Evaluation of Vessel CO_2 Emissions Methods using AIS Trajectories

Song Wu
Université Libre de Bruxelles
Brussels, Belgium
Aalborg University
Aalborg, Denmark
song.wu@ulb.be
songw@cs.aau.dk

Kristian Torp
Aalborg University
Aalborg, Denmark
torp@cs.aau.dk

Mahmoud Sakr
Université Libre de Bruxelles
Brussels, Belgium
mahmoud.sakr@ulb.be

Esteban Zimányi
Université Libre de Bruxelles
Brussels, Belgium
esteban.zimanyi@ulb.be

ABSTRACT

Accurate estimation of shipping CO_2 emissions is important for developing regulations to combat the greenhouse effect. Many shipping CO_2 emissions models have been proposed in the past decades. However, most of them are only validated for a few specific ships, and there is a lack of data-driven validation and comparison of these models on a large scale. To fill this gap, this study proposes a general evaluation framework to quantitatively validate and compare different emission models. This framework is based on data integration of three types of data sources: ship technical details, AIS trajectory, and weather. Along with emission models, these data are fed into three carefully-designed modules that perform analysis at both grid and trajectory level as well as use annually aggregated fuel consumption ground truth. Extensive experiments are conducted on one-month data from 1,571 ships passing Danish waters to demonstrate the utility of the framework and insights into the accuracy of five popular CO_2 emission models are presented.

CCS CONCEPTS

• **Information systems → Data analytics**; **Mobile information processing systems**.

KEYWORDS

AIS, trajectories, emission, CO_2

ACM Reference Format:
Song Wu, Kristian Torp, Mahmoud Sakr, and Esteban Zimányi. 2023. Evaluation of Vessel CO_2 Emissions Methods using AIS Trajectories. In *Symposium on Spatial and Temporal Data (SSTD '23), August 23–25, 2023, Calgary, AB, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3609956.3609960

SSTD '23, August 23–25, 2023, Calgary, AB, Canada

ACM ISBN 979-8-4007-0899-2/23/08.
https://doi.org/10.1145/3609956.3609960

1 INTRODUCTION

The Paris Agreement, a global collaborative effort to combat climate change, sets long-term goals to limit global warming. An important sector that is not regulated in the Paris Agreement is maritime shipping. Maritime shipping fulfills about 90% of global trade[1] and is relatively CO_2 friendly compared with other means of transport, e.g., aircraft and truck [5]. Nevertheless, the study in [22] shows that shipping CO_2 emissions increased from 962 million tons in 2012 to 1,056 million tons in 2018. Fuel consumption also makes up a large percentage (25%) of maritime shipping costs [5]. Improving ship energy efficiency and reducing CO_2 emissions is thus becoming an urgent issue in the maritime domain.

Policies are being put in place for a greener shipping industry. Among the initiatives are the Energy Efficiency Design Index (EEDI) by the International Maritime Organization (IMO) and the EU's Monitoring, Reporting, and Verification (MRV) system. EEDI is a CO_2 efficiency measure that came into effect in 2013 and concerns all newly-built ships. Through its four progressive phases, EEDI ensures that ships are more and more energy-efficient. For example, Phase 3 of EEDI requires a reduction of CO_2 emissions by 30% for ships built after 2025 [5]. Based on the reference formula for each ship type [3], a ship will not be approved by authorities if it fails to meet the EEDI requirements [5]. In 2018, the IMO adopted an "Initial Strategy" aimed at halving shipping emissions by 2050 [17] when compared with the year 2008. Introduced in 2017, the EU's MRV system demands that ships above 5,000 gross tonnage report CO_2 emissions data for their maritime transport activities in the EU area [5]. A verified annual report for each ship is then released on the MRV website[2] for public use. For example, the MRV dataset was used to study CO_2 emissions from ferries in [18]. In this work, the MRV dataset is used as quasi ground-truth to validate emission results estimated by different models. Without causing confusion, the '*comparison and validation*' of models will be simply referred to as '*comparison*' hereafter. Recently, some routing services have also been developed to help reduce shipping CO_2 emissions. For

[1] https://www.oecd.org/ocean/topics/ocean-shipping/
[2] https://mrv.emsa.europa.eu/#public/emission-report

example, the GUTTA-VISIR system[3] has been deployed to compute least-CO_2 ferry routes in the Adriatic Sea [17] and bulk carrier routes in the Pacific Ocean [16, 19].

One important issue that remains unaddressed is how to accurately estimate a ship's CO_2 emissions [20, 21]. Fuel consumption and CO_2 emissions of a ship depend on multiple factors such as ship size, sailing speed, fuel type, and weather conditions. Therefore, many emission models have been proposed in the past decades, and they take into account different factors. Two main limitations exist for most of these models. Firstly, they are only validated on a few specific ships, so their applicability to other ships is unclear. Secondly, these models usually require access to ground truth fuel consumption data, which is generally not open to the public and cannot be obtained at a large scale.

More importantly, there is a lack of detailed experimental comparison of different CO_2 emissions models in the literature. As far as we are concerned, comparative analysis of CO_2 emissions models is only conducted in [20, 21]. However, both of them are restricted to a small area around the Strait of Gibraltar. The study in [20] assumes a ship's speed to be constant in the study area, and the study in [21] only considers one Ro-Pax ship (roll-on/roll-off passenger ship). Therefore, a more detailed comparison in a larger scope and on more ships is needed to better investigate different CO_2 emissions models.

To fill this gap, this study aims to provide a systematic evaluation framework to compare different CO_2 emission models. The main contributions in this work are:

- A general evaluation framework based on data integration is proposed to compare different CO_2 emission models.
- Up to our knowledge, this is the first work to validate multiple ship CO_2 emission models that exist in the literature against a ground-truth. We use the MRV dataset as quasi ground-truth for many ships for this purpose.
- Extensive experiments are conducted using data from 1,571 ships to show the utility of the framework.

2 RELATED WORK

Since CO_2 emissions are directly determined by fuel type and the amount of fuel consumed [22], this section briefly reviews the existing fuel consumption models in the literature.

The shipping CO_2 emissions mainly come from three sources: main engines, auxiliary engines, and boilers [22]. These onboard machinery serve different purposes: main engines are used to propel the ship forward, auxiliary engines are used to generate electrical power, and boilers are used to produce heat. In terms of fuel types, heavy fuel oil (HFO) is the dominant fuel (79%) in the shipping industry [22], due to its competitive prices [5]. However, HFO has a high sulphur content and is harmful for the environment. Other cleaner fuel types exist, such as marine diesel oil (MDO) and liquefied natural gas (LNG); but they only make up a small share in the industry, compared with HFO [22].

The existing fuel consumption models fall into three categories [4, 26, 29, 30]: white-box models (WBM), black-box models (BBM) and grey-box models (GBM).

