. The scale of this study in itself is, in fact, unprecedented and thus resembles the primary novelty of $[J_4]$.

5.2   Discussion

Tsujimoto *et al.* [123] point out generally a poor agreement between theoretical/numerical and experimental added resistance results for slender container vessels in short regular waves. This is primarily owed to non-linear effects induced by the hull shape, such as wave breaking due to pronounced bow flare or the impact of a protruding bulb. Nabergoj and Prpić-Oršić [92] confirm these difficulties using five different approaches for determining the R_{AW} transfer function for a slender ferry and a standard design spectrum for the application in irregular waves. In fact, they found a spread within the individual methods by a factor of 3.5 in higher sea states, and the Jinkine and Ferdinande [60] approach gave the most conservative estimates. When it comes to in-service analyses of $\bar{R}_{AW}$, the degree of variance and bias can be of an even larger extent, e.g. Shigunov [111].

   In the context of $[J_4]$, the considered fleet experienced only moderate sea states due to good seamanship and weather routing. Still, there were indications for deviations in higher sea states in the case of the spectral method in comparison to the indirect estimate. Having said this, a considerable degree of uncertainty can be attributed to the sea state data obtained via the ERA5 database, as seen in the previous chapter. Moreover, applying the wave buoy analogy for deriving sea state parameters was unfeasible since the wave-induced responses (e.g. motions and hull girder strain) had not been captured during the observation period. Another possible reason for the deviations could be that the method by Faltinsen *et al.* [32] (as part of the $[J_2]$ method) may be stretched towards its application limits since it is derived for vertical walls, i.e. no existing flare, and lower forward speeds ($Fn < 0.2$). Therefore, the implementation of the NMRI correction for bow wave reflection could be a potential aspect for extending the method in $[J_2]$ since Tsujimoto *et al.* [123] showed reasonable agreement for a container ship. However, there is generally a significant knowledge gap and lower result quality in very short waves – regardless of the approach. As mentioned in Ch. 2, Sprenger *et al.* [115] found only weak confirmation of the second-order relationship between R_{AW} and wave amplitude. In this light, Kuroda and Takagi [72] conducted added resistance model tests in irregular waves and incorporated a 4th-order term for considering the effect of the hull shape above the waterline. In the case of a VLCC, good agreement was found during model tests using a higher-order expression; however, more systematic studies are needed for confirmation.

   For assessing the overall uncertainty of the employed transfer function, the calculation according to Eq. 2.6 is carried out for all three estimates given by the method of $[J_2]$, i.e. the mean and the two uncertainty bounds. It is noted that this reflects an inconsistent approach as it only takes the uncertainty of the transfer function into account and disregards the possible uncertainty pertaining to the sea state. The results are given in Fig. 5.2a for the same case vessel class examined in Fig. 1.3, i.e. the 15,000 TEU class. Expectedly, the uncertainty – expressed by the prediction interval – is dependent on the magnitude of $\bar{R}_{AW}$

and, therefore, negligible in lower sea states. Given Fig. 5.2a, there is an indication that the scatter for $H_s > 3.0$ m is closer to the lower uncertainty bound, but the general lack of samples in higher sea states impedes drawing any conclusions. Overall, it can be seen that the transfer function-related uncertainty is only of minor importance in view of the variance of the relative empirical added resistance obtained through the indirect methodology.
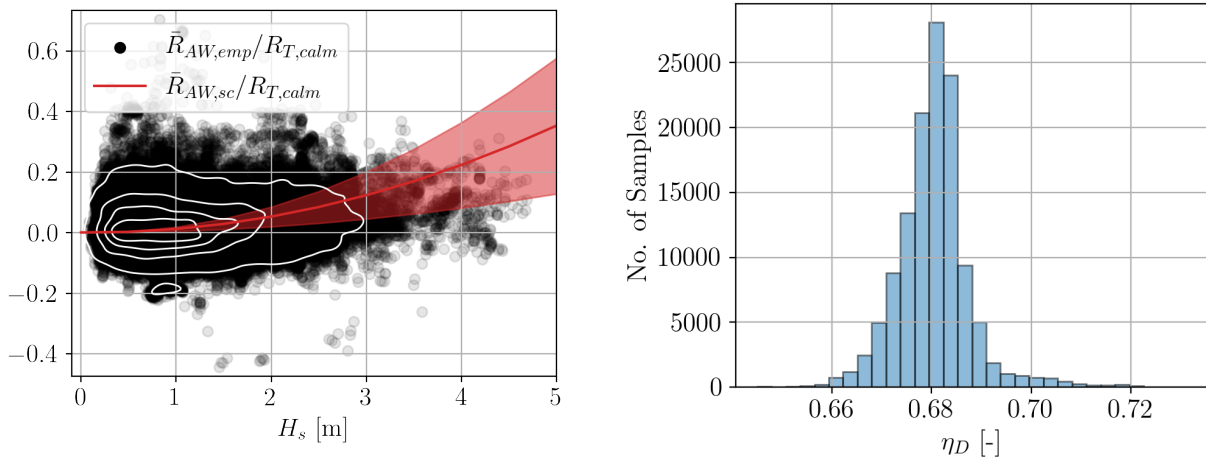
(a) The relative empirical and theoretical added resistance where the latter is approximated by three second order polynomials.

(b) Determination of η_D following the method of Kristensen and Lützen [71] assuming a Wageningen-B series propeller.

Figure 5.2: Two individual studies on (a) uncertainty pertaining to the R_{AW} transfer function and (b) estimating the propulsive efficiency η_D. Data from eight sister vessels of the 15,550 TEU vessel class is taken.

Regarding the indirect estimation of added resistance, several uncertainty sources have already been pointed out in $[J_4]$, and a considerable contribution can be ascribed to using an empirical formula for the dominant calm water resistance. Empirical approaches are dedicated to estimating $R_{T,calm}$ for design speed and draft. Their extrapolation capability may be, in fact, limited, and hence it seems worthwhile to estimate a quasi-calm water speed-power baseline by a piece-wise regression methodology following Berthelsen and Nielsen [9]. Another contribution to the observed uncertainty of the empirical added resistance in Fig. 5.2a may result from disregarding hull fouling. However, considering the application of special fouling-release paints, lower idling periods, higher forward speeds, and great efforts towards hull/propeller cleaning by the shipping company, it is believed that the disregard of fouling in $[J_4]$ is of little relevance.

The assumption of the propulsive efficiency as 70%, as suggested by ISO 19030, can be considered a relatively crude simplification in $[J_4]$. For this reason, the approach for determining the propulsive efficiency η_D by Kristensen and Lützen [71] is herein applied to the 15,000 TEU vessel class data. The method builds on the relationship between the thrust loading coefficient and the open-water propeller efficiency of the Wageningen-B propeller series. In Fig. 5.2b, a histogram of the calculated η_D values is presented, and the distribution arranges around 0.68. Considering that the optimized propeller will exhibit slightly higher

efficiency than a standard Wageningen-B propeller, the used value of 0.7 can be regarded as somewhat reasonable. However, applying this calculation procedure for η_D requires detailed information, such as the propulsive coefficients, which are determined according to Guldhammer and Harvald [42] as speed, draft, and seaway independent. Additionally, for calculating the required propeller thrust, an estimate of $\bar{R}_{AW}$ is required contradicting the actual purpose of the indirect methodology by Nielsen *et al.* [97]. In fact, the relative residuals of theoretical and empirical added resistance showed a reduced variance by 10% when using the corresponding η_D values from Fig. 5.2b, but this is mostly for instances in higher sea states, which is then owed to the inclusion of the spectral estimate when determining the required propeller thrust. In fact, in the work of Nielsen *et al.* [97], the empirical resistance and propulsion model was even more detailed, including the method by Fujiwara *et al.* [34] for wind resistance and the actual open water propeller diagram for the respective case ship. However, the result quality is relatively similar to that of [J_4]. Moreover, it was observed that the added wind resistance was generally higher in magnitude than the added wave resistance due to the low sea states combined with large windage areas of container ships. Apart from the measurement uncertainty of wind speed and direction, as shown in [J_4], the uncertainty regarding actual loading conditions and arrangement of containers on deck can be considered a significant contribution to the variance of the indirect estimation approach of added resistance. The random loading pattern induces local separation bubbles and can, in fact, surpass the wind resistance for a fully loaded condition, Andersen [5].

It is concluded that the added resistance is of lower importance for the performance monitoring of container vessels since it resembles only a minor fraction of the total ship resistance. However, in the case of blunt-type ships and for routing optimization regarding all vessel types, $\bar{R}_{AW}$ has to be estimated with high accuracy. As concluded by Bertram [11], the added *power* due to waves is of even greater practical relevance because it also includes the propulsive and engine-related losses in waves. In addition, maneuvering aspects, such as steering forces, have to be considered – especially in oblique waves with an increase in yaw and sway. This motivates a unified approach combining both seakeeping and maneuvering, as shown by Skejic and Faltinsen [114]. Ultimately, Lang *et al.* [75] compare the results of an empirical framework to machine learning methods for the prediction of shaft power in service conditions and found higher accuracy for the machine learning approach. Hence, the next chapter will focus on predicting a ship's propulsive power using neural networks and eventually extracting the biofouling-related power increase.

$[J_4]$ Assessment of added resistance estimates based on monitoring data from a fleet of container vessels

The paper entitled "Assessment of added resistance estimates based on monitoring data from a fleet of container vessels" is published in the Journal of Ocean Engineering as:

Assessment of added resistance estimates based on monitoring data from a fleet of container vessels

Malte Mittendorf [a,*], Ulrik Dam Nielsen [a], Harry B. Bingham [a], Jesper Dietz [b]

[a] Department of Civil and Mechanical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
[b] A.P. Møller-Mærsk, Energy Efficiency & Decarbonization, DK-1263 Copenhagen, Denmark

ARTICLE INFO

ABSTRACT

A practical estimation methodology of the mean added resistance in irregular waves is shown, and the present paper provides statistical analyses of estimates for ships in actual conditions. The study merges telemetry data of more than 200 in-service container vessels with ocean re-analysis data from ERA5. Theoretical estimates relying on spectral calculations of added resistance are made for both long- and short-crested waves and are based on a combination of a parametric expression for the wave spectrum and a semi-empirical formula for the added resistance transfer function. The theoretical estimates are compared to predictions from an indirect calculation of added resistance relying on shaft power measurements and empirical estimates of the remaining resistance components. Overall, the comparison reveals a bias in bow oblique waves and higher sea states of the spectral estimates as well as the large variance of the empirically derived predictions — particularly in beam-to-following waves. One of the study's main findings, confirming previous studies but based on a much larger dataset than in earlier similar studies, is that added resistance assessment based on in-service data is complex due to significant associated uncertainties.

1. Introduction

Seaworthiness is considered the ultimate criterion in ship design. However, standard practice is to optimize the hull shape for calm water conditions with a service allowance of around 15% to account for the added resistance in actual operating conditions, Shigunov (2017). However, the added resistance is not only vital in ship design but even more so in ship operation: The impact of added resistance on ship safety, economy, and overall life cycle is unquestionable. Therefore, vessel performance analysis and voyage optimization need reliable and precise information about this second order force in actual ship operations. In this line, Tsujimoto et al. (2008) conclude that estimates of added resistance due to waves are essential for correctly evaluating a vessel's emissions and thus its carbon footprint. With tightening rules for the efficient design and operation of ships, the relevance of added resistance is amplified. Recently, IMO (2021) stipulated the EEXI (Energy Efficiency eXisting ship Index) and CII (Carbon Intensity Indicator) regulations, which, in fact, may lead to a new regime of *involuntary* slow steaming due to Shaft Power Limitations (ShaPoLi) in case of non-compliance. In parallel, rising fuel prices and port congestion may be driving factors for *voluntary* slow steaming. Hence, added resistance could be of increasing relevance in the years to come due to reduced operating speeds. It is stressed that added resistance is

of higher relative magnitude under lower forward speeds. Over the long term, it is thought that climate change will lead to a greater interest in the influence of wind and waves on ship operations and, consequently, the fidelity of weather routing methods due to the rising likelihood of extreme weather events, as pointed out by Bitner-Gregersen et al. (2014). Moreover, Bertram (2016) concludes that knowledge about the prevailing sea state (or rather the lack of it) is an "Achilles heel" for the determination of the added resistance in actual, in-service conditions. Hence, modeling the wave environment is a key feature of the present work, which relies on re-analysis data retrieved from ERA5 (Hersbach et al., 2018).

1.1. Literature review

Tsujimoto et al. (2008) lay out a clear methodology for assessing added resistance in short-crested waves and the associated speed loss for a container vessel. For this reason, they propose the NMRI (National Maritime Research Institute) formula for the determination of the transfer function and the ITTC (International Towing Tank Conference) spectrum combined with the cosine power-type spreading function is used for wave modeling. Tsujimoto et al. (2008) also focus on the accurate calculation of additional resistances due to wind and steering

---

* Corresponding author.
E-mail address: mamit@dtu.dk (M. Mittendorf).

impacts. Shigunov (2017) estimates the added power on ships in a seaway by utilizing a Rankine panel method to calculate the transfer functions of wave drift forces/moments in 3 degrees of freedom. In addition, the consideration of propulsive and engine-related characteristics is a crucial aspect of his work. The calculation procedure for a container carrier is verified using noon report data of two sister vessels. Kim et al. (2017) present a study combining CFD (Computational Fluid Dynamics) results of the added resistance transfer function for the S175 container vessel and the estimation of the speed loss in irregular waves. The latter has been accomplished by applying a unidirectional Pierson Moskowitz spectrum. Unlike the already presented simulation-based studies, the increasing availability of in-service vessel monitoring data allows for more in-depth analyses of added resistance in actual conditions. For instance, Lakshmynarayanana and Hudson (2017) derive the calm water shaft power from high-frequency monitoring data of three sister vessels by strictly filtering for environmental conditions. Their analysis of the added power is then carried out using the initially derived calm water power as a baseline and under the same draft conditions. The results reveal considerable uncertainty in the derived added power and a weak dependency on the sea state. Nielsen et al. (2019) utilize auto-logged high-frequency data of a container vessel in worldwide service and determine the added resistance in an indirect approach. Similarly to the former study, the added resistance is considered the surplus to the empirically derived calm water shaft power compared to the actual measured shaft power. In general, reasonable results have been obtained; however, significant variance and possibly erroneous data in following waves were observed. Furthermore, Dalheim and Steen (2020a) evaluate the in-situ added resistance of an offshore supply vessel operating in the North Sea by considering model test data in both calm water and irregular waves as references. The analysis is based on one year of high-frequency data, and the added resistance has been predicted as non-conservative according to simulated data based on strip theory results. Lang and Mao (2020) propose a semi-empirical method for the prediction of the added resistance transfer function in head waves. In parallel, they apply this approach to high-frequency monitoring data of two separate vessels and use the JONSWAP spectrum together with hindcast data from ERA5 for wave modeling. Similar to Dalheim and Steen (2020a), they found an underestimation and subsequently derived a wave height correction factor determined as $\sqrt[3]{H_s}$. Taskar and Andersen (2021) assess several known procedures for predicting the added resistance transfer function and verify the obtained results in irregular waves with noon reported and high-frequency data of an oil and a gas tanker. The JONSWAP spectrum was used for modeling the wave environment, and bimodal seaways were considered by means of ERA5 sea state parameters. It stood out that significant variance was visible in the residuals – in higher sea states in particular – despite the detailed data pre-processing methodology.

Overall, considerable uncertainty is a common and dominant issue among all the above-mentioned publications. Both large variance *and* bias are observed making the isolated analysis of in-service added resistance challenging. Prpić-Oršić et al. (2018) report about several possible sources of uncertainty ranging from erroneous/biased sensor data to limited knowledge of the prevailing sea state. Nonetheless, it is emphasized that with the deployment of continuous monitoring infrastructure onboard vessels, the degree of measurement uncertainty (including human error) reduces drastically compared to manually recorded noon reports (Aldous et al., 2015; Christensen et al., 2018). In addition, the quality of met-ocean data, particularly wind and waves, has improved significantly compared to earlier analyses, often relying on visual observations (de Hauteclocque et al., 2020). Despite improving data quality, the isolation of added resistance in in-service data is still subject to substantial uncertainty.

1.2. Motivation and objective

Given the presented literature, it stands out that the majority of studies is limited to one vessel class and, in some cases, even to only a few voyages. Hence, a significant contribution of the present paper lies in the paper's consideration of more than two years of continuous high-frequency monitoring data from 32 vessel classes with 228 container vessels. However, no detailed hull shape or propeller information has been made available, which may be seen as a drawback in terms of accuracy, but this is considered the default case in general vessel performance monitoring by the respective industry. The proposed practical estimation methodology is based on a semi-empirical framework, and only information on the main particulars of the vessels and sea state parameters is required. The approach presented by Mittendorf et al. (2022b) is chosen for the determination of the added resistance transfer function, and the modeling of the ambient wave scenario follows from a parametric directional wave spectrum shaped with parameters obtained from re-analysis data for wind and swell using the public domain ERA5 service.

The paper consists of two parts: (1) Initially, the estimation of the *theoretical* added resistance in short- and long-crested waves is performed in parallel and compared. In general seakeeping calculations, long-crested waves are typically assumed due to reduced complexity and conservative results (Lloyd, 1989). Thus, it will be studied whether the same holds for the higher order force of added resistance. (2) The second part of the paper is about the comparison of spectral results to empirical predictions, which are derived from in-service shaft power measurements. This *empirical* added resistance is determined by an indirect methodology, as presented in Nielsen et al. (2019), i.e. by resistance decomposition using empirical methods. Due to the large windage area of container vessels, the wind resistance is considered through experimental data from Andersen (2013) according to ISO 15016 (ISO, 2015). In addition, the sea margins of the individual vessel classes will be presented and analyzed to link ship operation to design. In fact, the respective sea margins of theoretical and empirical methods are shown and discussed. In a dedicated section, different uncertainty sources will be pointed out and partly set into the context of ISO 19030 (ISO, 2016). Ultimately, the present study aims to provide a practical estimation methodology of the mean added resistance in irregular waves based on publicly available wave climate data and an established semi-empirical method for predicting the associated quadratic transfer function.

1.3. Composition

In Section 2, both the fleet performance data and metocean hindcast data will be introduced and analyzed. The data pre-processing and filtering makes up the majority of this section. Section 3 has its focus on the spectral estimation methodology and the in-direct empirical calculation procedure. The results are presented and discussed in Section 4, while Section 5 provides dedicated remarks about the inherent uncertainty. Finally, Section 6 summarizes the paper and discusses possibilities for extending work.

2. Data analysis and processing

2.1. General overview of data

The extensive database of fleet telemetry data includes 32 vessel classes with 228 vessels in total and has been acquired over two years between January 2019 and March 2021. As a side note, the time of the measurement campaign is characterized by severe disruptions in the global supply chains due to the COVID-19 pandemic. An overview of the fleet's two major characteristics, i.e. length between perpendiculars L_{pp} and TEU (Twenty-foot Equivalent Unit) capacity, can be seen in Fig. 1. In view of the regression analyses taken from Kristensen (2015),
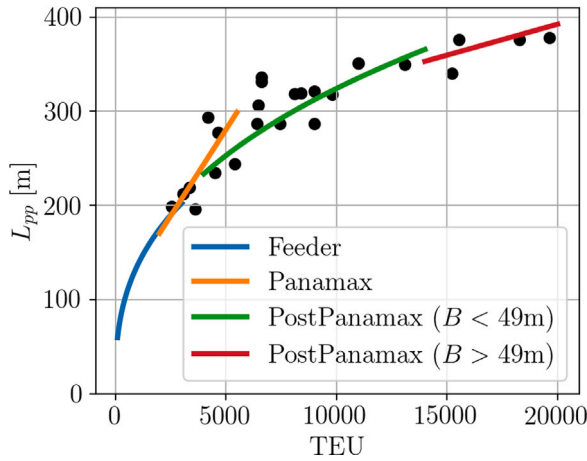
Fig. 1. Relationship between ship length (L_{pp}) and capacity in terms of TEU considering the individual vessel classes. It is noted that the regression curves are taken from Kristensen (2015) and that B denotes the beam of the vessel.

it is observed that the considered vessels resemble the world's fleet satisfactorily for both Panamax and PostPanamax vessels. The main particulars and other characteristics of the individual vessel classes have been compiled from class information and in-house data sheets. On a separate note, the present fleet shows varying geometric characteristics, such as vertical stems, twin-screw arrangements, and mounted cranes for cargo handling. All of this is useful in a potential discussion about the relevance of the present study and the generalization of its results with respect to ships of other fleets.

In view of Fig. 1, it is appreciated that the considered vessel classes are mostly characterized by L_{pp} larger 200 m, which underlines the importance of added resistance for smaller values of the relative wave length λ/L_{pp} with λ being the (absolute) wave length. Söding and Shigunov (2015) report that moderate sea states, characterized by wave lengths of 50 m to 150 m, are most commonly encountered by merchant ships. Hence, the bulk of experienced relative wave lengths arranges in the interval $\lambda/L_{pp} \in [0.125, 0.75]$, acknowledging that this regime of wave lengths is of higher complexity for the calculation of added resistance due to the dominant diffraction component.

The GPS position histories of all considered vessels are depicted in Fig. 2 and show a broad spatial coverage. Moreover, the route path projections of all vessels have been processed by the algorithm presented in Ikonomakis et al. (2022) for correcting any erroneous vessel positions. Still, minor inaccuracies are visible in the GPS data, e.g. in the southern part of Africa or Spain. It is stressed that the majority of the fleet operates the Europe-Far East trade via the Suez canal.

2.2. Sensor readings and filtering

The individual sensor readings being part of the CAMS[1] (Central Alarming and Monitoring System) have a sample frequency of 1 Hz. The data aggregated into 10-min samples is sent to shore whenever it is possible to establish a satellite connection. The basic sensor readings comprise GPS information, such as Longitude, Latitude, and Speed Over Ground (SOG). Additionally, engine-related data has been acquired, such as shaft power P_B and shaft revolutions rpm. Interquartile range filtering has been applied to the distribution of P_B/rpm^3 for the removal of potential outliers in the shaft power measurements. Moreover, sea and air temperatures have been measured on a number

of vessels, but for consistency, these two sensor readings were disregarded in this work entirely. Speed Through Water (STW) has been obtained by the Doppler velocity log, and it is noted that different types and manufacturers of this sensor have been used for different vessel classes. Ikonomakis et al. (2021) review several known issues of STW measurements and establish a filtering threshold in use of SOG as a reference, basically given as $|STW - SOG| < 2$ kts, in order to account for possible sea currents. The proposed threshold is also applied herein, and samples with SOG or STW below 5kts are filtered out, as this speed regime is considered maneuvering. It is noted that sequences with frozen sensor recordings or with an apparent offset have been filtered out manually in the case of STW and P_B. Moreover, the vessels were equipped with anemometers in the bow to measure relative wind speed $V_{w,rm}$ and direction ψ_r. The measured relative wind speed has been adjusted to a reference height of 10 m, and the recommended equations of ISO 19030 have been used for mapping the relative wind speed $V_{w,r}$ and direction ψ_r to the absolute wind speed V_w and direction ψ. The loading condition, i.e. the draft at bow T_f and stern T_a position, were also measured; however, high-frequency draft measurements are known for their reduced data quality under forward speed, as shown by Gunkel et al. (2018). Therefore, draft information from the loading computer is utilized in this work exclusively for robustness. It is noted that the draft data was linearly interpolated to match the 10-min interval of the remaining dataset. Moreover, large trim angles were removed, and a threshold of $\Delta T = T_a - T_f$ is defined so that only data for which $|\Delta T| < 2.5$ m is kept. In Fig. 3, the distribution of the relative mean draft is depicted for all time instances and considering the entire fleet (228 vessels).

As can be inferred from Fig. 3, the actual experienced mean drafts are highly skewed towards the scantling draft, as the relative draft is normalized by the design draft of the vessels. The average Froude number Fn of the entire fleet was 0.16 in the observation period; however, the average design Froude number is in the range of 0.22. It becomes clear that some of the vessels operate far off their actual design and contract conditions.

The richness of data allows for rigorous filtering thresholds, which may subsequently enhance the quality of results. Crucially, the data is filtered for consideration of steady conditions exclusively. Dalheim and Steen (2020b) present a detailed methodology for detecting unsteady sequences in high-frequency monitoring data based on the window-wise slope of the considered samples. However, in the present case, 10-min samples are available; therefore, the relative variances of STW, SOG, rpm, heading, and rudder angle are used to determine steady sailing conditions. The individual thresholds are empirically derived, and in Fig. 4, a sample result of the employed filtering techniques is presented for one 15,550 TEU vessel. It stands out that only legs with steady conditions are kept and that acceleration phases and port calls are disregarded.

The additional resistance due to steering is neglected by filtering out rudder angles $|\delta - \bar{\delta}| > 1.5$ deg. It is stressed that $\bar{\delta}$ denotes the global mean because the rudder position sensor had a minor offset in a few cases and was not precisely centered around 0 deg. To filter for possible shallow water resistance, samples with a depth-based Froude number $Fn_H = \frac{U}{\sqrt{gH}}$ larger than 0.5 were excluded. It is noted that H is the water depth below the keel, g is the gravitational acceleration, and U is the ship's advance speed. Even though water depth has been part of the sensor readings on most vessels, bathymetry data from ERA5 has been used for the robust and consistent computation of Fn_H. In addition, several vessel classes were disregarded due to the lack of essential sensor readings, such as GPS position or wind data. As a result of the filtering procedure, the remaining part of the study considers 166 ships in 22 vessel classes.

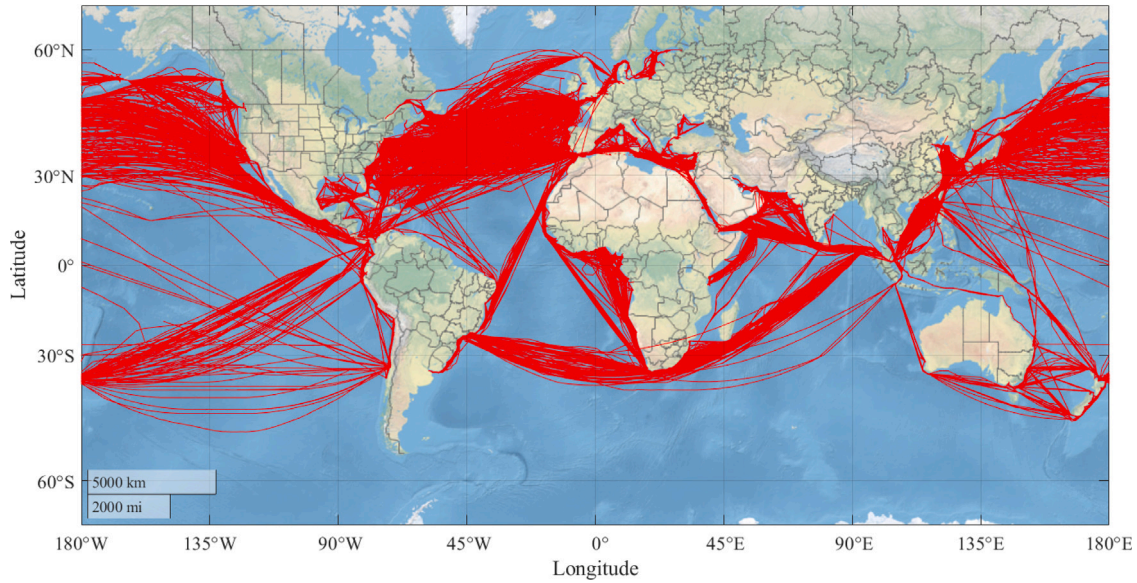[1] CAMS is the onboard data acquisition system of Mærsk Line.

Fig. 2. GPS position projection of 228 vessels in the time between January 2019 and March 2021.
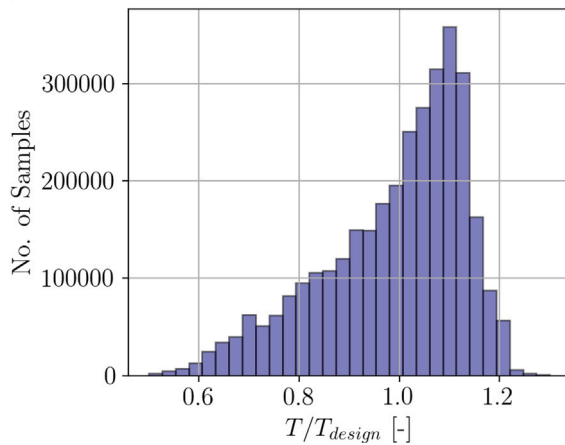

Fig. 3. Histogram of the relative draft according to the respective design drafts of the vessel classes.
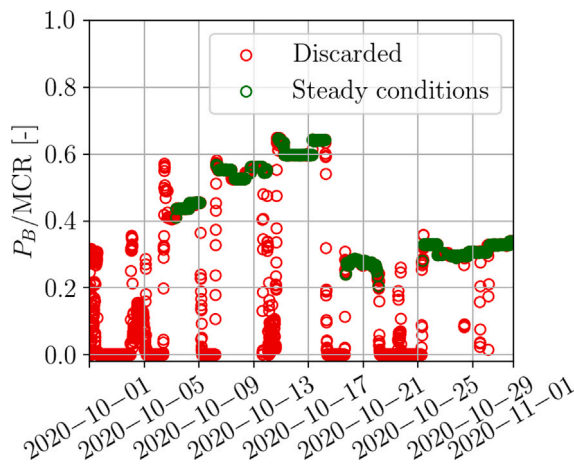

Fig. 4. Filtering methodology applied to one case ship (15,550 TEU) in October 2020. It is noted that MCR refers to the maximum continuous rating of the main engine.

2.3. Sea state information

In the final preprocessing step, the filtered data has been merged with wave data, based on ERA5, according to the corrected GPS data and UTC timestamps. It is stressed that samples with unavailable wave data were disregarded; thus, samples with erroneous GPS data (cf. Fig. 2) are dropped. The ERA5 database is part of the Copernicus EU program (Copernicus, 2020) and provides met-ocean data in hourly intervals at a 0.5 deg spacing in both Longitude and Latitude. In principle, the directional wave spectrum is available at every grid point. However, because of the ample required storage space and increased processing time due to the scale of the present dataset, it has been decided to work exclusively with integral wave parameters. Therefore, the directional spectrum $E(\omega, \beta)$ will be reconstructed using a parametric expression for the wave spectrum (at run time). Herein, it is of interest to model bimodal sea states and, thus, parameters of the wind-wave and swell partitions, as well as of the total mean system, were retrieved. Altogether, 10 parameters were downloaded per valid sample: significant wave height $H_{s,i}$ (total, wind and swell partition), peak period $T_{p,i}$ (total, wind and swell partition), mean zero up-crossing period T_z and wave direction v_i (total, wind and swell partition). It is noted that the wave direction v_i is converted to the relative wave direction, as seen from the ship, using the following equation: $\beta_i = v_i - \alpha$, where α denotes the ship's heading, which a gyrocompass has logged. It is noted that $\beta_0 = 180$ deg. defines head waves and 0 deg. represents following waves. For the spatiotemporal interpolation, a bilinear scheme has been used in the case of all 10 parameters. The fundamentals of interpolating sea state parameters along ships' routes are presented in Nielsen (2021). In Fig. 5, the joint distribution of the total H_s and T_z extracted from the present dataset is compared to global wave climate statistics. It is noted that the plot in Fig. 5(a) is an aggregated scatter, kernel density mapping, whereas Fig. 5(b) shows a contour plot.

The comparison of Figs. 5(a) and 5(b) reveals that the observed sea states are skewed towards lower values, notably for both H_s and T_z, in contrast to the long-term design probability distribution (DNV-GL, 2018). This finding is in line with the work of Nielsen and Ikonomakis (2021), and Miratsu et al. (2022), in which they list the use of weather routing and good seamanship as reasons for the lower experienced sea states. It is also mentioned that both distributions are characterized by different sources and degrees of epistemic (or systematic) uncertainty. Based on the isolines in Fig. 5(a), it is stated that roughly 80% of the
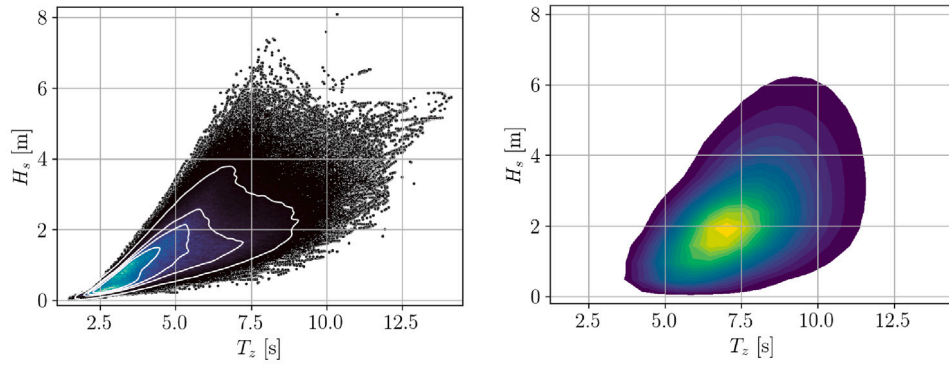
(a) Wave climate data retrieved from ERA5 based on GPS data of 165 ships in the time of interest.

(b) Long-term statistics from BMT (British Maritime Technology) as presented in DNV-GL (2018).

Fig. 5. Joint distribution of H_s and T_z (a) according to the present dataset and (b) to long-term statistics.

entire data has been acquired under conditions with $H_s < 4$ m. With that said, it is emphasized that ERA5 data is, indeed, assumed to be the ground truth in this work, acknowledging the fact that re-analysis data is generally an *estimate* of the encountered sea state. There are plenty of studies in the literature investigating the agreement of ERA5 data to wave buoy data or to other hindcast vendors, e.g. de Hauteclocque et al. (2020), or Nielsen et al. (2022). Therefore the validation of ERA5 data and the discussion of accuracy is beyond the scope of this work. Nonetheless, known related uncertainty sources will be presented in the later discussion herein.

