

Cammin, Philip; Yu, Jingjing; Voß, Stefan

Article — Published Version

Tiered prediction models for port vessel emissions inventories

Flexible Services and Manufacturing Journal

Provided in Cooperation with:

Springer Nature

Suggested Citation: Cammin, Philip; Yu, Jingjing; Voß, Stefan (2022) : Tiered prediction models for port vessel emissions inventories, Flexible Services and Manufacturing Journal, ISSN 1936-6590, Springer US, New York, NY, Vol. 35, Iss. 1, pp. 142-169, <https://doi.org/10.1007/s10696-022-09468-5>

This Version is available at:

<https://hdl.handle.net/10419/312853>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Tiered prediction models for port vessel emissions inventories

Philip Cammin¹ · Jingjing Yu^{1,2} · Stefan Voß¹

Accepted: 27 August 2022 / Published online: 2 October 2022
© The Author(s) 2022

Abstract

Albeit its importance, a large number of port authorities do not provide continuous or publicly available air emissions inventories (EIs) and thereby obscure the emissions contribution of ports. This is caused by, e.g., the economic effort generated by obtaining data. Therefore, the performance of abatement measures is not monitored and projected, which is specifically disadvantageous concerning top contributors such as container ships. To mitigate this issue, in this paper we propose port vessel EI prediction models by exploring the combination of different machine-learning algorithms, data from the one-off application of an activity-based bottom-up methodology and vessel-characteristics data. The results for this specific case show that prediction models enable acceptable trade-offs between the prediction performance and data requirements, promoting the creation of EIs.

Keywords Air emissions · Inventory · Green ports · Machine learning

1 Introduction

Air emissions are among the top environmental issues of port authorities (PAs) (COGEA 2017) and citizens (Ortiz et al. 2019), threatening to increase climate change and direct negative health impacts for approximately 230 million people by the top 100 ports (Merk 2014). Specifically shipping can contribute to the

✉ Philip Cammin
philip.cammin@uni-hamburg.de

Jingjing Yu
jingjing.yu@uni-hamburg.de

Stefan Voß
stefan.voss@uni-hamburg.de

¹ Institute of Information Systems, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany

² State Key Laboratory of Coastal and Offshore Engineering, Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116023, Liaoning, China

largest share of port emissions (Merk 2014). Yet, less than half of the world's top 49 container ports provide emissions reporting (Cammin et al. 2022) and thus lack transparency concerning the performance of emissions abatement measures for becoming more green ports. PAs, being the regulators of ports (Notteboom et al. 2021) justify the lack of EIs with their generated economic effort, as shown in case studies (Cammin et al. 2020). The motivation to create EIs, based on, e.g., political, regulatory or stakeholder requirements, may not be sufficient to justify this effort; for instance, through obtaining and further processing data, or engaging or integrating external (data) service providers. The vested interest to minimize data costs conforms with our observations in practice and academia. The shared and accepted solution is the use of assumptions for unavailable data in EI methodologies.

Although this practice is accepted, it increases the imprecision of the EIs. Therefore, it seems advisable to carefully scrutinize the capability of methodologies to determine the performance of emissions abatement measures. This includes the proposal in this study.

An example for vessel characteristics and activity-based assumptions is the use of average engine power and average speed per vessel type, respectively. Among the works that surmount the unavailability of data with assumptions are activity-based bottom-up methodologies. For instance, POLA (2019) and Ekmekçioğlu et al. (2021) adapt the EPA and ENTEC methodologies, respectively, i.e., using assumptions by relying on historic data on average speeds in certain areas. Likewise, academics suffer from the unavailability of data, trying to find a trade-off between the accuracy and costs, which is reasonable specifically for multi-port EIs. For instance, Wan et al. (2020) use multi-port Automatic Identification System (AIS) data and assume engine powers with linear regression based on deadweight tonnage (DWT), and Merk (2014) extrapolates data from one month to a year for budgetary reasons to create multi-port vessel EIs. As the necessary data scales with the number of ports assessed, less data-demanding methodologies can foster the (successive) creation of port-related vessel EIs.

To this end, the objective of this paper is the development of tiered prediction models to create port-related container ship EIs taking into account data requirements. In this sense, lower-tier models rely on less features, potentially resolving data dependencies as well as reducing costs to adhere to budgetary constraints. A range of alternative models can then be input to a trade-off analysis, which has not received much attention in previous studies that focus on the prediction performance.

In this paper, the development of prediction models is based on (1) data from the one-off application of an activity-based bottom-up methodology that incorporates detailed vessel-characteristics data and, (2) the use of vessel-characteristics data, both as input for (3) different machine-learning algorithms. While aiming to foster the creation of EIs in order to promote green ports, we also point out limitations and pitfalls of this approach.

The remainder of this paper is organized as follows. Section 2 reviews the application of machine learning for fuel consumption and emissions prediction of vessels. Next, Sect. 3 describes the methodology of the proposed approach and Sect. 4 addresses the numerical experiments. Section 5 discusses the key results and

limitations of the presented approach. Finally, Sect. 6 concludes this study and outlines limitations and potential future research.

2 Literature review

Academia has taken advantage of data-driven models and machine learning for a variety of problems in fuel consumption prediction. Solving this problem is an intermediate step towards emissions calculation, i.e., the necessary next steps are determining the used fuel type, applying the dependent emissions factors with respect to exhaust gas cleaning systems, such as scrubbers that are installed on vessels or barge-mounted to serve vessels at berth. Regarding the prediction of vessel emissions and fuel consumption, the literature can be classified by its focus on different machine-learning algorithms (e.g., artificial neural network (ANN), multiple linear regression (MLR), support vector regression (SVR), least absolute shrinkage and selection operator (LASSO), random forest (RF), nonlinear regression (NLR)), the focus on different ship types (e.g., container or cruise ships), the scale of application (e.g., single to multiple vessels), and the feature set. Notably, many of the works that employ vessel-specific and voyage-specific data, e.g., shaft revolution per minute (RPM) speed from on-board sensors or cargo weight, do so for a single vessel. From the perspective of PAs or public institutions, it is questionable if such a data coverage of all vessels operating in a port is feasible as it would generate high effort, whereas obtaining vessel-characteristics data from the usual maritime data providers is a common choice in literature and practice. Many of the works employ noon report data for the features and the response.

The examined literature is categorized in three groups. The first two groups refer to shipping-related literature that address the prediction of fuel consumption, and emissions. The summary is presented in Table 1. The third group refers to shipping-unrelated references that concern the prediction of emissions, which is summarized in Table 2. Due to the narrower domain, we found fewer shipping-related references concerning emissions, for which more insights are provided in the following.

First, Ekmekçioğlu et al. (2021) apply the regression analysis using the feature gross tonnage (GT) and the response emissions for container ships in the activity modes of cruising, maneuvering, and berthing at four ports in Turkey. Prior, the emissions are calculated using the ENTEC activity-based bottom-up method. The study generalizes time for the cruising and maneuvering modes, but appears to use actual berthing times for the calculation. Notably, times are not used as a feature for training. The model's performance is reported using the coefficient of determination (R^2) with 0.93 for the cruising mode, 0.92 for the maneuvering mode and 0.80 for the berthing mode.

The weaker performance of the berthing mode indicates that GT does not sufficiently correlate with the berthing time, although a loose correlation is reasonable, i.e., larger vessels potentially berth longer due to longer overall container handling operations. High variations of berthing times for similar-sized vessels are adverse for the performance of the model. This emergence could be based on inconsistent container volume handling across similar-sized vessels (see Cullinane et al. (2006))

Table 1 Summary of shipping-related literature concerning the prediction of fuel consumption and emissions

Reference	ANN	MLR	SVR	LASSO	RF	NLR	Vessel type	Vessel count	Features
Karagiannidis and Themelis (2021)	✓						Container	1	Vessel speed through water (STW), trim, draft mean, rudder angle, propeller shaft power, significant wave height, fouling, wind effect, sea current
Kim et al. (2021)	✓	✓					Container	1	Main engine RPM, vessel speed over ground (SOG), STW, relative wind speed and direction, rudder angle, trim, sea current, wetted surface area, main draught, displacement
Zhou et al. (2021)	✓		✓	✓			Oil tanker	1	Main engine speed, fore and aft draft, wind speed, relative wind direction
Zhu et al. (2021)	✓	✓	✓				Ro-Pax	1	STW, trim, draught, wind speed and direction, rudder angle, propeller pitch
Le et al. (2020)	✓						Container	143	Average vessel speed, sailing time, maximum capacity, cargo weight, number of containers, displacement
Panapakidis et al. (2020)	✓						Ro-Pax	1	Past fuel consumption (on days $d - 1$ and $d - 8$), average vessel speed, wind force, number of passengers, main engine hours and total fuel consumption, sailing distance (all on days d and $d - 1$)
Peng et al. (2020)	✓	✓			✓		Container, bulk, general cargo	7788	Net tonnage, DWT, actual cargo handling volume, efficiency of facilities, trade type; vessel arrival time, attribute of enterprise, berth, goods, import/export, ship length, ship arrival month and weekday, ship nation, operation department
Uyanik et al. (2020)			✓	✓			Container	1	Main engine power, RPM, cooler temperature, main lubricating oil, etc.
Yan et al. (2020)					✓		Bulk	1	Bad weather ratio, sailing speed, relative sea swell direction, relative wind direction, relative wind wave direction, sea swell height, sea current type and value, wind force, total cargo weight, combined wind waves and swell height