- Based on physics laws and hydrodynamics, the main idea of WBMs is to estimate various types of resistances that a ship encounters during sailing. Then combined with estimated propulsion efficiency, these resistances are used to compute a ship's instantaneous power demand [8, 13, 23]. For example, the study in [13] considers two types of resistances: total resistance in ideal conditions and additional resistance caused by wind and wave. Since a ship's speed is widely believed to be the most important factor for fuel consumption [29], many studies (e.g., [7, 11]) also simplify the impact of resistance as a fixed coefficient. They thus consider the fuel consumption rate to be proportional to the cubic of a ship's speed. The advantage of WBMs is that they can be applied at an early stage of ship design and all their parameters are known a priori [29]. However, the accuracy of WBMs depends a lot on the various assumptions [4, 29].
- Different from WBMs that are based on assumptions, BBMs are data-driven and predict fuel consumption using statistical approaches and machine learning models [9, 27, 31]. For example, the study in [31] uses a Gaussian process to predict fuel consumption using seven variables, such as speed and draught. The advantage of BBMs is that they usually achieve a higher accuracy than WBMs, but they have high data requirements and ground truth fuel consumption data are difficult to obtain at large scale [29].
- GBM is a kind of model that combines the advantages of WBMs and BBMs. It can be constructed in two ways [29]. The first way is to use a BBM to fine-tune the parameters in a WBM [30], whereas the second way is to integrate a WBM's a priori knowledge into a BBM [2]. However, developing GBMs is a non-trivial task, and GBMs are not used as frequently as BBMs [4].

From the data perspective, three types of data sources are usually used in fuel consumption models: ship technical details, noon reports, and Automatic Identification System (AIS) data.

- A ship's fuel consumption depends on its technical information, such as engine power, fuel type, and maximum speed. Most studies (e.g., [12, 22, 24]) have used a commercial data source called IHS[4] for ship technical information. However, it would be very expensive to use a commercial ship database for sizeable fleets [23]. Instead, one highlight of this study is the integration of public technical information from multiple websites.
- Noon report data is used in many studies (e.g., [30–32]) as ground truth fuel consumption data. As its name suggests, noon report data is usually recorded by the crew every 24 hours at noon. Although there is no standard format for noon report data [30], typically the following data fields are included: date and time, ship position, cargo weight, sea and weather conditions, and daily fuel consumption of main and auxiliary engines. The main limitation of this data source is its low frequency [9, 29]. Also its quality may be subject to human errors [29]. Therefore, this data source is not used in this study, and it is mentioned here for the sake of completeness.
- AIS data. AIS is a mandatory tracking system in the maritime domain for general ships above a certain size and all passenger ships, irrespective of their size [10]. AIS data contains a ship's detailed movement history, due to its high reporting frequency

[3]https://www.gutta-visir.eu/

[4]https://ihsmarkit.com/products/ships-full-data-lake.html

from 2 seconds to 3 minutes, depending on ship speed and rate of turn [28]. So it is a suitable data source to create fine-grained emission inventories [7, 24] and has been increasingly used in recent studies [8, 22].

Although there are several review studies on ship fuel consumption models in the literature [4, 26, 29], none of these studies are data-driven, and they do not conduct a quantitative comparison between different models. Therefore, this work aims to fill this gap and provide a systematic evaluation framework to quantitatively compare different models.

3 SYSTEM DESIGN

An overview of the proposed system is first presented, then each of the required input data for the system is introduced.

3.1 System Overview

Fig. 1 shows an overview of the evaluation framework, which includes three layers: raw data layer, input layer, and analysis layer. The arrows in Fig. 1 indicate the data flow.

The raw data layer contains three types of raw data: ship technical specifications, AIS data, and weather data. Ships differ in many aspects, such as size, engine power, and design speed. These factors impact the fuel consumption of a ship. The technical specifications of ships are used in this framework to consider the differences among ships. Since different locations correspond to different amounts of traffic, detailed ship movement data is necessary to model spatio-temporal distributions of fuel consumption and CO_2 emissions in a fine-grained level. AIS data is used for this purpose. Lastly, the fuel consumption of ships is affected by the weather conditions: harsh conditions often lead to higher fuel consumption. Thus the inclusion of weather data allows for more accurate fuel consumption estimations.

Using data integration, two datasets are created from the three raw data sources: the vessel dataset contains the detailed technical information of each ship, and the trajectory dataset contains the movement history of each ship enriched with weather conditions. Together with the CO_2 emission models, these two datasets are fed as input into the analysis modules introduced next.

The evaluation framework includes three analysis modules. The grid-based analysis module analyses the spatio-temporal distributions of fuel consumption and CO_2 emissions, which can be useful for decision-makers to develop regulation policies. The trajectory-based analysis module targets representative individual trajectories, which is of interest to relevant practitioners such as ship owners. The MRV-based validation module aims to cross-check the CO_2 emission estimations with the MRV ground-truth dataset.

The framework in Fig. 1 is inspired by the EcoMark 2.0 framework in [6], which investigated the utility of 11 vehicle fuel consumption models for assigning eco-weights to road segments. However, there are several significant differences between our framework and EcoMark 2.0: (1) there is no road network in the sea, (2) weather data is not used in EcoMark 2.0, (3) height information does not make much sense in the maritime domain, compared with uphill/downhill road segments, and (4) there is a larger difference in size/weight for ships than for vehicles.

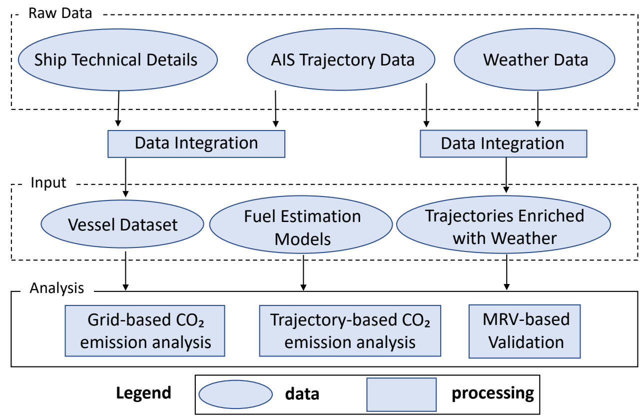


Figure 1: Overview of the Evaluation Framework

3.2 Raw Data in the System

- Ship Technical Data. Fuel consumption models require technical information of a ship to accurately estimate fuel consumption and CO_2 emissions. In this study, ship technical details are collected from six public sources[5] (Table 1): ① BalticShipping, ② Bureau Veritas, ③ FleetMon, ④ MarineTraffic, ⑤ ShipAtlas, and ⑥ VesselTracker. Among the parameters, Tons Per Centimeter (*TPC*) [1] is the required tons to lower a ship by one centimeter, and Deadweight (DWT) represents the maximum carrying capacity of a ship.