3. Applied methodologies

The longitudinal wave drift force – the added resistance R_{AW} – is a second order quantity, i.e. $R_{AW} \propto \zeta^2$, where ζ denotes the wave amplitude. In general, added resistance is considered a function of ship speed U, ship draft T, mean relative wave heading β_0, wave frequency ω and wave amplitude ζ (Strøm-Tejsen et al., 1973). As Tsujimoto et al. (2008) pointed out, the added resistance of slender-type vessels, e.g. container vessels, is of high complexity and variance due to pronounced flare angles, possibly leading to non-linear wave breaking effects, the influence of a protruding bulb, and also due to higher forward speeds. Several uncertainty sources in experiments and numerical modeling are presented in Shigunov et al. (2018) and Park et al. (2015). The wave correction in ISO 15016 suggests two calculation methods for the assessment of the mean added resistance in irregular waves: (1) A spectral calculation method is described, which relies on an estimate of the added resistance transfer function and a representation of the wave environment in terms of a wave energy density spectrum. (2) The application of the empirical formula STAwave-1 (ITTC, 2014), which expresses a dependency on H_s and the ship's main particulars, is also mentioned in the respective standard. Overall, the formula has a major resemblance to the Kreitner (1939) formula and reflects a similarly narrow definition range in terms of β_0, U, H_s, and most importantly, ship type. Herein, the more general spectral calculation procedure (1) is chosen, and its intricacies are elaborated in the following subsection.

3.1. Spectral calculation of added resistance

In most publications about added resistance on ships in service conditions, the prevalent sea state is represented by a unimodal wave spectrum, assuming long-crested waves, when, in fact, real sea states and their representing wave spectra contain several partitions resulting from wind and swell. Moreover, the waves propagate from directions statistically distributed around a mean vector, i.e. waves are short-crested. That being said, average parameters as such – the relative *mean* wave direction in particular – have a reduced physical meaning in

multimodal seaways. In fact, only 22% of all samples within the used dataset, cf. Section 2, show a sea state that can be represented by a unimodal wave spectrum[2] highlighting the importance of considering multimodal conditions. Therefore, calculating the mean added resistance in short-crested waves, assuming a bimodal wave spectrum, is carried out using Eq. (1). The expression is based on the assumption of stationarity and superposition, i.e. linear wave theory in infinite water depth.

$$\bar{R}_{AW,sc} = 2 \int_0^{2\pi} \int_0^\infty \frac{R_{AW}(\omega, \beta)}{\zeta^2} E(\omega, \beta) d\omega d\beta \tag{1}$$

Given Eq. (1), it is said that the mean added resistance in short-crested waves $\bar{R}_{AW,sc}$ is computed by integrating the product of the quadratic transfer function R_{AW} and the directional wave spectrum E w.r.t. wave frequency ω and the (relative) wave direction β. In contrast, the mean added resistance in long-crested waves $\bar{R}_{AW,lc}$ is calculated with E integrated for all β and the transfer function R_{AW} is only computed for the relative mean wave direction β_0. Lloyd (1989) states that considering short-crested waves, in general, smooths the extreme variations of the responses. In addition, it leads to an increased roll response in following waves and increased pitch in beam waves. Due to the effect resulting from the consideration of short-crested waves on the first order responses, it is thought that similar observations may show in the case of the second order added resistance.

The semi-empirical formula proposed by Liu and Papanikolaou (2020) has been developed mainly for practical purposes, for instance, when the detailed hull geometry is unavailable, not to mention that the formula requires low computational effort. Its distinct advantage over other approaches, such as the STAwave-2 (ITTC, 2014), is the validity for all wave heading angles and thus the applicability in short-crested waves. Moreover, the underlying approach of Liu and Papanikolaou (2020) underwent a rigorous validation study (Wang et al., 2021) and is recommended by the ITTC Ship Operation at Sea (SOS) specialist committee for the wave correction during sea trials, ITTC (2021). The Liu and Papanikolaou (2020) formulation applies a correction term for wave reflection in short waves (R_{AWR}) to the added resistance caused by the motions of the vessel, i.e. radiation (R_{AWM}). The necessary predictors for the semi-empirical approach are the main particulars, i.e. L_{pp}, B, T_a, T_f, the block coefficient C_B and the length of run L_R and length of entrance L_E. Additionally, the operating conditions are required, i.e. the longitudinal radius of gyration k_{yy} and the Froude number Fn. Lastly, the wave length λ and wave direction β have to be specified as discrete vectors for computing $R_{AW}(\omega, \beta)$. In the

---

[2] In this case, a sea state and its corresponding wave spectrum is considered as unimodal, when either the significant wave height of the wind or the swell partition is smaller than 0.1 m.
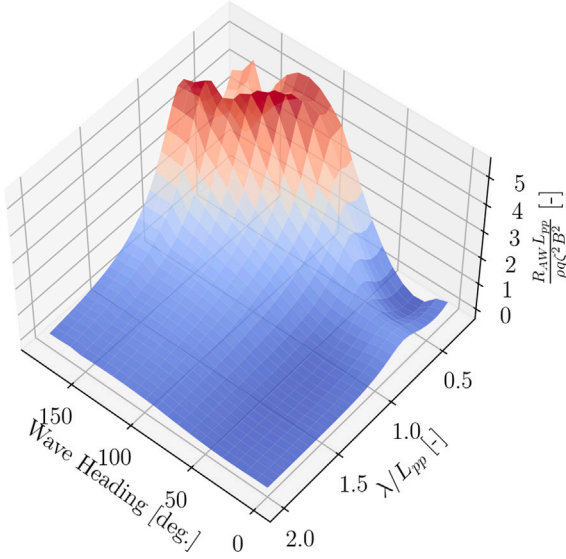
Fig. 6. The non-dimensional R_{AW} transfer function for $Fn = 0.16$ for one case vessel class (15,550 TEU) according to the formulation of Mittendorf et al. (2022b).

present work, the adaptation by Mittendorf et al. (2022b) of the original formula (Liu and Papanikolaou, 2020) is employed. The formula's parameter vector was calibrated for both slender and blunt-type vessels based on particle swarm optimization as well as a database of experimental data comprising 25 different ships and around 1100 samples. The computed non-dimensional R_{AW} transfer function is presented in Fig. 6 as a function of the relative wave length λ/L_{pp} and relative wave direction β. Note that the vessel class with a capacity of 15,550 TEU is taken as a case study for visualization.

In view of Fig. 6, it is observed that the formula by Mittendorf et al. (2022b) has a broad definition range and that the added resistance is more pronounced in head-to-beam waves. As it has been stated in Section 2, the added resistance in relatively short waves is of particular interest in this work, and in the formulation of Liu and Papanikolaou (2020), the reflection component is calculated by a modified version of Faltinsen's asymptotic formula. The theoretical details of this near-field approach are provided in Faltinsen et al. (1980). Söding and Shigunov (2015) state that Faltinsen's formula provides reasonable agreement in very short waves compared to other (semi-)empirical methods. From the plot Fig. 6, a minor discontinuity in the vicinity of $\beta = 165$ deg. is observed, which may be caused by the piece-wise calculation of the reflection component. It is noted that the spectral calculations were performed with 36 discrete wave directions $\beta \in [-\pi, \pi]$ rad and a frequency discretization of 50 discrete frequencies $\omega \in [0.01, 2\pi]$ rad/s. The upper cut-off frequency is justifiable by the low pass filtering effect of ships, i.e. they are irresponsive to high-frequency waves. Following the findings of Mittendorf et al. (2022b), there is an evident lack of experimental data beyond the threshold of $\lambda/L_{pp} < 0.15$. For this reason, the transfer function is padded in this particular regime using the value of $R_{AW}(\lambda/L_{pp} = 0.15)$.

The directional wave spectrum is calculated using the exact same frequency and direction discretization as the transfer function and is defined as $E(\omega, \beta) = S(\omega) D(\omega, \beta)$, i.e. a unidirectional parametric wave spectrum S is multiplied by a spreading function D. Herein, a JONSWAP (Joint North Sea Wave Observation Project) type spectrum is used for calculating S to be able to capture the ambient wave spectrum also in regions with limited fetch. Thereby it is understood that two separate spectra for wind and swell sea are superimposed for modeling bimodal sea states. In short, the present approach is inspired by the ten-parameter spectrum of Hogben and Cobb (1986), and the used

expression is defined in Eq. (2).

$$E(\omega, \beta) = \sum_{i=1}^{2} S_{J,i}(\omega | H_{s,i}, \omega_{p,i}) \times \frac{2^{2s_i-1} \Gamma^2(s_i + 1)}{\pi \Gamma(2s_i + 1)} \cos^{2s_i}\left(\frac{\beta - \beta_i}{2}\right). \quad (2)$$

As can be inferred from Eq. (2), a cosine-power type spreading function is applied to the underlying JONSWAP type spectrum S_J. The spreading function acts as a weighting function for preserving the energy of the spectrum and distributing it around a mean direction according to the spreading parameter s_i. Moreover, it is noted that $\omega_{p,i} = 2\pi/T_{p,i}$ and that Γ denotes the Gamma function. The directional spreading parameter s_i is defined in Eq. (3) and is a function of wave frequency ω and s_{max}, which is 10 in wind and 25 in swell conditions, as recommended by Goda (2000).

$$s_i = \begin{cases} \text{ceil}\left[(\omega/\omega_{p,i})^5 s_{max}\right] & \text{for } \omega \leq \omega_{p,i} \\ \text{ceil}\left[(\omega/\omega_{p,i})^{-2.5} s_{max}\right] & \text{for } \omega > \omega_{p,i}. \end{cases} \quad (3)$$

The unidirectional JONSWAP type spectrum S_J is an adaptation of the two parameter spectrum by Pierson and Moskowitz (1964), which is defined for fully developed seas. However, Hasselmann et al. (1973) found that sea states are never really fully developed, but continue to grow due to non-linear wave–wave interaction. For this reason, the peak enhancement factor γ has been introduced to the Pierson Moskowitz spectrum for better agreement to measurement data. The following expressions were adopted from the class guidelines as provided in DNV-GL (2018) and the JONSWAP type spectrum S_J is defined in Eq. (4).

$$S_J(\omega) = A_\gamma S_{PM}(\omega) \gamma^{\exp\left(-0.5\left(\frac{\omega - \omega_p}{\sigma \omega_p}\right)^2\right)} \text{ with } A_\gamma = 1 - 0.287 \ln \gamma. \quad (4)$$

The normalizing factor A_γ is applied for conservation of energy. The standard JONSWAP spectrum assumes $\gamma = 3.3$; and turns into the PM spectrum when $\gamma = 1$. In addition, the spectral width parameter σ is part of the exponent, and the underlying Pierson Moskowitz spectrum S_{PM} is defined in Eq. (5).

$$S_{PM}(\omega) = \frac{5}{16} H_s^2 \omega_p^4 \omega^{-5} \exp\left(-\frac{5}{4}\left(\frac{\omega}{\omega_p}\right)^{-4}\right). \quad (5)$$

As can be seen, the Pierson Moskowitz including the JONSWAP spectrum adopt a ω^{-5} decaying behavior in their tail part. The peak enhancement factor γ is in most publications a discrete value, e.g. $\gamma = 3.3$. However, in the present work, an expression dependent on the H_s and T_p is used and defined in Eq. (6), which has been adopted from DNV-GL (2018). It is noted that separate γ values are determined for wind and swell partitions.

$$\gamma = \begin{cases} 5 & \text{for } T_p/\sqrt{H_s} \leq 3.6 \\ \exp\left(5.75 - 1.15\frac{T_p}{\sqrt{H_s}}\right) & \text{for } 3.6 < T_p/\sqrt{H_s} < 5 \\ 1 & \text{for } T_p/\sqrt{H_s} \geq 5. \end{cases} \quad (6)$$

It stands out that Eq. (6) is a piecewise function and follows the definition range of the JONSWAP spectrum, i.e. $3.6 < T_p/\sqrt{H_s} < 5$. The spectral width parameter is defined in Eq. (7) separately for wind and swell partition.

$$\sigma = \begin{cases} 0.07 & \text{for } \omega \leq \omega_{p,i} \\ 0.09 & \text{for } \omega > \omega_{p,i}. \end{cases} \quad (7)$$

For illustrative purposes, the reconstructed directional spectrum, as obtained from Eq. (2), is compared to the directional spectrum from ERA5 for a specific space–time point. In the given case, a test site a few kilometers off the coast of the Faroe islands is used, and the result is shown in Fig. 7. The data applies to the time 3rd April 2016 00:00:00 at 60 deg of Latitude and 350 deg of Longitude, both to the exact degree. Hence, the presented spectra are unaffected by possible uncertainty due to interpolation.
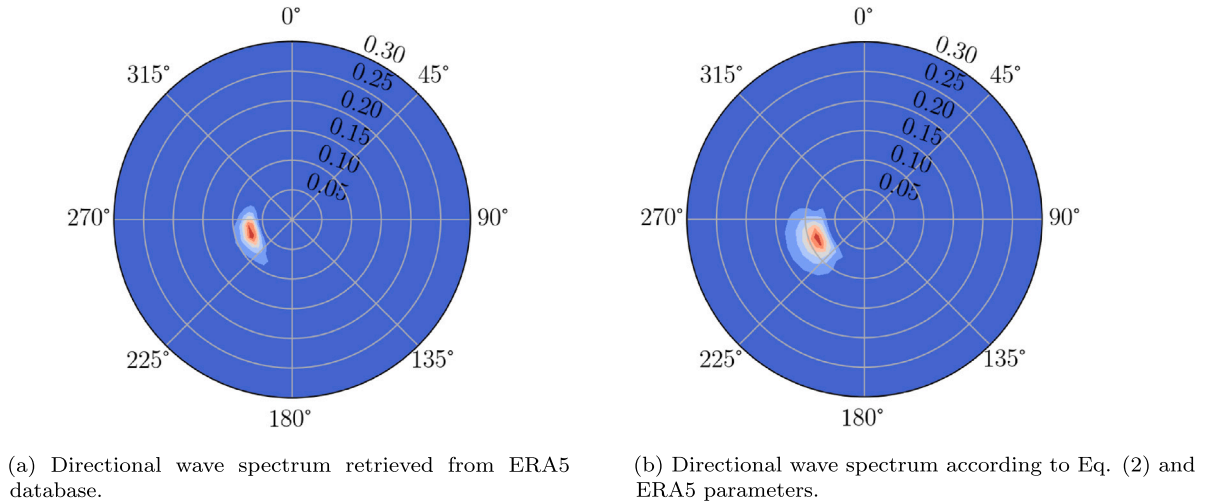
(a) Directional wave spectrum retrieved from ERA5 database.

(b) Directional wave spectrum according to Eq. (2) and ERA5 parameters.

Fig. 7. It is noted that both spectra correspond to the same time and position ($H_s = 3.7$ m). Also, it is stressed that $f = \frac{\omega}{2\pi}$ is displayed and that the spectral density is normalized.

In general, good agreement in terms of peak frequency and directional spreading is visible in Fig. 7, even though the directional spectrum from ERA5 shows minor asymmetric directional spreading. Modeling this is obviously not possible using the expression presented in Eq. (2). In addition, the spectral spreading is slightly larger in Fig. 7(b). However, it is concluded that the presented parametric expressions are capable of approximating the actual ERA5 2D wave spectrum sufficiently.

3.2. Indirect calculation of added resistance

The calculation framework adopted herein is taken partly from Nielsen et al. (2019) and builds upon semi-empirical resistance decomposition, i.e. superposition. The added resistance is calculated under steady conditions and derived by subtracting the calm water resistance from the actual measured in-service resistance. Thus, the determination of the empirical mean added resistance $\bar{R}_{AW,emp}$, i.e. based on measurement data, is summarized in Eq. (8).

$$\bar{R}_{AW,emp} = \frac{P_B}{STW} \times \eta_D \eta_M - R_{T,calm} - R_{wind} \qquad (8)$$

As can be seen, the actually experienced resistance force is obtained by multiplying the propulsive force $\frac{P_B}{STW}$ with the propulsive and transmission efficiencies. Following Kristensen and Lützen (2013), we assume $\eta_M = 0.98$ for the mechanical losses in, e.g. bearings. On the other hand, the determination of $\eta_D = \eta_0 \eta_H \eta_R$ is generally more complex and depends on the sea state. The available data of propulsive coefficients in waves (as a function of H_s, T_p, and β_0) is very scarce — for slender container vessels in particular. The experimental studies of Saettone et al. (2021) for a container ship and Yu et al. (2021) for an oil tanker are mentioned for the sake of completeness. The fundamental work of Nakamura and Naito (1977) shows model test data of propulsive coefficients of the S175 container vessel in regular head waves and indicates a quadratic decay of η_D with increasing wave amplitude. Additional contributors to the decrease in propulsive efficiency are cavitation and wake fluctuations in waves. In irregular waves, Taskar et al. (2019) show that the propulsive efficiency may decrease under wave impact using simulation data. Moreover, Shigunov (2017) assumes η_H and η_R to be constant, whereas η_0 has a dependency on H_s and β. It is, however, it may be sufficient to assume η_0 to be a constant as well for the higher forward speed and relatively low sea states. Additional driving factors for reduced propeller efficiency in waves are possible ventilation and an asymmetric wake profile in quartering waves, as shown by Prpić-Oršić and Faltinsen (2012) and Mikkelsen et al. (2022), respectively. Lastly, the open water propeller curves were not available,

and for this reason, the ISO 19030 default recommendation of $\eta_D = 0.7$ is adopted herein.

The studied fleet, cf. Section 2, shows a wide variety regarding hull form and outfitting characteristics. For this reason, the Hollenbach (1998) method has been used for the prediction of the calm water resistance $R_{T,calm}$, which is mainly due to the applicability for twin-screw vessels. Within the Hollenbach approach, $R_{T,calm}$ is calculated by the sum of the frictional resistance R_F, which is based on the ITTC'57 correlation line, and the residual resistance R_R, which is determined by a set of empirical formulae. Their parameters were derived from data of resistance and propulsion model tests of 433 displacement type vessels. The experiments were conducted in the Vienna model basin, and the formula was validated using experimental data from the Hamburg Ship Model Basin (HSVA). In Hollenbach (1997), it is stated that the method shows significantly less variance in contrast to other known methods, which stands out, particularly for twin-screw designs. Additional key advantages of the Hollenbach method include an uncertainty-aware estimate and the relatively small number of required inputs. Moreover, it is valid for different draft conditions and considers the entirety of possible appendages, such as shaft bossings or bow thrusters. The code implementation has been made according to Birk (2019).

Shigunov (2017) shows that the wind resistance may surpass the added-wave resistance in lower sea states, which is due to the slender hull shape and the large windage area of container vessels. Shielding effects and large separation bubbles caused by accommodation and container stacks amplify the wind resistance, Molland et al. (2011). Herein, the wind resistance is calculated according to ISO 15016 in Eq. (9).

$$R_{wind} = \frac{1}{2}\rho_{air}C_{AA}(\psi_r)V_{w,r}^2 A_x - \frac{1}{2}\rho_{air}C_{AA}(0 \text{ deg.})SOG^2 A_x \qquad (9)$$

It is stressed that A_x denotes the frontal area, the density of air is $\rho_{air} = 1.2$ kg/m^3 and that C_{AA} is taken from Kristensen and Lützen (2013). The wind resistance coefficients $C_x(\psi_r)$ are extracted from model test results presented in Andersen (2013) for the fully loaded vessel since this is the most conservative approach and also justified by Fig. 3. Lastly, it is stated that the presented approach disregards added resistance due to steering, nor is any effect of fouling of hull/propeller considered. Nonetheless, for the purpose of comparing the spectral approach, represented by Eq. (1) to the indirect, empirical prediction in terms of Eq. (8), the above-described methodology appears as reasonable, considering the general degree of uncertainty of in-service analyses of ships.
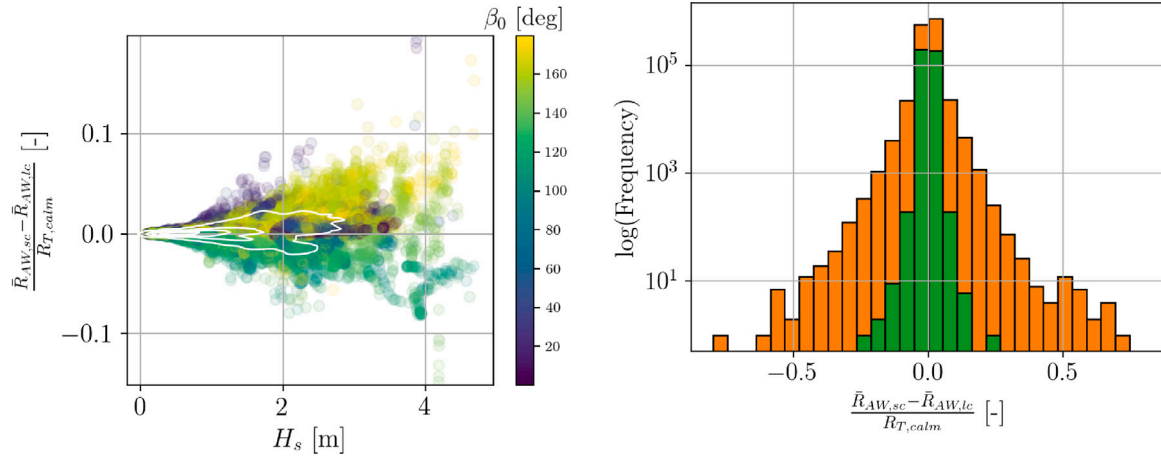
(a) Dependency of the relative residuals on H_s and β_0 and the joint distribution is indicated by a kernel density estimate using white isolines for the 15,550 TEU vessel class.

(b) Residual distributions in bimodal (orange) and unimodal wave conditions (green) for the entire fleet. Note that the ordinate is in logarithmic scale and a stacked histogram is depicted.

Fig. 8. Relative residuals of added resistance according to short- and long-crested waves.

4. Results

The following section is divided into three parts: (1) In Section 4.1, results of comparisons of the theoretical long- and short-crested wave computations for added resistance are presented. (2) In Section 4.2, comparisons between the short-crested theoretical estimates and corresponding predictions based on shaft power measurements and empirical resistance decomposition are shown. (3) Lastly, in Section 4.3, the sea margins based on the theoretical and empirical added resistance estimates are conveyed. Due to the different sizes of the individual vessel classes, the results are presented in a non-dimensional form throughout this section. As a definition, from this point, 'mean added resistance' is referred to as just 'added resistance' if not otherwise stated.

4.1. Comparison of added resistance estimates based on long- and short-crested waves

In Fig. 8(a), the deviations between long- and short-crested wave calculations of added resistance are shown in dimensionless form by normalizing the values by the empirical calm water resistance $R_{T,calm}$ (Hollenbach, 1998). Specifically, the relative residuals are presented as a function of significant wave height H_s, and the samples are colored according to the mean wave heading angle β_0. It is stated that only samples from the 15,550 TEU vessel class are shown, and a corresponding plot for the whole fleet is located in Appendix A.1. In the plot in Fig. 8(a), a heteroskedastic behavior of the residuals can be seen, i.e. their variance increases with higher H_s values, which underlines the non-linear character of added resistance. Additionally, it becomes clear that the magnitude of $\bar{R}_{AW,sc}$ is higher in head and following waves relative to $\bar{R}_{AW,lc}$, whereas it is vice versa in the regime of bow oblique and beam waves. This particular structure of the residuals is dependent on β_0 and is a result of the directional smoothing in short-crested wave calculations, i.e. 'averaging' over multiple wave headings. Moreover, in a few cases, the residuals exceed 10% of $R_{T,calm}$, and large scatter manifests in approximately 20% of the residuals, which has been estimated from the isolines. In fact, the four isolines segregate the kernel density estimate into five layers, which applies to all KDE plots in this paper.

It is found that the variance increases not only with H_s but it is also dependent on the number of modes (peaks) in the assumed wave spectrum representing the ambient sea state. This particular

observation is conveyed in Fig. 8(b) in a stacked histogram, which is applicable to the entire dataset. As described earlier, in this study, wave conditions are referred to as unimodal when either H_s of wind or swell waves are smaller than 0.1 m. Fig. 8(b) shows the relative residuals based on a unimodal wave spectrum (green bars) and a bimodal wave spectrum (orange bars), respectively. In view of Fig. 8(b), it is observed that the residuals between short- and long-crested wave calculations are characterized by more considerable variance in bimodal sea states. Two possible reasons are: (1) In contrast to unimodal seas, the mean wave direction, as a parameter, β_0 is physically less meaningful in a bimodal spectrum, as such a spectrum is composed of two partitions, most likely with different directions of the respective wave systems. However, in the case of $\bar{R}_{AW,lc}$, the added resistance transfer function is calculated for β_0, i.e. the wave energy is concentrated on an 'artificial' mean wave heading, which explains the significant variance in bimodal conditions between long- and short-crested calculations (considering actual wave directions). (2) Swell partitions usually show higher peak periods and, thus, their energy distribution is – at least for the considered ship sizes – closer to the resonance region, which in turn increases added resistance and underlines that swell waves cannot be neglected when calculating added resistance in actual conditions. An advantage of considering short-crested waves is the reduced relevance of the discontinuity in bow oblique waves within the utilized semi-empirical method, which stood out in Fig. 6. Similar observations in short-crested waves have been made in Liu et al. (2022) for the original method (Liu and Papanikolaou, 2020).

In Fig. 9(a), the non-dimensional added resistance in short-crested waves is depicted as a function of the mean relative wave direction, and samples are colored according to STW. Whereas, in Fig. 9(b), the non-dimensional added resistance is presented for T_z and colored according to the respective β_0 value. It is noted that the 15,550 TEU vessel class is used, and the sea state parameters are taken from the ERA5 hindcast data, whereas STW is part of the monitoring data.

In view of Fig. 9(a), it is found that $\bar{R}_{AW,sc}$ exhibits clear dependencies on β_0 and STW. Qualitatively, a similar behavior in magnitude is visible for β_0 when compared to regular waves in Fig. 6, as expected. On the other hand, in Fig. 9(b), it becomes clear that the ships encounter relatively short to medium wave lengths exclusively, since $T_z = 10$ s corresponds to $\lambda/L_{pp} = 0.42$ (considering infinite water depth). Even though computations in long-crested waves require only 34% of CPU time in direct comparison to the short-crested case, the latter calculation method is chosen for the remainder of this work. These particular
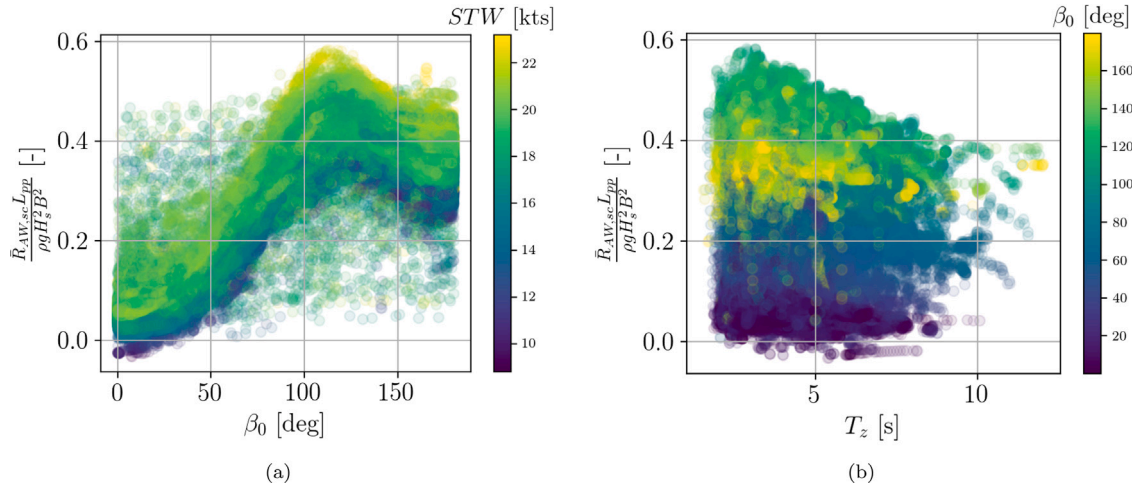
Fig. 9. Dependency of the non-dimensional added resistance on forward speed (STW), mean relative wave heading β_0 (a) and zero up-crossing period T_z (b).

computations do not only include the directional smoothing effect but they also distinguish wind and swell waves with their actual direction, which, in principle, yields a more accurate physical representation of prevailing wave conditions. For the sake of completeness, the residual distributions resulting from short- and long-crested wave calculations are presented in box plots for all vessel classes in Appendix A.1.

### 4.2. Comparison of spectral and empirical estimates of added resistance

In the following, data from the 15,550 TEU class with 8 vessels and 12.3×10^4 samples in total is considered. In Fig. 10(a), the relative empirical added resistance, resulting from the methodology presented in Section 3.2, together with the theoretical estimate in short-crested waves, are depicted as a function of the significant wave height H_s. Moreover, the difference (ΔU) between STW and SOG is shown in a colormap. It is appreciated that the empirical added resistance is subject to pronounced variance or, rather, uncertainty. It is striking that the corresponding values do take not only positive values but also negative ones, which is seen as unphysical in head-to-beam waves and underlines the large degree of uncertainty within the empirical predictions. Conversely, the theoretical added resistance shows roughly a quadratic dependency on wave height, as expected. Additionally, an outlier cluster becomes evident in the proximity of $H_s = 0.8$–0.9 m, possibly caused by frozen sensors. Furthermore, a dependency of the empirical added resistance on $\Delta U = STW - SOG$ is appreciated, which appears reasonable and may be caused by three aspects: (1) The effect of sea currents, (2) reduced sensor accuracy, and (3) insufficient filtering for steady conditions, i.e. some samples influenced by acceleration are still included. In addition, a variance decrease with increasing significant wave height (heteroskedasticity) is visible in Fig. 10(a), which is caused by unbalanced data, i.e. epistemic uncertainty. Multiple sources of uncertainty related to the estimation methodologies and ship telemetry data, in general, will be discussed in Section 5.

In Fig. 10(b), the relative residuals between the short-crested and empirical predictions are assessed corresponding to Fig. 10(a). It is stressed that a similar plot for the entire fleet is attached to this paper in Appendix A.2. To show the actual distribution of the residuals, a kernel density estimate is included. As an immediate observation, major uncertainty stands out in 20% of the data (estimated by isolines). Furthermore, a skewness appears within the residuals dependent on β_0, i.e. the residuals tend to be positive in head-to-bow oblique waves, whereas it is vice versa in beam-to-following waves. In fact, a large variance stands out in the following wave regime. The added resistance in following waves is of high complexity and, for this reason, often

neglected both experimentally and numerically. Container ships have a pronounced transom stern, which affects the prediction of added resistance in following waves: Transom sterns may lead to flow separation and breaking waves, which are generally challenging to model, especially by potential flow theory, e.g. Faltinsen's asymptotic formula. Additional challenges in following waves are not only the possible non-linear effect of broaching, i.e. an induced yaw moment due to surf-riding but also the ambiguity in wave loads due to the dispersion of waves, i.e. long waves overtake the ship, whereas short waves are overtaken by the ship. It is stated that conservative estimates of added resistance are obtained in higher sea states via the spectral method, but the large variance in the empirical added resistance and the reduced data availability in the regime $H_s > 3$ m impede drawing any firm conclusions. However, Shigunov (2017) shows similar results for two container vessels, i.e. with notable variance and possible overprediction of theoretical estimates in higher sea states.

In Fig. 11, box plots present the basic quantities of the underlying residual distributions of the added resistance calculations separately for each vessel class. It is noted that box plots indicate the interquartile range with its body, the median as well as the maximum and minimum of the distribution, i.e. the 0th percentile and 100th percentile, respectively. As in the previous section, the residuals are calculated relative to $R_{T,calm}$ and for ensuring confidentiality, the individual vessel classes are labeled by an ID number and ID14 corresponds to the 15,550 TEU class.

As can be inferred from Fig. 11, the two present estimation methodologies of added resistance yield values of similar magnitude for the majority of vessel classes, indicated by a residual distribution centered around zero. On the other hand, some vessel classes have a notable offset, i.e. the residuals are heavily biased. Several reasons for the bias are plausible, such as a dependency on ship size, slenderness, additional outfitting equipment (e.g. cranes), or propulsion arrangement (i.e. the number of propellers). However, no common reason has been identified among the outlier vessel classes. To add to the list of reasons, initially, it was thought that the poor agreement in the case of the ID6 class was due to relatively low data availability. However, the ID19 class shows a larger share (2.73%) of the total dataset, and the agreement is still considered relatively poor. Moreover, the variance is not constant for all vessel classes, and significant variance in the estimates appears for some ship types (e.g. ID5), but yet again, no common reason can be connected to that. The conclusion of Tsujimoto et al. (2008) is confirmed herein, i.e. obtaining reliable estimates in beam-to-following waves is challenging — especially for fast and slender vessels. However, in contrast, it is concluded that added resistance in stern quartering waves is, in fact, not negligible due to its magnitude and effect on ship safety. Ultimately, it is reported that the obtained results of the
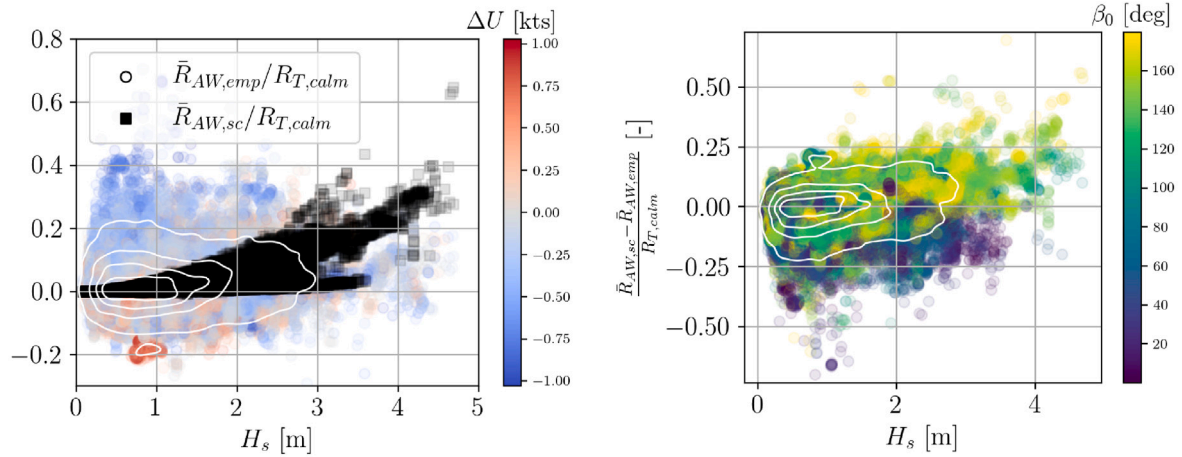
(a) Comparison of relative empirical and theoretical added resistance and the colormap indicates $\Delta U = STW - SOG$.