Table 1 (continued)

Reference	ANN	MLR	SVR	LASSO	RF	NLR	Vessel type	Vessel count	Features
Yuan et al. (2020)	✓	✓	✓	✓	✓		Inland cargo	1	SOG, course over ground (COG), left and right engine speed and temperature, water level and speed, wind speed and angle, segment ID
Hu et al. (2019)	✓						Container	1	Ship shaft, speed, average draft, trims, current speed and direction, wind speed and direction, wave height and direction
Kee and Simon (2019)		✓					Tugboat	2	Travelled distance and hours, vessel speed, DWT, consumption metric, wind speed
Zheng et al. (2019)	✓						Cruise	1	Vessel speed, load
Kee et al. (2018)	✓	✓					Tugboat	2	Traveled distance and hours, vessel speed, DWT, wind speed
Wang et al. (2018)	✓		✓	✓			Container	97	Length overall, beam, cargo weight, average draft, trim angle, initial metacentric height, RPM, main engine load, SOG, current factor, performance speed, beaufort scale, wind waves height, swell height and direction, CO ₂ efficiency, fuel efficiency, ship course, wind direction, significant wave height, displacement
Du et al. (2019)	✓						Container	2	Sailing speed, displacement, trim, wave height and direction, wind force and direction, current speed and direction, water temperature
Gkerekos et al. (2019)	✓		✓	✓	✓		Reefer, bulk	2	Vessel speed, engine speed, current speed, wind speed and direction, daily traveled distance, sea state and direction, slip, draft force and aft
Fletcher et al. (2018)		✓*					Cargo	2	Shaft speed, engine power
Ekmekçioglu et al. (2021)					✓		Container	1024	GT

Table 1 (continued)

Reference	ANN	MLR	SVR	LASSO	RF	NLR	Vessel type	Vessel count	Features
Fabregat et al. (2021)	✓		✓		✓		Cruise	–	Wind alignment port/pollutant station and airport/pollutant station, wind velocity, precipitation, temperature, atmospheric pressure, solar irradiance, relative humidity, traffic intensity, number of vessels and cruise ships, median length of vessels and cruise ships, flight numbers (all hourly), binary variables for each weekday, year, day of year, hour of day, daily cloud cover
Schaub et al. (2019)	✓						Test bed engine	–	Engine torque, speed and fuel consumption by time unit
This paper	✓	✓	✓				Container	104	Main engine and auxiliary engine power, auxiliary engine number, GT, IMO Tier, vessel design speed (DS), minutes in activity mode

* Principal Component Regression Model

Table 2 Summary of shipping-unrelated literature concerning the prediction of emissions

Reference	ANN	MLR	SVR	LASSO	RF	NLR	Features	Target
Li et al. (2017)	✓		✓				RPM, intake air temperature, manifold absolute pressure, ambient air temperature, idling duration	Real-time exhaust emission rates of carbon monoxide (CO), carbon dioxide (CO ₂), nitrogen oxides (NO _x), and hydrocarbon (HC) for light-duty vehicles
Bhowmik et al. (2018)	✓						Engine load, kerosene share, ethanol share	Brake thermal efficiency, brake specific energy consumption, NO _x , unburned HC and CO for diesel engine fueled with diesosanol blends
Ceylan and Bulkan (2018)	✓	✓					Temperature, visibility distance, dew point, wind velocity, pressure and relative humidity for a two-year period	Daily mean PM ₁₀ concentration levels in Sakarya city
Hosseini et al. (2019)		✓					Population, CO ₂ intensity, gross domestic product (GDP) per capita, share of fossil fuels in electricity production, energy consumption per capita	Total CO ₂ emissions per year in Iran
Zuo et al. (2019)			✓				Equivalence ratio, blend ratio	Brake thermal efficiency, brake specific fuel consumption, CO, unburned hydrocarbon, and NO _x for spark ignition engine fueled with butanol-gasoline blends
Xie et al. (2020)		✓	✓		✓		Passenger cars per 1,000 inhabitants, stock of vehicles, volumes of freight transport and passenger transport relative to GDP	Traffic-related emissions of non-methane volatile organic compounds for 28 European Union member countries
Altikat (2021)		✓	✓				Crop species, soil temperature, soil moisture content, photosynthetic active radiation, and soil oxygen exchange, plant type	CO ₂ flux from soil to atmosphere
Do et al. (2021)		✓	✓				Throttle position, air intake pressure, RPM, cooling water temperature, intake air temperature, exhausted emissions temperature, exhaust gas recirculation temperature, exhaust gas recirculation rate, and lambda	Torque and NO _x emissions of diesel engine fueled with biodiesel blends

Table 2 (continued)

Reference	ANN	MLR	SVR	LASSO	RF	NLR	Features	Target
Saha et al. (2021)				✓			Soil moisture, days after fertilization, soil texture, air temperature, soil carbon, precipitation, and nitrogen fertilizer rate	Daily N ₂ O flux for a corn rotation
Shams et al. (2021)	✓	✓					Urban traffic and green space parameters (e.g., number of vehicles at intersections, avg. distance between parks and measuring station), time and meteorological parameters (e.g., humidity, rainfall, air temperature, wind speed)	SO ₂ concentration level in Tehran
Umar et al. (2021)	✓	✓	✓				Hourly NO _x , nitrogen monoxide (NO), nitrogen dioxide (NO ₂), CO, SO ₂ emissions, wind speed, temperature, volume of cars and taxis	Concentration levels of PM _{2.5} and PM ₁₀ in central London
Ağbulut (2022)	✓		✓				GDP per capita, population, vehicle kilometer, and year	CO ₂ emissions and energy demand of transportation sector in Turkey

or disturbances in terminal operations such as the breakdown of quay cranes (Nour-mohammadzadeh and Voß 2022). If the effort to obtain the berthing time is acceptable, a regression model including time and GT could improve the performance for the berthing mode. Concerning cruising and maneuvering times, including the time as a feature would strengthen the model's performance in ports with different sailing routes, and thus different sailing durations within the geographic domain of the EI.

Second, Fabregat et al. (2021) use machine-learning algorithms to predict the hourly local pollutant concentration and estimate the impact of cruise ships' activities on the air quality of the port of Barcelona. To this end, twenty-five features and the target from eight stations are employed. The application of six algorithms shows that the day of the year and the traffic intensity are the two most important features. The gradient boosting machine (GBM) has the best prediction performance with the R^2 of 0.80 and the root mean squared error (RMSE) of 14%.

Third, Fletcher et al. (2018) use regression algorithms to predict the emissions concentration of two cargo ships' main engines and auxiliary engines in the activity modes of cruising, maneuvering, and berthing. The selected features are engine power and shaft speed. The results show that all five regression algorithms have a similar prediction performance for the cruising mode, while using supervised mixture probabilistic latent factor regression has the best prediction performance for the maneuvering and berthing modes.

Fourth, Schaub et al. (2019) apply the ANN to simulate the time series of the particulate matter (PM) concentration during a transient engine operation. The input features include the values of RPM and fuel consumption by time unit, while the prediction target is the PM concentration by time unit. The simulation results by using ANN are overall consistent with those by measurement while there are a lot of small amplitude oscillations and underestimation of the peak values of the PM concentration.

Another recent review of fuel consumption prediction (which also addresses optimization) is carried out by Yan et al. (2021a). In line with the findings by Yan et al. (2021b), this review indicates that machine learning has not been applied extensively for port vessel EIs, showing potential in its further exploration. Moreover, the application of machine-learning algorithms to predict emissions unrelated to shipping, as exemplified in Table 2, corroborates the idea of emissions prediction. It shows that machine-learning algorithms recently have received much attention for the prediction of emissions in various fields. In general, the popular applied methods include but are not limited to ANN, MLR and SVR. It is observed that literature focuses on improving the prediction performance of single models, whereas the present study examines the prediction performances of models considering different levels of data scarcity.

3 Methodology

In this section, the methodology carried out is presented. Section 3.1 outlines the data acquisition and Sect. 3.2 presents the steps concerning data preprocessing. Section 3.3 presents and justifies the feature engineering. In Sect. 3.4,

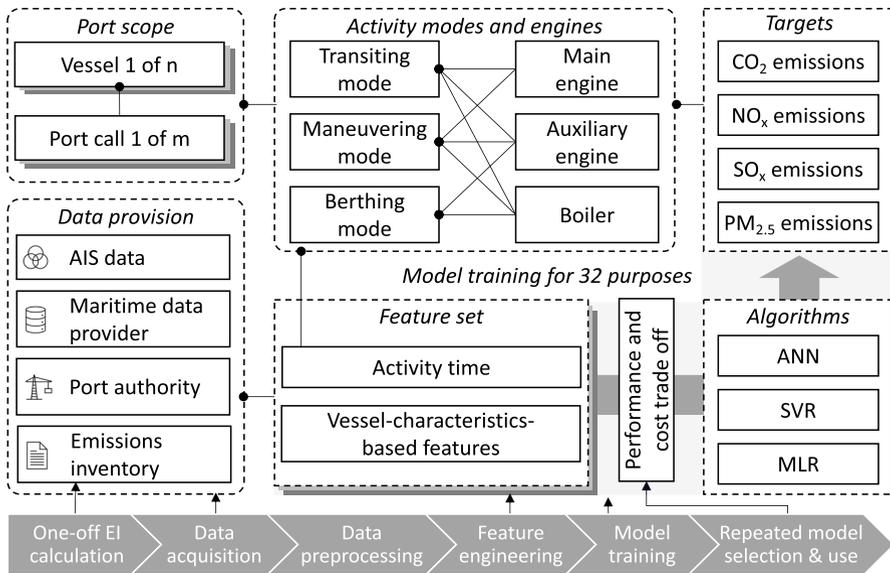


Fig. 1 Framework

the prediction models are defined and the algorithmic exposition is provided in Sect. 3.5. The performance assessment metrics are outlined in Sect. 3.6. The results are shown in the succeeding Sect. 4. The overall framework employed is compiled in Fig. 1.