Table 1: Data sources of ship technical information

Notation	Parameter	Unit	Sources
L	length	meters	④ ⑤
B	beam	meters	④ ⑤
D	maximum draught	meters	④ ⑤
DWT	deadweight	metric tons	④ ⑤
GT	gross tonnage	register tons	④
P	maximum power	kW	② ④ ⑤ ①
S	maximum speed	knots/hour	② ④ ⑥ ③
TPC	tons per centimeter	metric tons	⑤
RPM	revolutions per minute	r/min	② ④ ⑤
$Year$	year built	integer	④

- AIS Data. In this study, the following columns in AIS data are used: the Maritime Mobile Service Identity (MMSI) number, the IMO number, timestamp, longitude, latitude, and draught. Different ships can be distinguished by their MMSI and IMO numbers, and the ship's speed can be computed using two consecutive AIS messages.
- Weather Data. Three kinds of weather data are often considered in the literature: (1) wind data including wind speed and wind direction, (2) wave data including significant wave height and wave direction, and (3) ocean current data including current speed and current direction. Depending on the area of interest, weather data can be obtained from several sources, such as Copernicus Marine

[5]These sources were accessed on Oct 2, 2022.

Environment Monitoring Service (CMEMS) [15, 17] and National Oceanic and Atmospheric Administration (NOAA). Although some estimation models (e.g. [13]) use all of wind, wave and current data, they are too complex for implementation. Therefore, only wave data is used in this study for the moment.

4 MODEL ANALYSIS

This study covers five estimation models for CO_2 emissions. The main criterion for choosing these models is that they are not specific to certain geospatial regions, e.g., certain port.

4.1 The Baseline

The main idea of this model is to assume that a certain amount of CO_2 are emitted per metric ton of freight per kilometer of transport, so the model needs to estimate the weight of cargo carried by a ship. In this study, the value of 3 grams of CO_2 / (ton · km) is used, which can be achieved by a modern ship [5]. Specifically, the following formula is used to estimate CO_2 emissions (in kg):

$$CO_{2\_Baseline} = Mass_{cargo} * Distance * 3$$

where $Distance$ is the traveling distance (in kilometers) of a ship, and $Mass_{cargo}$ is the weight of cargo (in metric tons) carried by a ship, which is estimated as:

$$Mass_{cargo} = DWT - (D - draught_{AIS}) * TPC * 100$$

Note that "100" is used to convert centimeter to meter.

4.2 The Gross Tonnage Based Approach

In [14], Carol et al. proposed a CO_2 emission model based on a ship's gross tonnage and maximum speed. In this model, the daily fuel consumption C (in metric tons per day) of general cargo ships at full power is calculated as:

$$C = 9.8197 + 0.00143 * GT$$

The CO_2 emissions are then estimated as:

$$CO_{2\_GT} = C * (T/24) * F$$

where T is the operational hours of a ship, and F is the average emission factor of CO_2 (in kg CO_2 / ton Fuel) and taken as a piecewise function of the operation mode:

$$F = 3173 * \begin{cases} 1.0, & \text{cruising if } 0.8 < \frac{speed_{AIS}}{S} \leq 1 \\ 0.48, & \text{maneuvering if } 0.2 < \frac{speed_{AIS}}{S} \leq 0.8 \\ 0.03, & \text{hotelling if } \frac{speed_{AIS}}{S} \leq 0.2 \end{cases}$$

Note that 3,173 kg CO_2 / ton Fuel is used in [14] as the maximum emission factor during cruising mode.

4.3 The Speed-Cubic Approach

In [7], a CO_2 emission model is proposed in which a ship's instantaneous power is assumed to be proportional to the cubic of its instantaneous speed. The CO_2 emissions are estimated as:

$$CO_{2\_Cubic} = P * (\frac{speed_{AIS}}{S})^3 * T * EF$$

where T is the operational hours of a ship, and EF is the emission factor (in gCO_2/kWh) depending on the engine type and fuel type.

4.4 The IMO Approach

In the fourth GHG report by the IMO [22], the CO_2 emissions are based on the instantaneous speed and draught of a vessel and its technical details. First, the demanded propulsive power when sailing at a particular speed and draught is calculated as:

$$P_{demanded} = \frac{\delta_w \cdot P \cdot (\frac{draught_{AIS}}{D})^{0.66} \cdot (\frac{speed_{AIS}}{S})^3}{\eta_w \cdot \eta_f}$$

where (1) δ_w is the speed-power correction factor for certain ship types and sizes [22], (2) η_w is the weather correction factor to represent additional power requirement caused by weather conditions, and it is taken to be 0.909 for coastal ships and 0.867 for ocean-going ships [22], (3) η_f is the correction factor to indicate the impacts of hull fouling, and taken as 0.917 [22]. Note that δ_w is set to 1 in later experiments since the current study uses different data sources for technical details.

Based on the demanded propulsive power, the CO_2 emissions are estimated as follows:

$$CO_{2\_IMO} = P_{demanded} * T * SFC * EF$$

where (1) T is the operational hours of a ship, (2) SFC is the hourly specific fuel consumption (in gFuel/kWh), and depends on the engine type, fuel type, engine load, and engine generations, (3) EF is the CO_2 emission factor (in gCO_2/gFuel) depending on the fuel type used.

Although fuel consumption and emissions from the main engine are assumed in [22] to be zero when the engine load is below 7%, this assumption is not applied in this study.

4.5 The STEAM Approach

In [11], Jalkanen et al. proposed a model called STEAM to estimate exhaust emissions of marine traffic. Different from the previous four models, the STEAM model also considers the impact of weather conditions, by using wave data to compute a speed penalty. Specifically, the angle between the wave direction and the ship is divided into four ranges: below 30°, between 30° and 60°, between 60° and 150°, and above 150°. For each of four ranges, there is a different formula to calculate the directional part (μ) of the speed penalty [11]. The STEAM model estimates the CO_2 emissions as:

$$CO_{2\_STEAM} = P * (\frac{(1 + \mu * \frac{\Delta V}{V}) * speed_{AIS}}{S + V_{safety}})^3 * T * EF$$

where (1) $\mu * \frac{\Delta V}{V}$ is the speed penalty based on the displacement of a ship and wave data, and detailed formulas can be found in [11], (2) V_{safety} is a safety margin taken as 0.5 knots/hour, (3) EF is the emission factor (in gCO_2/kWh) depending on fuel type. Note that the STEAM model assumes a fuel consumption of 200 gFuel/kWh for all engines and restricts the speed penalty to a maximum of 50%. To evaluate the wave effect on CO_2 emissions, two versions of the STEAM model (with and without wave data) are applied in this study, and they are referred to as *STEAM* and *STEAM Without Wave* in the rest of this paper, respectively.

