International Association of Maritime Universities

IAMU 2021 Research Project
(No. 20210205)

Data fusion and machine learning for ship fuel efficiency analysis: a small but essential step towards green shipping through data analytics

Theme 2: Future opportunities and challenges of the sustainability of maritime industry

By

Australian Maritime College, University of Tasmania

August 2022

IAMU
International Association of Maritime Universities

This report is published as part of the 2021 Research Project in the 2021 Capacity Building Project of International Association of Maritime Universities, which is fully supported by The Nippon Foundation.

The text of the paper in this volume was set by the author. Only minor corrections to the text pertaining to style and/or formatting may have been carried out by the editors.

All rights reserved. Due attention is requested to copyright in terms of copying, and please inform us in advance whenever you plan to reproduce the same.

The text of the paper in this volume may be used for research, teaching and private study purposes.

No responsibility is assumed by the Publisher, the Editor and Author for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in this book.

Editorial

IAMU Academic Affairs Committee (AAC)

Head of Committee : Professor Dr. Adam Weintrit
Rector, Gdynia Maritime University (GMU)

Editorial committee : Adam Przybylowski (GMU)
Avtandil Gegenava (BSMA)
Christian Matthews (LJMU)

Contractor : Nigel Blundell

Research Coordinator : Yuquan (Bill) Du

Published by the International Association of Maritime Universities (IAMU) Secretariat
Meiwa Building 8F, 1-15-10 Toranomon, Minato-ku,
Tokyo 105-0001, JAPAN
TEL : 81-3-6257-1812 E-mail : info@iamu-edu.org URL : http://www.iamu-edu.org

Copyright ©IAMU 2022
All rights reserved
ISBN978-4-907408-38-1

International Association of Maritime Universities

IAMU 2021 Research Project
(No. 20210205)

Data fusion and machine learning for ship fuel efficiency analysis: a small but essential step towards green shipping through data analytics

Theme 2: Future opportunities and challenges of the sustainability of maritime industry

By
Australian Maritime College, University of Tasmania

Contractor : Nigel Blundell
Research Coordinator : Yuquan (Bill) Du
Research Partner : Shu-Ling Chen, AMC
Nataliya Nikolova, AMC
Prashant Bhaskar, AMC
Jiangang Fei, AMC
Alessandro Schönborn, WMU
Zhuo Sun, DMU

Data fusion and machine learning for ship fuel efficiency analysis: a small but essential step towards green shipping through data analytics

Theme 2:

Future opportunities and challenges of the sustainability of maritime industry

Australian Maritime College, University of Tasmania

Research Coordinator

Dr Yuquan (Bill) Du
Senior Lecturer, Australian Maritime College, University of Tasmania, Australia
yuquan. du@utas.edu.au

Abstract The shipping industry is concerned about ship fuel/energy efficiency due to the motivation of saving bunker fuel cost and mitigating ship emissions. A foundation for various energy/emission-efficient measures is the accurate quantification of bunker fuel consumption of a ship in one day or hour given its sailing speed, draft/displacement, trim, weather conditions, and sea conditions. This study takes advantage of four industry data sources including voyage report data, AIS data, sensor data, and meteorological data, and fuses these data sources to find the best datasets for ship fuel efficiency analysis. Based on fused datasets, we experimented with state-of-the-art machine learning models to quantify a ship's daily/hourly bunker fuel consumption, over eight 8,100-TEU to 14,000-TEU containerships of a global shipping company. When voyage report data is used as the basis for ship fuel/emission analysis, meteorological data and AIS data can be combined into voyage report data to improve the data quality. The fit errors of best machine learning models over the recommended datasets are normally within 5 ton/day, and can be as low as less than 1 ton/day. When sensor data is considered, combining meteorological data (waves, sea currents, sea water temperature) into sensor data will significantly improve the modeling accuracy. The best machine learning models achieve their R^2 at 0.999 or 1.000 on the training sets, and their R^2 values over the test sets are also all above 0.966. Their fit errors are below 0.75 ton/day (RMSE), or below 0.52 ton/day (MAE). The proposed datasets and models would be useful for sailing speed optimization, trim optimization, weather routing, voyage planning, and virtual (just-in-time) arrivals. We also published our computer code in Python and trained machine learning models in GitHub which is accessible to the public.

Keywords: *Ship fuel efficiency, Fuel consumption rate, Voyage report, Sensor data, AIS data, Meteorological data, Data fusion, Machine learning*

Executive Summary

With promotions of the International Maritime Organization (IMO) and governmental organizations, the shipping industry has been implementing operational measures to save bunker fuel and mitigate emissions from ships, including sailing speed optimization, trim optimization, weather routing, and the virtual arrival policy. Many frustrations have been emerging during the process of implementation of these measures. These frustrations are boiled down, if not fully, to how we can quantify the synergetic contributions of many factors (speed, draft/displacement, trim, weather conditions, sea conditions) on a ship's bunker fuel consumption rate (ton/day, ton/hour). A latest review paper points out that the basis of all operational measures for ship bunker fuel savings and emission mitigation is quantitatively modeling the relationship between fuel consumption rate and many determinants, including sailing speed, draft/displacement, trim, weather conditions, and sea conditions. **This project addresses this theoretical challenge** that restricts the implementation of energy-efficient operational measures by investigating the complementary roles of different data sources available to a shipping company, fusing these data sources, and employing state-of-the-art machine learning techniques.

We collected **voyage report data** and **sensor data** of eight 8,100-TEU to 14,000-TEU containerships from a global shipping company, purchased the **AIS data** of these ships from *MarineTraffic* with the financial support of the International Association of Maritime Universities (IAMU), and downloaded **meteorological data** from European Centre for Medium-range Weather Forecasts (ECMWF) and Copernicus Marine Service (CMEMS). Based on the information contained in these four data sources, we designated **three data fusion solutions (DFSs)**: **DFS1** fuses voyage report data and meteorological data, by considering the inaccurate information of weather and sea conditions recorded by voyage report; **DFS2** further fuses AIS data into voyage report data and meteorological data because AIS data helps find the actual sailing trajectory of the ship and thus helps retrieve more accurate information of weather and sea conditions from meteorological data; **DFS3** approaches sensor data as the main data source of a ship's fuel consumption rate, and overcomes the limitation of sensor data by taking advantage of the complete information of weather and sea conditions contained in meteorological data. **For each of the data fusion solutions, eight to nine datasets are constructed**.

Over these datasets from three data fusion solutions, **a large range of widely adopted machine learning models were experimented with**, including decision tree-based models, artificial neural network (ANN), support vector machine (SVM), ridge regression (Ridge), and LASSO. Tree-based models include the basic decision tree (DT) model and models produced by two ensemble strategies: Extremely randomized trees (ET) and random forest (RF) from the bagging ensemble strategy; AdaBoost (AB), gradient tree boosting (GB), XGBoost (XG), and LightGBM (LB) from the boosting ensemble strategy. During the experiments with these machine learning models, the impacts of data normalization, hyperparameter optimization, and the randomness in splitting training sets and test sets are well addressed.

Extensive experiments were conducted to answer three research questions regarding the choice of datasets from three data fusion solutions and the choice of machine learning models. A voting scheme is developed to break down the impacts of dataset choice and model choice. When dataset choice is considered, the original voyage report dataset ***Set1*** has a decent quality for ship fuel efficiency modeling; if more effort is paid to fuse voyage report data and meteorological data, data quality improves slightly and $\boldsymbol{Set3_{precise}}$ can be adopted. When AIS data is available, further including AIS data might also be beneficial, which suggests the adoption of the dataset $\boldsymbol{AIS5_{precise.}}$ Overall, the best datasets found with DFS1 and DFS2, including ***Set1***, $\boldsymbol{Set3_{precise}}$, and $\boldsymbol{AIS5_{precise}}$, ensure accurate fit performances of best ML models: R^2 on the training set is above 0.96 and even reaches 0.99 to 1.00, and R^2 on the test set is between 0.74 and 0.90; the fit errors measured by RMSE (Root Mean Square Error) and MAE (Mean Absolute Error) are between 0.5 and 4.5 ton/day. When sensor data, rather than voyage report data, is used as the main data source of ship bunker fuel consumption analysis, it elevates the modeling accuracy to a higher level, possibly the highest level if meteorological data is fused in. With DFS3, given the best dataset ***Sensor2***, the best ML models achieve their R^2 values over the training set at 0.999 or 1.000, and their R^2 values over the test set are all above 0.966. Their fit errors with RMSE values are below 0.75 ton/day, and with MAE below 0.52 ton/day.

As far as ML model choice is concerned, we **recommend the installation of four decision-tree based models including ET, AB, GB, and XG** because they usually possess the highest fit performance and good generalization performance. Their performances are also quite robust against random splits of a dataset into training and test sets. Our experiments with DFS1, DFS2, and DFS2 reach consistent findings about the performances of ML models and rank their performances into four tiers.

- Tier 1: ET, AB, GB, and XG.
- Tier 2: RF, LB
- Tier 3: DT, SVM, ANN
- Tier 4: Ridge, LASSO.

Apart from this **technical report**, as the **research outcomes**, this project produces three research papers that have been submitted to a peer-review journal. We have also developed **course material (teaching slides) for a three-hour teaching module for IMO's TTT course** on Energy Efficient Ship Operations, titled "Understanding ship fuel efficiency with real data". To broadcast our research findings to the maritime industry, we delivered **three industry presentations to industry professionals in May 2022 in Europe, Australia, and Asia**, respectively. **Computer code in Python in this study is published in GitHub as a software infrastructure** to reduce the exploration efforts of industry professionals. **Best trained machine learning models are also published in GitHub**, which enables maritime researchers to estimate the bunker fuel consumption rates of different sizes of mega containerships in different sailing speed, draft, trim and weather/sea conditions. Reader can find our Python code and trained machine learning models in the URL below:

https://github.com/yuquandu/Data-driven-Ship-Fuel-Efficiency-Modeling

Our recommendations for industry application are summarized in the following table.

Industry applications	• Sailing speed optimization • Weather routing • Virtual (just-in-time) arrival	Trim optimization
Industry stakeholders	• Shipping companies • Weather information service providers (WISPs) • Ship classification societies (such as ClassNK) • Shipping associations (such as BIMCO)	Shipping companies
Recommended data sources and datasets	• DFS1: Voyage report data + meteorological data • DFS2: Voyage report data + meteorological data + AIS data	DFS3: sensor data + meteorological data
Recommended models	Extremely randomized trees (ETs), Gradient tree boosting (GB), or XGBoost (XG)	Extremely randomized trees (ETs), Gradient tree boosting (GB), or XGBoost (XG)

Acknowledgements

This study is a part of the IAMU (International Association of Maritime Universities) research project titled "Data fusion and machine learning for ship fuel efficiency analysis: a small but essential step towards green shipping through data analytics" (Research Project No. 20210205_AMC). The materials and data in this publication have been obtained through the support of IAMU and The Nippon Foundation in Japan. Jean-Louis Bertholier developed the Python code of collecting meteorological data for ships during his Assistant Engineer internship at World Maritime University. This study has been conducted using E.U. Copernicus Marine Service Information; https://doi.org/10.48670/moi-00050. Hersbach et al. (2018) was downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store. The results of this study and trained machine learning models published contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

The content of this technical report was submitted as the following three papers to *Communications in Transportation Research*. This report is written based on our preprint version of these three papers but does not incorporate any of the modifications made as a result of the review process.

Xiaohe Li, Yuquan Du, Yanyu Chen, Son Nguyen, Wei Zhang, Alessandro Schönborn, Zhuo Sun, 2022. "Data fusion and machine learning for ship fuel efficiency modeling: Part I – voyage report data and meteorological data". Submitted to Communications in Transportation Research.

Yuquan Du, Yanyu Chen, Xiaohe Li, Alessandro Schönborn, Zhuo Sun, 2022a. "Data fusion and machine learning for ship fuel efficiency modeling: Part II – voyage report data, AIS data and meteorological data". Submitted to Communications in Transportation Research.

Yuquan Du, Xiaohe Li, Yanyu Chen, Alessandro Schönborn, Zhuo Sun, 2022b. "Data fusion and machine learning for ship fuel efficiency modeling: Part III – sensor data and meteorological data". Submitted to Communications in Transportation Research.

1. Introduction

1.1 Background and Research Questions

Reducing bunker fuel consumption of ships is paramount for the shipping industry with both commercial and environmental implications. Shipping companies have been always striving to reduce their bunker fuel costs of their fleets in marine operations because bunker fuel cost typically accounts for about 20% to 61% of a ship's operating costs (Meng et al., 2016; Soner et al., 2018). Meanwhile, reduction in bunker fuel consumption lies in the core of progressively stricter regulations on ship emissions proposed by the International Maritime Organization (IMO, 2020) and other international or national organizations such as European Union (EU, 2021), because ship emissions, especially CO2, NOx and SOx, are proportional to the bunker fuel consumption (Adland et al., 2019).

Shipping industry stakeholders, such as shipping companies, IMO, EU, and other governmental organizations, are making unprecedented efforts to reduce bunker fuel consumptions of ships and the accompanying emissions. Due to the expensiveness of technical solutions, shipping companies have been passionate in adopting various operational measures to reduce bunker fuel consumption, including weather/environmental routing, speed optimization, and trim optimization, and virtual (just-in-time) arrival policy (IMO, 2012; Coraddu et al., 2017; Li et al., 2018; Wan et al, 2018; Merkel et al., 2022). IMO has been calling on the shipping industry to implement the Data Collection System (EEOI, AER, DIST, TIME), Energy Efficiency Design Index (EEDI), Ship Energy Efficiency Management Plan (SEEMP), and in-progress Energy Efficiency eXisting ship Index (EEXI) and Carbon Intensity Index (CII) (Wang et al., 2021; Yan et al., 2021). EU also rolled out its Monitoring, Reporting and Verification (MRV) system from 2018.

However, during this process, many frustrations are heard from the shipping industry. In sailing speed optimization, a ship's fuel efficiencies, usually measured as its fuel consumption rate in terms of metric ton (MT) per hour, or MT per day, in different weather and sea conditions are hard to captured by deck officers. Therefore, a simple sprint-and-loiter practice is widely adopted (Johnson and Andersson, 2011; C-MAP, 2022). Regarding trim optimization, it is believed that trim optimization can save 4-6% (even up to 15%) of bunker fuel, according to various reports issued by IMO and DNV. However, our collaboration with some shipping companies received many complaints about the current trim optimization practice. Captains at sea feel that trim charts/tables/matrices based on model ship tests or computational fluid dynamics (CFD) calculation are not convincing, because these trim charts/tables/matrices cannot reflect the influence of weather and sea conditions on trim optimization and the suggested optimal trim value sometimes even cannot guarantee the full submergence of the propellor in sea water. Third, our discussion with seafarers also saw their complaints about the weather routing services provided by Weather Information Service Providers (WISPs). The weather routing services of WISPs are expensive, and the data transferred is often outdated or delayed. Therefore, many deck officers having been relying more and more on manual voyage/route planning with the assistance of real-time weather information websites, such as Windy.com. Fourth, when it turns to the virtual (just-in-time) arrival policy, Rehmatulla et al. (2017), Adland et al. (2020) and Merkel et al. (2022) report that a major barrier to this policy is quantitively assessing the bunker fuel consumption in different speed-weather scenarios and precisely calculating the cost savings of the policy for each voyage.

All these frustrations are boiled down, if not fully, to how we can quantify the synergetic contributions of many factors (speed, draft/displacement, trim, weather conditions, sea conditions) on a ship's bunker fuel consumption rate. A latest review paper, Yan et al. (2021), also points out that the basis of all operational measures for ship bunker fuel savings and emission mitigation is the quantitatively modeling the relationship between fuel consumption rate and many determinants, including sailing speed, draft/displacement, trim, weather conditions, and sea conditions, but it is not a trivial work.

As stated by Yan et al. (2021), there are two elementary factors that determine the accuracy of ship fuel efficiency modeling: choice of data, and choice of models. There are several sets of data sources that can support ship fuel efficiency modelling of a shipping company: voyage report data, sensor data, automatic identification system (AIS) data, ship mechanical data, ship maintenance data, and meteorological data. Haranen et al. (2016) and Yan et al. (2021) categorize ship fuel efficiency models as three classes: white-box models (WBMs), black-box models (BBMs), and grey-box models (GBMs), and discusses the advantages and disadvantages of each model class and the importance of selecting specific models.

The systematic review of Yan et al. (2021) summarizes the existing research efforts of data collection and ship fuel efficiency analysis with varieties of models, especially machine learning (ML) models. However, few of them consider the complementary role of different data sources. For instance, the quality of voyage report data about snapshotted weather and sea conditions is questionable, but this might be remedied by the publicly accessible meteorological data, such as the data of wind, waves, and sea water temperature from European Centre for Medium-range Weather Forecasts (ECMWF), and the data of sea currents from Copernicus Marine Service (CMEMS). Meanwhile, through AIS data, we can access the sailing trajectory of a ship over a day and the data about the positions of the ship might help us to find more accurate weather and sea conditional data from meteorological data. As another example, sensor data provides high-quality information on a ship's sailing profile including wind conditions, but the information of waves, sea water temperature, and sea currents is often absent. This may be complemented by the detailed meteorological data that are publicly accessible.

Therefore, the following research questions (RQs) could be asked by both academics and industry professionals:

- **RQ1**. Is it possible to combine/fuse different but complementary data sources for the sake of modeling accuracy for ship fuel efficiency analysis? And how these data sources can be fused?
- **RQ2**. Compared to a single data source, what are the benefits of fusing different data sources in terms of modeling accuracy and generalization?
- **RQ3**. Selection of datasets and choice of models are two different decision dimensions but they rely on each other. When these two decisions are interwoven, how can we select the best datasets and best models?

1.2 Research Outputs

This project produces the following research outcomes.

- **Three papers** submitted to the peer-reviewed journal *Communications in Transportation Research*.

  Xiaohe Li, Yuquan Du, Yanyu Chen, Son Nguyen, Wei Zhang, Alessandro Schönborn, Zhuo Sun, 2022. "Data fusion and machine learning for ship fuel efficiency modeling: Part I – voyage report data and meteorological data". Submitted to Communications in Transportation Research.

  Yuquan Du, Yanyu Chen, Xiaohe Li, Alessandro Schönborn, Zhuo Sun, 2022a. "Data fusion and machine learning for ship fuel efficiency modeling: Part II – voyage report data, AIS data and meteorological data". Submitted to Communications in Transportation Research.

Yuquan Du, Xiaohe Li, Yanyu Chen, Alessandro Schönborn, Zhuo Sun, 2022b. "Data fusion and machine learning for ship fuel efficiency modeling: Part III – sensor data and meteorological data". Submitted to Communications in Transportation Research.

- **This final report** that is written based on the preprint version of the above three papers.
- **Software infrastructure and trained machine learning models**. Computer code in Python in this study is published in GitHub as a software infrastructure to reduce the exploration efforts of industry professionals. Best trained machine learning models are also published in GitHub, which enables maritime researchers to estimate the bunker fuel consumption rates of different sizes of mega containerships in different sailing speed, draft, trim and weather/sea conditions, though our raw data are confidential. The machine learning models published are completely black boxes, and one cannot conduct reverse engineering to access the original datasets. Readers can find the computer code and trained machine learning models in the URL: https://github.com/yuquandu/Data-driven-Ship-Fuel-Efficiency-Modeling.
- **Course material contribution to IMO's TTT course**. To enhance the impacts of this IAMU project, we have developed course material (teaching slides) for a three-hour teaching module for IMO's TTT course on Energy Efficient Ship Operations, titled "Understanding ship fuel efficiency with real data". All the major methodologies and experimental findings have been included in these slides. We have asked our research partner at World Maritime University (WMU) to bring the knowledge found by this project to IMO's TTT Course. It is intended that the knowledge created by this research project will be passed onto future generations through teaching.
- **Three industry presentations**. The research findings in this project were also broadcast to industry professionals through three industry presentations. **Presentation 1** was conducted by our research partner at WMU on 10 May 2022 to CETENA (a maritime research & training company based in Italy). **Presentation 2** was conducted by our AMC (Australian Maritime College) side on 27 May 2022 to Australian Maritime Logistics Research Network. This presentation was joined by senior professionals from Marine Operations Department of Australian National Line and industry professionals from Shipping Australia Limited. **Presentation 3** was conducted by our research partner at DMU (Dalian Maritime University) on 29 May 2022 through a webinar to a wide cohort of maritime industry professionals in Asia.

1.3 Structure of the Report

The remainder of this report is organized as follows. Section 2 conducts a literature review and clarifies the research gaps. Section 3 summarizes our research efforts, discusses our contributions to existing literature, and clearly defines our research scope/boundary. Section 4 discusses the four data sources utilized by this project, including voyage report data, sensor data, AIS data, and meteorological data. Section 5 presents the technical details of machine learning models used by this project. Section 6 proposes the approach of fusing voyage report data and meteorological data and discusses the experimental findings. Section 7 proposes the approach of fusing voyage report data, AIS data, and meteorological data and discusses the experimental findings. Section 8 proposes the approach of fusing sensor data and meteorological data and discusses the experimental findings. Section 9 summarizes the findings of the whole project, proposes the recommendations for industry applications, and discusses the limitations of the studies in this project.

2. Literature Review and Research Gaps

2.1 Literature Review

Our studies focus on accurately modeling the relationship between ship fuel consumption rate (MT/h or MT/day) and several determinants, including sailing speed, draft/displacement, trim, weather conditions, and sea conditions, by using machine learning models. In this regard, Yan et al. (2021) conduct a systematic literature review for academic papers and technical reports published from 2008 (one year before the implementation of IMO EEOI) and 2021. In this taxonomy, machine learning models is one of the two types of BBMs, in parallel with statistical BBMs. To avoid duplicating the systematic review of Yan et al. (2021), we will only have a quick review about the BBM literature that involve two or more data sources, because our studies are addressing the research questions about the benefits of fusing several data sources and machine learning models.

Bocchetti et al. (2015) collect the data of a cruise ship from voyage report (a.k.a. noon report) and onboard sensors about ship maintenance and operations and sea and weather conditions, and develop a multiple linear regression (MLR) model. Their research purpose is to predict the fuel consumption of this cruise ship in a voyage, rather than that in a day or hour. Meanwhile, a systematic query is absent to how to select the best dataset by considering all the possible datasets that can be produced by voyage report and sensor data.

Adland et al. (2018) consider voyage report and hull maintenance data of a fleet of eight sister Aframax crude oil tankers, and perform a MLR analysis on fuel consumption rate. Their research purpose is to assess the impact of hull cleaning on ship fuel efficiency and thus combine the voyage reports of eight ships together. This is different from our studies that aim to build ship-specific fuel efficiency models for the applications of daily marine operations at sea, relying on daily operational data sources including voyage reports, sensor data, AIS data, and meteorological data.

As far as we know, Lee et al. (2018) is the first attempt to combine two daily operational data sources at sea for ship fuel consumption rate estimation. They fuse the data about voyages and meteorological data from CMEMS and develop a data mining algorithm that mines the impact of wind on ship fuel consumption rate. However, the adopted data about voyages is not voyage report data but "voyage abstract data" in which there is only one data entry for each voyage. This limitation on data availability makes the authors rely on a polynomial regression model of ship fuel efficiency proposed by Yao et al. (2012).

Gkerekos et al. (2019) utilize voyage report data and the data from an automated data logging & monitoring (ADLM) system. The data from the ADLM system is sensor based but its frequency, hourly, is lower than traditional sensor data which generally has a data entry about every 10-15 minutes. Meanwhile, they regard voyage report and ADLM data as two independent data sources and their purpose is to compare the performance of machine learning models on these two different data sources. The possibility of fusing different data sources is not discussed.

Man et al. (2020) make pioneering efforts to fuse different data sources by considering five ferries and collecting their sensor data, AIS data, meteorological data, and the captains' log on the estimated time of arrival (ETA) and summarized fuel consumption in each journey. Though four data sources are mentioned, their study mainly combines sensor data and meteorological data. Their AIS data from a Swedish company is not reliable to track the ship probably because the voyages of ferries between Gothenburg and Kiel are rather short compared to commercial cargo ships at open sea. This paralyzes the main advantage of AIS and make them approach a linear interpolation method to calculate the sailing trajectory of these ferries. Six datasets are produced after data fusion and they are tested with a multi-layer perceptron model and a self-organizing map model. The prediction target of their machine learning models is fuel consumption in a journey, rather than fuel consumption rate, which is different from our studies and from most studies reviewed by Yan et al. (2021). Their data structure and the nature of short sea sailing of the five ferries under investigation could challenge the applicability of their data fusion plans and experimental findings to the shipping practice of cargo ships such as containerships and oil tankers.

Farag and Ölçer (2020) adopt an artificial neural network (ANN) model to estimate a tanker ship's brake power based on serval determinants such as sailing speed and weather and sea conditions. They utilize a dataset provided by NAPA Group that is extracted from the ship's automatic continuous monitoring system (ACMS), AIS data, and meteorological data, but NAPA hides the details on how these data sources are combined.

Uyanık et al. (2020) combine voyage report and sensor data and populate 75 variables/features into their machine learning models. This is appropriate because their research purpose is to monitor engine performance and their models will be used by engine rooms. This is significantly different from our studies that target ship fuel efficiency model to be used by deck officers and captains for their daily sailing planning.

2.2 Research Gaps

Contrasting the research questions proposed in Section 1 and literature review conducted in Section 2.1, we can easily see the following research gaps posed by existing literature:

- Existing studies of ship fuel efficiency analysis that combine/fuse multiple data sources and explore their complementary roles are rare.
- Among these rare studies, only Lee et al. (2018), Man et al. (2020), and Uyanık et al. (2020) propose clear data fusion solutions and fuel efficiency models/algorithms from the perspective of a ship's daily sailing operation.
- To address the industry frustrations in speed optimization, trim optimization, water routing, and virtual arrival policy, a reliable model is needed that can accurately estimate a ship's bunker fuel consumption rate (MT/day, MT/h) based on several determinants outside of a ship's engine (sailing speed, draft/displacement, trim, weather conditions, and sea conditions). None of Lee et al. (2018), Man et al. (2020), and Uyanık et al. (2020) achieve this, not to mention a systematic research effort to construct promising fused datasets from voyage report, AIS data, sensor data, and meteorological data and to select the best datasets according to the fit and generalization performances of multiple machine learning models.

3. Research efforts, contributions, and scope/boundary

3.1 Research efforts and contributions

To address the research questions and gaps identified and build reliable fuel consumption rate forecast models that can be used in energy-efficient operational measures (speed optimization, trim optimization, water routing, and virtual arrival policy), we approached different industry stakeholders and collected/purchased all the four most relevant data sources that a shipping company can access, for eight modern mega containerships in different sizes: voyage report data, sensor data, AIS data, and meteorological data.

Then we analyzed the data structure of these data sources and proposed the following three possible data fusion/combination solutions, by discussing with a global shipping company, envisaging the possible industry application scenarios, and considering the endogeneity issue pointed by Yan et al. (2021):

- Data fusion solution 1 (DFS1): voyage report data + meteorological data.
- Data fusion solution 2 (DFS2): voyage report data + meteorological data + AIS data.
- Data fusion solution 3 (DFS3): sensor data + meteorological data.

For each data fusion solution, we constructed the all the possible datasets by taking into account the industry applications and the impact of endogeneity on feature/variable selection. Then we tested the fit and generalization performances of machine learning models widely adopted in literature over these possible datasets. When the decisions of dataset selection and model choice are interwoven, we adopted a voting scheme to enable machine learning models vote for best datasets.

Experiments with these industry data and machine learning models revealed many useful insights on the benefits of fusing these different data sources, selection of the best datasets, and choice of best machine learning models. Using the same ships, it also allowed us to compare the benefits of different data sources and compare the benefits of different data fusion solutions.

Towards the three data fusion solutions DFS1, DFS2, DFS3, for the first time, this project provides industry professionals with clear answers to RQ1 to RQ3 with extensive and intensive experimental evidence from different sizes of mega containerships. This project lays a solid theoretical foundation to accurately quantify a ship's fuel consumption rate in the energy-efficient operational measures being promoted by IMO, including sailing speed optimization, trim optimization, route selection (weather routing), and the virtual arrival policy.

3.2 Research scope

To avoid possible confusions, we define our research scope/boundary as follows.

(a) We only consider the fuel consumption of the main engine (M/E) of a ship, but will not consider its auxiliary engines and boilers.
(b) This project targets ship-specific fuel efficiency models, which means every model built is for a specific ship. This is different from the study that combines the data of a fleet of ships and develop a model for the fleet (Adland et al., 2018).
(c) For the purpose of applications of models in sailing speed optimization, trim optimization, route optimization, and the virtual arrival policy, our studies only adopt the features outside of a ship's mechanical system (engine and propulsion) as the input variables of a model, including sailing speed, draft/displacement, trim, and factors about weather and sea conditions. We will not consider the technical features regarding engine and propellor performance such as engine RPM (rotations per minute), engine power, shaft power, and propellor pitch. See the discussion of Yan et al. (2021) on the endogeneity issue and application scenarios of different types of models.
(d) The output/dependent variable of our model, i.e., the prediction target, is the fuel consumption rate in terms of MT/day (or equivalently MT/h), rather than fuel consumption in a voyage or journey in term of MT or specific fuel oil consumption (SFOC) in terms of g/kWh.
(e) Accordingly, only data sources relevant to a ship's voyage management and sailing behaviours will be utilised, including voyage report data, sensor data, AIS data, and meteorological data. Other data sources discussed by Yan et al. (2021) and ship fuel efficiency models based on those sources are not relevant to energy-efficiency operational measures for voyage management (speed optimization, trim optimization, route selection/weather routing, virtual arrival policy).

(f) We only test the machine learning models, especially those widely adopted in literature. We will not consider WBMs, statistical BBMs or GBMs that are discussed in Yan et al. (2021). See Yan et al. (2021) for a detailed discussion about the pros and cons of each type of models.

4. Data sources

Voyage report data, sensor data, AIS data, and meteorological data are the major data sources that a shipping company can access for the purpose of ship fuel efficiency analysis. This section discusses how we approached these four data sources and the information structure of each of data source. The particulars of eight ships whose data are utilized for experiments throughout this project are tabulated in **Table 1**.

Table 1. Particulars of eight ships used for experiments

Ship	Year built	Capacity (TEU)	Size (length/beam)	Draft recorded: Avg/Max (m)	Speed recorded: Avg/Max (knots)
S1	2013	14,000	398m/51m	13.5/25.1	13.9/23.3
S2	2013	14,000	398m/51m	14.1/21.5	13.8/23.3
S3	2012	11,000	347m/45m	13.7/23.8	12.7/23.6
S4	2012	11,000	347m/45m	12.1/15.7	12.9/24.4
S5	2013	9,200	328m/45m	11.7/19.3	12.4/24.0
S6	2014	9,200	328m/45m	12.6/23.5	12.8/22.3
S7	2013	9,200	328m/45m	12.4/17.4	12.3/23.1
S8	2013	8,100	320m/46m	12.0/22.3	12.4/23.9

Source: FleetMon.com. Accessed on 8 February 2022.

4.1 Voyage Report Data

Voyage report of a ship is a summary of the daily sailing situation submitted by the captain to the onshore officers so that the onshore officers can understand the ship's real sailing conditions. Usually, the captain will report the data at noon every day, and thus voyage report data is also called noon report data. Ship voyage reports are usually filled out manually by the crew based on the readings of the instruments on board or eye inspection based on personal experience. Voyage report data includes many sailing features of the ship, such as displacement, draft, trim, speed, true course, geographic location, Greenwich Mean Time, the fuel consumptions of the main engine, auxiliary engines, and boilers, weather conditions, and sea conditions.

Voyage report data of eight mega containerships shown in **Table 1** is provided by a global container shipping company. The sailing period recorded by the data spans from February 2014 to March 2016. A data preprocessing procedure that removes invalid data entries was employed to ensure the quality of datasets. Particularly, the data entries with N/A values, speeds below 12 knots or above 30 knots, sailing time less than 10 hours, or ship status being not "sailing at sea" were all deleted in data preprocessing. For the sailings of about two years, after preprocessing, ships S1 to S8 have 320, 296, 389, 380, 329, 402, 407 and 440 data entries, respectively, in their voyage reports.

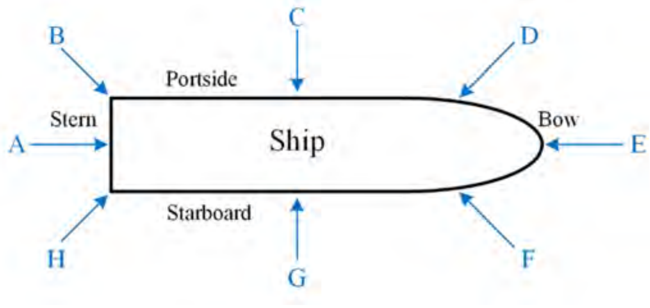
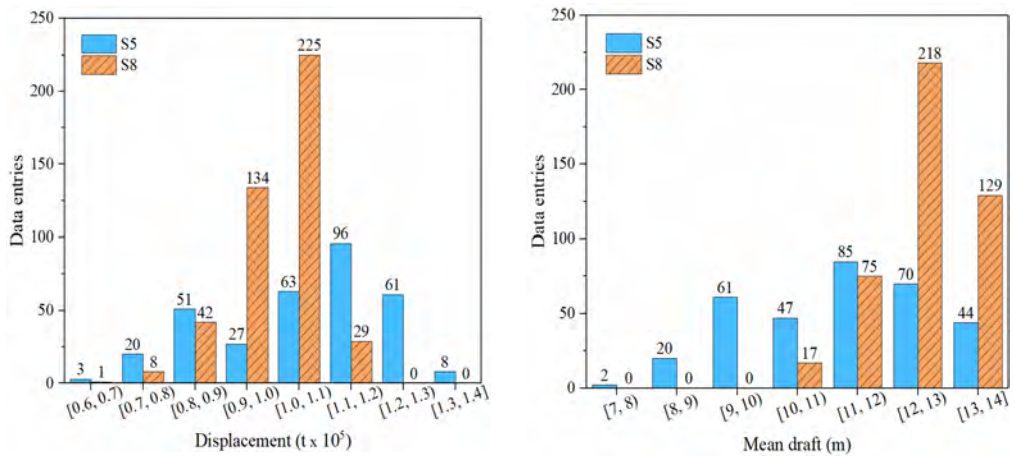


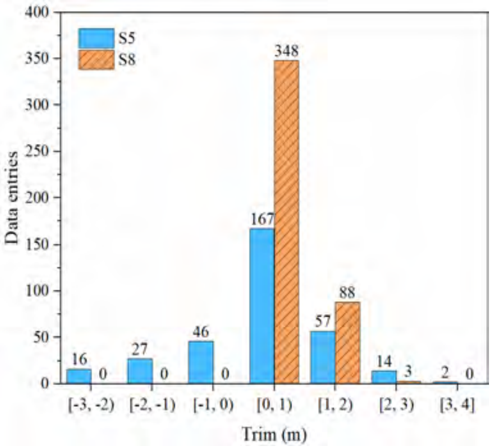
Figure 1. Illustration of wind/wave/sea current directions. Source: Meng et al. (2016b).

Motivated by the studies by Du et al. (2019), this study selects the fuel consumption rate of the main engine (t/day) in the voyage report as the output variable (target) of the ship fuel efficiency model. The input/independent variables (features) of the model includes displacement (MT) (equivalent to draft (m)), trim (m), sailing speed (knots), sea water temperature (°C), wind direction, wind force (Beaufort scale number), wave (swell) direction, wave (swell) height (m), sea current direction, and sea current speed (knots). The directions of wind, waves, and sea currents in the voyage report are recorded by the crew as fuzzy numbers denoting their approximate directions relative to the ship's heading, which are illustrated in **Figure 1**. For readers who are interested in the distributions of our voyage report data entries over these important features, see **Figure 2** for ships S5 and S8 as examples.

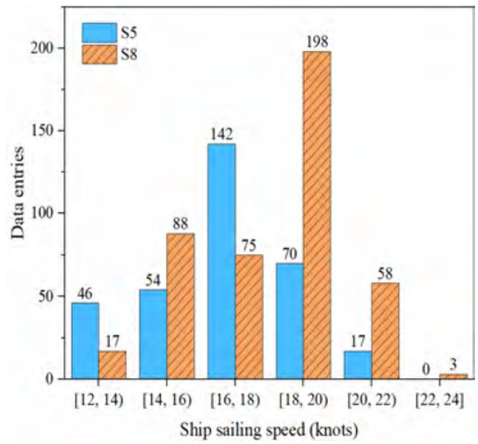


(a) Distribution of displacement

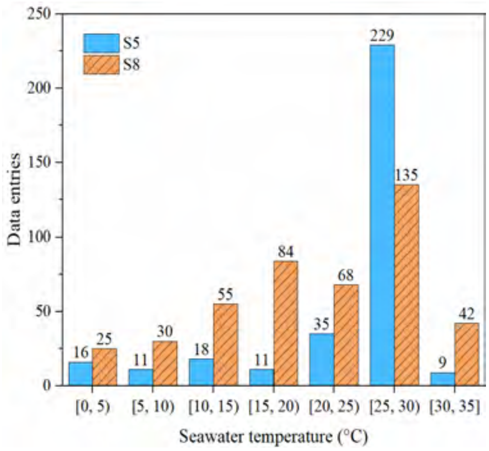
(b) Distribution of draft



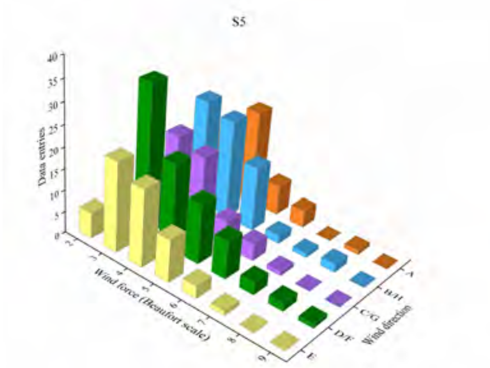
(c) Distribution of trim



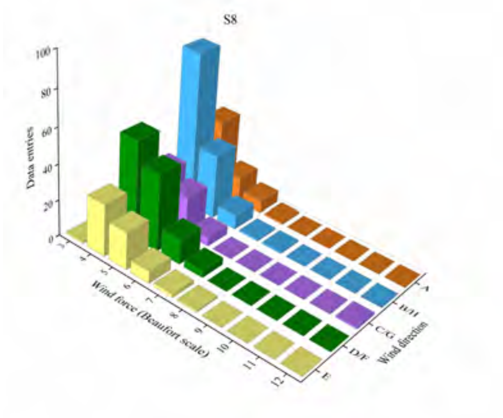
(d) Distribution of sailing speed



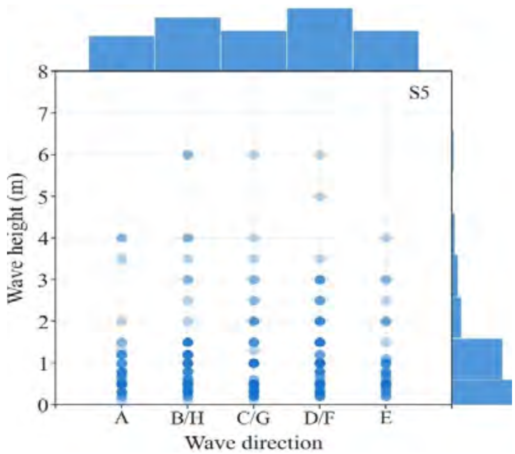
(e) Distribution of seawater temperature



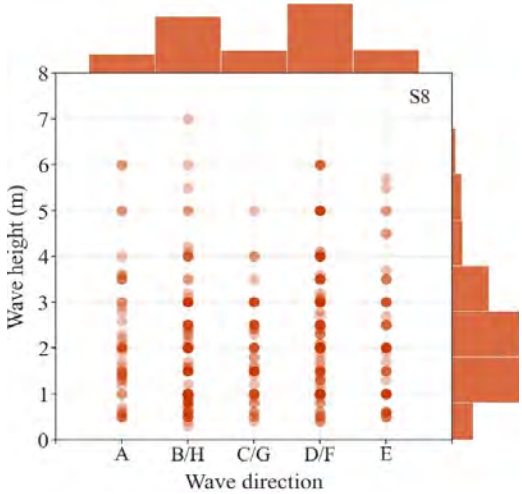
(f) Wind direction and wind force distribution of S5



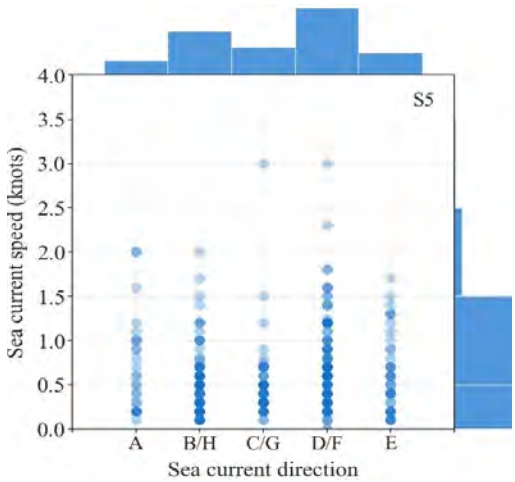
(g) Wind direction and wind force distribution of S8



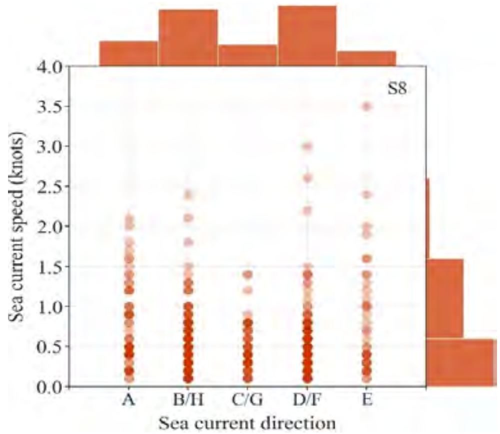
(h) Wave direction and wave height distribution of S5



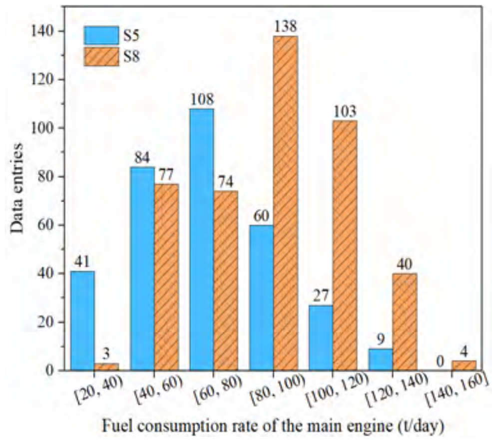
(i) Wave direction and wave height distribution of S8



(j) Sea current direction and sea current speed distribution of S5



(k) Sea current direction and sea current speed distribution of S8



(l) Distribution of fuel consumption rate of the main engine

Figure 2. Distribution of the voyage report data entries of ships S5 and S8

4.2 Meteorological Data

The information of weather and sea conditions at the hands of a shipping company may not complete or accurate. Therefore, our industry collaborators suggested us approaching some publicly accessible meteorological data sources to retrieve more reliable data of weather and sea conditions. Our research shows that European Centre for Medium-range Weather Forecasts (ECMWF) provides the finest data for wind, waves, and sea water temperature in the granularity of 0.25^{o} (longitude)A?0.25^{o} (latitude)A? A?1 hour (time), while Copernicus Marine Service (CMEMS, also a.k.a. "Copernicus") provides the finest data for sea currents in the granularity of 0.25^{o} (longitude) A? 0.25^{o} (latitude)A? A? 3 hour (time). These data sources are also adopted by *Windy.com* which is widely used by deck officers in the world for manual voyage planning.

ECMWF data on 12 variables/parameters are retrieved, including "*Significant height of combined wind waves and swell*" (paramId: 140229), "*Mean wave direction*" (paramId: 140230), "*Mean wave period*" (paramId: 140232), "*Significant height of wind waves*" (paramId: 140234), "*Mean direction of* *wind waves*" (paramId: 140235), "*Mean period of wind waves*" (paramId: 140236), "*Significant height of total swell*" (paramId: 140237), "*Mean direction of total swell*" (paramId: 140238), "*Mean period of total swell*" (paramId: 140239), "*10 metre U wind component*" (paramId: 165), "*10 metre V wind component*" (paramId: 166), "*Sea surface temperature*" (paramId: 34). Note that waves consist of two components: swell and wind waves, and ECMWF provide the information about swell, wind waves, and the combined waves calculated from these two components. 3-hourly data on sea currents are retrieved from CMEMS (Copernicus), involving two variables: *eastward_sea_water_velocity* and *northward_sea_water_velocity*.

4.3 AIS Data

With the financial support of IAMU, we purchased the AIS data of eight ships shown in Table 1 from *MarineTraffic*. The purchased AIS data has 15 columns. Apart from the identification and particulars of the ship (*MMSI*, *Call Sign*, *Ship Name*, *Flag Country*, *Draft Designed*, *Length*) and the information about the voyage (*ETA*, *Destination Port*), the AIS data contains the detailed navigation data including "*Timestamp (UTC)*", "*Navigation Status*", "*Longitude Position*", "*Latitude Position*", "*Ship Course*", "*Ship Heading*", and "*Sailing Speed*". There is a data entry every 3-5 minutes. "*Sailing Speed*" appears to be useful in our study. However, our study considers voyage report or sensor data as the main data sources of ship fuel consumption. When voyage report data is used which records information on a daily basis, "*Sailing Speed*" information in the time interval of 3-5 minutes from AIS data does not help. When sensor data is used, it already contains the accurate information (every 15 minutes) of sailing speeds of the ship, and "*Sailing Speed*" information in AIS data will not provide additional benefits.

4.4 Sensor Data

Sensor data of two mega containerships (ships S5 and S6 shown in **Table 1**) is provided by a global container shipping company. The time span of the data ranges from May to November of 2015. One sensor data entry was returned every 15 minutes, and the useful information for ship fuel efficiency modeling includes "*fuel consumption rate* (MT/day, or kg/h)", "*Sailing speed*", "*Draft*", "*Trim*", "*Wind speed*", "*Wind direction*". Ships S5 and S6 have 11,901 and 12,484 sensor data entries, respectively. The sensor data from this global shipping company, consisting of 100 columns, does not contain the information about waves, sea currents, and sea water temperature.

5. Machine Learning Models

5.1 Adopted Machine Learning Models

There are various ML methods in the field of data prediction and analysis. This study covers a large range of the current most popular and practical methods, including tree-based methods, ANN, Support Vector Machine (SVM), ridge regression (Ridge), and least absolute shrinkage and selection operator regression (LASSO). Tree-based methods can be further divided into decision tree (DT), extremely randomized trees (ETs), random forest (RF), AdaBoost (AB) (Freund and Schapire, 1995; Drucker, 1997), gradient tree boosting (GB) (Friedman, 2001), XGBoost (XG) (Chen and Guestrin, 2016), and LightGBM (LB) (Ke et al., 2017).

Model training algorithms in our studies are implemented using Python 3.7.6. The XG model is developed using the XGBoost 1.2.0 library, the LB model is developed using the LightGBM 2.3.1 library, and the remaining models are developed using the Scikit-learn 0.22.1.

5.1.1 Tree-based Methods

DT is a ML method for classification or regression (Breiman et al., 1984). The method creates a tree-structured model to learn simple decision rules from the data features (variables) to predict the value of a target variable. A DT model contains three types of nodes: root node (the topmost internal node), internal nodes, and leaf nodes (also known as terminal nodes). In a DT model, each internal node represents a judgment (test) of an attribute (i.e., values of a variable), and the judgment result is its output. These outputs are represented by branches of the tree. The judgment and output process of internal nodes is termed as splitting. In the regression algorithm, the splitting criterion of nodes is mean square error (MSE). The splitting termination condition of nodes is determined by three parameters: the tree's maximum depth (max_depth), the minimum number of samples used in the decision of splitting an internal node (min_samples_split), and the minimum number of samples contained in a leaf (min_samples_leaf). The model overfitting issue can be alleviated by setting these three parameters. Finally, each leaf node represents a classification/regression result. A simple DT structure is shown in **Figure. 3**.

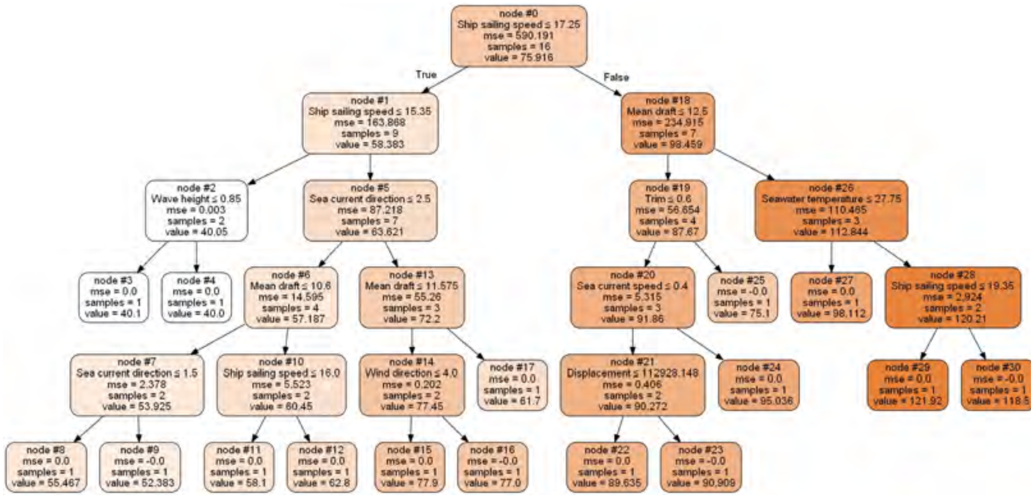
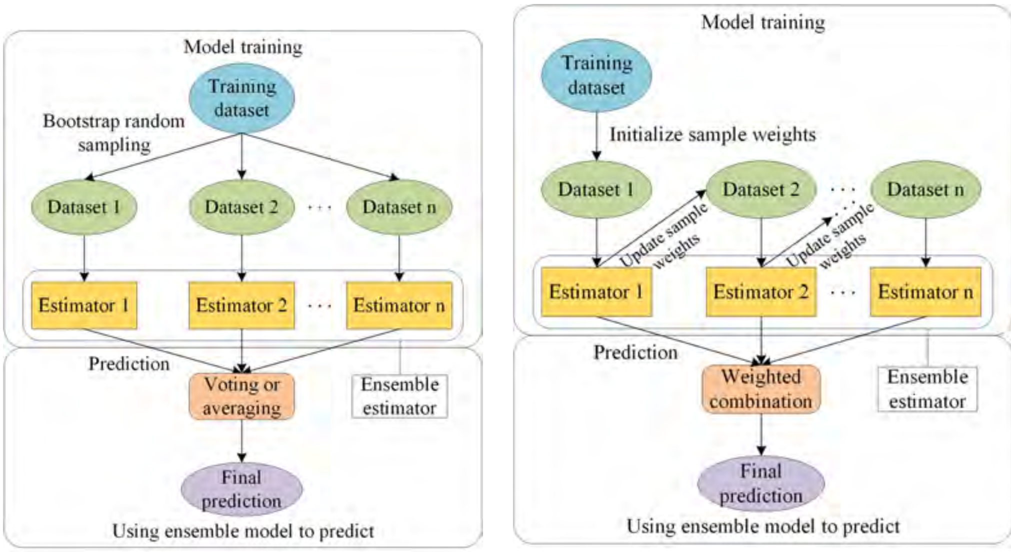


Figure 3. Visualization of DT structure.