(b) Relative residuals of both methods presented for H_s and the colormap shows β_0.

Fig. 10. Analysis of relative theoretical and empirical added resistance for the 15,550 TEU vessel class in direct comparison (a) and their residuals (b).
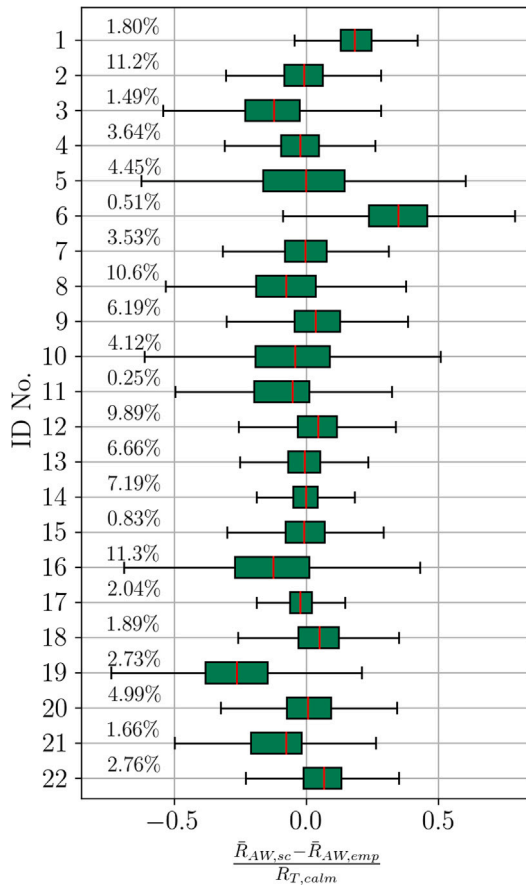

Fig. 11. Box plot of the relative residual distribution for all 22 considered vessel classes. Note that the relative samplesize is shown on the left hand side in the plot and that outliers are not depicted.

comparison show major variance but a minor bias, albeit of increasing magnitude in more severe sea states. This is also concluded in studies from the presented literature, e.g. Dalheim and Steen (2020a). The most central and noteworthy point from the previous results and discussions is the large variance in comparing theoretical and empirical added resistance. Generally, ship performance assessments, based on

in-service data, are typically associated with large uncertainties, and, as mentioned earlier, in Section 5, multiple uncertainty categories will be presented and discussed.
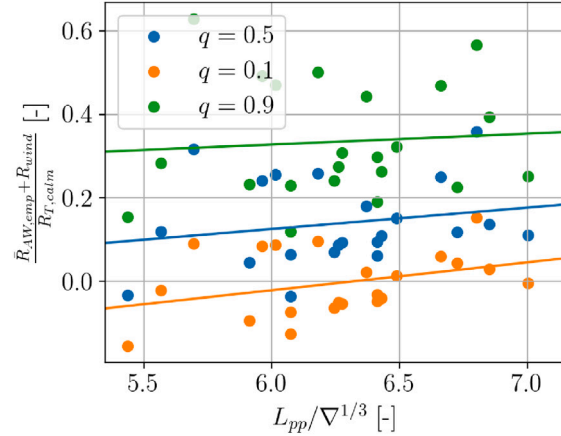
4.3. Calculation of the sea margin

The present study shows not only ramifications to vessel performance analysis or sea trial corrections, but it is also believed that analysis of in-service vessel data may impact the design of new vessels for optimizing several characteristics, such as required engine power, for their actual operational profile. Söding and Shigunov (2015) state that the hull shapes of slender and fast ships (e.g. container ships) are commonly not optimized for added resistance, but the calculation of viable engine margins requires the consideration of added resistance(s). Especially during the initial design phases, a sea margin is needed to estimate the required main engine power correctly. As mentioned in Section 1.1, constant values of approximately 15% are most frequently employed for this purpose. Due to the richness of data, it is attempted to postulate general hypotheses for container vessel design and their requirements for engine power in actual conditions. However, seamanship, maintenance scheduling, and the fidelity of the weather routing software are company-specific, which introduces certain biases into the present dataset. In the present case, a dependency of the empirical and theoretical sea margins on the vessels' slenderness ratio is shown in Fig. 12 using the 10th percentile, 50th percentile, and 90th percentile of the individual distributions of all 22 vessel classes. Herein, the sea margin is calculated by dividing the sum of added resistance in short-crested waves and wind resistance by the empirical calm water resistance, i.e. $\frac{\bar{R}_{AW} + R_{wind}}{R_{T,calm}}$. As a side note, the excess resistances due to marine growth and shallow water are neglected in the present analysis.

As shown in Fig. 12, both the theoretical- and empirical-based sea margins show dependencies on the slenderness ratio. In general, linear relationships between sea margins and $L_{pp}/\nabla^{1/3}$ are visualized using regression lines for all three quantiles q. Overall, it appears that theoretically and empirically derived sea margins show an opposing dependency on the slenderness ratio. In general, it is believed that the added resistance, and thus the sea margin, will be reduced for more slender vessels, which can be seen in Fig. 12(a). Moreover, the variance is relatively large in relation to the slope of the regression lines in Fig. 12(b). Hence, no defensible conclusions are possible in the case of the empirically derived sea margin. Moreover, the distribution of the sea margin appears to be asymmetric and skewed towards lower values resulting from the exponential H_s distribution. The 90th percentiles
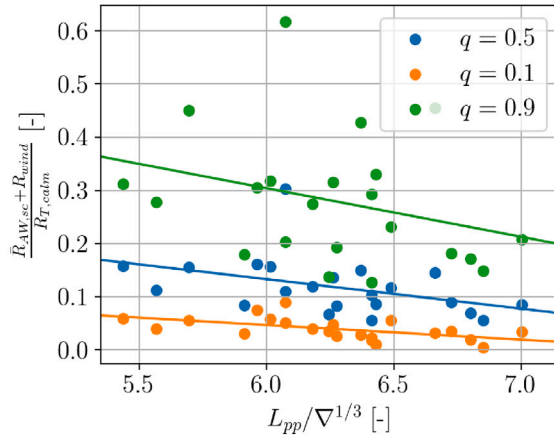
(a) Sea margin based on the theoretical estimate of added resistance.

(b) Sea margin based on the empirical estimate of added resistance.

Fig. 12. Three quantiles q of the sea margin distributions as a function of the slenderness ratio for all 22 considered vessel classes. It is noted that ∇ is the volume displacement at design draft.

exhibit significant scatter and samples in both Figs. 12(a) and 12(b), which are exceeding a 60% sea margin. Simply speaking, the lump sum corrections recommended for the use in the first design iterations appear reasonable given the average sea margin in Fig. 12 noting that Kristensen and Lützen (2013) suggest sea margins of around 20%–25% for the trade from Southeast Asia to Europe. It goes without saying that in advanced design iterations, more detailed approaches, as presented in Liu et al. (2022) or a simulation-driven design process, as shown in Harries et al. (2019), are needed for viable ship designs. In the end, the analysis of sea margins is of even greater importance in the case of slower and blunt-type ships, such as bulk carriers and tankers, due to the risk of losing maneuverability in severe sea states and thus for the minimum propulsion power requirements.

5. Discussion of uncertainty sources

Shigunov (2017) states that a large set of influencing factors of added resistance is associated with significant uncertainty or may even be completely unknown. It has been stressed that it was not possible to pinpoint *one* overall reason for the relatively large uncertainty in the respective estimates of added resistance and their comparison. Hence, several possible contributing reasons are discussed in the following, and several potential sources of uncertainty are presented, with the spectral calculation and the indirect empirical methodology addressed separately. First, however, common influencing factors of uncertainty are pointed out.

The empirical methodology is based on resistance decomposition, and the spectral method relies on the superposition of several wave components. Therefore, both adhere to linear theory and neglect interaction effects per definition. In reality, a ship's behavior in a seaway is non-linear, especially in higher sea states, and, therefore, superposition, in general, is not strictly valid. As a matter of fact, non-linear phenomena, such as propeller load fluctuations, propeller ventilation, and slamming (resulting in whipping-induced vibrations), increase in likelihood and magnitude in higher sea states. Moreover, the wave-induced transverse drift force and the yaw moment have been disregarded in the present work, but Shigunov (2017) reported a minor effect of an unrestrained heading on the results of added power in waves. The chosen physical parameters may also introduce uncertainty in both methods. For instance, seawater properties for a temperature of 15 °C, i.e. ρ_{water} = 1025.9 kg/m^3 and kinematic viscosity ν = 1.19298 × 10^{-6} m^2/s, were assumed to be constant, according to ISO 19030, when in fact both are dependent on temperature and salinity. Therefore, it is
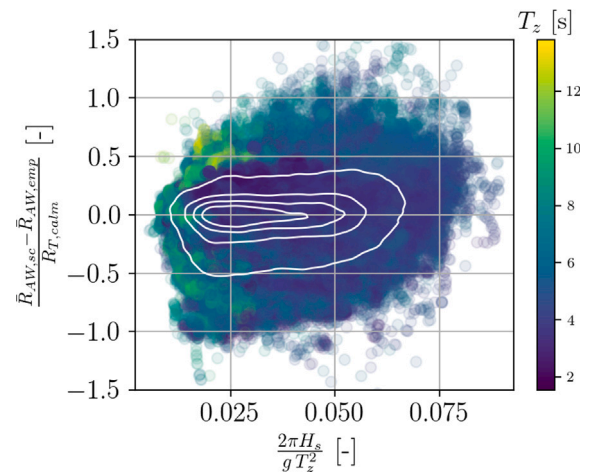

Fig. 13. Sensitivity of the relative residuals to the wave steepness irregular waves.

believed that the viscous resistance component is subject to scatter and possibly overestimated in regions with higher temperatures. Furthermore, in this study, all ships are represented by only their basic ship main particulars; since no details about hull or propeller shape or exact loading conditions were accessible.

5.1. Uncertainty sources within the spectral method

Inherently, the spectral calculations of added resistance do not capture the interaction effects and unsteady phenomena due to the underlying assumptions of Eq. (1). Moreover, linear wave theory assumes moderate wave steepness and it has been shown experimentally in regular waves in Park et al. (2019) that wave steepness has a significant effect on the added resistance transfer function in short wavelengths and on the occurrence of wave breaking. Therefore, the effect of wave steepness, $S_S = \frac{2\pi H_s}{gT_z^2}$, in irregular waves on the residuals has been studied in Fig. 13.

In fact, Fig. 13 reveals no significant correlation between residuals and steepness, but the variance of the relative residuals increases notably with higher S_S values. A clear error source within this analysis is that a mean steepness aggregating both wind and swell partition is used. Furthermore, the application of simplified expressions for both
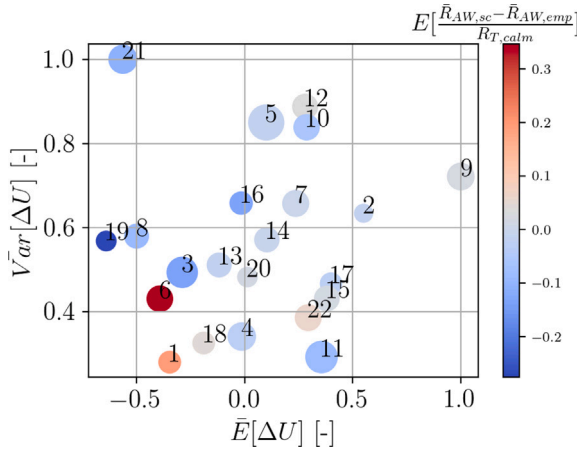
Fig. 14. Normalized variance and bias of $\Delta U = STW - SOG$. It is noted that the dots are colored according to the bias of the residuals and their size corresponds to the residuals' variance, as presented in Fig. 11.
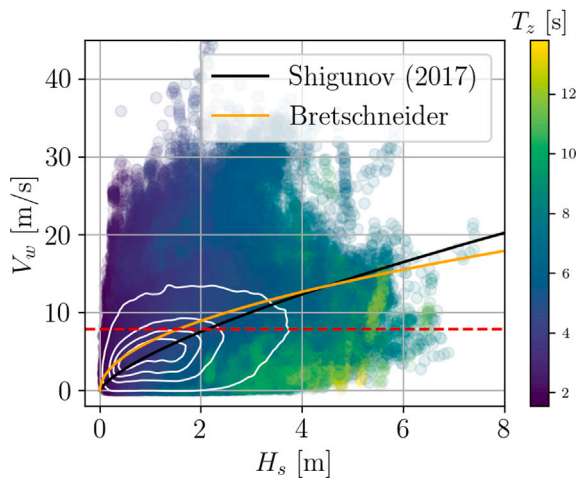

Fig. 15. Absolute wind speed V_w as a function of significant wave height H_s including two regression analyses and the ISO 19030 filtering threshold (red).

the quadratic R_{AW} transfer function and the parametric wave spectrum come with their own uncertainties. For instance, the present parametric directional spectrum is limited to symmetric directional spreading, while directional wave spectra of natural ocean waves rarely appear symmetric in their energy density directional distribution (Barstow et al., 2005). The retrieved sea state data and its interpolation introduce additional variance, as discussed by Nielsen (2021). Moreover, de Hauteclocque et al. (2020) compare several hindcast vendors (including ERA5) to data from 9 wave buoys and show that ERA5 data tends to underestimate the wave height by approximately 4%. In the same light, the comparison of ERA5 to altimeter (satellite) data reveals a reduced underestimation of only 1.6%; however, more significant deviations have been identified in harsh conditions, i.e. $H_s > 8$ m. In the present context, such conditions are mostly circumvented by weather routing; therefore, the associated discrepancies are not necessarily influential. In Section 4.1, different behavior in the residual variance has been observed in uni- and bimodal sea states, and, in fact, similar findings as in Fig. 8(b), but with larger variance, have been made in case of the comparison of theoretical and empirical added resistance. The corresponding histograms are placed in Appendix A.2. The employed semi-empirical method of Mittendorf et al. (2022b)

is already characterized by a reduced model parameter uncertainty due to the conducted calibration procedure based on particle swarm optimization and experimental data of slender and blunt-type ships. Still, notable variance and underestimation were observed in Mittendorf et al. (2022b) while validating the method for a fast and slender ship. As reported by Tsujimoto et al. (2008), the estimation of added resistance of slender vessels is problematic due to their pronounced flare angles (possibly leading to wave breaking), their higher operating speeds, and their distinct bulbous bow. Furthermore, the exact loading conditions and, thus, the metacentric height and the longitudinal radius of gyration k_{yy} are unknown. As shown by Holt and Nielsen (2021), the latter has a sizable impact on the added resistance transfer function and has been approximated with $k_{yy} = 0.25 L_{pp}$ throughout this study. In fact, Dalheim and Steen (2020a) conclude that both limited knowledge of the actual k_{yy} and nonlinear effects, such as wave breaking, are leading contributors to the deviations in their comparison of theoretical and empirical added resistance.

5.2. Uncertainty sources within the empirical method

The indirect empirical predictions are based on superimposing several resistance components, and the underlying calm water resistance method relies on a similarity to the parent hulls of the Hollenbach (1998) method. However, specific geometric characteristics, such as vertical bow shapes, as in the case of the ID1 class, were not considered among these underlying ship models. In the future, it could be interesting to study the sensitivity to several different empirical calm water resistance methods, such as the approach by Guldhammer and Harvald (1974). Furthermore, changes in engine characteristics in waves were neglected throughout this study, even though transient behavior, such as ventilation and propeller racing, may lead to engine overload, as described in Prpić-Oršić and Faltinsen (2012). Clearly, calculating the propulsive efficiency η_D is essential for the empirical calculation methodology. Even though Shigunov (2017) found that the influence of η_D on actual engine power in a seaway is one order of magnitude less than the influence of a resistance increase due to waves, which in turn justifies, to some extent, that the current focus is merely attributed to the accurate modeling of the wave-induced added resistance. Still, it has been shown by Nakamura and Naito (1977) that η_D decreases quadratically with increasing H_s, and consequently, the ISO 19030 default value of 0.7 loses its validity in higher sea states. Therefore, the found deviations in higher sea states in Fig. 10(b) may be partly attributed to the application of the constant value of η_D. More accurate modeling of propulsive coefficients and propeller characteristics is an essential area for future work. On a separate note, it is thought that data quality of monitoring data is another significant driving factor of the observed uncertainty — despite the rigorous filtering procedure. When it comes to the measurement uncertainty of ship performance data, Aldous et al. (2015) conclude that both speed and power measurements are most affected by measurement error and sensor drift when compared to other essential sensor readings. From Eq. (8), it can be inferred that both are vital for the empirical determination of added resistance. Hence, sensor maintenance and quality are of great importance for obtaining sufficient information on added resistance from in-situ data. In Fig. 10(a), we see that the empirical added resistance depends on $\Delta U = STW - SOG$. Thus, it is herein attempted to analyze the variance and bias of the different Doppler velocity logs with respect to SOG as a reference. Obviously, the analysis assumes both quantities to be time-invariant when, in fact, both quantities of STW are subject to sensor drift and calibration. However, an estimate of bias and variance may still be indicative of deviations observed in Fig. 11. In Fig. 14, the normalized variance and mean of ΔU are shown for each vessel class. In parallel, the variance and mean (or bias) of the residuals are visualized by the size of the dots and the colormap, respectively.

From Fig. 14, it is appreciated that there is a minor coupling between the variance and bias of ΔU and the residuals of theoretical and empirical added resistance. For instance, the vessel classes of ID6 and ID19 are characterized by large biases in both cases. It is striking that vessel classes with a large bias of the residuals show a negative bias of ΔU. Conversely, vessels with a positive ΔU bias show a negligible bias of the residuals. Broadly speaking, the sizes of the dots (i.e. the residual variance) also increase with higher variance of ΔU. However, ID11 is considered an outlier. All in all, the sensor quality, possible sea currents, and the applied method for filtering for steady samples may be responsible for these observations. It is believed that similar observations can be made for the torsion meter, i.e. the sensor for the brake power P_B. However, in this case, there is no direct reference value for a similar investigation like for STW. Data quality and sensor maintenance are critical issues in the field of performance monitoring. Therefore, it could be an interesting aspect to follow the approach of Sogihara (2021) and study the effect of individual sensor readings as well as wave parameters on the overall uncertainty within the added resistance estimate using a Monte Carlo simulation methodology.

It has been stated in Shigunov (2017) that the wind resistance of container ships can be of higher magnitude than added resistance due to waves in lower sea states, which is due to the large frontal area of container vessels. In fact, ISO 19030 imposes a filtering threshold based on the absolute wind speed V_w alone for disregarding any instances with more severe environmental conditions, which may lead to higher added resistance due to wind *and* waves. For fully developed wind seas with unlimited fetch, the relationship $H_s \propto V_w^2$ holds, and there are well-known formulae from, e.g. Bretschneider $H_s = 0.0248 V_w^2$ (Michel, 1999) and $H_s = 0.115 V_w^{1.41}$ (Shigunov, 2017), which are visualized in Fig. 15. In the following, the absolute wind speed V_w, as derived from relative wind measurements on all of the considered vessels, and corrected to the reference height of 10 m, is depicted for the significant wave height H_s. The samples are colored according to the zero up-crossing period T_z, and the wind threshold (i.e. 7.9 m/s) of ISO 19030 is included as a broken red line. It is noted that the two sea state parameters are subject to potential inaccuracies due to spatiotemporal interpolation and the inherent uncertainty of the ERA5 database.

Generally, it can be seen in Fig. 15 that the kernel density, i.e. the bulk of the data, clusters around the two empirical formulae in the lower sea states, i.e. $H_s < 4$ m. Still, significant variance stands out, and even possible erroneous samples are visible, e.g. samples with $V_w > 20$ m/s and $H_s < 1$ m. In fact, wind measurements are subject to major uncertainty, which is often a result of the placement of the anemometer and the resulting wake from, e.g. container stacks or the accommodation. Sogihara (2021) reports a measurement accuracy of $\pm 5\%$ and ± 5.0 deg. in case of relative wind speed and direction, respectively. In fact, too high wind resistance values were also reported by Schmode et al. (2018) and may be responsible for the negative bias of the residuals seen in Fig. 10(b). In this light, it may also be argued that the ISO 19030 wind threshold could be deemed inappropriate not only because of the weaker relationship of V_w and H_s in higher sea states but also due to the fact that swell waves are independent of the prevailing wind speed but yet may be significant in terms of added resistance (indicated by higher T_z values).

ISO 19030 is heavily discussed in the practical and academic domain and Schmode et al. (2018) suggest several improvements for ISO 19030 and point out drawbacks of the standard procedure, such as the speed and draft dependency of the ISO 19030 performance indicators. Moreover, Schmode et al. (2018) state that the strict filtering mandated by ISO 19030 may lead, in some cases, to the fact that a great chunk of the data has to be disregarded, which directly impacts the reliability of the obtained performance indicators. In view of Figs. 15 and 10(b), it appears favorable to filter samples according to shaft power measurements and a speed-power baseline, e.g. a sea trial curve. Bertram (2016) suggests disregarding any samples with a relative added power (dividing by the reference power from the baseline) that exceeds 100%.

In doing so, not only higher values of added wave and wind resistances (and possibly non-linear behavior) are neglected, but also potential measurement errors are filtered out. Another filtering technique could be based on the estimate of hindcasted H_s values and thus on the Beaufort scale. In ISO 15016, it is recommended (for $L_{pp} > 100$ m) to perform sea trials only if Bft < 6, i.e $H_s < 4$ m and $V_w < 13.8$ m/s. A similar threshold appears as reasonable in the case of ship performance analysis, i.e. ISO 19030.

Wind forces as such and their exploitation received increased focus in recent years, both due to the Suez canal blockage in 2021 and due to the re-emergence of wind(-assisted) ship propulsion. Hence, it is believed that dedicated studies for accurate modeling of the wind forces on container (and other) vessels and their uncertainty quantification could be of significant scientific and practical relevance. As pointed out, wind speed and direction measurements may be subject to considerable error due to the distorted airflow around the ship's hull, accommodation, and container stacks. Moreover, major differences in data quality of measurements obtained from different wind anemometers have been found in, e.g., the study of Oh et al. (2018). Therefore, a similar sensitivity study to that of Ikonomakis et al. (2021) for STW appears to be highly relevant for ship-based wind measurements and could be carried out using ERA5 wind data as a reference. Such studies may enhance understanding and corresponding estimation procedures of wind loads for both ship performance and sea trial assessment.

6. Conclusions and future work

This paper assesses theoretical and empirical estimates of added resistance using in-service data from a fleet comprising more than 200 container vessels. The calculation of the theoretical added resistance has been carried out in the spectral domain using an established semi-empirical formula for arbitrary wave heading angles (Mittendorf et al., 2022b). The sea state, on the other hand, has been modeled by a parametric expression for a directional sea state based on a JONSWAP type spectrum (Hasselmann et al., 1973), and the required parameters for both wind and swell partitions have been retrieved from the ERA5 database (Hersbach et al., 2018). For comparison, an empirical added resistance prediction was made using resistance decomposition and ship telemetry data, including shaft power measurements. The ship telemetry data underwent a rigorous filtering procedure beforehand to disregard any instances with, e.g. shallow water or unsteady conditions. Similar to other contributions from the literature, pronounced variance, and uncertainty in the comparison of theoretical and empirical added resistance estimates have been found, and multiple uncertainty sources have been discussed. It has been shown that data quality is vital for reliable predictions of R_{AW}, and the fidelity of both sensors and filtering techniques are essential. Ultimately, as the main finding of the present study, the conclusion of Bertram (2016) is confirmed, i.e. added resistance is generally difficult to predict in actual conditions — particularly in short and oblique waves.

In conclusion, the assessment made in this study calls for an uncertainty-aware estimation methodology of added resistance. In fact, the method of Mittendorf et al. (2022b) for the added resistance transfer function, and the approach by Hollenbach (1998) for the calm water resistance are both capable of providing transparent uncertainty bounds. However, for a consistent estimation procedure, reliable uncertainty information of sea state data is necessary. Following the work of Bos (2018), an ensemble of hindcast providers could form the basis of uncertainty quantification of wave data. Moreover, information on the prevailing wave environment in terms of a wave spectrum with increased accuracy is an essential aspect for calculating the mean added resistance in irregular waves. The application of the wave buoy analogy (e.g., Nielsen, 2017) – as presented in Mittendorf et al. (2022c) – for data acquisition (and for real-time decision support) shows potential for enhancing the quality of sea state information and thereby of in-situ added resistance estimates. On a separate note, it is
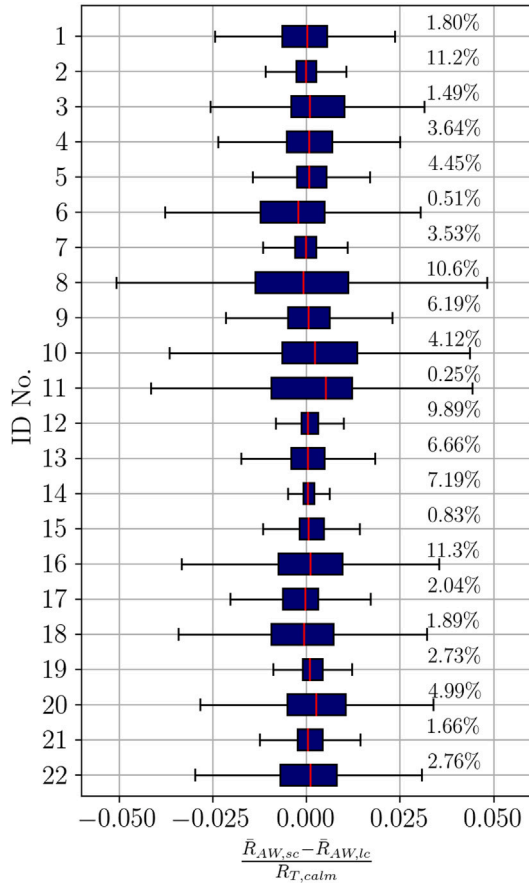
Fig. 16. Box plot of the relative residual distribution for all 22 considered vessel classes. Note that the relative samplesize is shown on the right hand side in the plot. Note that outliers are disregarded.
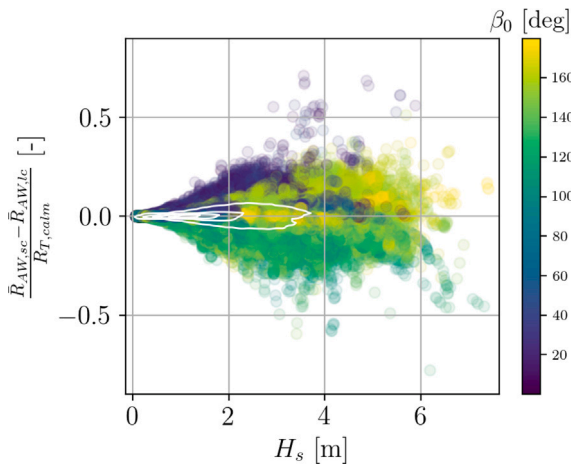

Fig. 17. Residual plot of the relative residuals between long- and short-crested results as function of H_s for all vessels.

Fig. 18. Residual plot of the relative residuals between empirical and theoretical method as function of H_s for all vessels with β_0 indicated by a colormap.

replacing the spectral calculation procedure. In their study, simulated training data has been used to develop an efficient tool for ship performance assessment. Hence, a similar method based on monitoring data appears favorable for future work. In addition, the use of unsupervised learning, i.e. clustering, appears to be promising for finding patterns and structures within the data and displaying them on a lower-dimensional manifold, as shown by Górski et al. (2021), appears as an essential next step. On a final note, the entire analysis is based on linear spectral analysis; however, as shown by Hasselmann (1966), the use of higher order spectral methods appears to be interesting when focusing on second order quantities, such as the added resistance. In a power density spectrum, phase information (and interaction) is lost, but with bi-spectral analysis and their generalizations, it is maintained (Iseki, 2010). Nevertheless, it is conceded that the use of higher order spectral analysis is relatively immature in the field of ship seakeeping and added resistance. It is, however, believed that the application of machine learning and image recognition methods, i.e. convolutional neural networks, may facilitate the use of bispectra for in-service ship analysis — including the field of added resistance due to waves.

CRediT authorship contribution statement

Malte Mittendorf: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualization, Writing – original draft, Editing. **Ulrik Dam Nielsen:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Resources, Funding acquisition. **Harry B. Bingham:** Conceptualization, Methodology, Writing – review & editing, Supervision, Resources, Project administration. **Jesper Dietz:** Conceptualization, Methodology, Data curation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

thought that numerical calculations combined with machine learning for approximating results from computationally expensive methods by regression algorithms will be vital in case of the prediction of added resistance, as shown in Mittendorf et al. (2022a), which focused on regular waves. The same holds for application to the mean added resistance in irregular waves – as presented by Guo et al. (2021) – for
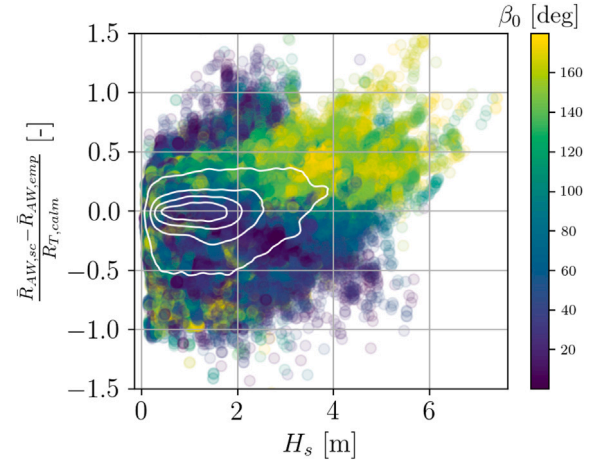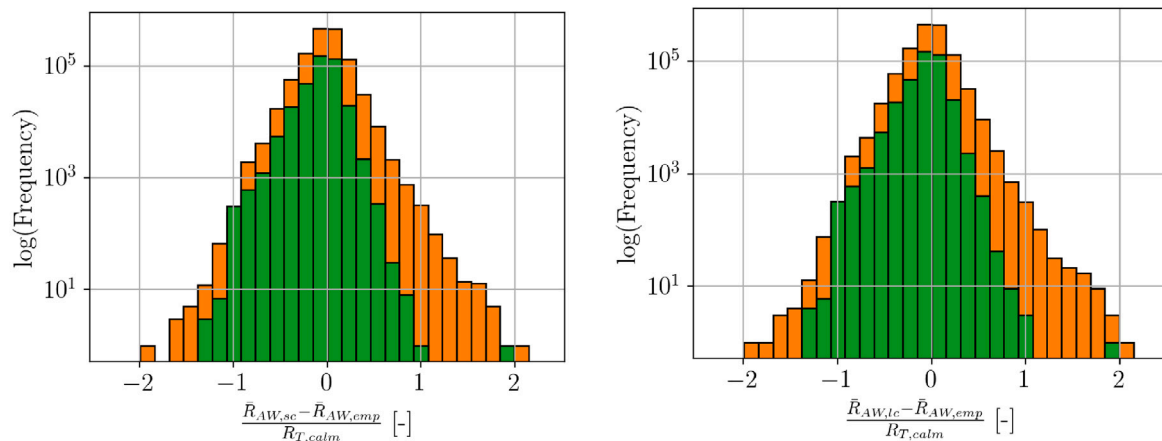
(a) Residual distribution in case of short-crested wave calculations and empirical added resistance.

(b) Residual distribution in case of long-crested wave calculations and empirical added resistance.

Fig. 19. Note that the ordinate is in logarithmic scale and that bimodal (orange) and unimodal conditions (green) are depicted separately in stacked histograms.

Acknowledgments

Appendix

A.1. Additional figures for the comparison of long- and short-crested wave calculations

The residual distribution $\frac{\bar{R}_{AW,sc} - \bar{R}_{AW,lc}}{R_{T,calm}}$ for all vessel classes is shown in box plots in Fig. 16 and a combined scatter plot as a function of H_s is depicted in Fig. 17.

A.2. Additional figures for the comparison of spectral and empirical calculations

The residuals between short-crested wave and empirical calculations are presented as a function of H_s in Fig. 18 for all 22 classes.

Also, the residual distributions in both short- and long-crested waves in comparison to the empirical added resistance are depicted in Fig. 19.

References

Aldous, L., Smith, T., Bucknall, R., Thompson, P., 2015. Uncertainty analysis in ship performance monitoring. Ocean Eng. 110, 29–38.

Andersen, I.M.V., 2013. Wind loads on post-panamax container ship. Ocean Eng. 58, 115–134. http://dx.doi.org/10.1016/j.oceaneng.2012.10.008.

Barstow, S., Bidlot, J., Caires, S., Donelan, M., Drennan, W., Dupuis, H., Graber, H., Green, J., Grønlie, Ø., Guérin, C., Gurgel, K.-W., Günther, H., Hauser, D., Hayes, K., Hessner, K., Hoja, D., Icard, D., Kahma, K., Keller, W., Wyatt, L., 2005. Measuring and analysing the directional spectrum of ocean waves. In: COST 714.

Bertram, V., 2016. Added power in waves – time to stop lying (to ourselves). In: Proc. of 1st Hull and Performance Insight Conference (HullPIC), Pavone.

Birk, L., 2019. Fundamentals of Ship Hydrodynamics: Fluid Mechanics, Ship Resistance and Propulsion. John Wiley & Sons.

Bitner-Gregersen, E.M., Bhattacharya, S.K., Chatjigeorgiou, I.K., Eames, I., Ellermann, K., Ewans, K., Hermanski, G., Johnson, M.C., Ma, N., Maisondieu, C., Nilva, A., Rychlik, I., Waseda, T., 2014. Recent developments of ocean environmental description with focus on uncertainties. Ocean Eng. 86, 26–46. http://dx.doi.org/10.1016/j.oceaneng.2014.69403.002.

Bos, M., 2018. An ensemble prediction of added wave resistance to identify the effect of spread of wave conditions on ship performance. In: Proc. of the Third Hull Performance and Insight Conference (HullPIC'18), Redworth.

Christensen, L.B., Thomas, G., Calleya, J., Nielsen, U.D., 2018. The effect of operational factors on container ship fuel performance.. In: Proc. of Full Scale Ship Performance (RINA), London, UK.

Copernicus, 2020. Copernicus climate change service information. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. (Accessed 02-03-2022). doi:https://cds.climate.copernicus.eu/cdsapp#!/home.

Dalheim, Ø.Ø., Steen, S., 2020a. Added resistance and speed loss of a ship found using onboard monitoring data. J. Ship Res. 64, 99–117.