4.6 Summary

A summary of the input data for each model is shown in Table 2, where ✓means that a parameter/variable is used in a model. Clearly,

Table 2: Input data for each CO_2 estimation model

Models	TPC	DWT	D	P	S	GT	RPM	Year	Length	Beam	Significant Wave Height	Wave Direction
$CO_{2\_Baseline}$	✓	✓	✓									
CO_{2_GT}					✓	✓						
$CO_{2\_SpeedCubic}$				✓	✓		✓					
CO_{2_IMO}			✓	✓	✓		✓	✓				
CO_{2_STEAM}				✓	✓				✓	✓	✓	✓

maximum speed and maximum power are the most popular parameters, since they are needed by the majority of models.

5 EMPIRICAL STUDIES

This section conducts a comprehensive comparison of CO_2 emission models by applying the framework in Fig. 1 to the shipping traffic around Danish waters. Firstly, the experimental setup is presented. Then a detailed application of the analysis modules is given. Since the area of study belongs to the Emission Control Area (ECA) in the North Sea and the Baltic Sea [22], all ships are assumed to use the marine diesel oil rather than the heavy fuel oil.

In terms of implementation: (1) The logic of emission models are written in Java; (2) Python is used to write scripts to collect ship technical details data; (3) All datasets are stored as CSV files; and (4) The Tableau software is used for visualization of results.

5.1 Setup

5.1.1 AIS Data. Theoretically, the evaluation framework can be used anywhere as long as the required data is available. To be more focused in this study, we apply the framework to the shipping traffic around Danish waters. Specifically, AIS data from May of 2022 is downloaded from the website of Danish Maritime Authority[6]. The following preprocessing steps are applied:

- Ship Type Filtering. As shown in the Fourth IMO Greenhouse Gas Study [22], container ships and bulk carriers contribute the largest proportion of international shipping CO_2 emissions among all ship types. So this study uses only AIS messages reported by cargo ships. This includes containers ships and bulk carriers as sub-categories.
- Spatial Filtering. The scope of area in this study was a rectangular area with a longitude range of [5, 20] and a latitude range of [52, 60], so only AIS messages located in this area are retained.
- Time-based Segmentation. In the AIS data, the time gap between two consecutive AIS messages sometimes can be several hours, this may be caused by the low AIS coverage in some regions or AIS switching-off by the crew. To facilitate later analysis, for each vessel, its AIS messages are segmented when a time gap larger than one hour is observed.
- Draught Correction. Draught information is important in several estimation models, and it is manually entered into the AIS system by the crew. However, this information is missing in about 2.3% of the AIS messages. So for each a of these AIS messages, we first found its temporally-closest neighbor n that has draught information available, then the draught of a is filled using n's draught.
- Outlier Removal. To remove outliers, the AIS messages are discarded that have a speed larger than 50 knots/hour [24, 25] *w.r.t.* the previous AIS message.
- Draught-based Segmentation. In this step, consecutive AIS messages that have the same draught values are grouped together, as needed by the Baseline and IMO models.

Therefore, the generated trajectories have two properties: (1) the time gap between any two consecutive AIS messages in a trajectory is less than one hour and (2) all AIS messages in a trajectory have the same draught values.

Table 3: The number of AIS messages after each step

step	# of AIS messages
raw AIS messages from cargo ships	63,352,466
spatial filtering	61,097,446
outlier removal	60,832,556
AIS messages from the final 1571 ships	42,039,748
removal of short or non-moving trajectories	41,024,724

Table 4: Statistics on the 8,369 trajectories

min. / avg. / max. length (km)	0.21 / 282.43 / 1,609.87
min. / avg. / max. duration (hours)	0.02 / 18.30 / 466.84
min. / avg. / max. number of AIS messages	10 / 4,982 / 148,690

5.1.2 Ship Technical Details. The trajectories after preprocessing involves 2,130 different MMSI/IMO pairs, and each pair is identified as a distinct ship in this study. The technical data for each ship was collected from the six public sources as mentioned before. Note that some key parameters can be found in multiple sources, such as maximum power and maximum speed. So after a consistency check between the sources, such parameters are determined as follows: the maximum power P is chosen in the order of Bureau Veritas, MarineTraffic, ShipAtlas, and BalticShipping. The engine power from Bureau Veritas is first checked and used when available. When it is missing in Bureau Veritas, the engine power from MarineTraffic is checked, and so on so forth. In a similar manner, the parameters L, B, D, DWT, S, and RPM are determined in their respective order, as shown in the column *Sources* in Table 1.

Since this study focuses on CO_2 emissions from main engines, two simple filters are applied to select the trajectories of this analysis: (1) a trajectory should contains at least 10 AIS messages, (2) the average speed of the whole trajectory should be larger than 1 knot/hour.

[6]https://web.ais.dk/aisdata/

In the end, 1,596 ships have all of their technical information available. However, 25 of these ships contain only short or non-moving trajectories. So experiments are conducted on 8,369 trajectories from the remaining 1,571 ships to ensure a fair comparison between the different CO_2 emission models.

Table 3 shows the number of AIS messages retained after each step. Overall, 64.8% of raw AIS data from cargo ships are used in the experiments. Table 4 shows some statistics of the 8,369 trajectories, and these trajectories differ a lot in terms of duration, length, and the number of AIS records.

To illustrate, Fig. 2 depicts the results for the ship #209316000 (MMSI) before and after preprocessing. Blue points in Fig. 2a represents the AIS messages with missing draught values. So the draught of each blue point is corrected based on its neighbor with closest timestamp. There are 12 green points in Fig. 2a with a draught of 6.3 meters. However, they are invisible because they are located in the port of Aarhus and overlap with other points. Because the trajectory formed by these 12 green points is almost still, it is excluded from later analysis. The ship's draught values are 10.1, 6.7 and 6.5 meters (Fig. 2b.) for the final 3 trajectories respectively.

5.1.3 Weather Data. The wave data required by the STEAM model are obtained from two CMEMS products: (1) The Baltic Sea product[7] has a spatial resolution of one nautical mile and a temporal resolution of one hour; (2) The North Sea product[8] has a spatial resolution of 3km (longitude) by 1.5 km (latitude) and a temporal resolution of one hour.

In this study, the AIS data is enriched with wave data that has the closest timestamp and geographical coordinates. The Baltic Sea product is used when the longitude of an AIS position is larger than 10°, otherwise the North Sea product is used.