(a) Bagging method. (b) Boosting method.

Figure 4. Ensemble strategies used in tree-based ML models. Source: KDnuggets (2017).

Ensemble methods are widely used in ML. The principle of ensemble methods is to combine the prediction results of multiple base estimators (trees) constructed using a given learning method to obtain better generalization ability/robustness than a single estimator. Bagging, also known as bootstrap aggregating, is an ensemble strategy proposed by Breiman (1996) to improve unstable estimation or classification scheme. ETs (Geurts et al., 2006) and RF (Breiman et al., 2001) are two ML models with the bagging ensemble strategy. The key idea of the bagging strategy is to build multiple independent estimators and then average their predictions, as illustrated in **Figure 4(a)**. Boosting is an ensemble strategy primarily used to reduce the model prediction bias of any given learning method. AB (Freund and Schapire, 1995; Drucker, 1997), GB (Friedman, 2001), XG (Chen and Guestrin, 2016), and LB (Ke et al., 2017) models adopt the boosting ensemble strategy. Both XG and LB are optimized gradient boosting methods, which are highly efficient implementations of GB. In the boosting strategy, the base estimator is built in sequence, as illustrated in **Figure 4(b)**.

5.1.2 ANN

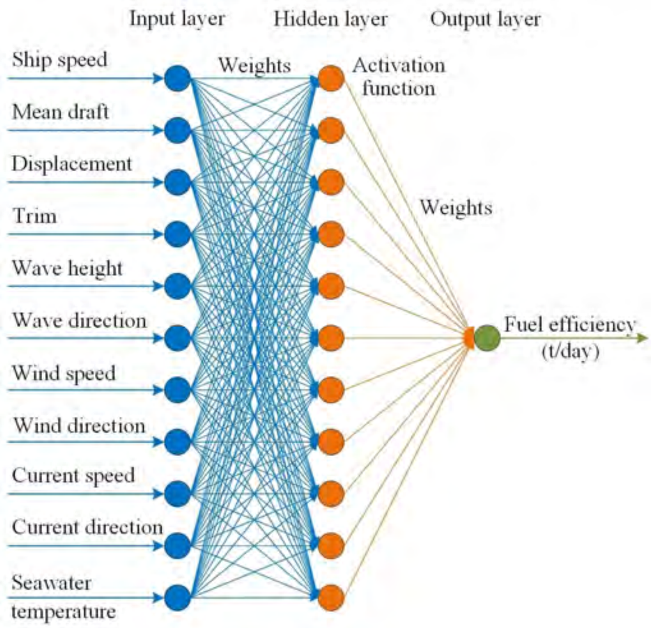


Figure 5. ANN model structure. Source: Du et al. (2019).

ANN is a widely used computing method composed of a large number of interconnected nodes (neurons). The working mechanism of these nodes imitates that of neurons in the nervous system. ANN usually distributes the neurons into three layers. The input layer receives input variables and passes the variable values to the hidden layer. The neurons in the hidden layer make a weighted linear summation of the output values from the previous layer, and then use a nonlinear activation function to transform the weighted results and transfer them to the next layer. The output layer weights and transforms the values from the last hidden layer to obtain the final output values. The working principle of the neurons in the hidden and output layers can be expressed by the following formula:

$$y_i = f(\sum_{j=1}^{m} w_{ij} \cdot x_j + b_i) \tag{1}$$

where y_i is the neuron output value; $f(\cdot)$ is the activation function; m is the number of neurons in the previous layer; w_{ij} is the network connection weight; x_j is the output values of the previous layer; b_i is the bias. The learning process of ANN is essentially the process of finding the best weight set $\{w_{ij}\}$. The most commonly used weight learning method in ANN is the gradient descent method, which continuously adjusts the weights of the network with a back-propagation algorithm to minimize the sum of squared errors (Cai et al., 2019). Readers are referred to Haykin (2008) for the technical details of ANN.

The structure of the ANN model established in existing studies related to ship fuel efficiency is usually a three-layer network structure (i.e., only one hidden layer; Hornik et al., 1989; Kolmogorov, 1957), such as the models established by Pedersen and Larsen (2009), Beşikçi et al. (2016), and Du et al. (2019). Therefore, this study also adopts a typical feedforward ANN model with a three-layer network structure (**Figure 5**). Inspired by Du et al. (2019) setting the number of neurons in the hidden layer equal to the number of input variables, and considering that the input layer has eleven input variables in this study, the number of neurons in the hidden layer is set to eleven. As described by Du et al. (2019), too many neurons in the hidden layer usually result in serious overfitting problems. Pedersen and Larsen (2009) also verified that eleven neurons in the hidden layer are sufficient with respect to fit performance.

5.1.3 Support Vector Machine

SVM (Boser et al., 1992) is a ML method that is widely used in classification or regression tasks. Its core idea is to use a mapping function ϕ to map the samples into a high-dimensional space, and then find a hyperplane (function) for linearly segmenting the samples in the high-dimensional space. The segmentation is to maximize the margin between the hyperplane and the samples. Since the mapping function is often complicated and difficult to calculate, in practice, a kernel function is usually used to perform the corresponding mapping calculation instead. The kernel function used in this study is the most commonly used gaussian radial basis function, which performs well without prior knowledge about the data. The SVM used for regression analysis is called support vector regression (SVR). In SVR, the margin maximization problem can be transformed into an equivalent convex quadratic programming problem (Equation 2). Since it is impossible to ensure that all samples are linearly separable, slack variables ξ_e, ξ_e^* are introduced for each sample in the constraints to solve the linear inseparability problem. At the same time, a penalty must be made for each slack variable introduced. The penalty parameter C is used in the objective function to adjust the penalty intensity for slack variables. Solving this problem returns the parameters of the hyperplane, including the normal vector w, the intercept b , and the slack variables ξ_e, ξ_e^*. Therefore, the objective function solved by SVR can be expressed by the following formula:

$$\min_{w,\, b,\, \xi_e,\, \xi_e^*} \left\{ \frac{1}{2} w^T w + C \sum_{e=1}^{n} (\xi_e + \xi_e^*) \right\}$$
$$s.\,t. \begin{cases} y_e - w^T \phi(x_e) - b \leq \varepsilon + \xi_e, \\ w^T \phi(x_e) + b - y_e \leq \varepsilon + \xi_e^*, \\ \xi_e,\, \xi_e^* \geq 0,\; e = 1, \cdots, n \end{cases} \tag{2}$$

where n is the number of samples; y_e is the actual target value; x_e is the training variables vector; ε is the maximum deviation between the target function value and the actual target value. Readers are referred to Smola and Schölkopf (2004) for the technical details of SVR.

5.1.4 Ridge regression

Ridge regression (Hoerl and Kennard, 1970) is a biased estimation regression method, which is an improvement of the ordinary least squares (OLS) method. The OLS method is an unbiased estimation method, which fits a linear model with coefficients w_c to minimize the sum of squared residuals between the actual target values in the sample data and the target values (estimated values) predicted by the linear model. The objective function solved by the OLS method can be expressed by the following formula:

$$\min_{w_c} \left\{ \left\| w_c \cdot X - Y \right\|_2^2 \right\} \tag{3}$$

where X is the covariates matrix; Y is the actual target vector.

When the number of variables used to construct a regression model is large and the sample size is relatively small, the OLS method could easily lead to overfitting issues. In addition, the coefficient estimates of the OLS method rely on the mutual independence of the covariates. When there is a multicollinearity problem between the covariates, the covariates matrix will become close to singular, which will make the OLS method highly sensitive to outliers of the observed target and produce a large

IAMU International Association of Maritime Universities (AMC)

variance. To solve this problem, ridge regression adds a certain degree of bias to the regression estimate to obtain a more reliable target prediction value (**Equation 4**). This approach of introducing bias is called regularization. The degree of bias added in the ridge regression is adjusted using the penalty parameter α in **Equation 4**:

$$\begin{aligned} &\min_{w_c} \left\{ \|w_c \cdot X - Y\|_2^2 + \alpha \|w_c\|_2^2 \right\} \\ &s.\ t.\ \|w_c\|_2^2 \le t \end{aligned} \tag{4}$$

where $t \ge 0$ is a pre-specified free parameter used to determine the amount of regularization, and its relationship with α depends on the input data.

The penalty for the regression coefficients increases the bias of the regression model but reduces the variance. The reduction in variance can usually compensate for the increase in bias, thereby improving the overall prediction performance (Lepore et al., 2017). Readers are referred to Hoerl and Kennard (1970) for the technical details of ridge regression.

5.1.5 LASSO

The LASSO is a regression analysis method proposed by Tibshirani (1996). Both the interpretability and prediction accuracy of the model can be strengthened in this approach because regularization and variable selection are performed simultaneously. LASSO is conceptually very similar to Ridge, both of them introduce a penalty for the regression coefficients. Ridge penalizes the sum of squared coefficients (L2 penalty), while LASSO penalizes the sum of the coefficients' absolute values (L1 penalty). The objective function solved by LASSO can be expressed by the following formula:

$$\begin{aligned} &\min_{w_c} \left\{ \|w_c \cdot X - Y\|_2^2 + \alpha \|w_c\|_1 \right\} \\ &s.\ t.\ \|w_c\|_1 \le t \end{aligned} \tag{5}$$

The biggest difference between the L1 penalty and the L2 penalty is that the L1 penalty both regularizes the function and eliminates the features that do not have sufficient impacts on the target (Coraddu et al., 2017). With high values of α, the regression coefficients of variables with low correlation with the output target (or variables with multicollinearity with other variables) will be exactly zeroed, thus achieving variable selection in LASSO. However, the coefficients of these variables in Ridge can only be close to zero.

5.2 Data Normalization

Different ML methods have different requirements for data preprocessing. The main difference is whether to use data normalization. To clarify the impact of data normalization on the performances of ML models, the performances (R^2) of ML models before and after data normalization were compared in a preliminary study. This preliminary study reveals that the performances of SVM and ANN models after data normalization are significantly better than those before normalization, while other models do not see a significant difference. See **Figure 6**. Therefore, our studies use data normalization for SVM and ANN but not for other models.

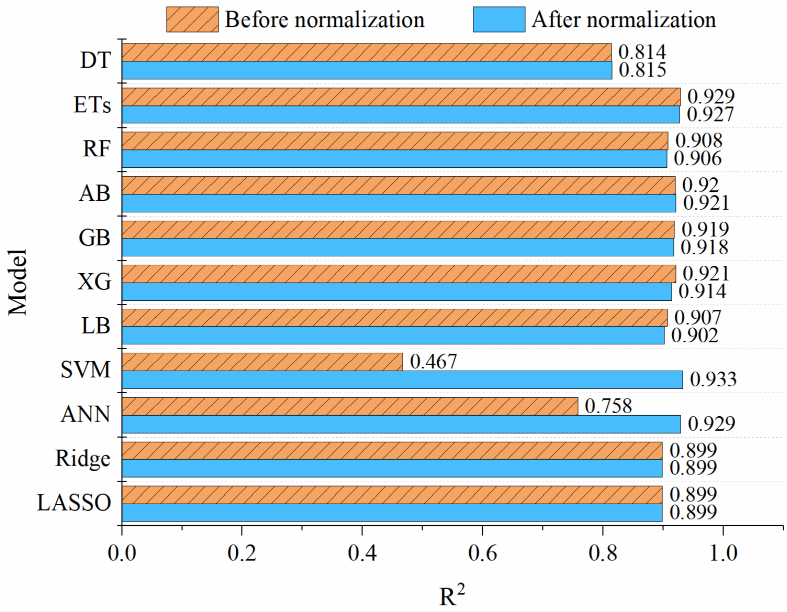


Figure 6. The impact of data normalization on model performance of ship S5 over a dataset adopted in a preliminary study

5.3 Hyperparameter optimization

In ML, some parameters' values need to be set prior to the learning process because they determine the structure of a ML model. These parameters are termed as hyperparameters. To maximize the performance of ML models, in the implementation of the above eleven ML methods, it is necessary to adjust the corresponding hyperparameters according to the training dataset. **Table 2** lists the hyperparameters that need to be optimized for the eleven ML models. When experimenting with optimization approaches for hyperparameter optimization, the Bayesian Optimization (BO) method was identified as the best. In a preliminary study, we further experimented with the BO based on tree-structured Parzen Estimators of hyperopt 0.2.2 library (Hyperopt) (Bergstra et al., 2013), the BO based on extra trees regressor of scikit-optimize 0.7.4 library (Skopt), and the multi-step grid search method of scikit-learn 0.22.1 library (Msgs). Showing superior accuracy and least time consumption, Hyperopt was finally selected as the method to optimize model hyperparameters. See **Figure 7**.

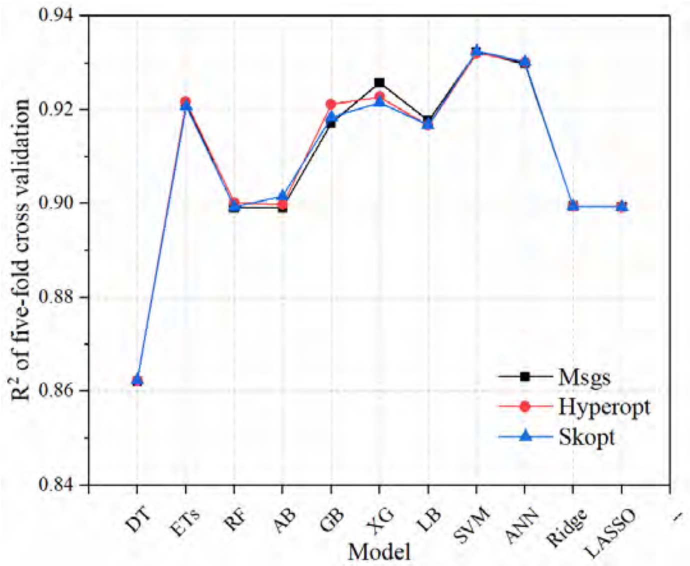
Table 2. Model hyperparameters to be optimized

Model	Hyperparameters	Package/Library	Package reference
DT	*max_depth* [2, 30], *min_samples_leaf* [1, 20], *min_samples_split* [2, 20], *max_features* [1, 15]	scikit-learn	scikit-learn, 2020

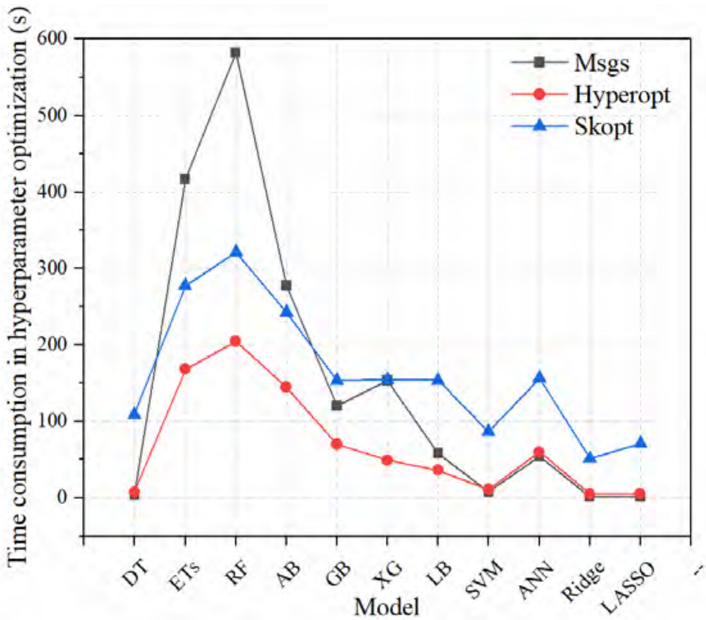
IAMU International Association of Maritime Universities (AMC)

ETs	*max_depth* [2, 30], *min_samples_leaf* [1, 20], *min_samples_split* [2, 20], *max_features* [1, 15], *n_estimators* [10, 300]	scikit-learn	scikit-learn, 2020
RF	*max_depth* [2, 30], *min_samples_leaf* [1, 20], *min_samples_split* [2, 20], *max_features* [1, 15], *n_estimators* [10, 300]	scikit-learn	scikit-learn, 2020
AB	*max_depth* [2, 10], *min_samples_leaf* [1, 20], *min_samples_split* [2, 20], *max_features* [1, 15], *n_estimators* [10, 300], *learning_rate* [0.00001, 1]	scikit-learn	scikit-learn, 2020
GB	*max_depth* [2, 10], *min_samples_leaf* [1, 20], *min_samples_split* [2, 20], *max_features* [1, 15], *n_estimators* [10, 300], *learning_rate* [0.00001, 1], *subsample* [0.4, 1]	scikit-learn	scikit-learn, 2020
XG	*max_depth* [2, 10], *n_estimators* [10, 300], *learning_rate* [0.00001, 1], *min_child_weight* [0, 10], *gamma* [0, 2], *colsample_bytree* [0.1, 1], *subsample* [0.4, 1], *reg_alpha* [0, 2], *reg_lambda* [0, 2]	XGBoost	XGBoost, 2020
LB	*max_depth* [2, 10], *n_estimators* [10, 300], *learning_rate* [0.00001, 1], *min_child_weight* [0, 10], *min_child_samples* [2, 100], *colsample_bytree* [0.1, 1], *subsample* [0.4, 1], *reg_alpha* [0, 2], *reg_lambda* [0, 2], *num_leaves* [5, 127], *min_split_gain* [0, 2]	LightGBM	LightGBM, 2020
SVM	*C* [0.00001, 100], *gamma* [0.00001, 1]	scikit-learn	scikit-learn, 2020
ANN	*Activation* [‘identity’, ‘tanh’, ‘logistic’, ‘relu’], *solver* [‘lbfgs’, ‘sgd’, ‘adam’], *alpha* [0.00001, 2], *learning_rate_init* [0.00001, 1], *beta_1* [0, 0.999], *beta_2* [0, 0.999]	scikit-learn	scikit-learn, 2020
Ridge	*alpha* [0, 10]	scikit-learn	scikit-learn, 2020
LASSO	*alpha* [0, 10]	scikit-learn	scikit-learn, 2020

Note: The brackets after the hyperparameter names list the value ranges of the hyperparameters.



(a) R^2 (accuracy) comparison of hyperparameter optimization methods



(b) Time consumption comparison of hyperparameter optimization methods

Figure 7. Comparison of three hyperparameter optimization methods for ship S8, over a dataset adopted in a preliminary study

5.4 Performance Metrics

Performance metrics that gauge the fit performances of ML models are defined in the following over the training set. The R^2 value over the test set, referred to as *R^2 (test)*, is used to measure the generalization performance of a ML model.

$$R^2 = 1 - \frac{\sum_{t=1}^{k}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{k}(y_t - \bar{y})^2} \tag{6}$$

$$MSE = \frac{1}{k}\sum_{t=1}^{k}(y_t - \hat{y}_t)^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{k}\sum_{t=1}^{k}(y_t - \hat{y}_t)^2} \tag{8}$$

$$MAE = \frac{1}{k}\sum_{t=1}^{k}\left|y_t - \hat{y}_t\right| \tag{9}$$

$$MAPE = \frac{100\%}{k}\sum_{t=1}^{k}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{10}$$

where y_t is the target value - actual ship fuel consumption rate (t/day); $\hat{y}_t$ is the predicted output value - predicted ship fuel consumption (t/day); $\bar{y}$ is the average of target values - average of actual ship fuel consumption rate (t/day); k is the number of samples in the data set; p is the number of input variables of the model.

6. Fusion Solution 1 (DFS1): Voyage Report Data + Meteorological Data

6.1 Rationale of Fusing Voyage Report Data and Meteorological Data

Yan et al. (2021) point out that weather and sea conditions recorded by voyage report are snapshot information by the deck officer. For instance, the wind speed/force and direction in a voyage report data entry are from the deck officer's one read of their anemometer, and the time of the deck officer's reading the anemometer can be random on the given day. Apart from the snapshotting method, our conversation with industry collaborators show that wave and sea current conditions recorded in voyage report depend highly on the deck officer's eye inspection and personal experience. These issues could all erode the data quality of voyage report on weather and sea conditions. To remedy the data quality issue of voyage report on weather and sea conditions, our industry collaborators suggested us approaching some publicly accessible meteorological data sources to retrieve more reliable data of weather and sea conditions. As discussed in Section 4.2, this project utilizes the meteorological data from ECMWF and CMEMS ("Copernicus").

6.2 Approach to Fusing Voyage Report Data and Meteorological Data

The first key step of fusing voyage report data and meteorological data is to estimate the sailing trajectory (hourly geographical positions) of the ship in a day. This estimation can be performed with the famous great circle route. In the actual voyage of a ship, the great circle route is the shortest economic route in terms of distance. However, following the great circle route often requires the deck team to constantly change the course of the ship. Therefore, to facilitate navigation, the great circle route is usually divided into several segments and then the ship sails along the rhumb line (or loxodrome) on each segment (Weintrit and Kopacz, 2011). Based on this, the latitude and longitude of each position the ship passes are calculated according to the rhumb line formulas (Bennett, 1996) shown below:

$$S = V \cdot h \tag{11}$$

$$\Delta\varphi = S \cdot \cos\vec{C} \tag{12}$$

$$\varphi_2 = \varphi_1 + \Delta\varphi \tag{13}$$

$$\varphi_m = \frac{\varphi_1 + \varphi_2}{2} \tag{14}$$

$$\Delta\lambda = S \cdot \sin\vec{C} \cdot \sec\varphi_m \tag{15}$$

$$\lambda_2 = \lambda_1 + \Delta\lambda \tag{16}$$

$$t_2 = t_1 + h \tag{17}$$

In these formulas, S is the sailing distance (n mile); V is the sailing speed (knots); h is the sailing time (hour); $\Delta\varphi$ is the latitude difference (°); $\vec{C}$ is the ship's course (°), which should be converted to the range of 0° - 90°, from north and south (e.g., courses 150° and 300° should be converted to 30° and 60° respectively.); φ_1 and φ_2 are the latitudes of the departure and arrival positions, respectively (°); φ_m is the average latitude between them (°); λ_1 and λ_2 are the longitudes of the departure and arrival positions, respectively (°); $\Delta\lambda$ is the longitude difference (°); t_1 and t_2 are the times of departure and arrival, respectively.

Second, the weather and sea conditions at each hourly position can be retrieved from ECMWF data on 12 variables and CMEMS (Copernicus) data on 2 variables. The wind/waves/sea currents direction obtained from meteorological data is the absolute direction. To obtain the directional information of wind/waves/sea currents relative to the bow of the ship, the "true course" information from the voyage report is used. Due to the symmetric structure of the ship, the relative wind/wave direction is between 0° and 180°. 0° represents the wind/waves/sea currents coming to the bow, and 180° represents the wind/waves/sea currents coming to the stern.

Due to the nature of voyage report data, it usually contains only one data entry per day. For a specific day (corresponding to a specific data entry of voyage report), meteorological data are used for the purpose of correcting the possibly inaccurate information of weather and sea conditions contained in this voyage report data entry. Therefore, it is necessary to average the weather/sea conditions along hourly geographical positions travelled through by the ship, and to use this daily average as the substitute for weather/sea condition information in this data entry corresponding to this specific day. The average method used is as follows:

$$\overline{W} = \frac{1}{M}\sum_{i=1}^{M} W_i \tag{18}$$

where $\overline{W}$ is the daily average weather/sea condition data; $M = 24$ is the number of hourly weather/sea condition data entries per day; W_i is the hourly weather/sea condition data. Note that the averaging method is widely adopted by meteorological services such as ECMWF to conduct data conversions between different granularities of longitude A? latitude A? time.

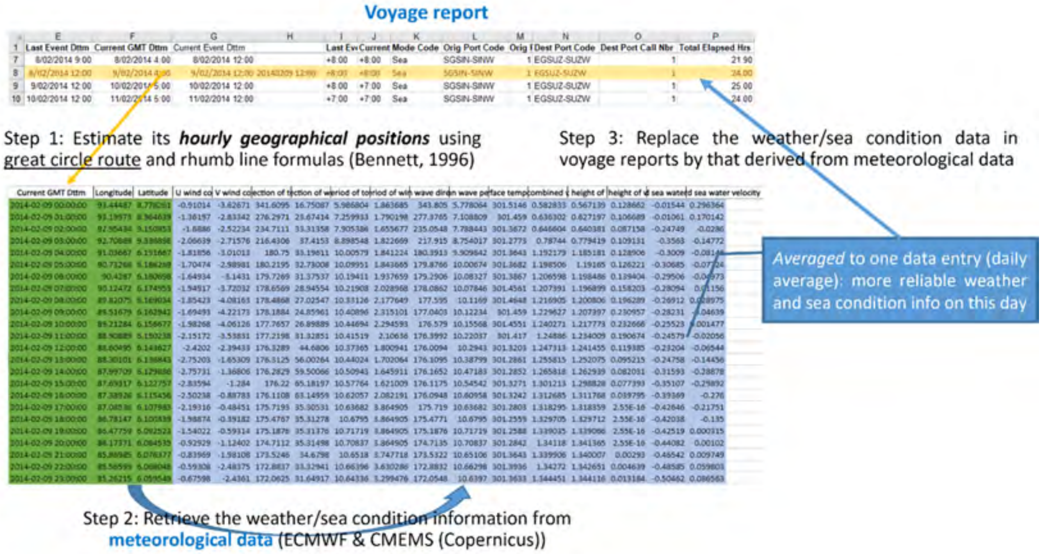


Figure 8. Approach of fusing voyage report data and meteorological data

The whole process of fusing voyage report data and meteorological data is illustrated in **Figure 8**. Until now, all the information derived from meteorological data about weather and sea conditions are in the form the precise values. Specifically, the relative directions of wind/waves/sea currents are represented as the degrees relevant to the ship's bow, and wind speed is in the unit of m/s. However, voyage reports use fuzzy values for these data. For the convenience of comparison experiments between precise values and fuzzy values, **Tables 3 and 4** can convert precise values of weather and sea conditions to fuzzy values.

We generate nine possible datasets using voyage report data and meteorological data, by considering the target application scenarios in energy-efficient operational measures for voyage management, the endogeneity issue discussed by Yan et al. (2021), the fact that waves consist of swells and wind waves, and the experimental choice of using precise or fuzzy values for weather and sea conditions. See **Table 5** for the details of these nine datasets.

Table 3. Conversion of relative wind/wave/sea current direction data from precise values to fuzzy values

Relative wind/wave direction angle (**precise value**)	Approximate wind/wave direction (**fuzzy value**)
$0^o \sim 30^o$	E
$30^o \sim 60^o$	D/F
$60^o \sim 120^o$	C/G
$120^o \sim 150^o$	B/H
$150^o \sim 180^o$	A

Table 4. Wind force scale corresponding to different wind speeds (ISO 15016: 2015(E)).

Wind speed (m/s) – **precise value**	Wind force (Beaufort scale) – **fuzzy value**
0.0 ~ 0.2	0
0.3 ~ 1.5	1
1.6 ~ 3.3	2
3.4 ~ 5.4	3
5.5 ~ 7.9	4
8.0 ~ 10.7	5
10.8 ~ 13.8	6
13.9 ~ 17.1	7
17.2 ~ 20.7	8
20.8 ~ 24.4	9

Table 5. Features contained in each dataset by fusing voyage report data and meteorological data

Original datasets	Data source	Features	Dataset								
			Set1	$Set2_{precise}$ [b]	$Set2_{fuzzy}$ [c]	$Set3_{precise}$ [b]	$Set3_{fuzzy}$ [c]	$Set4_{precise}$ [b]	$Set4_{fuzzy}$ [c]	$Set5_{precise}$ [b]	$Set5_{fuzzy}$ [c]
Voyage report data	Shipping company	Fuel consumption rate	√	√	√	√	√	√	√	√	√
		Sailing speed	√	√	√	√	√	√	√	√	√
		Displacement	√	√	√	√	√	√	√	√	√
		Trim	√	√	√	√	√	√	√	√	√
		Wind speed	√								
		Wind direction (Rel.)	√								
		Swell height	√								
		Swell direction (Rel.)	√								
		Sea currents speed	√								
		Sea currents direction (Rel.)	√								
		Sea water temperature	√								
Meteorological data	European Centre for Medium-range Weather Forecasts (ECMWF)	Wind speed		√	√	√	√	√	√	√	√
		Wind direction (Rel.) [a]		√	√	√	√	√	√	√	√
		Swell height		√	√	√	√	√	√		
		Swell direction (Rel.) [a]		√	√	√	√	√	√		
		Swell period									
		Wind wave height				√	√	√	√		
		Wind wave direction (Rel.) [a]				√	√	√	√		
		Wind wave period									
		Combined wave height				√	√			√	√
		Combined wave direction (Rel.) [a]				√	√			√	√
		Combined wave period									
		Sea water temperature		√	√	√	√	√	√	√	√
	Copernicus Marine Service	Sea current speed		√	√	√	√	√	√	√	√
		Sea current direction (Rel.) [a]		√	√	√	√	√	√	√	√

Notes:

a Relative directions of wind/waves/sea currents are calculated based on ship's "true course" information from voyage report data because "heading" information is absent from voyage report data.

b The subscript "precise" means the directions of wind/waves/sea currents are calculated as the angles relative to ship's heading measured by degrees.

c The subscript "fuzzy" means the precise information of directions of wind/waves/sea currents is converted to fuzzy data as per Table 3, and wind speed is represented by Beaufort scale numbers as per Table 4.

6.3 Experimental Results and Discussion

6.3.1 Performance of Eleven ML Models over Nine Datasets and Selection of the Best Datasets

To evaluate the model performance as comprehensively as possible, each of nine datasets of each ship (for instance, *Set1* of ship S1) is randomly divided into two subsets, where the training set contains 80% of the data entries, and the test set contains 20% of the data entries. The training set is used for model hyperparameter optimization and model fitting, and the test set is used to assess the generalization performance of the model. On the training set, *Hyperopt*, a Bayesian optimization method, is used to optimize model hyperparameters, and the optimization objective is to maximize the R^2 value of *five-fold cross-validation*.

To obtain statistical comparison results and ensure the robustness of the comparison results, the random split of each dataset of each ship (for instance, *Set1* of ship S1) into a training set and a test set is conducted 20 times. For instance, *Set1* of ship S1 has 20 different splits of training set and test set. For each random split, the hyperparameters of the ML model under investigation are re-optimized and a model with the best hyperparameter values is trained. Therefore, 20 random splits of a dataset necessitate 20 runs of hyperparameter optimization and model training, resulting in 20 trained models (the same type of ML model with different hyperparameter values). The average performance of the 20 runs (trained models) is taken as the final result for model evaluation to eliminate the impact of disturbance caused by randomness in dataset division/split. The values of performance metrics of eleven ML models over nine datasets for ship S1 are tabulated in **Table 6**. As stated above, the figure in each cell of Table 6 is the average result of 20 runs corresponding to 20 random splits of the dataset. For instance, R^2 of the model DT over the dataset *Set1*, 0.846, is the average of 20 R^2 values corresponding to 20 runs of the DT model over *Set1*. The fourth column of Table 6 (labelled as "R^2 (test)") is the R^2 values on the test set. The results for ships S2 to S8 can be found in **Tables A1 to A7 in Appendices**.

One may have been aware that performance of ML models and quality of datasets are interwoven together and the job of selecting the best datasets from the results of eleven ML models, nine datasets, and eight ships (shown in Table 6 and Tables A1 to A7) is overwhelming, not to mention the possible contrasts of R^2 values over the training set versus the test set. To overcome this, we develop a *voting scheme* to select the best datasets. In this scheme, every ML model is a voter and votes for the best datasets, by considering R^2 (with two decimal places) as the first priority and R^2 (test) (with two decimal places) as the secondary performance metric. For instance, in Table 6 for ship S1, the DT model finds the best R^2 value with two decimal places is 0.85 which is achieved over datasets *Set1*, $Set3_{precise}$, $Set3_{fuzzy}$, and $Set4_{precise}$. Over these four datasets, it finds the best R^2 (test) with two decimal places is 0.64 which is achieved over *Set1*. Therefore, the DT model of ship S1 votes for *Set1* as the best dataset. Similarly, we allow other ML models to vote for their best datasets and apply this voting scheme to all the eight ships. Voting results are shown in **Table 7**. The number of votes received by each of nine datasets under investigation is shown in **Figure 9**.

Figure 9 is the Tally sheet that counts the votes received by each dataset: Figure 9(a) consider all models as voters; Figure 9(b) does not consider DT, SVM, ANN, Ridge, and LASSO as voters because their fit performances are significantly worse than ET, RF, AB, GB, XG and LB and thus will not be preferred by industry applications; Figure 9(c) further removes RF and LB from the voter list because they are "*dominated*" by ET, AB, GB, and XG against both R^2 and R^2 (test). For instance, in Table 7, RF is dominated by ET because neither of R^2 and R^2 (test) of the RF model is better than the ET model.

It can be seen from Figure 9 that $Set3_{precise}$ and *Set1* receive the largest numbers of votes from best models. $Set3_{precise}$ receives 34 votes from all models, 17 votes from ET, RF, AB, GB, XG, and LB, and 13 votes from ET, AB, GB and XG. *Set1* receives 24 votes from all models, 18 votes from ET, RF, AB, GB, XG, and LB, and 10 votes from ET, AB, GB and XG. Therefore, it will be wise to choose *Set3*$_{precise}$ and *Set1* as the best datasets: *Set3*$_{precise}$ is the best; but the quality of the voyage report data *Set1* is also quite high. The advantage of *Set3*$_{precise}$ over *Set1* reveals the benefits of fusing/combining voyage report data and meteorological data.

Table 6. The fit performance of eleven machine learning models for ship S1 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	**0.846**	**0.643**	81.022	8.934	6.851	7.995
	Set2$_{precise}$	0.828	0.640	82.878	9.051	6.940	8.279
	Set2$_{fuzzy}$	0.836	0.642	78.921	8.821	6.792	8.085
	Set3$_{precise}$	**0.847**	0.617	73.848	8.532	6.522	7.697
	Set3$_{fuzzy}$	**0.848**	0.627	73.402	8.479	6.495	7.662
	Set4$_{precise}$	**0.853**	0.613	71.091	8.348	6.369	7.558
	Set4$_{fuzzy}$	0.838	0.628	77.915	8.728	6.692	7.953
	Set5$_{precise}$	0.834	0.628	80.418	8.896	6.781	8.033
	Set5$_{fuzzy}$	0.828	0.640	82.894	9.035	6.922	8.213
ET	*Set1*	0.992	0.781	4.001	1.525	1.090	1.255
	Set2$_{precise}$	0.931	0.762	33.569	5.674	4.330	5.239
	Set2$_{fuzzy}$	0.934	0.757	32.173	5.444	4.137	4.981
	Set3$_{precise}$	0.965	0.762	17.043	3.524	2.699	3.245
	Set3$_{fuzzy}$	0.939	0.766	29.313	5.012	3.862	4.698
	Set4$_{precise}$	0.956	0.764	20.951	3.918	2.993	3.612
	Set4$_{fuzzy}$	0.950	0.759	24.199	4.495	3.471	4.198
	Set5$_{precise}$	0.947	0.769	25.433	4.623	3.520	4.237
	Set5$_{fuzzy}$	0.943	0.764	27.454	4.842	3.693	4.442
RF	*Set1*	0.964	0.761	18.837	4.321	3.194	3.721
	Set2$_{precise}$	0.940	0.754	28.914	5.304	3.978	4.747
	Set2$_{fuzzy}$	0.944	0.764	27.225	5.174	3.867	4.607
	Set3$_{precise}$	0.936	0.756	30.736	5.506	4.112	4.911
	Set3$_{fuzzy}$	0.941	0.765	28.610	5.310	3.965	4.721
	Set4$_{precise}$	0.942	0.758	27.841	5.210	3.875	4.612
	Set4$_{fuzzy}$	0.935	0.763	31.277	5.535	4.138	4.929
	Set5$_{precise}$	0.940	0.760	29.131	5.331	3.971	4.713
	Set5$_{fuzzy}$	0.938	0.765	30.079	5.418	4.035	4.804
AB	*Set1*	0.955	0.758	23.482	4.687	4.036	4.940
	Set2$_{precise}$	0.931	0.753	33.603	5.671	4.810	5.910
	Set2$_{fuzzy}$	0.942	0.756	28.226	5.008	4.124	5.025
	Set3$_{precise}$	0.938	0.752	29.988	5.180	4.370	5.371
	Set3$_{fuzzy}$	0.926	0.752	36.202	5.814	4.843	5.928
	Set4$_{precise}$	0.942	0.749	28.430	5.117	4.333	5.324
	Set4$_{fuzzy}$	0.940	0.753	29.483	5.111	4.246	5.189
	Set5$_{precise}$	0.953	0.759	22.810	4.475	3.728	4.565
	Set5$_{fuzzy}$	0.952	0.763	23.416	4.491	3.657	4.450
GB	*Set1*	0.987	0.764	6.570	2.238	1.633	1.893
	Set2$_{precise}$	0.942	0.725	27.933	4.962	3.778	4.485
	Set2$_{fuzzy}$	0.943	0.750	27.623	5.067	3.856	4.569
	Set3$_{precise}$	0.962	0.743	18.367	3.776	2.825	3.330
	Set3$_{fuzzy}$	0.963	0.753	18.109	3.775	2.839	3.361
	Set4$_{precise}$	0.951	0.730	23.216	4.268	3.205	3.816
	Set4$_{fuzzy}$	0.960	0.743	19.340	4.084	3.115	3.716
	Set5$_{precise}$	0.946	0.741	26.335	4.731	3.567	4.221
	Set5$_{fuzzy}$	0.953	0.761	22.634	4.487	3.474	4.054

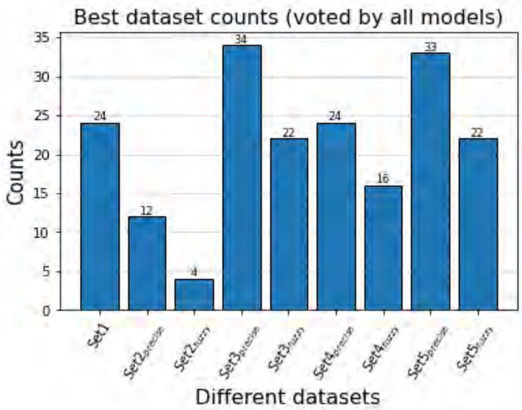
Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.995	0.771	2.805	1.392	1.008	1.168
	$Set2_{precise}$	0.959	0.740	19.763	4.102	3.055	3.544
	$Set2_{fuzzy}$	0.958	0.742	20.247	3.983	3.016	3.505
	$Set3_{precise}$	0.953	0.734	22.403	4.236	3.177	3.695
	$Set3_{fuzzy}$	0.947	0.747	25.851	4.557	3.419	3.985
	$Set4_{precise}$	0.956	0.734	21.189	3.938	2.921	3.407
	$Set4_{fuzzy}$	0.947	0.742	25.472	4.665	3.576	4.195
	$Set5_{precise}$	0.950	0.743	24.275	4.412	3.404	3.978
	$Set5_{fuzzy}$	0.944	0.756	26.767	4.732	3.619	4.196
LB	*Set1*	0.989	0.755	5.857	2.183	1.652	1.924
	$Set2_{precise}$	0.942	0.722	28.076	4.938	3.764	4.463
	$Set2_{fuzzy}$	0.927	0.732	34.990	5.685	4.376	5.161
	$Set3_{precise}$	0.943	0.723	27.467	4.806	3.609	4.272
	$Set3_{fuzzy}$	0.945	0.728	26.553	4.756	3.628	4.271
	$Set4_{precise}$	0.937	0.723	30.701	5.313	4.035	4.803
	$Set4_{fuzzy}$	0.940	0.720	28.687	5.168	3.921	4.654
	$Set5_{precise}$	0.937	0.738	30.654	5.365	3.983	4.713
	$Set5_{fuzzy}$	0.931	0.741	33.492	5.687	4.279	5.080
SVM	*Set1*	0.861	0.784	73.082	8.540	6.365	7.156
	$Set2_{precise}$	0.861	0.792	66.834	8.149	6.039	6.934
	$Set2_{fuzzy}$	0.858	0.779	68.637	8.253	6.125	7.051
	$Set3_{precise}$	0.858	0.786	68.382	8.263	6.143	7.059
	$Set3_{fuzzy}$	0.854	0.782	70.155	8.367	6.227	7.172
	$Set4_{precise}$	0.859	0.789	68.027	8.237	6.119	7.043
	$Set4_{fuzzy}$	0.854	0.779	70.517	8.384	6.239	7.201
	$Set5_{precise}$	0.859	0.795	68.012	8.240	6.139	7.042
	$Set5_{fuzzy}$	0.857	0.791	68.653	8.279	6.145	7.050
ANN	*Set1*	0.869	0.781	68.911	8.290	6.391	7.296
	$Set2_{precise}$	0.829	0.744	83.006	8.767	6.810	8.060
	$Set2_{fuzzy}$	0.855	0.772	70.144	8.315	6.429	7.503
	$Set3_{precise}$	0.854	0.778	70.184	8.366	6.437	7.518
	$Set3_{fuzzy}$	0.846	0.780	73.937	8.593	6.608	7.710
	$Set4_{precise}$	0.832	0.751	80.837	8.611	6.638	7.968
	$Set4_{fuzzy}$	0.868	0.768	64.080	7.921	6.128	7.159
	$Set5_{precise}$	0.857	0.777	68.894	8.256	6.373	7.439
	$Set5_{fuzzy}$	0.826	0.755	83.604	8.970	6.959	8.300
Ridge	*Set1*	0.814	0.774	97.422	9.868	7.725	8.932
	$Set2_{precise}$	0.825	0.782	84.128	9.170	7.087	8.291
	$Set2_{fuzzy}$	0.824	0.778	84.830	9.208	7.104	8.321
	$Set3_{precise}$	0.830	0.784	81.939	9.050	6.993	8.192
	$Set3_{fuzzy}$	0.829	0.784	82.300	9.070	6.989	8.183
	$Set4_{precise}$	0.827	0.782	83.165	9.117	7.029	8.213
	$Set4_{fuzzy}$	0.826	0.779	83.945	9.160	7.068	8.272
	$Set5_{precise}$	0.827	0.785	83.282	9.124	7.022	8.215
	$Set5_{fuzzy}$	0.827	0.784	83.482	9.135	7.014	8.202
LASSO	*Set1*	0.814	0.773	97.552	9.875	7.711	8.917
	$Set2_{precise}$	0.825	0.781	84.185	9.173	7.100	8.309
	$Set2_{fuzzy}$	0.823	0.777	85.087	9.222	7.120	8.339
	$Set3_{precise}$	0.829	0.786	82.204	9.064	6.997	8.191
	$Set3_{fuzzy}$	0.828	0.785	82.832	9.099	7.002	8.184
	$Set4_{precise}$	0.827	0.785	83.345	9.127	7.044	8.235

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	Set4$_{fuzzy}$	0.825	0.779	84.165	9.172	7.079	8.289
	Set5$_{precise}$	0.827	0.784	83.329	9.126	7.040	8.242
	Set5$_{fuzzy}$	0.826	0.783	83.567	9.139	7.031	8.226

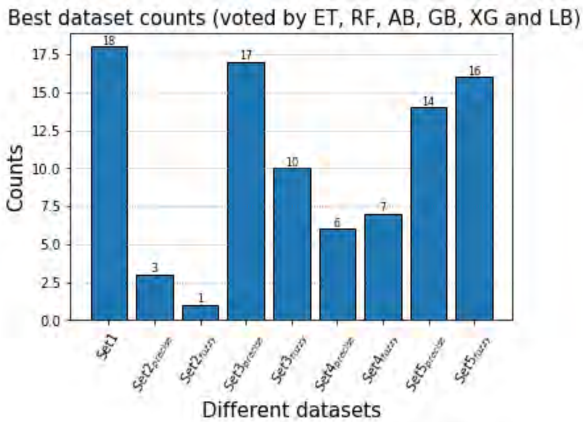
Table 7. DFS1. Best performance of each machine learning model from nine datasets and the datasets that achieve the best performance. R^2 (with two decimal places) is considered as the first priority, and R^2 (test) (with two decimal places) is the secondary performance metric.

Ship	Model	Best R^2	Best R^2 (test)	Datasets
S1	DT	0.85	0.64	*Set1*
	ET	0.99	0.78	*Set1*
	RF	0.96	0.76	*Set1*
	AB	0.96	0.76	*Set1*
	GB	0.99	0.76	*Set1*
	XG	1.00	0.77	*Set1*
	LB	0.99	0.76	*Set1*
	SVM	0.86	0.80	*Set5*$_{precise}$
	ANN	0.87	0.78	*Set1*
	Ridge	0.83	0.79	*Set5*$_{precise}$
	LASSO	0.83	0.79	*Set3*$_{precise}$, *Set3*$_{fuzzy}$, *Set4*$_{precise}$,
S2	DT	0.87	0.61	*Set2*$_{fuzzy}$
	ET	0.98	0.76	*Set4*$_{precise}$,
	RF	0.96	0.77	*Set1*
	AB	0.98	0.74	*Set4*$_{precise}$,
	GB	0.99	0.76	*Set3*$_{precise}$, *Set3*$_{fuzzy}$, *Set4*$_{precise}$, *Set4*$_{fuzzy}$
	XG	0.99	0.77	*Set3*$_{precise}$
	LB	0.98	0.75	*Set3*$_{precise}$
	SVM	0.87	0.81	*Set2*$_{precise}$, *Set4*$_{precise}$, *Set4*$_{fuzzy}$
	ANN	0.91	0.80	*Set2*$_{precise}$, *Set4*$_{precise}$, *Set5*$_{precise}$
	Ridge	0.83	0.80	*Set3*$_{precise}$, *Set4*$_{precise}$
	LASSO	0.82	0.80	*Set2*$_{precise}$, *Set3*$_{precise}$, *Set3*$_{fuzzy}$, *Set4*$_{precise}$, *Set4*$_{fuzzy}$, *Set5*$_{precise}$, *Set5*$_{fuzzy}$
S3	DT	0.87	0.7	*Set5*$_{precise}$
	ET	0.99	0.82	*Set3*$_{precise}$, *Set3*$_{fuzzy}$, *Set5*$_{fuzzy}$
	RF	0.96	0.81	*Set2*$_{precise}$, *Set5*$_{precise}$, *Set5*$_{fuzzy}$,
	AB	1.00	0.81	*Set4*$_{precise}$
	GB	0.98	0.82	*Set5*$_{precise}$
	XG	0.96	0.81	*Set3*$_{precise}$, *Set3*$_{fuzzy}$
	LB	0.96	0.81	*Set5*$_{precise}$
	SVM	0.85	0.82	*Set3*$_{fuzzy}$
	ANN	0.87	0.81	*Set2*$_{precise}$, *Set5*$_{precise}$
	Ridge	0.80	0.80	*Set3*$_{precise}$, *Set3*$_{fuzzy}$, *Set4*$_{precise}$, *Set4*$_{fuzzy}$, *Set5*$_{precise}$
	LASSO	0.80	0.80	*Set3*$_{precise}$, *Set3*$_{fuzzy}$, *Set4*$_{precise}$, *Set4*$_{fuzzy}$, *Set5*$_{precise}$
S4	DT	0.93	0.77	*Set4*$_{fuzzy}$
	ET	1.00	0.88	*Set5*$_{precise}$
	RF	0.98	0.86	*Set2*$_{precise}$, *Set4*$_{fuzzy}$, *Set5*$_{precise}$, *Set5*$_{fuzzy}$
	AB	0.99	0.87	*Set3*$_{precise}$, *Set3*$_{fuzzy}$
	GB	0.99	0.87	*Set3*$_{precise}$, *Set3*$_{fuzzy}$, *Set4*$_{precise}$, *Set4*$_{fuzzy}$, *Set5*$_{precise}$, *Set5*$_{fuzzy}$
	XG	1.00	0.87	*Set3*$_{precise}$

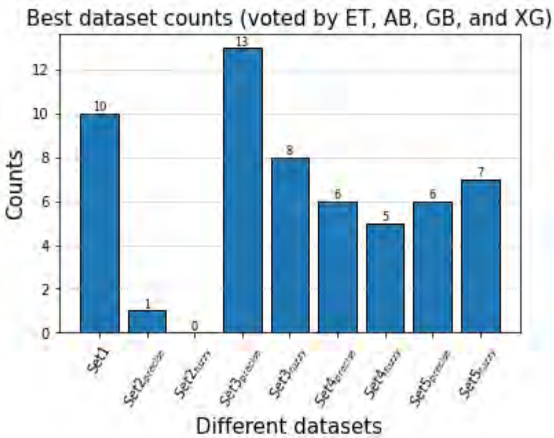
Ship	Model	Best R^2	Best R^2 (test)	Datasets
	LB	0.99	0.87	$Set5_{precise}$, $Set5_{fuzzy}$
	SVM	0.92	0.86	$Set3_{precise}$, $Set4_{precise}$, $Set5_{precise}$, $Set5_{fuzzy}$
	ANN	0.95	0.86	$Set3_{precise}$, $Set3_{fuzzy}$
	Ridge	0.83	0.82	*Set1*
	LASSO	0.83	0.81	$Set3_{precise}$, $Set3_{fuzzy}$, $Set4_{precise}$, $Set4_{fuzzy}$, $Set5_{precise}$, $Set5_{fuzzy}$
S5	DT	0.95	0.8	$Set3_{fuzzy}$
	ET	1.00	0.90	*Set1*
	RF	0.98	0.88	*Set1*, $Set2_{fuzzy}$, $Set3_{fuzzy}$, $Set4_{fuzzy}$, $Set5_{precise}$, $Set5_{fuzzy}$
	AB	1.00	0.89	$Set3_{precise}$, $Set4_{fuzzy}$, $Set5_{precise}$, $Set5_{fuzzy}$
	GB	1.00	0.89	$Set2_{precise}$, $Set4_{precise}$, $Set4_{fuzzy}$, $Set5_{precise}$
	XG	0.99	0.89	*Set1*, $Set3_{fuzzy}$, $Set5_{fuzzy}$
	LB	0.99	0.88	*Set1*, $Set3_{fuzzy}$, $Set5_{fuzzy}$
	SVM	0.93	0.88	*Set1*
	ANN	0.94	0.88	$Set2_{precise}$, $Set3_{precise}$, $Set4_{precise}$
	Ridge	0.89	0.88	$Set5_{fuzzy}$
	LASSO	0.89	0.87	$Set3_{precise}$, $Set3_{fuzzy}$, $Set4_{precise}$, $Set4_{fuzzy}$, $Set5_{precise}$
S6	DT	0.85	0.53	$Set4_{precise}$
	ET	0.99	0.77	*Set1*
	RF	0.96	0.77	*Set1*
	AB	0.98	0.76	$Set3_{precise}$
	GB	0.97	0.79	*Set1*
	XG	0.97	0.79	*Set1*
	LB	0.96	0.75	$Set3_{precise}$, $Set5_{precise}$, $Set5_{fuzzy}$
	SVM	0.85	0.77	$Set2_{precise}$
	ANN	0.88	0.76	$Set5_{precise}$,
	Ridge	0.78	0.75	$Set3_{precise}$
	LASSO	0.77	0.75	$Set3_{fuzzy}$
S7	DT	0.88	0.69	$Set5_{precise}$,
	ET	0.99	0.81	$Set3_{precise}$
	RF	0.97	0.80	$Set5_{precise}$, $Set5_{fuzzy}$
	AB	0.99	0.78	$Set4_{fuzzy}$, $Set5_{fuzzy}$
	GB	0.99	0.79	$Set3_{precise}$
	XG	0.99	0.78	$Set3_{precise}$
	LB	0.98	0.79	$Set3_{precise}$, $Set5_{fuzzy}$
	SVM	0.91	0.79	*Set1*
	ANN	0.90	0.77	$Set2_{precise}$, $Set4_{precise}$,
	Ridge	0.82	0.76	$Set2_{precise}$, $Set2_{fuzzy}$, $Set3_{precise}$, $Set3_{fuzzy}$, $Set4_{precise}$, $Set4_{fuzzy}$, $Set5_{precise}$, $Set5_{fuzzy}$
	LASSO	0.82	0.76	$Set2_{precise}$, $Set2_{fuzzy}$, $Set3_{precise}$, $Set3_{fuzzy}$, $Set4_{precise}$, $Set4_{fuzzy}$, $Set5_{precise}$, $Set5_{fuzzy}$
S8	DT	0.92	0.77	*Set1*, $Set3_{precise}$
	ET	1.00	0.88	*Set1*, $Set3_{precise}$, $Set5_{precise}$, $Set5_{fuzzy}$
	RF	0.98	0.86	*Set1*, $Set3_{precise}$, $Set5_{precise}$, $Set5_{fuzzy}$
	AB	1.00	0.87	$Set5_{fuzzy}$
	GB	1.00	0.86	$Set3_{fuzzy}$
	XG	1.00	0.85	$Set3_{fuzzy}$
	LB	0.98	0.87	*Set1*
	SVM	0.91	0.87	$Set3_{precise}$, $Set4_{precise}$, $Set5_{precise}$
	ANN	0.92	0.86	$Set3_{precise}$, $Set4_{precise}$, $Set5_{precise}$
	Ridge	0.88	0.86	$Set5_{precise}$
	LASSO	0.88	0.85	$Set3_{precise}$, $Set4_{precise}$, $Set5_{precise}$



(a) Best dataset counts (voted by all models)



(b) Best dataset counts (voted by ET, RF, AB, GB, XG and LB)



(c) Best dataset counts (voted by ET, AB, GB, and XG)

Figure 9. Best datasets voted by machine learning models

6.3.2 Performance Comparison of ML models

One may have found the performance differences of 11 ML models from Table 7. To further articulate the performance of these ML models over all the performance metrics, **Table 8** is presented for the ML models over the best dataset $Set3_{precise}$.