Dalheim, Ø.Ø., Steen, S., 2020b. Preparation of in-service measurement data for ship operation and performance analysis. Ocean Eng. 212, 107703.

de Hauteclocque, G., Zhu, T., Austefjord, H., Bitner-Gregersen, E.M., Johnson, M., 2020. Assessment of global wave datasets for long term response of ships. In: Proc. of the 39th International Conference on Ocean, Offshore and Arctic Engineering (OMAE), Fort Lauderdale (Virtual).

DNV-GL, 2018. Class Guidelines: Wave Loads (DNVGL-CG-0130). Technical Report.

Faltinsen, O.M., Minsaas, K.J., Liapis, N., Skjordal, S.O., 1980. Prediction of resistance and propulsion of a ship in a seaway. In: Proc. of 13th Symposium on Naval Hydrodynamics. Tokyo (Japan): the Shipbuilding Research Association of Japan. pp. 505–529.

Goda, Y., 2000. Random Seas and Design of Maritime Structures. In: Advanced Series on Ocean Engineering, vol. 15, World Scientific.

Górski, W., Michniewicz, J., Szlendak, A., 2021. Using unsupervised machine learning for building ship performance reference model. In: Proc. of 6th Hull Performance and Insight Conference (HullPIC), Pontignano.

Guldhammer, H.E., Harvald, S.A., 1974. Ship Resistance - Effect of Form and Principal Dimensions. Danmarks Tekniske HøJskole, Akademisk Forlag.

Gunkel, J., Dietz, J., Purup, N., 2018. Trim optimization in an operational environment using existing sensor installations. In: Proc. of the 3rd Hull Performance and Insight Conference (HullPIC'18), Redworth.

Guo, B., Rognebakke, O., Tvete, H.A., Adal, C., Storhaug, G., Schmidt, M., Bruset, T., Prytz, G., 2021. Setting the standard for evaluation of in-service technical ship performance. In: Proc. of 6th Hull Performance and Insight Conference (HullPIC), Pontignano.

Harries, S., Dafermos, G., Kanellopoulou, A., Florean, M., Gatchell, S., Kahva, E., Macedo, P., 2019. Approach to holistic ship design – methods and examples. In: Proc. of 18th Conf. on Computer and IT Applications in the Maritime Industries (COMPIT), Tullamore.

Hasselmann, K., 1966. On non-linear ship motions in irregular waves. J. Ship Res. 10, 64–68.

Hasselmann, K., Barnett, T.P., Bouws, E., Carlson, H., Cartwright, D.E., Enke, K., Ewing, J., Gienapp, A., Hasselmann, D., Kruseman, P., et al., 1973. Measurements of wind-wave growth and swell decay during the joint north sea wave project (JONSWAP). Ergaenz. Dtsch. Hydrogr. Z. Reihe A.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N., 2018. ERA5 hourly data on single levels from 1979 to present. http://dx.doi.org/10.24381/cds.adbb2d47, Copernicus climate change service (C3S) climate data store (CDS). (Accessed on 24-Feb-2022).

Hogben, N., Cobb, F.C., 1986. Parametric modelling of directional wave spectra. In: Proc. of the 18th Offshore Technology Conference, Houston, Texas. pp. 489–498.

Hollenbach, K.U., 1997. Beitrag zur Abschätzung von Widerstand und Propulsion von ein- und Zweischraubenschiffen im Vorentwurf (Ph.D. thesis). Institut Für Schiffbau, Universität Hamburg, Hamburg, Germany.

Hollenbach, K.U., 1998. Estimating resistance and propulsion for single-screw and twin-screw ships. J. Ship Technol. Res. 45 (2), 72–76.

Holt, P., Nielsen, U.D., 2021. Preliminary assessment of increased main engine load as a consequence of added wave resistance in the light of minimum propulsion power. Appl. Ocean Res. 108, 102543.

Ikonomakis, A., Nielsen, U.D., Holst, K.K., Dietz, J., Galeazzi, R., 2021. How good is the STW sensor? An account from a larger shipping company. J. Mar. Sci. Eng. 9 (465).

Ikonomakis, A., Nielsen, U.D., Holst, K.K., Dietz, J., Galeazzi, R., 2022. Validation and correction of auto-logged position measurements. Commun. Transp. Res. 2, 100051.

IMO, 2021. IMO working group agrees guidelines to support new GHG measures. https://Www.Imo.Org/En/MediaCentre/PressBriefings/Pages/ISWG-GHG-8.Aspx (Accessed on 11-07-2022).

Iseki, T., 2010. Real-time analysis of higher order ship motion spectrum. In: Proc. of the 29th Int. Conf. on Ocean, Offshore and Arctic Engineering (OMAE'10), Shanghai, China.

ISO, 2015. ISO 15016: Ships and Marine Technology — Guidelines for the Assessment of Speed and Power Performance by Analysis of Speed Trial Data. International Organization for Standardization, Geneva.

ISO, 2016. ISO 19030: Ships and Marine Technology – Measurement of Changes in Hull and Propeller Performance. International Organization for Standardization, Geneva.

ITTC, 2014. Recommended procedures and guidelines - preparation, conduct and analysis of speed/power trials. In: International Towing Tank Conference.

ITTC, 2021. The specialist committee on ships in operation at sea (SOS): Final report and recommendations to the 29th ITTC. In: International Towing Tank Conference.

Kim, M., Hizir, O., Turan, O., Day, S., Incecik, A., 2017. Estimation of added resistance and ship speed loss in a seaway. Ocean Eng. 141, 465–476.

Kreitner, J., 1939. Heave, pitch and resistance of ships in a seaway. Trans. R. Inst. Nav. Archit. 81.

Kristensen, H.O.H., 2015. Revision of statistical analysis and determination of regression formulas for main dimensions of container ships based on data from clarkson. In: Project no. 2016-108: Update of Decision Support System for Calculation of Exhaust Gas Emissions. Technical University of Denmark.

Kristensen, H.O.H., Lützen, M., 2013. Prediction of resistance and propulsion power of ships. In: Project no. 2010-56, Emissionsbeslutningsstøttesystem Work Package 2, Report no. 04.

Lakshmynarayanana, P.A., Hudson, D., 2017. Estimating added power in waves for ships through analysis of operational data. In: Proc. of 2nd Hull Performance and Insight Conference (Hullpic), Ulrichshusen.

Lang, X., Mao, W., 2020. A semi-empirical model for ship speed loss prediction at head sea and its validation by full-scale measurements. Ocean Eng. 209, 107494.

Liu, S., Chen, H., Papanikoloau, A., 2022. A fast and transparent method for setting powering margins in ship design. In: Proc. of the 41st Int. Conf. on Ocean, Offshore & Arctic Engineering (OMAE), Hamburg.

Liu, S., Papanikolaou, A., 2020. Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. Ocean Eng. 206.

Lloyd, A.R.J.M., 1989. Seakeeping: ship behaviour in rough weather. In: Admiralty Research Establishment, Haslar, Gosport. Publisher Ellis Horwood Ltd, John Wiley & Sons, New York.

Michel, W.H., 1999. Sea spectra revisited. Mar. Technol. 36 (4), 211–227.

Mikkelsen, H., Shao, Y., Walther, J.H., 2022. Numerical study of nominal wake fields of a container ship in oblique regular waves. Appl. Ocean Res. 119, 102968. http://dx.doi.org/10.1016/j.apor.2021.102968.

Miratsu, R., Sasmal, K., Kodaira, T., Fukui, T., Zhu, T., Waseda, T., 2022. Evaluation of ship operational effect based on long-term encountered sea states using wave hindcast combined with storm avoidance model. Mar. Struct. 86, 103293.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., 2022a. Data-driven prediction of added-wave resistance on ships in oblique waves - a comparison between tree-based ensemble methods and artificial neural networks.. Appl. Ocean Res. 118, 102964.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., Liu, S., 2022b. Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading. Ocean Eng. 251, 111040.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., Storhaug, G., 2022c. Sea state identification using machine learning - a comparative study based on in-service data from a container vessel. Mar. Struct. 85, 103274.

Molland, A.F., Turnock, S., Hudson, D., 2011. Ship Resistance and Propulsion: Practical Estimation of Ship Propulsive Power. Cambridge University Press, New York.

Nakamura, S., Naito, S., 1977. Propulsive performance of a container ship in waves. J. Soc. Nav. Archit. Jpn. 15.

Nielsen, U.D., 2017. A concise account of techniques available for shipboard sea state estimation. Ocean Eng. 129, 352–362.

Nielsen, U.D., 2021. Spatio-temporal variation in sea state parameters along virtual ship route paths. J. Oper. Oceanogr. 1–18. http://dx.doi.org/10.1080/1755876X.2021.1872894.

Nielsen, U.D., Ikonomakis, A., 2021. Wave conditions encountered by ships—A report from a larger shipping company based on ERA5. Ocean Eng. 237, 109584. http://dx.doi.org/10.1016/j.oceaneng.2021.109584.

Nielsen, U.D., Ikonomakis, A., Dietz, J., 2022. Sea states encountered by ships in the maersk fleet - an assessment based on reanalysis data (ERA5). In: Proc. 15th Int. Symp. Practical Design of Ships and Other Floating Structures (PRADS'22), Dubrovnic, Croatia.

Nielsen, U.D., Johannesen, J.R., Bingham, H.B., Blanke, M., Joncquez, S., 2019. Indirect measurements of added-wave resistance on an in-service container ship. In: Proc. 14th Int. Symp. Practical Design of Ships and Other Floating Structures (PRADS'19), Yokohama, Japan.

Oh, S., Kang, D., Choi, S., Lee, D., Kim, B., 2018. Characteristics of wind speed and direction during sea trials of ships. In: Proc. of the 28th Int. Ocean and Polar Engineering Conference, Sapparo, Japan.

Park, D.M., Lee, J., Jung, Y., Lee, J., Kim, Y., Gerhardt, F., 2019. Experimental and numerical studies on added resistance of ship in oblique sea conditions. Ocean Eng. 186, 106070.

Park, D.M., Lee, J., Kim, Y., 2015. Uncertainty analysis for added resistance experiment of KVLCC2 ship. Ocean Eng. 95, 143–156.

Pierson, W., Moskowitz, L., 1964. A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii. J. Geophys. Res. 69 (24), 5181–5190.

Prpić-Oršić, J., Faltinsen, O.M., 2012. Estimation of ship speed loss and associated CO2 emissions in a seaway. Ocean Eng. 44, 1–10.

Prpić-Oršić, J., Sasa, K., Valčić, M., Faltinsen, O.M., 2018. Uncertainties of ship speed loss evaluation under real weather conditions.. In: Proc. of the 37th International Conference on Ocean, Offshore and Arctic Engineering (OMAE2018), Madrid, Spain.

Saettone, S., Taskar, B., Steen, S., Andersen, P., 2021. Experimental measurements of propulsive factors in following and head waves. Appl. Ocean Res. 111, 102639. http://dx.doi.org/10.1016/j.apor.2021.102639.

Schmode, D., Hympendahl, O., Gundermann, D., 2018. Hull performance prediction beyond ISO 19030. In: Proc. of the 3rd Hull Performance and Insight Conference (HullPIC'18), Redworth.

Shigunov, V., 2017. Added power in seaway. Ship Technol. Res. 64 (2), 65–75. http://dx.doi.org/10.1080/09377255.2017.1331953.

Shigunov, V., el Moctar, O., Papanikolaou, A., Potthoff, R., Liu, S., 2018. International benchmark study on numerical simulation methods for prediction of maneuverability of ships in waves. Ocean Eng. 165, 365–385.

Söding, H., Shigunov, V., 2015. Added resistance of ships in waves. Ship Technol. Res. 62 (1), 2–13.

Sogihara, N., 2021. Uncertainty assessment in ship performance evaluation by Monte Carlo simulation using onboard monitoring data. J. Jpn. Soc. Nav. Archit. Ocean Eng. 33, 35–45. http://dx.doi.org/10.2534/jjasnaoe.33.35.

Strøm-Tejsen, J., Yeh, H.Y.H., Moran, D.D., 1973. Added resistance in waves. Soc. Nav. Archit. Mar. Eng. Trans. 81, 109–143.

Taskar, B., Andersen, P., 2021. Comparison of added resistance methods using digital twin and full-scale data. Ocean Eng. 229, 108710.

Taskar, B., Regener, P.B., Andersen, P., 2019. The impact of propulsion factors on vessel performance in waves. In: Proc. 6th Int. Symp. on Marine Propulsors, Rome, Italy.

Tsujimoto, M., Shibata, K., Kuroda, M., Takagi, K., 2008. A practical correction method for added resistance in waves. Jpn. Soc. Nav. Arch. Ocean Eng. 8, 177–184.

Wang, J., Bielicki, S., Kluwe, F., Xin, G., Orihara, H., Kume, K., Oh, S., Liu, S., Feng, P., 2021. Validation study on a new semi-empirical method for the prediction of added resistance in waves of arbitrary heading in analyzing ship speed trial results. Ocean Eng. 240, 109959.

Yu, J.-W., Lee, C.-M., Seo, J.-H., Chun, H.H., Choi, J.-E., Lee, I., 2021. Comparative study on the prediction of speed-power-rpm of the KVLCC2 in regular head waves using model tests. Int. J. Nav. Archit. Ocean Eng. 13, 24–34.

Chapter 6

Added Resistance in the Context of Ship Performance Analysis

One everlasting problem plaguing the maritime industry is marine growth. The resulting added frictional resistance caused by the attachment of microorganisms on the ship's hull and propeller leads to a substantial increase in required engine power and hence carbon emissions. In contrast to $\bar{R}_{AW}$, biofouling contributes almost exclusively to the second term in Eq. 2.2. Moreover, the accumulated biofouling layer allows for the spread of invasive species. Thus, after IMO's regulation on ballast water treatment, it was the next logical step to initiate a joint project focusing on biofouling and its mitigation. For this reason, the GloFouling initiative has been launched by a consortium of intergovernmental bodies, such as IMO, GloFouling [38]. Stressing the significance of marine growth, a review from I-Tech – a publicly-held Swedish biotechnology company – found that 44% of the entire world fleet shows hard fouling of 10%, which in turn corresponds to a 36% increase in required propulsion power, I-Tech [49]. Under the consideration of typical saving potentials of energy-saving devices, it becomes clear that biofouling control and proactive hull cleaning will be a cornerstone for reducing the carbon footprint of the maritime industry in the coming years. Thus, the present chapter is dedicated to machine learning-driven vessel performance monitoring. Fundamentals of in-service ship analyses can be found in Carlton [20].

6.1 Introduction

The de-facto industry standard for in-service ship analyses – ISO 19030 [58] – plays an important role, primarily for paint manufacturers, for reproducibly and transparently assessing the degree of biofouling on ships. In fact, the standard appends an additional term to Eq. 2.3, namely, the added resistance due to fouling. ISO 19030 is known for a rigorous filtering procedure and presupposes a speed-power baseline of some kind. However, both availability and extrapolation capability of, e.g. sea trial results, appear questionable in practice, Berthelsen and Nielsen [9]. Additional drawbacks of ISO 19030 include the speed and draft dependency of the performance indicators as well as questionable applicability to noon reports due to the draconian filter methodology.

Ship operations are in the midst of a transition from the reporting stages to monitoring

and finally to predictive analysis, where the latter will be facilitated by data analysis and machine learning techniques. In fact, the increasing acquisition of in-service vessel performance data and higher availability of computational resources allow for machine learning-based vessel performance monitoring. Pedersen and Larsen [101] determine the effect of biofouling in parallel to a regression approach using a neural network for predicting shaft power in the case of a tanker. In contrast, Coraddu *et al.* [23] estimate the ship speed as well as the speed loss due to fouling using monitoring data from two tankers. The comparison of the estimated speed loss to the result of the ISO 19030 procedure is a core novelty of their work. However, the majority of studies in the contemporary literature regarding digital representations of ships consider a finite dataset at once, i.e. as steady, when in fact, the monitoring data is subject to concept drift *and* is recorded in a data stream. Concept drift refers to a time-dependent change in the relationship between features and target, which subsequently increases generalization error (model drift). For example, biofouling leads to a different power value over time for given speed and draft conditions, but this trend cannot be inferred from any feature. Given Eq. 3.2 and empirical risk minimization, the data-generating joint distribution can be subject to change, which makes the regression problem unsteady and violates the statistical IID (independent and identically distributed) assumption. Concept drift can have incremental or sudden rates. The former corresponds to biofouling accumulation and potential sensor drift in the context of hydrodynamic ship performance. On the other hand, sudden concept drift results from hull or propeller cleaning events or an extended idling period, leading to a considerable increase in marine growth. From a practical perspective, certain silicone-based fouling-release coatings show a relatively stable performance plateau for several years before a drop in performance, which can also be seen as a sudden concept drift scenario. Hence, quantifying the degree of concept drift as a proxy variable for biofouling is the overarching goal and novelty of the present chapter.
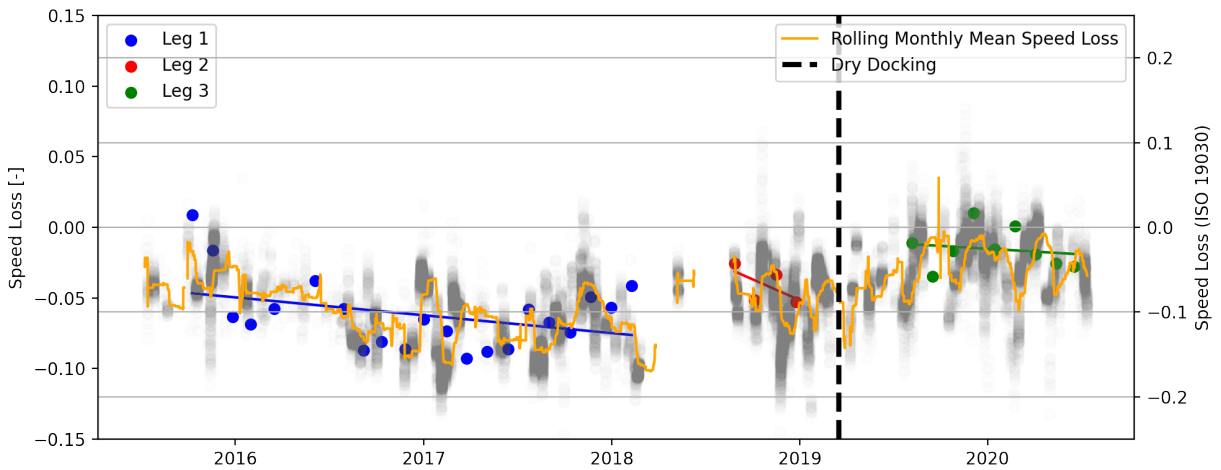

Figure 6.1: Speed loss caused by biofouling according the adaptive model on the primary axis and according to ISO 19030 on the secondary axis, taken from [C_2].

In-service sensor data is acquired in a datastream, which is an additional motivation for an *adaptive* training methodology. Thus, the proposed method draws inspiration from

a state-space approach presented in Antola *et al.* [6], where they merge noon reports and high-frequency data for performance monitoring. Moreover, it is herein hypothesized that there is, in fact, no sufficient feature combination that allows for capturing the performance decay in an explicit regression approach, which is due to the high complexity of marine growth patterns. Hence, $[C_2]$ and $[J_5]$ build on an implicit procedure for determining the performance loss at consistent conditions for a range of speeds. The overall concept has been partly developed during the author's external research stay at Wärtsilä Voyage and in collaboration with Hempel. In $[C_2]$, an initial proof-of-concept study was presented based on five years of high-frequency data from an LNG carrier. A neural network was re-trained incrementally in a scheduled rhythm for predicting STW. It turned out that the application of incremental learning can result in catastrophic forgetting, i.e. the loss of information obtained in previous training instances. For further details concerning catastrophic forgetting or interference, see Goodfellow *et al.* [40]. As a consequence, several hidden layers were frozen to make the model less susceptible to catastrophic forgetting, i.e. more stable. Overall, the performance estimate provided by the model found good qualitative agreement with the ISO 19030 speed loss indicator, which can be seen in Fig. 6.1. The speed loss according to the adaptive model is given on the primary axis, whereas the ISO 19030 estimates are on the secondary axis. The observed consistent offset is due to the use of different speed-power baselines. Another novelty of $[C_2]$ and $[J_5]$ is that the methodology derives an initial speed-power baseline in the warm-up period for the performance analysis. The fouling assessment is handled in a post-processing step and hence can also consider external data, such as sea trial speed-power baselines underlining the modularity of the method. Interestingly, in the second dry-docking interval in Fig. 6.1, a performance increase followed by a stable plateau can be observed, which results likely from the application of a fouling-release paint. Another aspect of both papers is the application of feature engineering methods, as presented in $[J_1]$. In fact, the analysis of collinearity in the field of machine learning-driven performance monitoring is considered rare. In order to study both the effect of data quality (or uncertainty) and to evaluate the individual models more rigorously, it was decided to pursue a simulation-based study in $[J_5]$ since, in this case, an actual ground truth of the added power due to biofouling is available. Therefore, a simulation framework for the KVLCC2 is set up, and the individual datasets are made publicly available for possible benchmarking and enhancing transparency, Mittendorf *et al.* [88]. Lastly, three different methods for incremental learning are compared to a default method, i.e. re-training without any restrictions. The application and systematic comparison of the individual methods is considered a core novelty of $[J_5]$.

6.2   Discussion

In $[C_2]$, an adaptive training methodology for a neural network has been proposed, which was then rigorously analyzed and extended in $[J_5]$ under the consideration of simulation data. Overall, it turns out that data quality – as already seen in $[J_4]$ – is paramount. However, not only the degree of stochastic variance but also the inherent steadiness of the dataset turned out to be vital when assessing vessel performance by machine learning. In the initial study in $[C_2]$, the speed loss has been determined, in contrast to the study in

$[J_5]$, where the power increase was estimated as a proxy variable of biofouling. Hence, it will be part of future research to determine which performance indicator is better suited in a machine learning-driven approach regarding robustness and accuracy. Moreover, the interval for re-training the model can be optimized and is ideally subject to automatic model drift detection. However, it becomes clear that not only the deteriorating model performance should be considered but also the availability of new unseen data. Even though the backshift (or lag) operator may be introduced into the data extraction procedure, it may lead to a lagging behavior of the corresponding performance estimate. In addition, it seems worthwhile to compare the derived speed-power baseline to an estimate by piece-wise regression using sensor data, e.g. Berthelsen and Nielsen [9], or sea trial curves. As pointed out in $[J_5]$, the overall methodology is also applicable for assessing the potential fuel savings of energy-saving devices, such as propeller boss cap fins, in comparison to established statistical methods by, e.g. Christensen [22].

Thus far, the present chapter had a dominant focus on the added power due to marine growth; however, the following paragraphs are directed towards applying the proposed methodology to added power due to the seaway – including added wave and wind resistance. A crucial advantage of a machine learning approach is that it imposes no assumptions on the overall calculation in contrast to Eq. 2.6. On the other hand, it is conceded that even though the prediction may show higher accuracy, the transparency and reproducibility are reduced when using a machine learning framework. For estimating the added *environmental* power, a dataset corresponding to scenario C (cf. $[J_5]$) is set up, i.e. showing realistic parameter distributions and sensor noise. In the simulation process, a required engine power for actual conditions and a corresponding value for calm water conditions is simulated, where the latter acts as a reference value for determining the ground truth of the added power.
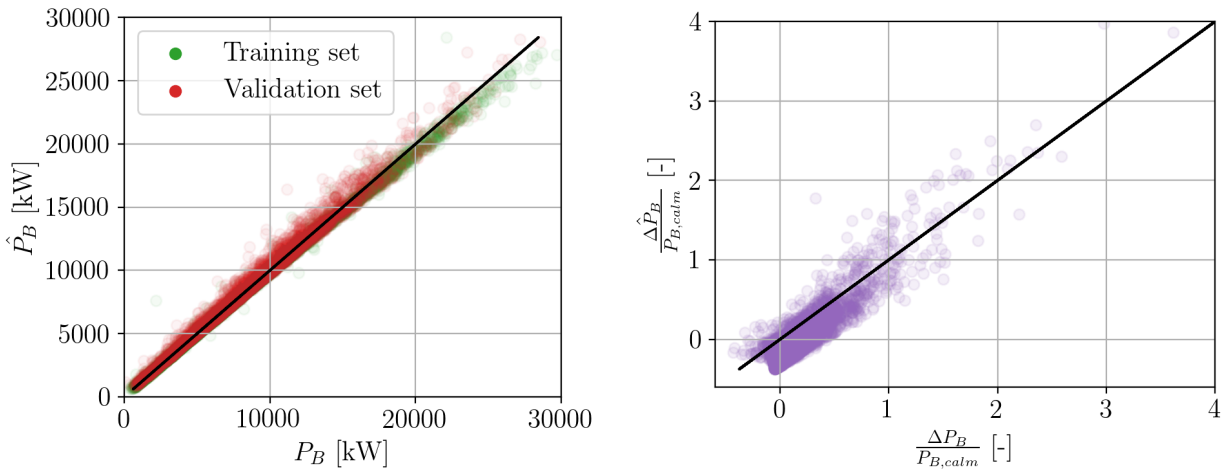
(a) Correlation plot for predictions on the training and validation set for predicting the required engine power.

(b) Comparison of derived relative added power by the model and simulated data (on the abscissa).

Figure 6.2: Correlation plots for the prediction of the added power due to wind and waves derived from simulated data for the KVLCC2, as shown in $[J_5]$.

It is noted that this dataset is not subject to concept drift because the roughness of the hull and propeller is kept constant at $k_S = 30\mu$m. In general, satisfactory accuracy and good generalization can be seen in Fig. 6.2a. It is noted that overhat indicates model predictions. Moreover, the variance increases slightly with the magnitude of the target variable (P_B), and erroneous samples can be identified for very large P_B values corresponding to instances with severe sea state conditions. Moreover, it is appreciated that all samples are bounded by the imposed engine limits ($P_B < 30,000$ kW); however, for a more realistic simulation framework, an actual engine load diagram with its speed/torque limits should be considered. As a side note the model underwent 30 epochs for training following the same training setup described in [J_5].

In contrast to the derived biofouling indicator, the added power due to the seaway is only computed for each instance and not for consistent (simulated) conditions, as in [J_5]. In Fig. 6.2b, the estimated added power normalized by the required engine power in calm weather $P_{B,calm}$ is shown in a correlation plot. It can be seen that even under the consideration of sea state parameter distributions biased towards lower values, the power increase due to service in wind and waves can be more than twofold in contrast to $P_{B,calm}$. This is likely caused by high sea states in combinations with low advance speeds. Moreover, the model exhibits somewhat biased results for lower magnitudes, despite a satisfactory correlation with the ground truth. Possible reasons for that are that machine learning models generally have a lower extrapolation capability, and the experience of precisely calm water conditions is doubtful given the used parameter distributions. Another cause could be that the same feature combination as in [J_5] is used, i.e. no information about the relative wind/wave direction and peak period. However, these parameters are highly relevant for estimating the added wind and wave resistance, as discussed earlier. Hence, it is believed that including these features may improve the obtained results. Still, when applying this approach to actual sensor data, the accuracy is expected to be lower due to multiple uncertainty sources, which are covered in [J_4]. However, it could be worthwhile to combine shaft power prediction and sea state identification in one monolithic model with multimodal input branches considering lower-frequent performance data and higher-frequent ship response data. Noting that when the ship sails mostly in diffraction-dominated seaways, it remains questionable whether the inclusion of ship motion information (causing radiation) will be beneficial for estimating the added power due to wind and waves.

Lastly, the possible application of curriculum learning, i.e. training a model in multiple training instances with increasing complexity, may be beneficial in the case of shaft power prediction for making the model more physics-informed. For example, the method could be pre-conditioned on simulation data with uniform distributions (scenario A in [J_5]) to explore the entire parameter search space evenly and thereby minimize model uncertainty due to low epistemic uncertainty within the training data. In the second step, actual in-service monitoring data can be used for re-training. Bengio *et al.* [8] provides a theoretical background for curriculum learning. Another aspect for mitigating catastrophic interference is the application of dilution (or dropout) in combination with other regularization techniques or alone, as shown by Goodfellow *et al.* [40]. In fact, it turned out that dilution is a robust and well-working countermeasure against catastrophic interference during incremental learning. Ultimately, we came full circle and found that in the end, it always comes down

to Eq. 3.2 and the difference between the empirical data, on which the model is trained, and the underlying data-generating process, which can be subject to change. The empirical training data is just a (possibly insufficient) representation of the data-generating process, leading to the statistical adage: Garbage in, garbage out.

[J_5] Capturing the effect of marine growth on ships by adaptive machine learning

The manuscript entitled "Capturing the effect of marine growth on ships by adaptive machine learning" has been submitted and is included herein as:

MITTENDORF, M.; NIELSEN, U.D.; BINGHAM, H.B. (2023). *Capturing the effect of marine growth on ships by adaptive machine learning.* (Under Review)

Capturing the effect of marine growth on ships by adaptive machine learning

Malte Mittendorf[a,*], Ulrik Dam Nielsen[a], Harry B. Bingham[a]

[a]*Department of Civil and Mechanical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

Abstract

Performance data from ships is subject to distributional shifts, sometimes referred to as concept drift. In this study, synthetic monitoring data is simulated for the KVLCC2, considering publicly available reference data and a semi-empirical simulation framework. Neural networks are trained to predict the required shaft power and to overcome the deterioration in model accuracy due to concept drift, three methods of incremental learning are applied and compared: (1) Layer freezing, (2) L_2 regularization, and (3) elastic weight consolidation. Furthermore, an implicit methodology for quantifying the changing hull and propeller performance is presented. In addition, a generic feature engineering framework is used for eliminating both collinear and insignificant features. In two investigations, sudden and incremental concept drift scenarios are examined, and the effect of different uncertainty categories on model performance is studied in parallel based on three different datasets. As a main finding, it is confirmed that data quality is of great importance for accurate machine learning-driven performance monitoring - even in simulated environments. Furthermore, the study shows that freezing layers during incremental learning proves to be the most robust and accurate, but it will be part of future work to examine this on actual sensor data.

Keywords: Hull biofouling, Concept drift, Incremental learning, Synthetic data, KVLCC2

1. Introduction

Seaborne transport is considered the most efficient transport mode, but it still emits approximately 3% of anthropogenic carbon emissions, as reported by IMO (2020). For reaching carbon-neutral shipping within this century, adopting a wide range of technologies and measures is indispensable - including energy-saving devices and alternative propulsion techniques. In the same

---

vein, biofouling mitigation and hull cleaning methods have received increased interest in recent years - not least through the GloFouling[1] initiative, which aims at limiting the spread of invasive species and abating carbon emissions caused by the accumulation of marine growth resulting in increased hydrodynamic resistance. In fact, a 55% increase in CO_2 emissions has been reported for a larger tanker with heavy calcareous fouling in a study presented in GloFouling (2022). This in turn underscores the importance of adequate hull and propeller maintenance for achieving fuel and corresponding emission savings, as also pointed out by Adland et al. (2018). For this reason, analysis of a ship's performance deterioration due to biofouling has been an active research field ever since. However, the need for performance monitoring has been amplified due to the recently adopted regulation of the Carbon Intensity Indicator (CII) (IMO, 2021). The CII is controversial since it is aimed at minimizing carbon emissions relative to the *potential* transport work, which is determined by considering the maximum (or design) cargo capacity instead of the actual loading condition, Wang et al. (2021). In addition to increased carbon emissions, Liu et al. (2021) showed that the added frictional resistance due to fouling might - in extreme cases - contribute to the loss of maneuverability in adverse weather conditions, which may pose a safety concern and, thus, has to be considered when determining the minimum propulsion power for blunt-type vessels, such as bulk carriers and tankers, cf. IMO (2017).

1.1. Literature review

Driven by international regulations and digitalization, availability (and quality) of in-situ data from ships increased significantly during the last decade; thus, machine learning-based monitoring experienced growing interest and adoption. Traditionally, the in-service performance of vessels has been analyzed by superimposing the individual resistance components determined by, e.g. empirical or numerical methods, as shown by Andersen et al. (2005) and Hansen (2011). Significant experience has been acquired in this field, and transparency as well as reproducibility are key advantages of these methodologies. Moreover, associated approaches are even applicable in case of scarce data, e.g. for noon reported data. However, the traditional way of performance monitoring is limited by its inherent approximations, which may show a lack of accuracy - particularly in off-design conditions, as reported by Berthelsen and Nielsen (2021) in a study of data from a large group

---

[1] GloFouling has been initiated by the International Maritime Organization (IMO), the United Nations Development Programme (UNDP) and the Global Environment Facility (GEF).

of tankers. In addition, Lang et al. (2022) conclude that a semi-empirical framework performs far worse for shaft power prediction than multiple machine learning methods in the case of two different vessels. Having said this, Aldous et al. (2015) showed that automatic data acquisition via a sensor framework onboard a ship and the resulting high-frequency data exhibit less uncertainty than manually collected data. Hence, the amount of sensor data from ships has been increasing in recent years, which enables establishing digital representations of vessels or in other words *cyber-physical systems*.

Most machine learning (ML) studies on ship performance monitoring rely on a data fusion concept, i.e. data from several different sources and domains are combined into one dataset. Some specific studies are listed in the following: Pedersen and Larsen (2009) merge noon report, high-frequency monitoring, and metocean hindcast data for estimating shaft power and the effect of biofouling for a tanker using neural networks. Petersen et al. (2012) employ neural networks for estimating the main engine power of a RoPax vessel based on several months of publicly available, high-frequency monitoring data. Coraddu et al. (2019) use an extreme learning machine, i.e. a neural network trained by a matrix inversion technique, for determining the speed loss due to fouling of two Handymax tankers in direct comparison to the de facto industry standard for vessel performance monitoring ISO 19030 (ISO, 2016). Laurie et al. (2021) compare multiple machine learning models for predicting the propulsion power of several container vessels and deriving the power increase due to fouling by imposing a time dependency. A key aspect of their work is the evaluation of feature importance, i.e. the sensitivity of the individual sensor recordings to the prediction model. Tsompopoulou et al. (2022) use a synthetic ship performance dataset from a bulk carrier for an uncertainty-aware estimate of the speed-power-relationship under varying conditions using probabilistic deep learning. It is emphasized that generalization to unseen operational conditions is negatively affected by the model's inductive bias, i.e. the model's assumptions based on the training data. Mittendorf et al. (2022c) show that performance data is subject to concept drift mainly caused by marine growth and thus violates pivotal assumptions of ML. Consequently, a neural network is trained *adaptively* on high-frequency data of an LNG carrier, and a proxy variable for biofouling is derived implicitly. It is found that the proposed methodology can reflect the temporal decay in vessel performance satisfactorily since the estimates have trends similar to an ISO 19030 indicator.