5.2 Grid-based Emissions Analysis

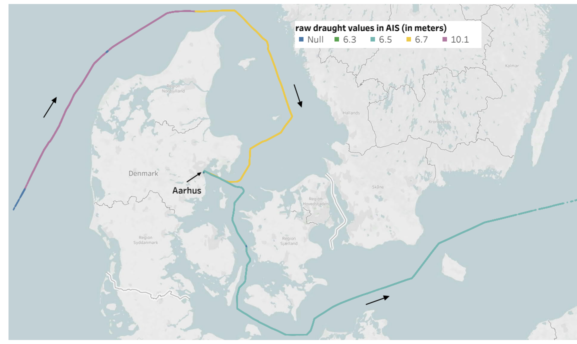
One advantage of AIS data is that it enables the fine-grained modeling of spatiotemporal distribution of CO_2 emissions. In this section, the area of study is split into grids with a size of 0.05° by 0.05° (approximately 5×5 km) and the following analysis are conducted:

5.2.1 Comparison of absolute emissions by each model. Fig. 3 shows that the Baseline approach gives the lowest estimations among all methods, whereas the GrossTonnage approach has the highest estimations. The other four models have similar estimations. Notably, the estimation by the GrossTonnage approach is almost 3 times of that by the Baseline approach. The accuracy of these models w.r.t. the MRV ground-truth will be presented later in Sec. 5.4.

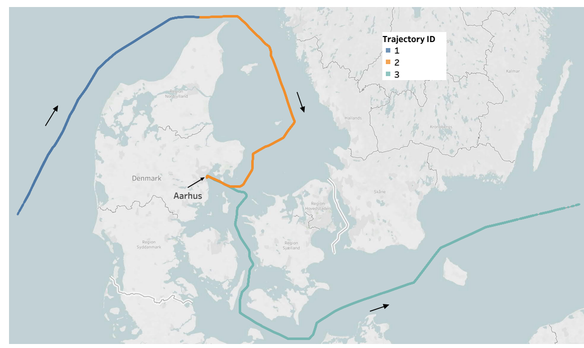
Fig. 4 depicts the spatial distribution of CO_2 emissions by each model. Note that the four models SpeedCubic, IMO, STEAM Without Wave, and STEAM give similar results. Therefore, only the result of the IMO model is shown for space reasons. The dotted box in Fig. 4 indicates the area of study. Clearly, the majority of CO_2 emissions is concentrated along the main route for passing the Danish waters.

5.2.2 Wave effects on CO_2 emissions. The effect of waves on CO_2 emissions can be observed by comparing the last two columns in Fig. 3. Surprisingly, taking wave data into account only leads to an overall 1.98% increase in CO_2 emissions. This is marginal compared with the IMO approach [22], which assumes that weather has a 15% impact for ocean-going ships and a 10% impact for coastal ships. Nevertheless, this observation agrees with the finding in [11] that reported the importance of waves to be around 2%.

It is interesting to know how the wave effect varies when the significant wave height becomes larger. To this end, the STEAM model under various wave conditions is applied by assuming that the wave direction and significant wave height are the same everywhere in the area of study. The studied wave heights range from 0 to 10 meters, increased with one meter each time. Let $E_{r,h}$ be the total emissions by the STEAM model under wave angle range r and wave height h, then the wave effect on CO_2 emissions can be



(a) raw AIS data of the ship in May of 2022



(b) the generated three trajectories of the ship after preprocessing

Figure 2: AIS data from the ship #209316000 (MMSI) before and after preprocessing

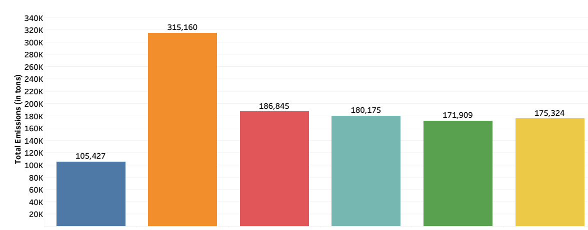
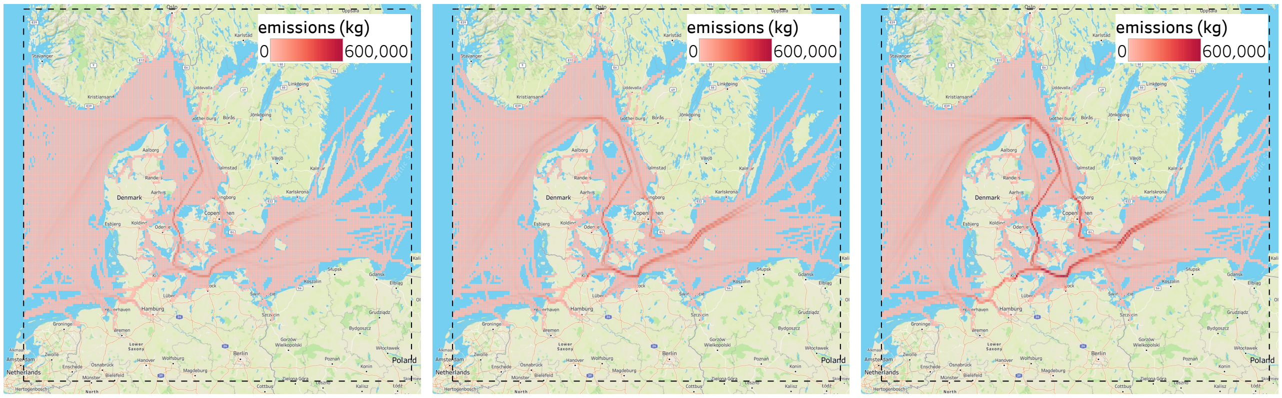


Figure 3: Total CO_2 emissions of the 1,571 ships by each model

[7] https://goo.by/FKzLj
[8] https://goo.by/vPwnP



(a) Baseline **(b) IMO** **(c) GrossTonnage**

Figure 4: Spatial distribution of CO_2 emissions by each model

evaluated as:

$$\text{Wave Effect Ratio} = \frac{E_{r,h}}{E_{*,0}}$$

where $E_{*,0}$ is the amount of emissions when the STEAM model is applied without wave data. Fig. 5 shows that the wave effect becomes larger when the wave height increases. A wave height of 10 meters can lead to an 81.47% increase in CO_2 emissions when the wave encounters the ship from the front end. Interestingly, when the wave height is above 3 meters, the effect of waves from the angle range (150, 180] becomes larger than that from the range (30, 60]. Such results are useful when the framework is applied to the regions that have more harsh weather conditions.