Tables 7 and 8 both confirm that ET, RF, AB, GB, XG and LB are good candidate models that can be adopted by the shipping industry: their R^2 values over the best datasets are all above 0.96 and even reach the level of 0.99 to 1.00, while their R^2 performance over test data is in the range from 0.74 to 0.90. The remaining models, including DT, SVM, ANN, Ridge, and LASSO, are not recommended for industry applications because their R^2 values are usually below 0.90, while the values of performance metric R^2 over test data are not better or even worse than ET, RF, AB, GB, XG and LB.

Further, the fit performance of RF and LB are usually slightly dominated by ET, AB, GB, and XG, against both R^2 and R^2 (test), which makes it safe for industry specialists to only install ET, AB, GB and XG into their machine learning model arsenal for ship energy efficiency modeling. Their fit errors on daily bunker fuel consumption, measured by RMSE and MAE, are usually between 0.5 to 4.0 ton/day, though fit errors might be over 4.0 ton/day occasionally if datasets are not carefully chosen.

The experimental results reported in Tables 7 and 8 also rank the performances of eleven machine learning models into four different tiers. The performances of the models in the same tier are quite close, while those of the models in different tiers are significantly different.

- Tier 1: ET, AB, GB, XG, and LB;
- Tier 2: RF;
- Tier 3: DT, SVM, ANN; and
- Tier 4: Ridge, LASSO.

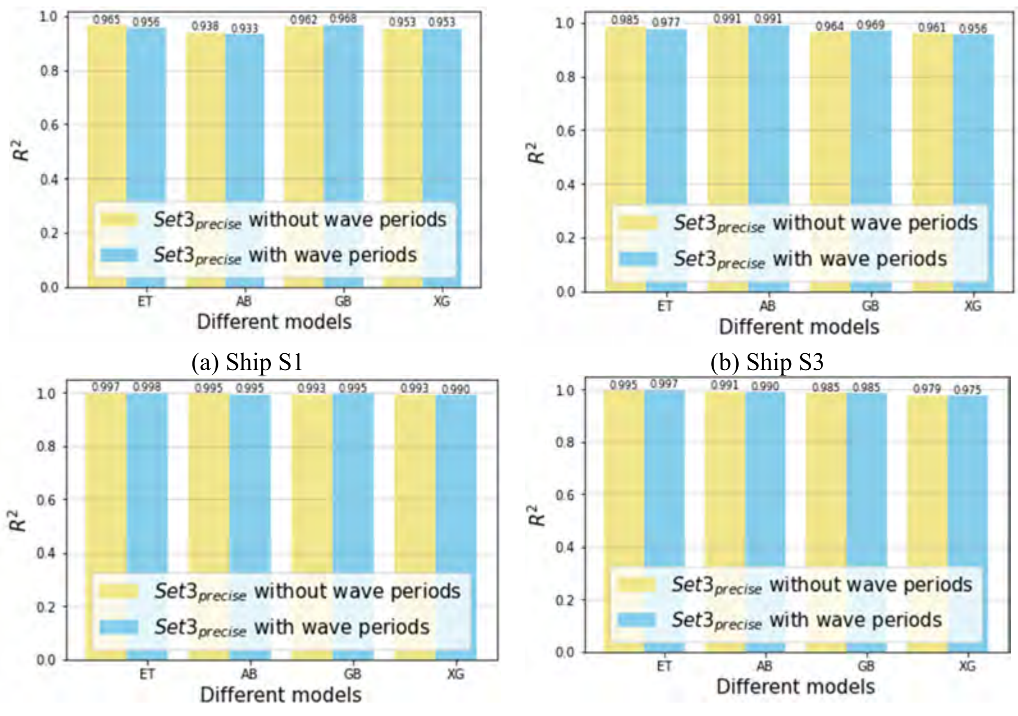
Table 8. The performance of eleven machine learning models over dataset $Set3_{precise}$ (DFS1)

Ship	Model	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
S1	DT	0.847	0.617	73.848	8.532	6.522	7.697
	ET	0.965	0.762	17.043	3.524	2.699	3.245
	RF	0.936	0.756	30.736	5.506	4.112	4.911
	AB	0.938	0.752	29.988	5.180	4.370	5.371
	GB	0.962	0.743	18.367	3.776	2.825	3.330
	XG	0.953	0.734	22.403	4.236	3.177	3.695
	LB	0.943	0.723	27.467	4.806	3.609	4.272
	SVM	0.858	0.786	68.382	8.263	6.143	7.059
	ANN	0.854	0.778	70.184	8.366	6.437	7.518
	Ridge	0.830	0.784	81.939	9.050	6.993	8.192
	LASSO	0.829	0.786	82.204	9.064	6.997	8.191
S2	DT	0.820	0.589	112.089	10.461	7.916	9.230
	ET	0.974	0.765	15.842	3.377	2.445	2.780
	RF	0.950	0.740	31.494	5.541	4.007	4.662
	AB	0.961	0.743	24.755	4.778	4.073	4.729
	GB	0.992	0.760	5.008	1.817	1.234	1.378
	XG	0.991	0.765	5.421	1.949	1.186	1.277
	LB	0.980	0.748	12.589	3.053	2.179	2.442
	SVM	0.864	0.812	84.860	9.176	6.608	7.210
	ANN	0.908	0.791	56.693	7.365	5.581	6.171

Ship	Model	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	Ridge	0.826	0.802	108.847	10.429	8.011	9.055
	LASSO	0.824	0.796	110.162	10.492	8.034	9.042
S3	DT	0.865	0.684	98.572	9.705	7.042	8.343
	ET	0.985	0.821	10.758	2.846	1.716	2.181
	RF	0.956	0.802	31.781	5.576	3.587	4.463
	AB	0.991	0.812	6.328	2.183	1.712	1.998
	GB	0.964	0.819	26.559	4.694	2.836	3.642
	XG	0.961	0.810	28.714	5.030	3.052	3.828
	LB	0.947	0.804	38.795	5.845	3.853	4.853
	SVM	0.844	0.820	113.000	10.591	6.627	8.167
	ANN	0.874	0.798	91.583	9.475	6.480	7.992
	Ridge	0.801	0.796	144.061	11.987	8.329	10.615
	LASSO	0.799	0.796	145.425	12.043	8.323	10.619
S4	DT	0.916	0.746	68.063	8.094	6.036	6.523
	ET	0.998	0.872	1.434	0.901	0.627	0.687
	RF	0.975	0.853	20.349	4.497	3.331	3.618
	AB	0.986	0.865	11.021	3.144	2.591	2.905
	GB	0.989	0.866	8.845	2.500	1.838	1.957
	XG	0.995	0.869	3.758	1.585	1.140	1.201
	LB	0.987	0.855	10.943	2.871	2.200	2.340
	SVM	0.921	0.857	63.718	7.972	5.848	6.146
	ANN	0.947	0.856	42.555	6.513	5.034	5.502
	Ridge	0.833	0.811	135.334	11.629	9.033	9.406
	LASSO	0.832	0.809	135.961	11.656	9.053	9.417
S5	DT	0.947	0.785	29.488	5.182	3.764	5.625
	ET	0.997	0.892	1.413	0.854	0.619	0.935
	RF	0.981	0.874	10.498	3.225	2.390	3.663
	AB	0.995	0.886	2.543	1.525	1.209	2.217
	GB	0.993	0.887	3.519	1.359	1.021	1.610
	XG	0.993	0.878	3.601	1.605	1.133	1.749
	LB	0.987	0.873	7.382	2.350	1.758	2.725
	SVM	0.916	0.873	46.421	6.785	4.917	7.472
	ANN	0.935	0.879	36.157	5.956	4.544	7.075
	Ridge	0.889	0.868	61.610	7.846	5.934	9.109
	LASSO	0.888	0.868	61.988	7.870	5.953	9.129
S6	DT	0.832	0.576	69.684	8.275	6.119	8.113
	ET	0.979	0.752	8.706	2.743	2.010	2.678
	RF	0.953	0.740	19.498	4.382	3.173	4.211
	AB	0.980	0.755	8.175	2.647	2.186	3.210
	GB	0.971	0.770	11.917	3.111	2.384	3.226
	XG	0.959	0.771	17.299	3.835	2.890	3.902
	LB	0.963	0.754	15.520	3.514	2.682	3.646
	SVM	0.843	0.767	65.144	8.045	5.755	7.629
	ANN	0.859	0.772	58.184	7.599	5.750	7.603
	Ridge	0.775	0.745	93.218	9.652	7.454	9.977
	LASSO	0.774	0.744	93.502	9.667	7.443	9.960
S7	DT	0.880	0.683	48.319	6.903	5.173	6.749
	ET	0.987	0.805	5.176	1.848	1.259	1.639
	RF	0.961	0.794	15.501	3.920	2.867	3.740

Ship	Model	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	AB	0.982	0.777	7.272	2.415	1.888	2.558
	GB	0.986	0.785	5.466	2.156	1.442	1.880
	XG	0.986	0.784	5.731	2.093	1.424	1.808
	LB	0.982	0.785	7.152	2.366	1.742	2.283
	SVM	0.871	0.748	51.533	7.113	5.173	6.591
	ANN	0.892	0.771	43.321	6.515	5.071	6.587
	Ridge	0.820	0.758	72.381	8.498	6.520	8.315
	LASSO	0.819	0.758	72.827	8.524	6.550	8.374
S8	DT	0.916	0.769	50.649	6.985	4.922	5.949
	ET	0.995	0.876	2.783	1.404	0.907	1.120
	RF	0.976	0.855	14.566	3.798	2.624	3.187
	AB	0.991	0.863	5.365	2.114	1.693	2.148
	GB	0.985	0.860	9.102	2.427	1.670	2.075
	XG	0.979	0.856	12.821	2.974	2.114	2.589
	LB	0.976	0.852	14.749	3.261	2.338	2.882
	SVM	0.910	0.869	54.154	7.349	5.117	6.123
	ANN	0.924	0.862	46.222	6.733	4.964	5.959
	Ridge	0.879	0.853	72.818	8.529	6.512	7.959
	LASSO	0.878	0.852	73.581	8.573	6.525	7.966

6.3.3 The Impact of Wave Periods



(a) Ship S1 (b) Ship S3

(c) Ship S5 (d) Ship S8

Figure 10. Fit performance of four best models (ET, AB, GB, XG) over dataset *$Set3_{precise}$* with and without wave period information

IAMU International Association of Maritime Universities (AMC)

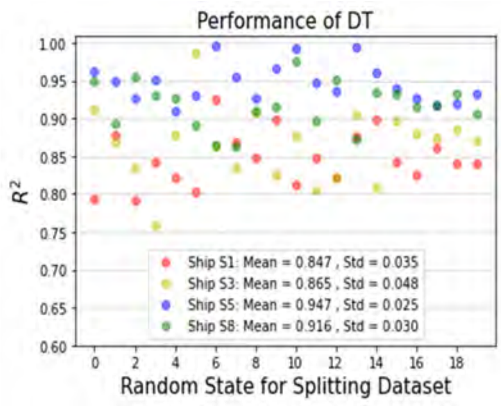
None of the nine datasets under investigation considers the impact of wave periods. To further assess whether the introduction of wave period information could improve the fit performance of ML models, we added the features about periods ("Swell period", "wind wave period", and "Combined wave period" in Table 5) into the best dataset $Set3_{precise}$, and re-experimented with four best models (ET, AB, GB, XG) for ships S1, S3, S5 and S8. Their fit performances over $Set3_{precise}$ with and without wave period information are shown in **Figure 10**.

Figure 10 indicates that including wave period information into models will not improve and even slightly reduces the fit performances of models. This might be explained by the fact that the impact of wave period on a mega containership's fuel efficiency at sea is negligible and adding it to models might introduce additional noise associated with its data. In another word, the impact of wave period on a big containership's fuel efficiency at sea could be covered by the random errors or noises of machine learning models, when voyage report data and meteorological data are used as the data sources.

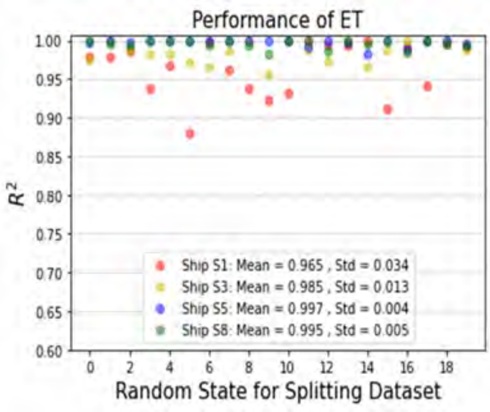
6.3.4 Robustness of ML Models' Performance

Sections 6.3.1 and 6.3.2 report the fit performance of eleven ML models, and the experiment result reported for each ML model over each dataset of each ship is based on the average of 20 runs corresponding to 20 random splits of the dataset into training set and test set. One may further ask a question 'does the fit performances of the models vary too much across the 20 runs?'. To answer this question about the robustness of ML models' performance against random splits of a dataset, we present the R^2 values of eleven ML models over the best dataset $Set3_{precise}$ for ships S1, S3, S5, and S8 in **Figure 11**.

It can be seen from Figure 11 that except DT, LB and ANN, the robustness of the remaining machine learning models is acceptable. RF possesses the highest robustness. The performances of the best models we recommended, including ET, AB, GB and XG, are robust enough for industry applications.

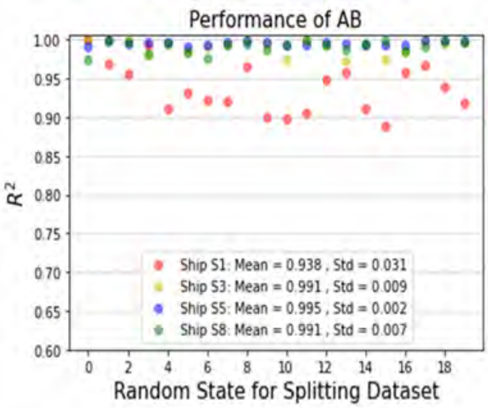


(a) DT

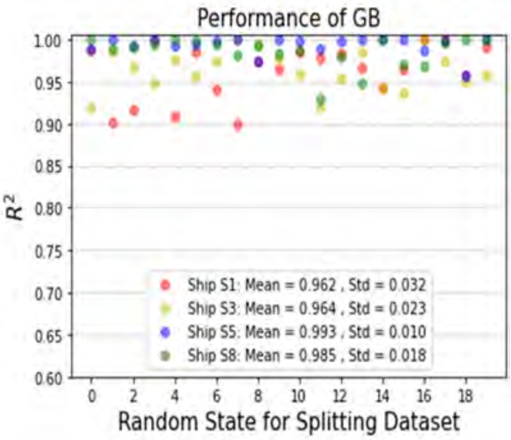


(b) ETs

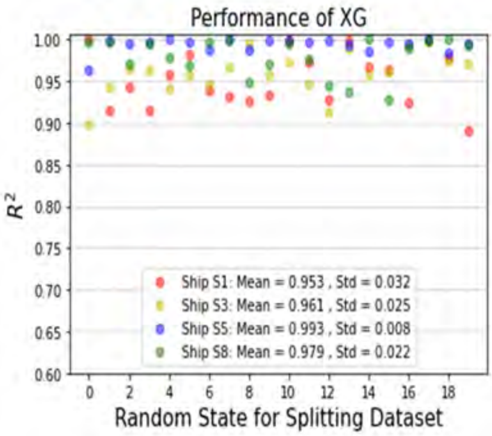
(c) RF



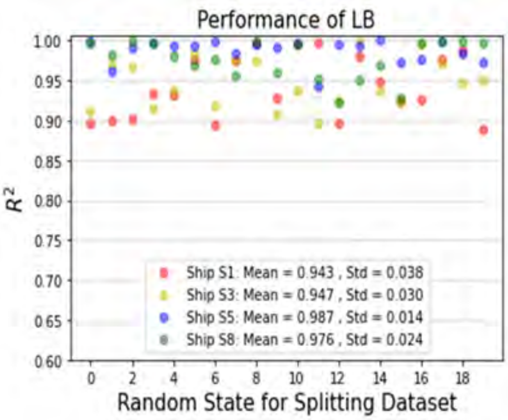
(d) AB



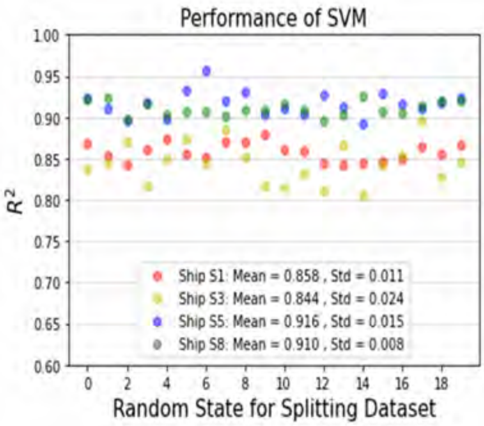
(e) GB



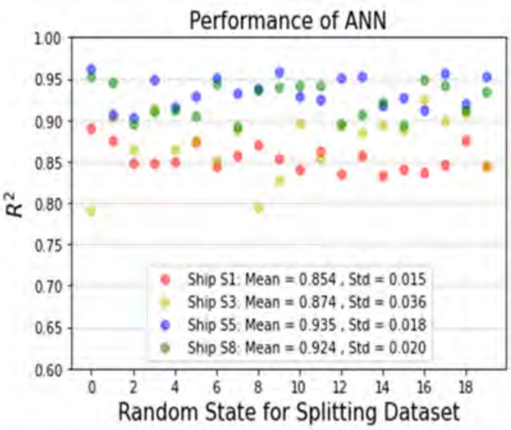
(f) XG



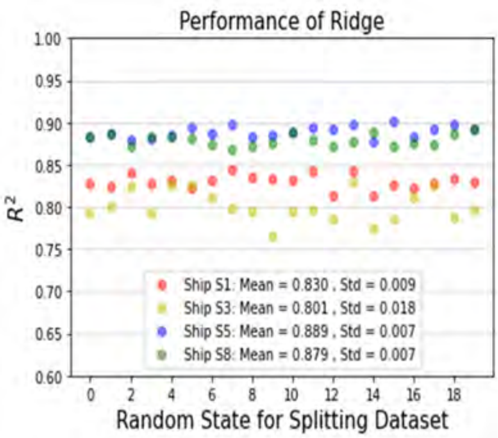
(g) LB



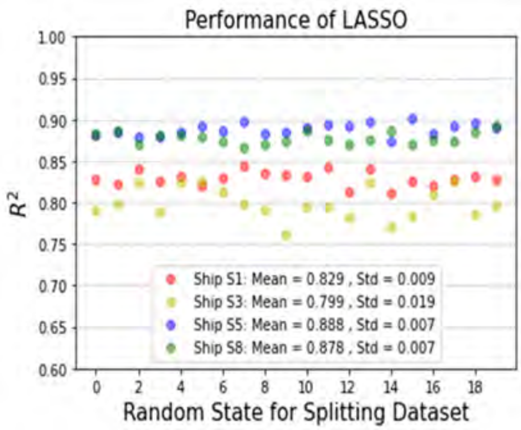
(h) SVM



(i) ANN



(j) Ridge



(k) LASSO

Figure 11. The R^2, mean and standard deviation of the models (DFS1)

6.3.5 Relative Importance of Each Determinant to Ship Fuel Efficiency

Yan et al. (2021) point out that one of the major drawbacks of ML models is poor interpretability. However, one of the exceptions is that tree-based models possess the ability to quantitatively explain the relevant importance of each input variable of the model to the dependant/output variable. The best ML models found by this study, including ET, AB, GB and XG, are all decision tree-based models. Therefore, we conducted the analyses of relevant importance of each feature/determinant to ship fuel consumption rate, based on these four models over the best dataset $Set3_{precise}$ of ships S1, S3, S5 and S8, and collected the results in **Figure 12**.

The first three subfigures (a-c) of Figure 12 reveal that sailing speed is the most important determinant of fuel consumption rate whose importance is between 0.6 and 0.8. This is consistent with the findings in ship propulsion theories.

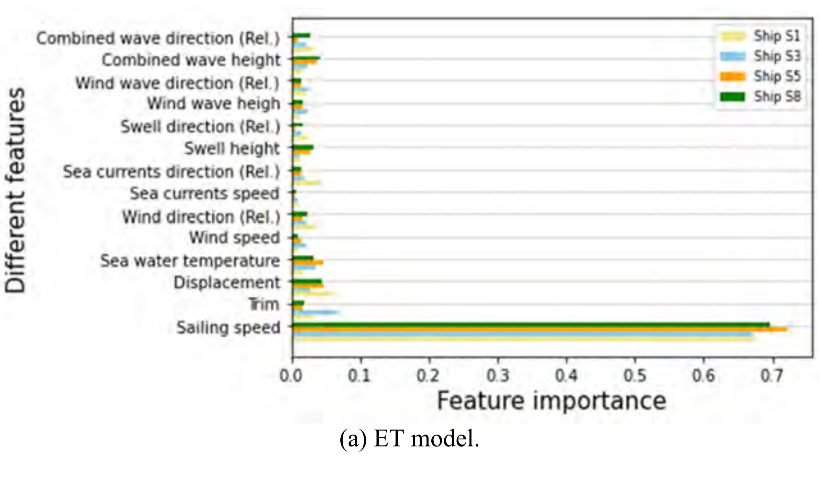
Though displacement/draft is usually considered as the second important determinant in ship propulsion principles, such as the *Admiralty coefficient*, its impact on ship fuel efficiency at sea is basically lower than wave conditions if both swell and wind waves are considered. Apparently, the impact of displacement is significantly lower than the total impact of sea and weather conditions in shipping reality. This finding does not falsify the significant importance of displacement to ship fuel efficiency in calm waters, but ships eventually sail at sea with different weather and sea conditions rather than stay in calm waters.

When sea and weather conditions are considered, waves, consisting of swells and wind waves, play the most significant role. The impact of sea water temperature could be close to that of displacement/draft, which might be beyond the imagination of seafarers at sea. The impact of wind conditions (wind speed and direction) is close to that of sea water temperature and thus also close to the impact of displacement/draft. These results all confirm the importance of weather routing practice in saving bunker fuel and reduce ship emissions.

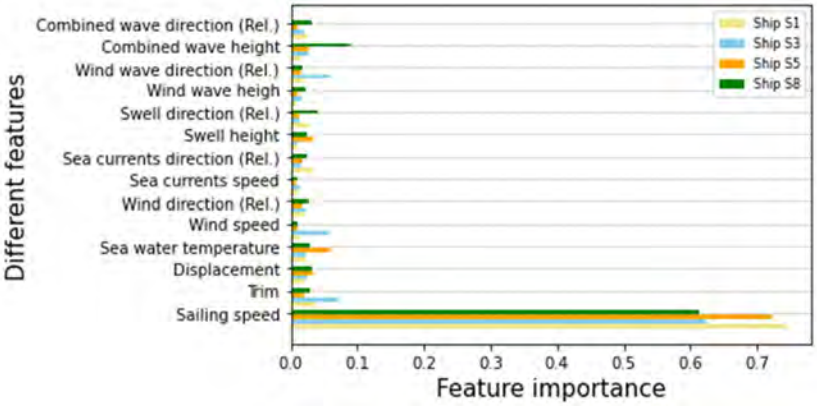
Seafarers at sea attach much importance to sea currents, but their impact on a ship's fuel efficiency in reality could be not comparable to other sea or weather conditions, such as waves, wind, or sea water temperature.

Trim's importance for ship fuel efficiency is usually less than 0.05 but sometimes can reach 0.1, which confirms the rationality of conducting trim optimization for ships. This result is consistent with that reported by the literature on trim optimization.

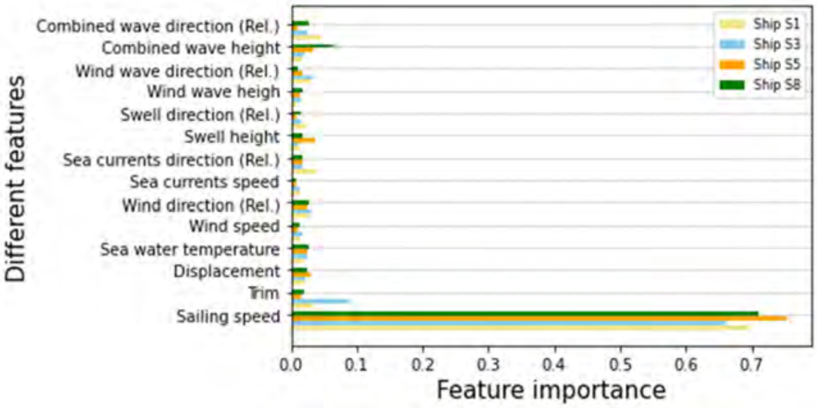
As shown in Figure 12(d), compared to ET, AB, and GB, the XG model reduces the polarization of relative importance allocated to different variables. For instance, in XG's result, the importance of sailing speed decreases and that of weather and sea conditions increases. This could be related to the model structure of XG that introduces a regularization term to avoid overfitting and prevents one variable from attracting too much importance. This characteristic of XG model could have caused the inconsistence of its findings on relative importance of variables/features with other decision tree- based models such as ET, AB, and GB. Therefore, this study leans more on the consistent results of ET, AB and GB during the analysis towards feature importance.



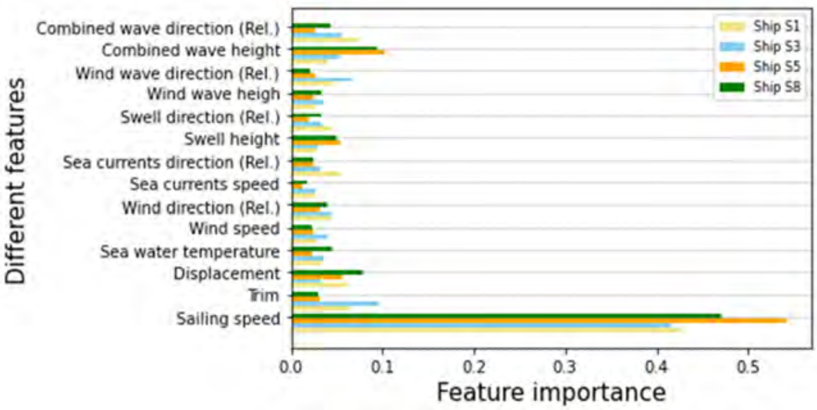
(a) ET model.



(b) AB model



(c) GB model



(d) XG model.

Figure 12. The average relative importance of models input variables (DFS1)

6.4. Summary

Motivated by the data quality issue of voyage reports on weather and sea conditions caused by snapshotting and human eye inspection, this study with DFS1 fuses voyage report data and meteorological data, and constructs nine datasets from this data fusion solution. We experimented with these nine datasets and eleven widely-adopted ML models to quantify the relationship between a ship's bunker fuel consumption rate (MT/day, or MT/h) and its determinants, including sailing speed, displacement/draft, trim, wind, waves (swells and wind waves), sea currents, and sea water temperature, over eight 8100-TEU to 14,000-TEU containerships from a global shipping company.

The best dataset we found, $Set3_{precise}$, reveal the benefits of fusing voyage report data and meteorological data and replacing the information of weather and sea conditions in voyage report by that from meteorological data. However, $Set3_{precise}$ is only sightly better than the original voyage report (*Set1*) which indicates that voyage report has rather acceptable (*hard-to-be-improved*) data quality for many application scenarios, which somewhat disapproves our industry collaborator's conjecture that retrieval of accurate information of weather and sea conditions from meteorological data sources would "*significantly*" improve the data quality for ship fuel efficiency analysis.

Among elven ML models, decision tree-based ensemble models, especially ET, AB, GB and XG, present the best fit and generalization performances. Their R^2 values over the best datasets are all above 0.96 and even reach the level of 0.99 to 1.00, while their R^2 performance over test data is in the range from 0.74 to 0.90. Their fit errors on daily bunker fuel consumption, measured by RMSE and MAE, are usually between 0.5 to 4.0 ton/day. Their performances against random divisions of the dataset into training and test sets are also quite robust. Therefore, it is safe for industry specialists to only install ET, AB, GB and XG into their machine learning model arsenal for ship energy efficiency analysis.

These four tree-based models are recommended also because of their ability to interpret the relative importance of different determinants/factors/features to a ship's fuel consumption rate. Our findings on the relative importance of sailing speed and trim are consistent with existing literature. However, all the tree-based models confirm that the impact of weather and sea conditions is significantly higher than that of the actual displacement/draft of a ship, which indicates the higher practical importance of weather routing studies compared to the studies that seek a sailing route of a ship to optimize its cargo load based on the Admiralty coefficient for the purpose of saving bunker fuel.

This is a pioneering study that combines several data sources to improve the accuracy of ship fuel consumption rate forecast targeting the industry applications in energy-efficient operational measures promoted by IMO, including speed optimization, trim optimization, weather routing, and the virtual arrival policy. The research scope/boundary discussed in Section 3.2 reflects our research limitations.

7. Data fusion solution 2 (DFS2): voyage report data + meteorological data + AIS data

7.1 Rationale of Fusing Voyage Report Data, Meteorological Data, and AIS Data

In our previous study with DFS1, we noticed the data quality issue of voyage report caused by the deck officers' practice of snapshotting and eye inspecting weather and sea conditions. To remedy this issue, we developed a solution DFS1 of fusing voyage report data and publicly accessible meteorological data by replacing the information of snapshotted weather and sea conditions in voyage report with accurate hourly weather and sea conditions retrieved from meteorological data. Over the nine datasets

IAMU
International Association of Maritime Universities
(AMC)

from data fusion for eight 8100-TEU to 14,000-TEU containerships, several ship-specific ML models of forecasting ship fuel consumption rate achieve high fit performances with R^2 values all above 0.96 and even reaching 0.99 to 1.00 for training sets, while their R^2 values for test sets are also promising between 0.74 and 0.90.

In DFS1, a key step before retrieving exact information of weather and sea conditions from meteorological data is calculating the ship's hourly geographical positions (<Timestamp, latitude, longitude>) along its sailing trajectory. DFS1 assumes the ship follows the great circle route approximated by the widely adopted rhumb line and adopts the rhumb line formulas (Bennett, 1996; Weintrit and Kopacz, 2011) to calculate the geographical locations the ship passes in a day. Several ship captains we consulted commented that the great circle route may not be followed in sailing for several reasons and using the geographical positions derived from the great circle route or the rhumb line may introduce inaccuracy when weather and sea conditions are retrieved from meteorological data. This is a prominent limitation of DFS1.

To address this limitation, we approached *MarineTraffic* headquartered in Greece and purchased the AIS data of the eight containerships shown in Table 1, because AIS data provides the detailed geographical positions of the ship forming its actual sailing trajectory. Meanwhile, AIS data also provides the information of the ship's heading at each geographical position, and this may make the calculation of the directions of wind/waves/sea currents relevant to the ship's heading more reliable. The objective of this study is to investigate whether the introduction of actual geographical positions in AIS data will improve the information quality of weather and sea conditions retrieved from meteorological data and therefore further improve the fit performances of ML models when meteorological data and voyage report data are combined.

7.2 Approach to Fusing Voyage Report Data, Meteorological Data, and AIS Data

The information in AIS data about "*Timestamp (UTC)*", "*Longitude Position*", and "*Latitude Position*" could be quite useful in that it helps us find the actual geographical positions of the ship in a day and recover its actual sailing trajectory on that day. Further, accurate detailed information of weather and sea conditions the ship sails through can be retrieved from meteorological data, according to the actual sailing trajectory. "*Ship Heading*" information of AIS data could be also useful because it helps convert the (absolute) directions of wind, waves, and sea currents reported by meteorological data to relative directions of wind, waves, and sea currents against the ship's heading, which is desired in ship fuel efficiency modeling. DFS1 had to utilize "True Course" information in voyage report in the calculation of relative directions of wind, waves, and sea currents as a workaround because voyage report does not record the heading of the ship.

The approach of fusing voyage report data, AIS data, and meteorological data is illustrated in **Figure 13**. First, for a given day recorded by voyage report, the ship's hourly geographical positions are retrieved from AIS data. Second, according to these geographical positions, the hourly weather and sea condition information are queried and obtained from meteorological data including ECMWF (wind, waves, sea water temperature) and Copernicus (sea currents). Then the directions of wind, waves, and sea currents are converted to the relative directions to the ship's heading. Third, these hourly weather and sea conditions are aggregated and produce their daily averages. At last, daily average conditions of wind, waves, sea water temperature, and sea currents are used to replace the meteorological record in the voyage report. In this data fusion approach, the noises of AIS data are not a concern because only hourly geographical positions of the ship are needed for the sake of retrieval of weather and sea conditions. Finer positions of the ship from AIS data are meaningless because our target in this study is the daily average weather and sea conditions the ship sails through for each day recorded by the voyage report. Even if there were noises in sampling the ship's hourly geographical positions from AIS data, they would not cause a problem in calculating the daily average weather and sea conditions confronting the ship.

Similar to DFS1, this study with DFS2 also allows the conversion of the precise values representing wind speed and relative directions of wind, waves, and sea currents to fuzzy values. See **Tables 3 and 4** and **Figure 1**. This is because voyage reports usually adopt fuzzy values and our preliminary experiments show that fuzzy values sometimes overcome data noises/inaccuracy and improve fit performance of ML models. Overall, nine datasets are constructed from this data fusion approach, and features of each dataset are listed in **Table 9**. "*Set1*" in Table 9 is exactly the same "*Set1*" in Table 5, which represents the voyage report.

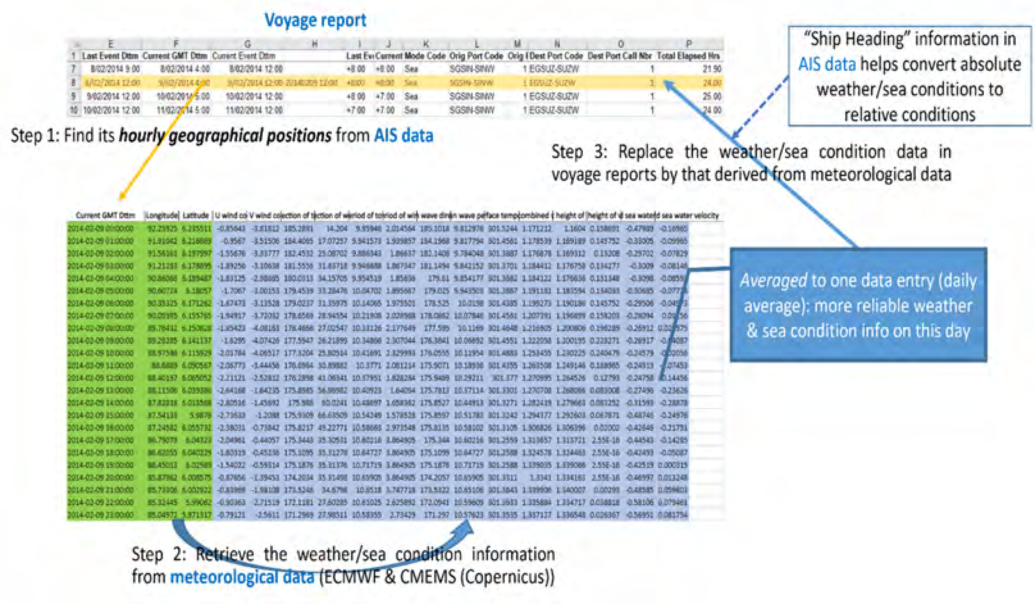


Figure 13. Approach of fusing voyage report data, AIS data, and meteorological data

Table 9. Features contained in each dataset (DFS2)

Original datasets	Data source	Features	Dataset								
			Set1	$AIS2_{precise}$ [b]	$AIS2_{fuzzy}$ [c]	$AIS3_{precise}$ [b]	$AIS3_{fuzzy}$ [c]	$AIS4_{precise}$ [b]	$AIS4_{fuzzy}$ [c]	$AIS5_{precise}$ [b]	$AIS5_{fuzzy}$ [c]
Voyage report data	Shipping company	Fuel consumption rate	√	√	√	√	√	√	√	√	√
		Sailing speed	√	√	√	√	√	√	√	√	√
		Displacement	√	√	√	√	√	√	√	√	√
		Trim	√	√	√	√	√	√	√	√	√
		Wind speed	√								
		Wind direction (Rel.)	√								
		Swell height	√								
		Swell direction (Rel.)	√								
		Sea currents speed	√								
		Sea currents direction (Rel.)	√								
		Sea water temperature	√								
AIS + Meteorological data	AIS+ European Centre for Medium-range Weather Forecasts (ECMWF)	Wind speed		√	√	√	√	√	√	√	√
		Wind direction (Rel.) [a]		√	√	√	√	√	√	√	√
		Swell height		√	√	√	√	√	√		
		Swell direction (Rel.) [a]		√	√	√	√	√	√		
		Swell period									
		Wind wave height				√	√	√	√		
		Wind wave direction (Rel.) [a]				√	√	√	√		
		Wind wave period									
		Combined wave height				√	√			√	√
		Combined wave direction (Rel.) [a]				√	√			√	√
		Combined wave period									
		Sea water temperature		√	√	√	√	√	√	√	√
	AIS+ Copernicus Marine Service	Sea current speed		√	√	√	√	√	√	√	√
		Sea current direction (Rel.) [a]		√	√	√	√	√	√	√	√

Notes:

a Relative directions of wind/waves/sea currents are calculated based on ship's "heading" information from AIS data.

b The subscript "precise" means the directions of wind/waves/sea currents are calculated as the angles relative to ship's heading measured by degrees.

c The subscript "fuzzy" means the precise information of directions of wind/waves/sea currents is converted to fuzzy data as per Table 3, and wind speed is represented by Beaufort scale numbers as per Table 4.

7.3 Experimental Results and Discussion

7.3.1 Performance of Eleven ML Models over Nine Datasets and Selection of the Best Datasets

Same to Li et al. (2022), for each dataset in Table 9, we randomly divided it to a training set (80% of data entries) and test set (20% data entries), which result in a *split* of the dataset. For each split of the dataset, we experimented with a given ML model involving a process of *five-fold cross-validation* based hyperparameter optimization with the Bayesian Optimization method using the tree-structured Parzen Estimators of hyperopt 0.2.2 library (*Hyperopt*) (Bergstra et al., 2013), which is called a *run*. For each ML model over each dataset, we have 20 random splits of the dataset and thus 20 runs of experiments. Each performance metric (R^2, *MSE*, *RMSE*, *MAE* and *MAPE* for training set, R^2 (*test)* for test set, see definition in Section 6) takes the average of 20 runs to overcome the influence of random splitting of the dataset. Experimental results of ship S1 are reported in **Table 10**, while the results of ships S2 to S8 can be found in **Tables A8 to A14 in Appendices**. Note that the performances of the best datasets with DFS1, including *Set1* and $Set3_{precise}$, are also reported in Tables 10 and A8 to A14, for the convenience of comparison with DFS1 in Section 6.

When quality of datasets and performance of ML model are interwoven, shown in Tables 10 and A8 to A14, a voting scheme same to Section 6 is adopted. Each ML model acts as a voter and votes for best datasets (candidates) by considering R^2 (with two decimal places) as the first priority and R^2 (test) (with two decimal places) as the secondary performance metric. The voting result is collated in **Table 11** in which the last column is the votes of the corresponding ML models (voters). **Figure 14** is the Tally sheet that counts the votes received by each dataset: Figure 14(a) consider all models as voters; Figure 14(b) does not consider DT, SVM, ANN, Ridge, and LASSO as voters, because their fit performances are significantly worse than ET, RF, AB, GB, XG and LB and thus they will not be preferred by industry applications; Figure 14(c) further removes RF, GB and LB from the voter list because they are dominated by ET, AB, and XG against both R^2 and R^2 (test).

It can be seen from Figure 14 that $AIS5_{precise}$ receives the largest number of votes, followed by $Set3_{precise}$ and *Set1*. $AIS5_{precise}$ receives 18 votes from ET, RF, AB, GB, XG and LB, according to Figure 14(b), and 9 votes from ET, AB, and XG according to Figure 14(c). This reveals that when AIS data is available for ship fuel efficiency analysis, $AIS5_{precise}$ is the best, and this dataset is better than *Set1* and $Set3_{precise}$ from DFS1 in Section 6. This demonstrates the benefits of further fusing AIS data to voyage report data and meteorological data considered in DFS1. Therefore, we recommend using $AIS5_{precise}$ in practice by fusing voyage report data, AIS data and meteorological data. When AIS data are not available, we can combine voyage report data and meteorological data and utilize $Set3_{precise}$, or even adopt voyage report data *Set1* directly.

Looking at the results here, one may ask why other datasets in Table 9 combine voyage report data, meteorological data, and AIS data but are not competitive with the original voyage report data *Set1*, and even the best dataset $AIS5_{precise}$ cannot always win the original voyage report dataset *Set1*. Similarly, regarding the results reported in Section 6, one may ask why many datasets in Table 5 combine voyage report data and meteorological data but are not competitive with the original voyage report data *Set1*, and even the best dataset $Set3_{precise}$ with DFS1 cannot always win the original voyage report dataset *Set1*. Our deep investigation into the data reveals the following possible reasons. First, as reported by ECMWF and Copernicus Marine Service in their websites, their meteorological data cannot avoid inaccuracy and errors, because these data rely on many types of collection equipment and the calculation of many models. As the evidence, we will see in Section 8 that the wind conditions contained in ECMWF are quite different from the actual wind conditions captured by the sensors on board the ships.

Second, the power of "average" calculation plays a critical role in reducing the quality of data used for model training. Specifically, the weather condition for a given day (corresponding to a voyage report data entry) is estimated by taking the average of the weather conditions at 24 waypoints (hourly waypoints) during the day. However, even accurate weather conditions at these waypoints cannot guarantee their daily average is closer to the actual weather condition (the reality). To provide an analogy for the sake of understanding, consider a situation in which we are estimating the actual average/mean value of a random variable through several observations. Assume the actual average value of this random variable is 10. Consider two different samples of observations: *Sample 1* = {9.5, 9.5, 9.5, 11, 12, 13} and *Sample 2* = {9, 7, 5, 11, 13, 15}. The deviation of data in *Sample 1* from the real average ({0.5, 0.5, 0.5, 1, 2, 3} is much smaller than that of *Sample 2* ({1, 3, 5, 1, 3, 5}). However, the average value estimated through *Sample 1* is 10.75, which is worse than that estimated from *Sample 2* (i.e., 10, the same as the actual average).

Third, the quality of voyage report data might already be good enough. Specifically, when it turns to the snapshot weather and sea condition data, a ship captain we consulted pointed out "***though the snapshot weather and sea condition data is not desired, if you snapshotted 8-meter waves/swells, it is almost impossible that your ship sailed through good weather and sea conditions on average on that day***". This comment indicates that the snapshot weather and sea condition data might be representative, though to unknow degrees, for the actual weather conditions the ship sails through in a day.

Table 10. The fit performance of eleven machine learning models for ship S1 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.846	0.643	81.022	8.934	6.851	7.995
	AIS2$_{precise}$	0.840	0.630	77.176	8.694	6.705	7.959
	AIS2$_{fuzzy}$	0.822	0.623	85.984	9.211	7.093	8.448
	AIS3$_{precise}$	0.827	0.624	83.223	9.057	6.908	8.190
	AIS3$_{fuzzy}$	0.837	0.630	78.690	8.719	6.714	7.970
	AIS4$_{precise}$	0.841	0.641	76.360	8.633	6.604	7.779
	AIS4$_{fuzzy}$	0.841	0.625	76.849	8.681	6.688	7.928
	AIS5$_{precise}$	0.838	0.618	78.187	8.788	6.765	7.982
	AIS5$_{fuzzy}$	0.835	0.635	79.655	8.869	6.857	8.166
	Set3$_{precise}$ [a]	0.847	0.617	73.848	8.532	6.522	7.697
ET	*Set1*	0.992	0.781	4.001	1.525	1.090	1.255
	AIS2$_{precise}$	0.958	0.773	20.546	4.176	3.152	3.765
	AIS2$_{fuzzy}$	0.955	0.766	21.556	4.174	3.164	3.797
	AIS3$_{precise}$	0.960	0.767	19.295	4.067	3.079	3.691
	AIS3$_{fuzzy}$	0.945	0.772	26.555	4.972	3.829	4.610
	AIS4$_{precise}$	0.959	0.768	19.646	4.095	3.091	3.719
	AIS4$_{fuzzy}$	0.966	0.769	16.417	3.578	2.716	3.266
	AIS5$_{precise}$	0.951	0.773	23.833	4.511	3.428	4.084
	AIS5$_{fuzzy}$	0.952	0.771	23.140	4.393	3.374	4.034
	Set3$_{precise}$ [a]	0.965	0.762	17.043	3.524	2.699	3.245
RF	*Set1*	0.964	0.761	18.837	4.321	3.194	3.721
	AIS2$_{precise}$	0.940	0.757	29.138	5.322	3.997	4.760
	AIS2$_{fuzzy}$	0.934	0.757	31.834	5.575	4.183	4.994
	AIS3$_{precise}$	0.932	0.753	32.837	5.657	4.221	5.028
	AIS3$_{fuzzy}$	0.943	0.756	27.491	5.186	3.895	4.635
	AIS4$_{precise}$	0.932	0.754	32.987	5.663	4.216	5.041
	AIS4$_{fuzzy}$	0.940	0.758	29.145	5.335	3.985	4.762
	AIS5$_{precise}$	0.938	0.751	30.021	5.416	4.040	4.798
	AIS5$_{fuzzy}$	0.949	0.766	24.816	4.914	3.678	4.368

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	$Set3_{precise}^{a}$	0.936	0.756	30.736	5.506	4.112	4.911
AB	$Set1$	0.955	0.758	23.482	4.687	4.036	4.940
	$AIS2_{precise}$	0.956	0.762	21.288	4.333	3.661	4.495
	$AIS2_{fuzzy}$	0.947	0.759	25.519	4.648	3.801	4.620
	$AIS3_{precise}$	0.947	0.755	25.740	4.879	4.148	5.082
	$AIS3_{fuzzy}$	0.958	0.759	20.443	4.176	3.395	4.129
	$AIS4_{precise}$	0.946	0.761	26.012	4.816	4.091	5.023
	$AIS4_{fuzzy}$	0.951	0.751	23.966	4.460	3.733	4.568
	$AIS5_{precise}$	0.950	0.763	24.422	4.732	3.966	4.854
	$AIS5_{fuzzy}$	0.963	0.765	17.825	3.861	3.142	3.804
	$Set3_{precise}^{a}$	0.938	0.752	29.988	5.180	4.370	5.371
GB	$Set1$	0.987	0.764	6.570	2.238	1.633	1.893
	$AIS2_{precise}$	0.958	0.740	20.367	4.158	3.130	3.722
	$AIS2_{fuzzy}$	0.943	0.756	27.321	5.079	3.867	4.574
	$AIS3_{precise}$	0.961	0.749	19.024	3.972	2.993	3.552
	$AIS3_{fuzzy}$	0.955	0.759	21.837	4.113	3.167	3.757
	$AIS4_{precise}$	0.952	0.746	23.273	4.533	3.398	4.068
	$AIS4_{fuzzy}$	0.957	0.752	20.695	4.328	3.319	3.954
	$AIS5_{precise}$	0.941	0.738	28.137	4.950	3.745	4.445
	$AIS5_{fuzzy}$	0.955	0.754	21.469	4.385	3.360	3.967
	$Set3_{precise}^{a}$	0.962	0.743	18.367	3.776	2.825	3.330
XG	$Set1$	0.995	0.771	2.805	1.392	1.008	1.168
	$AIS2_{precise}$	0.964	0.755	17.318	3.687	2.753	3.217
	$AIS2_{fuzzy}$	0.951	0.759	23.599	4.375	3.297	3.850
	$AIS3_{precise}$	0.955	0.753	21.321	4.017	2.952	3.439
	$AIS3_{fuzzy}$	0.951	0.766	23.136	4.388	3.304	3.850
	$AIS4_{precise}$	0.959	0.757	19.655	4.101	3.045	3.564
	$AIS4_{fuzzy}$	0.945	0.756	26.670	4.796	3.626	4.259
	$AIS5_{precise}$	0.940	0.755	28.831	5.231	3.827	4.389
	$AIS5_{fuzzy}$	0.957	0.759	20.713	4.086	3.018	3.475
	$Set3_{precise}^{a}$	0.953	0.734	22.403	4.236	3.177	3.695
LB	$Set1$	0.989	0.755	5.857	2.183	1.652	1.924
	$AIS2_{precise}$	0.941	0.732	28.704	4.895	3.705	4.403
	$AIS2_{fuzzy}$	0.927	0.737	34.903	5.699	4.347	5.157
	$AIS3_{precise}$	0.931	0.742	33.072	5.560	4.152	4.962
	$AIS3_{fuzzy}$	0.922	0.739	37.613	5.982	4.551	5.421
	$AIS4_{precise}$	0.929	0.723	34.023	5.658	4.281	5.131
	$AIS4_{fuzzy}$	0.914	0.723	41.501	6.174	4.671	5.577
	$AIS5_{precise}$	0.927	0.731	34.888	5.579	4.095	4.907
	$AIS5_{fuzzy}$	0.938	0.736	29.667	5.060	3.760	4.486
	$Set3_{precise}^{a}$	0.943	0.723	27.467	4.806	3.609	4.272
SVM	$Set1$	0.861	0.784	73.082	8.540	6.365	7.156
	$AIS2_{precise}$	0.868	0.795	63.462	7.956	5.915	6.810
	$AIS2_{fuzzy}$	0.865	0.796	64.961	8.047	6.059	6.973
	$AIS3_{precise}$	0.864	0.794	65.598	8.088	6.076	7.012
	$AIS3_{fuzzy}$	0.876	0.796	59.629	7.692	5.732	6.604
	$AIS4_{precise}$	0.865	0.793	64.842	8.037	6.034	6.961
	$AIS4_{fuzzy}$	0.870	0.786	62.681	7.882	5.928	6.838
	$AIS5_{precise}$	0.863	0.799	65.964	8.114	6.080	6.999
	$AIS5_{fuzzy}$	0.867	0.798	64.071	7.994	5.977	6.866
	$Set3_{precise}^{a}$	0.858	0.786	68.382	8.263	6.143	7.059

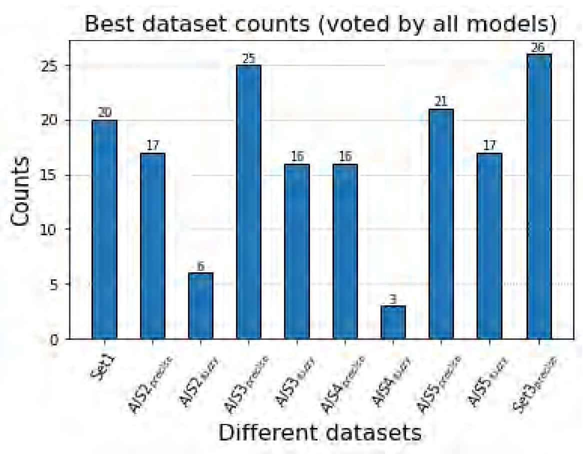
Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
ANN	*Set1*	0.869	0.781	68.911	8.290	6.391	7.296
	$AIS2_{precise}$	0.876	0.773	59.980	7.662	5.914	6.866
	$AIS2_{fuzzy}$	0.901	0.778	48.121	6.838	5.360	6.232
	$AIS3_{precise}$	0.865	0.784	65.208	8.036	6.231	7.285
	$AIS3_{fuzzy}$	0.864	0.781	66.001	8.030	6.270	7.379
	$AIS4_{precise}$	0.859	0.758	68.527	7.974	6.165	7.174
	$AIS4_{fuzzy}$	0.878	0.775	58.615	7.552	5.914	6.949
	$AIS5_{precise}$	0.871	0.780	62.329	7.848	6.054	7.041
	$AIS5_{fuzzy}$	0.868	0.773	63.823	7.814	6.053	7.071
	$Set3_{precise}$ [a]	0.854	0.778	70.184	8.366	6.437	7.518
Ridge	*Set1*	0.814	0.774	97.422	9.868	7.725	8.932
	$AIS2_{precise}$	0.826	0.786	83.624	9.143	7.097	8.337
	$AIS2_{fuzzy}$	0.823	0.783	85.424	9.241	7.252	8.528
	$AIS3_{precise}$	0.835	0.792	79.647	8.923	6.999	8.239
	$AIS3_{fuzzy}$	0.833	0.793	80.613	8.977	7.083	8.330
	$AIS4_{precise}$	0.832	0.790	80.693	8.981	7.014	8.239
	$AIS4_{fuzzy}$	0.828	0.788	82.842	9.100	7.205	8.459
	$AIS5_{precise}$	0.828	0.788	82.760	9.096	7.029	8.250
	$AIS5_{fuzzy}$	0.825	0.786	84.246	9.177	7.121	8.359
	$Set3_{precise}$ [a]	0.830	0.784	81.939	9.050	6.993	8.192
LASSO	*Set1*	0.814	0.773	97.552	9.875	7.711	8.917
	$AIS2_{precise}$	0.826	0.784	83.984	9.162	7.109	8.353
	$AIS2_{fuzzy}$	0.822	0.782	85.539	9.247	7.256	8.536
	$AIS3_{precise}$	0.834	0.793	79.815	8.932	6.999	8.238
	$AIS3_{fuzzy}$	0.832	0.791	81.130	9.005	7.087	8.327
	$AIS4_{precise}$	0.832	0.792	80.977	8.997	7.027	8.257
	$AIS4_{fuzzy}$	0.828	0.789	83.030	9.110	7.212	8.473
	$AIS5_{precise}$	0.828	0.788	82.784	9.097	7.041	8.271
	$AIS5_{fuzzy}$	0.825	0.784	84.263	9.178	7.124	8.369
	$Set3_{precise}$ [a]	0.829	0.786	82.204	9.064	6.997	8.191

Note: $Set3_{precise}$ is the best dataset with DFS1.