3

1.2. Motivation and objective

Following Tsompopoulou et al. (2022), synthetic ship performance data is simulated; thus, an actual ground truth of the fouling-related performance decrease is available. This is a decisive advantage compared to studies using in-service vessel data since it allows for a more rigorous model assessment. The KVLCC2 (KRISO Very Large Crude Oil Carrier 2) is taken as a case ship, and reference data, e.g. experimental propeller data, from the literature is adopted in a semi-empirical simulation framework. A key aspect is to consider a wide variety of different operating conditions – including environmental influences – and the effect of biofouling which is introduced to both hull and propeller through the 1978 ITTC (International Towing Tank Conference) method (ITTC, 2017), i.e. by an equivalent sand roughness, as proposed by Schultz (2007). However, this study is limited to lighter fouling conditions and disregards heavy barnacle or seaweed fouling entirely. In the study by Mittendorf et al. (2022c), several hidden layers of the neural network are frozen to preserve information from previous training instances. In the current study, more advanced incremental learning methods are applied and compared. The overarching goal of this paper is to determine the most accurate and practical framework among several methods. Hence, their sensitivity to data quality receives particular interest since in-service ship data is often characterized by larger variance, i.e. reduced data quality. In addition, the effect of different uncertainty sources on model behavior and estimated hull performance proxy variable are examined. This study investigates three scenarios, each characterized by a different type of uncertainty. Moreover, sudden and incremental, i.e. prolonged continuous, concept drift are examined in an isolated way. Similarly to Tsompopoulou et al. (2022) and Petersen et al. (2012), the individual datasets are made publicly available for reproducibility and transparency reasons; see Mittendorf et al. (2023).

1.3. Organization of the paper

The sections of the present paper are ordered in the following way: In the next Sec. 2, the calculation of synthetic monitoring data and the individual datasets will be presented. In Sec. 3, the applied incremental learning methods, as well as the implicit determination of the fouling proxy variable, will be described. Moreover, the obtained results are depicted and discussed in Sec. 4. In the final Sec. 5, the paper's main findings will be reiterated, and opportunities for future work are outlined.

2. Data generation

In contrast to slender ships, the hydrodynamic performance of blunt-type vessels is generally more affected by marine growth, which is due to their lower forward speeds, i.e. a greater viscous resistance component, and their operational profile with longer idling periods. An extreme example of the latter was the COVID-19 pandemic, where a large number of tankers were utilized for temporary oil storage, which led to increased biofouling on the corresponding vessels (Liu et al., 2021).

2.1. Case ship: KVLCC2

Similar to Tsompopoulou et al. (2022), we define a simulation framework for the generation of performance data pertaining to the KVLCC2. This vessel has been the subject of multiple numerical and experimental benchmark studies, e.g. SIMMAN (2014). Hence, plenty of reference data is available from the literature, on which the present work will build. The KVLCC2 as well as its propeller are designed for an advance speed of $U = 15.5$kts, which corresponds to $Fn = \frac{U}{\sqrt{gL_{pp}}} = 0.142$. In addition, the ballast condition is taken from Park et al. (2016) with a draft in the aft position $d_a = 11.9$m and draft in the fore position $d_f = 6.7$m. The main particulars and the hull shape itself are presented in Tab. 1 and Fig. 1, respectively. It is noted that the open water propeller curves (in model scale) are appended to this paper in appendix A1.

Table 1: Main particulars of hull and propeller of the KVLCC2 in design conditions for $Fn = 0.142$, cf. SIMMAN (2014).

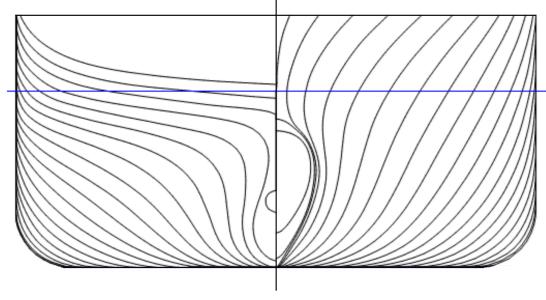| Definition	Symbol	Unit	Magnitude
Length between perpendiculars	L_{pp}	[m]	320.0
Beam	B	[m]	58.0
Mean draft	d_m	[m]	20.8
Block coefficient	C_B	[-]	0.808
Propeller diameter	D	[m]	9.86
Blade number	Z	[-]	4
Expanded blade area ratio	A_e/A_0	[-]	0.431
Pitch–diameter ratio	P/D	[-]	0.721

Figure 1: Body plan of the KVLCC2 with the design water line (blue line).

¹¹⁰ *2.2. Shaft power calculation*

¹¹¹   Initially, several assumptions are made, which include disregarding cavitation, propeller slip,

¹¹² currents and swell waves. Moreover, additional resistances due to shallow water and steering are

¹¹³ neglected entirely, since these particular occurrences would typically be filtered in reality anyhow,

¹¹⁴ as was shown in Mittendorf et al. (2022c). It is also assumed that biofouling is homogeneous

¹¹⁵ and can be approximated by the equivalent sand roughness, as proposed by Schultz (2007). In

¹¹⁶ addition, the degree of fouling is of similar extent on both ship hull and propeller, i.e. $k_S = k_P$. For

¹¹⁷ the determination of the added resistance due to roughness, we make use of the ITTC'78 power

¹¹⁸ prediction method (ITTC, 2017). Herein, the total hull resistance R_T, which has to be overcome by

¹¹⁹ the propeller and main engine, is decomposed into three main contributions, as can be seen in Eq.

¹²⁰ (1).

$$R_T = R_{T,calm} + \bar{R}_{AW} + R_{wind} \tag{1}$$

¹²¹   The total calm water resistance $R_{T,calm}$ is calculated by the Hollenbach (1998) method, which

¹²² splits the resistance into frictional and residual resistance, i.e. $R_{T,calm} = R_F + R_R$. The residual

¹²³ component is determined using a large set of empirical parameters and also includes the resistance of

¹²⁴ appendages, such as hull bossings. The frictional resistance is calculated by the ITTC'57 correlation

¹²⁵ line and the roughness allowance taken from ITTC (2017) adds the influence of k_S. Both the

¹²⁶ kinematic viscosity ν and the density ρ of seawater are determined by its temperature ϑ_{sea} for a

¹²⁷ constant salinity of 35g/kg. Additionally, it is noted that S corresponds to the wetted surface area
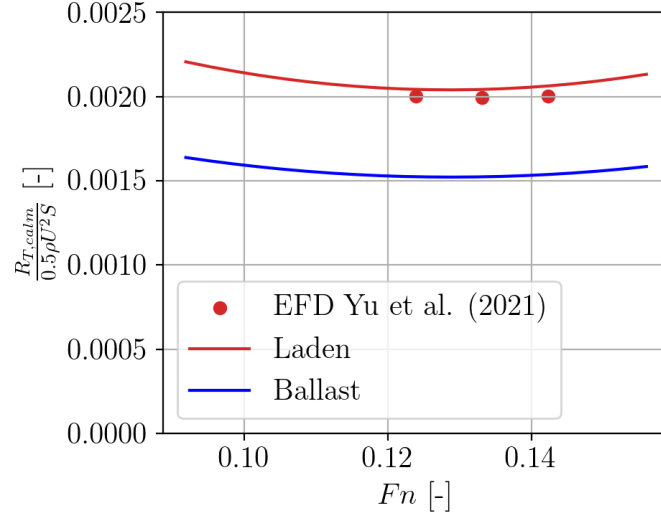
¹²⁸ of the ship.

6

Figure 2: The non-dimensional (full-scale) calm water resistance in laden and ballast conditions as a function of the Froude number following Hollenbach (1998).

In Fig. 2, good agreement between non-dimensional estimated and experimental calm water resistance can be seen in the proximity of design conditions, i.e. laden draft and $Fn = 0.142$. However, it is uncertain, whether the method can reflect $R_{T,calm}$ at ballast draft and below design speed as accurately, since no experimental data is available for these regimes. For keeping resemblance to Mittendorf et al. (2022c), conditions with $U < 5$kts (or $Fn < 0.046$) are not considered herein.

The mean added resistance in short-crested waves $\bar{R}_{AW}$ is determined from the spectral formulation presented in Eq. (2), i.e. the product of a quadratic transfer function, R_{AW}, and a directional wave spectrum, E, is integrated over both wave frequency ω and wave direction β.

$$\bar{R}_{AW} = 2 \int_0^{2\pi} \int_0^\infty \frac{R_{AW}(\omega, \beta)}{\zeta^2} E(\omega, \beta) d\omega d\beta \tag{2}$$

It Eq. 2, ζ denotes the wave amplitude. The transfer function is determined using the method by Mittendorf et al. (2022d), which is an adaptation of Liu and Papanikolaou (2020). The agreement with experimental data can be seen in Fig. 3 and it stands out that the semi-empirical formulation is able to approximate the non-dimensional added resistance not only in laden conditions, but also in ballast conditions, which are characterized by a notable trim angle. In view of Fig. 3, it is said that the agreement in the resonance region is considered as satisfactory in both laden and ballast conditions, but larger deviations are notable in shorter relative wavelengths. Note that g is the

¹⁴⁵ gravitational acceleration and λ denotes the wavelength.

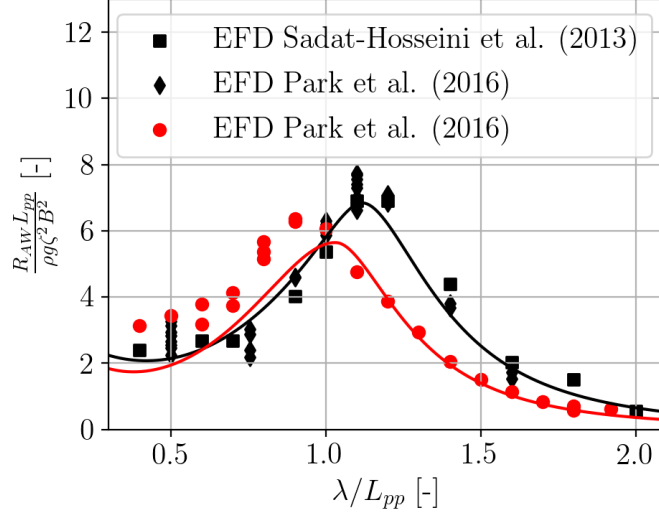

Figure 3: Non-dimensional added-wave resistance transfer function for $Fn = 0.142$ and $\beta_0 = 180$deg. in laden (black) and ballast (red) conditions as a function of the relative wave length, according to Mittendorf et al. (2022d).

¹⁴⁶ The directional wave energy spectrum $E(\omega, \beta)$ is approximated by a unimodal Pierson and
¹⁴⁷ Moskowitz (1964) spectrum combined with a cos-squared spreading function. For this reason, 36
¹⁴⁸ discrete wave directions $\beta \in [-\pi, \pi]rad$ and a frequency discretization of 50 discrete frequencies
¹⁴⁹ $\omega \in [0.01, 2\pi]$rad/s are employed. The significant wave height H_s, peak period T_p, mean relative
¹⁵⁰ wave heading β_0 and the spreading parameter ($s = 10$ for wind waves) are needed for the calculation
¹⁵¹ of the directional wave spectrum. Additional mathematical details are described in DNV-GL (2018).
¹⁵² It is emphasized that only wind waves are assumed and a direct relationship between the
¹⁵³ significant wave height and the absolute wind speed V_w is employed. Thus, the empirical formula
¹⁵⁴ $H_s = 0.115V_w^{1.41}$ from Shigunov (2017) is used. Moreover, based on the assumption of wind waves,
¹⁵⁵ wave and wind directions coincide. However, it is stressed that the directions have opposing
¹⁵⁶ definitions. In fact, for the wave heading angle $\beta_0 = 180$deg. corresponds to head and 0deg. to
¹⁵⁷ following directions, whereas it is vice versa for wind. The wind resistance is calculated following
¹⁵⁸ Eq. (3) after mapping wind speed V_w and direction ψ from the absolute into the relative domain
¹⁵⁹ (with subscript r). The required procedure is described in ISO 15016 ISO (2015). Moreover, a set
¹⁶⁰ of wind resistance coefficients ($C_{AA} = -C_X$) for a larger tanker are adopted from ISO 15016 and
¹⁶¹ presented in appendix A1.

$$R_{wind} = \frac{1}{2}\rho_{air}C_{AA}(\psi_r)V_{w,r}^2 A_x - \frac{1}{2}\rho_{air}C_{AA}(0\text{deg.})U^2 A_x \tag{3}$$

In Eq. (3), the air density is $\rho_{air} = 1.22\text{kg/m}^3$ and the frontal area of the vessel A_x is calculated according to Kristensen and Lützen (2013). It is noted that different sets of wind resistance coefficients are available for laden and ballast conditions.

For accurate modeling of hull and propeller interaction, both wake fraction and thrust deduction coefficients have been adopted from experimental data presented in Yu et al. (2021). In order to extrapolate propeller data from model to full scale, the Reynolds number correction in ITTC (2017) is used. In fact, the roughness due to fouling k_S is introduced the same way. Trivially, increased roughness increases the frictional losses of the propeller and thus reduces its efficiency. The determination of required shaft speed n as well as corresponding shaft torque Q are determined under thrust identity and the applied procedure together with the required equations is laid out in greater detail in Liu et al. (2021). Lastly, the engine's brake power P_B is obtained using Eq. (4), where η_M is the transmission efficiency including the losses of, e.g. bearings, and K_Q corresponds to the torque coefficient,

$$P_B = \frac{2\pi n Q}{\eta_M} \text{ with } Q = K_Q \rho n^2 D^5 \tag{4}$$

In the following, two different fouling scenarios are depicted in comparison to a clean hull ($k_S = 30\mu\text{m}$) for the same reference conditions, i.e. $H_s = 1.0\text{m}$, $T_p = 5.0\text{s}$ $V_w = 4.6\text{m/s}$ in head winds and waves at design draft. Moreover, according to Schultz (2007) only lighter biofouling conditions are investigated with $k_S = 30\mu\text{m}$ as applied coating, $k_S = 150\mu\text{m}$ reflects lighter slime and $k = 300\mu\text{m}$ shows the effect of heavy slime. As can be seen, extreme calcareous fouling conditions have been disregarded entirely herein.
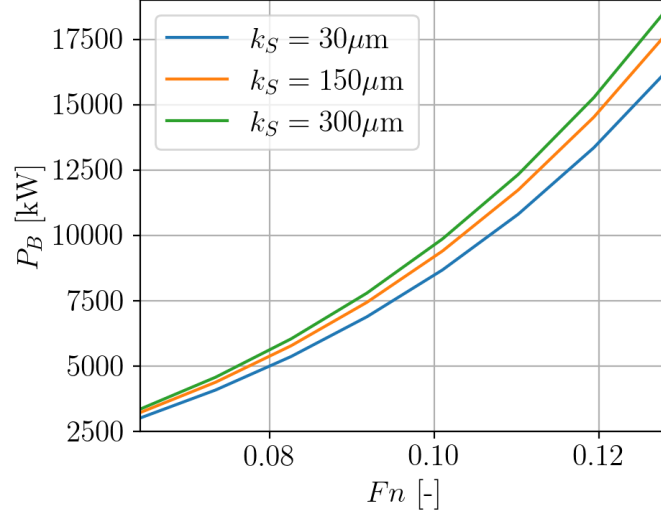
Figure 4: Three speed-power-curves for the same set of conditions, but different degrees of biofouling indicated by k_S.

Following the classification by Schultz (2007), it stands out in Fig. 4 that even lighter fouling ($k_s \leq 300$) conditions lead to a notable power increase. In addition, it is mentioned that fouling can be theoretically determined by k_S, but the actual growth pattern is highly complex and subject to both known and unknown parameters. A number of them are discussed in Yebra et al. (2004). Lastly, the herein defined simulation framework requires the following quantities as inputs: U, d_a, d_f, H_s, T_p, β_0, ϑ_{sea}, k_S and the ship main particulars.

2.3. Individual datasets

Similar to Mittendorf et al. (2022b), a wave breaking constraint is incorporated in order to enforce moderate wave steepness. The maximum steepness in irregular wave $\epsilon = \frac{2\pi H_s}{g T_z^2}$ is used as a threshold and presented in Eq. (5) (DNV-GL, 2018).

$$\epsilon = \begin{cases} \frac{1}{10} & \text{for } T_z \leq 6s \\ \frac{1}{15} & \text{for } T_z \geq 12s \end{cases} \tag{5}$$

It is noted that the relationship $T_p = 1.41 T_z$ holds due to the use of the Pierson Moskowitz spectrum. Also, T_z denotes the zero up-crossing period and linear interpolation is performed in between the two thresholds in Eq. (5). Furthermore, limitations of a potential main engine are imposed at 30,000kW and 82rpm as two additional constraints, following Taskar and Andersen (2019). As already mentioned, any speeds $U < 5$kts are dropped.

In the present work, random sampling has been applied, which completely disregards that ship performance data is temporally coherent. In both scenarios, 30,000 initial random samples are defined, but the imposed constraints (mainly the wave steepness constraint) had a significant effect and the final sample size is in the range of 15,000. Herein, two different concept drift scenarios will be investigated: (1) Sudden and (2) Incremental concept drift. The former is attributed to hull/propeller cleaning or sensor calibration, whereas the latter may result from the time-dependent accumulation of biofouling or from sensor drift. From a study presented in GloFouling (2022), it was found that in equatorial regions, an increase from newly applied paint ($k_S = 30\mu$m) to heavy slime ($k_S = 300\mu$m) is possible within one year. Hence, in case of sudden concept drift, three different cases of roughness were calculated for the same set of parameters using the k_S values from Fig. 4, i.e. 30, 150 and 300μm. Conversely, for incremental concept drift, a linear increase of 0.009μm per sample was introduced. Thus, from a statistical point of view, the samples in the sudden concept drift scenario are, in fact, stationary considering instances with the same roughness in a step-wise approach. In contrast, this assumption does not hold for an incremental concept drift, since roughness is linearly increased. Due to this transient character, the entire dataset is binned into smaller windows, in which it is assumed that the samples are quasi-stationary.

The synthetic data is free of any distributional shifts due to changes in the operational profile. In addition, any disturbing influences caused by acceleration (due to speed or course changes), wave drift or significant rudder angles are disregarded entirely. From the operational profile of a VLCC, it is determined that 50% of the samples are in laden and 50% in ballast conditions. In addition, the mean wave encounter angle is uniformly distributed in the range $\beta_0 \in [0, 360]$deg. for the sake of simplicity. The simulation of the remaining parameters is described for each dataset individually.

A crucial advantage of using synthetic data is isolating different uncertainty categories and studying their effect on the individual models. Initially, two uncertainty categories are defined: (1) *Epistemic uncertainty* results from limited data availability, i.e. unobserved conditions, and will eventually vanish with increasing the sample size. (2) *Aleatoric uncertainty*, on the other hand, results from the stochastic nature of the data and is unaffected by data availability. Herein, three different datasets per concept drift scenario will be simulated:

1. Following Mittendorf et al. (2022b), dataset $\mathcal{D}_A$ is characterized by uniform parameter distributions and thus exhibits no epistemic uncertainty. Moreover, the samples resemble

deterministic conditions due to the employed simulation framework and therefore show no

aleatoric uncertainty either. In fact, the following definition ranges have been used for

simulation: $H_s \in [0.1, 7.4]$m, $T_p \in [2.1, 15.5]$s and $U \in [6, 16]$kts. Besides, the sea water

temperature was held constant at 15°C. In relation to the other two datasets, $\mathcal{D}_A$ shows an

exaggerated likelihood of severe sea states (due to the uniform parameter distribution of H_s)

and hence large variance due these sea states.

2. Dataset $\mathcal{D}_B$ is consists of realistic parameter distributions and an unobserved parameter (by

the model), i.e. the sea water temperature, which has a minor impact on the total resistance

due to the change in ν. In other words, this dataset is characterized by epistemic uncertainty.

In addition, variance has been introduced to the draft (before simulation) caused by the

change in sea water density or water ballast for, e.g. the possible avoidance of slamming

in ballast conditions. For this reason, a normal distribution around the draft from ballast

and laden conditions with a standard deviation of 0.4m is assumed. In fact, the realistic

distributions of H_s, T_p, U and ϑ_{sea} are adopted from 5 years of data from a gas carrier in

Mittendorf et al. (2022c) and are depicted in Fig. 5. The used Probability Density Functions

(PDF) as well as the fitted parameters are included in the respective caption. As a side note,

the scipy[2] library has been used for fitting the parametric PDFs. In contrast to $\mathcal{D}_A$, it can be

seen that the PDFs of H_s and T_p are heavily skewed towards lower values. Moreover, it has

to be stressed that the distributional mean of U, as obtained from Mittendorf et al. (2022c),

has been reduced to 12.3kts, in order to match the service speed of a typical VLCC.

3. Dataset $\mathcal{D}_C$ is characterized by epistemic *and* aleatoric uncertainty due to the injection of

sensor noise according to Sogihara (2021) into the previous dataset $\mathcal{D}_B$. The sensor accuracy

found in this survey is presented in Tab. 2. Since, no accuracy value for torque Q was

available, the corresponding P_B accuracy was taken. It is also mentioned that no noise was

introduced to the draft measurements, which is a fair assumption when data from the loading

computer is utilized, since draft measurements obtained from sensors usually show larger

variance (in sailing conditions). From Tab. 2, it turns out that the last three environmental

parameters (ψ_r, H_s, β_0) show a homoscedastic behavior, in contrast to the other quantities,

where the extent of variance depends on the parameter's magnitude. In addition, it shows
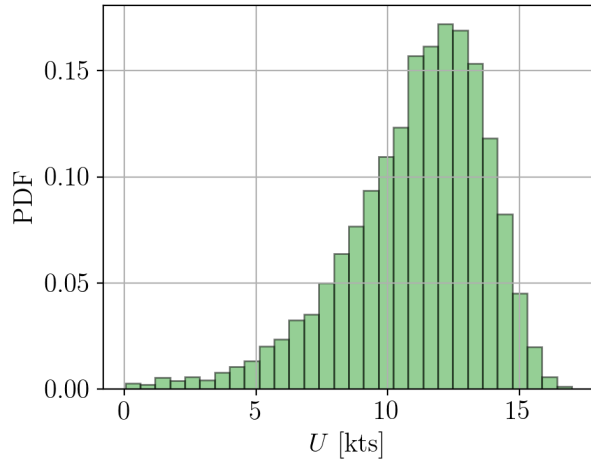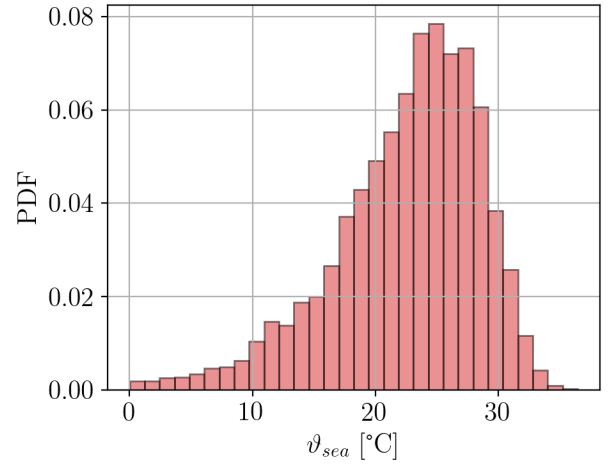
---

[2]https://scipy.org/
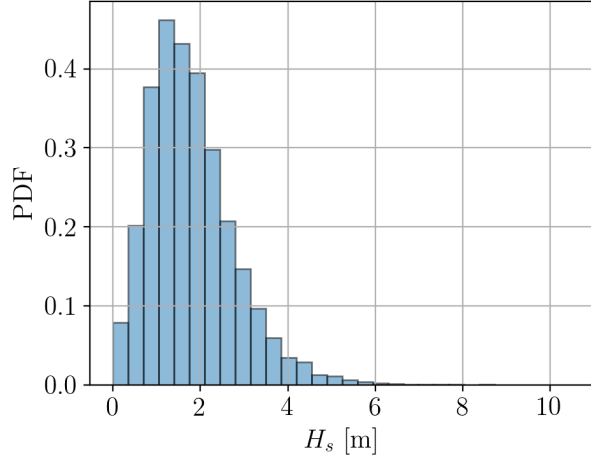
(a) Gumbel (right) distribution of the significant wave height with $\mu = 1.35$ and $\sigma = 0.785$.
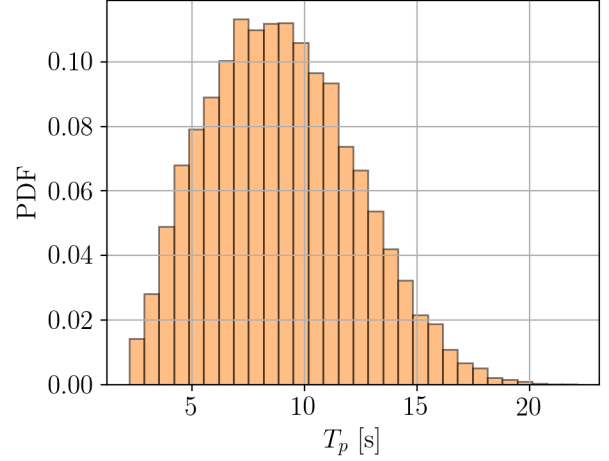
(b) Exponentiated Weibull distribution of the peak period with $a = 0.655$, $c = 2.895$, $\mu = 2.106$ and $\sigma = 9.013$.

(c) Gumbel (left) distribution of the ship advance speed with $\mu = 12.3$ and $\sigma = 2.11$.

(d) Gumbel (left) distribution of the sea temperature with $\mu = 25.3$ and $\sigma = 4.75$.

Figure 5: Probability density functions (PDF) of the individual quantities. It is noted that μ denotes the mean and σ is the standard deviation.

that the wind measurements are generally of lower quality. The shown accuracy of sea state data is considered as presumably unrealistic, considering the reported accuracy of an X-band wave radar in Mittendorf et al. (2022e).

Table 2: Measurement accuracy of sensor equipment, taken from Sogihara (2021).

Sensor reading	Equipment	Accuracy
U	Doppler log	1%
rpm	Revolution counter	1%
P_B	Torsion meter	0.5%
$V_{w,r}$	Anemometer	5%
ψ_r	Anemometer	5deg.
H_s	Hindcast, wave radar	0.1m
β_0	Hindcast, wave radar	5deg.

3. Methodology

In monitoring problems, the dataset is typically not fixed, but samples are acquired as a sequential data stream. Mathematically, a data stream is defined as a possibly infinite sequence $Z = (z_1, \ldots, z_t)$ of tuples $z_t = \{\mathbf{x}_i, y_i\}$, where t represents time, and y_i is the target (herein P_B) for a given feature vector $\mathbf{x}_i$ (herein operational parameters). In IoT (Internet of Things) practice, data and, thus, the learning task is not time-invariant. Assuming so may eventually lead to a mismatch between training and application instances over time. This phenomenon is known as *concept drift* and can be expressed theoretically by a joint distribution p with a dependency on t, i.e. $p_{t_0}(\mathbf{x}, y) \neq p_{t_1}(\mathbf{x}, y)$. However, in ML applications, the assumption of IID (independently and identically distributed) and stationary samples is indispensable (sine qua non) and, therefore, generally implicitly assumed. Concept drift has different patterns usually associated with sudden, incremental, and/or reoccurring drift rates (Losing et al., 2018). In fact, all three types can be observed in ship performance analysis: (1) Hull and propeller fouling and sensor drift correspond to incremental concept drift, (2) hull/propeller cleaning, sensor calibration, or an extensive idle period may lead to sudden concept drift. (3) When considering several dry-docking intervals, a reoccurring pattern in ship performance will stand out. As a side note, it is stressed that hull deterioration can be both temporary and permanent, e.g. due to corrosion or mechanical damages. After all, continual learning is needed to alleviate the decay in model accuracy caused by concept drift.

14

3.1. Incremental learning

Continual learning generally acts as an umbrella term for a wide range of different methods, and in the present context, an incremental learning paradigm is considered to correspond to repeated (and batch-wise) transfer learning on newly acquired data. For more detailed information about continual learning using neural networks, cf. Parisi et al. (2019). Broadly speaking, incremental learning resembles human intelligence and memory consolidation, i.e. the time-dependent transformation of sensory information into long-term memory (Kirkpatrick et al., 2017). Similar to human beings, the herein proposed model learns in a *sequential* manner, which is depicted in Fig. 6, where the cyclical methodology of incremental learning in the present context is visualized.
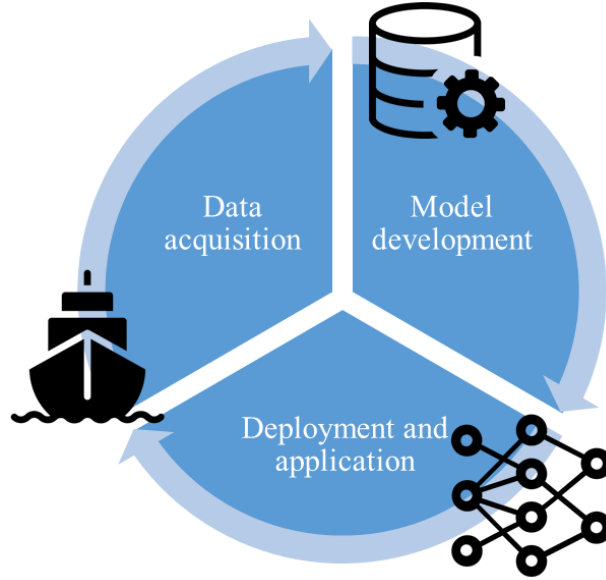

Figure 6: Concept of incremental learning applied to ship performance monitoring.

In Fig. 6, it is illustrated that data is collected onboard the vessel in the initial stage, while the second stage contains data preprocessing, which includes outlier detection and filtering for, e.g. unsteadiness or shallow water influence. Moreover, the model - a neural network - is trained in a regression task for predicting the required engine power P_B in this stage. In the final stage, the model is applied and used for, e.g. decision support both onboard and ashore. It is stressed that the initial and the final stages happen in parallel in actual practical applications. In fact, the presented methodology has a major resemblance to the procedure of MLOps, which is a coinage

15

made of ML and DevOps[3], which refers back to the maintenance of software products. Ultimately, the practical relevance of the present work is emphasized since the digital representation of the ship is constantly updated, and the drop in model accuracy is mitigated. As such, the present study stands out compared to many other similar studies in the literature, most of which treat the gathered monitoring data as time-invariant by considering the *entire* dataset at once.

Incremental learning is considered a sub-field of deep learning, but it is emphasized that there are similar methods in different domains, such as the incremental decision tree. Similar to traditional machine learning, deep learning techniques are optimized for application in stationary and self-contained environments. Artificial Neural Networks (ANN) have a resemblance to dendrites of human brains and are commonly used as universal function approximators. Mathematically, the stacking of affine functions (or layers) eventually leads to the universal approximation theorem (in theory). It is noted that the recent breakthrough in deep learning is not only attributed to deep architectures but even more so to special layer types, such as recurrent or convolutional layers. Theoretical details of deep learning can be found in Goodfellow et al. (2016). In Eq. (6), it can be seen that an ANN produces a mapping from feature vector $\mathbf{x}$ to dependent variable y or rather its approximation $\hat{y}$.

$$\hat{y} = f_L[\Theta_L, f_{L-1}(\Theta_{L-1}, \ldots, f_1\{\Theta_1, \mathbf{x}\})] \tag{6}$$

As can be inferred from Eq. (6), neural networks act as composite functions, and each affine layer with index L consists of parameters Θ and a non-linear activation function f. It is noted that bias terms are omitted in Eq. (6). The optimal set of parameters is obtained by gradient-based optimization, which assumes samples to be IID, of an arbitrary loss function $\mathcal{L}$. It is desired to find a global minimum of the loss function, but it has been shown that there is no unique parameter combination θ yielding this. It is noted that θ corresponds to a vector concatenating all Θ_L. In actual fact, there is a region of optimal parameter combinations around the global optimum of the loss landscape. This multiplicity has been found by Sussmann (1992) and eventually allows for continual learning, i.e. optimizing for two different tasks at once without (or marginally) sacrificing accuracy. During training, the aim is to learn a parameterization θ_1 that leads to optimal performance on the first task. Afterwards, the model consisting of the prior parametrization is retrained on task

---

[3]DevOps is a term combining the fields of software development (Dev) and IT operations (Ops).

2, leading to a posterior parameter combination $\theta_{1,2}$, which yields the best performance on task 2 without losing significant accuracy on task 1. Mathematically speaking, the process can be described by the posterior $p(\theta|\mathcal{D}) = p(p(\theta|\mathcal{D}_1)|\mathcal{D}_2)$, where $\mathcal{D}_i$ refers to the dataset of the associated task. In fact, Kirkpatrick et al. (2017) utilizes Bayes' theorem for deriving corresponding regularized loss functions for incremental learning. Crucially, methods for incremental learning suffer from the plasticity-stability dilemma. In other words, from training on task 1, the parameter vector θ_1 lies on a manifold corresponding to optimal accuracy on task 1. However, an updated parameterization θ_2 will detach from the initial manifold and change to a manifold tailored to task 2. Hence, the posterior parameter vector θ_2 unlearned all information on task 1, which was encoded in the prior parameter combination θ_1. Therefore, models are supposed to be flexible to acquire new information but rigid enough to avoid catastrophic interference (or forgetting). Ramasesh et al. (2021) provides a detailed study on catastrophic interference and its mitigation. In the present work, three methods for incremental learning are applied and compared to the **Default** method, which is considered retraining the entire model without any restrictions (or regularization) and thus corresponds to a highly flexible model. The three examined methods are presented in the following:

1. **Freezing** has been presented in Jung et al. (2016) and is frequently applied in the field of transfer learning, i.e. the application of a pre-trained model in a similar domain. This procedure is considered as straightforward for implementation and efficient during training since only the parameters in the unfrozen layers are adaptable. In contrast to the default method, this method is considered as relatively rigid. Interestingly, Ramasesh et al. (2021) observed that the upper hidden layers contribute most to the associated decay in accuracy and hence the first two hidden layers are frozen when using this method. In Mittendorf et al. (2022c), layer freezing has been applied in a practical setting and showed good applicability, since it does not require a custom loss function, in contrast to the following two methods.