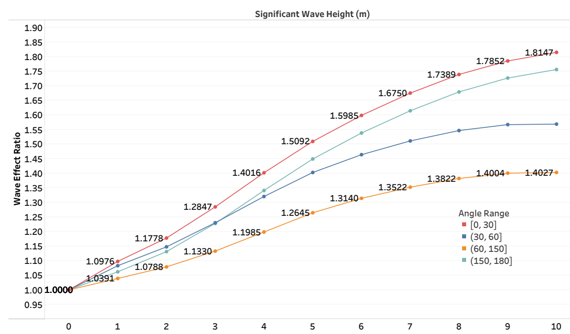


Figure 5: The effect of wave angle and wave height on the total CO_2 emissions

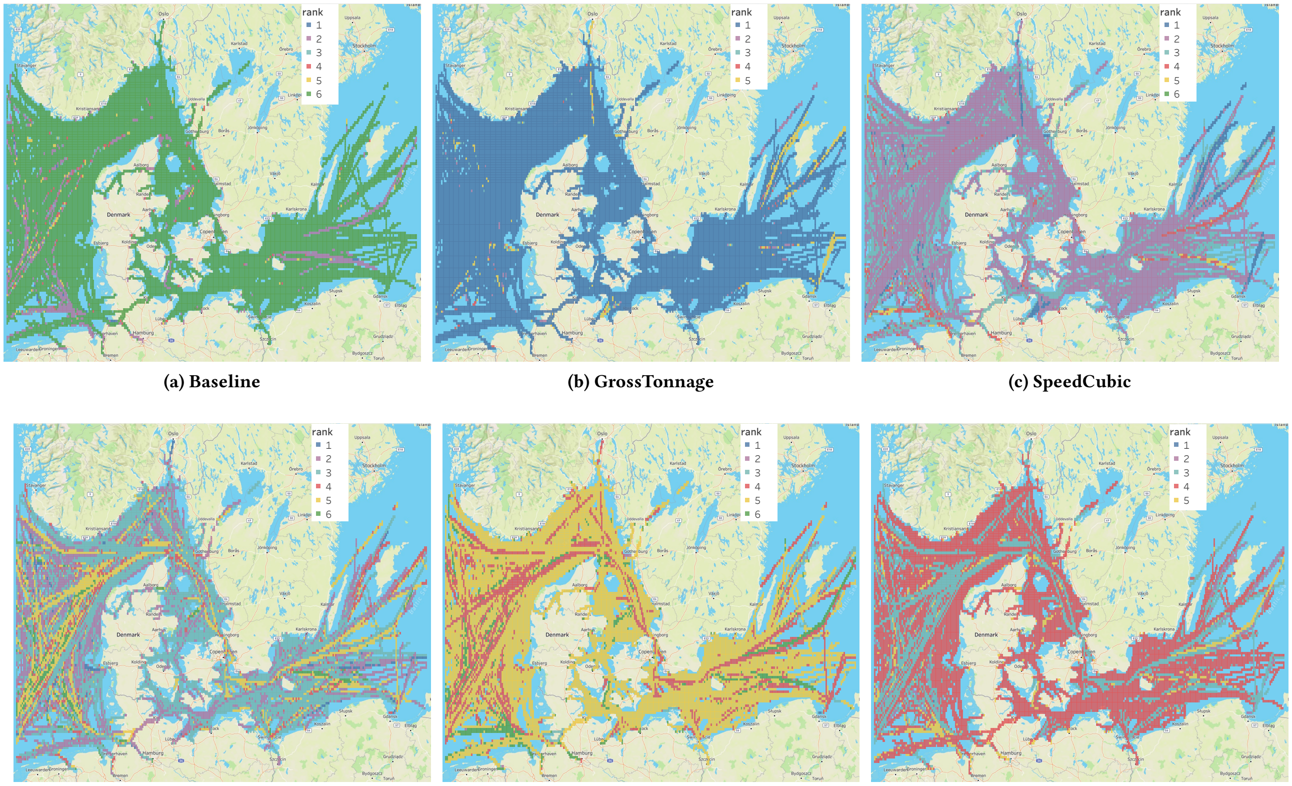
5.2.3 Grid-level ranking of each model. Instead of the overall difference in Fig. 3, this step investigates the difference between models at a grid level. Such analysis may reveal some spatial patterns. A promising way is to compute the ranking of models in each spatial grid based on their CO_2 emissions. Fig. 6 shows the resulting spatial ranking of each model, where "1" means highest emissions and "6" means lowest emissions. For the Baseline approach, it ranks 6^{th} in most grids, as expected (Fig. 3). However, its ranking is as high as 2^{nd} for some grids in the North Sea and in the southeast of Sweden. For the GrossTonnage approach, it ranks 1^{st} in most grids, as expected (Fig. 3). Nonetheless, it has a low ranking of 5^{th} for some grids between Norway and Denmark, and for some grids around the Swedish Gotland island and in the north of Poland. For the SpeedCubic approach, it ranks 2^{nd} in most grids, and ranks 3^{rd} in many other grids as well. For the remaining three approaches from Figs. 6d to 6f, a clear pattern can be observed. Their ranking is different from that in Fig. 3 for the grids connecting the North Sea and the Baltic Sea through the Sound Belt passing Copenhagen. For example, the IMO approach ranks 3^{rd} in Fig. 3, whereas its ranking decreases to 5^{th} in Fig. 6 along the Sound Belt grids. In contrast, the ranking for the STEAM Without Wave approach and the STEAM approach increases from 5^{th} to 4^{th} and from 4^{th} to 3^{rd}, respectively.

5.3 Trajectory-based Emissions Analysis

This section investigates the CO_2 emissions of representative trajectories. Such analysis is important for ship owners to know the performance of their fleets. To this end, we conduct a case study of trips that transit between Skagen (the northernmost of Denmark) and Bornholm (the easternmost of Denmark) via Copenhagen. In addition, these trips should be performed by ships above a certain size and do not stop too long in between. After some statistical analysis, the trips of interest are selected using the following conditions: (1) The deadweight of the corresponding ship is above 2,000 tons; (2) The length of the trip is below 600 kilometers; and (3) The average speed for each hour of the day during the trip is above 5 knots/hour. As a result, 192 trips from 162 ships meet these conditions, as shown in Fig. 7. Table 5 shows that the lengths of the 192 trips are close to each other. However, there is a large difference in their passing speeds, which indicates diversity in the 192 trips.

Taking the IMO approach for example, an equivalent CO_2 efficiency can be computed for each trip i by the following formula:

$$E_{CO_2,i} = \frac{CO_{2\_IMO}}{Cargo_i * Length_i}$$



(a) Baseline (b) GrossTonnage (c) SpeedCubic

(d) IMO (e) STEAM Without Wave (f) STEAM

Figure 6: Spatial ranking distribution of each model based on their CO_2 emissions

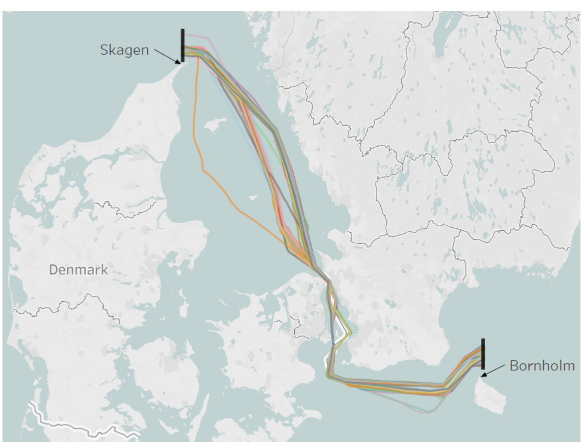


Figure 7: Routes of the 192 trips that pass Danish waters

Table 5: Statistic of the 192 trips that pass Danish waters

min. / avg. / max. length (km)	478.4 / 488.8 / 507.9
min. / avg. / max. duration (hours)	13.8 / 23.3 / 32.9
min. / avg. / max. passing speed (knots/hour)	8.0 / 11.7 / 19.0

where CO_{2_IMO} is the total CO_2 emissions (in grams) estimated by the IMO approach for a trip i, $Cargo_i$ is the weight of cargo carried (in tons) in a trip i, and $Length_i$ is the length (in km) of a trip i.