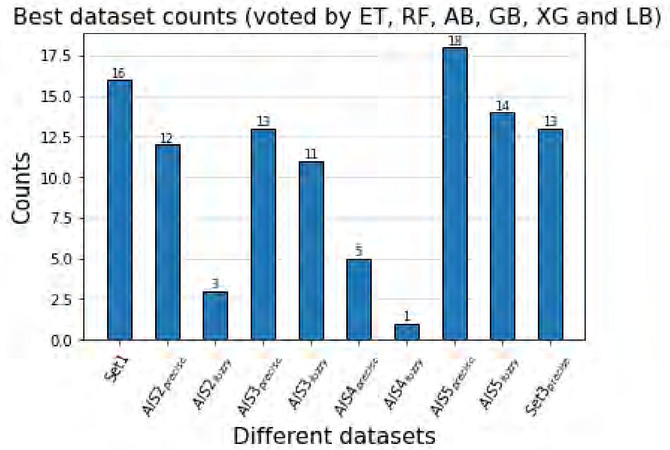
Table 11. DFS2. Best performance of each machine learning model from ten datasets and the datasets that achieve the best performance. R^2 (with two decimal places) is considered as the first priority, and R^2 (test) (with two decimal places) is the secondary performance metric.

Ship	Model	Best R^2	Best R^2 (test)	Datasets
S1	DT	0.85	0.64	*Set1*
	ET	0.99	0.78	*Set1*
	RF	0.96	0.76	*Set1*
	AB	0.96	0.77	$AIS5_{fuzzy}$
	GB	0.99	0.76	*Set1*
	XG	1.00	0.77	*Set1*
	LB	0.99	0.76	*Set1*
	SVM	0.88	0.80	$AIS3_{fuzzy}$
	ANN	0.90	0.78	$AIS2_{fuzzy}$
	Ridge	0.84	0.79	$AIS3_{precise}$
	LASSO	0.83	0.79	$AIS3_{precise}$
S2	DT	0.85	0.65	$AIS2_{precise}$
	ET	0.98	0.78	$AIS5_{precise}$
	RF	0.96	0.77	*Set1*
	AB	0.98	0.75	$AIS2_{precise}$, $AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS4_{precise}$, $AIS5_{precise}$
	GB	0.99	0.77	$AIS4_{precise}$
	XG	0.99	0.77	$Set3_{precise}$
	LB	0.98	0.75	$Set3_{precise}$
	SVM	0.88	0.82	$AIS3_{precise}$, $AIS4_{precise}$
	ANN	0.91	0.79	$Set3_{precise}$
	Ridge	0.84	0.81	$AIS3_{precise}$, $AIS4_{precise}$
	LASSO	0.84	0.81	$AIS3_{precise}$, $AIS4_{precise}$
S3	DT	0.87	0.71	$AIS3_{precise}$
	ET	0.99	0.82	$AIS2_{precise}$, $Set3_{precise}$
	RF	0.96	0.81	$AIS2_{precise}$, $AIS5_{precise}$, $AIS5_{fuzzy}$
	AB	1.00	0.82	$AIS5_{precise}$
	GB	0.97	0.82	$AIS2_{precise}$, $AIS3_{precise}$, $AIS5_{precise}$, $AIS5_{fuzzy}$
	XG	0.98	0.82	$AIS5_{precise}$
	LB	0.95	0.80	$AIS2_{precise}$, $AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS5_{precise}$, $AIS5_{fuzzy}$, $Set3_{precise}$
	SVM	0.85	0.83	$AIS4_{precise}$
	ANN	0.87	0.80	$AIS3_{precise}$, $Set3_{precise}$
	Ridge	0.80	0.80	$AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS4_{precise}$, $AIS4_{fuzzy}$, $AIS5_{precise}$, $Set3_{precise}$
	LASSO	0.80	0.80	$AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS4_{precise}$, $AIS4_{fuzzy}$, $AIS5_{precise}$, $Set3_{precise}$
S4	DT	0.93	0.73	$AIS3_{fuzzy}$
	ET	1.00	0.87	$AIS2_{precise}$, $AIS3_{precise}$, $AIS4_{precise}$, $AIS5_{precise}$, $AIS5_{fuzzy}$, $Set3_{precise}$
	RF	0.98	0.86	$AIS2_{precise}$, $AIS5_{precise}$, $AIS5_{fuzzy}$
	AB	0.99	0.87	$AIS2_{precise}$, $AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS5_{precise}$, $AIS5_{fuzzy}$, $Set3_{precise}$
	GB	1.00	0.87	$AIS3_{precise}$
	XG	1.00	0.87	$AIS3_{fuzzy}$, $Set3_{precise}$
	LB	0.99	0.87	$AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS5_{precise}$, $AIS5_{fuzzy}$
	SVM	0.94	0.85	$AIS2_{fuzzy}$
	ANN	0.95	0.86	$Set3_{precise}$

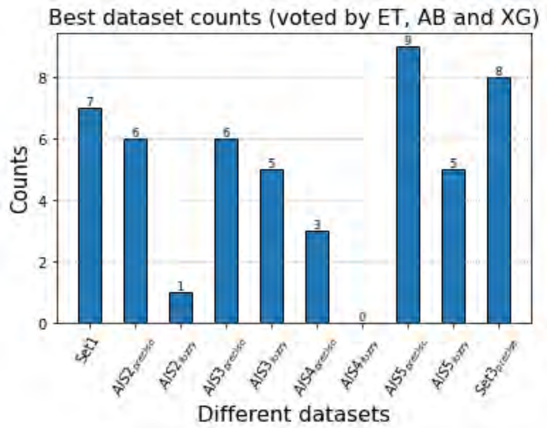
Ship	Model	Best R^2	Best R^2 (test)	Datasets
	Ridge	0.83	0.82	*Set1*
	LASSO	0.83	0.81	$AIS3_{precise}$, $AIS4_{precise}$, $AIS5_{precise}$, $Set3_{precise}$
S5	DT	0.95	0.83	$AIS5_{fuzzy}$
	ET	1.00	0.90	*Set1*, $AIS2_{precise}$, $AIS3_{precise}$, $AIS4_{precise}$
	RF	0.98	0.89	$AIS3_{fuzzy}$, $AIS4_{fuzzy}$, $AIS5_{fuzzy}$
	AB	1.00	0.90	$AIS3_{fuzzy}$, $AIS5_{fuzzy}$
	GB	1.00	0.89	$AIS2_{precise}$, $AIS2_{fuzzy}$, $AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS4_{precise}$, $AIS5_{precise}$
	XG	1.00	0.89	$AIS3_{precise}$
	LB	0.99	0.88	*Set1*, $AIS2_{precise}$, $AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS5_{precise}$
	SVM	0.93	0.88	*Set1*
	ANN	0.94	0.89	$AIS2_{fuzzy}$
	Ridge	0.89	0.88	$AIS2_{precise}$, $AIS5_{fuzzy}$
	LASSO	0.89	0.88	$AIS2_{precise}$
S6	DT	0.86	0.57	$AIS4_{precise}$
	ET	0.99	0.77	*Set1*, $AIS2_{precise}$, $AIS2_{fuzzy}$, $AIS3_{fuzzy}$
	RF	0.96	0.77	*Set1*
	AB	0.99	0.75	$AIS3_{precise}$
	GB	0.97	0.79	*Set1*
	XG	0.97	0.79	*Set1*
	LB	0.97	0.77	$AIS2_{fuzzy}$
	SVM	0.86	0.77	$AIS2_{precise}$
	ANN	0.88	0.76	$AIS2_{precise}$
	Ridge	0.79	0.75	$AIS3_{precise}$, $AIS3_{fuzzy}$, $AIS4_{precise}$
	LASSO	0.79	0.75	$AIS3_{precise}$, $AIS4_{precise}$
S7	DT	0.88	0.68	$Set3_{precise}$
	ET	0.99	0.81	$Set3_{precise}$
	RF	0.97	0.82	$AIS5_{fuzzy}$
	AB	0.99	0.83	$AIS5_{precise}$
	GB	0.99	0.79	$Set3_{precise}$
	XG	0.99	0.78	$Set3_{precise}$
	LB	0.98	0.81	$AIS3_{precise}$, $AIS3_{fuzzy}$
	SVM	0.91	0.79	*Set1*
	ANN	0.90	0.82	$AIS4_{precise}$
	Ridge	0.82	0.76	$Set3_{precise}$
	LASSO	0.82	0.76	$Set3_{precise}$
S8	DT	0.93	0.78	$AIS5_{fuzzy}$
	ET	1.00	0.88	*Set1*, $AIS5_{precise}$, $Set3_{precise}$
	RF	0.98	0.86	*Set1*, $AIS5_{precise}$, $AIS5_{fuzzy}$, $Set3_{precise}$
	AB	1.00	0.87	$AIS5_{precise}$, $AIS5_{fuzzy}$
	GB	0.99	0.86	$AIS5_{precise}$, $AIS5_{fuzzy}$, $Set3_{precise}$
	XG	0.99	0.88	*Set1*
	LB	0.98	0.87	*Set1*
	SVM	0.91	0.87	$Set3_{precise}$
	ANN	0.92	0.86	$Set3_{precise}$
	Ridge	0.88	0.85	$Set3_{precise}$
	LASSO	0.88	0.85	$Set3_{precise}$



(a) Best dataset counts (voted by all models)



(b) Best dataset counts (voted by ET, RF, AB, GB, XG and LB)



(c) Best dataset counts (voted by ET, AB, and XG)

Figure 14. Best datasets voted by machine learning models (DFS2)

7.3.2 Performance Comparison of ML Models

While Table 11 reveals the performances of different ML models, we further report their performances over the best dataset $AIS5_{precise}$ of eight ships in **Table 12**. Tables 11 and 12 both confirm that ET, RF, AB, GB, XG and LB are good candidate models that can be adopted by the shipping industry. Their R^2 values over the best datasets are all above 0.95 and even reach the level of 0.99 to 1.00, while their R^2 performance over the test sets is in the range from 0.75 to 0.90. The remaining models, including DT, SVM, ANN, Ridge, and LASSO, are not recommended for industry applications because their R^2 values on the training sets are usually comparatively low, while the values of R^2 over the test sets have not shown any advantages compared to ET, RF, AB, GB, XG and LB.

Further, the fit performances of RF and LB are usually slightly dominated by ET, AB, GB, and XG, against both R^2 and R^2 (test), which confirms the sufficiency of only installing ET, AB, GB and XG in industry applications related to ship fuel efficiency analysis. GB can also be removed from industry installation once XG has already be installed because GB and XG have close fit performances. Fit errors of ET, AB, GB, and XG on daily bunker fuel consumption, measured by RMSE and MAE, are usually between 0.8 to 4.5 ton/day, though fit errors might be over 4.5 ton/day occasionally if datasets are not carefully chosen.

The experimental results reported in Tables 11 and 12 also rank the performances of eleven ML models into the following four different tiers. The performances of the models in the same tier are quite close, while those of the models in different tiers are significantly different. All the experimental findings for fit performance of ML models are consistent with those from Section 6 with DFS1.

- Tier 1: ET, AB, GB, and XG.
- Tier 2: RF, LB
- Tier 3: DT, SVM, ANN
- Tier 4: Ridge, LASSO.

Table 12. The fit performance of eleven machine learning models over dataset *$AIS5_{precise}$*

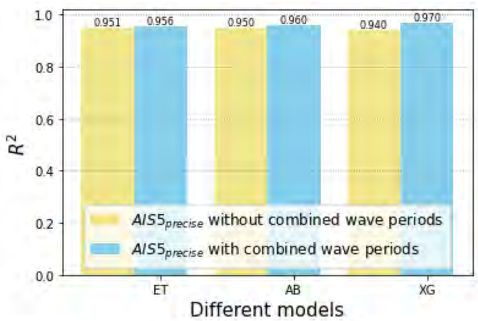
Ship	Model	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
S1	DT	0.838	0.527	78.187	8.788	6.765	7.982
	ET	0.951	0.719	23.833	4.511	3.428	4.084
	RF	0.938	0.692	30.021	5.416	4.040	4.798
	AB	0.950	0.706	24.422	4.732	3.966	4.854
	GB	0.941	0.676	28.137	4.950	3.745	4.445
	XG	0.940	0.696	28.831	5.231	3.827	4.389
	LB	0.927	0.667	34.888	5.579	4.095	4.907
	SVM	0.863	0.752	65.964	8.114	6.080	6.999
	ANN	0.871	0.728	62.329	7.848	6.054	7.041
	Ridge	0.828	0.738	82.760	9.096	7.029	8.250
	LASSO	0.828	0.737	82.784	9.097	7.041	8.271
S2	DT	0.838	0.546	100.776	9.878	7.326	8.621
	ET	0.979	0.717	13.486	3.239	2.390	2.760
	RF	0.948	0.693	32.466	5.645	4.072	4.748
	AB	0.975	0.690	15.834	3.830	3.227	3.765
	GB	0.964	0.705	22.248	4.288	3.136	3.537
	XG	0.965	0.700	21.869	4.287	2.959	3.263
	LB	0.959	0.662	25.714	4.671	3.276	3.798
	SVM	0.880	0.758	74.547	8.599	6.191	6.738
	ANN	0.895	0.744	65.299	8.036	6.139	6.847
	Ridge	0.829	0.760	106.864	10.332	7.780	8.791

Ship	Model	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	LASSO	0.829	0.760	106.939	10.336	7.779	8.774
S3	DT	0.867	0.630	97.066	9.534	6.948	8.222
	ET	0.982	0.799	13.029	3.188	1.858	2.364
	RF	0.963	0.777	27.121	5.168	3.376	4.170
	AB	0.995	0.789	3.588	1.728	1.292	1.513
	GB	0.969	0.787	22.552	4.221	2.710	3.366
	XG	0.976	0.789	17.745	3.884	2.439	2.960
	LB	0.952	0.770	35.323	5.432	3.468	4.412
	SVM	0.840	0.795	115.821	10.727	6.746	8.262
	ANN	0.860	0.781	101.322	10.029	6.763	8.262
	Ridge	0.796	0.771	147.564	12.133	8.425	10.745
	LASSO	0.796	0.770	147.612	12.135	8.426	10.746
S4	DT	0.904	0.706	78.681	8.637	6.425	6.897
	ET	0.998	0.849	1.642	0.927	0.651	0.696
	RF	0.975	0.835	20.097	4.472	3.292	3.590
	AB	0.988	0.847	9.701	2.956	2.466	2.774
	GB	0.991	0.851	7.631	2.412	1.810	1.933
	XG	0.992	0.848	6.441	2.074	1.548	1.640
	LB	0.992	0.842	6.859	2.321	1.771	1.925
	SVM	0.927	0.836	59.875	7.686	5.642	6.018
	ANN	0.939	0.846	50.219	7.062	5.524	5.962
	Ridge	0.827	0.775	141.409	11.888	9.244	9.534
	LASSO	0.827	0.775	141.586	11.895	9.244	9.530
S5	DT	0.948	0.764	28.458	5.104	3.741	5.634
	ET	0.997	0.875	1.475	0.901	0.652	0.988
	RF	0.983	0.857	9.594	3.090	2.281	3.497
	AB	0.995	0.869	2.723	1.476	1.172	2.123
	GB	0.997	0.874	1.628	1.102	0.823	1.310
	XG	0.991	0.871	4.860	1.909	1.382	2.183
	LB	0.991	0.858	4.875	2.049	1.523	2.390
	SVM	0.918	0.856	45.274	6.711	4.874	7.385
	ANN	0.935	0.855	36.276	5.973	4.538	6.997
	Ridge	0.887	0.851	62.515	7.903	5.941	9.040
	LASSO	0.887	0.851	62.689	7.914	5.950	9.050
S6	DT	0.847	0.521	63.834	7.896	5.826	7.701
	ET	0.984	0.729	6.604	2.439	1.780	2.368
	RF	0.959	0.711	17.057	4.116	2.974	3.950
	AB	0.983	0.714	6.959	2.393	1.958	2.888
	GB	0.954	0.731	19.294	4.244	3.282	4.469
	XG	0.948	0.731	21.685	4.533	3.501	4.755
	LB	0.956	0.713	18.441	3.987	3.023	4.106
	SVM	0.846	0.738	64.572	8.017	5.703	7.479
	ANN	0.868	0.731	55.181	7.401	5.673	7.526
	Ridge	0.778	0.707	92.805	9.631	7.393	9.895
	LASSO	0.777	0.704	93.139	9.648	7.387	9.878
S7	DT	0.865	0.633	54.511	7.334	5.473	7.099
	ET	0.978	0.811	8.811	2.497	1.753	2.266
	RF	0.964	0.787	14.703	3.799	2.760	3.604
	AB	0.988	0.802	4.812	2.046	1.675	2.298
	GB	0.975	0.802	10.330	3.084	2.147	2.810
	XG	0.973	0.803	10.967	3.204	2.208	2.823

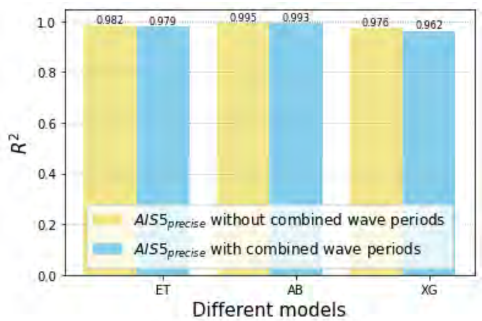
Ship	Model	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	LB	0.981	0.775	7.624	2.614	1.804	2.364
	SVM	0.854	0.789	58.893	7.641	5.411	6.831
	ANN	0.879	0.778	48.929	6.903	5.287	6.805
	Ridge	0.809	0.771	77.312	8.789	6.635	8.431
	LASSO	0.808	0.769	77.734	8.813	6.664	8.483
S8	DT	0.908	0.738	55.429	7.369	5.275	6.362
	ET	0.998	0.860	1.223	0.864	0.549	0.687
	RF	0.975	0.835	15.054	3.848	2.645	3.229
	AB	0.995	0.847	3.047	1.544	1.182	1.516
	GB	0.988	0.839	7.367	2.097	1.438	1.781
	XG	0.973	0.843	16.064	3.747	2.646	3.231
	LB	0.973	0.827	16.500	3.523	2.521	3.138
	SVM	0.897	0.838	62.015	7.865	5.503	6.604
	ANN	0.911	0.824	53.888	7.282	5.345	6.458
	Ridge	0.867	0.817	80.108	8.945	6.728	8.344
	LASSO	0.867	0.818	80.433	8.963	6.737	8.348

7.3.3 The Impact of Wave Period

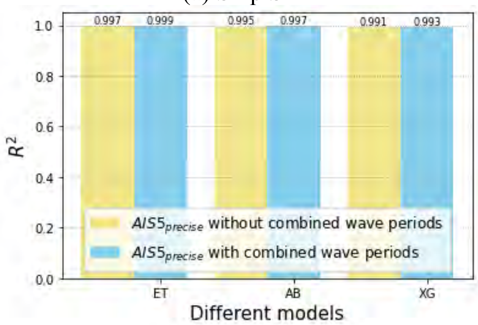
We further added "combined waves period" to the best dataset $AIS5_{precise}$ to see whether adding wave period information improves the experimental result. The experimental results of three best models (ET, AB, and XG) for ships S1, S3, S5, and S8 are shown in **Figure 15**.



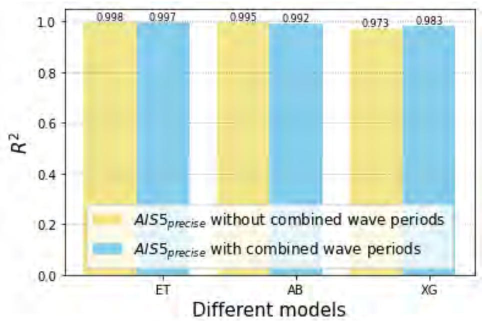
(a) Ship S1



(b) Ship S3



(c) Ship S5

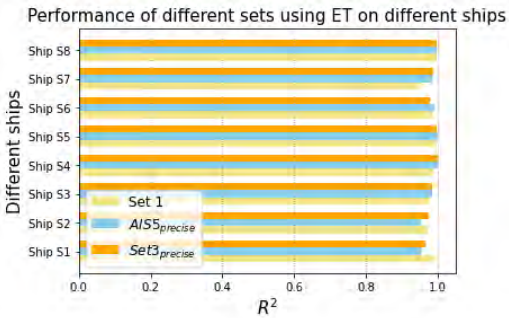


(d) Ship S8

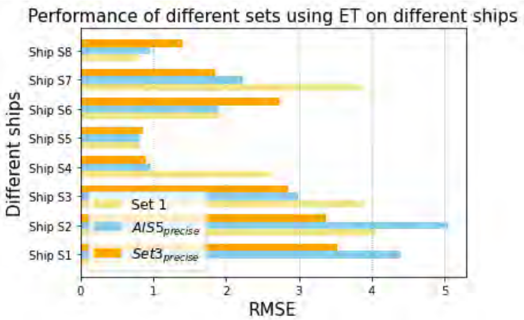
Figure 15. Fit performance of three best models (ET, AB, XG) over dataset *$AIS5_{precise}$*, with and without wave period information

Figure 15 reveals that including wave period information into models might improve the fit performance of models (Ships S1 and S5) but this improvement is often negligible. It might also slightly reduces the fit performance of models. This indicates that the influence of wave period on the fuel consumption rate of a mega conainership at sea is negligible and could be explained by the noises associated with the training data. By considering the consistent result with Section 6, we do not recommend including wave period into models, if voyage report data and meteriological data are combined, no matter whether AIS data is involved.

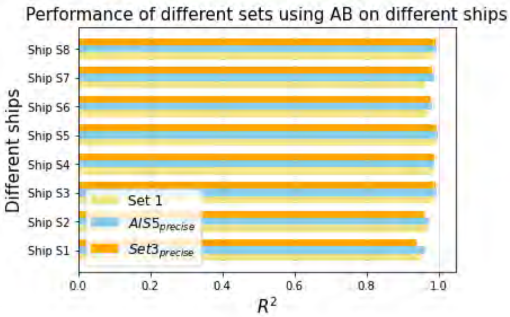
7.3.4 An Experimental Summary of DFS1 and DFS2



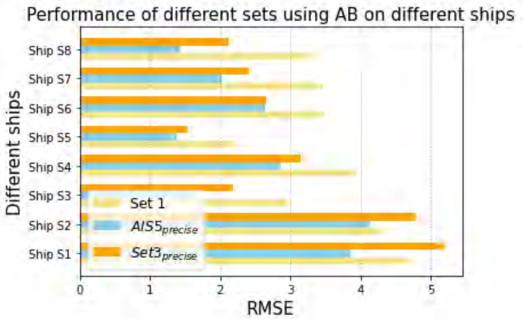
(a) Model: ET; Performance metric: R^2



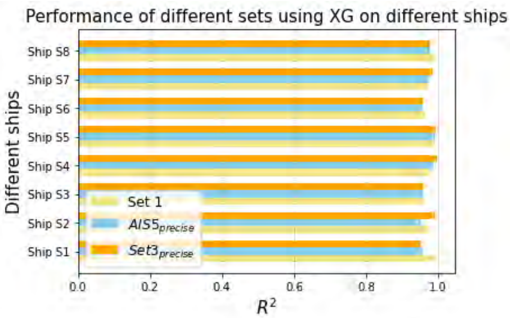
(b) Model: ET; Performance metric: RMSE



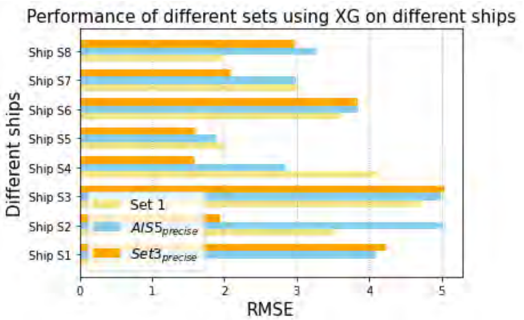
(c) Model: AB; Performance metric: R^2



(d) Model: AB; Performance metric: RMSE



(e) Model: XG; Performance metric: R^2



(f) Model: XG; Performance metric: RMSE

Figure 16. Fit performance (R^2 and RMSE) of three best models (ET, AB, XG) on three best datasets (*Set 1*, *$Set3_{precise}$*, *$AIS5_{precise}$*)

This section summarizes the experimental findings in Section 6 with DFS1 and Section 7 with DFS2. **Figure 16** illustrates the fit performances (R^2 and RMSE) of three best models (ET, AB and XG) over three best datasets: ***Set1*** is the original voyage report data**,** ***Set3**$_{precise}$* represents the best dataset by fusing voyage report data and meteorological data, and ***AIS5**$_{precise}$* represents the best dataset by fusing voyage report data, meteorological data, and AIS data. Overall, as shown in the Tally sheet in Figure 14, ***AIS5**$_{precise}$* is slightly better than ***Set3**$_{precise}$* which in turn is slightly better than ***Set1***. The fit errors of ET, AB and XG over these datasets are normally within 5 ton/day and can be as low as less than 1 ton/day.

Figures 14 and 16 also reveal that the decision of selecting good ML models is interwoven with the decision of selecting good datasets. For instance, in Figure 16, when the model AB is adopted, ***AIS5**$_{precise}$* demonstrates the quality of the best dataset. However, when ET or XG is adopted, ***Set1*** and ***Set3**$_{precise}$* have some chance to win.

7.4. Summary

This study was motivated by a limitation of our previous study in Section 6 that weather and sea condition information derived from the great circle sailing route (suggested by industrial professionals) might be inaccurate. In this study, AIS data is further fused to voyage report data and meteorological data in that AIS data provides actual geographical positions of the ship which further help to retrieve more accurate weather and sea condition information from meteorological data.

To summarize Sections 6 and 7, when dataset choice is considered, the original voyage report dataset ***Set1*** has a decent quality for ship fuel efficiency modeling; if more effort is paid to fuse voyage report data and meteorological data, data quality improves slightly and ***Set3**$_{precise}$* can be adopted. When AIS data is available, further including AIS data might also be beneficial, which suggests the adoption of the dataset ***AIS5**$_{precise}$*. As far as ML model choice is concerned, we recommend the installation of four decision-tree based models including ET, AB, GB, and XG because they usually possess the highest fit performance and good generalization performance. Their performances are also quite robust against random splits of a dataset into training and test sets.

Overall, the best datasets found, including ***Set1***, ***Set3**$_{precise}$*, and ***AIS5**$_{precise}$*, ensure accurate fit performances of best ML models: R^2 on the training set is above 0.96 and even reaches 0.99 to 1.00, and R^2 on the test set is between 0.74 and 0.90; the fit errors measured by RMSE and MAE are between 0.5 and 4.5 ton/day. This accuracy is sufficient for many industry applications and energy-efficient operational measures for shipping companies, including sailing speed optimization, weather routing, and virtual arrivals.

8. Data fusion solution 3 (DFS3): sensor data + meteorological data

8.1 Rationale of Fusing Sensor Data and Meteorological Data

Our previous studies in Sections 6 and 7 address the three RQs by exploring the benefits of fusing voyage report data, meteorological data, and AIS data with widely adopted machine learning (ML) models. Sections 6 and 7 report that given the best datasets found from data fusion, the state-of-the-art decision tree-based ML models achieve a high fit performance with R^2 values above 0.96 and mostly from 0.99 to 1.00, and a good generalization performance with R^2 values on test sets from 0.75 to 0.90. The average fit errors, measured by RMSE and MAE, are between 0.8 and 4.5 ton/day. The selected datasets and ML models are competent for many voyage-based fuel-saving measures such as sailing speed optimization, weather routing, and virtual arrival. Meanwhile, Section 7 points out this might be highest performance we could achieve for mega containerships when voyage report data is used as the main source of bunker fuel consumption. The originates from the fact that voyage report data provides a ship's "daily" fuel consumption information and this data granularity/resolution limits the possibility of further improving the accuracy of ship fuel consumption rate models.

Therefore, it will be interesting to explore some data sources of bunker fuel consumption with a finer data granularity, such as sensor data, and the benefits of combing these data sources with other data sources that provide complementary information. This study makes effort in this direction by fusing sensor data and meteorological data, constructing nine datasets from this data fusion, experimenting with widely adopted ML models over two 9,200-TEU containerships (ships S5 and S6 in Table 1), and revealing the benefits of fusing sensor data and meteorological data.

The wind condition information contained in sensor data is an important indicator of weather conditions confronting a ship. However, conditions of waves, sea water temperature, and sea currents are absent from sensor data. Therefore, we approach publicly accessible meteorological data provided by European Centre for Medium-Range Weather Forecasts (ECWMF) and Copernicus Marine Service (CMEMS, "Copernicus"). The finest meteorological datasets from ECWMF with the resolution of 0.25º (longitude) A?0.25º (latitude)A? A? 1 hour (time) are adopted. ECWMF does not provide the data about sea current conditions. Therefore, the finest datasets for sea currents from Copernicus Marine Service are adopted whose resolution is 0.25º (longitude) A? 0.25º (latitude)A? A? 3 hour (time). For a detailed description of data from ECWMF and CMEMS, see Section 4.

8.2 Approach of Fusing Sensor Data and Meteorological Data

To fuse sensor data and meteorological data, the information contained in sensor data about "*Timestamp (UTC)*", "*Longitude Position*", "*Latitude Position*", and a ship's "*heading (º)*" is used. First, given a sensor data entry, the ship's "*Timestamp (UTC)*", "*Longitude Position*", "*Latitude Position*" is retrieved. Based on this piece of information, the weather and sea conditions the ship experienced at this particular geographical position and time are queried and obtained from the meteorological data sources provided by ECWMF and CMEMS. The weather and sea condition information (wind speed, wind direction, wave direction, sea current speed, sea current direction) from the meteorological data are absolute information regardless of a ship's sailing course and heading. Therefore, second, the ship's "*heading*" information in sensor data is used to convert the absolute weather/sea condition information to relative information to a ship's heading, because weather and sea conditions relative to a ship's heading are more meaningful for fuel efficiency analysis. **Figure 17** illustrates this data fusion approach.

With this data fusion approach, nine datasets shown in **Table 13** are constructed by considering the research purpose of developing fuel consumption rate models for voyage-based energy-efficient operational measures and the endogeneity issue discussed by Yan et al. (2021). Data distributions of the features in Table 13 are presented in **Figures 18 and 19**, for ships S5 and S6, respectively. Before merging sensor data and meteorological data, ships S5 and S6 have 11,901 and 12,484 sensor data entries, respectively. After data fusion with meteorological data, there are 11,410 and 11,968 data entries in total, respectively, by removing the data entries with absent values.

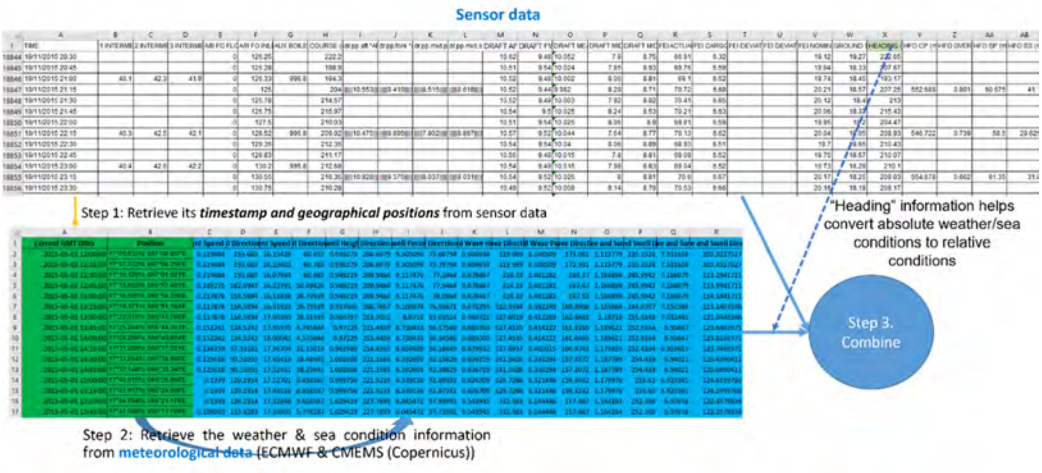


Figure 17. Approach of fusing sensor data and meteorological data

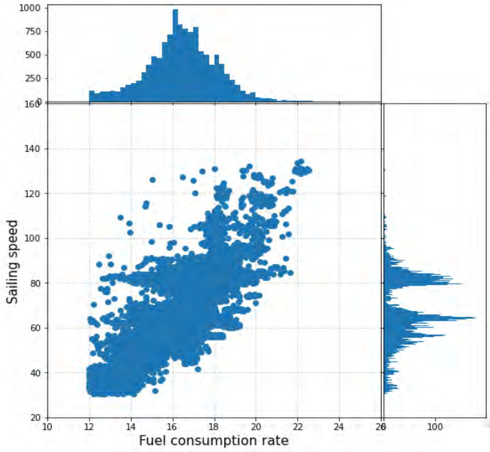
Table 13. Features contained in each dataset (DFS3)

Original datasets	Data source	Features	Dataset								
			Sensor1	*Sensor2*	*Sensor3*	*Sensor4*	*Sensor5*	*Sensor6*	*Sensor7*	*Sensor8*	*Sensor9*
Sensor dataset	Shipping company	Fuel consumption rate	√	√	√	√	√	√	√	√	√
		Sailing speed	√	√	√	√	√	√	√	√	√
		Draft	√	√	√	√	√	√	√	√	√
		Trim	√	√	√	√	√	√	√	√	√
		Wind speed (Rel.)	√	√		√		√		√	
		Wind direction (Rel.)	√	√		√		√		√	
Meteorological data	European Centre for Medium-range Weather Forecasts (ECMWF) [b]	Wind speed (Rel. [a])			√		√		√		√
		Wind direction (Rel. [a])			√		√		√		√
		Combined wave height		√	√	√	√	√	√	√	√
		Combined wave direction (Rel. [a])		√	√	√	√	√	√	√	√
		Combined wave period		√	√			√	√	√	√
		Sea water temperature		√	√	√	√			√	√
	Copernicus Marine Service	Sea current speed (Rel. [a])		√	√	√	√	√	√		
		Sea current direction (Rel. [a])		√	√	√	√	√	√		

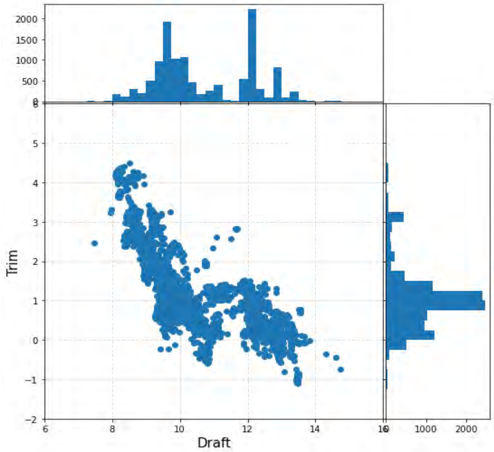
Note:

[a] Relative directions and speeds of wind/waves/sea currents are calculated based on ship's "heading" information from sensor data.

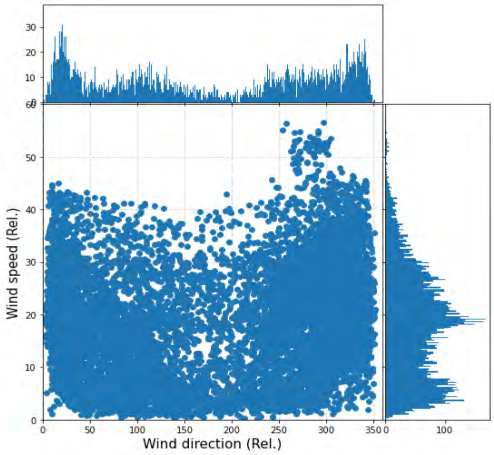
[b] Only combined waves are considered in this study because they represent the impacts of both swell and wind waves, and our preliminary study reveal this is sufficient and additionally including swell conditions and wind waves conditions into this study will not bring any benefits.



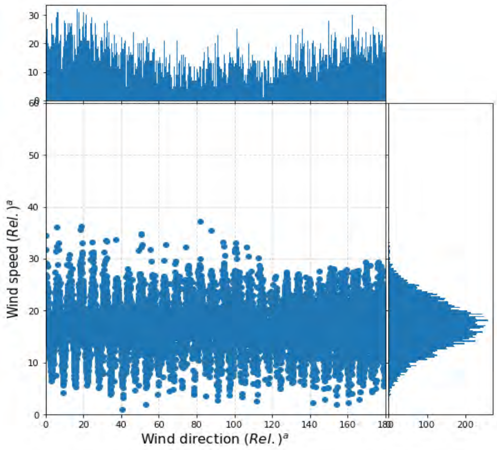
(a) Distribution of ship fuel consumption rate and sailing speed



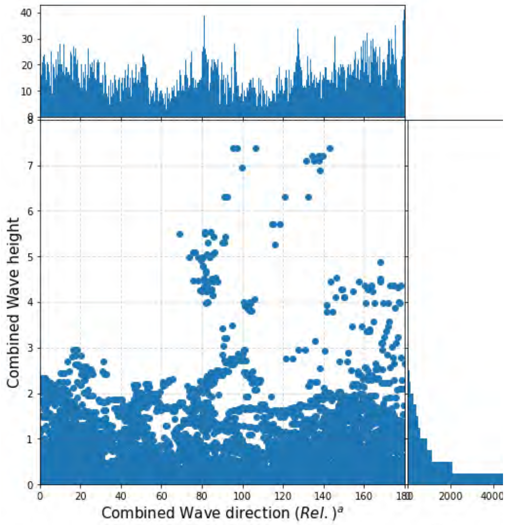
(b) Distribution of trim and draft.



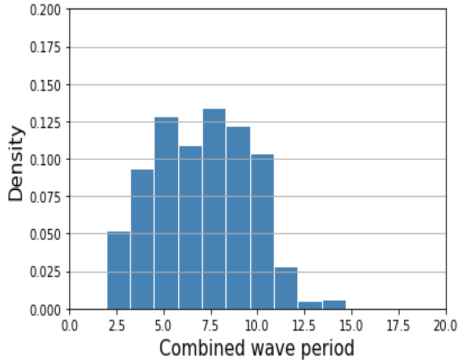
(c) Distribution of wind speed (Rel.) and (Rel.) – Sensor Data.



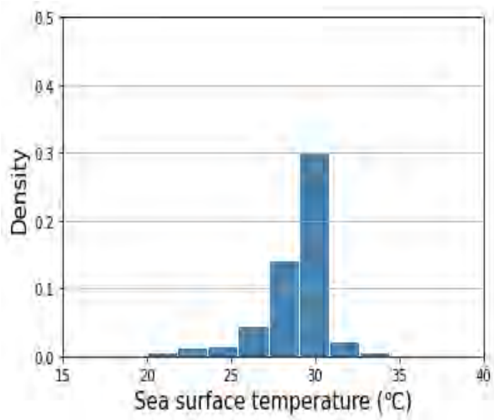
(d) Distribution of wind speed (Rel.) and wind direction (Rel) – ECMWF .



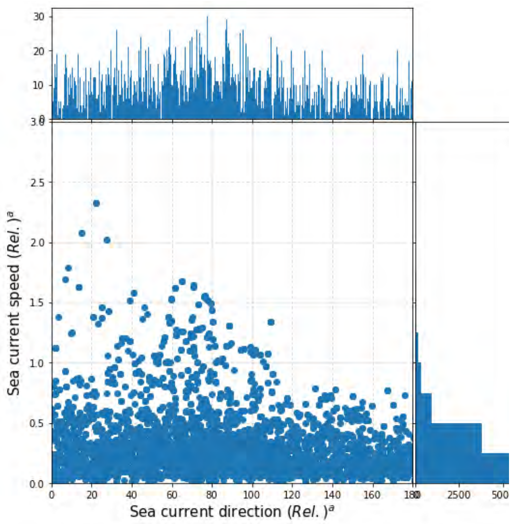
(e) Distribution of wave height and wave direction (Rel.) - ECMWF.



(f) Distribution of wave period - ECMWF.



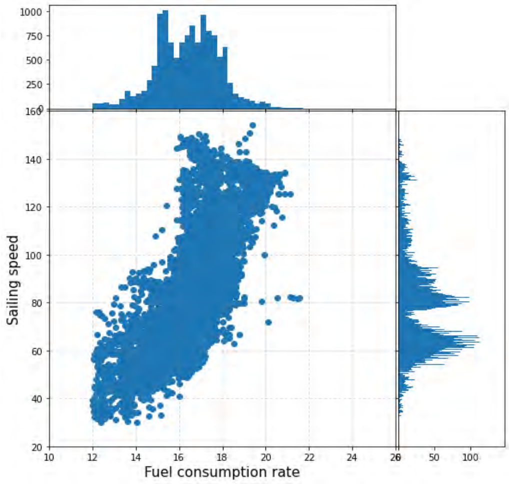
(g) Distribution of seawater temperature - ECMWF.



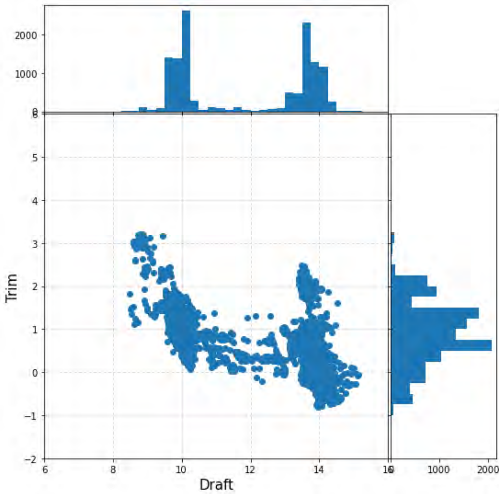
(h) Distribution of current speed and current direction - Copernicus.

Figure 18. The distribution of dataset features in Table 13 (Ship S5)

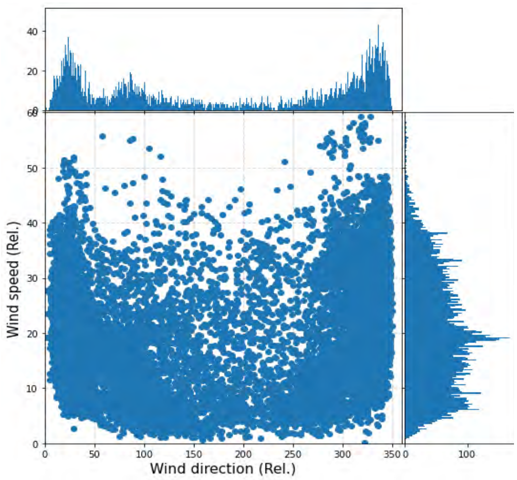
IAMU International Association of Maritime Universities (AMC)



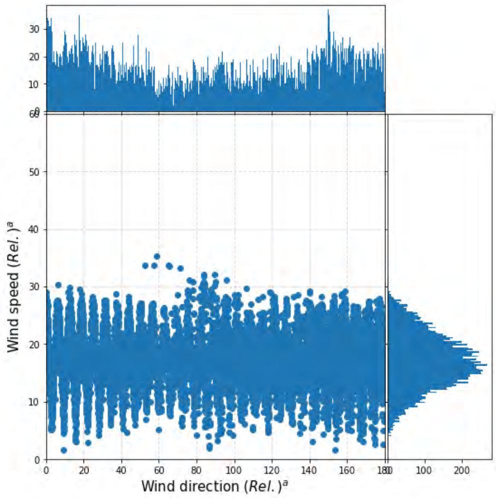
(a) Distribution of ship fuel consumption rate and sailing speed



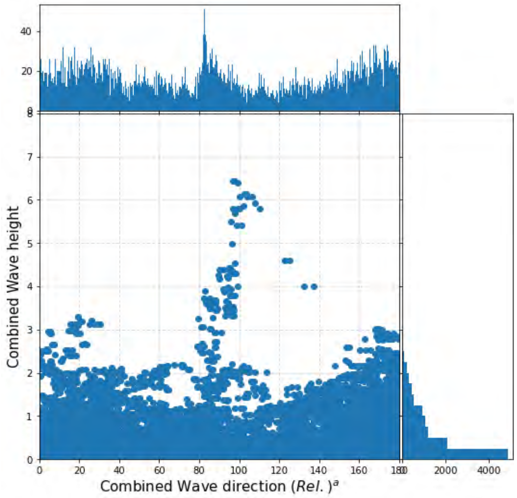
(b) Distribution of trim and draft.



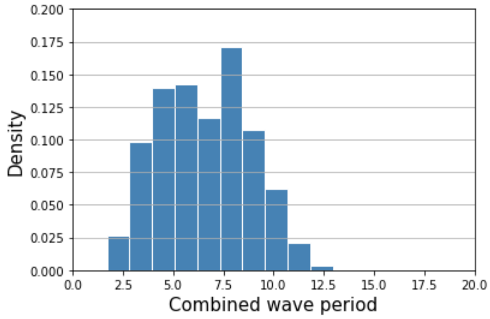
(c) Distribution of wind speed (Rel.) and (Rel.) – Sensor Data.



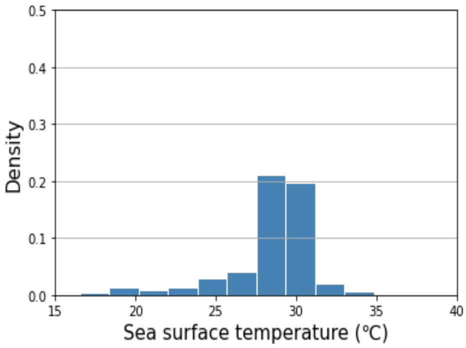
(d) Distribution of wind speed (Rel.) and wind direction (Rel) – ECMWF.



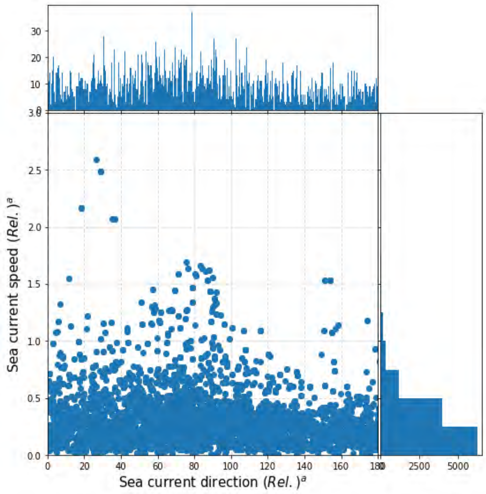
(e) Distribution of wave height and wave direction (Rel.) - ECMWF.



(f) Distribution of wave period - ECMWF.



(g) Distribution of seawater temperature - ECMWF.



(h) Distribution of current speed and current direction - Copernicus.