2. **L$_2$** regularization is widely used during the training of neural networks (and other ML methods) and aims at minimizing complexity for mitigating overfitting. The conventional definition penalizes the actual magnitude of weights, but the implementation of Kirkpatrick et al. (2017) punishes large differences between θ_i and θ_{i-1}, i.e. the change from prior to posterior parametrization. By using the L_2 norm, the posterior parameter vector stays in Euclidean proximity to the prior one.

17

$$\mathcal{L}_{L2}(\theta_i) = \mathcal{L}(\theta_i) + \lambda \sum_j (\theta_{i,j} - \theta_{i-1,j})^2 \tag{7}$$

In Eq. (7), i corresponds to the index of the task and j is the index of the parameter vector θ. The method relies on structural risk minimization, i.e. the actual loss function for the current task is extended by a penalty term, which is multiplied by the hyperparameter λ for weighting of the regularization term.

3. Elastic Weight Consolidation (**EWC**) has been proposed by Kirkpatrick et al. (2017) and mimics an information-based regularization technique. In fact, each parameter has an assigned information-weight, which is calculated by the Fisher information matrix, or rather its diagonal. A theoretical derivation of EWC and the intricacies of the calculation of the Fisher information matrix are presented in Aich (2021). EWC follows also structural risk minimization, which can be seen in Eq. (8).

$$\mathcal{L}_{EWC}(\theta_i) = \mathcal{L}(\theta_i) + \frac{\Lambda}{2} \sum_j F_j (\theta_j - \theta_{i-1,j})^2 \tag{8}$$

In Eq. (8), Λ is a hyperparameter determining the strength of the regularization, F_j is the diagonal of the Fisher information matrix and θ_{i-1} are the model parameters resulting from the previous training instance. In the present context, the hyperparamter $\alpha = \Lambda/2$ is defined and eventually optimized. The diagonal of the Fisher information matrix is approximated by using the average of squared gradients of a subset of the training data. It is noted that the calculation of F_j leads to an increase in CPU time.

In theory, Eqs. (7) and (8) could both be extended towards considering even a task before the prior. However, from a Bayesian perspective, it is assumed that the parameters, which are important to the prior, are subsequently also important to the prior of the prior. Lastly, the four investigated methods have a varying degree of plasticity and stability. The default method is considered the most plastic, whereas the freezing method is the most stable. The two regularization techniques are related, since the L_2 method punishes large changes in magnitude of the parametrization and EWC punishes the same but, according to an importance estimate, obtained through the Fisher information matrix. In addition, both hyperparameters λ and α act as stiffness and higher values attribute more weight to the previous training instance, i.e. the prior.

3.2. Determination of the performance indicator

In Fig. 4, it was observed that fouling leads to a shift in the relationship between operational/environmental conditions and shaft power, i.e. a higher engine power is needed for the same conditions due to biofouling. In addition, when this shift remains unexplained, i.e. cannot be inferred from any feature, it is a clear indication of concept drift. In fact, some papers from the literature attempt to obtain the fouling-related added power in an explicit methodology by forcing a temporal relationship onto marine growth by, e.g. using "days since last dry-docking" as a feature. This is presumably inspired by explicit methods from the field of predictive maintenance, such as the prediction of remaining useful life. However, time is only one of many predictors of biofouling, which include (among others) salinity, temperature, paint type, and sun exposure. For more detailed information on the driving factors of marine growth and its technical as well as economic ramification, cf. Yebra et al. (2004). Due to the high complexity of marine growth patterns, an implicit methodology for deriving a biofouling indicator is pursued in the present context. For reference, concept drift may be detected by a windowing approach via, e.g. the Kolmogorov-Smirnov test, in a univariate problem, but ship performance monitoring is a multivariate problem, i.e. propulsion power depends on a number of different features, such as speed and draft. Additionally, added power due to biofouling is a latent variable, i.e. not directly measurable; therefore, it has to be derived using similar reference conditions. In this work, the relative power increase $\Delta\tilde{P}_B$ is used as a proxy variable for biofouling and defined in Eq. (9).

$$\Delta\tilde{P}_B = \frac{P_{B,t} - P_{B,t_0}}{P_{B,t_0}} \tag{9}$$

In Eq. (9), P_{B,t_0} denotes the engine power from an underlying baseline with no (or minor) biofouling and $P_{B,t}$ is the instantaneous value under the current conditions at time t including the effect of fouling. In fact, the speed and draft dependency of ISO 19030 fouling indicators (and performance indicators in general) are some of the major drawbacks of the standard, as shown in Mittendorf et al. (2022c). The dependency of the relative added power $\Delta\tilde{P}_B$ on non-dimensional advance speed Fn and sea state in terms of H_s is shown in Fig. 7 using synthetic data. It is noted that the dataset with uniformly distributed samples is employed, i.e. the plot shows an exaggerated likelihood of severe sea states for the sake of clarity.
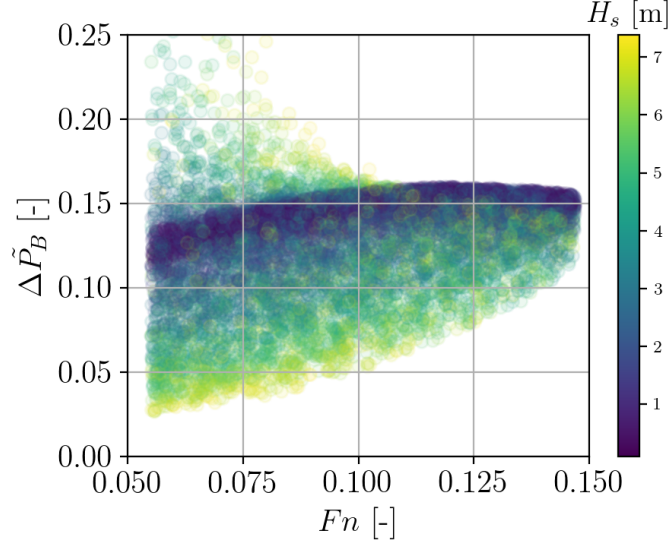
19

Figure 7: The relative added power $\Delta \tilde{P}_B$ caused by added roughness as a function of ship speed Fn and sea state H_s using dataset $\mathcal{D}_A$ for sudden concept drift. The reference conditions are for $k_S = 30\mu m$ and the fouled conditions correspond to $k_S = 300\mu m$, i.e. heavy slime. It is noted that the average power increase is 13.2%.

The total ship resistance R_T is composed of a frictional R_F and a pressure component R_P, with an interaction term, i.e. the viscous pressure resistance R_{VP} (Molland et al., 2011). It is stated that biofouling and thus $\Delta \tilde{P}_B$ are mainly dependent on the R_F fraction of the total resistance, and it is emphasized that Fig. 7 results from a simplified simulation framework disregarding possible effects of, e.g. propeller ventilation or cavitation. Initially, it stands out that $\Delta \tilde{P}_B$ increases with speed, which is also reflected in Fig. 4. However, the increase saturates for $Fn > 0.1$, where (roughly) the wave-making resistance, i.e. the pressure resistance fraction, starts to increase. In addition, the added-wave resistance contributes significantly to R_P, which is reflected by the decrease of the fouling related added power in higher sea states. Interestingly, for $Fn < 0.1$, a possible increase in $\Delta \tilde{P}_B$ is visible due to a relative decrease of R_P caused by a negative added-wave resistance in following waves, i.e. when $\bar{R}_{AW}$ is a pushing force. This phenomenon is a function of the encounter frequency and is thus limited to lower Fn, cf. Liu and Papanikolaou (2020). In addition, the effect of the main engine constraints can be inferred in Fig. 7 since the occurrence of higher sea states reduces with higher forward speed. Lastly, Fig. 7 underlines the importance of evaluating hull performance under consistent reference conditions to avoid biases resulting from the vessel's operational profile and encountered sea states.

The present work follows a similar implicit approach, as shown by Schmode and Antola (2020),

i.e. the added power is evaluated under the same *simulated* reference conditions using an adaptive model. In other words, this model is used for approximating the temporary relationship between operational conditions and required shaft power P_B. As has been proposed in Mittendorf et al. (2022c), the proxy for biofouling is determined using a baseline from an initial model, and an *instantaneous* estimate of that obtained from an adaptive model, and the power increase is then obtained from Eq. (9) in a post-processing step. The relative added power $\Delta \tilde{P}_B$ is calculated for 8 speeds $U \in [7.0, 14.0]$kts and eventually averaged to minimize its speed dependency and possible variance caused by modeling errors. The reference conditions correspond to the ones used in the case of Fig. 4, i.e. $H_s = 1.0$m, $T_p = 5.0$s $V_w = 4.6$m/s in head wind and waves at design draft. For mitigating biases due to limited data availability, a lower sea state has been chosen, cf. Fig. 5a. In a nutshell, the overall machine learning regression task is to predict the required shaft power in different training instances, and the actual fouling indicator is obtained using Eq. (9) for consistent reference conditions in a separate post-processing step.

3.3. Training setup

The used model is a multilayer perceptron with 3 hidden layers and 50 neurons per layer, i.e. is of lower complexity (5,401 trainable parameters) and without additional regularization techniques. In a preliminary sensitivity study, it appeared that the mean squared error and log-cosh error are too sensitive to the variance within the used datasets when using them as the loss function. Hence, the mean absolute error (MAE) is chosen as the loss function in the present case and defined in Eq. (10),

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i| \tag{10}$$

Here N corresponds to the sample size, $\hat{y}$ is the model prediction and y is the ground truth (in the present case P_B). For optimizing the loss function, the Adam optimizer (Kingma and Ba, 2015) is used and the batch size has been set to 64. The used sample sizes, as well as the number of epochs, are mentioned separately for the two different scenarios in the following section. In order to mitigate the look-ahead bias, the data is normalized according to the extrema of the warm-up period. The employed deep learning library is TensorFlow 2.6, which was proposed by Abadi et al. (2015). Due to the smaller model complexity, the training procedures were performed on a single

445 Intel© Core i7-8565U CPU, 1.80 GHz with 16 GB physical memory (RAM). It is noted that a global
446 seed was defined for the Python environment including each library, in order to obtain reproducible
447 results throughout the conducted studies.

4. Results and discussion

449 This section presents the obtained results and their interpretation. We consider three major
450 aspects: (1) Feature engineering in Subsection 4.1, (2) Analysis of two concept drift scenarios in
451 Subsection 4.2 as well as (3) the critical assessment of the findings in Subsection 4.3. Throughout,
452 the root mean squared error normalized by the mean of the P_B values (indicated by overbar)
453 $\widetilde{RMSE}$ is used as a metric, cf. Eq. 11.

$$\widetilde{RMSE} = \frac{RMSE}{\bar{P_B}} \text{ with } RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{11}$$

454 The individual models are evaluated on the $\widetilde{RMSE}$ of the holdout or validation set and the
455 values for the training set are listed in appendix A2. In doing so, the generalization capability of
456 the methods is assessed. Due to the application of incremental learning, there are multiple training
457 instances and in each of them, the data is split into training and validation sets, where the latter
458 corresponds to 20% of the entire set. Hence, the presented $\widetilde{RMSE}$ reflects a mean for all validation
459 sets throughout the considered instances. In doing so, all training procedures and their local models
460 receive equal weight regardless of the sample size.

4.1. Feature engineering

462 In several older studies pertaining to ML-based performance monitoring, no feature engineering
463 has been carried out, which led at times to peculiar combinations of predictors including, e.g.
464 the vessel's GPS position or the *absolute* wave heading angle. In more recent studies, correlation
465 analysis is conducted or the built-in feature importance method (mean decrease impurity) in case
466 of tree-based ensemble methods is employed. However, it turns out that the elimination of collinear
467 features is rarely addressed. It is emphasized that the model's feature space is supposed to be
468 as efficient as possible due to the negative consequences caused by the curse of dimensionality,
469 i.e. the more dimensions, the sparser the data becomes (Hastie et al., 2009). In Mittendorf et al.
470 (2022c), a feature vector was determined using domain knowledge, but herein, a similar framework

22

Table 3: Variance inflation factor (VIF) analysis for the three different datasets showing one iteration per row.

Dataset	U	H_s	T_p	β_0	$V_{w,r}$	ψ_r	Q	rpm	d_m
	36.32	7.202	9.601	38.76	22.54	27.91	76.40	219.89	2.538
$\mathcal{D}_A$	14.18	7.036	7.527	35.86	19.27	27.09	31.70	elim.	2.272
	13.92	7.021	7.237	elim.	18.90	2.700	31.50	elim.	2.235
	3.240	6.997	6.793	elim.	9.796	2.677	elim.	elim.	1.783
	355.5	6.111	6.678	30.59	18.38	20.99	111.1	761.1	7.963
$\mathcal{D}_B$	57.93	5.485	6.148	29.33	11.18	20.73	88.74	elim.	4.342
	6.162	5.383	5.510	28.52	6.642	20.54	elim.	elim.	2.146
	5.942	5.363	5.387	elim.	6.597	2.491	elim.	elim.	2.126
	355.5	6.111	6.678	30.59	18.38	20.99	111.1	761.1	7.963
$\mathcal{D}_C$	57.93	5.485	6.148	29.33	11.18	20.73	88.74	elim.	4.342
	6.162	5.383	5.510	28.52	6.642	20.54	elim.	elim.	2.146
	5.942	5.363	5.387	elim.	6.597	2.491	elim.	elim.	2.126

as in Mittendorf et al. (2022a) will be applied, i.e. collinear features are eliminated in an initial iteration and the Mean Decrease Accuracy (MDA) method, as proposed by Breiman (2001), is used afterwards.

The Variance Inflation Factor (VIF) is a measure for an increase in variance of parameter estimates in linear regression if an additional variable i is added to the model. Thus, the VIF acts as an indicator of collinearity, which refers to highly correlated predictors, i.e. those that do not add any unique or independent information. The VIF is defined in Eq. (12).

$$VIF = \frac{1}{1 - R_i^2} \text{ with } R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y}_i)^2} \tag{12}$$

In view of Eq. (12), it is stated that R^2 is the coefficient of determination and an overbar denotes the mean. In essence, the feature with the highest VIF is eliminated in every iteration and the procedure is repeated until all features have a $VIF < 10$. The results from this iterative procedure for the initial feature vector $\mathbf{x}_i = \{U, H_s, T_p, \beta_0, V_{w,r}, \psi_r, Q, rpm, d_m\}$ are presented in Tab. 3 for all three datasets. It is noted that data from the sudden concept drift scenario is taken for $k_S = 30\mu$m and that engine power P_B is not included, as it is the target variable. As a side note, the statsmodels[4] library is used for the calculation of the VIF.

Initially, it stands out that both rpm and Q are eliminated in all three datasets, similarly to

---

[4]https://www.statsmodels.org/

Mittendorf et al. (2022c). Physically, both quantities are related (through Eq. (4)) to the shaft power, which is highly dependent on forward speed U. Hence, rpm and Q exhibit no additional information under the considered conditions. Additionally, the assumption of only wind waves and the simplified relationship between H_s and V_w has an effect on collinearity, which would not occur on actual in-service data. As can be seen in Tab. 3, the relative wave direction is considered collinear in all cases, since it can be easily determined from the relative wind speed and direction. It is noted that the absolute wind speed is given implicitly through H_s. Moreover, it is appreciated that both $\mathcal{D}_B$ and $\mathcal{D}_C$ show the exact same VIF values, which indicates that the procedure is unaffected by introducing stochastic noise. Eventually, rpm, Q and β_0 are eliminated from all datasets after the VIF procedure and an updated feature vector is used during the MDA study. This model-agnostic algorithm is iterative and quantifies the importance of one particular feature on model accuracy by disregarding it during training. For the sake of clarity it is noted that the engine power P_B is the target variable and that the sklearn[5] function is used in the present work.
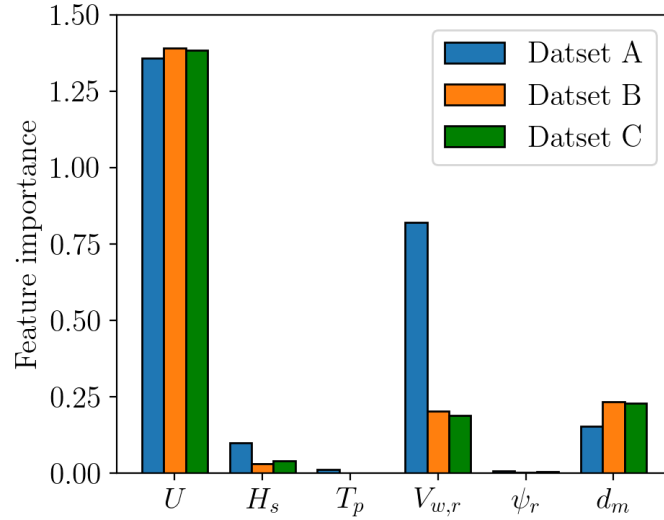

Figure 8: Sensitivity study for permutation feature importance based on the datasets for sudden concept drift and $k_S = 30\mu$m, i.e. clean conditions.

The immediate observation from Fig. 8 is that advance speed U is by far the most important predictor of shaft power, which is due to the proportionality $P_B \propto U^3$. Moreover, it can be seen that both U and d_m increase slightly in importance when comparing $\mathcal{D}_B$ and $\mathcal{D}_C$ to $\mathcal{D}_A$. This may be attributed to the change in likelihood of severe sea states, since both H_s and especially $V_{w,r}$

---

[5]https://scikit-learn.org/

24

⁵⁰³ lose importance in case of $\mathcal{D}_B$ and $\mathcal{D}_C$. Lastly, peak period T_p and relative wind direction $V_{w,r}$

⁵⁰⁴ exhibit negligible feature importance and are therefore eliminated in this iteration. As possible

⁵⁰⁵ explanations, T_p is related to H_s through the wave steepness constraint and ψ_r can be inferred,

⁵⁰⁶ since all vector components for calculating the relative wind speed (and direction) are available, as

⁵⁰⁷ noted above. As a side note, the transformations used in Mittendorf et al. (2022e) have been tested,

⁵⁰⁸ i.e. using sine/cosine in case of directions and the logarithm for H_s, but no increase in accuracy or

⁵⁰⁹ feature importance could be seen (considering all three datasets). The final shape of the feature

⁵¹⁰ vector and the target variable are visible in Eq. (13).

$$\mathbf{x}_i = \{U, H_s, V_{w,r}, T_m\} \text{ with } y_i = P_B \tag{13}$$

⁵¹¹ Essentially, the composition of feature vector $\mathbf{x}_i$ is greatly influenced by the simplified simulation

⁵¹² procedure and therefore not applicable under real conditions. However, the proposed framework as

⁵¹³ such is universally applicable and it is emphasized that collinearity has to be considered during

⁵¹⁴ feature engineering in the field of ML-based ship performance monitoring.

⁵¹⁵ *4.2. Comparison of different incremental learning strategies*

⁵¹⁶ In this subsection, two concept drift scenarios will be investigated. In both cases, hyperparameter

⁵¹⁷ studies, model assessment and the model sensitivity to different uncertainty categories are addressed.

⁵¹⁸ *4.2.1. Sudden concept drift*

⁵¹⁹ In most Bayesian (or iterative) approaches, finding a starting point (or value) is a delicate

⁵²⁰ matter. In the present context, an initial model is trained in a so-called warm-up period without

⁵²¹ considering any of the presented incremental learning procedures. In the succeeding training periods,

⁵²² the model is trained adaptively. In the warm-up period, the initial model is trained for 30 epochs,

⁵²³ whereas it is adapted in 10 epochs in the following instances. Similarly, 5600 samples are used in

⁵²⁴ the initial training period and 1750 in the two adaptive training periods.

⁵²⁵ Both L_2 and EWC methods require the optimization of a hyperparater, cf. Sec. 3.1. Hence, a

⁵²⁶ sensitivity study of λ and α for L_2 and EWC, respectively, is presented in Fig. 9. For this purpose,

⁵²⁷ the results of the hyperparameters in the range of $\{0.1, 1, 2, 5, 10, 15\}$ are presented for $\widetilde{RMSE}$ in
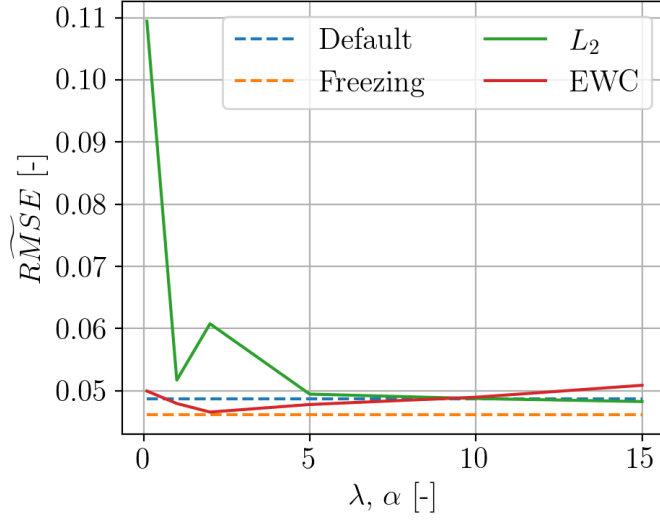
⁵²⁸ Fig. 9.

Figure 9: Hyperparameter study for the L_2 and EWC method using $\mathcal{D}_A$ (sudden concept drift) and the $\widetilde{RMSE}$ of the validation sets.

From Fig. 9, it can be seen that EWC shows a more robust performance compared to the L_2 method. In addition, moderate values of λ and α lead to higher out of sample accuracy. The same study was conducted for the remaining two datasets for sudden concept drift and it turns out that the hyperparameters vary in their optimal values. This in turn, indicates that hyperparameter studies are needed in any case and one can not rely on "best practice" or default values, which complicates the use of the associated methods in a practical context.

The corresponding loss curves for the models with optimized hyperparameters are visible in Fig. 10a for $\mathcal{D}_A$, where the $\widetilde{RMSE}$ of the validation set is depicted for the number of epochs. In the warm-up period, only the results for the Default method are shown, since all four models are based on one initial model. In the two succeeding adaptive training instances, it stands out that the Default method shows a volatile behavior in relation to the other methods. In addition, the Freezing and EWC methods yield relatively stable performance with lower $\widetilde{RMSE}$ values. The performance of the Default method on the validation set of the warm-up period (shown by +) is supposed to indicate the degree of concept drift, i.e. the difference in accuracy, and thereby the necessity for incremental learning. It is noted that Fig. 10a has a resemblance to a study in Kirkpatrick et al. (2017), but the discrepancy in performance of the individual methods was larger in their case, which could be due to the fact that the individual tasks are sequential and related in this context.

Overall, relatively similar $\widetilde{RMSE}$ values have been obtained for all four methods, which are

26
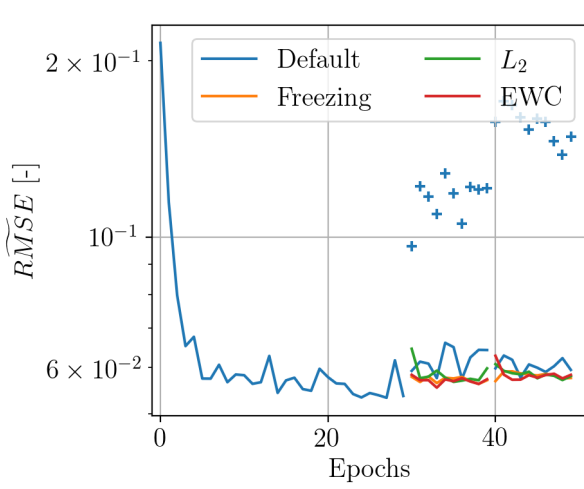
(a) Training loss for $\mathcal{D}_A$ including samples indicating the error of the instantaneous model on the validation set of the warm-up period (+).
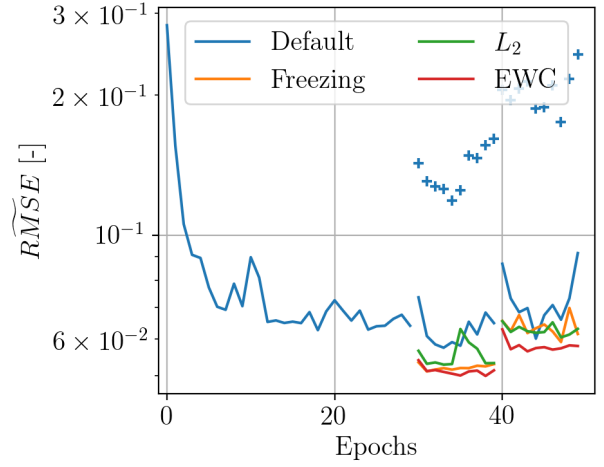
(b) Training loss for $\mathcal{D}_C$ including systematic biases resembling shifts in operational profile with samples showing the error of the instantaneous model on the validation set of the warm-up period (+).

Figure 10: Loss curves for all three instances under sudden concept drift. It is noted that the ordinates are in logarithmic scale.

also given for all three datasets in Tab. 4. Due to random sampling, the synthetic dataset shows large variation in conditions, i.e. the search space of parameters is well explored. This is in contrast to reality, where the ship may experience changes in its operational profile. Hence, in Fig. 10b systematic biases or distributional shifts are introduced following Tsompopoulou et al. (2022). In order to achieve greater a resemblance to real data, dataset $\mathcal{D}_C$ is used (with aleatoric and epistemic uncertainty) and in the second training period all ballast conditions are filtered out, whereas speeds $U > 10.5$kts are disregarded in the third interval. In fact, it turns out that the Default method, i.e. the most adaptive method, shows large scatter in accuracy and even convergence issues, which may indicate the occurrence of catastrophic interference. For consistency reasons, the same number of epochs were used for all methods, but in future studies it will be important to define a termination criterion for ensuring sufficient model convergence. Moreover, the difference in accuracy of the remaining three methods turns out to be larger in comparison to the previous study. Similarly to Kirkpatrick et al. (2017), the L_2 method shows minor scatter in its loss, and the Freezing as well as EWC methods yield stable and higher performance in case of the latter training instances. Overall, it is underscored that incremental learning methods are required in case of real data due to the possible change in operational conditions.

The effects of epistemic and aleatoric uncertainty on model performance, assessed in terms

Table 4: $\widetilde{RMSE}$ of predictions using the validation sets in case of sudden concept drift.

	Default	Freezing	L_2	EWC
$\mathcal{D}_A$	0.0549	**0.0520**	0.0544	0.0525
$\mathcal{D}_B$	0.0271	0.0242	**0.0217**	0.0231
$\mathcal{D}_C$	0.0295	0.0252	**0.0249**	0.0260

of $\widetilde{RMSE}$, are shown in Tab. 4 and Fig. 11. Initially, it is emphasized that the models do not perform better on $\mathcal{D}_B$ than on $\mathcal{D}_A$, as could be inferred by the lower $\widetilde{RMSE}$ values, but $\mathcal{D}_A$ shows a higher likelihood of severe states. This in turn leads to a higher (physical) variance, to which $\widetilde{RMSE}$ is very sensitive in case of modeling discrepancies.
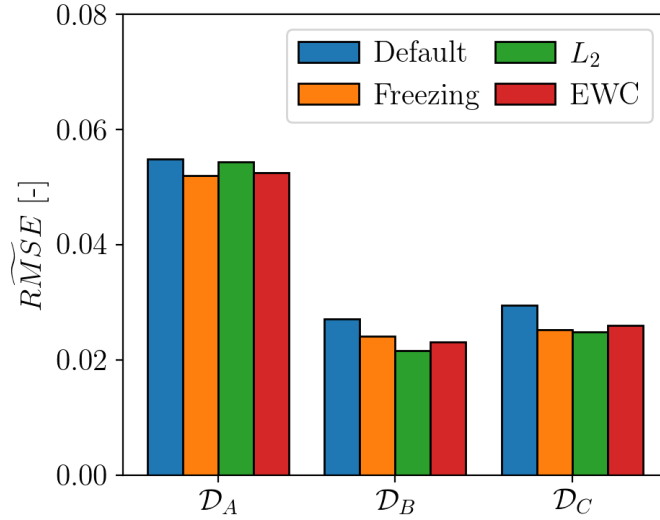

Figure 11: Performance on validation set of all three datasets with sudden concept drift using $\widetilde{RMSE}$ as a metric.

In the case of $\mathcal{D}_A$, it stands out that all methods perform equally well, but the Freezing method shows the lowest $\widetilde{RMSE}$ value. In view of $\mathcal{D}_B$, the Default method shows a lower accuracy than the other three methods, which can be attributed to the introduction of epistemic uncertainty, i.e. the skewed parameter distributions and the disturbance due to the unobserved change in ϑ_{sea}. It shows that the L_2 method performs best, but the injection of white noise, i.e. aleatoric uncertainty, into the datasets in case of $\mathcal{D}_C$, leads to a relative drop in accuracy of the L_2 method. Ultimately, it is indicated that all methods, but the L_2 method in particular, are sensitive to aleatoric uncertainty, which has also been observed in Fig. 10b. Finally, both Freezing and EWC showed stable, but moderate performance on all three considered datasets.

4.2.2. Incremental concept drift

578 In preliminary studies, it turned out that the considered models require more epochs for 579 successful convergence on the transient dataset. Hence, the number of epochs of the warm-up period 580 was increased to 40 and in the following instances, re-training was carried out for 20 epochs. The 581 dataset is binned into 10 windows without any overlap, in contrast to Mittendorf et al. (2022c), as it 582 introduced a lagging behavior of the performance indicator. This leads then to approximately 1500 583 samples per window and the same sample size as before (5600) is used in the warm-up period, where 584 the data is taken from the sudden concept drift scenario (with $k_S = 30\mu$m). In a real application, 585 binning of the data as well as the introduction of overlap is dependent on data quality as well as 586 availability and should be carefully considered.

587 Similarly to the previous subsection, the hyperparameter study is carried out using $\mathcal{D}_A$. In 588 contrast to Fig. 9, a relatively clear trend becomes apparent in Fig. 12, i.e. the higher the 589 hyperparameter, i.e. the weight on the prior, the higher the model performance of the L_2 and EWC 590 methods. In fact, exclusively values larger than 5 have been obtained for the other two datasets 591 ($\mathcal{D}_B$ and $\mathcal{D}_C$). In Fig. 12, it is also appreciated that the accuracy of L_2 and EWC are consistently 592 lower than the Default (and Freezing), which might indicate that the former two methods are more 593 affected by unsteadiness.
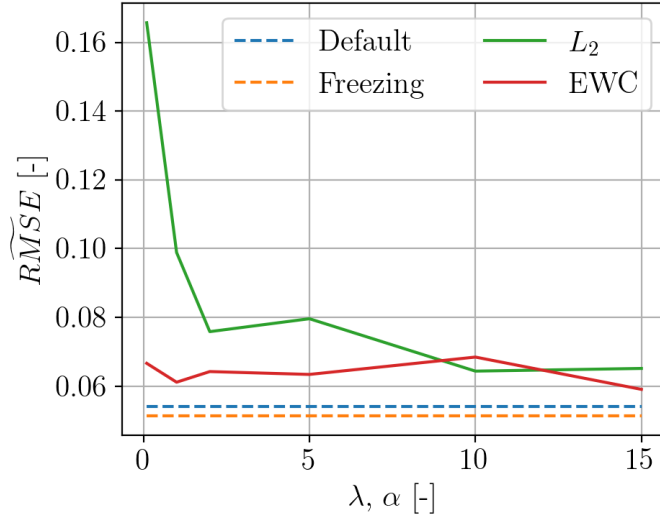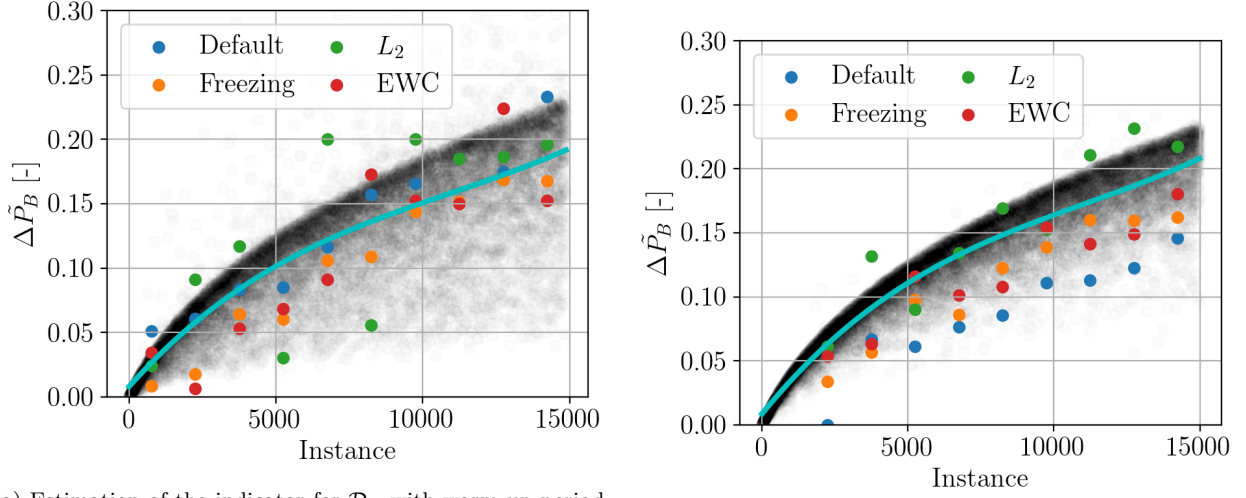

Figure 12: Hyperparameter study for the L_2 and EWC method using $\mathcal{D}_A$ (incremental concept drift) and $\widetilde{RMSE}$ of the validation sets.

594 In Fig. 13, the estimated $\Delta \tilde{P}_B$ indicator is depicted for the number of data instances, which

(a) Estimation of the indicator for $\mathcal{D}_A$ with warm-up period using the corresponding dataset from the sudden concept drift scenario.

(b) Estimation of the indicator $\Delta \tilde{P}_B$ for $\mathcal{D}_C$ using the first bin of the data as warm-up period.

Figure 13: Development of the ship performance proxy according to the individual methods.

is another measure for the development in time. The simulated ground truth is indicated in grey (transparent black) and a second order polynomial has been fitted to the relative power increase of $\mathcal{D}_A$. Once again, the scatter of the performance decay due to the environmental and operational conditions stands out, cf. Fig. 7. As a side note it is seen in Fig. 13 that a linear increase in roughness leads to a non-linear increase in added power, which is also confirmed by a study focused on CO_2 emissions presented in GloFouling (2022).