3.1 Data acquisition

For this study, a three-month data set covering August to October 2019 from a large North American PA, residing in an emissions control area, is used. The data set includes the activity data per vessel and generated emissions. For each vessel and call, the operational time in different activity modes is provided. Moreover, the data set includes the volume of emitted emissions (disaggregated by gas type) by each engine type during those activity modes. The activity modes comprise transiting, maneuvering and berthing. The engines considered are the main engine, auxiliary engine and boiler. Finally, the gas types include CO_2 , SO_x , NO_x and $\text{PM}_{2.5}$. The emissions abatement measure shore-side electricity is not considered in the calculations; thus, the data set is considered as a preliminary EI. The majority of vessels calling at this port are container ships and bulk carriers, which is reflected in the EI. Vessel-characteristics data, such as the power of different engines are employed and, in this study, we focus on 104 container ships that are addressed in the preliminary EI for which we could acquire the data for all the features (see Sec. 3.3). These are responsible for 119 vessel port calls. The EI methodology of the PA follows an activity-based bottom-up calculation

approach with assumptions concerning the load factor in different areas of the geographic domain and is based on AIS data gathered via the PA's own terrestrial antennas, as well as satellite-AIS and vessel-characteristics data from external maritime data providers.

3.2 Data preprocessing

Concerning activity data preprocessing we differentiate between the common tasks of data smoothing and data transformation. Data smoothing is considered as an essential step for the data preparation, to address missing data and suppress outliers by data denoising schemes, such as the Empirical Mode Decomposition (Chen et al. 2021; Zhao et al. 2022). This has been taken into account with the preparation of the provided preliminary EI; thus, data smoothing is not further implemented here.

The data set of the EI comprises different data tables that store the information for 119 port calls. These tables are transformed into a single table. The resulting fields are vessel identifier (IMO number), vessel port call, minutes, mode, engine, gas, and emissions in tons. In total, 3,808 samples are generated, which is based on the 119 vessel port calls and 32 combinations (purposes). Therefore, each sample refers to one of the 32 combinations based on the three engines, three modes and four gas types, where the main engine does not emit during berthing. A combination expresses the specific purpose of a model (see Fig. 1). The gas type further specifies the target variable *tons* of emissions. Furthermore, the data is merged with vessel-characteristics data fields based on the vessel identifier. The resulting data set is the input to the feature engineering which is described in the subsequent Sect. 3.3. Notably, a combination is used to filter the samples accordingly for model training, hence, Engine, Mode and Gas are considered technical constraint variables.

3.3 Feature engineering

Feature engineering can be differentiated into feature generation and feature selection. As we do not merely aim to provide one best model but several best model alternatives for each tier, the feature selection for dimensionality reduction is limited to excluding redundant features (see Venkatesh and Anuradha (2019)).

The feature selection could be driven by known physical (causal) relationships from domain knowledge, or by statistical correlations (Karagiannidis and Themelis 2021; Lee and Lee 2021). The feature *Min* (minutes) relates to the operational runtime in the activity *Mode* and is suspected to have a major effect on a vessel's fuel consumption. Dynamic activity times in all modes per vessel and port are possible and thus, *Min* could be beneficial in a port-independent prediction model. For instance, the terminal locations in the port EI geographic domain could require north- or south-bound transiting sailing routes with different sailing distances as found in this paper's case port. The constraint variable *Mode* is included because it might reflect average load factors assumed in the activity-based bottom-up methodology. Moreover, the engine-power-related features, *TKW ME* (main engine power in kw), *THP ME* (main engine power in hp), *TKW AE* (auxiliary engines power in

kw), as well as *AE Num* (auxiliary engines quantity) and vessel *DS* (design speed) can be utilized. *TKW ME* and *THP ME* correlate by about 0.75 (compare Fig. 2), hence, *TKW ME* is selected arbitrarily for model training to reduce redundancy. The feature *GT* correlates with a vessel's size, similar to DWT which is used, e.g., in Wan et al. (2020), to assume the engine power by using regression. Hence, *GT* is regarded as a potential option to substitute the aforementioned engine-power-related features. Table 4 presents the correlation of features. The strongest correlations can be observed between *TKW ME* and *DS* as well as between *TKW ME* and *GT*. In this study, the categorical feature *IMOT* (IMO Tier) is generated based on the vessel build year, relating to the nitrogen oxide (NO_x) control requirements of MARPOL Annex VI (IMO 2019). The feature *IMOT* is label-encoded to reflect the ordinal dependency (see Stauder and Kühn (2021)). Furthermore, we apply the zero mean standardization to normalize the feature data (Wang et al. 2018).

The variable *VT* (vessel type) could be used as a feature or as a constraint variable to filter out samples for training multiple models each for a different type of vessels. In this study, *VT* is used merely as a constraint variable and is fixed to container ships. Thereby, the trained models potentially only respect potential vessel-type-specific methodologies and assumptions such as operating patterns. However, in this study, no further categorization or voyage data is employed. This would be beneficial if the training data would respect voyage-related data; for instance, the number of powered reefer containers.

Although both *GT* and *VT* are somewhat visible by nature, options to obtain mass data legally, conveniently, and reliably in one batch, are most likely only accessible through paid services.¹

Given the unavailability of the exact EI methodology employed by the PA, statistical correlation serves as an indicator for the usefulness of the features. However, in this paper, all feature set combinations are explored by computations with different algorithms so that alternative models can be suggested. Table 3 provides an overview of the used features and constraint variables.

3.4 Prediction models

Two sets of feature combination sets are defined. The following notation is used:

F Set of n features, $F = \{f_1, f_2, \dots, f_n\}$.

F_1 Set of feature sets based on select- k -best, $F_1 = \{(f_1), (f_1, f_2), \dots, (f_1, \dots, f_n)\}$, the elements in set F are in descending order based on their mutual information.

¹ See latest terms of use of International Maritime Organisation (IMO)'s Global Integrated Shipping Information System (GISIS) that offers, e.g., MMSI, VT, build year and GT.

Table 3 Features and constraints

Code	Feature	Constraint	Unit	Categorical values	Description
Engine		✓	–	Mai, Aux, Boi	Engine type
Mode		✓	–	TRN, MAN, BRT	Mode type
Gas		✓	–	CO ₂ , SO _x , NO _x , PM _{2.5}	Gas type
VT		✓	–	Container ship	Vessel type
DS	✓		kn	–	Design speed
GT	✓		t	–	Gross tonnage
IMOT	✓		–	0, 1, 2, 3	IMO tier
Min	✓		min	–	Minutes
TKW ME	✓		kw	–	Main engine power
TKW AE	✓		kw	–	Auxiliary engines power
AE Num	✓		–	–	Auxiliary engines qty.

Main (Mai), auxiliary (Aux), boiler (Boi)

Transiting (TRN), maneuvering (MAN), berthing (BRT)

Table 4 Correlation of features

	DS	GT	IMOT	Min	TKW ME	TKW AE	AE Num
DS	1.00	0.68	−0.42	−0.19	0.91	0.62	0.32
GT	0.68	1.00	−0.24	−0.11	0.81	0.48	0.13
IMOT	−0.42	−0.24	1.00	0.10	−0.42	−0.21	0.02
Min	−0.19	−0.11	0.10	1.00	−0.18	−0.11	−0.00
TKW ME	0.91	0.81	−0.42	−0.18	1.00	0.59	0.14
TKW AE	0.62	0.48	−0.21	−0.11	0.59	1.00	0.24
AE Num	0.32	0.13	0.02	−0.00	0.14	0.24	1.00

F_2 Set of feature sets, $F_2 = \{\hat{F} \mid \hat{F} \subseteq F\}$.

$F_i(e, m, g)$ Models using features of set F_i constraint by engine e , mode m and gas g .

T_i Tier of models with i being the number of used features.

The impact assessments refer to the applied algorithms, the combinations (purposes) of the models, and the feature sets. For instance, $F_1(\text{Mai}, \text{MAN}, \text{CO}_2)$ denotes models based on the feature sets of F_1 with the purpose to predict CO₂ emissions generated by the main engine when vessels maneuver (see, e.g., the use of this notation in Figs. 2 and 3).