Figure 19. The distribution of dataset features in Table 13 (Ship S6)

8.3 Experimental Results and Discussion

8.3.1 A Correlation Analysis Towards the Dataset Features

We conducted a pairwise correlation analysis towards the 14 features/variables in Table 13, using Spearman's correlation coefficient which explains how well a variable can be modeled as a monotonical function of the other. **Figures 20 and 21** report the result. The Spearman's correlation coefficient between fuel consumption rate and sailing speed is 0.73 and 0.77 for ships S5 and S6, respectively, which indicates it is mostly acceptable that fuel consumption rate can be modeled as a monotonically increasing function of sailing speed. Similarly, with the confidence level of 0.54, we can expect that fuel consumption rate increases when wind speed (sensor data) increases. Similarly, the statement that an increase in draft results in an increase in fuel consumption rate only achieves a confidence level of 0.32 for ship S5 and 0.24 for ship S6.

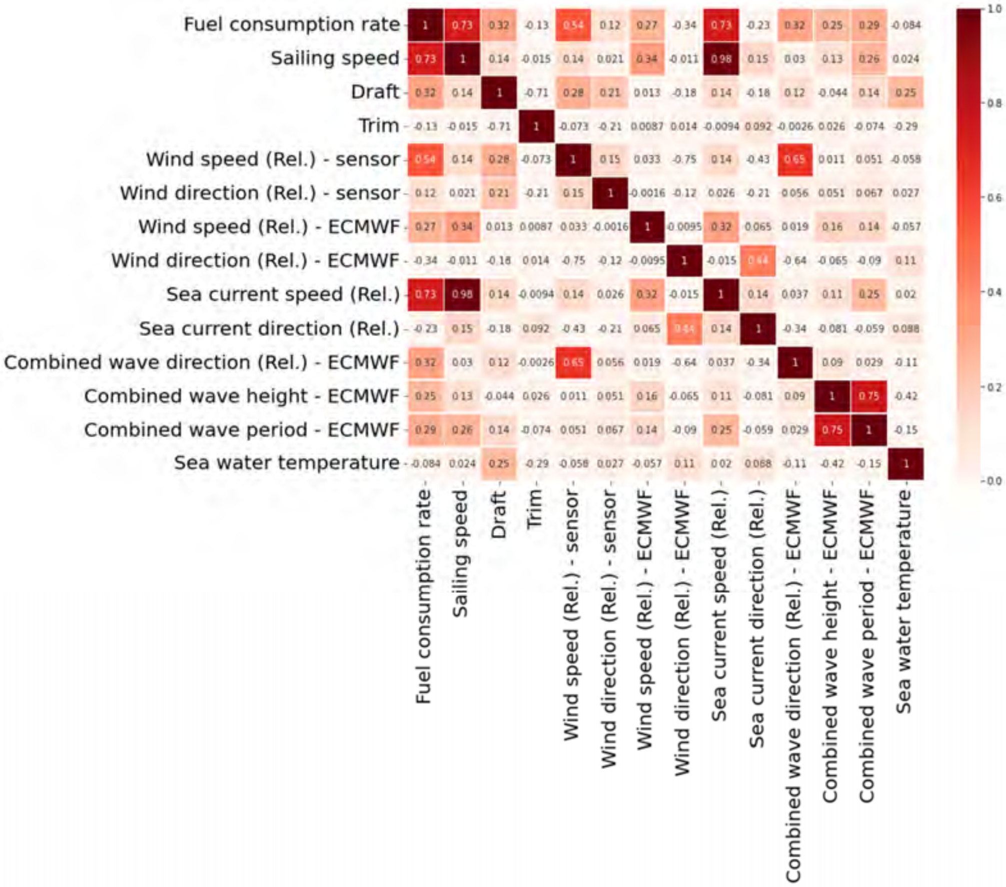


Figure 20. Spearman's correlation coefficients of 14 features (Ship S5)

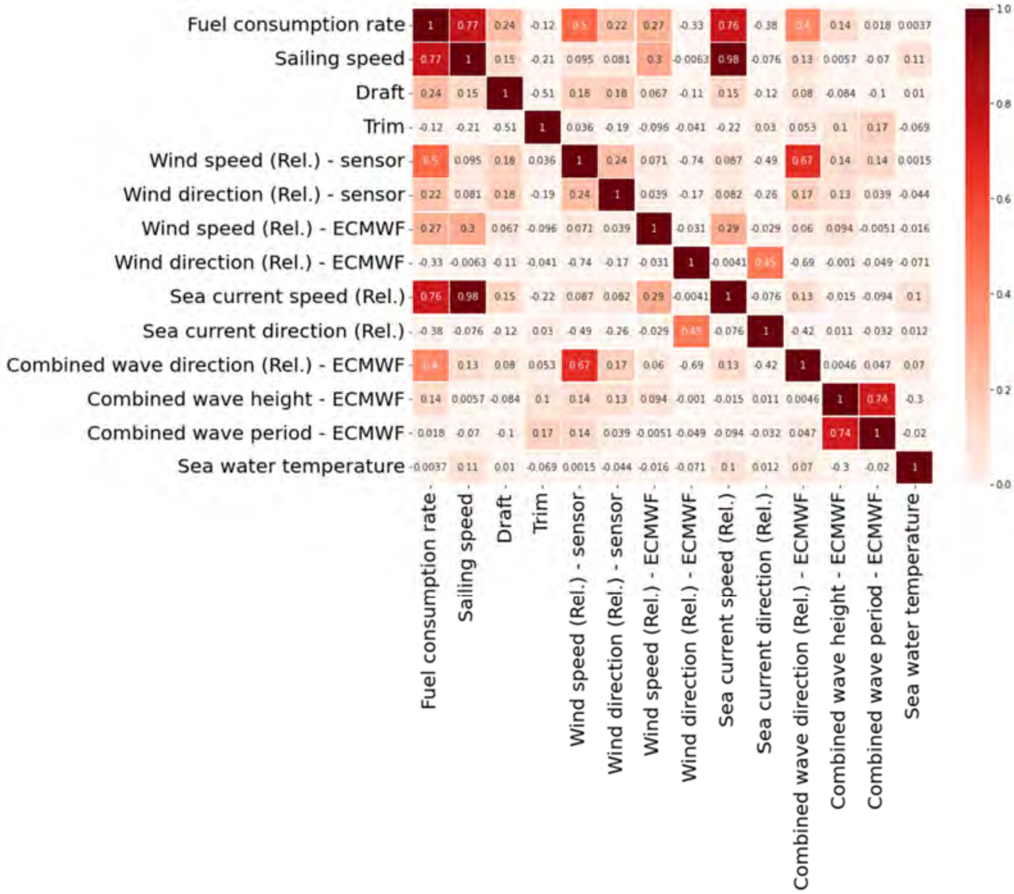


Figure 21. Spearman's correlation coefficients of 14 features (Ship S6)

However, it is inappropriate to utilize the values of Spearman's correlation coefficient to rank the relative importance of different features/variables to ship fuel efficiency (fuel consumption rate), because Spearman's correlation coefficient only indicates how well the relationship between two variables can be reflected by a monotonical function, which is irrelevant to the importance of one variable to the other. The relative importance of different features/variables to ship fuel efficiency will be discussed later.

It is worth noting that "*Wind speed (Rel.) – sensor*" has a rather low correlation with "*Wind speed (Rel.) - ECMWF*", reflected by the coefficient values of 0.033 and 0.071, respectively, for ships S5 and S6. Similarly, "*Wind direction (Rel.) – sensor*" has a rather low and odd (negative) correlation with "*Wind direction (Rel.) - ECMWF*", demonstrated by the coefficient values of -0.12 and -0.17, respectively, for ships S5 and S6. Apparently, the wind condition returned from sensors on board the ship is more reliable than that derived from the hourly data reported by ECMWF which relies on different types of equipment and estimation of some theoretical models. This reflects the inaccuracy of ECMWF data and the possible noises introduced by converting the hourly ECMWF data with absolute wind speeds and directions to 15-min data with relative wind speeds and directions to a ship's heading. This may explain why the best dataset found by this study, *Sensor2*, contains the wind condition data from sensor data rather than from ECMWF. This might also partially explain why DFS1 and DFS2 of this series of studies reported in Sections 6 and 7 have not produced a combined dataset that is always significantly better than the original voyage report data (*Set1*).

8.3.2 Performance of ML models over Nine Datasets and Selection of the Best Datasets

For the two mega containerships S5 and S6, seven widely adopted ML models are experimented with over nine datasets from Table 13. These ML models include Extremely randomized trees (ET) (Geurts et al., 2006), gradient tree boosting (GB) (Friedman, 2001), XGBoost (XG) (Chen and Guestrin, 2016), LightGBM (LB) (Ke et al., 2017), artificial neural network (ANN) (Haykin, 2008), random forest (RF) (Breiman et al., 2001), and support vector machine (SVM) (Boser et al., 1992). Compared to Li et al. (2022) and Du et al. (2022), ridge regression (Ridge) (Hoerl and Kennard, 1970), LASSO (Tibshirani, 1996) and the basic decision tree (DT) model (Breiman et al., 1984) are not considered because these three models had the worst performance in Sections 6 and 7. AdaBoost (AB) (Freund and Schapire, 1997; Drucker, 1997) is not considered in this study as well because (a) GB and XG are considered as the advances towards AB in machine learning theory development; and (b) our preliminary experiments revealed the performance of AB is close to but slightly worse than GB and XG, with the large amount of sensor data in this study. All the experimental settings are the same as Sections 6 and 7.

Same to Sections 6 and 7, for each dataset in Table 13, we randomly divide it into a training set with 80% of its data entries and a test set with 20% of its data entries. This is termed as a *split* of this dataset. To overcome the impact of randomness in data splitting, we produce 20 splits for each dataset in Table 13, and average the performances of each ML model over 20 random splits as the performance indicator of this ML model for this dataset. We also adopt the same performance metrics for a ML model as Sections 6 and 7, including R^2, *MSE*, *RMSE*, *MAE* and *MAPE* for training set, R^2 *(test)* for test set. **Table 14** reports the performances of seven ML models over nine datasets for ship S5, while **Table 15** reports the result for ship S6.

Table 14. The fit performance of seven machine learning models for ship S5 (DFS3)

Model [a]	Dataset	R^2 [b]	R^2 (test) [c]	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
ET	*Sensor1*	0.998	0.924	20.790	0.706	0.385	0.589
	Sensor2	1.000	0.969	4.371	0.306	0.166	0.263
	Sensor3	1.000	0.968	3.348	0.258	0.131	0.207
	Sensor4	1.000	0.968	4.155	0.273	0.147	0.232
	Sensor5	1.000	0.965	5.212	0.327	0.173	0.271
	Sensor6	1.000	0.968	3.943	0.266	0.141	0.223
	Sensor7	0.999	0.966	8.352	0.395	0.210	0.330
	Sensor8	1.000	0.969	5.148	0.345	0.184	0.291
	Sensor9	1.000	0.967	2.704	0.212	0.109	0.171
GB	*Sensor1*	0.990	0.925	111.231	1.586	1.051	1.617
	Sensor2	0.999	0.969	13.092	0.492	0.301	0.474
	Sensor3	0.999	0.967	11.210	0.455	0.279	0.437
	Sensor4	0.999	0.967	9.303	0.414	0.243	0.384
	Sensor5	0.999	0.965	5.941	0.341	0.207	0.325
	Sensor6	0.999	0.967	8.743	0.423	0.250	0.392
	Sensor7	0.998	0.963	27.591	0.777	0.537	0.844
	Sensor8	0.999	0.967	11.513	0.465	0.289	0.455
	Sensor9	0.999	0.965	8.100	0.400	0.245	0.383
LB	*Sensor1*	0.981	0.917	215.192	2.240	1.460	2.244
	Sensor2	0.996	0.965	46.974	1.021	0.694	1.089
	Sensor3	0.997	0.964	36.110	0.888	0.594	0.929

Model [a]	Dataset	R^2 [b]	R^2 (test) [c]	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	Sensor4	0.997	0.963	34.926	0.878	0.598	0.936
	Sensor5	0.996	0.959	45.018	0.984	0.672	1.050
	Sensor6	0.996	0.962	45.985	1.002	0.673	1.053
	Sensor7	0.996	0.960	45.506	0.989	0.662	1.036
	Sensor8	0.996	0.962	50.098	1.066	0.727	1.146
	Sensor9	0.995	0.959	55.632	1.110	0.760	1.192
ANN	*Sensor1*	0.841	0.834	1782.833	6.541	4.696	7.164
	Sensor2	0.923	0.917	860.785	4.545	3.229	4.946
	Sensor3	0.902	0.892	1092.319	5.120	3.670	5.609
	Sensor4	0.918	0.910	914.236	4.684	3.343	5.116
	Sensor5	0.896	0.888	1161.034	5.278	3.768	5.735
	Sensor6	0.917	0.911	924.206	4.709	3.355	5.143
	Sensor7	0.898	0.890	1140.669	5.232	3.744	5.721
	Sensor8	0.909	0.903	1014.211	4.933	3.562	5.458
	Sensor9	0.886	0.878	1273.935	5.529	4.014	6.157
RF	*Sensor1*	0.986	0.924	153.059	1.909	1.047	1.588
	Sensor2	0.995	0.965	60.053	1.200	0.626	0.975
	Sensor3	0.994	0.964	63.344	1.233	0.648	1.004
	Sensor4	0.995	0.964	60.261	1.202	0.635	0.987
	Sensor5	0.994	0.960	69.331	1.288	0.692	1.065
	Sensor6	0.995	0.964	60.744	1.206	0.640	0.994
	Sensor7	0.994	0.961	69.599	1.291	0.691	1.064
	Sensor8	0.994	0.964	63.108	1.230	0.630	0.980
	Sensor9	0.994	0.962	66.961	1.267	0.646	1.002
SVM	*Sensor1*	0.960	0.909	446.793	3.268	1.945	2.950
	Sensor2	0.987	0.963	141.353	1.838	1.254	1.943
	Sensor3	0.986	0.955	158.722	1.944	1.359	2.101
	Sensor4	0.986	0.961	158.958	1.949	1.312	2.026
	Sensor5	0.982	0.949	202.504	2.200	1.462	2.245
	Sensor6	0.986	0.961	153.639	1.915	1.281	1.982
	Sensor7	0.984	0.950	179.111	2.071	1.406	2.172
	Sensor8	0.986	0.961	158.008	1.941	1.300	2.018
	Sensor9	0.985	0.950	166.134	1.992	1.376	2.130
XG	*Sensor1*	0.988	0.918	133.212	1.743	1.191	1.828
	Sensor2	0.999	0.966	13.331	0.529	0.350	0.548
	Sensor3	0.998	0.966	18.250	0.637	0.413	0.644
	Sensor4	0.998	0.966	17.460	0.613	0.400	0.626
	Sensor5	0.999	0.963	15.971	0.566	0.364	0.566
	Sensor6	0.999	0.965	12.433	0.507	0.323	0.504
	Sensor7	0.999	0.963	15.904	0.586	0.365	0.568
	Sensor8	0.999	0.965	15.598	0.569	0.367	0.575
	Sensor9	0.999	0.964	13.206	0.529	0.343	0.535

Notes.

[a] AdaBoost (AB) was not included in this study because (a) GB and XG are considered as the advances towards AB in machine learning theory development; (b) our preliminary experiments revealed the performance of AB is close to but slightly worse than GB and XG, with the large amount of sensor data in this study.

[b] R^2 for training set.

[c] R^2 for test set.

Table 15. The fit performance of seven machine learning models for ship S6 (DFS3)

Model [a]	Dataset	R^2 [b]	R^2 (test) [c]	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
ET	*Sensor1*	0.998	0.943	47.029	0.987	0.646	0.841
	Sensor2	1.000	0.977	5.862	0.326	0.211	0.281
	Sensor3	1.000	0.974	7.144	0.348	0.230	0.307
	Sensor4	1.000	0.976	7.773	0.378	0.248	0.333
	Sensor5	1.000	0.972	9.945	0.462	0.305	0.406
	Sensor6	1.000	0.975	6.061	0.347	0.222	0.295
	Sensor7	1.000	0.972	9.847	0.471	0.305	0.404
	Sensor8	1.000	0.975	8.721	0.433	0.290	0.387
	Sensor9	1.000	0.971	8.460	0.442	0.289	0.384
GB	*Sensor1*	0.990	0.940	202.349	2.128	1.540	2.049
	Sensor2	0.999	0.976	15.356	0.558	0.376	0.507
	Sensor3	0.999	0.973	15.331	0.557	0.391	0.525
	Sensor4	0.999	0.975	15.925	0.597	0.389	0.524
	Sensor5	1.000	0.972	9.389	0.424	0.283	0.381
	Sensor6	0.999	0.974	11.789	0.486	0.319	0.430
	Sensor7	0.998	0.969	48.980	1.067	0.781	1.044
	Sensor8	0.998	0.974	44.638	0.989	0.667	0.888
	Sensor9	0.999	0.970	18.288	0.570	0.388	0.517
LB	*Sensor1*	0.982	0.935	381.070	2.939	2.120	2.795
	Sensor2	0.997	0.973	72.018	1.284	0.939	1.254
	Sensor3	0.996	0.970	77.293	1.322	0.967	1.295
	Sensor4	0.996	0.971	86.995	1.408	1.033	1.379
	Sensor5	0.996	0.967	82.839	1.337	0.975	1.305
	Sensor6	0.995	0.970	111.632	1.591	1.160	1.547
	Sensor7	0.996	0.967	84.585	1.388	1.006	1.344
	Sensor8	0.996	0.970	89.087	1.412	1.032	1.374
	Sensor9	0.996	0.965	78.583	1.308	0.954	1.273
ANN	*Sensor1*	0.866	0.863	2832.667	8.245	5.935	7.615
	Sensor2	0.935	0.931	1372.573	5.739	4.260	5.594
	Sensor3	0.926	0.922	1562.461	6.123	4.554	5.983
	Sensor4	0.932	0.928	1445.245	5.889	4.335	5.694
	Sensor5	0.921	0.917	1668.811	6.328	4.672	6.134
	Sensor6	0.930	0.926	1485.849	5.971	4.408	5.763
	Sensor7	0.917	0.913	1751.915	6.484	4.790	6.260
	Sensor8	0.916	0.913	1764.069	6.506	4.801	6.223
	Sensor9	0.900	0.897	2103.728	7.105	5.209	6.732
RF	*Sensor1*	0.989	0.939	223.262	2.308	1.529	1.984
	Sensor2	0.996	0.973	89.147	1.462	0.936	1.227
	Sensor3	0.995	0.970	98.361	1.536	0.974	1.275
	Sensor4	0.995	0.971	96.748	1.523	0.981	1.285
	Sensor5	0.995	0.968	106.136	1.595	1.014	1.324
	Sensor6	0.995	0.970	98.826	1.539	0.989	1.291
	Sensor7	0.995	0.967	109.389	1.619	1.033	1.347
	Sensor8	0.995	0.970	103.479	1.575	1.016	1.326
	Sensor9	0.995	0.967	113.163	1.647	1.042	1.355
SVM	*Sensor1*	0.958	0.923	877.202	4.585	3.047	3.889
	Sensor2	0.989	0.972	237.134	2.384	1.851	2.444
	Sensor3	0.985	0.961	316.781	2.753	2.048	2.689
	Sensor4	0.988	0.970	257.851	2.485	1.909	2.517
	Sensor5	0.983	0.958	357.214	2.925	2.146	2.814

Model [a]	Dataset	R^2 [b]	R^2 (test) [c]	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	Sensor6	0.987	0.969	274.848	2.564	1.923	2.531
	Sensor7	0.984	0.956	338.096	2.841	2.081	2.729
	Sensor8	0.988	0.968	245.671	2.426	1.872	2.453
	Sensor9	0.984	0.950	331.498	2.812	2.085	2.712
XG	*Sensor1*	0.987	0.937	269.027	2.486	1.799	2.379
	Sensor2	0.999	0.975	24.510	0.740	0.517	0.690
	Sensor3	0.999	0.972	25.761	0.735	0.516	0.686
	Sensor4	0.999	0.974	25.744	0.743	0.525	0.701
	Sensor5	0.999	0.971	27.194	0.758	0.526	0.701
	Sensor6	0.999	0.973	29.003	0.776	0.545	0.726
	Sensor7	0.999	0.970	22.467	0.675	0.471	0.627
	Sensor8	0.998	0.973	32.380	0.852	0.599	0.798
	Sensor9	0.999	0.970	26.362	0.761	0.533	0.707

Notes.

[a] AdaBoost (AB) was not included in this study because (a) GB and XG are considered as the advances towards AB in machine learning theory development; (b) our preliminary experiments revealed the performance of AB is close to but slightly worse than GB and XG, with the large amount of sensor data in this study.

[b] R^2 for training set.

[c] R^2 for test set.

Table 16. DFS3. Best performance of each machine learning model from nine datasets and the datasets that achieve the best performance. R^2 (train) (with three decimal places) is considered as the first priority, and R^2 (test) (with three decimal places) is the secondary performance metric.

Ship	Model	Best R^2	Best R^2 (test)	Datasets
S5	ET	1.000	0.969	*Sensor2, Sensor8*
	GB	0.999	0.969	*Sensor2*
	LB	0.997	0.964	*Sensor3*
	ANN	0.923	0.917	*Sensor2*
	RF	0.995	0.965	*Sensor2*
	SVM	0.987	0.963	*Sensor2*
	XG	0.999	0.966	*Sensor2*
S6	ET	1.000	0.977	*Sensor2*
	GB	1.000	0.972	*Sensor5*
	LB	0.997	0.973	*Sensor2*
	ANN	0.935	0.931	*Sensor2*
	RF	0.996	0.973	*Sensor2*
	SVM	0.989	0.972	*Sensor2*
	XG	0.999	0.975	*Sensor2*

In Tables 14 and 15, the performance of ML models and quality of datasets are interwoven together. To decouple the performance of ML and quality of datasets, we adopt the same voting scheme as Sections 6 and 7 and present the voting result in **Table 16**. In Table 16, each ML model acts as a voter and votes for best datasets (candidates) by considering R^2 (with three decimal places) as the first priority and R^2 (test) (with three decimal places) as the secondary performance metric. The last column is the votes of each ML model (voter).

It can be seen from Table 16 that *Sensor2* is the best dataset voted by all models. This confirms the benefits of fusing sensor data and meteorological data. As an insight for industry application, when these two datasets are combined, as far as weather and sea conditions are considered, wind condition information from sensor dataset is preferred. Then the information about wave conditions, sea water temperature and sea currents from meteorological data sources can be utilized.

When the results from *Sensor1* and those from *Sensor2* are compared, Tables 14 and 15 also demonstrate the benefits of fusing sensor data and meteorological data over all the ML models. For instance, for ship S6, the XG model produces a RMSE of 2.486 ton/day with the dataset *Sensor1*, and a RMSE of 0.740 ton/day with the dataset *Sensor2*. The XG model also generates a MAE of 1.799 ton/day with the dataset *Sensor1*, and a MAE of 0.517 ton/day with the dataset *Sensor2*. Similarly, with the ANN model, ship S6's RMSE and MAE are 8.245 ton/day and 5.935 ton/day, respectively, with the dataset *Sensor1*. The dataset *Sensor2* reduces the values of these two metrics to 5.739 ton/day and 4.260 ton/day, respectively.

8.3.3 Performance Comparison of ML Models

To further compare the performances of ML models, we tabulate the performance metrics of seven ML models over the best dataset *Sensor2* in **Table 17**. Table 17 finds ET, GB and XG are the best machine learning in terms of all the performance metrics, which is consistent with the results of Sections 6 and 7. These three models achieve their R^2 at 0.999 or 1.000 on the training sets, and their R^2 values over the test sets are also all above 0.966. Their RMSE values are below 0.75 ton/day, and MAE below 0.52 ton/day. These results are rather beyond the requirements of most industry applications for ship fuel efficiency analysis.

Compared to results in Table 6 with those in Sections 6 and 7, it can be seen that modeling accuracy (fit performance) and generalization performance of every ML model have been improved with sensor data, compared to voyage report data. This is because sensor data is superior to voyage report data in both the size and quality.

Table 17. The fit performance of seven machine learning models over dataset *Sensor2*

Ship	Model	$R^{2\ b}$	R^2 (test) [c]	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
S5	ET	1.000	0.969	4.371	0.306	0.166	0.263
	GB	0.999	0.969	13.092	0.492	0.301	0.474
	LB	0.996	0.965	46.974	1.021	0.694	1.089
	ANN	0.923	0.917	860.785	4.545	3.229	4.946
	RF	0.995	0.965	60.053	1.200	0.626	0.975
	SVM	0.987	0.963	141.353	1.838	1.254	1.943
	XG	0.999	0.966	13.331	0.529	0.350	0.548
S6	ET	1.000	0.977	5.862	0.326	0.211	0.281
	GB	0.999	0.976	15.356	0.558	0.376	0.507
	LB	0.997	0.973	72.018	1.284	0.939	1.254
	ANN	0.935	0.931	1372.573	5.739	4.260	5.594
	RF	0.996	0.973	89.147	1.462	0.936	1.227
	SVM	0.989	0.972	237.134	2.384	1.851	2.444
	XG	0.999	0.975	24.510	0.740	0.517	0.690

[a] R^2 for training set.
[b] R^2 for test set.

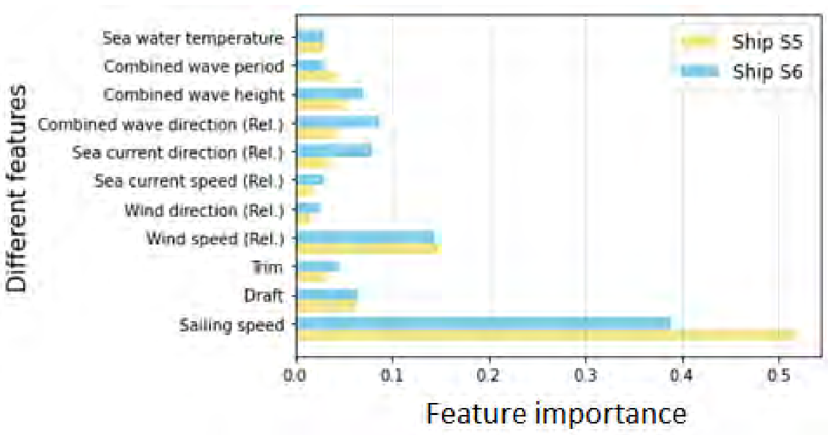
8.3.4 Relative Importance of Each Determinant to Ship Fuel Efficiency

A decision tree-based model possesses a good interpretability in explaining the relative importance of input variables of the model (features of the dataset) to the output/target variable. Therefore, over the best dataset identified, *Sensor2*, decision tree models, including ET, GB, XG, and RF are used to analyze the relative importance of each determinant (feature/input variable) to fuel consumption rate (the output/target variable). The results are shown in **Figure 22**.

Figure 22 shows the consistent results of four decision tree-based models towards the importance of relevant factors for ship fuel efficiency. Most results are also consistent with the findings in Section 6 with voyage report data and meteorological data. First, sailing speed is the most significant determinant of a ship's fuel efficiency. Second, the impact of draft/displacement (relative importance: around 0.06) is much lower than that of many factors about sea and weather conditions such as wind or waves, and of course cannot compete with the total impact of weather and sea conditions. Third, trim's impact is not negligible, whose relative importance is between 0.03 and 0.05, which justifies the necessity of conducting trim optimization in the shipping industry.

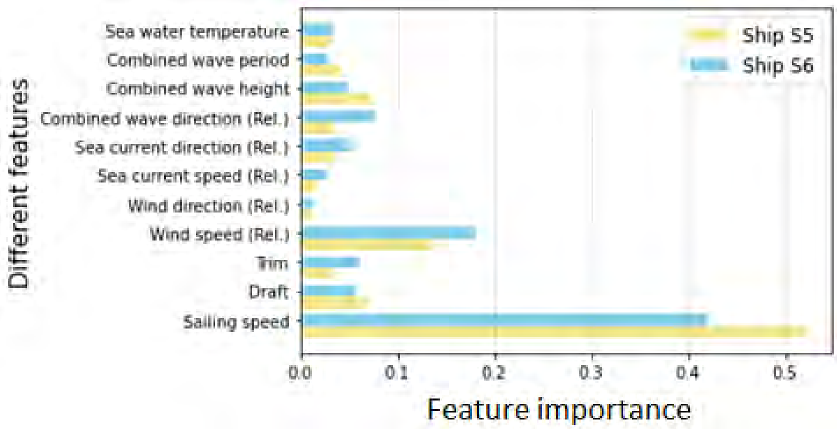
Some results do not fully agree with those from Section 6. First, wind plays the most critical role in weather/sea conditions whose importance reaches around 0.2 if wind speed and wind direction are both considered. Specifically, the importance of wind to ship fuel efficiency is higher than waves (relative importance: around 0.15 totaling the values of three relevant variables). This is not surprising because wind contributes to wind waves as well as air resistance which is particularly significant for a containership with containers on the deck. Second, the importance of sea currents revealed by this study, roughly from 0.05 to 1.0, is higher than that found by Section 6. Third, sea water temperature owns the least importance at 0.03-0.05. Compared to Section 6, the importances of sea currents and sea water temperture may be more convincing for seafarers.

Overall, all the three parts of our studies (DFS1 in Section 6, DFS2 in Section 7, and DFS3 in Section 8) demonstrate consistent results regarding the importance of these determinants/variables to ship fuel efficiency, including the most significant role of sailing speed, higher importance of weather and sea conditions than draft/displacement, and the minor but nonnegligible impact of trim. Regarding the inconsistent results of this study with DFS1 in Section 6 on the relative importances of wind, waves, sea currents, and sea water temperature, the superiority of sensor data to voyage report data in data size and quality makes the results of this study more convincing.

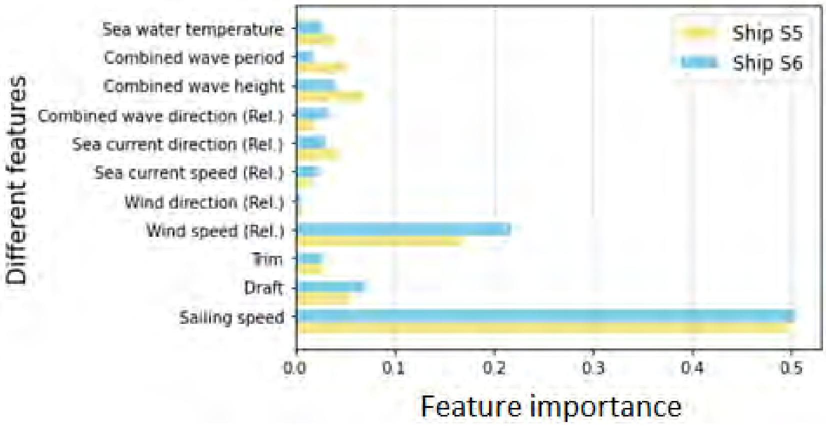


(a) ET model.

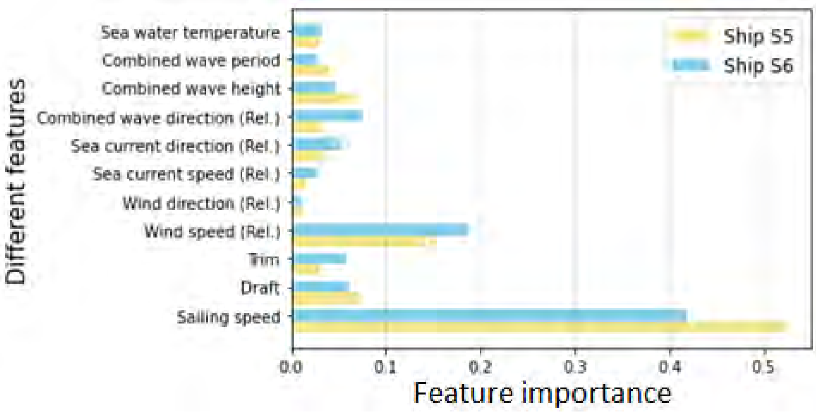
IAMU International Association of Maritime Universities (AMC)



(b) GB model.



(c) XG model.



(d) RF model.

Figure 22. The average relative importance of model's input variables (dataset features) with DFS3

8.3.5 A Rolling Horizon Approach in Practice

The above experiments adopted sensor data of about 6 months, with the data of about 5 months (80%) for training, and the data of about 1 month (20%) for testing. Given the decent size and quality of sensor data and the proven good performances of ML models in above experiments, one may ask a question '*how much sensor data is needed to achieve good fit and generalization preformance in practice?*'

The most possible application scenario in practice will be a rolling horizon approach. For instance, in a "3-month training + 1-month test/application" scenario, our sensor data from May to November might be involved in four trainings and four tests/applications if a rolling horizon principle is adopted. First, we utilize three-month data of May, June, and July to train the model (Model 1), and Model 1 will be adopted in the whole month of August at sea. Once August ends and actual data in August is accumulated/realized, we can verify the performance of Model 1 for August. This is equivalent to using the data of May, June, and July for training and the data of August for testing. Second, when time comes to the very beginning of September, we can utilize the data of most recent three months (June, July, and August) to train the model and obtain a new Model 2. Then we apply/test Model 2 in the whole September. Similarly, data for July, August, and September will produce a new Model 3 and October will verify/apply Model 3. At last, a new Model 4 is obtained using the data of August, September, and October, and applied/tested in November. **Table 18** summarizes this rolling horizon process.

Table 18. A rolling horizon process for a "3-month training + 1-month test/applicatoin" scenario with sensor data from May to November

		Data for training	Model trained	Time for application (data for test)
RollingSet1	1st rolling horizon	May, June, July	Model 1	August
RollingSet2	2nd rolling horizon	June, July, August	Model 2	September
RollingSet3	3rd rolling horizon	July, August, September	Model 3	October
RollingSet4	4th rolling horizon	August, September, October	Model 4	November

In a "3-month training + 1-month test/applicatoin" scenario, we report in **Figure 23** the performances of ET, GB and XG in a rolling horizon process for ships S5 and S6. Similarly, a "2-month training + 1-month test/applicatoin" scenario or a "1-month training + 1-month test/applicatoin" scenario is also possible. The corresponding results are reported in Appendices (**Figures A1, A2, A3 and A4**).

Figures 23, A1, A2, A3 and A4 all reveal rather high R^2 values for training sets, but unacceptable R^2 values for test sets (application periods). The R^2 values on test sets (application periods) can be lower than 0.5 and even negative. This indicates that none of "3-month training + 1-month test/applicatoin", "2-month training + 1-month test/application" and "1-month training + 1-month test/application" rolling horizon strategies is acceptable in practice.

By contrasting this finding and the good performance of ML models in Sections 8.3.2 and 8.3.3, for industry applications, we recommend using the sensor data of the most recent 5 months for model training, and the trained model can be applied in the coming one month, if a rolling horizon approach is adopted. Due to the unavailability of more sensor data, we cannot conduct more comparison experiments regarding the rolling horizon approach and make the recommended strategy more underpinned by additional experimental results. This is a limitation of this study.

3-month training + 1-month application/test

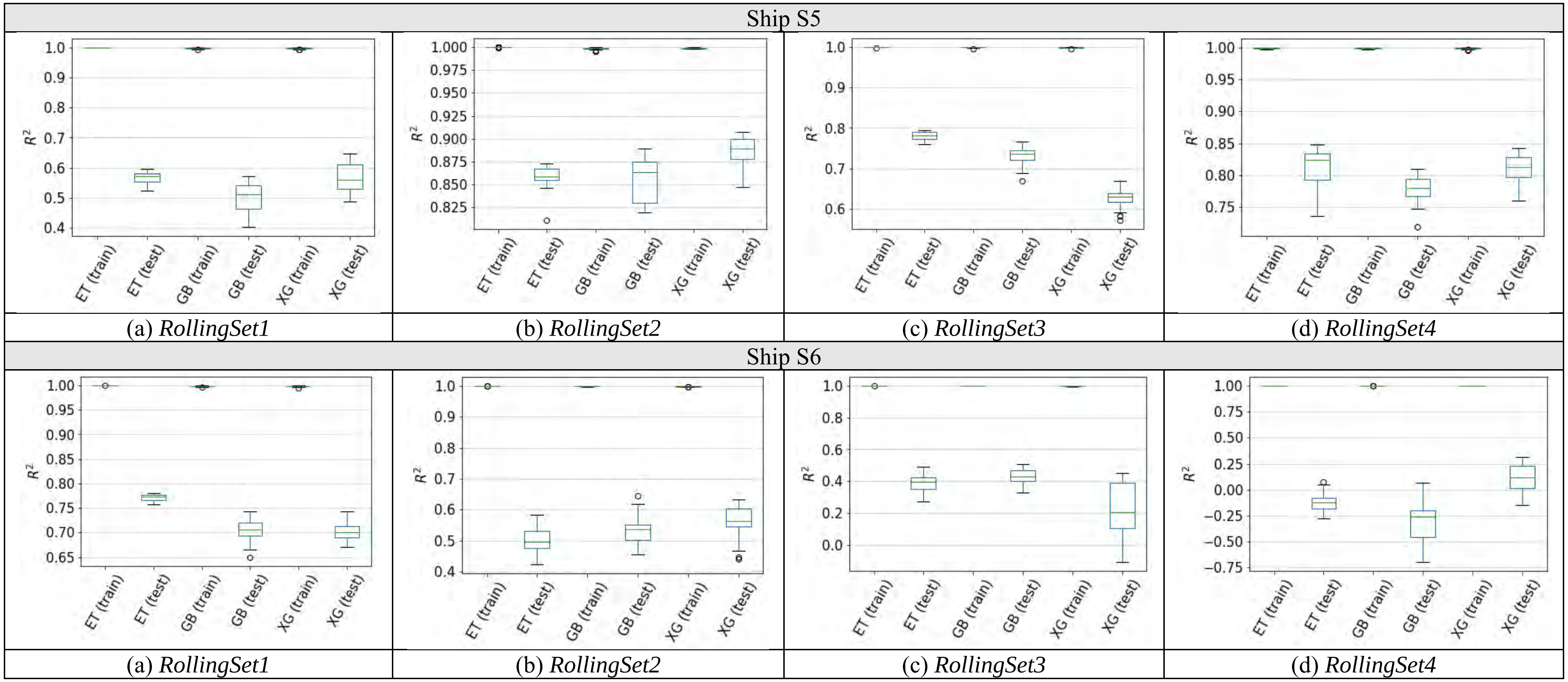


Figure 23. Fit performance (R^2) of models in a "3-month training + 1-month application/test" application scenario. "ET (train)" means the fit performance of ET on train set. "ET (test)" means the fit performance of ET on test/application set.

IAMU International Association of Maritime Universities (AMC)

8.4 Summary

This study fuses sensor data and meteorological data for the purpose of improving the accuracy of ML models that quantify ship fuel consumption rate based on sailing speed, draft, trim, weather conditions, and sea conditions. The best dataset found, *Sensor2*, reveals the benefits of combing sensor data and meteorological data. Experiments with two 9,200-TEU containerships show that compared to using sensor data as the sole data source, fusing sensor data and meteorological data will improve the fit performance of all ML models. The best ML models found are consistent with our previous studies in Sections 6 and 7, including ET, GB and XG. Given the best dataset *Sensor2*, their R^2 values over the training set are 0.999 or 1.000, and their R^2 values over the test set are all above 0.966. Their fit errors with RMSE values are below 0.75 ton/day, and with MAE below 0.52 ton/day. These promising results are well beyond the requirements of most industry applications for ship fuel efficiency analysis. We also verify the applicability of the selected datasets and ML models in a rolling horizon approach, and conjecture that a rolling horizon strategy of "5-month training + 1-month test/applicatoin" could work well in practice and sensor data of less than five months could be insufficient to train ML models.

In the data fusion approach proposed by this study, the information from sensor data about *timestamp*, *geographical positions*, and *ship heading* plays a critical role in retrieving weather and sea condition information from meteorological data. In Section 7 with DFS2, the information from AIS data about *timestamp*, *geographical positions*, and *ship heading* plays the same role. One may claim that DFS2 in Section 7 can substitute sensor data for AIS data. This is correct in principle, but we should realize that not every containership has a good sensor system. However, AIS data is always available for shipping companies.

9. Conclusions and Recommendations

9.1 Conclusions

With promotions of IMO and governmental organizations, the shipping industry has been implementing operational measures to save bunker fuel and mitigate emissions from ships, including sailing speed optimization, trim optimization, weather routing, and the virtual arrival policy. Many frustrations have been emerging during the process of implementation of these measures. These frustrations are boiled down, if not fully, to how we can quantify the synergetic contributions of many factors (speed, draft/displacement, trim, weather conditions, sea conditions) on a ship's bunker fuel consumption rate. A latest review paper, Yan et al. (2021), points out that the basis of all operational measures for ship bunker fuel savings and emission mitigation is quantitatively modeling the relationship between fuel consumption rate and many determinants, including sailing speed, draft/displacement, trim, weather conditions, and sea conditions. This project addresses this theoretical challenge that restricts the implementation of energy-efficient operational measures by investigating the complementary roles of different data sources available to a shipping company, fusing these data sources, and employing state-of-the-art machine learning techniques.

We collected **voyage report data** and **sensor data** of eight 8,100-TEU to 14,000-TEU containerships from a global shipping company, purchased the **AIS data** of these ships from *MarineTraffic* with the financial support of IAMU, and downloaded **meteorological data** from European Centre for Medium-range Weather Forecasts (ECMWF) and Copernicus Marine Service (CMEMS). Based on the information contained in these four data sources, we designated **three data fusion solutions**: **DFS1** fuses voyage report data and meteorological data, by considering the inaccurate information of weather and sea conditions recorded by voyage report; **DFS2** further fuses AIS data into voyage report data and meteorological data because AIS data helps find the actual sailing trajectory of the ship and thus retrieve more accurate information of weather and sea conditions from meteorological data; **DFS3** approaches sensor data as the main data source of a ship's fuel consumption rate, and overcomes the limitation of sensor data by taking advantage of the complete information of weather and sea conditions contained in meteorological data. **For each of the data fusion solutions, eight to nine datasets are constructed**.

Over these datasets from three data fusion solutions, a large range of widely adopted machine learning models were experimented with, including decision tree-based models, artificial neural network (ANN), support vector machine (SVM), ridge regression (Ridge), and LASSO. Tree-based models include the basic decision tree (DT) model and models produced by two ensemble strategies: Extremely randomized trees (ET) and random forest (RF) from the bagging ensemble strategy; AdaBoost (AB), gradient tree boosting (GB), XGBoost (XG), and LightGBM (LB) from the boosting ensemble strategy. During the experiments with these machine learning models, the impacts of data normalization, hyperparameter optimization, and the randomness in splitting training sets and test sets are well addressed.

Extensive experiments were conducted to answer three research questions regarding the choice of datasets from three data fusion solutions and the choice of machine learning models. A voting scheme is developed to break down the impacts of dataset choice and model choice. When dataset choice is considered, the original voyage report dataset ***Set1*** has a decent quality for ship fuel efficiency modeling; if more effort is paid to fuse voyage report data and meteorological data, data quality improves slightly and $\boldsymbol{Set3_{precise}}$ can be adopted. When AIS data is available, further including AIS data might also be beneficial, which suggests the adoption of the dataset $\boldsymbol{AIS5_{precise.}}$ Overall, the best datasets found with DFS1 and DFS2, including ***Set1***, $\boldsymbol{Set3_{precise}}$, and $\boldsymbol{AIS5_{precise}}$, ensure accurate fit performances of best ML models: R^2 on the training set is above 0.96 and even reaches 0.99 to 1.00, and R^2 on the test set is between 0.74 and 0.90; the fit errors measured by RMSE and MAE are between 0.5 and 4.5 ton/day. When sensor data, rather than voyage report data, is used as the main data source of ship bunker fuel consumption analysis, it will elevate the modeling accuracy to a higher level, possibly the highest level if meteorological data is fused in. With DFS3, given the best dataset ***Sensor2***, best ML models achieve their R^2 values of over the training set at 0.999 or 1.000, and their R^2 values over the test set are all above 0.966. Their fit errors with RMSE values are below 0.75 ton/day, and with MAE below 0.52 ton/day.

As far as ML model choice is concerned, we recommend the installation of four decision-tree based models including ET, AB, GB, and XG because they usually possess the highest fit performance and good generalization performance. Their performances are also quite robust against random splits of a dataset into training and test sets. Our experiments with DFS1, DFS2, and DFS2 reach consistent findings about the performances of ML models and rank their performances into four tiers.

- Tier 1: ET, AB, GB, and XG.
- Tier 2: RF, LB
- Tier 3: DT, SVM, ANN
- Tier 4: Ridge, LASSO.

9.2 Recommendations for Industry Applications

Voyage report data, meteorological data, sensor data, and AIS data are the major data sources that can be utilized by a shipping company and other industry stakeholders for ship energy efficiency analysis. These four data sources have different but complementary information for ship fuel efficiency analysis. The first insight delivered by this project is fusing these data sources is usually beneficial in terms of accuracy of ship fuel efficiency modeling.

Regarding dataset selection, voyage report data is usually sufficient for many industry applications based on ship energy/emission analysis. Fusing meteorological data into voyage report generally slightly improve the accuracy of ship fuel efficiency models. Furthermore, if AIS data is available, it describes the sailing trajectory of the ship and thus helps find more accurate information of weather and sea conditions the ship sailed through. Substituting this accurate information for the snapshotted weather and sea condition data in voyage report generally improves the performance of ship fuel efficiency models based on voyage report.

However, there is no guarantee that fusing AIS data and meteorological data into voyage report data must improve the performance of ship fuel efficiency analysis model. This can be explained by the fact that the snapshotted information of weather and sea conditions in voyage report data may have been representative, though to an unknown extent, and by the fact that accurate weather and sea condition data on the waypoints does not necessarily lead to a more accurate estimation of daily average weather and sea conditions the ship sailed through. This is a finding that might contract the imagination and intuition of industry professionals.

The reported fit and generalization performances of ET, AB, GB and XG (summarized in Section 7.4) are probably the highest level of accuracy we could achieve to model a mega containership's fuel consumption rate, if voyage report data is used as the main source of bunker fuel consumption. The main reasons why it is difficult, if not impossible, to further improve the modeling accuracy boil down to the fact that voyage report data reports the "daily" bunker fuel consumption of a ship, and this data granularity ("daily") restricts the model performance. However, this modelling accuracy is sufficient for many industry applications and energy-efficient operational measures for shipping companies, including sailing speed optimization, weather routing, and virtual arrivals.

The limitation of voyage report data can be overcome by sensor data. When sensor data is used as the main data source of ship bunker fuel efficiency analysis, it elevates the modeling accuracy to a higher level, possibly the highest level if meteorological data is fused in. With DFS3, given the best dataset ***Sensor2***, best ML models achieve their R^2 values of over the training set at 0.999 or 1.000, and their R^2 values over the test set are all above 0.966. Their fit errors with RMSE values are below 0.75 ton/day, and with MAE below 0.52 ton/day. The highly accurate performance of ML models with sensor data and meteorological data justifies their application in trim optimization. Our discussion with industry professionals conveys the information that it would be hard to imagine sensor data with the resolution of 15 minutes can be used in speed optimization, weather routing, and virtual arrivals, if not impossible. ML models with sensor data and meteorological data can be trained and utilized in a rolling-horizon approach by considering the large quality of senor data. Specifically, we recommend the shipping companies to employ the latest five-month sensor data and meteorological data to train the models and update the models on a monthly basis.

As far as ML model choice is concerned, we recommend the installation of four decision-tree based models including ET, AB, GB, and XG because they usually possess the highest fit performance and good generalization performance. Our experiments with DFS1, DFS2, and DFS2 reach consistent findings about the performances of ML models and rank their performances into four tiers. It would be safe for industry applications to only consider the ML models in Tier 1, including ET, AB, GB, and XG.

We summarize our recommendations on industry applications in Table 19. The importance of operational energy-saving measures can be ranked in the following

$$\text{Speed optimization and virtual arrival} \succ \text{weather routing} \succ \text{trim optimization.}$$

Our quantitative results in Sections 6.3.5 and 8.3.4 about the relative importance of determinants of ship fuel efficiency rate show that sailing speed is the first most important factor, and weather and sea conditions as a whole represent the second most significant factor. The impact of trim is the minor but nonnegligible. The impact of displacement/draft because of cargo load and ballast water can be ignored in reality, because it is well dominated by that of weather and sea conditions and largely outweighed by the profit margin of carrying more cargoes.

Apart from the applications in Table 19, we also recommend that industry stakeholders make more efforts on software/system development, based on our findings and Python code infrastructure published in GitHub website. Regulators such as IMO and EU can also improve their understanding of ship fuel/emission efficiency through the findings in this data-driven project.

Table 19. Summary of recommendations for industry applications

Industry applications	• Sailing speed optimization • Weather routing • Virtual (just-in-time) arrival	Trim optimization
Industry stakeholders	• Shipping companies • Weather information service providers (WISPs) • Ship classification societies (such as ClassNK) • Shipping associations (such as BIMCO)	Shipping companies
Recommended data sources and datasets	• DFS1: Voyage report data + meteorological data • DFS2: Voyage report data + meteorological data + AIS data	DFS3: sensor data + meteorological data
Recommended models	Extremely randomized trees (ETs), Gradient tree boosting (GB), or XGBoost (XG)	Extremely randomized trees (ETs), Gradient tree boosting (GB), or XGBoost (XG)

9.3 Limitations and Future Studies

The studies of this project have the following limitations which can be addressed in future studies.

- This project only considers containerships but not other ship types such as bulk ships and oil tankers.
- This project models the fuel consumption rate as the output of a ML model. If engine RPM is utilized as the output feature, the built ML models will enable the ship captain to determine required engine RPM when sailing the ship with different speed requirements, displacement, and weather and sea conditions.
- ANN is considered as the main approach of deep learning and has many variants with different model structures. This project only considers the traditional three-layer feedforward network. Future studies can investigate the performance of other variants of ANN.
- In Section 8.3.5, due to the unavailability of more sensor data, we could not conduct more comparison experiments regarding the rolling horizon approach and made the recommended strategy more underpinned by additional experimental results.

- The data fusion approach of DFS3 in Section 8 depends on the structure of sensor data. If the sensor data from another shipping company is used and the structure of their sensor data is different from this study, the data fusion approach needs to be revised. For instance, we also made a communication with another global shipping company, and their sample sensor data shared to us has a different structure. Their sensor data contains features about wind, waves, and sea currents, but their information on wind is less complete than the sensor data used in this study. Neither of their sensor data and the sensor data used in this study includes the information of wave period and sea water temperature.
- This project is driven by the data collected from manned mega containerships. During the design of an autonomous ship, the designer will have to develop automated algorithms for the ship's sailing speed optimization and weather routing. In future studies, more sensor data can be collected from autonamous ships which generally have much smaller sizes than the mega containerships considered by this project. Accordingly, it would be interesting to develop data fusion apporaches for these data and experiment on the performance of ML models.

References

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.

Adland, R., Cariou, P., & Wolff, F.C. (2019). When energy efficiency is secondary: The case of Offshore Support Vessels. *Transportation Research Part D: Transport and Environment*, *72*, 114-126.

Adland, R., Cariou, P., & Wolff, F. C. (2020). Optimal ship speed and the cubic law revisited: Empirical evidence from an oil tanker fleet. *Transportation Research Part E: Logistics and Transportation Review*, *140*, 101972.

Beşikçi, E. B., Arslan, O., Turan, O., & Ölçer, A.I. (2016). An artificial neural network based decision support system for energy efficient ship operations. *Computers & Operations Research,* 66, 393-401.

Bennett, G.G. (1996). Practical rhumb line calculations on the spheroid. *Journal of Navigation,* 49(1), 112-119.

Bergstra, J., Yamins, D., & Cox, D.D (2013). Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. In *Proceedings of the 30th International Conference on Machine Learning (ICML 2013)* (pp. 115-123).