In view of Fig. 13a, it can be said that generally, all four models are able to identify an incremental concept drift, but the degree of robustness, i.e. the variance of the performance indicator varies significantly. Similarly to Fig. 11, it turns out that the estimate obtained from the L_2 method is highly affected by variance. The same holds for the Default method, i.e. the least stable one. The Freezing and EWC methods emulate the performance decay sufficiently, but show minor outliers. In Fig. 13b, a more realistic case study is shown, where no initial model from a separate and steady warm-up period is available, i.e. the first bin of the transient dataset is taken as a warm-up period. In addition, $\mathcal{D}_C$, i.e. the dataset including both uncertainty categories, is used and $\Delta \tilde{P}_B$ shows a lower variance, which is due to the parameter distributions, which are skewed towards lower sea states. Overall, the robustness of the provided performance indicators increased in contrast to Fig. 13a due to the lower inherent (physical) variance. In turn, this underlines the possible benefit of filtering instances in higher sea states, as demanded by ISO 19030. Moreover,

30

Table 5: $\widetilde{RMSE}$ of predictions using the validation sets in case of incremental concept drift.

	Default	Freezing	L_2	EWC
$\mathcal{D}_A$	0.0540	**0.0514**	0.0643	0.0590
$\mathcal{D}_B$	0.0248	**0.0231**	0.0404	0.0337
$\mathcal{D}_C$	0.0258	**0.0230**	0.0375	0.0318

similar observations as in the case of the previous case study can be made, i.e. both EWC and Freezing yield the most robust estimates of the relative power increase. Still minor variance is visible and underlines that the first order approximation by linear regression in Mittendorf et al. (2022c) appears as reasonable, even though the actual performance drop has a non-linear character. Moreover, it can be seen that all $\Delta \tilde{P}_B$ estimates are slightly biased (non-conservative), in comparison to Fig. 13a. This is due to training on the first bin of transient data, which introduces a consistent offset. Lastly, it is unexpected that the models perform relatively well without a warm-up period on steady data and it will be important to confirm this on real data.

Similarly to Fig. 13, the Freezing method shows superior accuracy on the validation sets in Tab.5. However, as seen in Fig. 12, both L_2 and EWC exhibit a reduced performance in direct comparison to the Default method. As discussed, it might indicate that both methods are affected by the transient behavior of the data. As a side note, when taking the MCR (Maximum Continuous Rating) of the potential main engine as a reference, i.e. 30,000kW, then the original RMSE values on the validation sets arrange between 1-2% considering all methods. In Fig. 14, the values of Tab. 5 are shown and it is noted that the larger $\widetilde{RMSE}$ values in case of $\mathcal{D}_A$ are attributed to the higher variance, which results from the relatively higher likelihood of harsh sea states.
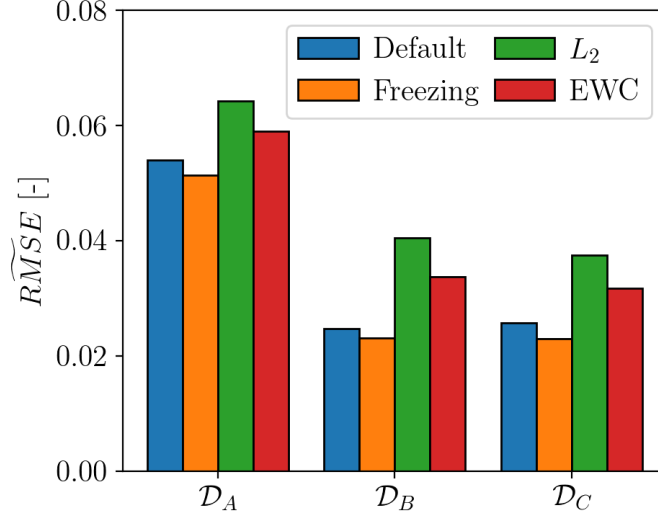
Figure 14: Performance on validation sets of all three datasets with incremental concept drift using $\widetilde{RMSE}$ as a metric.

629 From Fig. 14, it stands out that both L_2 and EWC perform relatively worse on transient
630 data (quasi-IID), while both Default and Freezing methods show consistent out-of-sample accuracy.
631 Moreover, it is striking that the performance of L_2 and EWC decreases relatively to the other two
632 methods in case of $\mathcal{D}_B$ (epistemic uncertainty), but increases again for $\mathcal{D}_C$ (compared to Default
633 and Freezing), cf. Tab. 5. As shown by Bishop (1995), injecting noise into training data may
634 act as additional regularization leading to increased model performance. This in turn indicates
635 that the combination of conventional regularization methods, e.g. Dropout, and regularization
636 aimed at incremental learning may enhance model accuracy in case of incremental concept drift.
637 In order to study in this more rigorously, it will be an interesting aspect to use additional more
638 sophisticated metrics dedicated for quantifying model uncertainty with higher accuracy, e.g. by
639 using F1 retention curves, as shown in Tsompopoulou et al. (2022).

640 *4.3. General Remarks and discussion*

641 Following the principle of Occam's razor: The simplest incremental learning model provides
642 the most robust and accurate performance. In fact, layer freezing exhibits great versatility and
643 stable estimates of the biofouling proxy variable in both concept drift scenarios. In addition, the
644 particular method does not require a regularized loss function or any additional computational
645 effort. In Mittendorf et al. (2022c), it has even been observed that layer freezing performs well on

in-service data. However, it will be part of future studies to evaluate and improve the performance of the L_2 and EWC methods. The L_2 norm, used in both regularization methods (i.e. L_2 and EWC), tends to be volatile when applied to data with significant variance. Hence, in theory, it could be interesting to use the L_1 norm since it is generally more robust. However, it can lead to parameter sparsity,i.e. forcing model parameters (or in this case, the parameter change) toward zero. In other words, L_1 regularization might make the model overly stable and thus unable to adapt to newly acquired data. As indicated in the case of incremental concept drift, additional regularization techniques may be beneficial for the more advanced methods. For instance, it has been shown that Dropout, i.e. omitting several neurons at random during training, acts satisfactorily as a mitigation technique against catastrophic interference and thus for incremental learning (Mirzadeh et al., 2020). From a theoretical perspective, the application of progressive neural networks seems advantageous since adding neurons and lateral connections for each additional task without altering the original parameters shows overall good performance (Rusu et al., 2016). Nevertheless, this approach requires increased computational effort and becomes increasingly complex due to the additive architecture.

In accordance with most other studies of ML-based performance monitoring, data quality is seen as the most influential factor for accurate ML-based performance monitoring. However, in contrast to other contributions, it is not the degree of variance but the IID condition that appears to be most critical, which is a finding from the incremental concept drift study. In fact, the degree of stochastic noise has an overall reduced effect on model accuracy and even turned out to be beneficial in the case of transient data. However, the performance indicator showed sensitivity to physical variance caused by higher sea states. It is thought that increasing training epochs and model complexity might not lead to more robust estimates, but filtering for harsh weather conditions, as proposed in ISO 19030, appears to be more promising. Another common data (or sensor) quality issue is sensor drift, which complicates differentiating between actual biofouling and the effect of a degrading sensor. Especially the Doppler velocity log for speed through water and the torsionmeter for measuring shaft power are known to be very susceptible to sensor drift, as shown by Ikonomakis et al. (2021) and Aldous et al. (2015). Sensor drift affects not only machine learning-based performance monitoring but the statistical adage 'Garbage in, garbage out' holds for ship in-service analyses in general.

In the present work, an aggregated proxy variable for both hull and propeller performance

33

is provided. However, their decomposition is desirable from a ship operator's point of view but notoriously difficult to realize. Park et al. (2018) achieve this decomposition via a parallel ML approach predicting both speed and power simultaneously. Interestingly, the proposed approach not only facilitates decomposition but also increases overall transparency. Gupta et al. (2022) apply Monte-Carlo dropout as a regularization method in an ML-based ship performance monitoring application. Crucially, Monte-Carlo dropout is seen as a deep Gaussian process regression allowing for an uncertainty-aware estimate, i.e. accompanied by prediction intervals. When facing reduced data quality and distributional shifts, an assessment of model uncertainty is indispensable in a practical scenario. Hence, Tsompopoulou et al. (2022) applied probabilistic deep learning for quantifying uncertainty at prediction time. In fact, all three mentioned methods show great potential for extending the presented methodology but may also lead to an increased computational effort.

5. Summary and future work

With market conditions dominated by increasingly stringent regulations combined with fluctuating fuel prices and freight rates, vessel operators are forced to improve their ships' fuel efficiency to reduce the associated cost and comply with stipulated legislation. In the present paper, it has been shown that digital representations of ships have to be continuously updated to avoid a drop in model accuracy and to capture the effect of biofouling. In the present context, synthetic data for a standard tanker (KVLCC2) was simulated for various operational and environmental conditions. Moreover, several methods for incremental learning were implemented and compared concerning accuracy as well as robustness. The investigated methods were applied in both sudden and incremental concept drift scenarios. Additionally, their sensitivity to aleatoric and epistemic uncertainty within the training data has been examined. Lastly, it appears that the layer freezing methodology, i.e. fixing the upper layers of a pre-trained model throughout re-training, is the most practical method. However, it will be essential to confirm this using actual in-service data from ships in future studies. Possible modifications of the original incremental learning methods and the combination of several regularization techniques will also be part of extended work.

Ultimately, the presented adaptive methodology can lead to greater predictability of marine growth and, thus, allows for proactive hull maintenance. Nevertheless, digital representations of ships have not only adaptability as a critical characteristic but also multi-modality. Hence, the

34

inclusion of additional (high-frequency) sensor data, e.g. related to the motion or structural response of the ship in waves, as seen in Mittendorf et al. (2022e), may lead to an increase in accuracy as well as applicability in other domains, such as the assessment of a vessel's structural fatigue. In fact, performance monitoring is just one of several application cases of digital representations of ships. It becomes clear that associated models for route or trim optimization have to be continuously updated via incremental learning, which is not only due to the accumulation of marine growth but also the possible change in the operational profile. The use of hybrid methods shows an enormous potential for further increasing model robustness. In the presented literature study, physical modeling (first principles) and machine learning were treated separately, but so-called *greybox* approaches combine both domains and are usually characterized by greater transparency, as shown by Haranen et al. (2016). Ultimately, the presented ML-based methodology for determining a change in vessel performance may also serve to determine fuel savings from the application of, e.g. wind-assisted propulsion or other energy-saving devices. In fact, Camilleri et al. (2022) compare an ML-based approach to ISO 19030 for the assessment of the achieved savings related to the implementation of an air lubrication device. It becomes clear that performance indicators are generally speed, draft, and seaway dependent. Therefore, it is crucial for transparency to evaluate the possible savings under consistent conditions, as shown in the present work.

CRediT authorship contribution statement

M. Mittendorf: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Validation, Visualization, Writing - original draft & editing. U.D. Nielsen: Conceptualization, Methodology, Writing - review & editing, Supervision, Project Administration, Resources, Funding Acquisition. H.B. Bingham: Conceptualization, Methodology, Writing - review & editing, Supervision, Resources, Project Administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The individual synthetic datasets are publicly available on DTU Data under the reference Mittendorf et al. (2023) and the DOI: `https://doi.org/10.11583/DTU.21750257`.

Acknowledgements

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. URL: `https://www.tensorflow.org/`. software available from tensorflow.org.

Adland, R., Cariou, P., Jia, H., Wolff, F.C., 2018. The energy efficiency effects of periodic ship hull cleaning. Journal of Cleaner Production 178, 1–13.

Aich, A., 2021. Elastic weight consolidation (ewc): Nuts and bolts. URL: `https://arxiv.org/abs/2105.04093`.

Aldous, L., Smith, T., Bucknall, R., Thompson, P., 2015. Uncertainty analysis in ship performance monitoring. Ocean Eng. 110, 29–38.

Andersen, P., Borrod, A.S., Blanchot, H., 2005. Evaluation of the service performance of ships. Marine Technology 42, 177–183.

Berthelsen, F.H., Nielsen, U.D., 2021. Prediction of ships' speed-power relationship at speed intervals below the design speed. Transportation Research Part D Transport and Environment 99, 102996.

Bishop, C.M., 1995. Training with noise is equivalent to tikhonov regularization. Neural Computation 7, 108—116.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. doi:doi:http://dx.doi.org/10.1023/A:1010933404324.

Camilleri, J., Connolly, D., Sobey, A., Hudson, D., 2022. Quantifying the effect of green technologies on ship performance using operational data – design of experiments and data collection. Proc. of the 14th Symp. of High Performance Marine Vehicles (HIPER), Cortona. .

Coraddu, A., Oneto, L., Baldi, F., Cipollini, F., Atlar, M., Savio, S., 2019. Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. Ocean Engineering 186, 106063.

DNV-GL, 2018. Class guidelines: Wave loads (dnvgl-cg-0130). Technical Report .

GloFouling, 2022. Analysing the impact of marine biofouling on the energy efficiency of ships and the ghg abatement potential of biofouling management measures. GloFouling Partnerships and GIA for Marina Biosafety, London .

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. The MIT Press, Cambridge.

Gupta, P., Rasheed, A., Steen, S., 2022. Ship performance monitoring using machine-learning. Ocean Engineering 254, 111094.

Hansen, S.V., 2011. Ship performance monitoring. Ph.D. thesis, Technical University of Denmark .

Haranen, M., Pakkanen, P., Kariranta, R., Salo, J., 2016. White, grey and black-box modelling in ship performance evaluation. Proc. of the 1st Hull Performance and Insight Conference (HullPIC), Pavone .

Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2 ed., Springer Science+Business Media, New York.

Hollenbach, K.U., 1998. Estimating resistance and propulsion for single-screw and twin-screw ships. J. Ship Tech. Res. 45, 72–76.

Ikonomakis, A., Nielsen, U.D., Holst, K., Dietz, J., Galeazzi, R., 2021. How good is the stw sensor? an account from a larger shipping company. Journal of Marine Science and Engineering 9.

IMO, 2017. Draft revised guidelines for determining minimum propulsion power to maintain the manoeuvrability of ships in adverse conditions. MEPC 71/INF.28. London, UK .

IMO, 2020. Fourth imo ghg study 2020, full report. `https://www.imo.org/en/OurWork/Environment/Pages/Fourth-IMO-Greenhouse-Gas-Study-2020.aspx` (Accessed on 16-08-2022) .

IMO, 2021. Imo working group agrees guidelines to support new ghg measures. `https://www.imo.org/en/MediaCentre/PressBriefings/pages/ISWG-GHG-8.aspx` (Accessed on 11-07-2022) .

ISO, 2015. Iso 15016: Ships and marine technology — guidelines for the assessment of speed and power performance by analysis of speed trial data. International Organization for Standardization, Geneva .

ISO, 2016. Iso 19030: Ships and marine technology – measurement of changes in hull and propeller performance. International Organization for Standardization, Geneva .

ITTC, 2017. 1978 ittc performance prediction method. International Towing Tank Conference - Recommended Procedures and Guidelines .

Jung, H., Ju, J., Jung, M., Kim, J., 2016. Less-forgetting learning in deep neural networks. arXiv doi:doi:10.48550/ARXIV.1607.00122.

Kingma, D.P., Ba, J., 2015. Adam: a method for stochastic optimization. Proc. of 3rd Int. Conf. on Learning Representations (ICLR '15), San Diego .

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D., Hadsell, R., 2017. Overcoming catastrophic forgetting in neural networks. Proc. of the National Academy of Sciences 114, 3521–3526. doi:doi:10.1073/pnas.1611835114.

Kristensen, H.O.H., Lützen, M., 2013. Prediction of Resistance and Propulsion Power of Ships. Project no. 2010-56, Emissionsbeslutningsstøttesystem Work Package 2, Report no. 04.

Lang, X., Wu, D., Mao, W., 2022. Comparison of supervised machine learning methods to predict ship propulsion power at sea. Ocean Engineering 245, 110387.

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.

Liu, S., Papanikolaou, A., 2020. Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. J. Ocean Eng. 206, 107357. doi:doi:https://doi.org/10.1016/j.oceaneng.2020.107357.

Liu, S., Papanikolaou, A., Bezunartea-Barrio, A., Shang, B., Sreedharan, M., 2021. On the effect of bio-fouling on the minimum propulsion power of ships for safe navigation in realistic conditions. Biofouling doi:doi:10.1080/08927014.2021.1890044.

Losing, V., Hammer, B., Wersing, H., 2018. Tackling heterogeneous concept drift with the self-adjusting memory (sam). Knowledge and Information Systems 54, 171–201.

Mirzadeh, S., Farajtabar, M., Ghasemzadeh, H., 2020. Dropout as an implicit gating mechanism for continual learning. CoRR abs/2004.11545. URL: https://arxiv.org/abs/2004.11545.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., 2022a. Data-driven prediction of added-wave resistance on ships in oblique waves - a comparison between tree-based ensemble methods and artificial neural networks. Applied Ocean Research 118, 102964. doi:doi:https://doi.org/10.1016/j.apor.2021.102964.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., 2022b. The prediction of sea state parameters by deep learning techniques using ship motion data. Proc. of 7th World Maritime Technology Conference (WMTC'22), Copenhagen, Denmark .

Mittendorf, M., Nielsen, U.D., Bingham, H.B., 2023. Synthetic performance data for the kvlcc2. Technical University of Denmark. Dataset. doi:doi:https://doi.org/10.11583/DTU.21750257.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., Gundermann, D., Schmode, D., Deymier, C., 2022c. Performance analysis of a gas carrier using continual learning in a data stream context. Proc. of the 7th Hull Performance and Insight Conference (HullPIC'22), Tullamore .

Mittendorf, M., Nielsen, U.D., Bingham, H.B., Liu, S., 2022d. Towards the uncertainty quantification of semi-empirical formulas applied to the added resistance of ships in waves of arbitrary heading. Ocean Engineering 251, 111040. doi:doi:10.1016/j.oceaneng.2022.111040.

Mittendorf, M., Nielsen, U.D., Bingham, H.B., Storhaug, G., 2022e. Sea state identification using machine learning - a comparative study based on in-service data from a container vessel. Marine Structures 85, 103274. doi:doi:10.1016/j.marstruc.2022.103274.

Molland, A., Turnock, A., Hudson, D., 2011. Ship Resistance and Propulsion - Practical Estimation of Ship Propulsive Power. Cambridge University Press.

Parisi, G., Kemker, R., L., P.J., Kanan, C., Wermter, S., 2019. Continual lifelong learning with neural networks: A review. Neural Networks 113, 54–71.

Park, D.M., Kim, Y., Seo, M.G., Lee, J., 2016. Study on added resistance of a tanker in head waves at different drafts. Ocean Engineering 111, 569–581.

Park, J., Kim, B., Shim, H., Ahn, K., Park, J., Jeong, K., Jeong, S., 2018. Hull and propeller fouling decomposition and its prediction based on machine learning approach. Proc. of the third Hull Performance and Insight Conference (HullPIC), Redworth .

Pedersen, B., Larsen, J., 2009. Prediction of full-scale propulsion power using artificial neural networks. Proc. of 8th COMPIT Conference, Budapest .

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, 64–72.

Pierson, W., Moskowitz, L., 1964. A proposed spectral form for fully developed wind seas based on the similarity theory of s. a. kitaigorodskii. Journal of geophysical research 69, 5181–5190.

Ramasesh, V., Dyer, E., Raghu, M., 2021. Anatomy of catastrophic forgetting: Hidden representations and task semantics. Proc. of 9th Int. Conf. on Learning Representations (ICLR'21), Virtual .

Rusu, A.A., Rabinowitz, N.C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., Hadsell, R., 2016. Progressive neural networks. URL: https://arxiv.org/abs/1606.04671.

Schmode, D., Antola, M., 2020. Fusion of high-frequency navigational data and noon-reported data to predict hull condition. Proc. of the 5th Hull Performance and Insight Conference (HullPIC'20), Hamburg .

Schultz, M.P., 2007. Effects of coating roughness and biofouling on ship resistance and powering. Biofouling 23, 331–341.

Shigunov, V., 2017. Added power in seaway. Ship Technology Research 64, 65–75. doi:doi:10.1080/09377255.2017.1331953.

SIMMAN, 2014. Preprints of workshop proceedings. SIMMAN 2014, Lyngby, Denmark, April 14-16 .

Sogihara, N., 2021. Uncertainty assessment in ship performance evaluation by monte carlo simulation using onboard monitoring data. Journal of the Japan Society of Naval Architects and Ocean Engineers 33, 35–45. doi:doi:10.2534/jjasnaoe.33.35.

Sussmann, H.J., 1992. Uniqueness of the weights for minimal feedforward nets with a given input-output map. Neural Networks 5, 589–593.

Taskar, B., Andersen, P., 2019. Ship simulation workbench: User manual. Technical University of Denmark, Department of Mechanical Engineering .

Tsompopoulou, E., Athanassopoulos, A., Sivena, E., Polymenakos, K., Tsarsitalidis, V., Nikitakis, A., 2022. On the evaluation of uncertainty of ai models for ship powering and its effect on power estimates for non-ideal conditions. Proc. of the 7th Hull Performance and Insight Conference (HullPIC'22), Tullamore .

Wang, S., Psaraftis, H.N., Qi, J., 2021. Paradox of international maritime organization's carbon intensity indicator. Communications in Transportation Research 1, 100005.

Yebra, D.M., Kiil, S., Dam-Johansen, K., 2004. Antifouling technology—past, present and future steps towards efficient and environmentally friendly antifouling coatings. Progress in Organic Coatings 50, 75–104.

Yu, J.W., Lee, C.M., Seo, J.H., Chun, H.H., Choi, J.E., Lee, I., 2021. Comparative study on the prediction of speed-power-rpm of the kvlcc2 in regular head waves using model tests. International Journal of Naval Architecture and Ocean Engineering 13, 24–34.
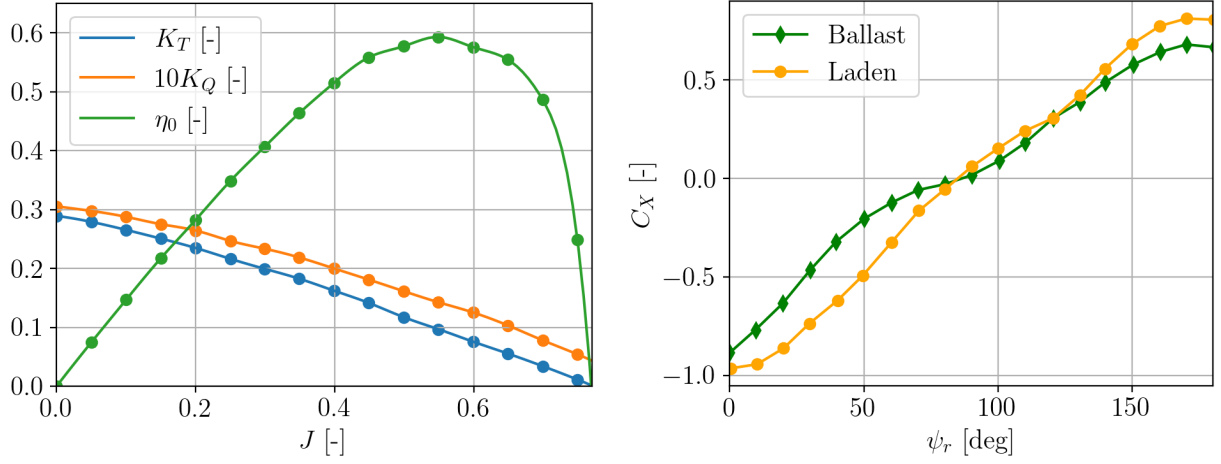
(a) Open water diagram of the KVLCC2 propeller in model scale with the advance ratio J, taken from SIMMAN (2014)

(b) Wind resistance coefficients C_X in laden and ballast conditions for the wind direction, taken from ISO (2015).

Figure 15: Additional reference data for the KVLCC2 from the public domain.

Table 6: $\widetilde{RMSE}$ of predictions using the training dataset in case of sudden concept drift.

	Default	Freezing	L_2	EWC
$\mathcal{D}_A$	0.0517	**0.0494**	0.0515	0.0514
$\mathcal{D}_B$	0.0263	0.0225	**0.0225**	0.0227
$\mathcal{D}_C$	0.0275	**0.0239**	0.0270	0.0273

Appendices

A1 Additional reference data for the KVLCC2

The open water propeller curves for the KVLCC2 are taken from SIMMAN (2014) and presented in Fig. 15a. In addition, the wind resistance coefficients are taken from ISO 15016 ISO (2015) and the non-dimensional data for a 280,000 tdw tanker (conventional bow) is depicted in Fig. 15b

A2 Performance on training data

The accuracy on the training sets for sudden and incremental concept drift are displayed in Tabs. 6 and 7, respectively.

Table 7: $\widetilde{RMSE}$ of predictions using the training dataset in case of incremental concept drift.

	Default	Freezing	L_2	EWC
$\mathcal{D}_A$	0.0454	**0.0454**	0.0617	0.0572
$\mathcal{D}_B$	0.0228	**0.0217**	0.0378	0.0384
$\mathcal{D}_C$	0.0226	**0.0215**	0.0365	0.0354

Chapter 7

Conclusion

Shipping is the backbone of the global economy and among the most important industries in Denmark. In particular, during the last couple of years and the severe disruptions in the worldwide supply chains, it became painfully evident how quintessential accessibility and affordability of maritime transport are for economic growth and globalization. Hence, decarbonizing shipping will be a balancing act and demands a wide range of innovative but cost-effective approaches. The optimization of ship operations is undoubtedly among these measures since it allows for sizable emission and cost savings with a low barrier to entry. Digitalization is and will be an accelerating driving factor for enhancing fleet operations, as demonstrated in the present thesis. Lastly, it is concluded that machine learning techniques combined with physical insights allow for greater predictability of the behavior of ocean-going ships, increasing safety and energy efficiency. Hence, it is firmly believed that the present thesis contributed to some degree to the Sustainability Development Goals (SDG) postulated by the United Nations. The work of the present thesis contributed in particular to (8) *Decent work and economic growth*, (9) *Industry, Innovation and Infrastructure*, (12) *Responsible Consumption and Production*, (13) *Climate action*, and (14) *Life below water*.

7.1 Summary

This thesis is founded on statistical analysis – including machine learning – and ship hydrodynamics in waves. The classical problem of added resistance on ships in seafaring conditions was addressed and several different aspects were elucidated – following Eq. 2.6 as the central equation. For this purpose, the present work in divided into four parts:

1. Initially, the practical estimation of the added resistance transfer function in arbitrary wave heading was examined in two separate studies. On the one hand, artificial neural networks were compared to tree-based ensemble methods using numerical results to predict R_{AW} in regular waves. In doing so, the significance of data processing and the optimization of hyperparameters stood out. Moreover, tree-based methods showed inferior characteristics in terms of generalization in comparison to neural networks. On the other hand, concepts from machine learning theory were applied for calibrating and extending an established semi-empirical formula for estimating the quadratic R_{AW}

transfer function. The methodology is considered a hybrid approach and led to increased transparency due to implementing prediction intervals via quantile regression. Considering both studies, it is concluded that the added wave resistance remains challenging – especially in short waves and beam-to-following waves demanding additional experimental and numerical investigations.

2. Deep neural networks were trained for estimating the integral sea state parameters $\{H_s, T_p, \beta_0\}$ based on in-situ wave and ship response data obtained on a container ship operating in the Northern Atlantic. Two main findings were the superior properties of employing frequency domain features as well as the application of multi-task learning for an additional increase in accuracy and efficiency. Overall, deep learning turned out to be accurate and versatile but computationally costly and greedy in terms of reliable training data, which was herein provided by an X-band wave radar.

3. The findings from the previous two parts were utilized for calculating the mean added resistance $\bar{R}_{AW}$ in a seaway using in-service data from a container ship fleet. Theoretical spectral estimates were compared to in-direct predictions derived from the measured shaft power and empirical resistance decomposition. The latter procedure is subject to large uncertainty and hampered by a scarcity of available reference data, e.g. sea trial curves. Altogether, it was confirmed that the added resistance of slender vessels – including container ships – is complex due to profound non-linear effects. Moreover, the use of weather routing by the respective shipping company reduced the likelihood of harsh weather conditions and thereby higher $\bar{R}_{AW}$ values drastically.

4. The final part was devoted to machine learning-based monitoring of the hydrodynamic performance of ships and resembles the industrial application case. In this respect, an adaptive training methodology for a neural network predicting shaft power was presented, and a proxy variable for biofouling accumulation was derived implicitly. The developed procedure builds on incremental learning and has been applied in two separate contributions. The use of simulation data in $[J_5]$ allowed for a more rigorous assessment of model behavior and uncertainty influences. Generally, the procedure yields satisfactory results and shows good potential for further extensions, e.g. assessing the added power in a seaway.

In his landmark paper "The two cultures", Breiman [16] compares data and algorithmic modeling. The former corresponds to traditional statistical methods, such as linear regression, and the latter refers to machine learning. Their benefits and shortcomings are outlined, and Breiman [16] supposes that science progresses from simple to more complex models with an increased gathering of knowledge and data. Another aspect is that model complexity and prediction accuracy are (to a certain extent) correlated; however, this comes at the cost of transparency, as transparency and model complexity are inversely proportional. It is believed that this led to the "black box" notion, which is thought (by the author) to be a rather unscientific expression. In fact, it has been shown throughout this thesis via, e.g. sensitivity studies, how predictive and inferential[1] transparency can be improved. Even

---

[1]Prediction refers to providing an estimate for one sample and inference means deriving basic knowledge about the data-generating process using the entire dataset.

though it is conceded that such efforts are generally more tedious and can be subject to uncertainties. Domain knowledge in both the field of application (ship hydrodynamics herein) *and* statistical theory are of utmost importance for producing reliable and trustworthy results by machine learning. In general, machine learning is viewed as an interdisciplinary field combining mathematics, computer science, and domain knowledge, where the latter is seen as *the* major aspect.

Interestingly, Breiman [16] treats data and algorithmic modeling as separate when in fact, both obey very similar assumptions and mathematics. For instance, in $[J_1]$, it was shown that the feature engineering aspects originating from linear regression turned out to be beneficial when using machine learning, such as using an exponential data transformation or eliminating collinear features. In $[J_5]$, it has been shown that the assumption of steady, i.e. independent and identically distributed, data is still relevant when using advanced deep learning concepts. Another common ground is the sensitivity to the availability and quality of data. Following the principles of empirical risk minimization, the data has to be considered a representation of the underlying data-generating process and therefore data-driven methods in general inherit possible insufficiencies of the training data. This, in turn, motivates approaches that are not solely data-driven but complemented by physics-guided mechanisms. Overall, machine learning is a rapidly evolving field. In the realm of engineering and ship hydrodynamics specifically, the application of physics-informed modeling and the use of explainable methods show great potential for mitigating issues caused by a lack of or corrupted data. Ultimately, the author is convinced that machine learning combined with physical procedures will enhance not only ship operations but also ship design.

7.2 Present Contributions

As previously mentioned, Breiman [16] states that science evolves from simple to more complex theories by gathering more knowledge (and data). Hence, one can draw parallels between numerical fluid dynamics and machine learning. In essence, the emergence of machine learning and CFD methods in the last decades are comparable under particular assumptions. It is believed that machine and deep learning concepts have to undergo a similar maturation process to find adoption and trust in research and, eventually, in the industry. It is believed that benchmark studies, such as the SHOPERA benchmark study (cf. Shigunov *et al.* [112]) and the development of best practice guidelines regarding discretization and used terminology led to increased acceptance. Additionally, the openness regarding mesh resolution, such as $y+$ or grid dependency studies, made CFD more trustworthy. In the end, both open-sourcing, e.g. OpenFOAM, and the experience in academia and practice turned CFD (and RANS) into a tool suitable for industrial use. As a matter of fact, an initial goal of this Ph.D. project was to contribute to the maturation process of machine learning techniques in the academic and industrial maritime field. Therefore, all publications and most of the developed methods/data are publicly available. In the following, the core novelties of the present thesis are listed with an emphasis on algorithmic transparency and reproducibility.

    – In $[J_1]$, the rigorous comparison of artificial neural networks to non-parametric en-

semble methods is considered novel for the prediction of the R_{AW} transfer function. In addition, a comprehensive feature engineering and hyperparameter optimization methodology was applied and presented.

– The main novelty in [J_2] is a generic procedure for the calibration and uncertainty quantification of existing empirical methods using a meta-heuristic optimizer and a sufficient set of experimental (or numerical) results. In addition, the derived adapted semi-empirical formula is publicly available[2] and was implemented in an updated version of the DTU in-house ship simulation workbench[3].

– Both the use of multi-task learning and the comparison of frequency and time domain features for machine learning-based sea state identification are original contributions in [J_3]. In addition, applying the wave buoy analogy using neural networks to in-situ data is considered rare in the present literature.

– The sheer scale of the study shown in [J_4] can be essentially seen as a novelty. Apart from that, the detailed modeling of the wave environment as well as the comparison of spectral and indirect estimation of the mean added resistance in irregular waves, are other important aspects of this journal article.

– In both [C_2] and [J_5], an adaptive incremental learning strategy is proposed, and the degree of biofouling is determined implicitly. The concept is, in fact, adopted from state space modeling but novel in the field of machine learning. Additionally, an alternative method for deriving the added power in a seaway using a similar approach is mentioned in the present thesis. The synthetic performance dataset for the case vessel is publicly available for benchmarking, and reproducibility, Mittendorf *et al.* [88].

Gibney [37] warns that machine learning-based research is on the verge of a "reproducibility crisis". Contributing factors are bad practices during data preprocessing or the lacking ability to verify code snippets during peer review, according to Gibney [37]. In fact, it is believed by the author that all of these aspects are just *symptoms* of a lack of maturity of machine learning in applied research areas, such as naval architecture. By many, the emergence (and increased adoption) of machine and deep learning is credited to the development of convolutional and recurrent layers allowing for multidimensional inputs and improved scalability. However, the presumably greatest breakthrough was that major tech companies open-sourced their machine learning frameworks, which sparked a vast range of applications in multiple domains. For instance, Google and Meta published TensorFlow in 2015 and PyTorch in 2016, respectively. However, this "gift" can be seen as a double-edged sword since the barrier to entry has been lowered substantially, but the application of machine learning demands a thorough understanding of the mathematical and statistical fundamentals, which has been partially omitted in earlier papers leading to reproducibility and subsequently credibility issues. Still, it is believed that open-sourcing of algorithms

---

[2]`https://gitlab.gbar.dtu.dk/mamit/RAW_Formula`
[3]`https://www.ssw.mek.dtu.dk/`

and knowledge – in the form of conference or journal articles – will eventually close these knowledge and experience gaps.