The set F_1 is defined based on the select- k -best method. The seven features are sorted (descending) based on their mutual information score which detects the statistical dependence between the feature and the target, as shown in Fig. 2. The

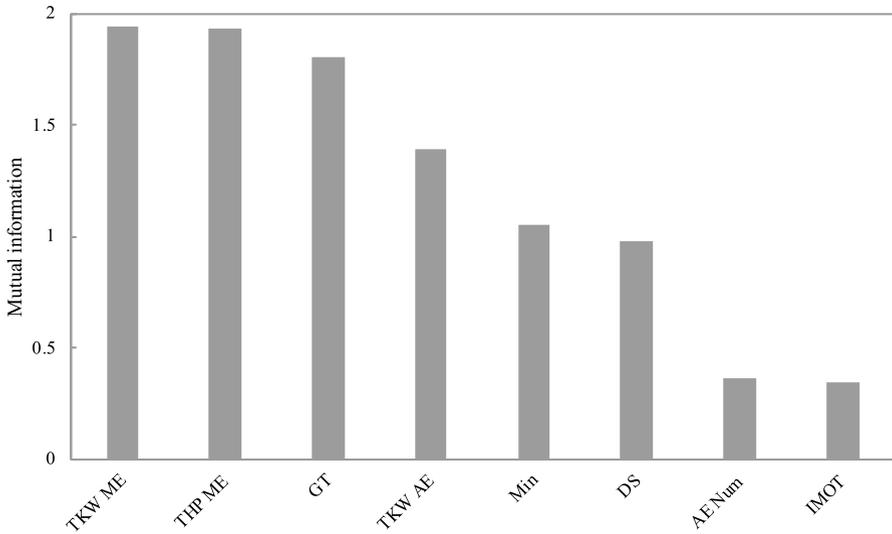


Fig. 2 Mutual information used for univariate feature selection based on $F_1(Mai, MAN, CO_2)$

Table 5 Constraints and features utilized for models based on $F_1(Mai, MAN, CO_2)$

Tier	Gas	Mode	Engine	VT	TKW ME	GT	TKW AE	DS	Min	AE Num	IMOT
1	✓	✓	✓	✓	✓						
2	✓	✓	✓	✓	✓	✓					
3	✓	✓	✓	✓	✓	✓	✓				
4	✓	✓	✓	✓	✓	✓	✓	✓			
5	✓	✓	✓	✓	✓	✓	✓	✓	✓		
6	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

mutual information is a non-negative value and estimated by the Shannon entropy from k -nearest neighbor distances (Kraskov et al. 2004; Ross 2014). If the mutual information is equal to zero, the two variables are independent; a higher value of the mutual information means a higher dependency between the two variables. Those features are then added cumulatively resulting in seven feature sets as shown in Table 5. The purpose of F_1 is to exemplify a statistical approach based on models using $F_1(Mai, MAN, CO_2)$. The results are presented in Sect. 4.1.

The set F_2 is defined by using the product of the seven features. Thereby, F_2 contains 127 feature sets.² The purpose of F_2 is fourfold:

² Based on $2^n - 1$, i.e., $2^7 - 1$, where -1 is necessary to include at least one feature.

- The best-performing models for all purposes $F_2(e, m, g)$ are identified to show the viability of the approach. The results are presented in Sect. 4.2.
- The performance of the models is differentiated by tier and algorithm for selected e, m and g . This may show the benefits of using different algorithms as well as the performance improvements with higher tiers. The results are presented in Sect. 4.3.
- The best models and alternatives for $F_2(Mai, MAN, CO_2)$, including the used features, adopted algorithm, and hyperparameters are shown. By doing so, substituting important features to resolve data unavailability is exemplified. The results are presented in Sect. 4.4.
- A comparison with the study of Ekmekçiöğlü et al. (2021), introduced in Sect. 2, who trained regression models with GT for container ships, is conducted. We exemplify the effect of using the feature Min on the prediction performance for time-dependent targets. The results are determined in Sect. 4.5.

3.5 Algorithms

This section briefly introduces the algorithms applied in this study. The selection of ANN, MLR and SVR is based on their widespread usage in the literature as presented in Sect. 2, and according to Yan et al. (2021), ANN is the most frequently used algorithm by far for fuel consumption prediction. Besides ANN, MLR and SVR are also popular supervised learning algorithms for the prediction of fuel consumption. In contrast to ANN and SVR that both can pick up nonlinear and linear relationships, MLR is only able to find linear relationships. Moreover, ANN, SVR, and MLR are representative algorithms of the neural-network-based models (Panapakidis et al. 2020; Karagiannidis and Themelis 2021), instance-based models (Zhou et al. 2021; Zhu et al. 2021), and statistical-learning-based models (Uyanik et al. 2020; Kim et al. 2021), respectively. In this study, to compare and analyze the prediction performance of the three different types of models, ANN, MLR, and SVR are selected to predict the vessel emissions.

The ANN algorithm is developed by imitating the information handling process in human brain neurons to figure out the complex relationships implied between the input and output pairs of data (Rojas 1996). In the structure of the ANN algorithm, there are input layers, hidden layers, and output layers. Its training process is based on the forward transfer information and combined with the backpropagated errors between the predicted values and the actual values. During the training process, the connection weights between any two neurons are adjusted continuously until the errors are small enough to be accepted (Heidari et al. 2016). In this paper, the architecture of the ANN consists of one input layer with the nodes corresponding to the input feature set, a diverse number of hidden layers by hyperparameter tuning, and one output layer which is the predicted emissions. Given a neuron with m inputs, x_i is the i th input, $i \in \{1, \dots, m\}$, ω_i is the weight on the connection, b_i is the bias, and $f()$ is the activation function, then y is the output according to Eq. (1). The training process of the ANN is to find the proper values of ω and b based on the loss function as shown in Eq. (2), where n is the number of samples, D_i is the actual value and y_i is the predicted value.

$$y = f\left(\sum_{i=1}^m (\omega_i x_i + b_i)\right) \quad (1)$$

$$\arg \min_{\omega, b} \frac{1}{n} \sum_{i=1}^n (D_i - y_i)^2 \quad (2)$$

The SVR algorithm is extended from the support vector machine (SVM) algorithm (Smola and Schölkopf 2004). In the SVR algorithm, the linearly inseparable features (x_1, x_2, \dots, x_m) are transferred to be linearly separable by mapping into a higher dimensional space as shown in Eq. (3). ω is the weight vector, b is the bias and $\varphi(x)$ is the kernel function.

$$y = \omega\varphi(x) + b \quad (3)$$

$$\min \frac{1}{2} \|\omega\|^2 \quad (4)$$

subject to:

$$|D_i - y_i| \leq \varepsilon, \quad i = 1, \dots, n \quad (5)$$

In the higher dimensional space, a hyperplane is set approximately. The training process of the SVR algorithm is to minimize the largest distance between the mapped points and the hyperplane to maximize the distance from the nearest mapped points to the hyperplane (Eq. (4)), with the constraint that the distance from any mapped point to the hyperplane is no more than a defined tolerant margin ε (Eq. (5)).

The MLR algorithm uses the regression analysis in mathematical statistics to identify the relationship between one dependent variable (y) and multiple independent variables (x_1, x_2, \dots, x_m) as formulated by Eq. (6) (Tranmer et al. 2020). The training process of the MLR algorithm aims to identify the proper values of $\theta_0, \theta_1, \dots, \theta_m$ to minimize the residual sum of squares between the actual values and the predicted values as shown in Eq. (7).

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_j x_j + \dots + \theta_m x_m \quad (6)$$

$$\arg \min_{\theta} \left\{ \sum_{i=1}^n (D_i - y_i)^2 \right\} \quad (7)$$

3.6 Performance assessment

To evaluate the performance of the prediction model, we adopt the metrics RMSE and mean absolute error (MAE) and R^2 . A five-fold cross-validation is applied with the performance assessment metrics being the average values. The RMSE is given by the standard deviation of the residuals that express the distance between the predicted

values and the actual values on the regression curve. The MAE is the absolute error between the predicted values and the actual values. Lower values for RMSE and MAE represent better prediction results. The R^2 provides an assessment of the accuracy rate of the prediction model with the best value of 1. The calculations of RMSE, MAE and R^2 are shown in Eqs. (8), (9), and (10), respectively, where y is the actual value, \hat{y} is the predicted value, \bar{y} is the mean of the y values, and n is the number of samples.

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

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

4 Numerical experiments

This section presents the numerical experiments based on the models using F_1 and F_2 , where the definitions are given in Sect. 3.4. The training uses subsets of the presented seven features on 3,808 samples and employs ANN, SVR and MLR. The numerical experiments are carried out using sklearn (Pedregosa et al. 2011) based on the default hyperparameters, although their optimization is exemplified in Sect. 4.4.

First, the results based on incremental select- k -best are exemplified with the models for a specific purpose in Sect. 4.1. Second, the best-performing models for all purposes are identified in Sect. 4.2. Subsequently, a differentiation by tier and algorithm is exemplified in Sect. 4.3. Next, the best-performing models and alternative models are exemplified in Sect. 4.4. Finally, the effect of the feature *Min* on a time-dependent target is assessed in Sect. 4.5.

4.1 Results for models based on incremental select- k -best

Using the seven feature sets (see Table 5) and a specific combination described by F_1 (*Mai*, *MAN*, *CO₂*) as well as three algorithms (MLR, SVR, ANN), 21 models are trained. Differentiated by tier and algorithm, the values for the metrics R^2 , MAE and RMSE are provided in Fig. 3.

Concerning the tiers, the models of this approach yield an acceptable performance starting from T_5 , from where ANN and MLR perform better than SVR. Concerning R^2 , Fig. 3a shows that the top best-performing models with an R^2 between 0.90 and 0.92 can be found in T_5 , T_6 and T_7 using ANN and MLR. Notably, once the feature *Min* is added starting from T_5 , a large performance increase can be observed.