Bocchetti, D., Lepore, A., Palumbo, B., & Vitiello, L. (2015). A statistical approach to ship fuel consumption monitoring. *Journal of Ship Research*, *59*(03), 162-171.

Boser B.E., Guyon I.M., & Vapnik V.N. (1992). A training algorithm for optimal margin classifiers. In: Haussler D. (Ed.), In *Proceedings of the Annual Conference on Computational Learning Theory* (pp. 144-152). ACM Press, Pittsburgh, PA.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees*. Wadsworth, Belmont, CA.

Cai, B., Xu, L.F., & Fu, F. (2019). Shear Resistance Prediction of post-fire reinforced concrete beams using artificial neural network. *International Journal of Concrete Structures and Materials,* 13 (1), 1-13.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).

C-MAP. (2022). VVOS: Vessel and Voyage Optimization Solution. https://www.oceanweather.com/forecast/VVOS/index.html. Access on 19 March 2022.

Coraddu, A., Oneto, L., Baldi, F., & Anguita, D. (2017). Vessels fuel consumption forecast and trim optimisation: A data analytics perspective. *Ocean Engineering*, 130, 351-370.

Drucker, H. (1997). Improving regressors using boosting techniques. In *Proceedings of the Fourteenth International Conference on Machine Learning* (pp. 107-115).

Du, Y., Meng, Q., Wang, S., & Kuang, H. (2019). Two-phase optimal solutions for ship speed and trim optimization over a voyage using voyage report data. *Transportation Research Part B: Methodological*, *122*, 88-114.

European Union (EU). (2021). Monitoring, reporting and verification of EU ETS emissions. https://ec.europa.eu/clima/eu-action/eu-emissions-trading-system-eu-ets/monitoring-reporting-and-verification-eu-ets-emissions_en. Accessed on 19 March 2022.

Farag, Y. B., & Ölçer, A. I. (2020). The development of a ship performance model in varying operating conditions based on ANN and regression techniques. *Ocean Engineering*, *198*, 106972.

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119-139.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232.

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*(1), 3-42.

Gkerekos, C., Lazakis, I., & Theotokatos, G. (2019). Machine learning models for predicting ship main engine Fuel Oil Consumption: A comparative study. *Ocean Engineering*, *188*, 106282.

Haykin, S.O. (2008). *Neural Networks and Learning Machines*. The third edition. Prentice Hall: New Jersey.

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. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 10-Sep-2021), 10.24381/cds.adbb2d47.

Hoerl, A.E., & Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics,* 12 (1), 55-67.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5), 359-366.

IMO. (2012). Guideline for Development of a Ship Energy Efficiency Management Plan (SEEMP), MEPC, 213 (63) Annex 9.

IMO. (2020). Fourth IMO GHG study 2020. https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/Fourth%20IMO%20GHG%20Study%202020%20Executive-Summary.pdf. Accessed on 19 March 2022.

Johnson, H., & Andersson, K. (2011). The energy efficiency gap in shipping-barriers to improvement. In *International Association of Maritime Economists (IAME) Conference*.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, 3146-3154.

Kolmogorov, A. N. (1957). On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Doklady Akademii Nauk,* 114(5), 953-956.

Lee, H., Aydin, N., Choi, Y., Lekhavat, S., & Irani, Z. (2018). A decision support system for vessel speed decision in maritime logistics using weather archive big data. *Computers & Operations Research*, *98*, 330-342.

Lepore, A., dos Reis, M.S., Palumbo, B., Rendall, R., & Capezza, C. (2017). A comparison of advanced regression techniques for predicting ship CO2 emissions. *Quality and Reliability Engineering International,* 33, 1281-1292.

Li, X., Sun, B., Zhao, Q., Li, Y., Shen, Z., Du, W., & Xu, N. (2018). Model of speed optimization of oil tanker with irregular winds and waves for given route. *Ocean Engineering,* 164, 628-639.

Pedersen, B.P., & Larsen, J. (2009). Prediction of full-scale propulsion power using artificial neural networks. In *COMPIT'09: 8th International Conference on Computer and IT Applications in the Maritime Industries* (pp. 10-12), Budapest.

Rehmatulla, N., Parker, S., Smith, T., & Stulgis, V. (2017). Wind technologies: Opportunities and barriers to a low carbon shipping industry. *Marine Policy*, *75*, 217-226.

Rio, M. H., Mulet, S., & Picot, N. (2014). Beyond GOCE for the ocean circulation estimate: Synergetic use of altimetry, gravimetry, and in situ data provides new insight into geostrophic and Ekman currents. *Geophysical Research Letters*, 41(24), 8918-8925.

Man, Y., Sturm, T., Lundh, M., & MacKinnon, S. N. (2020). From ethnographic research to big data analytics—a case of maritime energy-efficiency optimization. *Applied Sciences*, *10*(6), 2134.

Meng, Q., Zhang, Y., & Xu, M. (2017). Viability of transarctic shipping routes: a literature review from the navigational and commercial perspectives. *Maritime Policy & Management*, *44*(1), 16-41.

Merkel, A., Kalantari, J., & Mubder, A. (2022). Port call optimization and CO2-emissions savings–Estimating feasible potential in tramp shipping. *Maritime Transport Research*, *3*, 100054.

Smola, A.J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14 (3), 199-222.

Soner, O., Akyuz, E., & Celik, M. (2018). Use of tree based methods in ship performance monitoring under operating conditions. *Ocean Engineering*, *166*, 302-310.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Uyanık, T., Karatuğ, Ç., & Arslanoğlu, Y. (2020). Machine learning approach to ship fuel consumption: A case of container vessel. *Transportation Research Part D: Transport and Environment*, *84*, 102389.

Wan, Z., El Makhloufi, A., Chen, Y., & Tang, J. (2018). Decarbonizing the international shipping industry: Solutions and policy recommendations. *Marine Pollution Bulletin*, *126*, 428-435.

Wang, S., Psaraftis, H.N., & Qi, J. (2021). Paradox of international maritime organization's carbon intensity indicator. *Communications in Transportation Research*, *1*, 100005.

Weintrit, A., & Kopacz, P. (2011). A novel approach to loxodrome (rhumb-line), orthodrome (great circle) and geodesic line in ECDIS and navigation in general. *TransNav-International Journal on Marine Navigation and Safety of Sea Transportation,* 5(4), 507-517.

Yan, R., Wang, S., & Psaraftis, H.N. (2021). Data analytics for fuel consumption management in maritime transportation: Status and perspectives. *Transportation Research Part E: Logistics and Transportation Review*, *155*, 102489.

Yao, Z., Ng, S. H., & Lee, L. H. (2012). A study on bunker fuel management for the shipping liner services. *Computers & Operations Research*, *39*(5), 1160-1172.

Appendix

Table A1. The fit performance of eleven machine learning models for ship S2 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.833	0.668	113.854	10.580	7.934	8.951
	Set2$_{precise}$	0.820	0.591	113.281	10.459	7.954	9.321
	Set2$_{fuzzy}$	0.871	0.612	80.754	8.724	6.480	7.612
	Set3$_{precise}$	0.820	0.589	112.089	10.461	7.916	9.230
	Set3$_{fuzzy}$	0.819	0.575	112.765	10.428	7.896	9.219
	Set4$_{precise}$	0.808	0.595	120.097	10.818	8.149	9.543
	Set4$_{fuzzy}$	0.814	0.591	116.912	10.691	8.068	9.324
	Set5$_{precise}$	0.823	0.615	110.287	10.266	7.739	9.008
	Set5$_{fuzzy}$	0.833	0.596	103.989	9.909	7.434	8.724
ET	*Set1*	0.971	0.786	19.857	4.055	2.986	3.306
	Set2$_{precise}$	0.960	0.755	24.360	4.399	3.253	3.686
	Set2$_{fuzzy}$	0.958	0.757	25.878	4.553	3.366	3.839
	Set3$_{precise}$	0.974	0.765	15.842	3.377	2.445	2.780
	Set3$_{fuzzy}$	0.970	0.763	18.735	3.789	2.711	3.086
	Set4$_{precise}$	0.977	0.764	14.537	3.128	2.237	2.530
	Set4$_{fuzzy}$	0.966	0.753	20.670	3.685	2.710	3.126
	Set5$_{precise}$	0.962	0.761	23.530	4.202	3.100	3.525
	Set5$_{fuzzy}$	0.973	0.759	16.740	3.553	2.608	2.959
RF	*Set1*	0.959	0.766	27.622	5.205	3.750	4.227
	Set2$_{precise}$	0.953	0.739	29.359	5.350	3.843	4.436
	Set2$_{fuzzy}$	0.957	0.744	26.791	5.118	3.763	4.381
	Set3$_{precise}$	0.950	0.740	31.494	5.541	4.007	4.662
	Set3$_{fuzzy}$	0.946	0.743	33.716	5.699	4.096	4.743
	Set4$_{precise}$	0.957	0.740	26.572	5.118	3.734	4.336
	Set4$_{fuzzy}$	0.947	0.743	33.116	5.695	4.054	4.702
	Set5$_{precise}$	0.953	0.739	29.568	5.382	3.900	4.537
	Set5$_{fuzzy}$	0.955	0.750	28.280	5.224	3.773	4.393
AB	*Set1*	0.968	0.762	21.779	4.305	3.609	4.143
	Set2$_{precise}$	0.964	0.732	22.762	4.602	3.941	4.548
	Set2$_{fuzzy}$	0.959	0.732	25.577	4.816	4.018	4.635
	Set3$_{precise}$	0.961	0.743	24.755	4.778	4.073	4.729
	Set3$_{fuzzy}$	0.962	0.739	23.645	4.672	3.976	4.617
	Set4$_{precise}$	0.978	0.739	13.828	3.528	2.989	3.464
	Set4$_{fuzzy}$	0.972	0.732	17.180	3.860	3.257	3.764
	Set5$_{precise}$	0.970	0.735	18.789	4.114	3.503	4.068
	Set5$_{fuzzy}$	0.972	0.737	17.302	3.979	3.369	3.948
GB	*Set1*	0.964	0.781	24.457	4.564	3.429	3.793
	Set2$_{precise}$	0.984	0.756	10.221	2.615	1.849	2.072
	Set2$_{fuzzy}$	0.987	0.750	8.325	2.190	1.584	1.767
	Set3$_{precise}$	0.992	0.760	5.008	1.817	1.234	1.378
	Set3$_{fuzzy}$	0.985	0.762	9.472	2.472	1.716	1.938
	Set4$_{precise}$	0.990	0.763	6.143	1.963	1.364	1.529
	Set4$_{fuzzy}$	0.990	0.755	6.254	1.885	1.401	1.570
	Set5$_{precise}$	0.975	0.747	15.364	3.270	2.336	2.634
	Set5$_{fuzzy}$	0.976	0.756	14.869	3.182	2.389	2.720
XG	*Set1*	0.975	0.781	16.733	3.503	2.631	2.868
	Set2$_{precise}$	0.985	0.759	9.566	2.542	1.691	1.837
	Set2$_{fuzzy}$	0.990	0.755	6.058	2.038	1.246	1.328
	Set3$_{precise}$	0.991	0.765	5.421	1.949	1.186	1.277
	Set3$_{fuzzy}$	0.982	0.759	11.047	3.003	1.792	1.921

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	$Set4_{precise}$	0.984	0.770	10.314	2.587	1.638	1.760
	$Set4_{fuzzy}$	0.988	0.761	7.480	2.197	1.362	1.457
	$Set5_{precise}$	0.967	0.754	21.190	4.015	2.689	2.956
	$Set5_{fuzzy}$	0.977	0.755	14.591	3.195	2.170	2.360
LB	$Set1$	0.946	0.761	36.850	5.784	4.429	4.834
	$Set2_{precise}$	0.981	0.736	11.640	2.940	2.158	2.384
	$Set2_{fuzzy}$	0.981	0.727	12.277	3.145	2.224	2.455
	$Set3_{precise}$	0.980	0.748	12.589	3.053	2.179	2.442
	$Set3_{fuzzy}$	0.982	0.741	11.425	3.003	2.145	2.400
	$Set4_{precise}$	0.975	0.753	15.660	3.469	2.582	2.863
	$Set4_{fuzzy}$	0.976	0.737	14.911	3.522	2.568	2.839
	$Set5_{precise}$	0.974	0.724	16.238	3.488	2.447	2.769
	$Set5_{fuzzy}$	0.971	0.731	18.560	3.787	2.594	2.943
SVM	$Set1$	0.848	0.797	103.306	10.147	7.260	7.779
	$Set2_{precise}$	0.868	0.812	82.693	9.063	6.443	7.021
	$Set2_{fuzzy}$	0.868	0.802	82.818	9.066	6.404	7.072
	$Set3_{precise}$	0.864	0.812	84.860	9.176	6.608	7.210
	$Set3_{fuzzy}$	0.871	0.799	81.178	8.918	6.364	7.014
	$Set4_{precise}$	0.870	0.814	81.122	8.974	6.442	7.034
	$Set4_{fuzzy}$	0.870	0.808	81.347	8.990	6.402	7.062
	$Set5_{precise}$	0.859	0.807	88.173	9.339	6.736	7.391
	$Set5_{fuzzy}$	0.867	0.795	83.747	9.040	6.466	7.202
ANN	$Set1$	0.876	0.787	84.367	9.093	6.935	7.682
	$Set2_{precise}$	0.907	0.800	57.855	7.489	5.695	6.295
	$Set2_{fuzzy}$	0.897	0.789	64.406	7.958	6.055	6.742
	$Set3_{precise}$	0.908	0.791	56.693	7.365	5.581	6.171
	$Set3_{fuzzy}$	0.893	0.803	67.203	8.110	6.127	6.837
	$Set4_{precise}$	0.908	0.803	56.949	7.439	5.643	6.227
	$Set4_{fuzzy}$	0.892	0.805	67.986	8.160	6.176	6.905
	$Set5_{precise}$	0.909	0.798	56.575	7.417	5.624	6.182
	$Set5_{fuzzy}$	0.893	0.787	66.436	8.051	6.101	6.783
Ridge	$Set1$	0.822	0.786	121.419	11.016	8.454	9.312
	$Set2_{precise}$	0.820	0.801	112.767	10.614	8.029	9.059
	$Set2_{fuzzy}$	0.813	0.791	116.890	10.806	8.083	9.148
	$Set3_{precise}$	0.826	0.802	108.847	10.429	8.011	9.055
	$Set3_{fuzzy}$	0.823	0.798	110.559	10.511	8.004	9.066
	$Set4_{precise}$	0.825	0.803	109.502	10.460	8.011	9.081
	$Set4_{fuzzy}$	0.821	0.798	112.098	10.584	8.023	9.111
	$Set5_{precise}$	0.821	0.804	112.183	10.587	8.050	9.089
	$Set5_{fuzzy}$	0.815	0.796	115.730	10.753	8.099	9.178
LASSO	$Set1$	0.822	0.785	121.508	11.020	8.471	9.331
	$Set2_{precise}$	0.819	0.798	113.218	10.635	8.023	9.028
	$Set2_{fuzzy}$	0.811	0.786	117.999	10.857	8.090	9.146
	$Set3_{precise}$	0.824	0.796	110.162	10.492	8.034	9.042
	$Set3_{fuzzy}$	0.821	0.797	112.336	10.595	8.043	9.099
	$Set4_{precise}$	0.824	0.800	110.115	10.490	8.007	9.039
	$Set4_{fuzzy}$	0.820	0.795	112.895	10.621	8.032	9.095
	$Set5_{precise}$	0.820	0.801	112.578	10.606	8.047	9.060
	$Set5_{fuzzy}$	0.815	0.796	115.774	10.755	8.086	9.150

Table A2. The fit performance of eleven machine learning models for ship S3 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.857	0.684	105.672	10.125	7.259	8.643
	$Set2_{precise}$	0.853	0.713	107.107	10.167	7.422	8.762
	$Set2_{fuzzy}$	0.845	0.700	112.697	10.432	7.586	9.080
	$Set3_{precise}$	0.865	0.684	98.572	9.705	7.042	8.343
	$Set3_{fuzzy}$	0.868	0.692	95.656	9.586	6.903	8.258
	$Set4_{precise}$	0.864	0.694	99.016	9.741	7.126	8.471
	$Set4_{fuzzy}$	0.874	0.678	90.962	9.304	6.662	7.963
	$Set5_{precise}$	0.870	0.701	94.794	9.586	6.956	8.295
	$Set5_{fuzzy}$	0.858	0.695	102.704	9.931	7.215	8.621
ET	*Set1*	0.977	0.800	17.021	3.911	2.462	2.964
	$Set2_{precise}$	0.973	0.821	19.352	3.820	2.270	2.890
	$Set2_{fuzzy}$	0.969	0.820	22.459	4.479	2.719	3.433
	$Set3_{precise}$	0.985	0.821	10.758	2.846	1.716	2.181
	$Set3_{fuzzy}$	0.986	0.818	10.342	2.391	1.438	1.823
	$Set4_{precise}$	0.975	0.821	18.085	3.943	2.304	2.928
	$Set4_{fuzzy}$	0.976	0.819	17.269	3.827	2.328	2.940
	$Set5_{precise}$	0.984	0.830	11.712	2.940	1.661	2.141
	$Set5_{fuzzy}$	0.985	0.824	11.150	2.920	1.767	2.250
RF	*Set1*	0.960	0.768	29.573	5.369	3.497	4.234
	$Set2_{precise}$	0.959	0.809	29.986	5.388	3.453	4.290
	$Set2_{fuzzy}$	0.963	0.801	27.032	5.170	3.369	4.169
	$Set3_{precise}$	0.956	0.802	31.781	5.576	3.587	4.463
	$Set3_{fuzzy}$	0.952	0.805	34.613	5.786	3.654	4.563
	$Set4_{precise}$	0.959	0.804	29.892	5.406	3.491	4.325
	$Set4_{fuzzy}$	0.952	0.802	34.739	5.778	3.694	4.603
	$Set5_{precise}$	0.958	0.812	30.199	5.441	3.473	4.321
	$Set5_{fuzzy}$	0.959	0.806	29.588	5.388	3.409	4.241
AB	*Set1*	0.988	0.798	9.177	2.942	2.371	2.718
	$Set2_{precise}$	0.986	0.810	10.039	2.915	2.240	2.541
	$Set2_{fuzzy}$	0.984	0.805	11.278	3.202	2.508	2.796
	$Set3_{precise}$	0.991	0.812	6.328	2.183	1.712	1.998
	$Set3_{fuzzy}$	0.992	0.807	5.598	2.100	1.542	1.789
	$Set4_{precise}$	0.995	0.814	3.875	1.811	1.356	1.550
	$Set4_{fuzzy}$	0.992	0.801	5.657	2.166	1.647	1.869
	$Set5_{precise}$	0.994	0.813	4.175	1.738	1.313	1.514
	$Set5_{fuzzy}$	0.996	0.797	3.048	1.491	1.070	1.230
GB	*Set1*	0.962	0.776	28.220	4.726	3.221	3.841
	$Set2_{precise}$	0.968	0.814	23.569	4.167	2.747	3.412
	$Set2_{fuzzy}$	0.948	0.813	37.767	5.792	3.874	4.811
	$Set3_{precise}$	0.964	0.819	26.559	4.694	2.836	3.642
	$Set3_{fuzzy}$	0.968	0.812	23.652	4.569	2.741	3.501
	$Set4_{precise}$	0.956	0.818	32.004	5.489	3.562	4.440
	$Set4_{fuzzy}$	0.956	0.810	32.023	5.280	3.217	4.123
	$Set5_{precise}$	0.976	0.819	17.731	3.499	2.207	2.765
	$Set5_{fuzzy}$	0.969	0.816	22.829	4.478	3.030	3.732
XG	*Set1*	0.959	0.778	30.013	4.738	3.214	3.744
	$Set2_{precise}$	0.941	0.811	42.884	6.013	3.809	4.801
	$Set2_{fuzzy}$	0.950	0.799	36.417	5.386	3.565	4.427
	$Set3_{precise}$	0.961	0.810	28.714	5.030	3.052	3.828
	$Set3_{fuzzy}$	0.962	0.809	27.987	4.971	3.024	3.802

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	$Set4_{precise}$	0.952	0.816	34.801	5.579	3.556	4.392
	$Set4_{fuzzy}$	0.949	0.807	37.325	5.970	3.839	4.756
	$Set5_{precise}$	0.951	0.816	35.428	5.466	3.558	4.432
	$Set5_{fuzzy}$	0.953	0.808	34.093	5.267	3.457	4.250
LB	$Set1$	0.935	0.766	48.608	6.560	4.506	5.448
	$Set2_{precise}$	0.946	0.809	39.482	5.955	3.895	4.905
	$Set2_{fuzzy}$	0.921	0.799	57.989	7.204	4.865	6.100
	$Set3_{precise}$	0.947	0.804	38.795	5.845	3.853	4.853
	$Set3_{fuzzy}$	0.952	0.801	35.030	5.577	3.723	4.654
	$Set4_{precise}$	0.951	0.805	35.768	5.641	3.702	4.652
	$Set4_{fuzzy}$	0.935	0.803	47.096	6.576	4.433	5.520
	$Set5_{precise}$	0.963	0.808	26.595	4.681	2.977	3.758
	$Set5_{fuzzy}$	0.939	0.798	43.912	6.098	4.029	5.143
SVM	$Set1$	0.812	0.791	138.669	11.753	7.557	8.957
	$Set2_{precise}$	0.837	0.823	117.826	10.819	6.698	8.237
	$Set2_{fuzzy}$	0.830	0.818	123.105	11.072	6.843	8.436
	$Set3_{precise}$	0.844	0.820	113.000	10.591	6.627	8.167
	$Set3_{fuzzy}$	0.847	0.817	111.184	10.490	6.551	8.120
	$Set4_{precise}$	0.844	0.822	113.090	10.598	6.624	8.169
	$Set4_{fuzzy}$	0.843	0.821	113.745	10.618	6.643	8.199
	$Set5_{precise}$	0.840	0.823	115.818	10.727	6.701	8.227
	$Set5_{fuzzy}$	0.833	0.821	121.038	10.973	6.888	8.472
ANN	$Set1$	0.829	0.780	126.769	11.217	7.780	9.353
	$Set2_{precise}$	0.865	0.809	98.002	9.850	6.647	8.181
	$Set2_{fuzzy}$	0.859	0.807	102.467	10.070	6.665	8.284
	$Set3_{precise}$	0.874	0.798	91.583	9.475	6.480	7.992
	$Set3_{fuzzy}$	0.859	0.796	101.857	10.026	6.907	8.541
	$Set4_{precise}$	0.861	0.800	100.566	9.972	6.828	8.394
	$Set4_{fuzzy}$	0.848	0.796	110.236	10.430	7.158	8.877
	$Set5_{precise}$	0.865	0.809	97.761	9.821	6.634	8.116
	$Set5_{fuzzy}$	0.853	0.801	106.070	10.237	6.811	8.423
Ridge	$Set1$	0.780	0.778	162.676	12.739	9.007	11.114
	$Set2_{precise}$	0.792	0.799	150.342	12.247	8.523	10.908
	$Set2_{fuzzy}$	0.790	0.798	151.889	12.310	8.513	10.899
	$Set3_{precise}$	0.801	0.796	144.061	11.987	8.329	10.615
	$Set3_{fuzzy}$	0.801	0.797	143.736	11.974	8.307	10.617
	$Set4_{precise}$	0.798	0.797	146.525	12.089	8.366	10.721
	$Set4_{fuzzy}$	0.798	0.798	145.947	12.066	8.335	10.674
	$Set5_{precise}$	0.795	0.803	148.003	12.151	8.436	10.765
	$Set5_{fuzzy}$	0.793	0.802	149.528	12.214	8.459	10.815
LASSO	$Set1$	0.779	0.778	163.445	12.769	9.011	11.128
	$Set2_{precise}$	0.792	0.798	150.396	12.249	8.514	10.895
	$Set2_{fuzzy}$	0.790	0.797	151.947	12.313	8.502	10.883
	$Set3_{precise}$	0.799	0.796	145.425	12.043	8.323	10.619
	$Set3_{fuzzy}$	0.799	0.798	145.550	12.049	8.307	10.630
	$Set4_{precise}$	0.797	0.796	147.092	12.112	8.361	10.718
	$Set4_{fuzzy}$	0.798	0.796	146.291	12.080	8.324	10.657
	$Set5_{precise}$	0.795	0.803	148.077	12.154	8.433	10.758
	$Set5_{fuzzy}$	0.793	0.801	149.538	12.215	8.448	10.795

Table A3. The fit performance of eleven machine learning models for ship S4 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.906	0.758	81.312	8.851	6.363	6.681
	$Set2_{precise}$	0.926	0.750	59.829	7.503	5.663	6.124
	$Set2_{fuzzy}$	0.921	0.759	63.603	7.812	5.886	6.341
	$Set3_{precise}$	0.916	0.746	68.063	8.094	6.036	6.523
	$Set3_{fuzzy}$	0.921	0.755	63.698	7.776	5.796	6.220
	$Set4_{precise}$	0.917	0.758	67.295	7.920	5.904	6.372
	$Set4_{fuzzy}$	0.928	0.771	58.562	7.517	5.638	6.059
	$Set5_{precise}$	0.905	0.739	76.918	8.473	6.344	6.864
	$Set5_{fuzzy}$	0.918	0.764	66.324	7.897	5.900	6.326
ET	*Set1*	0.988	0.858	10.120	2.625	1.778	1.862
	$Set2_{precise}$	0.996	0.865	2.961	1.362	0.957	1.036
	$Set2_{fuzzy}$	0.998	0.862	1.882	1.077	0.738	0.796
	$Set3_{precise}$	0.998	0.872	1.434	0.901	0.627	0.687
	$Set3_{fuzzy}$	0.998	0.870	1.957	1.022	0.713	0.777
	$Set4_{precise}$	0.997	0.871	2.141	1.101	0.778	0.844
	$Set4_{fuzzy}$	0.997	0.867	2.092	0.994	0.710	0.779
	$Set5_{precise}$	0.999	0.875	1.183	0.904	0.623	0.675
	$Set5_{fuzzy}$	0.999	0.871	0.933	0.704	0.479	0.524
RF	*Set1*	0.974	0.848	22.794	4.752	3.335	3.501
	$Set2_{precise}$	0.977	0.855	18.989	4.350	3.226	3.528
	$Set2_{fuzzy}$	0.975	0.852	20.670	4.529	3.344	3.673
	$Set3_{precise}$	0.975	0.853	20.349	4.497	3.331	3.618
	$Set3_{fuzzy}$	0.974	0.856	20.789	4.535	3.341	3.631
	$Set4_{precise}$	0.974	0.855	21.029	4.568	3.359	3.660
	$Set4_{fuzzy}$	0.976	0.855	19.273	4.381	3.235	3.533
	$Set5_{precise}$	0.975	0.857	20.143	4.472	3.320	3.609
	$Set5_{fuzzy}$	0.975	0.859	20.015	4.457	3.275	3.570
AB	*Set1*	0.980	0.843	17.332	3.939	3.283	3.654
	$Set2_{precise}$	0.981	0.855	15.155	3.560	2.938	3.255
	$Set2_{fuzzy}$	0.980	0.856	16.402	3.758	3.082	3.408
	$Set3_{precise}$	0.986	0.865	11.021	3.144	2.591	2.905
	$Set3_{fuzzy}$	0.992	0.868	6.360	2.277	1.815	2.053
	$Set4_{precise}$	0.992	0.864	6.179	2.258	1.821	2.046
	$Set4_{fuzzy}$	0.992	0.864	6.345	2.272	1.806	2.021
	$Set5_{precise}$	0.991	0.864	7.066	2.371	1.903	2.139
	$Set5_{fuzzy}$	0.993	0.862	5.401	2.070	1.614	1.806
GB	*Set1*	0.977	0.851	19.591	4.196	3.176	3.352
	$Set2_{precise}$	0.986	0.863	11.541	2.974	2.340	2.505
	$Set2_{fuzzy}$	0.985	0.858	12.254	2.929	2.287	2.437
	$Set3_{precise}$	0.989	0.866	8.845	2.500	1.838	1.957
	$Set3_{fuzzy}$	0.986	0.869	11.433	2.985	2.229	2.395
	$Set4_{precise}$	0.991	0.870	7.282	2.380	1.819	1.934
	$Set4_{fuzzy}$	0.990	0.867	8.073	2.523	1.913	2.039
	$Set5_{precise}$	0.990	0.867	7.786	2.438	1.819	1.963
	$Set5_{fuzzy}$	0.986	0.868	11.156	2.861	2.190	2.362
XG	*Set1*	0.977	0.858	19.657	4.126	3.068	3.185
	$Set2_{precise}$	0.993	0.860	5.385	1.929	1.448	1.532
	$Set2_{fuzzy}$	0.993	0.864	5.871	2.111	1.576	1.681
	$Set3_{precise}$	0.995	0.869	3.758	1.585	1.140	1.201
	$Set3_{fuzzy}$	0.993	0.874	5.620	2.167	1.535	1.623

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	$Set4_{precise}$	0.994	0.871	4.730	1.636	1.209	1.273
	$Set4_{fuzzy}$	0.990	0.868	7.942	2.402	1.785	1.886
	$Set5_{precise}$	0.993	0.869	5.929	1.909	1.441	1.535
	$Set5_{fuzzy}$	0.986	0.871	11.464	2.905	2.212	2.369
LB	$Set1$	0.968	0.844	28.153	5.010	3.861	4.044
	$Set2_{precise}$	0.980	0.850	15.758	3.612	2.794	2.972
	$Set2_{fuzzy}$	0.978	0.851	17.859	3.894	3.036	3.245
	$Set3_{precise}$	0.987	0.855	10.943	2.871	2.200	2.340
	$Set3_{fuzzy}$	0.987	0.861	10.620	2.951	2.264	2.432
	$Set4_{precise}$	0.986	0.857	11.242	2.921	2.241	2.385
	$Set4_{fuzzy}$	0.977	0.863	18.771	4.087	3.166	3.364
	$Set5_{precise}$	0.992	0.866	6.305	2.107	1.627	1.771
	$Set5_{fuzzy}$	0.987	0.868	10.615	2.972	2.300	2.509
SVM	$Set1$	0.906	0.842	81.874	9.015	6.318	6.374
	$Set2_{precise}$	0.920	0.852	64.910	8.026	5.944	6.337
	$Set2_{fuzzy}$	0.917	0.847	67.218	8.157	5.995	6.449
	$Set3_{precise}$	0.921	0.857	63.718	7.972	5.848	6.146
	$Set3_{fuzzy}$	0.912	0.853	70.915	8.406	6.166	6.467
	$Set4_{precise}$	0.920	0.861	64.380	8.005	5.896	6.203
	$Set4_{fuzzy}$	0.913	0.855	70.163	8.353	6.150	6.479
	$Set5_{precise}$	0.921	0.863	64.323	8.003	5.923	6.291
	$Set5_{fuzzy}$	0.918	0.862	66.783	8.152	6.034	6.410
ANN	$Set1$	0.925	0.845	65.521	8.076	6.102	6.390
	$Set2_{precise}$	0.936	0.848	51.520	7.145	5.561	6.025
	$Set2_{fuzzy}$	0.939	0.851	49.215	6.999	5.433	5.884
	$Set3_{precise}$	0.947	0.856	42.555	6.513	5.034	5.502
	$Set3_{fuzzy}$	0.947	0.863	42.882	6.543	5.085	5.528
	$Set4_{precise}$	0.944	0.859	45.586	6.744	5.243	5.676
	$Set4_{fuzzy}$	0.942	0.855	47.334	6.865	5.320	5.759
	$Set5_{precise}$	0.939	0.866	49.545	7.025	5.477	5.914
	$Set5_{fuzzy}$	0.933	0.865	54.535	7.374	5.737	6.216
Ridge	$Set1$	0.825	0.821	152.631	12.351	9.343	9.548
	$Set2_{precise}$	0.824	0.805	142.173	11.919	9.220	9.569
	$Set2_{fuzzy}$	0.820	0.799	145.931	12.075	9.331	9.742
	$Set3_{precise}$	0.833	0.811	135.334	11.629	9.033	9.406
	$Set3_{fuzzy}$	0.828	0.806	138.677	11.771	9.128	9.537
	$Set4_{precise}$	0.833	0.812	135.132	11.620	9.021	9.387
	$Set4_{fuzzy}$	0.829	0.807	138.424	11.761	9.115	9.516
	$Set5_{precise}$	0.829	0.812	138.032	11.744	9.121	9.468
	$Set5_{fuzzy}$	0.826	0.808	140.883	11.865	9.207	9.606
LASSO	$Set1$	0.824	0.823	153.402	12.382	9.347	9.537
	$Set2_{precise}$	0.824	0.804	142.470	11.932	9.228	9.572
	$Set2_{fuzzy}$	0.819	0.799	146.410	12.095	9.344	9.750
	$Set3_{precise}$	0.832	0.809	135.961	11.656	9.053	9.417
	$Set3_{fuzzy}$	0.826	0.806	140.683	11.856	9.181	9.580
	$Set4_{precise}$	0.833	0.810	135.135	11.621	9.022	9.385
	$Set4_{fuzzy}$	0.828	0.806	139.214	11.793	9.132	9.526
	$Set5_{precise}$	0.829	0.811	138.350	11.758	9.127	9.471
	$Set5_{fuzzy}$	0.825	0.807	141.337	11.884	9.215	9.611

Table A4. The fit performance of eleven machine learning models for ship S5 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.939	0.821	33.699	5.588	4.144	6.259
	Set2$_{precise}$	0.938	0.795	34.454	5.745	4.239	6.417
	Set2$_{fuzzy}$	0.935	0.810	35.970	5.952	4.463	6.795
	Set3$_{precise}$	0.947	0.785	29.488	5.182	3.764	5.625
	Set3$_{fuzzy}$	0.948	0.798	28.634	5.139	3.789	5.696
	Set4$_{precise}$	0.938	0.786	34.047	5.667	4.152	6.186
	Set4$_{fuzzy}$	0.940	0.799	33.582	5.565	4.092	6.178
	Set5$_{precise}$	0.937	0.799	35.019	5.784	4.241	6.345
	Set5$_{fuzzy}$	0.941	0.811	32.605	5.613	4.060	6.073
ET	*Set1*	0.998	0.895	1.057	0.805	0.569	0.857
	Set2$_{precise}$	0.996	0.892	2.026	1.108	0.820	1.257
	Set2$_{fuzzy}$	0.994	0.889	3.403	1.580	1.182	1.787
	Set3$_{precise}$	0.997	0.892	1.413	0.854	0.619	0.935
	Set3$_{fuzzy}$	0.997	0.891	1.821	1.076	0.784	1.184
	Set4$_{precise}$	0.995	0.892	2.602	1.195	0.883	1.343
	Set4$_{fuzzy}$	0.997	0.888	1.705	0.950	0.681	1.028
	Set5$_{precise}$	0.998	0.890	0.845	0.785	0.560	0.856
	Set5$_{fuzzy}$	0.997	0.889	1.447	0.856	0.619	0.939
RF	*Set1*	0.982	0.884	9.951	3.140	2.354	3.594
	Set2$_{precise}$	0.981	0.874	10.785	3.268	2.396	3.663
	Set2$_{fuzzy}$	0.983	0.881	9.662	3.097	2.265	3.480
	Set3$_{precise}$	0.981	0.874	10.498	3.225	2.390	3.663
	Set3$_{fuzzy}$	0.981	0.882	10.352	3.195	2.354	3.614
	Set4$_{precise}$	0.982	0.873	9.889	3.137	2.295	3.509
	Set4$_{fuzzy}$	0.981	0.880	10.422	3.210	2.317	3.539
	Set5$_{precise}$	0.981	0.876	10.305	3.189	2.355	3.598
	Set5$_{fuzzy}$	0.982	0.881	10.256	3.184	2.356	3.630
AB	*Set1*	0.990	0.895	5.408	2.213	1.830	3.156
	Set2$_{precise}$	0.994	0.882	3.555	1.780	1.439	2.538
	Set2$_{fuzzy}$	0.992	0.890	4.604	1.967	1.620	2.822
	Set3$_{precise}$	0.995	0.886	2.543	1.525	1.209	2.217
	Set3$_{fuzzy}$	0.994	0.893	3.311	1.634	1.320	2.360
	Set4$_{precise}$	0.994	0.882	3.462	1.734	1.393	2.479
	Set4$_{fuzzy}$	0.995	0.890	2.588	1.508	1.191	2.135
	Set5$_{precise}$	0.995	0.886	2.965	1.629	1.315	2.379
	Set5$_{fuzzy}$	0.996	0.890	2.337	1.387	1.066	1.931
GB	*Set1*	0.993	0.895	3.885	1.743	1.360	2.158
	Set2$_{precise}$	0.996	0.887	2.381	1.188	0.926	1.492
	Set2$_{fuzzy}$	0.990	0.888	5.618	1.854	1.419	2.233
	Set3$_{precise}$	0.993	0.887	3.519	1.359	1.021	1.610
	Set3$_{fuzzy}$	0.994	0.888	3.316	1.398	1.074	1.699
	Set4$_{precise}$	0.997	0.889	1.727	1.039	0.805	1.275
	Set4$_{fuzzy}$	0.995	0.891	2.928	1.162	0.867	1.355
	Set5$_{precise}$	0.996	0.884	2.061	1.052	0.818	1.293
	Set5$_{fuzzy}$	0.993	0.886	3.950	1.562	1.211	1.917
XG	*Set1*	0.990	0.892	5.361	1.995	1.520	2.370
	Set2$_{precise}$	0.993	0.873	3.883	1.589	1.173	1.842
	Set2$_{fuzzy}$	0.987	0.881	7.018	2.245	1.655	2.554
	Set3$_{precise}$	0.993	0.878	3.601	1.605	1.133	1.749
	Set3$_{fuzzy}$	0.989	0.888	5.924	2.046	1.490	2.314

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	$Set4_{precise}$	0.994	0.878	3.149	1.496	1.041	1.616
	$Set4_{fuzzy}$	0.990	0.883	5.556	2.019	1.478	2.308
	$Set5_{precise}$	0.987	0.883	7.125	2.403	1.776	2.773
	$Set5_{fuzzy}$	0.993	0.886	3.759	1.666	1.228	1.927
LB	$Set1$	0.986	0.879	7.810	2.636	2.028	3.173
	$Set2_{precise}$	0.984	0.874	9.087	2.680	2.016	3.145
	$Set2_{fuzzy}$	0.984	0.882	9.066	2.728	2.073	3.215
	$Set3_{precise}$	0.987	0.873	7.382	2.350	1.758	2.725
	$Set3_{fuzzy}$	0.987	0.876	7.357	2.458	1.852	2.887
	$Set4_{precise}$	0.979	0.875	11.525	3.220	2.420	3.743
	$Set4_{fuzzy}$	0.980	0.877	11.114	3.131	2.393	3.696
	$Set5_{precise}$	0.984	0.871	8.646	2.662	1.998	3.104
	$Set5_{fuzzy}$	0.987	0.876	7.270	2.424	1.830	2.842
SVM	$Set1$	0.931	0.884	38.408	6.173	4.382	6.630
	$Set2_{precise}$	0.919	0.879	45.002	6.674	4.868	7.358
	$Set2_{fuzzy}$	0.919	0.883	44.677	6.633	4.835	7.319
	$Set3_{precise}$	0.916	0.873	46.421	6.785	4.917	7.472
	$Set3_{fuzzy}$	0.917	0.882	46.286	6.793	4.904	7.472
	$Set4_{precise}$	0.915	0.876	47.018	6.840	4.985	7.541
	$Set4_{fuzzy}$	0.917	0.879	45.834	6.758	4.928	7.496
	$Set5_{precise}$	0.924	0.878	41.942	6.444	4.689	7.114
	$Set5_{fuzzy}$	0.921	0.880	43.746	6.591	4.771	7.252
ANN	$Set1$	0.926	0.886	40.737	6.373	4.900	7.545
	$Set2_{precise}$	0.940	0.876	33.557	5.753	4.426	6.867
	$Set2_{fuzzy}$	0.930	0.876	38.794	6.188	4.724	7.295
	$Set3_{precise}$	0.935	0.879	36.157	5.956	4.544	7.075
	$Set3_{fuzzy}$	0.932	0.882	37.513	6.094	4.633	7.202
	$Set4_{precise}$	0.941	0.882	32.448	5.659	4.328	6.738
	$Set4_{fuzzy}$	0.929	0.878	39.150	6.229	4.760	7.381
	$Set5_{precise}$	0.928	0.876	39.757	6.269	4.802	7.409
	$Set5_{fuzzy}$	0.930	0.884	38.850	6.201	4.720	7.277
Ridge	$Set1$	0.875	0.868	69.368	8.325	6.341	9.937
	$Set2_{precise}$	0.883	0.873	65.112	8.066	6.112	9.419
	$Set2_{fuzzy}$	0.881	0.873	66.119	8.128	6.124	9.423
	$Set3_{precise}$	0.889	0.868	61.610	7.846	5.934	9.109
	$Set3_{fuzzy}$	0.888	0.870	62.092	7.876	5.983	9.191
	$Set4_{precise}$	0.887	0.871	62.716	7.916	6.011	9.210
	$Set4_{fuzzy}$	0.886	0.870	63.240	7.949	6.063	9.298
	$Set5_{precise}$	0.885	0.874	63.789	7.983	6.042	9.244
	$Set5_{fuzzy}$	0.885	0.875	63.975	7.995	6.045	9.248
LASSO	$Set1$	0.874	0.868	69.799	8.351	6.357	9.948
	$Set2_{precise}$	0.882	0.873	65.214	8.072	6.121	9.436
	$Set2_{fuzzy}$	0.881	0.873	66.225	8.135	6.131	9.439
	$Set3_{precise}$	0.888	0.868	61.988	7.870	5.953	9.129
	$Set3_{fuzzy}$	0.887	0.870	62.780	7.920	6.019	9.213
	$Set4_{precise}$	0.886	0.870	62.963	7.932	6.022	9.224
	$Set4_{fuzzy}$	0.886	0.871	63.365	7.957	6.070	9.298
	$Set5_{precise}$	0.885	0.874	63.959	7.994	6.054	9.256
	$Set5_{fuzzy}$	0.884	0.873	64.202	8.009	6.059	9.247

Table A5. The fit performance of eleven machine learning models for ship S6 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	Set1	0.837	0.636	67.292	8.143	5.917	7.777
	$Set2_{precise}$	0.812	0.542	77.709	8.758	6.526	8.686
	$Set2_{fuzzy}$	0.825	0.576	72.552	8.460	6.207	8.219
	$Set3_{precise}$	0.832	0.576	69.684	8.275	6.119	8.113
DT	$Set3_{fuzzy}$	0.813	0.579	77.414	8.738	6.468	8.593
	$Set4_{precise}$	0.852	0.530	60.986	7.701	5.653	7.536
	$Set4_{fuzzy}$	0.832	0.561	69.219	8.249	6.071	8.066
	$Set5_{precise}$	0.832	0.578	69.447	8.227	6.047	8.027
	$Set5_{fuzzy}$	0.816	0.589	76.002	8.649	6.372	8.434
	Set1	0.985	0.765	6.050	1.928	1.359	1.796
	$Set2_{precise}$	0.982	0.755	7.287	2.366	1.743	2.313
	$Set2_{fuzzy}$	0.982	0.749	7.640	2.515	1.877	2.489
	$Set3_{precise}$	0.979	0.752	8.706	2.743	2.010	2.678
ET	$Set3_{fuzzy}$	0.976	0.744	9.938	2.705	1.995	2.654
	$Set4_{precise}$	0.986	0.747	5.823	2.122	1.533	2.036
	$Set4_{fuzzy}$	0.973	0.743	10.997	2.925	2.161	2.860
	$Set5_{precise}$	0.986	0.750	5.796	2.088	1.517	2.009
	$Set5_{fuzzy}$	0.971	0.735	12.056	3.189	2.388	3.162
	Set1	0.956	0.766	18.155	4.225	3.016	4.012
	$Set2_{precise}$	0.957	0.743	18.016	4.231	3.057	4.053
	$Set2_{fuzzy}$	0.956	0.747	18.125	4.227	3.108	4.136
	$Set3_{precise}$	0.953	0.740	19.498	4.382	3.173	4.211
RF	$Set3_{fuzzy}$	0.954	0.746	19.075	4.326	3.152	4.198
	$Set4_{precise}$	0.956	0.741	18.314	4.255	3.102	4.122
	$Set4_{fuzzy}$	0.952	0.747	19.918	4.437	3.247	4.325
	$Set5_{precise}$	0.956	0.741	18.403	4.261	3.092	4.101
	$Set5_{fuzzy}$	0.955	0.744	18.502	4.271	3.128	4.161
	Set1	0.969	0.770	12.857	3.481	2.871	4.105
	$Set2_{precise}$	0.973	0.752	10.958	3.199	2.673	3.873
	$Set2_{fuzzy}$	0.968	0.758	13.404	3.558	2.994	4.289
	$Set3_{precise}$	0.980	0.755	8.175	2.647	2.186	3.210
AB	$Set3_{fuzzy}$	0.974	0.760	10.820	3.157	2.664	3.851
	$Set4_{precise}$	0.977	0.747	9.492	2.996	2.541	3.702
	$Set4_{fuzzy}$	0.977	0.749	9.524	2.962	2.484	3.578
	$Set5_{precise}$	0.983	0.748	6.868	2.445	2.005	2.945
	$Set5_{fuzzy}$	0.971	0.752	12.050	3.405	2.911	4.181
	Set1	0.965	0.786	14.509	3.538	2.597	3.507
	$Set2_{precise}$	0.962	0.784	15.864	3.689	2.868	3.907
	$Set2_{fuzzy}$	0.963	0.780	15.468	3.513	2.763	3.749
	$Set3_{precise}$	0.971	0.770	11.917	3.111	2.384	3.226
GB	$Set3_{fuzzy}$	0.963	0.780	15.292	3.572	2.725	3.693
	$Set4_{precise}$	0.968	0.776	13.322	3.271	2.530	3.425
	$Set4_{fuzzy}$	0.968	0.778	13.396	3.319	2.549	3.451
	$Set5_{precise}$	0.962	0.771	16.035	3.730	2.902	3.958
	$Set5_{fuzzy}$	0.950	0.771	20.910	4.247	3.327	4.513
	Set1	0.966	0.786	14.223	3.620	2.692	3.641
	$Set2_{precise}$	0.957	0.785	17.661	3.806	2.941	4.004
XG	$Set2_{fuzzy}$	0.945	0.786	22.931	4.633	3.606	4.878
	$Set3_{precise}$	0.959	0.771	17.299	3.835	2.890	3.902
	$Set3_{fuzzy}$	0.966	0.776	13.923	3.361	2.538	3.412

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	$Set4_{precise}$	0.958	0.770	17.405	3.889	2.959	3.993
	$Set4_{fuzzy}$	0.955	0.774	18.740	4.036	3.066	4.127
	$Set5_{precise}$	0.957	0.773	17.877	3.837	2.946	4.005
	$Set5_{fuzzy}$	0.940	0.777	24.886	4.800	3.711	5.024
LB	$Set1$	0.951	0.773	20.401	4.334	3.285	4.472
	$Set2_{precise}$	0.951	0.768	20.215	4.252	3.275	4.467
	$Set2_{fuzzy}$	0.936	0.772	26.454	4.902	3.810	5.175
	$Set3_{precise}$	0.963	0.754	15.520	3.514	2.682	3.646
	$Set3_{fuzzy}$	0.951	0.762	20.246	4.209	3.275	4.443
	$Set4_{precise}$	0.942	0.758	24.155	4.587	3.520	4.790
	$Set4_{fuzzy}$	0.936	0.768	26.469	4.900	3.800	5.151
	$Set5_{precise}$	0.962	0.752	15.655	3.592	2.706	3.679
	$Set5_{fuzzy}$	0.956	0.751	18.155	3.893	2.959	4.023
SVM	$Set1$	0.838	0.748	67.236	8.175	5.625	7.308
	$Set2_{precise}$	0.846	0.766	63.819	7.962	5.661	7.464
	$Set2_{fuzzy}$	0.832	0.754	69.485	8.311	5.956	7.829
	$Set3_{precise}$	0.843	0.767	65.144	8.045	5.755	7.629
	$Set3_{fuzzy}$	0.832	0.760	69.588	8.322	5.960	7.862
	$Set4_{precise}$	0.840	0.765	66.027	8.104	5.765	7.603
	$Set4_{fuzzy}$	0.828	0.762	71.019	8.420	5.973	7.859
	$Set5_{precise}$	0.844	0.767	64.564	8.020	5.739	7.570
	$Set5_{fuzzy}$	0.825	0.766	72.178	8.491	6.051	7.957
ANN	$Set1$	0.851	0.740	61.550	7.798	5.849	7.715
	$Set2_{precise}$	0.851	0.768	61.489	7.821	5.883	7.791
	$Set2_{fuzzy}$	0.847	0.759	63.370	7.935	6.000	7.927
	$Set3_{precise}$	0.859	0.772	58.184	7.599	5.750	7.603
	$Set3_{fuzzy}$	0.846	0.768	63.903	7.977	6.043	7.967
	$Set4_{precise}$	0.852	0.773	61.205	7.803	5.883	7.760
	$Set4_{fuzzy}$	0.849	0.760	62.489	7.871	5.952	7.852
	$Set5_{precise}$	0.875	0.759	51.893	7.155	5.453	7.254
	$Set5_{fuzzy}$	0.863	0.758	56.786	7.496	5.712	7.557
Ridge	$Set1$	0.758	0.729	100.434	10.018	7.588	10.192
	$Set2_{precise}$	0.762	0.736	98.605	9.927	7.566	10.137
	$Set2_{fuzzy}$	0.758	0.734	99.954	9.994	7.577	10.120
	$Set3_{precise}$	0.775	0.745	93.218	9.652	7.454	9.977
	$Set3_{fuzzy}$	0.772	0.743	94.465	9.716	7.434	9.927
	$Set4_{precise}$	0.774	0.744	93.718	9.678	7.439	9.949
	$Set4_{fuzzy}$	0.770	0.743	95.067	9.747	7.424	9.909
	$Set5_{precise}$	0.768	0.742	95.854	9.788	7.559	10.138
	$Set5_{fuzzy}$	0.764	0.739	97.477	9.870	7.573	10.120
LASSO	$Set1$	0.753	0.724	102.272	10.109	7.629	10.199
	$Set2_{precise}$	0.762	0.736	98.666	9.930	7.562	10.136
	$Set2_{fuzzy}$	0.758	0.733	99.969	9.995	7.573	10.118
	$Set3_{precise}$	0.774	0.744	93.502	9.667	7.443	9.960
	$Set3_{fuzzy}$	0.771	0.745	94.691	9.728	7.429	9.920
	$Set4_{precise}$	0.774	0.743	93.686	9.676	7.435	9.942
	$Set4_{fuzzy}$	0.770	0.743	95.199	9.754	7.411	9.879
	$Set5_{precise}$	0.768	0.742	95.856	9.788	7.560	10.146
	$Set5_{fuzzy}$	0.764	0.739	97.494	9.871	7.573	10.120