Fig. 8 depicts the results. The horizontal axis represents the deadweight (DWT) of the ship in a particular trip. The larger DWT is, the larger a ship can be considered. The vertical axis represents the average CO_2 efficiency for the corresponding trip. Overall, there is a downward trend in Fig. 8, meaning that larger ships tend to emit less CO_2 per unit of transport work.

Next, two trips of the same ship, trip#45 and trip#47 (the green and red circles in Fig. 8) are examined. This ship has a DWT of 6,410 tons and a maximum speed of 19 knots/hour. Since s is fully loaded at its DWT for the two trips, the cargo weight is the same for them. Furthermore, the IMO approach assumes a constant impact by weather conditions, thus speed becomes the only influencing factor for the CO_2 efficiency of the two trips. Table 6 gives a comparison of the two trips. It shows that a 5.2% decrease in speed (w.r.t. maximum speed) leads to a 8.7% increase in the CO_2 efficiency. This finding suggests that shipowners can probably improve the CO_2 efficiency of their fleet by speed optimization. In additional, s has a technical efficiency of 15.13 $gCO_2/t{\cdot}km$ based on the MRV dataset. This is sufficiently close to the efficiency values in Table 6. The gap between them is probably due to auxiliary engines, which are not in the scope of this study.

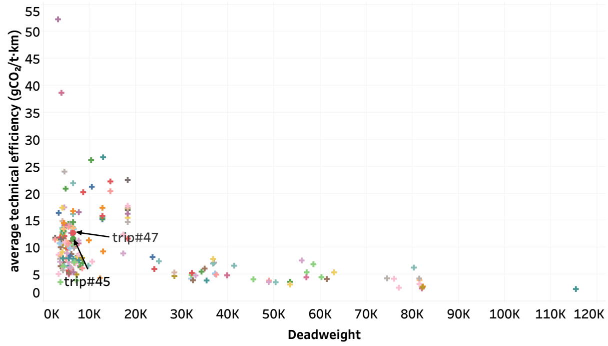


Figure 8: Average CO_2 efficiency of the 192 trips

Table 6: Comparison of trip#45 and trip#47 by the same ship

	trip#45	trip#47
length (km)	489.8	489.6
avg. passing speed (knots/hour)	13.17	14.15
CO_2 emissions (kg)	36,372	39,798
CO_2 efficiency (gCO_2/t·km)	11.58	12.68

5.4 Validation using the MRV dataset

Fig. 3 shows that the 3 grams of CO_2 / (ton · km) in the Baseline approach is probably too optimistic, thus more reasonable values should be used for each ship, e.g., based on ship length or deadweight. For this purpose, the MRV dataset is used in this study as a quasi ground truth which contains annually aggregated emission results for each included ship. At the time of writing, there are four emissions reports available in the MRV website, corresponding to each year from 2018 to 2021.

Based on the IMO number, 760 out of the 1,571 ships have matching entries in the MRV dataset. The key column in the MRV dataset is *Technical Efficiency* (in gCO_2/t·nm). This column is expressed as either EIV or EEDI of a ship in a reporting year. Since EIV is a simplified version of EEDI [3], this study gives precedence to EEDI. Namely, when both EIV and EEDI of a ship exist, EEDI will be used for that ship. Furthermore, whichever metric is chosen, the latest metric value is used. Last but not least, the EIV is multiplied by the ratio between the emission factor for diesel oil (3.206) and the emission factor for heavy fuel oil (3.114). This is because 3.114 (gCO_2/gFuel) is used in the EIV formula, whereas marine diesel oil is assumed in the area of study.

As a result, 172 ships (22.6%) use EEDI as their technical efficiency, and EIV are used for the remaining 588 ships (77.4%). Fig. 9 shows the technical efficiency of these ships in decreasing order. Clearly, the technical efficiency varies a lot from ship to ship. The horizontal line (5.556 gCO_2/t·nm) in Fig. 9 represents the Baseline approach. Among the 760 ships, 567 (74.6%) are above the line and thus less energy-efficient than suggested by the Baseline approach.

Using the MRV dataset, CO_2 emissions for the 760 ships are updated, and the result is shown as the last column in Fig. 10. Interestingly, the updated result is highly consistent with the speed-based models. Such consistency indicates that some stakeholders

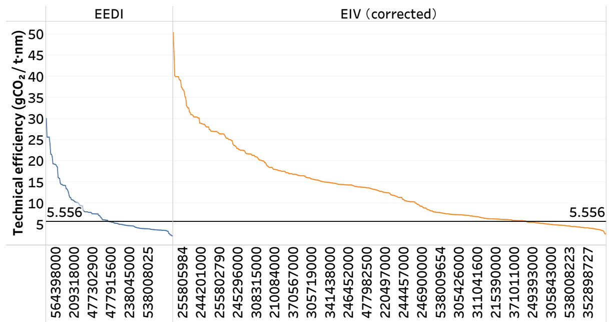


Figure 9: Technical efficiency of the 760 ships based on the MRV dataset

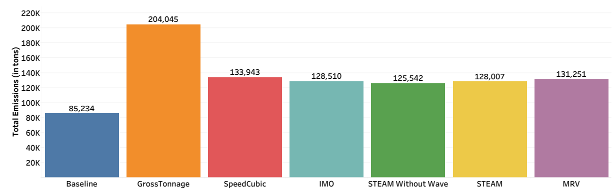


Figure 10: Total CO_2 emissions of the 760 ships by each model

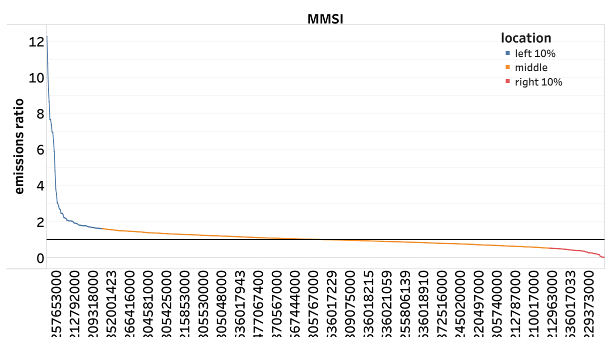


Figure 11: Ship-level emission ratio between the IMO and the MRV-based approaches