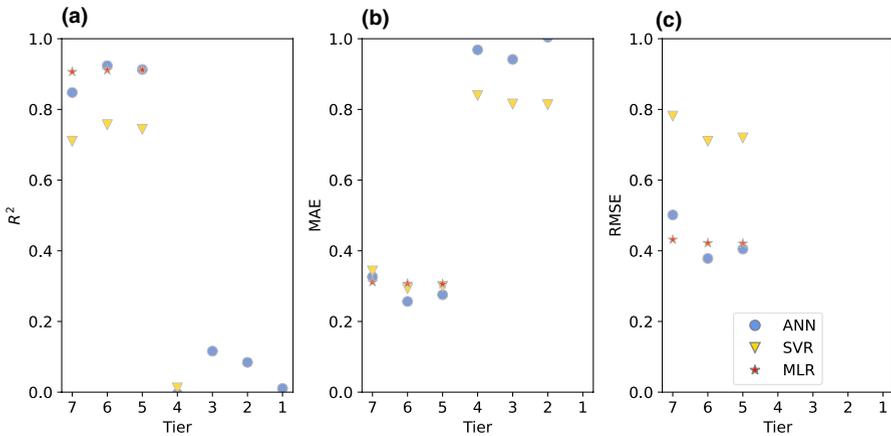


Fig. 3 Performance metrics for models based on $F_1(Mai, MAN, CO_2)$

Table 6 R^2 of each best-performing model based on $F_2(e, m, g)$

Engine	Mode	Gas			
		CO ₂	NO _x	PM _{2.5}	SO _x
Main	Transiting	0.8688	0.8724	0.8688	0.8688
	Maneuvering	0.9421	0.9145	0.9158	0.9158
Auxiliary	Transiting	0.2564	0.2268	0.2330	0.2330
	Maneuvering	0.6811	0.6556	0.6496	0.6496
	Berthing	0.7384	0.7420	0.7384	0.7384
Boiler	Transiting	0.8859	0.8859	0.8859	0.8859
	Maneuvering	0.9829	0.9829	0.9829	0.9829
	Berthing	1.0000	1.0000	1.0000	1.0000

Prior to decreasing the data dependency while trying to maximize the prediction performance, the next section presents the overall best-performing models for each purpose.

4.2 Best-performing models differentiated by engine, mode and gas

Based on the 127 feature sets and 32 combinations described by $F_2(e, m, g)$, the application of the three algorithms (MLR, SVR, ANN) results in 12,192 trained models. The subsequent Sects. 4.3, 4.4 and 4.5 make use of a subset of these models in the sense that the combinations and/or feature sets are limited with respect to different aims.

Let us now present the R^2 values for the best-performing models, irrespective of features or algorithms in Table 6. The results illustrate that the formulation of F_2 provides a competitive performance in regard of the main and boiler engines exceeding an arbitrary threshold of 0.85 R^2 . Notably, in the activity-based bottom-up

calculation, boiler emissions are based on a simple factor multiplication with the variable *Min*. In contrast, the main engine emissions depend on load factor changes during activity over time: This complication reduces the performance. Furthermore, the prediction performance regarding the auxiliary engine is significantly worse, ranging from about 0.2 to 0.75 R^2 . In general, a strong correlation between the prediction performances of different gases for the same combination of engine and mode is observed, indicating a calculation using factor multiplication.

4.3 Models differentiated by tier and algorithm

In this section, models are differentiated by tier and algorithm for selected $F_2(e, m, g)$, i.e., $F_2(Mai, MAN, CO_2)$ and $F_2(Mai, TRN, NO_x)$. The R^2 values are provided in Fig. 4a and b.

Figure 4a shows no significant performance increase after models in T_2 . In general, the sub-figures exemplify the performance diversification between algorithms depending on the constraints. For instance, Fig. 4a shows that MLR provides better results than SVR throughout most of the tiers. However, ANN achieves the best performance in most tiers. Notably, the sub-figure shows that most models reside in two clusters, either achieving an R^2 value greater than 0.7 or less than 0.2. In comparison to $F_1(Mai, MAN, CO_2)$ (see Fig. 3), the models with less features achieve a better performance. Furthermore, Fig. 4b exemplifies the advantage of MLR in a scenario, achieving the highest R^2 value. Moreover, models after T_3 show no significant increase in their performance. While the sub-figures provide an overview about the achievable prediction performance per tier, they lack the capability to show on which features the models rely. Therefore, the next section explores models, based on $F_2(Mai, MAN, CO_2)$, in more detail.

4.4 Best-performing and alternative models in each tier

In this section, the best models and alternative models for $F_2(Mai, MAN, CO_2)$, differentiated by tier and algorithm, are shown. A hyperparameter tuning is carried out for ANN and SVR. A simple grid search is applied; concerning ANN, six variants concerning the number of hidden layers and neurons are used, i.e., one or two layers with each layer having 100, 25, or 5 neurons. Regarding SVR, 16 variants concerning $C = \{0.01, 0.1, 1, 10\}$ and $\epsilon = \{0.01, 0.1, 1, 10\}$ are used. Table 7 shows the experimental results for models with a distinctive feature set. To be more clear, only the best models, trained with the same feature set but with different algorithms and hyperparameters, are shown. We exemplify the options to substitute features with one another in case of data unavailability as well as to add features in order to improve the prediction performance.

In general, one aims to minimize the number of features, preferably those requiring more effort to obtain, e.g., *Min* in different activity modes or *TKW ME* as vessel-characteristics data. This is in contrast to features such as *GT* and *IMOT* that may be easier to obtain (see Sect. 3.1). Thereby, after the training of models, a decision-making process per each port using a trade-off analysis should

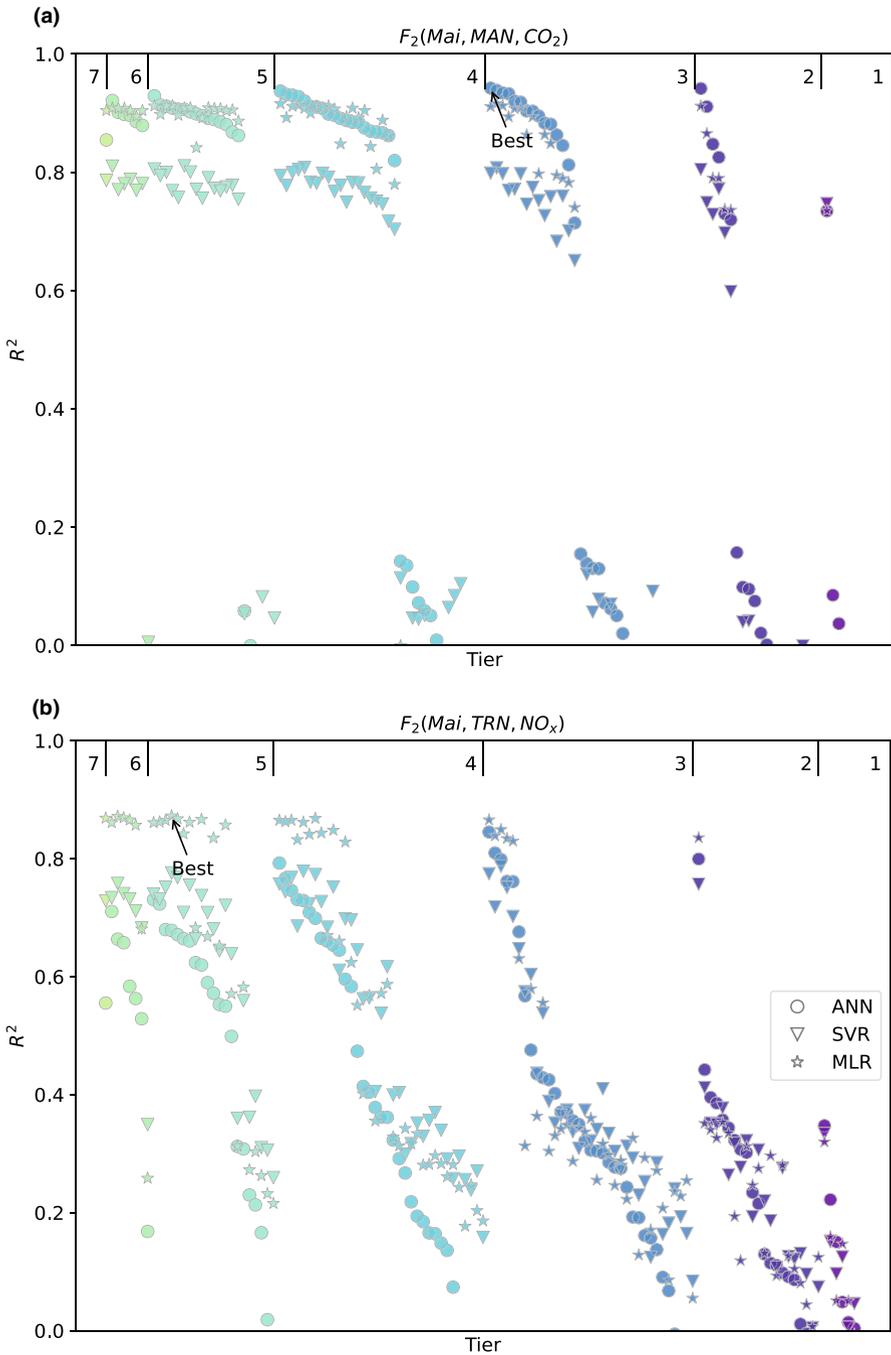


Fig. 4 R^2 for selected models in different tiers based on F_2

Table 7 Performance metrics for best-performing model and selected alternative models based on $F_2(Mai, MAN, CO_2)$, from T_1 to T_5 , ordered by tier and R^2