Table A6. The fit performance of eleven machine learning models for ship S7 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.828	0.680	69.472	8.260	6.302	8.155
	$Set2_{precise}$	0.857	0.682	57.167	7.500	5.712	7.424
	$Set2_{fuzzy}$	0.849	0.660	60.603	7.737	5.932	7.752
	$Set3_{precise}$	0.880	0.683	48.319	6.903	5.173	6.749
	$Set3_{fuzzy}$	0.869	0.690	52.514	7.183	5.401	7.043
	$Set4_{precise}$	0.875	0.656	50.042	7.032	5.307	6.936
	$Set4_{fuzzy}$	0.861	0.667	55.626	7.387	5.599	7.303
	$Set5_{precise}$	0.881	0.694	47.827	6.863	5.127	6.665
	$Set5_{fuzzy}$	0.867	0.700	53.253	7.244	5.425	7.077
ET	*Set1*	0.956	0.806	17.780	3.880	2.884	3.713
	$Set2_{precise}$	0.972	0.801	11.382	3.040	2.178	2.834
	$Set2_{fuzzy}$	0.963	0.790	14.758	3.560	2.603	3.391
	$Set3_{precise}$	0.987	0.805	5.176	1.848	1.259	1.639
	$Set3_{fuzzy}$	0.978	0.798	8.693	2.379	1.664	2.155
	$Set4_{precise}$	0.985	0.801	6.087	2.040	1.419	1.851
	$Set4_{fuzzy}$	0.983	0.793	6.706	2.149	1.522	1.983
	$Set5_{precise}$	0.989	0.804	4.334	1.623	1.156	1.507
	$Set5_{fuzzy}$	0.979	0.799	8.329	2.549	1.855	2.405
RF	*Set1*	0.964	0.793	14.369	3.774	2.813	3.649
	$Set2_{precise}$	0.962	0.791	15.123	3.842	2.826	3.694
	$Set2_{fuzzy}$	0.962	0.788	15.442	3.899	2.887	3.800
	$Set3_{precise}$	0.961	0.794	15.501	3.920	2.867	3.740
	$Set3_{fuzzy}$	0.960	0.793	15.963	3.978	2.931	3.838
	$Set4_{precise}$	0.961	0.791	15.742	3.947	2.898	3.795
	$Set4_{fuzzy}$	0.963	0.789	14.852	3.828	2.850	3.746
	$Set5_{precise}$	0.966	0.796	13.853	3.691	2.705	3.528
	$Set5_{fuzzy}$	0.967	0.797	13.384	3.644	2.715	3.551
AB	*Set1*	0.964	0.790	14.672	3.464	2.781	3.712
	$Set2_{precise}$	0.975	0.770	10.014	2.848	2.207	2.964
	$Set2_{fuzzy}$	0.975	0.777	10.055	2.981	2.462	3.326
	$Set3_{precise}$	0.982	0.777	7.272	2.415	1.888	2.558
	$Set3_{fuzzy}$	0.980	0.782	7.977	2.604	2.149	2.934
	$Set4_{precise}$	0.982	0.776	7.209	2.516	2.080	2.838
	$Set4_{fuzzy}$	0.986	0.775	5.620	2.132	1.717	2.345
	$Set5_{precise}$	0.984	0.783	6.493	2.337	1.897	2.589
	$Set5_{fuzzy}$	0.987	0.783	5.343	2.100	1.752	2.417
GB	*Set1*	0.962	0.803	15.408	3.756	2.777	3.605
	$Set2_{precise}$	0.966	0.777	13.669	3.347	2.513	3.299
	$Set2_{fuzzy}$	0.967	0.782	13.148	3.471	2.602	3.411
	$Set3_{precise}$	0.986	0.785	5.466	2.156	1.442	1.880
	$Set3_{fuzzy}$	0.978	0.782	9.078	2.764	1.973	2.582
	$Set4_{precise}$	0.979	0.781	8.487	2.562	1.839	2.398
	$Set4_{fuzzy}$	0.977	0.780	9.243	2.751	2.065	2.709
	$Set5_{precise}$	0.974	0.786	10.452	2.955	2.139	2.784
	$Set5_{fuzzy}$	0.977	0.791	9.274	2.812	2.038	2.663
XG	*Set1*	0.972	0.813	11.021	3.022	2.222	2.865
	$Set2_{precise}$	0.972	0.777	11.392	3.120	2.269	2.926
	$Set2_{fuzzy}$	0.966	0.784	13.695	3.461	2.551	3.314
	$Set3_{precise}$	0.986	0.784	5.731	2.093	1.424	1.808
	$Set3_{fuzzy}$	0.980	0.791	8.192	2.475	1.677	2.138

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	$Set4_{precise}$	0.979	0.784	8.576	2.711	1.867	2.394
	$Set4_{fuzzy}$	0.975	0.781	10.191	2.769	1.988	2.568
	$Set5_{precise}$	0.978	0.792	8.680	2.772	1.971	2.532
	$Set5_{fuzzy}$	0.973	0.798	11.025	3.167	2.272	2.932
LB	$Set1$	0.957	0.789	17.547	4.044	3.053	3.968
	$Set2_{precise}$	0.978	0.766	8.892	2.744	2.051	2.693
	$Set2_{fuzzy}$	0.967	0.781	13.285	3.426	2.603	3.407
	$Set3_{precise}$	0.982	0.785	7.152	2.366	1.742	2.283
	$Set3_{fuzzy}$	0.975	0.779	10.039	2.861	2.175	2.840
	$Set4_{precise}$	0.981	0.775	7.814	2.542	1.865	2.427
	$Set4_{fuzzy}$	0.961	0.774	15.690	3.755	2.837	3.692
	$Set5_{precise}$	0.979	0.781	8.598	2.522	1.855	2.431
	$Set5_{fuzzy}$	0.979	0.787	8.541	2.673	1.945	2.545
SVM	$Set1$	0.906	0.786	38.185	6.078	4.323	5.574
	$Set2_{precise}$	0.870	0.744	52.317	7.160	5.243	6.699
	$Set2_{fuzzy}$	0.873	0.744	50.939	7.070	5.115	6.582
	$Set3_{precise}$	0.871	0.748	51.533	7.113	5.173	6.591
	$Set3_{fuzzy}$	0.866	0.746	53.727	7.249	5.261	6.730
	$Set4_{precise}$	0.873	0.745	50.841	7.045	5.125	6.552
	$Set4_{fuzzy}$	0.876	0.745	49.716	6.983	5.055	6.504
	$Set5_{precise}$	0.867	0.752	53.427	7.236	5.264	6.699
	$Set5_{fuzzy}$	0.863	0.750	54.999	7.343	5.366	6.862
ANN	$Set1$	0.863	0.786	55.639	7.392	5.651	7.274
	$Set2_{precise}$	0.902	0.770	39.097	6.203	4.856	6.310
	$Set2_{fuzzy}$	0.888	0.764	45.229	6.656	5.210	6.807
	$Set3_{precise}$	0.892	0.771	43.321	6.515	5.071	6.587
	$Set3_{fuzzy}$	0.896	0.760	41.782	6.386	4.981	6.481
	$Set4_{precise}$	0.897	0.767	41.488	6.373	4.959	6.441
	$Set4_{fuzzy}$	0.891	0.765	43.762	6.559	5.110	6.643
	$Set5_{precise}$	0.895	0.756	42.257	6.425	5.003	6.480
	$Set5_{fuzzy}$	0.884	0.755	46.486	6.748	5.244	6.798
Ridge	$Set1$	0.790	0.781	85.163	9.224	6.955	8.817
	$Set2_{precise}$	0.817	0.761	73.490	8.564	6.612	8.463
	$Set2_{fuzzy}$	0.816	0.761	74.035	8.596	6.669	8.578
	$Set3_{precise}$	0.820	0.758	72.381	8.498	6.520	8.315
	$Set3_{fuzzy}$	0.818	0.756	72.910	8.530	6.596	8.451
	$Set4_{precise}$	0.819	0.759	72.799	8.523	6.552	8.361
	$Set4_{fuzzy}$	0.818	0.758	73.098	8.541	6.609	8.473
	$Set5_{precise}$	0.818	0.761	73.217	8.547	6.587	8.422
	$Set5_{fuzzy}$	0.816	0.760	73.809	8.583	6.660	8.550
LASSO	$Set1$	0.789	0.781	85.405	9.238	6.961	8.819
	$Set2_{precise}$	0.816	0.760	73.729	8.577	6.627	8.498
	$Set2_{fuzzy}$	0.815	0.760	74.215	8.606	6.673	8.595
	$Set3_{precise}$	0.819	0.758	72.827	8.524	6.550	8.374
	$Set3_{fuzzy}$	0.816	0.759	73.644	8.573	6.627	8.498
	$Set4_{precise}$	0.817	0.758	73.386	8.557	6.591	8.435
	$Set4_{fuzzy}$	0.817	0.759	73.478	8.563	6.620	8.500
	$Set5_{precise}$	0.817	0.761	73.504	8.564	6.606	8.462
	$Set5_{fuzzy}$	0.815	0.761	74.089	8.599	6.666	8.564

Table A7. The fit performance of eleven machine learning models for ship S8 (DFS1)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	$Set1$	0.916	0.774	54.181	7.305	5.213	6.441
	$Set2_{precise}$	0.912	0.759	52.806	7.194	5.129	6.172
	$Set2_{fuzzy}$	0.905	0.766	57.141	7.487	5.353	6.419
	$Set3_{precise}$	0.916	0.769	50.649	6.985	4.922	5.949
	$Set3_{fuzzy}$	0.919	0.764	48.884	6.889	4.885	5.920
	$Set4_{precise}$	0.912	0.746	52.752	7.168	5.070	6.092
	$Set4_{fuzzy}$	0.904	0.759	57.557	7.549	5.362	6.450
	$Set5_{precise}$	0.914	0.759	51.752	7.105	5.020	6.054
	$Set5_{fuzzy}$	0.909	0.770	54.614	7.268	5.163	6.237
ET	$Set1$	0.998	0.882	1.556	0.811	0.551	0.679
	$Set2_{precise}$	0.997	0.872	1.552	0.879	0.565	0.694
	$Set2_{fuzzy}$	0.998	0.866	1.288	0.841	0.540	0.661
	$Set3_{precise}$	0.995	0.876	2.783	1.404	0.907	1.120
	$Set3_{fuzzy}$	0.997	0.872	1.940	1.024	0.652	0.801
	$Set4_{precise}$	0.996	0.871	2.382	1.227	0.799	0.993
	$Set4_{fuzzy}$	0.995	0.865	2.894	1.392	0.879	1.077
	$Set5_{precise}$	0.999	0.883	0.612	0.629	0.398	0.486
	$Set5_{fuzzy}$	0.996	0.877	2.216	1.169	0.771	0.947
RF	$Set1$	0.978	0.859	13.895	3.707	2.535	3.124
	$Set2_{precise}$	0.974	0.846	15.712	3.941	2.668	3.233
	$Set2_{fuzzy}$	0.977	0.846	14.095	3.740	2.546	3.081
	$Set3_{precise}$	0.976	0.855	14.566	3.798	2.624	3.187
	$Set3_{fuzzy}$	0.975	0.854	15.158	3.868	2.676	3.254
	$Set4_{precise}$	0.976	0.847	14.789	3.811	2.615	3.173
	$Set4_{fuzzy}$	0.977	0.848	13.912	3.714	2.561	3.109
	$Set5_{precise}$	0.978	0.864	13.567	3.653	2.490	3.026
	$Set5_{fuzzy}$	0.976	0.861	14.658	3.788	2.569	3.128
AB	$Set1$	0.982	0.870	11.601	3.288	2.747	3.479
	$Set2_{precise}$	0.989	0.860	6.723	2.470	2.032	2.565
	$Set2_{fuzzy}$	0.990	0.863	5.990	2.329	1.896	2.390
	$Set3_{precise}$	0.991	0.863	5.365	2.114	1.693	2.148
	$Set3_{fuzzy}$	0.991	0.864	5.138	2.093	1.673	2.127
	$Set4_{precise}$	0.992	0.859	5.046	2.111	1.705	2.162
	$Set4_{fuzzy}$	0.992	0.861	4.835	2.069	1.654	2.102
	$Set5_{precise}$	0.993	0.870	4.374	1.789	1.402	1.780
	$Set5_{fuzzy}$	0.997	0.874	1.966	1.273	0.931	1.189
GB	$Set1$	0.983	0.875	10.771	3.062	2.188	2.750
	$Set2_{precise}$	0.978	0.855	13.587	3.035	2.111	2.607
	$Set2_{fuzzy}$	0.983	0.857	10.329	2.874	2.021	2.498
	$Set3_{precise}$	0.985	0.860	9.102	2.427	1.670	2.075
	$Set3_{fuzzy}$	0.995	0.857	3.004	1.287	0.842	1.048
	$Set4_{precise}$	0.988	0.852	7.318	2.176	1.474	1.838
	$Set4_{fuzzy}$	0.986	0.851	8.424	2.377	1.660	2.048
	$Set5_{precise}$	0.986	0.862	8.808	2.254	1.584	1.944
	$Set5_{fuzzy}$	0.988	0.866	7.600	2.117	1.463	1.809
XG	$Set1$	0.991	0.877	5.538	1.956	1.429	1.791
	$Set2_{precise}$	0.984	0.855	9.630	2.793	1.927	2.356
	$Set2_{fuzzy}$	0.986	0.856	8.638	2.520	1.735	2.124
	$Set3_{precise}$	0.979	0.856	12.821	2.974	2.114	2.589
	$Set3_{fuzzy}$	0.996	0.850	2.393	1.247	0.855	1.037

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
	Set4$_{precise}$	0.987	0.850	7.927	2.477	1.690	2.090
	Set4$_{fuzzy}$	0.989	0.844	6.722	2.123	1.467	1.806
	Set5$_{precise}$	0.975	0.862	15.510	3.650	2.627	3.198
	Set5$_{fuzzy}$	0.984	0.865	9.919	2.722	1.919	2.340
LB	*Set1*	0.979	0.871	13.718	3.540	2.601	3.309
	Set2$_{precise}$	0.976	0.841	14.612	3.308	2.379	2.931
	Set2$_{fuzzy}$	0.984	0.847	9.763	2.672	1.913	2.391
	Set3$_{precise}$	0.976	0.852	14.749	3.261	2.338	2.882
	Set3$_{fuzzy}$	0.982	0.853	10.653	2.726	1.918	2.392
	Set4$_{precise}$	0.972	0.846	16.669	3.865	2.817	3.487
	Set4$_{fuzzy}$	0.976	0.844	14.589	3.543	2.518	3.129
	Set5$_{precise}$	0.981	0.857	11.529	2.914	2.067	2.566
	Set5$_{fuzzy}$	0.968	0.855	19.211	3.849	2.753	3.397
SVM	*Set1*	0.900	0.862	64.371	8.014	5.742	6.905
	Set2$_{precise}$	0.903	0.862	58.473	7.635	5.275	6.257
	Set2$_{fuzzy}$	0.895	0.851	63.089	7.936	5.594	6.596
	Set3$_{precise}$	0.910	0.869	54.154	7.349	5.117	6.123
	Set3$_{fuzzy}$	0.901	0.858	59.486	7.706	5.436	6.468
	Set4$_{precise}$	0.910	0.870	54.276	7.358	5.123	6.137
	Set4$_{fuzzy}$	0.901	0.859	59.951	7.737	5.479	6.524
	Set5$_{precise}$	0.905	0.870	57.155	7.549	5.309	6.301
	Set5$_{fuzzy}$	0.898	0.860	61.411	7.828	5.547	6.547
ANN	*Set1*	0.914	0.857	55.217	7.398	5.605	6.809
	Set2$_{precise}$	0.916	0.849	50.726	7.075	5.203	6.214
	Set2$_{fuzzy}$	0.912	0.842	53.036	7.247	5.382	6.405
	Set3$_{precise}$	0.924	0.862	46.222	6.733	4.964	5.959
	Set3$_{fuzzy}$	0.910	0.858	54.260	7.342	5.454	6.491
	Set4$_{precise}$	0.920	0.862	48.397	6.914	5.080	6.086
	Set4$_{fuzzy}$	0.916	0.854	50.805	7.090	5.262	6.283
	Set5$_{precise}$	0.915	0.860	51.212	7.114	5.213	6.234
	Set5$_{fuzzy}$	0.910	0.856	54.312	7.342	5.411	6.426
Ridge	*Set1*	0.866	0.842	86.315	9.288	7.004	8.561
	Set2$_{precise}$	0.870	0.844	78.603	8.861	6.746	8.191
	Set2$_{fuzzy}$	0.865	0.839	81.580	9.027	6.944	8.384
	Set3$_{precise}$	0.879	0.853	72.818	8.529	6.512	7.959
	Set3$_{fuzzy}$	0.874	0.847	76.048	8.716	6.690	8.141
	Set4$_{precise}$	0.878	0.851	73.870	8.591	6.541	7.997
	Set4$_{fuzzy}$	0.872	0.846	76.952	8.768	6.731	8.177
	Set5$_{precise}$	0.879	0.855	73.221	8.552	6.522	7.973
	Set5$_{fuzzy}$	0.873	0.850	76.678	8.752	6.703	8.148
LASSO	*Set1*	0.865	0.842	87.140	9.332	7.023	8.576
	Set2$_{precise}$	0.869	0.843	78.883	8.876	6.752	8.189
	Set2$_{fuzzy}$	0.864	0.838	81.756	9.037	6.950	8.384
	Set3$_{precise}$	0.878	0.852	73.581	8.573	6.525	7.966
	Set3$_{fuzzy}$	0.872	0.848	77.013	8.771	6.702	8.135
	Set4$_{precise}$	0.877	0.850	74.067	8.602	6.544	7.999
	Set4$_{fuzzy}$	0.872	0.845	77.229	8.784	6.740	8.181
	Set5$_{precise}$	0.878	0.854	73.626	8.576	6.533	7.981
	Set5$_{fuzzy}$	0.872	0.849	77.215	8.782	6.711	8.148

Table A8. The fit performance of eleven machine learning models for ship S2 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.833	0.668	113.854	10.580	7.934	8.951
	$AIS2_{precise}$	0.848	0.647	95.131	9.497	7.026	8.169
	$AIS2_{fuzzy}$	0.850	0.635	94.066	9.585	7.155	8.275
	$AIS3_{precise}$	0.832	0.625	105.279	10.063	7.466	8.761
	$AIS3_{fuzzy}$	0.835	0.616	103.114	10.046	7.495	8.671
	$AIS4_{precise}$	0.854	0.634	91.071	9.389	6.915	8.058
	$AIS4_{fuzzy}$	0.832	0.638	104.706	10.121	7.549	8.763
	$AIS5_{precise}$	0.838	0.641	100.776	9.878	7.326	8.621
	$AIS5_{fuzzy}$	0.844	0.645	98.151	9.776	7.276	8.406
	$Set3_{precise}$ [a]	0.820	0.589	112.089	10.461	7.916	9.230
ET	*Set1*	0.971	0.786	19.857	4.055	2.986	3.306
	$AIS2_{precise}$	0.966	0.769	21.080	4.247	3.184	3.681
	$AIS2_{fuzzy}$	0.958	0.765	26.376	5.012	3.869	4.454
	$AIS3_{precise}$	0.973	0.775	16.741	3.697	2.653	3.048
	$AIS3_{fuzzy}$	0.968	0.769	19.826	4.047	2.961	3.398
	$AIS4_{precise}$	0.969	0.767	19.271	4.076	3.015	3.476
	$AIS4_{fuzzy}$	0.962	0.764	23.481	4.488	3.310	3.817
	$AIS5_{precise}$	0.979	0.776	13.486	3.239	2.390	2.760
	$AIS5_{fuzzy}$ [a]	0.952	0.766	29.638	5.037	3.783	4.339
	$Set3_{precise}$	0.974	0.765	15.842	3.377	2.445	2.780
RF	*Set1*	0.959	0.766	27.622	5.205	3.750	4.227
	$AIS2_{precise}$	0.947	0.753	32.873	5.683	4.092	4.765
	$AIS2_{fuzzy}$	0.955	0.751	27.856	5.237	3.790	4.396
	$AIS3_{precise}$	0.955	0.757	28.055	5.259	3.774	4.415
	$AIS3_{fuzzy}$	0.945	0.750	34.278	5.766	4.051	4.704
	$AIS4_{precise}$	0.951	0.753	30.772	5.477	3.879	4.491
	$AIS4_{fuzzy}$	0.946	0.748	33.751	5.721	4.078	4.720
	$AIS5_{precise}$	0.948	0.757	32.466	5.645	4.072	4.748
	$AIS5_{fuzzy}$	0.949	0.754	31.822	5.577	3.975	4.637
	$Set3_{precise}$ [a]	0.950	0.740	31.494	5.541	4.007	4.662
AB	*Set1*	0.968	0.762	21.779	4.305	3.609	4.143
	$AIS2_{precise}$	0.980	0.748	12.278	3.280	2.750	3.190
	$AIS2_{fuzzy}$	0.973	0.739	16.514	3.821	3.216	3.725
	$AIS3_{precise}$	0.980	0.749	12.642	3.279	2.739	3.205
	$AIS3_{fuzzy}$	0.977	0.746	14.184	3.472	2.922	3.409
	$AIS4_{precise}$	0.980	0.748	12.489	3.353	2.804	3.262
	$AIS4_{fuzzy}$	0.976	0.738	14.841	3.491	2.913	3.390
	$AIS5_{precise}$	0.975	0.754	15.834	3.830	3.227	3.765
	$AIS5_{fuzzy}$	0.971	0.754	18.197	4.128	3.549	4.162
	$Set3_{precise}$ [a]	0.961	0.743	24.755	4.778	4.073	4.729
GB	*Set1*	0.964	0.781	24.457	4.564	3.429	3.793
	$AIS2_{precise}$	0.980	0.764	12.651	3.166	2.496	2.842
	$AIS2_{fuzzy}$	0.964	0.749	22.561	4.354	3.304	3.719
	$AIS3_{precise}$	0.979	0.773	12.602	2.905	2.149	2.421
	$AIS3_{fuzzy}$	0.980	0.760	12.341	3.100	2.236	2.502
	$AIS4_{precise}$	0.988	0.772	7.589	2.274	1.677	1.884
	$AIS4_{fuzzy}$	0.973	0.763	16.682	3.628	2.707	3.057
	$AIS5_{precise}$	0.964	0.766	22.248	4.288	3.136	3.537
	$AIS5_{fuzzy}$	0.967	0.759	20.448	4.297	3.300	3.722
	$Set3_{precise}$ [a]	0.992	0.760	5.008	1.817	1.234	1.378

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.975	0.781	16.733	3.503	2.631	2.868
	$AIS2_{precise}$	0.959	0.757	25.467	4.457	3.004	3.294
	$AIS2_{fuzzy}$	0.962	0.754	23.570	4.567	3.274	3.615
	$AIS3_{precise}$	0.976	0.772	15.247	3.278	2.139	2.359
	$AIS3_{fuzzy}$	0.978	0.768	13.808	3.206	2.178	2.390
	$AIS4_{precise}$	0.966	0.765	20.767	4.157	2.743	3.008
	$AIS4_{fuzzy}$	0.960	0.762	24.749	4.576	3.229	3.544
	$AIS5_{precise}$	0.965	0.763	21.869	4.287	2.959	3.263
	$AIS5_{fuzzy}$	0.953	0.758	29.132	5.014	3.622	3.995
	$Set3_{precise}$ [a]	0.991	0.765	5.421	1.949	1.186	1.277
LB	*Set1*	0.946	0.761	36.850	5.784	4.429	4.834
	$AIS2_{precise}$	0.951	0.738	30.039	5.172	3.733	4.200
	$AIS2_{fuzzy}$	0.941	0.727	36.987	5.955	4.353	4.919
	$AIS3_{precise}$	0.972	0.745	17.363	3.757	2.709	3.079
	$AIS3_{fuzzy}$	0.961	0.730	24.200	4.627	3.333	3.789
	$AIS4_{precise}$	0.959	0.683	25.467	4.524	3.348	3.862
	$AIS4_{fuzzy}$	0.957	0.716	27.166	4.917	3.685	4.188
	$AIS5_{precise}$	0.959	0.733	25.714	4.671	3.276	3.798
	$AIS5_{fuzzy}$	0.941	0.733	36.457	5.684	4.108	4.702
	$Set3_{precise}$ [a]	0.980	0.748	12.589	3.053	2.179	2.442
SVM	*Set1*	0.848	0.797	103.306	10.147	7.260	7.779
	$AIS2_{precise}$	0.878	0.812	76.193	8.710	6.221	6.788
	$AIS2_{fuzzy}$	0.873	0.808	79.392	8.896	6.332	6.952
	$AIS3_{precise}$	0.875	0.816	78.296	8.829	6.359	6.937
	$AIS3_{fuzzy}$	0.871	0.810	80.452	8.950	6.470	7.131
	$AIS4_{precise}$	0.875	0.815	78.376	8.834	6.364	6.962
	$AIS4_{fuzzy}$	0.874	0.808	78.811	8.862	6.358	7.007
	$AIS5_{precise}$	0.880	0.809	74.547	8.599	6.191	6.738
	$AIS5_{fuzzy}$	0.869	0.804	81.630	9.012	6.541	7.180
	$Set3_{precise}$ [a]	0.864	0.812	84.860	9.176	6.608	7.210
ANN	*Set1*	0.876	0.787	84.367	9.093	6.935	7.682
	$AIS2_{precise}$	0.890	0.805	68.589	8.223	6.248	6.993
	$AIS2_{fuzzy}$	0.884	0.803	72.222	8.473	6.444	7.232
	$AIS3_{precise}$	0.893	0.815	66.528	8.111	6.172	6.926
	$AIS3_{fuzzy}$	0.891	0.811	67.646	8.181	6.182	6.952
	$AIS4_{precise}$	0.894	0.818	66.090	8.068	6.141	6.914
	$AIS4_{fuzzy}$	0.886	0.817	71.270	8.405	6.354	7.172
	$AIS5_{precise}$	0.895	0.797	65.299	8.036	6.139	6.847
	$AIS5_{fuzzy}$	0.887	0.802	70.576	8.362	6.365	7.100
	$Set3_{precise}$ [a]	0.908	0.791	56.693	7.365	5.581	6.171
Ridge	*Set1*	0.822	0.786	121.419	11.016	8.454	9.312
	$AIS2_{precise}$	0.829	0.810	107.128	10.345	7.775	8.763
	$AIS2_{fuzzy}$	0.822	0.803	110.983	10.529	7.876	8.873
	$AIS3_{precise}$	0.837	0.810	102.096	10.100	7.682	8.732
	$AIS3_{fuzzy}$	0.833	0.808	104.294	10.208	7.715	8.787
	$AIS4_{precise}$	0.836	0.813	102.598	10.125	7.695	8.759
	$AIS4_{fuzzy}$	0.833	0.809	104.538	10.220	7.723	8.798
	$AIS5_{precise}$	0.829	0.810	106.864	10.332	7.780	8.791
	$AIS5_{fuzzy}$	0.823	0.803	110.694	10.515	7.877	8.895
	$Set3_{precise}$ [a]	0.826	0.802	108.847	10.429	8.011	9.055
	Set1	0.822	0.785	121.508	11.020	8.471	9.331

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
LASSO	$AIS2_{precise}$	0.829	0.810	107.127	10.344	7.774	8.748
	$AIS2_{fuzzy}$	0.822	0.802	111.266	10.542	7.882	8.861
	$AIS3_{precise}$	0.836	0.809	102.493	10.119	7.676	8.700
	$AIS3_{fuzzy}$	0.831	0.809	105.332	10.259	7.719	8.779
	$AIS4_{precise}$	0.836	0.812	102.731	10.131	7.685	8.720
	$AIS4_{fuzzy}$	0.832	0.809	105.148	10.250	7.714	8.765
	$AIS5_{precise}$	0.829	0.810	106.939	10.336	7.779	8.774
	$AIS5_{fuzzy}$	0.823	0.803	110.874	10.524	7.884	8.887
	$Set3_{precise}$ [a]	0.824	0.796	110.162	10.492	8.034	9.042

Note: $Set3_{precise}$ is the best dataset with DFS1.

Table A9. The fit performance of eleven machine learning models for ship S3 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.857	0.684	105.672	10.125	7.259	8.643
	$AIS2_{precise}$	0.854	0.699	106.123	10.087	7.292	8.759
	$AIS2_{fuzzy}$	0.850	0.695	108.384	10.209	7.427	8.876
	$AIS3_{precise}$	0.871	0.706	93.959	9.334	6.759	8.074
	$AIS3_{fuzzy}$	0.874	0.681	90.614	9.273	6.729	7.972
	$AIS4_{precise}$	0.867	0.692	97.233	9.628	6.941	8.311
	$AIS4_{fuzzy}$	0.869	0.687	94.625	9.621	7.005	8.343
	$AIS5_{precise}$	0.867	0.684	97.066	9.534	6.948	8.222
	$AIS5_{fuzzy}$	0.853	0.704	106.927	10.083	7.267	8.664
	$Set3_{precise}$ [a]	0.865	0.684	98.572	9.705	7.042	8.343
ET	*Set1*	0.977	0.800	17.021	3.911	2.462	2.964
	$AIS2_{precise}$	0.986	0.822	10.267	2.769	1.681	2.113
	$AIS2_{fuzzy}$	0.977	0.814	16.566	3.552	2.175	2.756
	$AIS3_{precise}$	0.982	0.818	12.791	3.030	1.702	2.189
	$AIS3_{fuzzy}$	0.988	0.813	8.588	2.438	1.478	1.867
	$AIS4_{precise}$	0.981	0.816	13.605	3.179	1.861	2.371
	$AIS4_{fuzzy}$	0.981	0.811	13.638	3.326	2.026	2.557
	$AIS5_{precise}$	0.982	0.828	13.029	3.188	1.858	2.364
	$AIS5_{fuzzy}$	0.983	0.826	12.304	2.986	1.751	2.218
	$Set3_{precise}$ [a]	0.985	0.821	10.758	2.846	1.716	2.181
RF	*Set1*	0.960	0.768	29.573	5.369	3.497	4.234
	$AIS2_{precise}$	0.963	0.807	26.977	5.153	3.378	4.167
	$AIS2_{fuzzy}$	0.959	0.803	29.977	5.426	3.478	4.315
	$AIS3_{precise}$	0.951	0.803	35.971	5.896	3.790	4.743
	$AIS3_{fuzzy}$	0.948	0.807	37.592	6.052	3.787	4.741
	$AIS4_{precise}$	0.959	0.797	29.386	5.381	3.477	4.317
	$AIS4_{fuzzy}$	0.954	0.802	33.398	5.680	3.620	4.511
	$AIS5_{precise}$	0.963	0.809	27.121	5.168	3.376	4.170
	$AIS5_{fuzzy}$	0.960	0.805	29.142	5.343	3.415	4.241
	$Set3_{precise}$ [a]	0.956	0.802	31.781	5.576	3.587	4.463
AB	*Set1*	0.988	0.798	9.177	2.942	2.371	2.718
	$AIS2_{precise}$	0.990	0.815	7.319	2.589	2.060	2.335
	$AIS2_{fuzzy}$	0.989	0.805	7.569	2.576	1.975	2.218
	$AIS3_{precise}$	0.995	0.813	3.799	1.682	1.270	1.480
	$AIS3_{fuzzy}$	0.994	0.803	4.303	1.839	1.356	1.568
	$AIS4_{precise}$	0.992	0.810	5.579	2.186	1.699	1.933
	$AIS4_{fuzzy}$	0.989	0.801	7.743	2.614	2.101	2.367
	$AIS5_{precise}$	0.995	0.820	3.588	1.728	1.292	1.513
	$AIS5_{fuzzy}$	0.995	0.802	3.489	1.623	1.154	1.330
	$Set3_{precise}$ [a]	0.991	0.812	6.328	2.183	1.712	1.998
GB	*Set1*	0.962	0.776	28.220	4.726	3.221	3.841
	$AIS2_{precise}$	0.966	0.815	25.195	4.467	2.872	3.617
	$AIS2_{fuzzy}$	0.962	0.798	27.998	4.806	3.066	3.870
	$AIS3_{precise}$	0.974	0.817	18.985	3.721	2.311	2.932
	$AIS3_{fuzzy}$	0.968	0.810	23.467	4.597	2.809	3.582
	$AIS4_{precise}$	0.960	0.813	29.026	4.863	3.092	3.881
	$AIS4_{fuzzy}$	0.966	0.811	25.102	4.373	2.668	3.431
	$AIS5_{precise}$	0.969	0.818	22.552	4.221	2.710	3.366
	$AIS5_{fuzzy}$	0.971	0.817	21.037	4.028	2.586	3.238
	$Set3_{precise}$ [a]	0.964	0.819	26.559	4.694	2.836	3.642

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.959	0.778	30.013	4.738	3.214	3.744
	$AIS2_{precise}$	0.961	0.814	28.566	4.798	2.995	3.753
	$AIS2_{fuzzy}$	0.942	0.807	42.158	5.988	3.915	4.873
	$AIS3_{precise}$	0.964	0.809	26.403	4.516	2.727	3.416
	$AIS3_{fuzzy}$	0.969	0.806	22.552	4.331	2.511	3.237
	$AIS4_{precise}$	0.954	0.809	33.595	5.384	3.366	4.179
	$AIS4_{fuzzy}$	0.961	0.806	28.229	5.078	3.093	3.892
	$AIS5_{precise}$	0.976	0.819	17.745	3.884	2.439	2.960
	$AIS5_{fuzzy}$	0.959	0.808	30.356	4.995	3.159	3.921
	$Set3_{precise}$ [a]	0.961	0.810	28.714	5.030	3.052	3.828
LB	*Set1*	0.935	0.766	48.608	6.560	4.506	5.448
	$AIS2_{precise}$	0.945	0.802	40.487	5.912	3.868	4.848
	$AIS2_{fuzzy}$	0.925	0.802	53.776	7.068	4.742	5.928
	$AIS3_{precise}$	0.954	0.802	33.337	5.389	3.474	4.419
	$AIS3_{fuzzy}$	0.952	0.795	34.420	5.268	3.534	4.423
	$AIS4_{precise}$	0.938	0.801	44.899	6.546	4.381	5.462
	$AIS4_{fuzzy}$	0.947	0.784	38.969	5.866	3.946	4.969
	$AIS5_{precise}$	0.952	0.804	35.323	5.432	3.468	4.412
	$AIS5_{fuzzy}$	0.953	0.796	34.216	5.421	3.563	4.487
	$Set3_{precise}$ [a]	0.947	0.804	38.795	5.845	3.853	4.853
SVM	*Set1*	0.812	0.791	138.669	11.753	7.557	8.957
	$AIS2_{precise}$	0.842	0.825	114.633	10.675	6.588	8.092
	$AIS2_{fuzzy}$	0.829	0.817	123.943	11.116	6.904	8.485
	$AIS3_{precise}$	0.847	0.823	110.536	10.488	6.537	8.005
	$AIS3_{fuzzy}$	0.842	0.816	114.255	10.658	6.761	8.329
	$AIS4_{precise}$	0.845	0.828	112.062	10.563	6.588	8.077
	$AIS4_{fuzzy}$	0.842	0.822	114.254	10.668	6.734	8.244
	$AIS5_{precise}$	0.840	0.825	115.821	10.727	6.746	8.262
	$AIS5_{fuzzy}$	0.840	0.818	115.603	10.720	6.657	8.187
	$Set3_{precise}$ [a]	0.844	0.820	113.000	10.591	6.627	8.167
ANN	*Set1*	0.829	0.780	126.769	11.217	7.780	9.353
	$AIS2_{precise}$	0.849	0.812	109.169	10.413	7.022	8.680
	$AIS2_{fuzzy}$	0.849	0.813	109.194	10.415	6.917	8.561
	$AIS3_{precise}$	0.869	0.800	94.542	9.681	6.645	8.156
	$AIS3_{fuzzy}$	0.847	0.799	111.263	10.484	7.230	8.936
	$AIS4_{precise}$	0.860	0.799	101.451	10.008	6.840	8.418
	$AIS4_{fuzzy}$	0.847	0.800	110.709	10.470	7.192	8.881
	$AIS5_{precise}$	0.860	0.813	101.322	10.029	6.763	8.262
	$AIS5_{fuzzy}$	0.854	0.810	106.288	10.265	6.849	8.413
	$Set3_{precise}$ [a]	0.874	0.798	91.583	9.475	6.480	7.992
Ridge	*Set1*	0.780	0.778	162.676	12.739	9.007	11.114
	$AIS2_{precise}$	0.793	0.801	149.861	12.227	8.517	10.890
	$AIS2_{fuzzy}$	0.790	0.799	152.029	12.316	8.509	10.860
	$AIS3_{precise}$	0.801	0.797	143.924	11.981	8.319	10.627
	$AIS3_{fuzzy}$	0.802	0.798	143.536	11.966	8.330	10.651
	$AIS4_{precise}$	0.798	0.799	145.868	12.062	8.348	10.695
	$AIS4_{fuzzy}$	0.799	0.800	145.189	12.035	8.339	10.671
	$AIS5_{precise}$	0.796	0.804	147.564	12.133	8.425	10.745
	$AIS5_{fuzzy}$	0.793	0.802	149.516	12.214	8.459	10.779
	$Set3_{precise}$ [a]	0.801	0.796	144.061	11.987	8.329	10.615
	Set1	0.779	0.778	163.445	12.769	9.011	11.128

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
LASSO	$AIS2_{precise}$	0.793	0.800	149.869	12.227	8.508	10.879
	$AIS2_{fuzzy}$	0.790	0.799	152.048	12.317	8.502	10.852
	$AIS3_{precise}$	0.800	0.798	145.050	12.028	8.317	10.627
	$AIS3_{fuzzy}$	0.800	0.798	144.547	12.007	8.322	10.641
	$AIS4_{precise}$	0.798	0.799	146.172	12.075	8.346	10.694
	$AIS4_{fuzzy}$	0.799	0.799	145.246	12.037	8.329	10.657
	$AIS5_{precise}$	0.796	0.804	147.612	12.135	8.426	10.746
	$AIS5_{fuzzy}$	0.793	0.802	149.539	12.215	8.450	10.765
	$Set3_{precise}$ [a]	0.799	0.796	145.425	12.043	8.323	10.619

Note: $Set3_{precise}$ is the best dataset with DFS1.

Table A10. The fit performance of eleven machine learning models for ship S4 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.906	0.758	81.312	8.851	6.363	6.681
	$AIS2_{precise}$	0.916	0.736	69.034	8.033	5.999	6.463
	$AIS2_{fuzzy}$	0.910	0.730	73.824	8.275	6.130	6.586
	$AIS3_{precise}$	0.916	0.741	68.065	8.086	6.006	6.432
	$AIS3_{fuzzy}$	0.926	0.725	60.675	7.518	5.564	5.919
	$AIS4_{precise}$	0.921	0.745	64.090	7.807	5.747	6.160
	$AIS4_{fuzzy}$	0.899	0.741	82.735	8.913	6.661	7.152
	$AIS5_{precise}$	0.904	0.749	78.681	8.637	6.425	6.897
	$AIS5_{fuzzy}$	0.912	0.758	72.400	8.366	6.185	6.648
	$Set3_{precise}$ [a]	0.916	0.746	68.063	8.094	6.036	6.523
ET	*Set1*	0.988	0.858	10.120	2.625	1.778	1.862
	$AIS2_{precise}$	0.999	0.867	0.705	0.629	0.394	0.422
	$AIS2_{fuzzy}$	0.999	0.863	0.811	0.714	0.462	0.495
	$AIS3_{precise}$	0.997	0.869	2.131	1.151	0.807	0.876
	$AIS3_{fuzzy}$	0.997	0.864	2.431	1.115	0.779	0.846
	$AIS4_{precise}$	0.997	0.870	2.521	1.191	0.845	0.915
	$AIS4_{fuzzy}$	0.998	0.864	1.826	1.115	0.768	0.830
	$AIS5_{precise}$	0.998	0.871	1.642	0.927	0.651	0.696
	$AIS5_{fuzzy}$	0.998	0.866	1.390	0.973	0.659	0.716
	$Set3_{precise}$ [a]	0.998	0.872	1.434	0.901	0.627	0.687
RF	*Set1*	0.974	0.848	22.794	4.752	3.335	3.501
	$AIS2_{precise}$	0.976	0.856	19.824	4.444	3.279	3.600
	$AIS2_{fuzzy}$	0.976	0.854	19.766	4.429	3.273	3.607
	$AIS3_{precise}$	0.972	0.854	22.477	4.723	3.464	3.773
	$AIS3_{fuzzy}$	0.973	0.855	21.949	4.659	3.424	3.751
	$AIS4_{precise}$	0.973	0.855	21.804	4.644	3.421	3.742
	$AIS4_{fuzzy}$	0.974	0.858	21.363	4.599	3.381	3.708
	$AIS5_{precise}$	0.975	0.859	20.097	4.472	3.292	3.590
	$AIS5_{fuzzy}$	0.975	0.858	20.309	4.491	3.298	3.602
	$Set3_{precise}$ [a]	0.975	0.853	20.349	4.497	3.331	3.618
AB	*Set1*	0.980	0.843	17.332	3.939	3.283	3.654
	$AIS2_{precise}$	0.987	0.865	10.568	3.102	2.595	2.916
	$AIS2_{fuzzy}$	0.985	0.862	11.992	3.358	2.790	3.103
	$AIS3_{precise}$	0.990	0.867	8.454	2.664	2.206	2.484
	$AIS3_{fuzzy}$	0.989	0.867	8.670	2.766	2.285	2.560
	$AIS4_{precise}$	0.988	0.864	9.894	2.953	2.477	2.782
	$AIS4_{fuzzy}$	0.987	0.864	10.576	3.117	2.607	2.916
	$AIS5_{precise}$	0.988	0.869	9.701	2.956	2.466	2.774
	$AIS5_{fuzzy}$	0.988	0.868	9.410	2.864	2.401	2.674
	$Set3_{precise}$ [a]	0.986	0.865	11.021	3.144	2.591	2.905
GB	*Set1*	0.977	0.851	19.591	4.196	3.176	3.352
	$AIS2_{precise}$	0.989	0.868	9.077	2.587	1.960	2.099
	$AIS2_{fuzzy}$	0.988	0.868	9.667	2.653	2.046	2.189
	$AIS3_{precise}$	0.995	0.868	4.407	1.836	1.343	1.452
	$AIS3_{fuzzy}$	0.991	0.871	7.288	2.277	1.677	1.808
	$AIS4_{precise}$	0.993	0.866	5.462	1.884	1.438	1.531
	$AIS4_{fuzzy}$	0.991	0.863	6.981	2.207	1.677	1.803
	$AIS5_{precise}$	0.991	0.873	7.631	2.412	1.810	1.933
	$AIS5_{fuzzy}$	0.990	0.875	8.437	2.566	1.943	2.087
	$Set3_{precise}$ [a]	0.989	0.866	8.845	2.500	1.838	1.957

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.977	0.858	19.657	4.126	3.068	3.185
	$AIS2_{precise}$	0.991	0.864	7.508	2.341	1.721	1.808
	$AIS2_{fuzzy}$	0.992	0.870	6.889	2.321	1.715	1.816
	$AIS3_{precise}$	0.994	0.868	4.689	1.591	1.126	1.196
	$AIS3_{fuzzy}$	0.995	0.873	4.260	1.605	1.117	1.188
	$AIS4_{precise}$	0.993	0.861	5.846	1.821	1.344	1.426
	$AIS4_{fuzzy}$	0.989	0.865	8.965	2.311	1.733	1.832
	$AIS5_{precise}$	0.992	0.870	6.441	2.074	1.548	1.640
	$AIS5_{fuzzy}$	0.987	0.878	11.059	2.835	2.157	2.299
	$Set3_{precise}$ [a]	0.995	0.869	3.758	1.585	1.140	1.201
LB	*Set1*	0.968	0.844	28.153	5.010	3.861	4.044
	$AIS2_{precise}$	0.979	0.851	17.028	3.711	2.865	3.064
	$AIS2_{fuzzy}$	0.988	0.854	9.860	2.778	2.152	2.313
	$AIS3_{precise}$	0.989	0.866	9.039	2.625	2.009	2.161
	$AIS3_{fuzzy}$	0.986	0.865	11.633	3.097	2.388	2.564
	$AIS4_{precise}$	0.978	0.855	18.105	3.969	3.112	3.336
	$AIS4_{fuzzy}$	0.981	0.846	15.360	3.672	2.862	3.083
	$AIS5_{precise}$	0.992	0.865	6.859	2.321	1.771	1.925
	$AIS5_{fuzzy}$	0.987	0.873	10.530	2.901	2.250	2.440
	$Set3_{precise}$ [a]	0.987	0.855	10.943	2.871	2.200	2.340
SVM	*Set1*	0.906	0.842	81.874	9.015	6.318	6.374
	$AIS2_{precise}$	0.930	0.848	56.850	7.333	5.462	5.814
	$AIS2_{fuzzy}$	0.936	0.845	51.762	6.893	5.122	5.485
	$AIS3_{precise}$	0.929	0.846	57.822	7.548	5.488	5.811
	$AIS3_{fuzzy}$	0.917	0.842	68.045	8.187	5.956	6.291
	$AIS4_{precise}$	0.926	0.850	60.077	7.704	5.660	5.977
	$AIS4_{fuzzy}$	0.917	0.846	67.802	8.170	5.974	6.346
	$AIS5_{precise}$	0.927	0.860	59.875	7.686	5.642	6.018
	$AIS5_{fuzzy}$	0.919	0.854	66.517	8.094	5.967	6.387
	$Set3_{precise}$ [a]	0.921	0.857	63.718	7.972	5.848	6.146
ANN	*Set1*	0.925	0.845	65.521	8.076	6.102	6.390
	$AIS2_{precise}$	0.936	0.862	52.288	7.217	5.653	6.104
	$AIS2_{fuzzy}$	0.936	0.842	52.055	7.184	5.607	6.059
	$AIS3_{precise}$	0.944	0.859	46.191	6.779	5.331	5.782
	$AIS3_{fuzzy}$	0.941	0.848	48.047	6.911	5.366	5.842
	$AIS4_{precise}$	0.943	0.856	46.674	6.812	5.354	5.803
	$AIS4_{fuzzy}$	0.943	0.852	46.629	6.809	5.277	5.720
	$AIS5_{precise}$	0.939	0.868	50.219	7.062	5.524	5.962
	$AIS5_{fuzzy}$	0.930	0.858	57.295	7.541	5.889	6.371
	$Set3_{precise}$ [a]	0.947	0.856	42.555	6.513	5.034	5.502
Ridge	*Set1*	0.825	0.821	152.631	12.351	9.343	9.548
	$AIS2_{precise}$	0.822	0.802	145.669	12.066	9.341	9.625
	$AIS2_{fuzzy}$	0.816	0.796	150.467	12.263	9.500	9.837
	$AIS3_{precise}$	0.832	0.807	137.392	11.717	9.182	9.497
	$AIS3_{fuzzy}$	0.826	0.802	141.931	11.909	9.280	9.647
	$AIS4_{precise}$	0.831	0.806	138.025	11.744	9.187	9.487
	$AIS4_{fuzzy}$	0.826	0.802	142.348	11.926	9.279	9.628
	$AIS5_{precise}$	0.827	0.808	141.409	11.888	9.244	9.534
	$AIS5_{fuzzy}$	0.822	0.803	145.201	12.046	9.355	9.690
	$Set3_{precise}$ [a]	0.833	0.811	135.334	11.629	9.033	9.406
	Set1	0.824	0.823	153.402	12.382	9.347	9.537

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
LASSO	$AIS2_{precise}$	0.822	0.803	145.771	12.070	9.340	9.620
	$AIS2_{fuzzy}$	0.815	0.797	150.852	12.279	9.508	9.835
	$AIS3_{precise}$	0.831	0.806	137.956	11.741	9.191	9.499
	$AIS3_{fuzzy}$	0.824	0.804	143.555	11.977	9.303	9.641
	$AIS4_{precise}$	0.831	0.806	138.364	11.759	9.193	9.489
	$AIS4_{fuzzy}$	0.825	0.802	143.263	11.964	9.292	9.626
	$AIS5_{precise}$	0.827	0.808	141.586	11.895	9.244	9.530
	$AIS5_{fuzzy}$	0.822	0.805	145.687	12.067	9.364	9.692
	$Set3_{precise}$ [a]	0.832	0.809	135.961	11.656	9.053	9.417

Note: $Set3_{precise}$ is the best dataset with DFS1.