7.3 Outlook and Future Work

Initially, it is hypothesized that machine learning can be seen just as another tool, which is still in its infant stage – at least in the field of ship hydrodynamics. The previous chapters of this thesis outlined multiple shortcomings of the presented work and, in parallel, several aspects for extending work. In the present section, it is the aim to address how data- and physics-driven models can be further developed and merged into hybrid approaches. It is believed that machine learning methods used for digital representations of ships may benefit significantly from being interwoven with physical models for enhanced transparency and accuracy. Even in terms of sensitivity to data quality, a physical foundation might increase overall robustness. For example, in the case of $[J_3]$, an input from a model-based method could have led to improved results or, at least, to the observation that the wave radar was not calibrated correctly. In this respect, applying physics-informed methods, as shown by Karniadakis *et al.* [62], will presumably play a major role in the maturation process of machine learning within the realm of naval architecture. Another aspect for increasing acceptance and trust is the provision of uncertainty estimates and transparent metrics for a rigorous (and straightforward) evaluation of the presented method. The use of quantile regression in $[J_2]$ is believed to be a first step in this direction. Applying probabilistic machine learning models is another aspect of future work, but the increased computational cost may be a limiting factor in practical application. The field of explainable artificial intelligence tries to enhance overall interpretability by providing importance estimates of each feature for the current prediction.

According to LeCun *et al.* [76], progress in the field of deep unsupervised learning has been overshadowed by the great success (and simplicity) of supervised learning. In fact, humans mostly learn in an unsupervised manner, and in the present thesis, unsupervised methods have not been investigated; hence, it will be a promising part of future work to examine the suitability of un- or semi-supervised methodologies both in the context of sea state monitoring and added resistance. This is mainly because the ground truth is, in both cases, complex to measure or extract and most cases unavailable. However, applying unsupervised methods is not straightforward, and the degree of information that can be extracted may be limited (and uncertain). Meta-learning is another learning paradigm, which is receiving increasing attention, and aims at building a model capable of performing multiple tasks – comparable to $[J_3]$ – considering only minimal labeled training examples. In fact, the GPT3 model builds on a large transformer architecture trained via few-shot learning, which is a sub-field of meta-learning, Brown *et al.* [18]. Ultimately, these novel learning frameworks will allow not only for data *efficiency* but also increased generalization. In particular, the former is of great interest in the field of ship in-service analyses as reliable reference data is typically scarce.

For so-called digital twins of vessels (and in general), it becomes clear that not only adaptability, as shown in $[J_5]$, but also multi-modality regarding the output (as in $[J_3]$) and data sources (or input branches) bears sizable potential. For this reason, it could be inter-

esting to merge performance data and ship response data into one *monolithic* architecture, which not only estimates shaft power but also identifies the sea state. In essence, both tasks are physically interrelated and hence lead in combination to greater accuracy. In parallel to the hydrodynamic performance of the ship, it is a crucial aspect of ship safety also to cover structural health monitoring and the accumulation of structural fatigue – especially in the case of container vessels due to their large deck openings. Karvelis *et al.* [63] show an initial study in this respect, feeding acoustic sensor data into deep learning models for fatigue assessment. Moreover, the use of higher-order spectral analysis for the assessment of added wave resistance could be facilitated by using deep learning techniques. For instance, Hasselmann [43] examines non-linear aspects of added resistance in irregular waves using bi-spectra, which preserve the phase information, but are not commonly used in practice. Hence, this rather immature research niche could be addressed using convolutional neural networks for extracting more and human-readable information. Finally, it is concluded that ship hydrodynamics and machine learning are complementary and share that they build upon a scientific foundation, are driven by experience in practice, and results occasionally bear a resemblance to a form of art.

Bibliography

[1] AGARWALA, P., CHHABRA, S., AND AGARWALA, N. Using digitalisation to achieve decarbonisation in the shipping industry. *Journal of International Maritime Safety, Environmental Affairs, and Shipping 5*, 4 (2021), 161–174.

[2] ALDOUS, L., SMITH, T., BUCKNALL, R., AND THOMPSON, P. Uncertainty analysis in ship performance monitoring. *Ocean Engineering 110* (2015), 29–38.

[3] ALEXANDERSSON, M. A study of methods to predict added resistance in waves. *Master thesis at KTH Royal Institute of Technology, Stockholm* (2009).

[4] AMINI-AFSHAR, M. Salvesen's method for added resistance revisited. *J. Offshore Mech. Arct. Eng. 143*, 5 (2021), 051902.

[5] ANDERSEN, I. M. V. Wind loads on post-panamax container ship. *Ocean Engineering 58* (2013), 115–134.

[6] ANTOLA, M., SOLONEN, A., AND STABOULIS, S. The art of scarcity: Combining high-frequency data with noon reports in ship modelling. *Proc. of the 2nd Hull Performance and Insight Conference (HullPIC'17), Ulrichshusen* (2017).

[7] ARDHUIN, F., STOPA, J. E., CHAPRON, B., COLLARD, F., HUSSON, R., JENSEN, R. E., JOHANNESSEN, J., MOUCHE, A., PASSARO, M., QUARTLY, G. D., SWAIL, V., AND YOUNG, I. Observing sea states. *Front. Mar. Sci. 6* (2019), 124.

[8] BENGIO, Y., LOURADOUR, J., COLLOBERT, R., AND WESTON, J. Curriculum learning. *Proc. of 26th Annual Int. Conf. on Machine Learning, pp. 41–48* (2009).

[9] BERTHELSEN, F. H., AND NIELSEN, U. D. Prediction of ships' speed-power relationship at speed intervals below the design speed. *Transportation Research Part D Transport and Environment 99* (2021), 102996.

[10] BERTRAM, V. *Practical Ship Hydrodynamics*. Butterworth-Heinemann, 2000.

[11] BERTRAM, V. Added power in waves – time to stop lying (to ourselves). *Proc. of the 1st Hull Performance and Insight Conference (HullPIC'16), Castello di Pavone* (2016).

[12] BLOK, J. J. *The Resistance Increase of a Ship in Waves*. PhD Thesis at Technical University of Delft, 1993.

[13] Boese, P. *Eine einfache Methode zur Berechnung der Widerstandserhöhung eines Schiffes im Seegang.* Schriftenreihe Schiffbau, Vol. 258, pp. 1-9, 1970.

[14] Boom, H., Huisman, H., and Mennen, F. New guidelines for speed/power trials: Level playing field established for imo eedi. *SWZ Maritime, Schip En Werf de Zee Foundation, Rotterdam* (2013).

[15] Bouman, E. A., Lindstad, E., Rialland, A. I., and Strømman, A. H. State-of-the-art technologies, measures, and potential for reducing ghg emissions from shipping – a review. *Transportation Research Part D: Transport and Environment 52* (2017), 408–421.

[16] Breiman, L. The two cultures. *Statistical Science 16*, 3 (2001), 199–215.

[17] Brix, J. E. Manoeuvring technical manual. *Seehafen Verlag, Hamburg* (1993).

[18] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. Language models are few-shot learners. In *Advances in Neural Information Processing Systems* (2020), H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33, Curran Associates, Inc., pp. 1877–1901.

[19] Cames, M., Graichen, J., Siemons, A., and Cook, V. Emission reduction targets for international aviation and shipping. *Study for the ENVI Committee, European Parliament* (2015).

[20] Carlton, J. *Marine Propellers and Propulsion.* Butterworth-Heinemann, 2019.

[21] Cepowski, T. The prediction of ship added resistance at the preliminary design stage by the use of an artificial neural network. *Ocean Engineering 195* (2020), 106657.

[22] Christensen, L. B. *Evaluating the effect of retrofits on fuel performance.* MSc Thesis. Technical University of Denmark, Lyngby, 2019.

[23] Coraddu, A., Oneto, L., Baldi, F., Cipollini, F., Atlar, M., and Savio, S. Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. *Ocean Engineering Vol. 186 pp. 106063* (2019).

[24] D'Agostini, A., Bernardino, M., and Guedes Soares, C. Projected wave storm conditions under the rcp8.5 climate change scenario in the north atlantic ocean. *Ocean Engineering 266* (2022), 112874.

[25] Dalheim, Ø. Ø., and Steen, S. Preparation of in-service measurement data for ship operation and performance analysis. *Ocean Engineering 212* (2020), 107703.

[26] DNV. Energy transition outlook 2022. *A global and regional forecast to 2050* (2022).

[27] DRUMMEN, I., AND HAGEMAN, R. B. Valid 3 – analysis report. *MARIN report no. 31200-05-PaS* (2022).

[28] DÜZ, B., MAK, B., HAGMANN, N., AND GRASSO, N. Real time estimation of local wave characteristics from ship motions using artificial neural networks. In *Proc. 14th Int. Symp. Practical Design of Ships and Other Floating Structures, Yokohama, Japan* (2019).

[29] EFTEKHAR, S. F. *Prediction of fuel consumption based on a combined naval architecture and machine learning approach.* MSc. Thesis, Technical University of Denmark, Lyngby, 2023.

[30] FALTINSEN, O. M. *Sea Loads on Ships and Offshore Structures.* Cambridge University Press, 1990.

[31] FALTINSEN, O. M. *Hydrodynamics of High-Speed Marine Vehicles.* Cambridge University Press, 2005.

[32] FALTINSEN, O. M., MINSAAS, K. J., LIAPIS, N., AND SKJØRDAL, S. O. Prediction of resistance and propulsion of a ship in a seaway. *Proc. of 13th Symposium on Naval Hydrodynamics. Tokyo (Japan): The Shipbuilding Research Association of Japan* (1980), 505–529.

[33] FUJII, H., AND TAKAHASHI, T. Experimental study on the resistance increase of a large full ship in regular oblique waves. *J. Soc. Naval Archit. Japan 1975*, 137 (1975), 132–137.

[34] FUJIWARA, T., UENO, M., AND IKEDA, Y. A new estimation method of wind forces and moment acting on ships on the basis of physical component models. *Jpn. Soc. Nav. Archit. Ocean Eng. 2* (2005), 243–255.

[35] GANGESKAR, R. Verifying high-accuracy ocean surface current measurements by x-band radar for fixed and moving installations. *IEEE Transactions on Geoscience and Remote Sensing 56*, 8 (2018), 4845–4855.

[36] GERRITSMA, J., AND BEUKELMAN, W. Analysis of the resistance increase in waves of a fast cargo-ship. *Int. Shipbuild. Prog. 18*, 217 (1972).

[37] GIBNEY, E. Is ai fuelling a reproducibility crisis in science? *Nature 608* (2022), 250–251.

[38] GLOFOULING. Analysing the impact of marine biofouling on the energy efficiency of ships and the ghg abatement potential of biofouling management measures. *GloFouling Partnerships and GIA for Marina Biosafety, London* (2022).

[39] GOODFELLOW, I., BENGIO, Y., AND COURVILLE, A. *Deep Learning.* The MIT-Press, 2016.

[40] GOODFELLOW, I. J., MIRZA, M., XIAO, D., COURVILLE, A., AND BENGIO, Y. An empirical investigation of catastrophic forgeting in gradient-based neural networks. *Proc. of Int. Conf. on Learning Representations (ICLR)* (2014).

[41] GRIM, O. *Bewegungen und Belastungen des Schiffes im Seegang.* Insititut für Schiffbau der Universität Hamburg, 1974.

[42] GULDHAMMER, H. E., AND HARVALD, S. A. Ship resistance - effect of form and principal dimensions. *Danmarks Tekniske Højskole, Akademisk Forlag* (1974).

[43] HASSELMANN, K. On non-linear ship motions in irregular waves. *Journal of Ship Research 10* (1966), 64–68.

[44] HAVELOCK, T. H. The pressure of water waves upon a fixed obstacle. *Proc. of the Royal Soc. of London, Series A, Mathematical and Physical Sciences 175*, 963 (1940), 409–421.

[45] HENGELMOLEN, V., AND WELLENS, P. R. An experimental study on added resistance focused on the effects of bow wave breaking and relative wave measurements. *International Shipbuilding Progress 69* (2022), 61–89.

[46] HERSBACH, H., BELL, B., BERRISFORD, P., BIAVATI, G., HORÁNYI, A., MUÑOZ SABATER, J., NICOLAS, J., PEUBEY, C., RADU, R., ROZUM, I., SCHEPERS, D., SIMMONS, A., SOCI, C., DEE, D., AND THÉPAUT, J.-N. Era5 hourly data on single levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 24-Feb-2022)* (2018).

[47] HOLLENBACH, K. U. Estimating resistance and propulsion for single-screw and twin-screw ships. *Ship Tech. Res. 45*, 2 (1998), 72–76.

[48] HOLT, P., AND NIELSEN, U. D. Preliminary assessment of increased main engine load as a consequence of added wave resistance in the light of minimum propulsion power. *Applied Ocean Research 108* (2021), 102543.

[49] I-TECH. Quantifying the scale of the barnacle fouling problem on the global shipping fleet. *White paper.* `https://selektope.com/wp-content/uploads/2020/12/ITECH-WHITE-PAPER_June-2020-1.pdf` *(Accessed on 19-01-2023)* (2020).

[50] IKONOMAKIS, A., NIELSEN, U. D., HOLST, K. K., DIETZ, J., AND GALEAZZI, R. How good is the stw sensor? an account from a larger shipping company. *Journal of Marine Science and Engineering 9* (2021), 465.

[51] IMO. Draft revised guidelines for determining minimum propulsion power to maintain the manoeuvrability of ships in adverse conditions. *MEPC 71/INF.28. London, UK* (2017).

[52] IMO. Initial imo strategy on reduction of ghg emissions from ships. *Resolution MEPC.304(72)* (2018).

[53] IMO. Fourth imo ghg study 2020, full report. *https://www.imo.org/en/OurWork/Environment/Pages/Fourth-IMO-Greenhouse-Gas-Study-2020.aspx (Accessed on 16-08-2022)* (2020).

[54] IMO. Imo working group agrees guidelines to support new ghg measures. *https://www.imo.org/en/MediaCentre/PressBriefings/pages/ISWG-GHG-8.aspx (Accessed on 11-07-2022)* (2021).

[55] IMO. Imo's work to cut ghg emissions from ships. *https://www.imo.org/en/MediaCentre/HotTopics/Pages/Cutting-GHG-emissions.aspx (Accessed on 15-09-2022)* (2022).

[56] ISEKI, T., AND OHTSU, K. Bayesian estimation of directional wave spectra based on ship motions. *Control Eng. Practice 8* (2000), 215–219.

[57] ISO. Iso 15016: Ships and marine technology-guidelines for the assessment of speed and power performance by analysis of speed trial data. *International Organization for Standardization, Geneva* (2015).

[58] ISO. Iso 19030: Ships and marine technology – measurement of changes in hull and propeller performance. *International Organization for Standardization, Geneva* (2016).

[59] ITTC. 1978 ittc performance prediction method. *International Towing Tank Conference - Recommended Procedures and Guidelines* (2017).

[60] JINKINE, V., AND FERDINANDE, V. A method for predicting the added resistance of fast cargo ships in head waves. *Int. Shipbuild. Prog. 21*, 238 (1974), 149–167.

[61] JONCQUEZ, S. A. G., BINGHAM, H. B., ANDERSEN, P., AND KRING, D. Validation of added resistance computations by a potential-flow boundary-element method. *Proc. of the 27th Symp. on Naval Hydrodynamics, Office of Naval Research* (2008).

[62] KARNIADAKIS, G. E., KEVREKIDIS, I. G., LU, L., PERDIKARIS, P., WANG, S., AND YANG, L. Physics-informed machine learning. *Nature Reviews Physics 3* (2021), 422—-440.

[63] KARVELIS, P., GEORGOULAS, G., KAPPATOS, V., AND STYLIOS, C. Deep machine learning for structural health monitoring on ship hulls using acoustic emission method. *Ships and Offshore Structures 16*, 4 (2020), 440–448.

[64] KASHIWAGI, M. Prediction of surge and its effect on added resistance by means of the enhanced unified theory. *Trans West-Japan Soc Nav Arch 89* (1995), 77—89.

[65] KASHIWAGI, M. Hydrodynamic study on added resistance using unsteady wave analysis. *J. Ship Res. 57*, 4 (2013), 220–240.

[66] Kashiwagi, M., Ikeda, T., and Sasakawa, T. Effects of forward speed of a ship on added resistance in waves. *Int. Journal of Offshore and Polar Eng. 20*, 3 (2010), 196–203.

[67] Kempf, G. A study of ship performance in smooth and rough water. *Trans. SNAME 44* (1936), 195—227.

[68] Kim, T., Yoo, S., and Kim, H. J. Estimation of added resistance of an lng carrier in oblique waves. *Ocean Engineering 231* (2021), 109068.

[69] Kinsman, B. Wind waves: Their generation and propagation on the ocean surface. *Englewood Cliffs, Prentice-Hall Inc., N.J.* (1965).

[70] Kreitner, J. Heave, pitch and resistance of ships in a seaway. *Trans. Institution of Naval Architects 81* (1939).

[71] Kristensen, H. O. H., and Lützen, M. *Prediction of Resistance and Propulsion Power of Ships.* Project no. 2010-56, Emissionsbeslutningsstøttesystem Work Package 2, Report no. 04, 2013.

[72] Kuroda, M., and Takagi, K. Slowly-varying added resistance in irregular waves in consideration of higher order components. *Journal of the Japan Society of Naval Architects and Ocean Engineers 28* (2018), 37–43.

[73] Kwon, Y. J. Speed loss due to added resistance in wind and waves. *Naval Architect* (2008), 14—16.

[74] Lang, X., and Mao, W. A practical speed loss prediction model at arbitrary wave heading for ship voyage optimization. *Marine. Sci. Appl. 20* (2021), 410–425.

[75] Lang, X., Wu, D., and Mao, W. Comparison of supervised machine learning methods to predict ship propulsion power at sea. *Ocean Engineering 245* (2022), 110387.

[76] LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. *Nature 521* (2015), 436–444.

[77] Lee, J. H., Seo, M. G., Park, D. M., Yang, K. K., Kim, K. H., and Kim, Y. Study on the effects of hull form on added resistance. *Proc. of the 12th Int. Symp. on Practical Design of Ships and Other Floating Structures (PRADS), Changwon, Korea* (2013).

[78] Lindstad, E., Lagemann, B., Rialland, A., Gamlem, G., and Valland, A. Reduction of maritime ghg emissions and the potential role of e-fuels. *Transportation Research Part D: Transport and Environment 101* (2021), 103075.

[79] Liu, S., and Papanikolaou, A. Regression analysis of experimental data for added resistance in waves of arbitrary heading and development of a semi-empirical formula. *Ocean Engineering 206* (2020), 107357.

[80] LIU, S., PAPANIKOLAOU, A., AND WANG, J. Prediction of the added resistance of ships in natural seaways. *Proc. of 14th International Conference on Hydrodynamics (ICHD'22), Wuxi, China* (2022).

[81] LIU, S., PAPANIKOLAOU, A., AND ZARAPHONITIS, G. Prediction of added resistance of ships in waves. *Ocean Engineering 38* (2011), 641–650.

[82] LONGUET-HIGGINS, M. S., CARTWRIGHT, D. E., AND SMITH, N. D. Observations of the directional spectrum of sea waves using motions of a floating buoy. *Ocean Wave Spectra, Prentice Hall, New York.* (1963), 111–136.

[83] MARUO, H. The excess resistance of a ship in rough seas. *Int. Shipbuild. Prog. 4*, 35 (1957), 337–345.

[84] MARUO, H. The drift of a body floating on waves. *Journal of Ship Research 4*, 3 (1960), 1–10.

[85] MIKKELSEN, H., SHAO, Y., AND WALTHER, J. H. Numerical study of nominal wake fields of a container ship in oblique regular waves. *Applied Ocean Research 119* (2022), 102968.

[86] MITCHELL, T. M. *Machine Learning.* McGraw-Hill, New York, 1997.

[87] MITTENDORF, M., AND NIELSEN, U. D. Deep learning-based sea state estimation using sensor data of wave-induced ship responses. In *Proc. of 2nd Marine AI Seminar of WISE program, Tokyo University of Marine Science and Engineering* (2022).

[88] MITTENDORF, M., NIELSEN, U. D., AND BINGHAM, H. B. Synthetic performance data for the kvlcc2. *Technical University of Denmark. Dataset.* (2023).

[89] MITTENDORF, M., NIELSEN, U. D., BINGHAM, H. B., AND LIU, S. Uncertainty-aware prediction of added resistance using an adapted semi-empirical formula. In *DNV Nordic Maritime Universities Workshop, Horten* (2022).

[90] MITTENDORF, M., AND PAPANIKOLAOU, A. D. Hydrodynamic hull form optimisation of fast catamarans using surrogate models. *Ship Tech. Res. 68*, 1 (2021), 1–13.

[91] MOLLAND, A. F., TURNOCK, A. R., AND HUDSON, D. A. *Ship Resistance and Propulsion - Practical Estimation of Ship Propulsive Power.* Cambridge University Press, 2011.

[92] NABERGOJ, R., AND PRPIĆ-ORŠIĆ, J. A comparison of different methods for added resistance prediction. *Proc of 22nd Int. Workshop on Water Waves and Floating Bodies, Plitvice* (2007).

[93] NAKAMURA, S., AND NAITO, S. Propulsive performance of a containership in waves. *J. Soc. Nav. Archit. Jpn. 15* (1977), 24–48.

[94] NEWMAN, J. N. The drift force and moment on ships in waves. *Journal of Ship Research 11* (1967), 51–60.

[95] NIELSEN, U. D. A concise account of techniques available for shipboard sea state estimation. *Ocean Engineering 129* (2017), 352—-362.

[96] NIELSEN, U. D. Spatio-temporal variation in sea state parameters along virtual ship route paths. *Journal of Operational Oceanography* (2021), 1–18.

[97] NIELSEN, U. D., JOHANNESEN, J. R., BINGHAM, H. B., BLANKE, M., AND JON-CQUEZ, S. Indirect measurements of added-wave resistance on an in-service container ship. *Proc. 14th Int. Symp. Practical Design of Ships and Other Floating Structures (PRADS'19), Yokohama, Japan* (2019).

[98] PAPANIKOLAOU, A. Review of advanced marine vehicles concepts. *Proc. of 7th Int. High Speed Marine Vehicles Conf. (HSMV05), Naples* (2005).

[99] PAPANIKOLAOU, A., HARRIES, S., HOOIJMANS, P., MARZI, J., LE NÉNA, R., TORBEN, S., YRJÄNÄINEN, A., AND BODEN, B. A Holistic Approach to Ship Design: Tools and Applications. *Journal of Ship Research 66*, 01 (2022), 25–53.

[100] PARK, D. M., LEE, J., AND KIM, Y. Uncertainty analysis for added resistance experiment of kvlcc2 ship. *Ocean Engineering 95* (2015), 143–156.

[101] PEDERSEN, B., AND LARSEN, J. Prediction of full-scale propulsion power using artificial neural networks. *Proc. of 8th Conf. on Computer Applications and Information Technology in the Maritime Industries (COMPIT), Budapest* (2009).

[102] PERERA, L. P., AND GUEDES SOARES, C. Weather routing and safe ship handling in the future of shipping. *Ocean Engineering 130* (2017), 684–695.

[103] PEREZ ARRIBAS, F. Some methods to obtain the added resistance of a ship advancing in waves. *Ocean Engineering 34*, 7 (2007), 946–955.

[104] PIERSON, W., AND MOSKOWITZ, L. A proposed spectral form for fully developed wind seas based on the similarity theory of s. a. kitaigorodskii. *Journal of geophysical research 69*, 24 (1964), 5181–5190.

[105] RUMELHART, D., HINTON, G., AND WILLIAMS, R. Learning representations by back-propagating errors. *Nature 323* (1986), 533–536.

[106] SADAT-HOSSEINI, H., WU, P., CARRICA, P., KIM, H., TODA, Y., AND STERN, F. Cfd verification and validation of added resistance and motions of kvlcc2 with fixed and free surge in short and long head waves. *Ocean Engineering 59* (2013), 240–273.

[107] SAETTONE, S. *Ship Propulsion Hydrodynamics in Waves.* Ph.D. thesis, DTU Mechanical Engineering, 2020.

[108] SALVESEN, N. Added resistance of ships in waves. *J. Hydronautics 12*, 1 (1978), 21–34.

[109] SALVESEN, N., TUCK, E., AND FALTINSEN, O. M. Ship motions and sea loads. *Trans. SNAME 78* (1970), 250–287.

[110] SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM J. Res. Develop. 3*, 3 (1959), 210–229.

[111] SHIGUNOV, V. Added power in seaway. *Ship Technology Research 64*, 2 (2017), 65–75.

[112] SHIGUNOV, V., EL MOCTAR, O., PAPANIKOLAOU, A., POTTHOFF, R., AND LIU, S. International benchmark study on numerical simulation methods for prediction of maneuverability of ships in waves. *Ocean Engineering 165* (2018), 365–385.

[113] SIGMUND, S. *Performance of Ships in Waves*. PhD thesis, University of Duisburg-Essen, 2018.

[114] SKEJIC, R., AND FALTINSEN, O. M. A unified seakeeping and maneuvering analysis of ships in regular waves. *Journal of Marine Science and Technology 13* (2008), 371–394.

[115] SPRENGER, F., MARON, A., DELEFORTRIE, G., VAN ZWIJNSVOORDE, T., CURA-HOCHBAUM, A., LENGWINAT, A., AND PAPANIKOLAOU, A. Experimental studies on seakeeping and maneuverability of ships in adverse weather conditions. *Journal of Ship Research 61*, 3 (2017), 131–152.

[116] STORHAUG, G., MOE, E., AND PIEDRAS LOPES, T. A. Whipping measurements onboard a midsize container vessel operating in the north atlantic. *Marintec China Proceedings (RINA, CMP, and SNAME)* (2007).

[117] STRØM-TEJSEN, J., YEH, H. Y. H., AND MORAN, D. D. Added resistance in waves. *Society of Naval Architects and Marine Engineers Transactions 81* (1973), 109–143.

[118] SÖDING, H., AND SHIGUNOV, V. Added resistance of ships in waves. *Ship Tech. Res. 62*, 1 (2015), 2–13.

[119] SÖDING, H., SHIGUNOV, V., SCHELLIN, T. E., AND EL MOCTAR, O. A rankine panel method for added resistance of ships in waves. *Journal of Offshore Mechanics and Arctic Engineering 136*, 3 (2014), 031601.

[120] TASKAR, B., AND ANDERSEN, P. Ship simulation workbench: User manual. *Technical University of Denmark, Department of Mechanical Engineering* (2019).

[121] TASKAR, B., AND ANDERSEN, P. Benefit of speed reduction for ships in different weather conditions. *Transportation Research Part D: Transport and Environment 85* (2020), 102337.

[122] THORNHILL, D. C., AND STREDULINSKY, E. M. Ship motion and wave radar data fusion for shipboard wave measurement. *Ship Research 55*, 2 (2012), 73–85.

[123] TSUJIMOTO, M., SHIBATA, K., KURODA, M., AND TAKAGI, K. A practical correction method for added resistance in waves. *Japan Soc. Naval Arch. and Ocean Eng. 8* (2008), 177–184.

[124] TU, F., SAM GE, S., SANG CHOO, Y., AND CHIEH HANG, C. Sea state identification based on vessel motion response learning via multi-layer classifiers. *Ocean Engineering 147* (2018), 318–332.

[125] TURING, M. A. Computing machinery and intelligence. *Mind 59*, 236 (1950), 433–460.

[126] UHAREK, S. A. *Numerical Prediction of Ship Manoeuvering Performance in Waves*. Ph.D. Thesis, Technical University of Berlin, 2019.

[127] UNCTAD. Review of maritime transport 2022, full report. *https://unctad.org/ system/files/official-document/rmt2022_en.pdf (Accessed on 07-12-2022)* (2022).

[128] VALANTO, P., AND HONG, Y. Experimental investigation of ship waves added resistance in regular head, oblique, beam and following waves. *Proc. of 25th Int. Ocean and Polar Engineering Conf. (ISOPE'15), Hawaii, USA* (2015).

[129] VETTOR, R., AND GUEDES SOARES, C. Reflecting the uncertainties of ensemble weather forecasts on the predictions of ship fuel consumption. *Ocean Engineering 250* (2022), 111009.

[130] VITALI, N., PRPIĆ-ORŠIĆ, J., AND GUEDES SOARES, C. Coupling voyage and weather data to estimate speed loss of container ships in realistic conditions. *Ocean Engineering 210* (2020), 106758.

[131] WANG, S., PSARAFTIS, H. N., AND QI, J. Paradox of international maritime organization's carbon intensity indicator. *Communications in Transportation Research 1* (2021), 100005.

[132] YU, J. W., LEE, C. M., LEE, I., AND CHOI, J. E. Bow hull-form optimization in waves of a 66,000 dwt bulk carrier. *Int. J. Naval. Archit. Ocean. Eng. 9*, 5 (2017), 499–508.

[133] YU, J.-W., LEE, C.-M., SEO, J.-H., CHUN, H. H., CHOI, J.-E., AND LEE, I. Comparative study on the prediction of speed-power-rpm of the kvlcc2 in regular head waves using model tests. *International Journal of Naval Architecture and Ocean Engineering 13* (2021), 24–34.

List of Figures

List of Tables

Nomenclature

Symbol	Description	Dimension
B	Beam of the ship	m
$C_{AW} = \frac{R_{AW}L_{pp}}{\rho g \zeta_0^2 B^2}$	Added resistance coefficient	
C_B	Block coefficient	
D	Spreading function	
E	Directional wave energy density spectrum	m^2/rad Hz
f	Intrinsic wave frequency	Hz
F	Unidirectional wave energy density spectrum	m^2 Hz
$Fn = \frac{U}{\sqrt{gL}}$	Froude number	
g	Gravitational acceleration constant	m/s^2
G	Center of gravity	
H_s	Significant wave height	m
J	Cost function, Advance ratio	
k	Wave number	m^{-1}
L_{pp}	Length per perpendicular	m
m_n	Spectral moment of order n	
$\boldsymbol{n}$	Normal vector	
n	Number of features	
N	Yaw moment	Nm
m	Sample size	
P_B	Engine brake power	kW
p	Pressure, joint distribution	
R_{AA}	Air resistance in calm water	N
R_{AW}	Wave resistance in regular waves	N
$\bar{R}_{AW}$	Mean wave resistance in irregular waves	N
R_F	Frictional resistance	N
R_T	Total ship resistance	N
$R_{T,calm}$	Total ship resistance in calm water	N
R_{VP}	Viscous pressure resistance	N
R_W	Wave resistance in calm water	N
R_{wind}	Added wind resistance	N
$Rn = \frac{UL}{\nu}$	Reynolds number	
rpm	Revolutions per minute	min^{-1}
S	Wetted surface area	m^2

Symbol	Description	Dimension
T	Ship draft	m
T_a	Ship draft in aft position	m
T_f	Ship draft in fore position	m
T_m	Ship mean draft	m
T_p	Peak period	s
T_z	Zero up-crossing period	s
$t = 1 - \frac{R_T}{T}$	Thrust deduction coefficient	
U	Ship advance speed	m/s
V_w	Absolute wind speed	m/s
$w = 1 - \frac{U_A}{U_S}$	Wake fraction coefficient	
$\boldsymbol{x}$	Independent variable (Feature vector)	
y	Dependent variable	
$\hat{y}$	Approximation of independent variable	
β	Relative wave direction	deg.
β_0	Mean encounter wave direction	deg.
δ	Rudder angle	deg.
ϵ	Phase angle, residual	
ζ	Wave elevation	m
ζ_0	Wave amplitude	m
η	Efficiency	
$\boldsymbol{\theta}$	Model parameter vector	
λ	Wave length	m
μ	Absolute mean wave direction	deg.
ν	Kinematic viscosity	m^2/s
ξ	Drift angle	deg.
ρ	Seawater density	kg/m^3
ρ_{air}	Air density	kg/m^3
σ^2	Variance	
$\boldsymbol{\tau_w}$	Shear stress	MPa
ϕ	Activation function	
ψ	Absolute wind direction	deg.
ω	Intrinsic angular wave frequency	rad/s
ω_e	Encounter wave frequency	rad/s
∇	Hull volume, Gradient	
$\mathcal{L}$	Loss function	
$\mathbb{E}$	Expectation	

Abbreviations

ANN	Artificial Neural Network
AIS	Automatic Identification System
CFD	Computational Fluid Dynamics
CPU	Central Processing Unit
CII	Carbon Intensity Indicator
CL	Continual Learning
CNN	Convolutional Neural Network
CO$_2$	Carbon dioxide
DNN	Deep Neural Network
DOF	Degrees of Freedom
DTU	Danmarks Tekniske Universitet
ECMWF	European Centre for Medium-Range Weather Forecasts
EEDI	Energy Efficiency Design Index
EEXI	Energy Efficiency eXisting ship Index
ERA5	Fifth generation ECMWF atmospheric reanalysis of global climate
ESD	Energy Saving Device
EU	European Union
GHG	Greenhouse gas
GPS	Global Positioning System
GPU	Graphics Processing Unit
HSVA	Hamburgische Schiffbauversuchsanstalt, Germany
IMO	International Maritime Organization
IOT	Internet of Things
ISO	International Organization for Standardization
ITTC	International Towing Tank Conference
JIP	Joint Industry Project
JONSWAP	Joint North Sea Wave Project
KRISO	Korea Research Institute of Ships & Ocean Engineering
LNG	Liquefied Natural Gas
LSTM	Long Short Term Memory
MCR	Maximum Continuous Rating
MARPOL	International Convention for the Prevention of Pollution from Ships
MEPC	Maritime Environmental Protection Committee
ML	Machine Learning
NMRI	National Maritime Research Institute, Japan

PSO	Particle Swarm Optimization
RANS	Reynolds-averaged Navier Stokes
RAO	Response Amplitude Operator
SEEMP	Ship Energy Efficiency Management Plan
SHOPERA	Energy Efficient Safe Ship Operation
SOG	Speed Over Ground
SSW	Ship Simulation Workbench
STA	Sea Trial Analysis
STW	Speed Through Water
TEU	Twenty-foot-Equivalent Unit
TUMSAT	Tokyo University of Marine Science and Technology, Japan
UNCTAD	United Nations Conference on Trade and Development
VLCC	Very Large Crude-oil Carrier