Tier	Feature set	Algorithm	Hyperparameter	R^2	MAE	RMSE
1	Min	ANN	100, 100	0.7761	0.3735	0.6582
1 [‡]	GT	SVR	$C = 10.0, \epsilon = 1.0$	0.1916	1.0590	1.2520
1	DS	SVR	$C = 10.0, \epsilon = 1.0$	0.1251	1.0555	1.3087
1	TKW ME	SVR	$C = 10.0, \epsilon = 1.0$	0.0661	1.1419	1.3534
1	TKW AE	SVR	$C = 0.1, \epsilon = 1.0$	-0.0530	1.2376	1.4530
2 [†]	Min, TKW ME	ANN	100, 100	0.9450	0.1680	0.3357
2	GT, Min	ANN	100	0.9108	0.3055	0.4154
2	Min, TKW AE	ANN	100, 100	0.8640	0.3259	0.5095
2	DS, Min	ANN	25, 25	0.8381	0.3396	0.5515
2	AE Num, Min	ANN	100, 100	0.7661	0.4049	0.6746
2	IMOT, Min	MLR	-	0.7371	0.4603	0.6913
2 [‡]	GT, IMOT	ANN	100,	0.1570	1.0301	1.2628
3	Min, TKW AE, TKW ME	ANN	100, 100	0.9535	0.1708	0.3014
3	GT, IMOT, Min	ANN	100	0.9329	0.2634	0.3631
3	DS, GT, Min	ANN	100, 100	0.9128	0.2363	0.3989
3	DS, IMOT, Min	ANN	25, 25	0.8452	0.3261	0.5534
3 [‡]	AE Num, GT, IMOT	ANN	100, 100	0.2065	0.7849	1.2503
4*	IMOT, Min, TKW AE, TKW ME	ANN	100, 100	0.9577	0.1688	0.2867
5	GT, IMOT, Min, TKW AE, TKW ME	ANN	100, 100	0.9516	0.1801	0.2978
5 [‡]	AE Num, GT, IMOT, TKW AE, TKW ME	SVR	$C = 1.0, \epsilon = 0.1$	0.0819	0.9089	1.3277

* Model with the highest R^2 value

[†] First model in its tier with a significant increase compared to the lower-tier's best model

[‡] First model in its tier without the feature *Min*

be conducted, between the number of expensive features and the prediction performance.

Regarding T_1 models, the best two models have a large diverging performance; it is shown that the importance of the feature *Min* far exceeds those of the remaining features. In general, a comparison of models within each tier shows that models with a feature set lacking *Min* perform significantly worse than the best models (see [‡] in Table 7). At least in this case study, this cannot be balanced by increasing the feature set with the remaining features.

As discovered in Sect. 4.3, using two features for model training can increase the performance significantly. In doing so, the results show that the best six models include *Min*. Notably, for each of these models in T_2 , the importance rank of the additional feature changed, compared to the T_1 models; see, e.g., *TKW ME* and *GT*. Only three models in T_2 achieve a performance higher than the arbitrary threshold of $0.85 R^2$.

Considering *GT* and *IMOT* as being easier to obtain, their combination in conjunction with *Min* (T_3) achieves a better performance with $0.93 R^2$ compared to their

Table 8 Performance metrics for models based on $F_2(Mai, MAN, CO_2)$ using features *Min* and *GT*

Tier	Feature set	Algorithm	Hyperparameter	R^2	MAE	RMSE
1	GT	SVR	$C = 10.0, \epsilon = 1.0$	0.1916	1.0590	1.2520
1	GT	ANN	25, 25	0.0575	1.0694	1.3426
1	GT	MLR	–	– 0.0934	1.2257	1.4665
2	GT, Min	ANN	100	0.9108	0.3055	0.4154
2	GT, Min	MLR	–	0.8653	0.3752	0.5050
2	GT, Min	SVR	$C = 10.0, \epsilon = 0.1$	0.7809	0.3278	0.6520

separate utilization with *Min* (T_2) that achieves 0.91, 0.73 R^2 , respectively. The T_3 model could be considered a viable substitute that achieves a similar performance than the T_2 model using *Min* and *TKW ME* that achieves 0.945.

The model with the overall best performance achieves a R^2 value of 0.9577 and belongs to T_4 ; however, the difference to the best model in T_2 that achieves 0.9450 is minor. Hyperparameter tuning increases the best model's performance insignificantly.

4.5 Effect of the feature minutes on a time-dependent target

Based on the numerical results obtained in Sect. 4.3, models using the features *Min* and *GT*, based on $F_2(Mai, MAN, CO_2)$, are revisited by listing the R^2 values, the adopted algorithms and hyperparameters completely in Table 8. In regard of the activity-based bottom-up methodology, which considers the geographic environment with two main waterways of different sailing distances, the table shows the importance of the feature *Min* being included in the training of the models.

In doing so, the models achieve an increased performance using ANN with an R^2 value of 0.9108. Furthermore, the importance of experimenting with different algorithms is exemplified as significant performance gaps are observed.

5 Discussion

Taken together, we show that in the given case study, some prediction models for the same purpose provide a similar performance while requiring less data. In light of the numerical experiments, three aspects should be discussed, namely, the case-by-case use of prediction models, the feature dependency between models for different purposes, and the life cycle of prediction models.

Arguably, given inconsistent vessel-characteristics data pertinent to a single port and being the input to EI formulations, the use of different prediction models that fit the available data at hand on a case-by-case, i.e., vessel-by-vessel basis may constitute an advantage: The prediction models using ANN and SVR are able to capture strong nonlinearity using few features considering the port and calling vessels' characteristics and their typical trajectories. This may be difficult to realize with traditional calculation formulation methods (Le et al. 2020; Jassim et al. 2022). For

instance, if the essential input parameters of an activity-based bottom-up calculation formulation, such as the speed or load factor in small time intervals are missing, the formulation requires assumptions as substitutes. For nonlinear relationships, setting up groups of vessels to approximate, e.g., load factors, even by few of the available features, such as ship size or engine power, requires additional effort and may lead to inferior approximations than using ANN. Along this line of thinking, it is expected that establishing a range of prediction models with different feature sets, as presented here, takes less effort than adapting calculation formulations or setting up assumptions for the same. We exemplify this idea by exploring a range of prediction models for the purpose of predicting CO₂ emissions by the main engine in the activity mode maneuvering. Here, the prediction performance is 0.94 and 0.91 R² under the use of ANN and the features activity time in minutes in conjunction with the total power of the main engine and the GT, respectively. Moreover, it is shown that the prediction performance concerning emissions originating from the main engine and the boiler by far exceeds those concerning the auxiliary engine. Models for boiler emissions achieve the best performance, which can be explained with a simple linear relationship with the activity time in minutes that is easy to grasp by MLR. A major drawback is that the activity time is still required to achieve an acceptable performance, which in turn requires, e.g., obtaining and processing AIS data.

Furthermore, we demonstrate that higher tier models could achieve a higher performance until the relevant features are exhausted. Some features can hardly be substituted (*Min*) while other features may be substituted (*TKW ME*) with an acceptable loss in the prediction performance; thus, some solution space for a trade-off analysis is provided. However, it is important to highlight that such a trade-off analysis for one purpose cannot be carried out independently if models for other purposes are necessary; thus maximizing the overlap of features between those models has to be considered as well. One way to remedy this issue could be the simplification of the framework to reduce the number of models, by predicting activity mode emissions irrespectively of the engine or vice versa.

Another important concern, not limited to this study, represents the life cycle of prediction models. Generic EI formulations and their implementations in information systems can be adapted to the geographic and operational characteristics of any port. In contrast, the trained models are a black box that cannot be adapted. The performance of those trained models likely declines with the application in ports with changing characteristics. This even applies to the same port, e.g., through the implementation of speed reduction policies, or with the retrofitting of vessels with scrubbers. Similarly, caution is necessary if employing *volatile* vessel ratings as features that are influenced by policies, for instance, this relates to the operational carbon intensity rating of a vessel. The rating code (A–E) can change over time without changing a vessel's characteristics and fuel efficiency, that is, the rating code is planned to be influenced by both the carbon intensity indicator (carbon emissions per unit transport work), as well as by policy-based declining upper bounds for each rating code over the years (Wang et al. 2021). In these cases, carrying out the entire process presented in Fig. 1, including the EI calculation based on an updated methodology and model training, is required.

6 Conclusions

In this paper, we introduce a framework to create port-related vessel EIs, which is motivated by the lack of continuous EIs in ports. We investigate whether our approach yields acceptable prediction models with a similar performance while reducing data requirements. The managerial impact is that a range of acceptable prediction models can be subject to a trade-off analysis under the constraints discussed in Sect. 5. For carrying out the proposed approach, different algorithms, such as ANN, MLR and SVR, should be employed to achieve better prediction performances of the models.