Table A11. The fit performance of eleven machine learning models for ship S5 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.939	0.821	33.699	5.588	4.144	6.259
	$AIS2_{precise}$	0.947	0.795	29.259	5.181	3.809	5.761
	$AIS2_{fuzzy}$	0.946	0.799	30.200	5.379	3.954	5.947
	$AIS3_{precise}$	0.942	0.793	31.964	5.502	4.072	6.137
	$AIS3_{fuzzy}$	0.941	0.812	32.694	5.643	4.181	6.328
	$AIS4_{precise}$	0.940	0.793	33.054	5.628	4.130	6.221
	$AIS4_{fuzzy}$	0.941	0.800	32.814	5.558	4.121	6.247
	$AIS5_{precise}$	0.948	0.800	28.458	5.104	3.741	5.634
	$AIS5_{fuzzy}$	0.945	0.826	30.556	5.217	3.848	5.813
	$Set3_{precise}$ [a]	0.947	0.785	29.488	5.182	3.764	5.625
ET	*Set1*	0.998	0.895	1.057	0.805	0.569	0.857
	$AIS2_{precise}$	0.999	0.896	0.493	0.529	0.371	0.559
	$AIS2_{fuzzy}$	0.998	0.893	1.165	0.835	0.604	0.914
	$AIS3_{precise}$	0.998	0.895	0.890	0.707	0.495	0.748
	$AIS3_{fuzzy}$	0.997	0.893	1.963	1.134	0.815	1.234
	$AIS4_{precise}$	0.998	0.895	1.105	0.768	0.550	0.842
	$AIS4_{fuzzy}$	0.998	0.893	1.248	0.820	0.579	0.879
	$AIS5_{precise}$	0.997	0.894	1.475	0.901	0.652	0.988
	$AIS5_{fuzzy}$	0.998	0.891	1.055	0.808	0.584	0.883
	$Set3_{precise}$ [a]	0.997	0.892	1.413	0.854	0.619	0.935
RF	*Set1*	0.982	0.884	9.951	3.140	2.354	3.594
	$AIS2_{precise}$	0.982	0.878	9.753	3.116	2.290	3.509
	$AIS2_{fuzzy}$	0.983	0.883	9.496	3.072	2.249	3.461
	$AIS3_{precise}$	0.982	0.879	10.152	3.168	2.328	3.562
	$AIS3_{fuzzy}$	0.983	0.887	9.681	3.103	2.298	3.538
	$AIS4_{precise}$	0.981	0.875	10.565	3.238	2.356	3.594
	$AIS4_{fuzzy}$	0.983	0.885	9.200	3.026	2.224	3.413
	$AIS5_{precise}$	0.983	0.879	9.594	3.090	2.281	3.497
	$AIS5_{fuzzy}$	0.984	0.887	9.003	2.996	2.238	3.464
	$Set3_{precise}$ [a]	0.981	0.874	10.498	3.225	2.390	3.663
AB	*Set1*	0.990	0.895	5.408	2.213	1.830	3.156
	$AIS2_{precise}$	0.994	0.886	3.187	1.671	1.336	2.365
	$AIS2_{fuzzy}$	0.995	0.893	2.917	1.577	1.231	2.199
	$AIS3_{precise}$	0.996	0.892	2.337	1.372	1.073	1.942
	$AIS3_{fuzzy}$	0.997	0.897	1.684	1.136	0.841	1.554
	$AIS4_{precise}$	0.995	0.888	2.854	1.584	1.262	2.249
	$AIS4_{fuzzy}$	0.995	0.894	2.937	1.618	1.277	2.286
	$AIS5_{precise}$	0.995	0.889	2.723	1.476	1.172	2.123
	$AIS5_{fuzzy}$	0.996	0.895	2.299	1.392	1.081	1.951
	$Set3_{precise}$ [a]	0.995	0.886	2.543	1.525	1.209	2.217
GB	*Set1*	0.993	0.895	3.885	1.743	1.360	2.158
	$AIS2_{precise}$	0.996	0.892	2.273	1.265	0.946	1.496
	$AIS2_{fuzzy}$	0.997	0.893	1.801	1.015	0.782	1.257
	$AIS3_{precise}$	0.995	0.890	2.674	1.204	0.879	1.397
	$AIS3_{fuzzy}$	0.996	0.894	2.307	1.143	0.839	1.335
	$AIS4_{precise}$	0.997	0.888	1.936	1.010	0.720	1.132
	$AIS4_{fuzzy}$	0.994	0.893	3.367	1.426	1.047	1.656
	$AIS5_{precise}$	0.997	0.893	1.628	1.102	0.823	1.310
	$AIS5_{fuzzy}$	0.993	0.892	3.733	1.665	1.282	2.046
	$Set3_{precise}$ [a]	0.993	0.887	3.519	1.359	1.021	1.610

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.990	0.892	5.361	1.995	1.520	2.370
	$AIS2_{precise}$	0.997	0.884	1.919	1.190	0.830	1.304
	$AIS2_{fuzzy}$	0.993	0.886	3.701	1.539	1.118	1.750
	$AIS3_{precise}$	0.996	0.885	2.079	1.206	0.857	1.347
	$AIS3_{fuzzy}$	0.991	0.893	4.859	1.785	1.314	2.077
	$AIS4_{precise}$	0.990	0.884	5.595	2.020	1.420	2.218
	$AIS4_{fuzzy}$	0.992	0.888	4.440	1.860	1.368	2.169
	$AIS5_{precise}$	0.991	0.891	4.860	1.909	1.382	2.183
	$AIS5_{fuzzy}$	0.992	0.887	4.417	1.897	1.403	2.206
	$Set3_{precise}$ [a]	0.993	0.878	3.601	1.605	1.133	1.749
LB	*Set1*	0.986	0.879	7.810	2.636	2.028	3.173
	$AIS2_{precise}$	0.987	0.879	7.311	2.490	1.883	2.957
	$AIS2_{fuzzy}$	0.984	0.869	9.040	2.831	2.115	3.308
	$AIS3_{precise}$	0.991	0.882	5.103	2.066	1.528	2.380
	$AIS3_{fuzzy}$	0.985	0.884	8.471	2.658	1.990	3.143
	$AIS4_{precise}$	0.987	0.846	7.256	2.420	1.848	2.954
	$AIS4_{fuzzy}$	0.984	0.878	9.032	2.837	2.146	3.365
	$AIS5_{precise}$	0.991	0.880	4.875	2.049	1.523	2.390
	$AIS5_{fuzzy}$	0.983	0.879	9.198	2.795	2.104	3.278
	$Set3_{precise}$ [a]	0.987	0.873	7.382	2.350	1.758	2.725
SVM	*Set1*	0.931	0.884	38.408	6.173	4.382	6.630
	$AIS2_{precise}$	0.916	0.881	46.567	6.810	4.916	7.414
	$AIS2_{fuzzy}$	0.912	0.886	48.618	6.946	4.984	7.542
	$AIS3_{precise}$	0.912	0.880	48.897	6.986	5.055	7.619
	$AIS3_{fuzzy}$	0.913	0.882	48.270	6.942	5.017	7.597
	$AIS4_{precise}$	0.913	0.878	47.993	6.918	5.019	7.549
	$AIS4_{fuzzy}$	0.915	0.881	47.104	6.851	4.980	7.544
	$AIS5_{precise}$	0.918	0.878	45.274	6.711	4.874	7.385
	$AIS5_{fuzzy}$	0.920	0.880	44.342	6.608	4.801	7.317
	$Set3_{precise}$ [a]	0.916	0.873	46.421	6.785	4.917	7.472
ANN	*Set1*	0.926	0.886	40.737	6.373	4.900	7.545
	$AIS2_{precise}$	0.938	0.876	34.278	5.776	4.398	6.845
	$AIS2_{fuzzy}$	0.943	0.885	31.933	5.592	4.281	6.685
	$AIS3_{precise}$	0.942	0.881	32.259	5.619	4.257	6.634
	$AIS3_{fuzzy}$	0.935	0.881	35.988	5.948	4.524	7.035
	$AIS4_{precise}$	0.940	0.880	33.504	5.733	4.354	6.741
	$AIS4_{fuzzy}$	0.938	0.878	34.189	5.806	4.435	6.882
	$AIS5_{precise}$	0.935	0.878	36.276	5.973	4.538	6.997
	$AIS5_{fuzzy}$	0.939	0.881	33.836	5.787	4.393	6.840
	$Set3_{precise}$ [a]	0.935	0.879	36.157	5.956	4.544	7.075
Ridge	*Set1*	0.875	0.868	69.368	8.325	6.341	9.937
	$AIS2_{precise}$	0.886	0.875	63.247	7.949	5.972	9.145
	$AIS2_{fuzzy}$	0.884	0.874	64.133	8.004	5.987	9.116
	$AIS3_{precise}$	0.892	0.868	60.011	7.743	5.826	8.883
	$AIS3_{fuzzy}$	0.892	0.871	59.890	7.735	5.845	8.893
	$AIS4_{precise}$	0.890	0.872	60.879	7.799	5.867	8.914
	$AIS4_{fuzzy}$	0.890	0.873	60.924	7.802	5.897	8.940
	$AIS5_{precise}$	0.887	0.874	62.515	7.903	5.941	9.040
	$AIS5_{fuzzy}$	0.887	0.875	62.587	7.907	5.952	9.027
	$Set3_{precise}$ [a]	0.889	0.868	61.610	7.846	5.934	9.109
	Set1	0.874	0.868	69.799	8.351	6.357	9.948

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
LASSO	$AIS2_{precise}$	0.886	0.875	63.278	7.951	5.976	9.156
	$AIS2_{fuzzy}$	0.884	0.874	64.249	8.012	5.995	9.128
	$AIS3_{precise}$	0.891	0.869	60.385	7.766	5.839	8.900
	$AIS3_{fuzzy}$	0.891	0.870	60.464	7.771	5.876	8.925
	$AIS4_{precise}$	0.889	0.869	61.278	7.824	5.879	8.934
	$AIS4_{fuzzy}$	0.890	0.873	61.092	7.813	5.909	8.957
	$AIS5_{precise}$	0.887	0.874	62.689	7.914	5.950	9.050
	$AIS5_{fuzzy}$	0.887	0.874	62.829	7.923	5.967	9.032
	$Set3_{precise}$ [a]	0.888	0.868	61.988	7.870	5.953	9.129

Note: $Set3_{precise}$ is the best dataset with DFS1.

Table A12. The fit performance of eleven machine learning models for ship S6 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.837	0.636	67.292	8.143	5.917	7.777
	$AIS2_{precise}$	0.823	0.572	73.999	8.507	6.292	8.362
	$AIS2_{fuzzy}$	0.839	0.584	67.435	8.118	5.944	7.841
	$AIS3_{precise}$	0.847	0.568	63.921	7.913	5.799	7.668
	$AIS3_{fuzzy}$	0.841	0.581	66.494	8.095	5.969	7.923
	$AIS4_{precise}$	0.855	0.571	60.340	7.645	5.580	7.407
	$AIS4_{fuzzy}$	0.833	0.594	69.924	8.261	6.086	8.043
	$AIS5_{precise}$	0.847	0.581	63.834	7.896	5.826	7.701
	$AIS5_{fuzzy}$	0.832	0.591	70.398	8.318	6.145	8.148
	$Set3_{precise}$ [a]	0.832	0.576	69.684	8.275	6.119	8.113
ET	*Set1*	0.985	0.765	6.050	1.928	1.359	1.796
	$AIS2_{precise}$	0.991	0.767	3.917	1.699	1.229	1.633
	$AIS2_{fuzzy}$	0.992	0.765	3.275	1.519	1.124	1.491
	$AIS3_{precise}$	0.988	0.764	4.917	1.561	1.136	1.516
	$AIS3_{fuzzy}$	0.986	0.765	5.657	1.944	1.414	1.886
	$AIS4_{precise}$	0.988	0.762	4.791	1.801	1.318	1.760
	$AIS4_{fuzzy}$	0.987	0.763	5.498	1.871	1.381	1.838
	$AIS5_{precise}$	0.984	0.763	6.604	2.439	1.780	2.368
	$AIS5_{fuzzy}$	0.989	0.759	4.566	1.904	1.413	1.874
	$Set3_{precise}$ [a]	0.979	0.752	8.706	2.743	2.010	2.678
RF	*Set1*	0.956	0.766	18.155	4.225	3.016	4.012
	$AIS2_{precise}$	0.956	0.747	18.527	4.279	3.106	4.133
	$AIS2_{fuzzy}$	0.961	0.753	16.394	4.028	2.959	3.942
	$AIS3_{precise}$	0.954	0.746	19.175	4.366	3.148	4.187
	$AIS3_{fuzzy}$	0.954	0.752	19.080	4.337	3.155	4.211
	$AIS4_{precise}$	0.956	0.747	18.285	4.259	3.068	4.089
	$AIS4_{fuzzy}$	0.958	0.755	17.786	4.194	3.069	4.097
	$AIS5_{precise}$	0.959	0.747	17.057	4.116	2.974	3.950
	$AIS5_{fuzzy}$	0.962	0.751	16.017	3.985	2.922	3.886
	$Set3_{precise}$ [a]	0.953	0.740	19.498	4.382	3.173	4.211
AB	*Set1*	0.969	0.770	12.857	3.481	2.871	4.105
	$AIS2_{precise}$	0.984	0.752	6.749	2.491	2.043	2.988
	$AIS2_{fuzzy}$	0.980	0.759	8.310	2.799	2.346	3.410
	$AIS3_{precise}$	0.985	0.752	6.184	2.309	1.883	2.810
	$AIS3_{fuzzy}$	0.983	0.766	7.124	2.537	2.098	3.091
	$AIS4_{precise}$	0.983	0.755	6.924	2.537	2.113	3.111
	$AIS4_{fuzzy}$	0.975	0.756	10.579	3.112	2.619	3.768
	$AIS5_{precise}$	0.983	0.750	6.959	2.393	1.958	2.888
	$AIS5_{fuzzy}$	0.980	0.759	8.222	2.643	2.186	3.192
	$Set3_{precise}$ [a]	0.980	0.755	8.175	2.647	2.186	3.210
GB	*Set1*	0.965	0.786	14.509	3.538	2.597	3.507
	$AIS2_{precise}$	0.974	0.774	10.728	2.915	2.250	3.046
	$AIS2_{fuzzy}$	0.971	0.778	12.005	3.246	2.535	3.453
	$AIS3_{precise}$	0.968	0.768	13.413	3.271	2.531	3.437
	$AIS3_{fuzzy}$	0.974	0.766	10.865	2.713	2.076	2.823
	$AIS4_{precise}$	0.974	0.768	10.657	3.097	2.384	3.237
	$AIS4_{fuzzy}$	0.970	0.775	12.628	3.386	2.650	3.600
	$AIS5_{precise}$	0.954	0.765	19.294	4.244	3.282	4.469
	$AIS5_{fuzzy}$	0.960	0.765	16.628	3.852	2.995	4.059
	$Set3_{precise}$ [a]	0.971	0.770	11.917	3.111	2.384	3.226

IAMU International Association of Maritime Universities (AMC)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.966	0.786	14.223	3.620	2.692	3.641
	$AIS2_{precise}$	0.971	0.773	11.988	3.174	2.438	3.306
	$AIS2_{fuzzy}$	0.964	0.777	15.126	3.599	2.766	3.757
	$AIS3_{precise}$	0.964	0.762	15.014	3.485	2.673	3.639
	$AIS3_{fuzzy}$	0.965	0.774	14.662	3.570	2.728	3.697
	$AIS4_{precise}$	0.971	0.768	12.290	3.278	2.516	3.416
	$AIS4_{fuzzy}$	0.972	0.779	11.539	3.130	2.381	3.228
	$AIS5_{precise}$	0.948	0.765	21.685	4.533	3.501	4.755
	$AIS5_{fuzzy}$	0.957	0.758	17.993	3.859	2.929	3.955
	$Set3_{precise}$ [a]	0.959	0.771	17.299	3.835	2.890	3.902
LB	*Set1*	0.951	0.773	20.401	4.334	3.285	4.472
	$AIS2_{precise}$	0.957	0.769	18.000	4.086	3.154	4.284
	$AIS2_{fuzzy}$	0.965	0.774	14.771	3.662	2.829	3.877
	$AIS3_{precise}$	0.965	0.751	14.716	3.549	2.709	3.676
	$AIS3_{fuzzy}$	0.961	0.765	16.089	3.729	2.875	3.906
	$AIS4_{precise}$	0.961	0.756	16.273	3.800	2.915	3.961
	$AIS4_{fuzzy}$	0.962	0.764	15.650	3.683	2.832	3.853
	$AIS5_{precise}$	0.956	0.749	18.441	3.987	3.023	4.106
	$AIS5_{fuzzy}$	0.963	0.757	15.608	3.608	2.754	3.761
	$Set3_{precise}$ [a]	0.963	0.754	15.520	3.514	2.682	3.646
SVM	*Set1*	0.838	0.748	67.236	8.175	5.625	7.308
	$AIS2_{precise}$	0.859	0.768	58.991	7.655	5.491	7.220
	$AIS2_{fuzzy}$	0.840	0.762	66.704	8.154	5.866	7.724
	$AIS3_{precise}$	0.849	0.766	63.237	7.922	5.687	7.503
	$AIS3_{fuzzy}$	0.835	0.768	68.880	8.281	5.903	7.754
	$AIS4_{precise}$	0.840	0.769	66.795	8.150	5.817	7.656
	$AIS4_{fuzzy}$	0.842	0.766	66.097	8.114	5.750	7.527
	$AIS5_{precise}$	0.846	0.770	64.572	8.017	5.703	7.479
	$AIS5_{fuzzy}$	0.826	0.774	72.707	8.522	6.072	7.993
	$Set3_{precise}$ [a]	0.843	0.767	65.144	8.045	5.755	7.629
ANN	*Set1*	0.851	0.740	61.550	7.798	5.849	7.715
	$AIS2_{precise}$	0.875	0.763	52.267	7.189	5.479	7.270
	$AIS2_{fuzzy}$	0.855	0.767	60.467	7.758	5.885	7.788
	$AIS3_{precise}$	0.865	0.774	56.395	7.477	5.698	7.549
	$AIS3_{fuzzy}$	0.854	0.776	61.192	7.813	5.943	7.845
	$AIS4_{precise}$	0.868	0.775	55.403	7.419	5.662	7.506
	$AIS4_{fuzzy}$	0.859	0.769	58.976	7.652	5.795	7.653
	$AIS5_{precise}$	0.868	0.765	55.181	7.401	5.673	7.526
	$AIS5_{fuzzy}$	0.854	0.765	60.861	7.781	5.944	7.870
	$Set3_{precise}$ [a]	0.859	0.772	58.184	7.599	5.750	7.603
Ridge	*Set1*	0.758	0.729	100.434	10.018	7.588	10.192
	$AIS2_{precise}$	0.773	0.740	94.924	9.740	7.360	9.841
	$AIS2_{fuzzy}$	0.771	0.738	95.615	9.775	7.376	9.863
	$AIS3_{precise}$	0.787	0.749	89.046	9.434	7.255	9.691
	$AIS3_{fuzzy}$	0.785	0.747	89.799	9.473	7.247	9.679
	$AIS4_{precise}$	0.786	0.750	89.615	9.464	7.237	9.663
	$AIS4_{fuzzy}$	0.784	0.748	90.306	9.500	7.233	9.661
	$AIS5_{precise}$	0.778	0.743	92.805	9.631	7.393	9.895
	$AIS5_{fuzzy}$	0.777	0.741	93.183	9.651	7.389	9.899
	$Set3_{precise}$ [a]	0.775	0.745	93.218	9.652	7.454	9.977
	Set1	0.753	0.724	102.272	10.109	7.629	10.199

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
LASSO	$AIS2_{precise}$	0.772	0.738	95.454	9.767	7.342	9.801
	$AIS2_{fuzzy}$	0.770	0.735	96.087	9.799	7.368	9.835
	$AIS3_{precise}$	0.786	0.747	89.527	9.459	7.238	9.656
	$AIS3_{fuzzy}$	0.784	0.749	90.435	9.507	7.231	9.643
	$AIS4_{precise}$	0.785	0.747	89.820	9.475	7.231	9.647
	$AIS4_{fuzzy}$	0.783	0.749	90.743	9.523	7.206	9.600
	$AIS5_{precise}$	0.777	0.741	93.139	9.648	7.387	9.878
	$AIS5_{fuzzy}$	0.777	0.741	93.454	9.665	7.379	9.870
	$Set3_{precise}$ [a]	0.774	0.744	93.502	9.667	7.443	9.960

Note: $Set3_{precise}$ is the best dataset with DFS1.

Table A13. The fit performance of eleven machine learning models for ship S7 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.828	0.680	69.472	8.260	6.302	8.155
	$AIS2_{precise}$	0.863	0.680	55.328	7.389	5.573	7.264
	$AIS2_{fuzzy}$	0.853	0.670	59.404	7.669	5.797	7.542
	$AIS3_{precise}$	0.856	0.670	58.237	7.557	5.694	7.432
	$AIS3_{fuzzy}$	0.850	0.671	60.824	7.759	5.896	7.686
	$AIS4_{precise}$	0.855	0.667	58.606	7.606	5.720	7.432
	$AIS4_{fuzzy}$	0.856	0.671	58.321	7.622	5.775	7.511
	$AIS5_{precise}$	0.865	0.678	54.511	7.334	5.473	7.099
	$AIS5_{fuzzy}$	0.867	0.669	53.839	7.272	5.433	7.071
	$Set3_{precise}$ [a]	0.880	0.683	48.319	6.903	5.173	6.749
ET	*Set1*	0.956	0.806	17.780	3.880	2.884	3.713
	$AIS2_{precise}$	0.983	0.830	6.818	2.117	1.487	1.925
	$AIS2_{fuzzy}$	0.971	0.819	11.791	3.070	2.178	2.810
	$AIS3_{precise}$	0.979	0.830	8.690	2.410	1.676	2.170
	$AIS3_{fuzzy}$	0.973	0.826	10.900	2.886	2.007	2.593
	$AIS4_{precise}$	0.983	0.829	6.702	2.040	1.405	1.820
	$AIS4_{fuzzy}$	0.976	0.820	9.693	2.688	1.894	2.454
	$AIS5_{precise}$	0.978	0.834	8.811	2.497	1.753	2.266
	$AIS5_{fuzzy}$	0.983	0.828	6.851	2.236	1.591	2.050
	$Set3_{precise}$ [a]	0.987	0.805	5.176	1.848	1.259	1.639
RF	*Set1*	0.964	0.793	14.369	3.774	2.813	3.649
	$AIS2_{precise}$	0.962	0.808	15.226	3.874	2.845	3.726
	$AIS2_{fuzzy}$	0.959	0.806	16.605	4.019	2.964	3.884
	$AIS3_{precise}$	0.958	0.810	16.981	4.069	2.963	3.883
	$AIS3_{fuzzy}$	0.958	0.811	17.018	4.095	2.979	3.891
	$AIS4_{precise}$	0.957	0.809	17.317	4.118	3.023	3.958
	$AIS4_{fuzzy}$	0.957	0.806	17.549	4.147	3.037	3.972
	$AIS5_{precise}$	0.964	0.813	14.703	3.799	2.760	3.604
	$AIS5_{fuzzy}$	0.966	0.817	13.809	3.684	2.693	3.513
	$Set3_{precise}$ [a]	0.961	0.794	15.501	3.920	2.867	3.740
AB	*Set1*	0.964	0.790	14.672	3.464	2.781	3.712
	$AIS2_{precise}$	0.984	0.813	6.382	2.399	1.970	2.684
	$AIS2_{fuzzy}$	0.986	0.812	5.520	2.083	1.663	2.269
	$AIS3_{precise}$	0.991	0.813	3.818	1.747	1.405	1.959
	$AIS3_{fuzzy}$	0.988	0.816	4.669	1.917	1.519	2.098
	$AIS4_{precise}$	0.987	0.810	5.060	2.008	1.642	2.260
	$AIS4_{fuzzy}$	0.984	0.812	6.476	2.401	2.023	2.782
	$AIS5_{precise}$	0.988	0.826	4.812	2.046	1.675	2.298
	$AIS5_{fuzzy}$	0.987	0.820	5.055	2.014	1.624	2.245
	$Set3_{precise}$ [a]	0.982	0.777	7.272	2.415	1.888	2.558
GB	*Set1*	0.962	0.803	15.408	3.756	2.777	3.605
	$AIS2_{precise}$	0.971	0.815	11.694	3.275	2.287	2.984
	$AIS2_{fuzzy}$	0.963	0.811	14.984	3.741	2.665	3.503
	$AIS3_{precise}$	0.978	0.818	8.854	2.705	1.774	2.312
	$AIS3_{fuzzy}$	0.973	0.821	10.691	3.106	2.141	2.809
	$AIS4_{precise}$	0.973	0.811	10.692	3.036	2.085	2.729
	$AIS4_{fuzzy}$	0.975	0.814	9.991	2.994	2.068	2.705
	$AIS5_{precise}$	0.975	0.826	10.330	3.084	2.147	2.810
	$AIS5_{fuzzy}$	0.977	0.822	9.116	2.812	1.991	2.610
	$Set3_{precise}$ [a]	0.986	0.785	5.466	2.156	1.442	1.880

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.972	0.813	11.021	3.022	2.222	2.865
	AIS2$_{precise}$	0.972	0.810	11.116	3.156	2.168	2.779
	AIS2$_{fuzzy}$	0.968	0.814	12.937	3.415	2.437	3.161
	AIS3$_{precise}$	0.981	0.816	7.733	2.613	1.674	2.122
	AIS3$_{fuzzy}$	0.977	0.822	9.345	2.859	1.912	2.453
	AIS4$_{precise}$	0.973	0.813	10.730	3.019	2.111	2.714
	AIS4$_{fuzzy}$	0.972	0.821	11.364	3.182	2.236	2.883
	AIS5$_{precise}$	0.973	0.827	10.967	3.204	2.208	2.823
	AIS5$_{fuzzy}$	0.976	0.831	9.854	2.981	2.094	2.702
	Set3$_{precise}$ [a]	0.986	0.784	5.731	2.093	1.424	1.808
LB	*Set1*	0.957	0.789	17.547	4.044	3.053	3.968
	AIS2$_{precise}$	0.963	0.796	14.730	3.673	2.655	3.489
	AIS2$_{fuzzy}$	0.959	0.794	16.325	3.769	2.812	3.691
	AIS3$_{precise}$	0.976	0.809	9.713	2.740	1.932	2.536
	AIS3$_{fuzzy}$	0.975	0.805	10.022	3.034	2.163	2.840
	AIS4$_{precise}$	0.963	0.790	15.034	3.546	2.604	3.424
	AIS4$_{fuzzy}$	0.976	0.788	9.920	2.910	2.171	2.846
	AIS5$_{precise}$	0.981	0.803	7.624	2.614	1.804	2.364
	AIS5$_{fuzzy}$	0.973	0.813	10.759	2.940	2.095	2.762
	Set3$_{precise}$ [a]	0.982	0.785	7.152	2.366	1.742	2.283
SVM	*Set1*	0.906	0.786	38.185	6.078	4.323	5.574
	AIS2$_{precise}$	0.867	0.819	53.666	7.308	5.142	6.525
	AIS2$_{fuzzy}$	0.855	0.812	58.832	7.646	5.481	6.978
	AIS3$_{precise}$	0.861	0.820	56.159	7.463	5.273	6.678
	AIS3$_{fuzzy}$	0.860	0.820	56.543	7.490	5.326	6.747
	AIS4$_{precise}$	0.860	0.821	56.451	7.482	5.268	6.680
	AIS4$_{fuzzy}$	0.868	0.816	53.522	7.291	5.227	6.663
	AIS5$_{precise}$	0.854	0.815	58.893	7.641	5.411	6.831
	AIS5$_{fuzzy}$	0.855	0.811	58.735	7.637	5.447	6.893
	Set3$_{precise}$ [a]	0.871	0.748	51.533	7.113	5.173	6.591
ANN	*Set1*	0.863	0.786	55.639	7.392	5.651	7.274
	AIS2$_{precise}$	0.877	0.822	49.757	7.015	5.388	6.972
	AIS2$_{fuzzy}$	0.880	0.815	48.398	6.909	5.348	6.960
	AIS3$_{precise}$	0.886	0.818	46.000	6.707	5.154	6.669
	AIS3$_{fuzzy}$	0.891	0.816	44.189	6.596	5.071	6.574
	AIS4$_{precise}$	0.895	0.820	42.238	6.442	4.942	6.414
	AIS4$_{fuzzy}$	0.886	0.819	46.178	6.751	5.216	6.783
	AIS5$_{precise}$	0.879	0.805	48.929	6.903	5.287	6.805
	AIS5$_{fuzzy}$	0.869	0.806	53.158	7.238	5.571	7.180
	Set3$_{precise}$ [a]	0.892	0.771	43.321	6.515	5.071	6.587
Ridge	*Set1*	0.790	0.781	85.163	9.224	6.955	8.817
	AIS2$_{precise}$	0.807	0.797	77.989	8.828	6.677	8.485
	AIS2$_{fuzzy}$	0.806	0.793	78.557	8.860	6.742	8.596
	AIS3$_{precise}$	0.813	0.798	75.843	8.705	6.551	8.321
	AIS3$_{fuzzy}$	0.813	0.795	75.785	8.701	6.614	8.420
	AIS4$_{precise}$	0.811	0.800	76.446	8.740	6.580	8.334
	AIS4$_{fuzzy}$	0.811	0.797	76.688	8.754	6.640	8.432
	AIS5$_{precise}$	0.809	0.799	77.312	8.789	6.635	8.431
	AIS5$_{fuzzy}$	0.809	0.796	77.344	8.791	6.707	8.543
	Set3$_{precise}$ [a]	0.820	0.758	72.381	8.498	6.520	8.315
	Set1	0.789	0.781	85.405	9.238	6.961	8.819

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
LASSO	$AIS2_{precise}$	0.807	0.796	78.356	8.848	6.703	8.536
	$AIS2_{fuzzy}$	0.806	0.792	78.750	8.871	6.753	8.625
	$AIS3_{precise}$	0.811	0.796	76.696	8.753	6.608	8.410
	$AIS3_{fuzzy}$	0.811	0.796	76.732	8.756	6.644	8.459
	$AIS4_{precise}$	0.809	0.798	77.292	8.787	6.634	8.425
	$AIS4_{fuzzy}$	0.809	0.797	77.165	8.781	6.659	8.477
	$AIS5_{precise}$	0.808	0.798	77.734	8.813	6.664	8.483
	$AIS5_{fuzzy}$	0.808	0.796	77.682	8.811	6.719	8.568
	$Set3_{precise}$ [a]	0.819	0.758	72.827	8.524	6.550	8.374

Note: $Set3_{precise}$ is the best dataset with DFS1.

Table A14. The fit performance of eleven machine learning models for ship S8 (DFS2)

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
DT	*Set1*	0.916	0.774	54.181	7.305	5.213	6.441
	$AIS2_{precise}$	0.910	0.734	54.413	7.305	5.175	6.255
	$AIS2_{fuzzy}$	0.909	0.751	55.016	7.391	5.225	6.303
	$AIS3_{precise}$	0.904	0.764	58.172	7.506	5.430	6.557
	$AIS3_{fuzzy}$	0.910	0.757	54.043	7.255	5.193	6.272
	$AIS4_{precise}$	0.910	0.746	54.031	7.280	5.178	6.240
	$AIS4_{fuzzy}$	0.907	0.747	56.211	7.451	5.288	6.384
	$AIS5_{precise}$	0.908	0.771	55.429	7.369	5.275	6.362
	$AIS5_{fuzzy}$	0.925	0.777	45.620	6.622	4.665	5.687
	$Set3_{precise}$ [a]	0.916	0.769	50.649	6.985	4.922	5.949
ET	*Set1*	0.998	0.882	1.556	0.811	0.551	0.679
	$AIS2_{precise}$	0.998	0.866	0.927	0.774	0.471	0.588
	$AIS2_{fuzzy}$	0.995	0.862	2.789	1.167	0.761	0.941
	$AIS3_{precise}$	0.997	0.870	2.074	1.178	0.744	0.925
	$AIS3_{fuzzy}$	0.997	0.870	1.736	0.987	0.628	0.776
	$AIS4_{precise}$	0.996	0.865	2.660	1.228	0.795	0.986
	$AIS4_{fuzzy}$	0.997	0.863	1.868	1.000	0.639	0.791
	$AIS5_{precise}$	0.998	0.877	1.223	0.864	0.549	0.687
	$AIS5_{fuzzy}$	0.998	0.874	1.459	0.949	0.616	0.766
	$Set3_{precise}$ [a]	0.995	0.876	2.783	1.404	0.907	1.120
RF	*Set1*	0.978	0.859	13.895	3.707	2.535	3.124
	$AIS2_{precise}$	0.974	0.840	15.552	3.918	2.680	3.274
	$AIS2_{fuzzy}$	0.976	0.840	14.417	3.782	2.623	3.196
	$AIS3_{precise}$	0.970	0.848	17.980	4.202	2.858	3.508
	$AIS3_{fuzzy}$	0.975	0.848	15.355	3.895	2.688	3.294
	$AIS4_{precise}$	0.974	0.839	15.743	3.948	2.718	3.318
	$AIS4_{fuzzy}$	0.974	0.840	15.800	3.958	2.747	3.356
	$AIS5_{precise}$	0.975	0.856	15.054	3.848	2.645	3.229
	$AIS5_{fuzzy}$	0.976	0.860	14.471	3.781	2.619	3.204
	$Set3_{precise}$ [a]	0.976	0.855	14.566	3.798	2.624	3.187
AB	*Set1*	0.982	0.870	11.601	3.288	2.747	3.479
	$AIS2_{precise}$	0.993	0.852	4.466	2.032	1.640	2.081
	$AIS2_{fuzzy}$	0.990	0.860	6.002	2.344	1.935	2.443
	$AIS3_{precise}$	0.993	0.857	4.187	1.850	1.461	1.869
	$AIS3_{fuzzy}$	0.993	0.859	4.258	1.782	1.406	1.799
	$AIS4_{precise}$	0.992	0.853	4.887	2.044	1.651	2.115
	$AIS4_{fuzzy}$	0.994	0.856	3.765	1.844	1.442	1.826
	$AIS5_{precise}$	0.995	0.866	3.047	1.544	1.182	1.516
	$AIS5_{fuzzy}$	0.996	0.874	2.605	1.432	1.084	1.388
	$Set3_{precise}$ [a]	0.991	0.863	5.365	2.114	1.693	2.148
GB	*Set1*	0.983	0.875	10.771	3.062	2.188	2.750
	$AIS2_{precise}$	0.990	0.845	6.089	1.946	1.312	1.638
	$AIS2_{fuzzy}$	0.991	0.842	5.472	1.961	1.317	1.636
	$AIS3_{precise}$	0.989	0.851	6.797	2.127	1.416	1.777
	$AIS3_{fuzzy}$	0.985	0.850	9.273	2.484	1.694	2.119
	$AIS4_{precise}$	0.991	0.849	5.801	1.838	1.227	1.539
	$AIS4_{fuzzy}$	0.992	0.846	4.749	1.968	1.355	1.695
	$AIS5_{precise}$	0.988	0.859	7.367	2.097	1.438	1.781
	$AIS5_{fuzzy}$	0.986	0.863	8.510	2.330	1.622	2.011
	$Set3_{precise}$ [a]	0.985	0.860	9.102	2.427	1.670	2.075

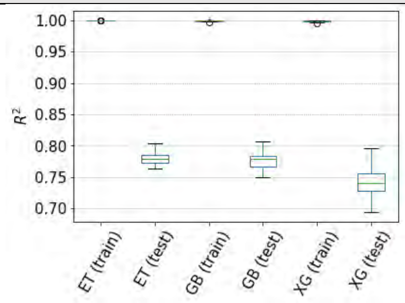
Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
XG	*Set1*	0.991	0.877	5.538	1.956	1.429	1.791
	$AIS2_{precise}$	0.979	0.841	12.914	3.086	2.202	2.713
	$AIS2_{fuzzy}$	0.989	0.850	6.647	2.089	1.460	1.815
	$AIS3_{precise}$	0.985	0.842	9.459	2.492	1.759	2.163
	$AIS3_{fuzzy}$	0.989	0.850	6.678	2.185	1.473	1.824
	$AIS4_{precise}$	0.982	0.839	11.047	2.796	1.970	2.439
	$AIS4_{fuzzy}$	0.987	0.847	7.669	2.271	1.552	1.925
	$AIS5_{precise}$	0.973	0.863	16.064	3.747	2.646	3.231
	$AIS5_{fuzzy}$	0.978	0.861	13.557	3.275	2.288	2.797
	$Set3_{precise}$ [a]	0.979	0.856	12.821	2.974	2.114	2.589
LB	*Set1*	0.979	0.871	13.718	3.540	2.601	3.309
	$AIS2_{precise}$	0.973	0.832	16.182	3.417	2.447	3.043
	$AIS2_{fuzzy}$	0.983	0.837	9.937	2.763	1.977	2.484
	$AIS3_{precise}$	0.982	0.844	10.828	2.743	1.936	2.441
	$AIS3_{fuzzy}$	0.979	0.848	12.946	3.069	2.203	2.759
	$AIS4_{precise}$	0.969	0.824	18.763	4.017	2.929	3.677
	$AIS4_{fuzzy}$	0.973	0.835	16.409	3.765	2.725	3.415
	$AIS5_{precise}$	0.973	0.849	16.500	3.523	2.521	3.138
	$AIS5_{fuzzy}$	0.978	0.860	13.784	3.213	2.262	2.836
	$Set3_{precise}$ [a]	0.976	0.852	14.749	3.261	2.338	2.882
SVM	*Set1*	0.900	0.862	64.371	8.014	5.742	6.905
	$AIS2_{precise}$	0.892	0.846	64.923	8.048	5.555	6.624
	$AIS2_{fuzzy}$	0.886	0.841	68.842	8.289	5.746	6.819
	$AIS3_{precise}$	0.906	0.861	56.735	7.523	5.223	6.371
	$AIS3_{fuzzy}$	0.901	0.856	59.608	7.712	5.314	6.426
	$AIS4_{precise}$	0.906	0.862	56.668	7.517	5.201	6.343
	$AIS4_{fuzzy}$	0.900	0.856	60.112	7.746	5.316	6.423
	$AIS5_{precise}$	0.897	0.858	62.015	7.865	5.503	6.604
	$AIS5_{fuzzy}$	0.892	0.854	65.006	8.054	5.607	6.685
	$Set3_{precise}$ [a]	0.910	0.869	54.154	7.349	5.117	6.123
ANN	*Set1*	0.914	0.857	55.217	7.398	5.605	6.809
	$AIS2_{precise}$	0.903	0.834	58.524	7.590	5.568	6.678
	$AIS2_{fuzzy}$	0.896	0.830	62.810	7.876	5.865	7.024
	$AIS3_{precise}$	0.913	0.854	52.371	7.201	5.280	6.420
	$AIS3_{fuzzy}$	0.910	0.850	54.338	7.351	5.383	6.496
	$AIS4_{precise}$	0.917	0.849	50.268	7.033	5.138	6.252
	$AIS4_{fuzzy}$	0.913	0.847	52.479	7.207	5.293	6.397
	$AIS5_{precise}$	0.911	0.846	53.888	7.282	5.345	6.458
	$AIS5_{fuzzy}$	0.912	0.848	53.388	7.276	5.335	6.431
	$Set3_{precise}$ [a]	0.924	0.862	46.222	6.733	4.964	5.959
Ridge	*Set1*	0.866	0.842	86.315	9.288	7.004	8.561
	$AIS2_{precise}$	0.858	0.829	85.639	9.248	7.008	8.573
	$AIS2_{fuzzy}$	0.856	0.826	87.056	9.324	7.077	8.631
	$AIS3_{precise}$	0.868	0.839	79.392	8.905	6.681	8.300
	$AIS3_{fuzzy}$	0.866	0.836	81.009	8.996	6.756	8.358
	$AIS4_{precise}$	0.867	0.838	79.957	8.936	6.693	8.299
	$AIS4_{fuzzy}$	0.864	0.835	81.764	9.037	6.780	8.364
	$AIS5_{precise}$	0.867	0.840	80.108	8.945	6.728	8.344
	$AIS5_{fuzzy}$	0.865	0.838	81.568	9.026	6.771	8.361
	$Set3_{precise}$ [a]	0.879	0.853	72.818	8.529	6.512	7.959
	Set1	0.865	0.842	87.140	9.332	7.023	8.576

Model	Dataset	R^2	R^2 (test)	MSE	RMSE (t/day)	MAE (t/day)	MAPE (%)
LASSO	$AIS2_{precise}$	0.857	0.828	85.994	9.267	7.011	8.567
	$AIS2_{fuzzy}$	0.855	0.826	87.359	9.340	7.076	8.621
	$AIS3_{precise}$	0.867	0.839	80.087	8.944	6.729	8.334
	$AIS3_{fuzzy}$	0.864	0.837	81.870	9.043	6.794	8.365
	$AIS4_{precise}$	0.867	0.838	80.157	8.948	6.700	8.306
	$AIS4_{fuzzy}$	0.864	0.834	81.947	9.047	6.787	8.369
	$AIS5_{precise}$	0.867	0.840	80.433	8.963	6.737	8.348
	$AIS5_{fuzzy}$	0.864	0.838	82.001	9.050	6.782	8.359
	$Set3_{precise}$ [a]	0.878	0.852	73.581	8.573	6.525	7.966

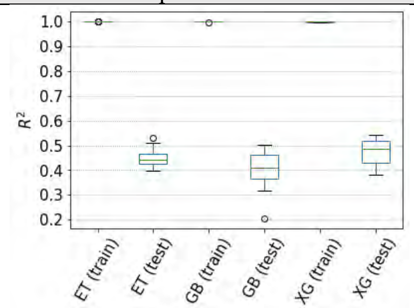
Note: $Set3_{precise}$ is the best dataset with DFS1.

2-month training + 1-month application/test (ship S5)

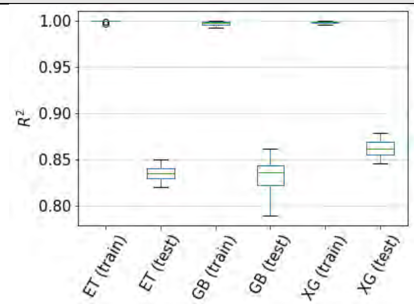
Ship S5



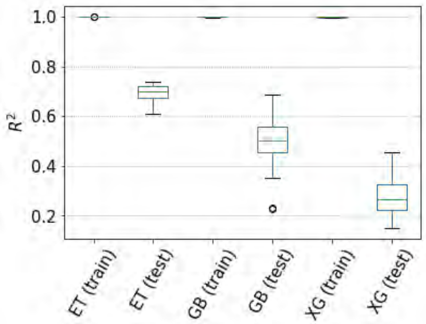
(a) *RollingSet1*



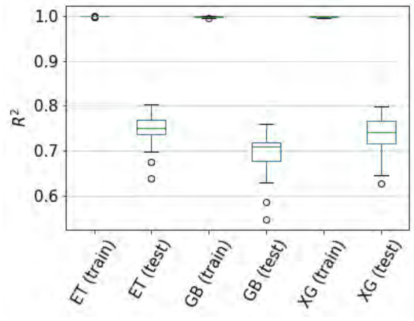
(b) *RollingSet2*



(c) *RollingSet3*



(d) *RollingSet4*



(e) *RollingSet5*

Figure A1. Fit performance (R^2) of models in a "2-month training + 1-month application/test" application scenario for ship S5. "ET (train)" means the fit performance of ET on train set. "ET (test)" means the fit performance of ET on test/application set.

2-month training + 1-month application/test (ship S6)

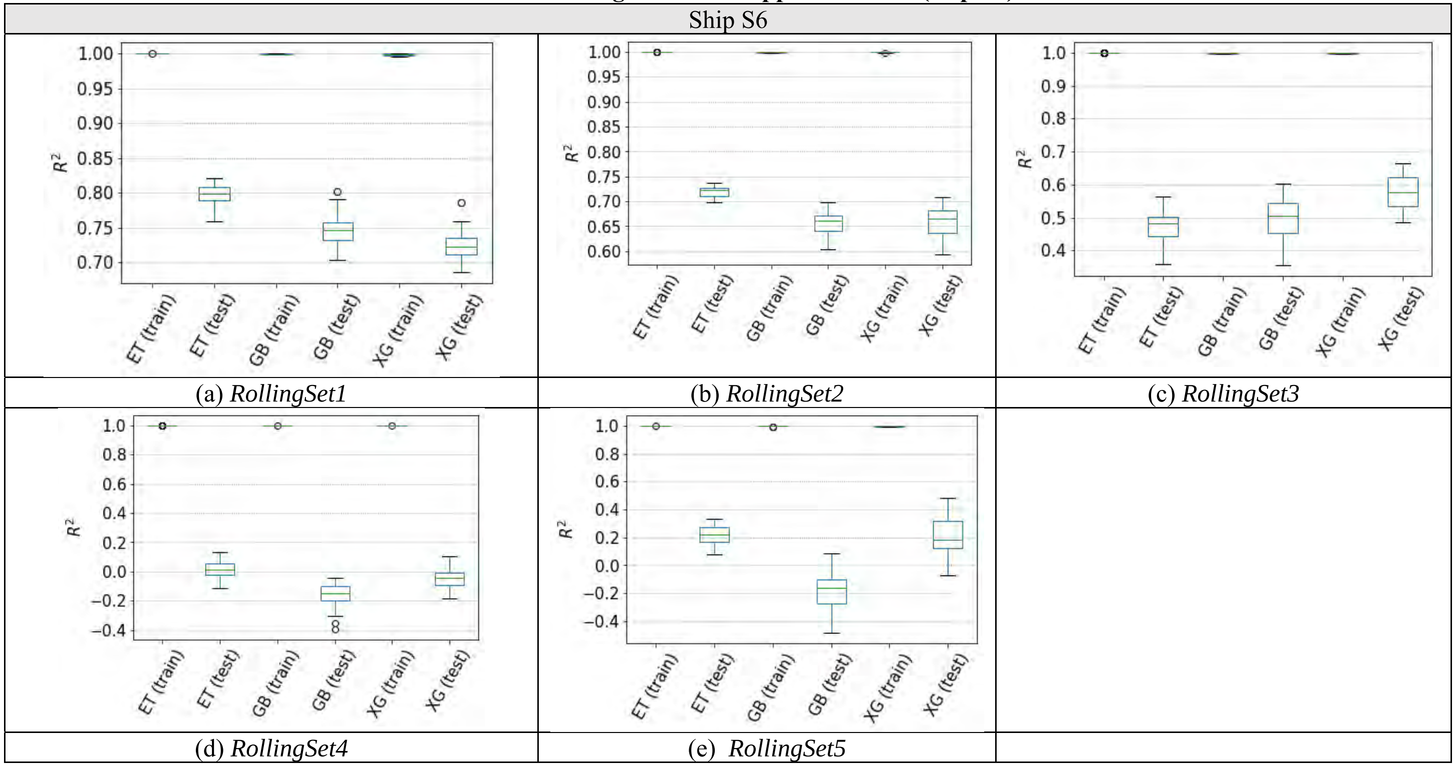
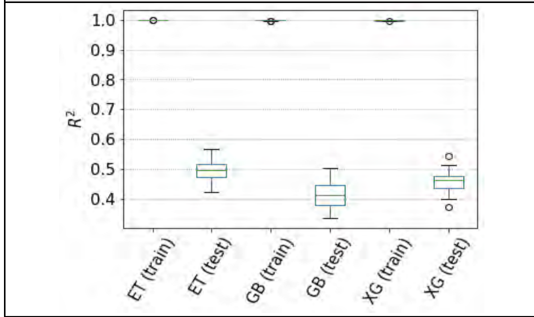


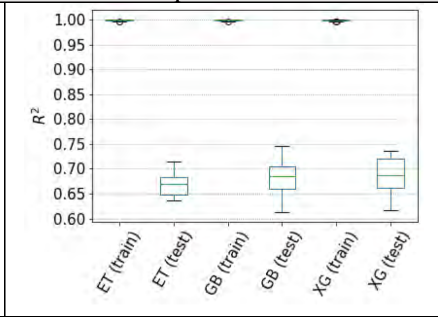
Figure A2. Fit performance (R^2) of models in a "2-month training + 1-month application/test" application scenario for ship S6. "ET (train)" means the fit performance of ET on train set. "ET (test)" means the fit performance of ET on test/application set.

1-month training + 1-month application/test (ship S5)

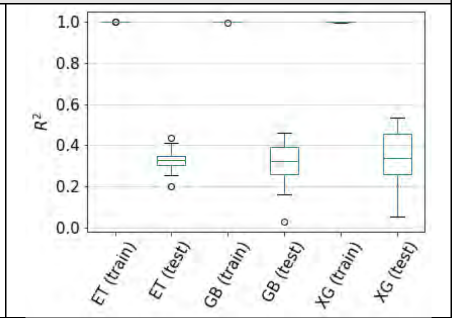
Ship S5



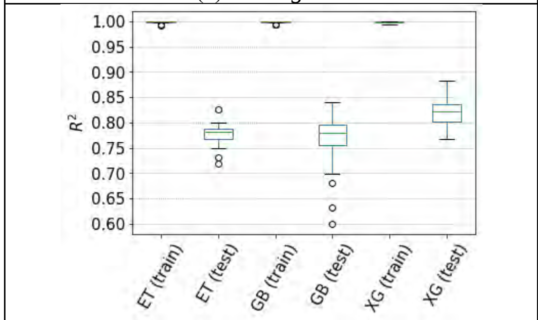
(a) *RollingSet1*



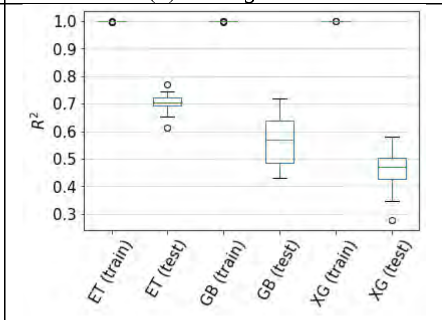
(b) *RollingSet2*



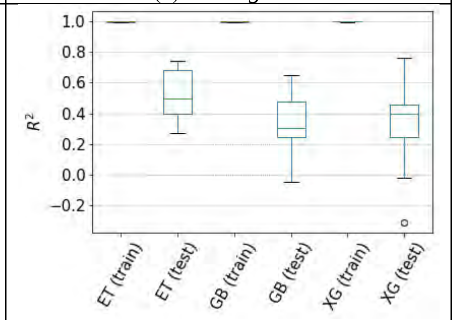
(c) *RollingSet3*



(d) *RollingSet4*



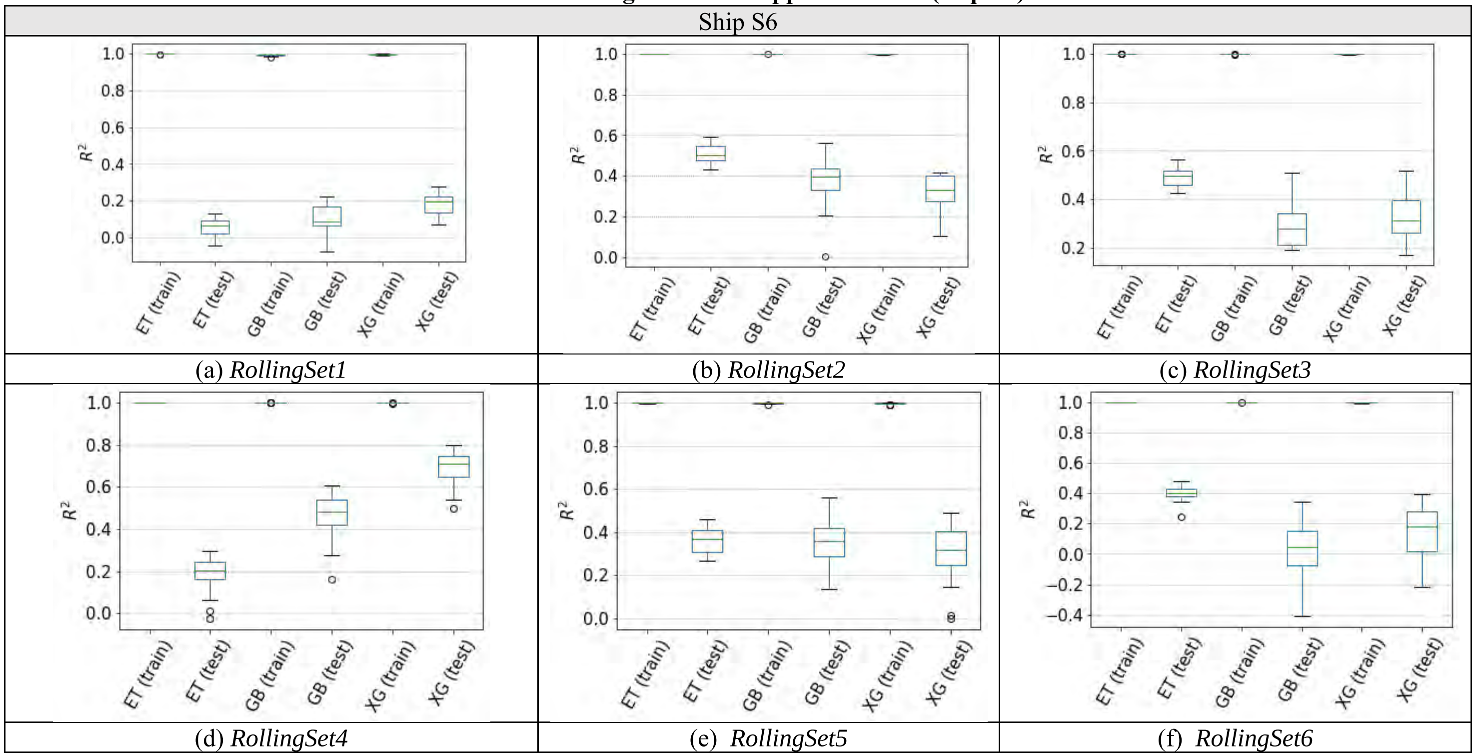
(e) *RollingSet5*



(f) *RollingSet6*

Figure A3. Fit performance (R^2) of models in a "1-month training + 1-month application/test" application scenario for ship S5. "ET (train)" means the fit performance of ET on train set. "ET (test)" means the fit performance of ET on test/application set.

1-month training + 1-month application/test (ship S6)



(a) *RollingSet1* (b) *RollingSet2* (c) *RollingSet3*

(d) *RollingSet4* (e) *RollingSet5* (f) *RollingSet6*

Figure A4. Fit performance (R^2) of models in a "1-month training + 1-month application/test" application scenario for ship S6. "ET (train)" means the fit performance of ET on train set. "ET (test)" means the fit performance of ET on test/application set.

International Association of Maritime Universities

Meiwa Building 8F, 1-15-10 Toranomon, Minato-ku, Tokyo 105-0001, Japan

Tel : 81-3-6257-1812 E-mail : info@iamu-edu.org URL : http://www.iamu-edu.org

ISBN No. 978-4-907408-38-1