One important limitation of this work is that the investigation is constrained to container ships and to a time scope of three months. This is in contrast to EIs in practice that usually exhibit an annual quantification of multiple vessel types; thus, by extending the scope, uncertainty about the applicability can be mitigated and relevance to practice can be improved. Second, while the prediction models for the main engine provide an acceptable performance, the auxiliary engine models require improved feature engineering based on the applied EI methodology that provides the target data. For instance, feature sets could be extended to account for the number of reefer containers that consume energy, i.e., the load factor of auxiliary engines for reefer container ships is about twice as high as for container ships (EPA 2009). Although this work explores a range of popular algorithms, including ANN and SVR that can pick up nonlinear patterns, the hyperparameter tuning could be improved by adopting more sophisticated methods such as the Gaussian process-based Bayesian optimization (Snoek et al. 2012), the tree-structured Parzen estimator approach (Bergstra et al. 2011) and the sequential model-based optimization (Hutter et al. 2011). Finally, algorithms other than ANN, MLR and SVR and their tailoring could be employed to improve the prediction performance.

While addressing the above-mentioned limitations will benefit the applicability of the approach within a port, an interesting research avenue concerns sharing trained models between ports. Future work could evaluate the performance of prediction models applied in one port but trained in another, to avoid the need to carry out another activity-based bottom-up calculation. This may spark interest in the community of ports and promote the use of predictive tools to create EIs, contributing to the development of green ports. Another avenue for future research could be the implementation of EIs for other modes of transportation such as rail and public transport (see, e.g., Chipindula et al. 2022).

Acknowledgements Open Access funding enabled and organized by Projekt DEAL. The authors thank the reviewers for their valuable comments on the previous versions of this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission

directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ağbulut Ü (2022) Forecasting of transportation-related energy demand and CO₂ emissions in Turkey with different machine learning algorithms. *Sustain Prod Consump* 29:141–157. <https://doi.org/10.1016/j.spc.2021.10.001>
- Altikat S (2021) Prediction of CO₂ emission from greenhouse to atmosphere with artificial neural networks and deep learning neural networks. *Int J Environ Sci Technol* 18(10):3169–3178. <https://doi.org/10.1007/s13762-020-03079-z>
- Bergstra J, Bardenet R, Bengio Y, Kégl B (2011) Algorithms for hyper-parameter optimization. In: Shawe-Taylor J, Zemel R, Bartlett P, Pereira F, Weinberger K (eds) *Advances in neural information processing systems*. 25th annual conference on neural information processing systems. Morgan Kaufmann, p 2546–2554
- Bhowmik S, Paul A, Panua R, Ghosh SK, Debroy D (2018) Performance-exhaust emission prediction of diesosenol fueled diesel engine: an ANN coupled MORSM based optimization. *Energy* 153:212–222. <https://doi.org/10.1016/j.energy.2018.04.053>
- Cammin P, Brüssau K, Voß S (2022) Classifying maritime port emissions reporting. *Maritime Transp Res* 3:100066. <https://doi.org/10.1016/j.martra.2022.100066>
- Cammin P, Yu J, Heilig L, Voß S (2020) Monitoring of air emissions in maritime ports. *Transp Res Part D: Transp Environ* 87:102479. <https://doi.org/10.1016/j.trd.2020.102479>
- Ceylan Z, Bulkan S (2018) Forecasting PM10 levels using ANN and MLR: a case study for Sakarya City. *Global NEST J* 20(2):281–290
- Chen X, Chen H, Yang Y, Wu H, Zhang W, Zhao J, Xiong Y (2021) Traffic flow prediction by an ensemble framework with data denoising and deep learning model. *Phys A: Stat Mech Appl* 565:125574
- Chipindula J, Du H, Botlaguduru, VSV, Choe D, Kommalapati RR (2022) Life cycle environmental impact of a high-speed rail system in the Houston-Dallas I-45 corridor. *Public Transp* 14:481–501. <https://doi.org/10.1007/s12469-021-00264-2>
- COGEA (2017) Study on differentiated port infrastructure charges to promote environmentally friendly maritime transport activities and sustainable transportation. <https://doi.org/10.13140/RG.2.2.31551.56488>
- Cullinane K, Wang TF, Song DW, Ji P (2006) The technical efficiency of container ports: comparing data envelopment analysis and stochastic frontier analysis. *Transp Res Part A* 40(4):354–374. <https://doi.org/10.1016/j.tra.2005.07.003>
- Do QH, Lo SK, Chen JF (2021) A comparative study of machine learning techniques in prediction of exhaust emissions and performance of a diesel engine fuelled with biodiesel blends. *Nature Environ Poll Technol* 20:2. <https://doi.org/10.46488/NEPT.2021.v20i02.049>
- 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. *Transp Res Part B: Methodol* 122:88–114. <https://doi.org/10.1016/j.trb.2019.02.004>
- Ekmekçiöğlü A, Ünlügençoğlu K, Çelebi UB (2021) Container ship emission estimation model for the concept of green port in Turkey. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment* 147509022110244. <https://doi.org/10.1177/14750902211024453>
- EPA (2009) Current methodologies in preparing mobile source port-related emission inventories: final report. U.S. Environmental Protection Agency, ICF International. <https://www.epa.gov/sites/production/files/2016-06/documents/2009-port-inventory-guidance.pdf>
- Fabregat A, Vázquez L, Vernet A (2021) Using machine learning to estimate the impact of ports and cruise ship traffic on urban air quality: the case of Barcelona. *Environ Modell Softw* 139:104995. <https://doi.org/10.1016/j.envsoft.2021.104995>
- Fletcher T, Garaniya V, Chai S, Abbassi R, Yu H, Van DC, Brown RJ, Khan F (2018) An application of machine learning to shipping emission inventory. *Int J Maritime Eng* 160:381–395. <https://doi.org/10.3940/rina.ijme.2018.a4.500>

- Gkerekos C, Lazakis I, Theotokatos G (2019) Machine learning models for predicting ship main engine fuel oil consumption: a comparative study. *Ocean Eng* 188:106282. <https://doi.org/10.1016/j.oceaneng.2019.106282>
- Heidari E, Sobati MA, Movahedirad S (2016) Accurate prediction of nanofluid viscosity using a multi-layer perceptron artificial neural network (MLP-ANN). *Chemometrics Intell Lab Syst* 155:73–85. <https://doi.org/10.1016/j.chemolab.2016.03.031>
- Hosseini SM, Saifoddin A, Shirmohammadi R, Aslani A (2019) Forecasting of CO₂ emissions in Iran based on time series and regression analysis. *Energy Rep* 5:619–631. <https://doi.org/10.1016/j.egy.2019.05.004>
- Hu Z, Jin Y, Hu Q, Sen S, Zhou T, Osman MT (2019) Prediction of fuel consumption for enroute ship based on machine learning. *IEEE Access* 7:119497–119505. <https://doi.org/10.1109/ACCESS.2019.2933630>
- Hutter F, Hoos HH, Leyton-Brown K (2011) Sequential model-based optimization for general algorithm configuration. In: Coello CAC (ed) *Learning and intelligent optimization*, volume 6683 of *Lecture notes in computer science*. Springer, Cham, pp 507–523
- IMO (2019) Nitrogen Oxides (NOx) - Regulation 13. International Maritime Organization. [https://www.imo.org/en/OurWork/Environment/Pages/Nitrogen-oxides-\(NOx\)-%E2%80%93Regulation-13.aspx](https://www.imo.org/en/OurWork/Environment/Pages/Nitrogen-oxides-(NOx)-%E2%80%93Regulation-13.aspx)
- Jassim MS, Coskuner G, Zontul M (2022) Comparative performance analysis of support vector regression and artificial neural network for prediction of municipal solid waste generation. *Waste Manage Res* 40(2):195–204. <https://doi.org/10.1177/0734242X211008526>
- Karagiannidis P, Themelis N (2021) Data-driven modelling of ship propulsion and the effect of data pre-processing on the prediction of ship fuel consumption and speed loss. *Ocean Eng* 222:108616. <https://doi.org/10.1016/j.oceaneng.2021.108616>
- Kee KK, Simon BYL (2019) Cloud-based IoT solution for predictive modeling of ship fuel consumption. In: *Proceedings of the 2019 8th International Conference on Software and Computer Applications*, New York, NY, USA, ACM, pp 44–49
- Kee KK, Simon BYL, Renco KH (2018) Artificial neural network back-propagation based decision support system for ship fuel consumption prediction. In: *IET Conference Proceedings*. Institution of Engineering and Technology, Kuala Lumpur, Malaysia, p 13
- Kim YR, Jung M, Park JB (2021) Development of a fuel consumption prediction model based on machine learning using ship in-service data. *J Marine Sci Eng* 9(2):137. <https://doi.org/10.3390/jmse9020137>
- Kraskov A, Stögbauer H, Grassberger P (2004) Estimating mutual information. *Phys Rev E* 69:066138. <https://doi.org/10.1103/PhysRevE.69.066138>
- Le LT, Lee G, Park KS, Kim H (2020) Neural network-based fuel consumption estimation for container ships in Korea. *Marit Policy Manage* 47(5):615–632. <https://doi.org/10.1080/03088839.2020.1729437>
- Lee H, Lee T (2021) Demand modelling for emergency medical service system with multiple casualties cases: k-inflated mixture regression model. *Flexible Serv Manufact J* 33(4):1090–1115. <https://doi.org/10.1007/s10696-020-09402-7>
- Li Q, Qiao F, Yu L (2017) A machine learning approach for light-duty vehicle idling emission estimation based on real driving and environmental information. *Environ Poll Clim Change* 1:1. <https://doi.org/10.4172/2573-458X.1000106>
- Merk O (2014). Shipping emissions in ports. <http://www.internationaltransportforum.org/jtrc/DiscussionPapers/DP201420.pdf>
- Notteboom T, Pallis A, Rodrigue JP (2021) *Port economics, management and policy*. Routledge, London
- Nourmohammadzadeh A, Voß S (2022) A robust multiobjective model for the integrated berth and quay crane scheduling problem at seaside container terminals. *Annals Math Artif Intell* 90:831–853. <https://doi.org/10.1007/s10472-021-09743-5>
- Ortiz AG, Guerreiro C, Soares J, Antognazza F, Gsella A, Houssiau M, Lükewille A, Öztürk E, Horálek J, Targa J, Colette A (2019) *Air quality in Europe - 2019 report*, Volume No 10/2019 of EEA report. Luxembourg: Publications Office of the European Union
- Panapakidis I, Sourtzi VM, Dagoumas A (2020) Forecasting the fuel consumption of passenger ships with a combination of shallow and deep learning. *Electronics* 9(5):776. <https://doi.org/10.3390/electronics9050776>
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830

- Peng Y, Liu H, Li X, Huang J, Wang W (2020) Machine learning method for energy consumption prediction of ships in port considering green ports. *J Cleaner Prod* 264:121564. <https://doi.org/10.1016/j.jclepro.2020.121564>
- POLA (2019) Inventory of air emissions for calendar year 2018. Port of Los Angeles. https://kentico.portoflosangeles.org/getmedia/0e10199c-173e-4c70-9d1d-c87b9f3738b1/2018_Air_Emissions_Inventory
- Rojas R (1996) *Neural networks*. Springer, Heidelberg
- Ross BC (2014) Mutual information between discrete and continuous data sets. *PLoS one* 9(2):e87357. <https://doi.org/10.1371/journal.pone.0087357>
- Saha D, Basso B, Robertson GP (2021) Machine learning improves predictions of agricultural nitrous oxide (N₂O) emissions from intensively managed cropping systems. *Environ Res Lett* 16(2):024004. <https://doi.org/10.1088/1748-9326/abd2f3>
- Schaub M, Finger G, Riebe T, Dahms F, Hassel E, Baldauf M (2019) Data-based modelling of ship emissions and fuel oil consumption for transient engine operation. In *OCEANS 2019 - Marseille*. IEEE, pp 1–5. <https://doi.org/10.1109/OCEANS.2019.8867061>
- Shams SR, Jahani A, Kalantary S, Moeinaddini M, Khorasani N (2021) The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models for predicting SO₂ concentration. *Urban Clim* 37:100837. <https://doi.org/10.1016/j.uclim.2021.100837>
- Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14(3):199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
- Snoek J, Larochelle H, Adams RP (2012) Practical Bayesian optimization of machine learning algorithms. In: Pereira F, Burges CJ, Bottou L, Weinberger KQ (eds) *Advances in neural information processing systems*. pp 2951–2959
- Stauder M, Kühl N (2021) AI for in-line vehicle sequence controlling: development and evaluation of an adaptive machine learning artifact to predict sequence deviations in a mixed-model production line. *Flexible Serv Manuf J*. <https://doi.org/10.1007/s10696-021-09430-x>
- Tranmer M, Murphy J, Elliot M, Pampaka M (2020) *Multiple Linear Regression (2nd Edition)*. Cathie Marsh Institute Working Paper 2020-01. <http://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/multiple-linear-regression.pdf>
- Umar IK, Nourani V, Gökçekuş H (2021) A novel multi-model data-driven ensemble approach for the prediction of particulate matter concentration. *Environ Sci Poll Res Int* 28(36):49663–49677. <https://doi.org/10.1007/s11356-021-14133-9>
- Uyanık T, Karatug Ç, Arslanoğlu Y (2020) Machine learning approach to ship fuel consumption: a case of container vessel. *Transp Res Part D: Transp Environ* 84:102389. <https://doi.org/10.1016/j.trd.2020.102389>
- Venkatesh B, Anuradha J (2019) A review of feature selection and its methods. *Cybern Inf Technol* 19(1):3–26. <https://doi.org/10.2478/cait-2019-0001>
- Wan Z, Ji S, Liu Y, Zhang Q, Chen J, Wang Q (2020) Shipping emission inventories in China's Bohai Bay, Yangtze River Delta, and Pearl River Delta in 2018. *Marine Pollution Bulletin* 151:110882. <https://doi.org/10.1016/j.marpolbul.2019.110882>
- Wang S, Ji B, Zhao J, Liu W, Xu T (2018) Predicting ship fuel consumption based on LASSO regression. *Transp Res Part D: Transp Environ* 65:817–824. <https://doi.org/10.1016/j.trd.2017.09.014>
- Wang S, Psaraftis HN, Qi J (2021) Paradox of international maritime organization's carbon intensity indicator. *Commun Transp Res* 1:100005. <https://doi.org/10.1016/j.commtr.2021.100005>
- Xie M, Wu L, Li B, Li Z (2020) A novel hybrid multivariate nonlinear grey model for forecasting the traffic-related emissions. *Appl Math Modell* 77:1242–1254. <https://doi.org/10.1016/j.apm.2019.09.013>
- Yan R, Wang S, Du Y (2020) Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship. *Transp Res Part E: Logist Transp Rev* 138:101930. <https://doi.org/10.1016/j.tre.2020.101930>
- Yan R, Wang S, Psaraftis HN (2021a) Data analytics for fuel consumption management in maritime transportation: status and perspectives. *Transp Res Part E: Logist Transp Rev* 155:102489. <https://doi.org/10.1016/j.tre.2021.102489>
- Yan R, Wang S, Zhen L, Laporte G (2021b) Emerging approaches applied to maritime transport research: past and future. *Commun Transp Res* 1:100011. <https://doi.org/10.1016/j.commtr.2021.100011>
- Yuan Z, Liu J, Liu Y, Yuan Y, Zhang Q, Li Z (2020) Fitting analysis of inland ship fuel consumption considering navigation status and environmental factors. *IEEE Access* 8:187441–187454. <https://doi.org/10.1109/ACCESS.2020.3030614>

- Zhao J, Lu J, Chen X, Yan Z, Yan Y, Sun Y (2022) High-fidelity data supported ship trajectory prediction via an ensemble machine learning framework. *Physica A: Stat Mech Appl* 586:126470. <https://doi.org/10.1016/j.physa.2021.126470>
- Zheng J, Zhang H, Yin L, Liang Y, Wang B, Li Z, Song X, Zhang Y (2019) A voyage with minimal fuel consumption for cruise ships. *J Cleaner Prod* 215:144–153. <https://doi.org/10.1016/j.jclepro.2019.01.032>
- Zhou T, Hu Q, Hu Z, Zhen R (2021) An adaptive hyper parameter tuning model for ship fuel consumption prediction under complex maritime environments. *J Ocean Eng Sci*. <https://doi.org/10.1016/j.joes.2021.08.007>
- Zhu Y, Zuo Y, Li T (2021) Modeling of ship fuel consumption based on multisource and heterogeneous data: case study of passenger ship. *J Marine Sci Eng* 9(3):273. <https://doi.org/10.3390/jmse9030273>
- Zuo Q, Zhu X, Liu Z, Zhang J, Wu G, Li Y (2019) Prediction of the performance and emissions of a spark ignition engine fueled with butanol-gasoline blends based on support vector regression. *Environ Progr Sustain Energy* 38(3):e13042. <https://doi.org/10.1002/ep.13042>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Philip Cammin is a Ph.D. candidate at the Institute of Information Systems at the University of Hamburg. He received a M.Sc. degree in Business Informatics from the University of Rostock. Previous positions include IT consultant in the automotive industry. His current research interests include green ports and approaches for establishing air emissions inventories on a large scale using machine learning.

Jingjing Yu is a Ph.D. candidate at the Institute of Information Systems at the University of Hamburg and the Faculty of Infrastructure Engineering at the Dalian University of Technology. Her current research is concerned with logistics optimization and data mining within the context of smart ports. Her research aims to improve the operation scheduling of ports to increase service efficiency and reduce energy consumption and emissions.

Stefan Voß is professor and director of the Institute of Information Systems at the University of Hamburg, Germany. Moreover, he serves as dean of the Hamburg Business School (School of Business Administration). Previous positions include full professor and head of the department of Business Administration, Information Systems and Information Management at the University of Technology Braunschweig (Germany) from 1995 up to 2002. He holds degrees in Mathematics (diploma) and Economics from the University of Hamburg and a Ph.D. and the habilitation from the University of Technology Darmstadt. His current research interests are in quantitative/information systems approaches to supply chain management and logistics including public mass transit and telecommunications. He is author and co-author of several books and several hundred papers in various journals. In the German *Handelsblatt* ranking he is continuously within the top 20 professors in business administration within the German-speaking countries. The most recent *Wirtschaftswoche* rankings list him among the top 10. Stefan Voß serves on the editorial board of some journals including being Editor of *Public Transport*. He is frequently organizing workshops and conferences. Furthermore, he is consulting with several companies.