such as government agencies can rely on these speed-based models to monitor regional CO_2 emissions. Then at an individual ship level, consistency between these speed-based methods and the MRV-based approach still exist. Taking the IMO approach as an example, Fig. 11 depicts the emissions ratio between the IMO approach and MRV-based approach, for each of the 760 ships. The horizontal axis in Fig. 11 shows the MMSI numbers of these ships in decreasing order of the corresponding ratio. The middle 80% of these ships have their ratio values close to 1. Despite this, shipowners should be cautious about the application of these speed-based models, because the consistency in Fig. 11 is not as strong as that in Fig. 10. By comparing the first columns in Fig. 3 and Fig. 10, another noteworthy point is that although the 760 ships make up only 48.4% of the 1,571 ships, their emissions in Fig. 10 cover about 80.8% of that in Fig. 3. This is because the MRV scheme only regulates relatively large ships above 5,000 gross tonnage [5].

5.5 Summary of the Results

The main findings from experiments include: (1) Among the five models covered in this study, the three speed-based models return similar results. (2) Overall, most of the CO_2 emissions are located along the main sea routes. (3) Large ships tend to be more CO_2 efficient than small ships. (4) Emission results from the three speed-based models are consistent with the MRV dataset, which is used as quasi ground truth in this study. (5) The comparison of two similar trips from the same ship suggests that a ship has the potential to improve its CO_2 efficiency through speed optimization.

6 CONCLUSION

In this work, we propose a general data-driven evaluation framework to quantitatively compare different estimation models for shipping CO_2 emissions. The utility of this framework is demonstrated through extensive experiments on a large dataset around Danish waters. For future work, we plan to investigate more models and conduct experiments on shipping traffic covering larger time periods and geographical regions.

ACKNOWLEDGMENTS

This publication has been developed under the framework of the "Data Engineering for Data Science" (DEDS) project that has received funding from the European Union's Horizon 2020 programme (call identified: H2020-MSCA-ITN-EJD-2020) under grant agreement No 955895.

REFERENCES

[1] Hongtao Chi, Giulia Pedrielli, Thomas Kister, Szu Hui Ng, and Stéphane Bressan. 2015. An AIS-based framework for real time monitoring of vessels efficiency. In *2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE, 1218–1222.

[2] Andrea Coraddu, Luca Oneto, Francesco Baldi, and Davide Anguita. 2017. Vessels fuel consumption forecast and trim optimisation: A data analytics perspective. *Ocean Engineering* 130 (2017), 351–370. https://doi.org/10.1016/j.oceaneng.2016.11.058

[3] JF Faber and MJJ Hoen. 2017. *Estimated Index Values of Ships 2009-2016: Analysis of the Design Efficiency of Ships that Have Entered the Fleet Since 2009.* CE Delft.

[4] Ailong Fan, Jian Yang, Liu Yang, Da Wu, and Nikola Vladimir. 2022. A review of ship fuel consumption models. *Ocean Engineering* 264 (2022), 112405.

[5] Jørgen Villy Fenhann. 2017. CO_2 emissions from international shipping. (2017).

[6] Chenjuan Guo, Bin Yang, Ove Andersen, Christian S Jensen, and Kristian Torp. 2015. Ecomark 2.0: empowering eco-routing with vehicular environmental models and actual vehicle fuel consumption data. *GeoInformatica* 19, 3 (2015), 567–599.

[7] Weiwei Han, Wenguo Yang, and Suixiang Gao. 2016. Real-time identification and tracking of emission from vessels based on Automatic Identification System Data. In *2016 13th International Conference on Service Systems and Service Management (ICSSSM)*. IEEE, 1–6.

[8] Tina Hensel, Constance Ugé, and Carlos Jahn. 2020. Green shipping: using AIS data to assess global emissions. In *Sustainability Management Forum| NachhaltigkeitsManagementForum*, Vol. 28. Springer, 39–47.

[9] Zhihui Hu, Tianrui Zhou, Mohd Tarmizi Osman, Xiaohe Li, Yongxin Jin, and Rong Zhen. 2021. A novel hybrid fuel consumption prediction model for ocean-going container ships based on sensor data. *Journal of Marine Science and Engineering* 9, 4 (2021), 449.

[10] Clément Iphar, Aldo Napoli, and Cyril Ray. 2015. Detection of false AIS messages for the improvement of maritime situational awareness. In *Oceans 2015-mts/ieee washington*. IEEE, 1–7.

[11] J-P Jalkanen, Anders Brink, Juha Kalli, Heidi Pettersson, Jussi Kukkonen, and Tapani Stipa. 2009. A modelling system for the exhaust emissions of marine traffic and its application in the Baltic Sea area. *Atmospheric Chemistry and Physics* 9, 23 (2009), 9209–9223.

[12] Lasse Johansson, Jukka-Pekka Jalkanen, and Jaakko Kukkonen. 2017. Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution. *Atmospheric Environment* 167 (2017), 403–415.

[13] Seong-Hoon Kim, Myung-Il Roh, Min-Jae Oh, Sung-Woo Park, and In-Il Kim. 2020. Estimation of ship operational efficiency from AIS data using big data technology. *International Journal of Naval Architecture and Ocean Engineering* 12 (2020), 440–454.

[14] See Chuan Leong and Prateek Singhal. 2015. Estimation of CO_2 emission from marine traffic in Singapore Straits using Automatic Identification Systems data. In *Environmental Science and Information Application Technology*. CRC Press, 177–184.

[15] Xiaohe Li, Yuquan Du, Yanyu Chen, Son Nguyen, Wei Zhang, Alessandro Schönborn, and Zhuo Sun. 2022. Data fusion and machine learning for ship fuel efficiency modeling: Part I–Voyage report data and meteorological data. *Communications in Transportation Research* 2 (2022), 100074.

[16] Gianandrea Mannarini. 2022. The GUTTA-VISIR system for operational CO_2 emission savings from ferries. https://doi.org/10.5281/zenodo.7400954

[17] Gianandrea Mannarini, Lorenzo Carelli, Josip Orović, Charlotte P Martinkus, and Giovanni Coppini. 2021. Towards least-CO_2 ferry routes in the Adriatic sea. *Journal of Marine Science and Engineering* 9, 2 (2021), 115.

[18] Gianandrea Mannarini, Lorenzo Carelli, and Amal Salhi. 2020. EU-MRV: An analysis of 2018's Ro-Pax CO_2 data. In *2020 21st IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 287–292.

[19] Gianandrea Mannarini, Amal Salhi, Nadia Pinardi, Mario Leonardo Salinas, Lorenzo Carelli, and Giovanni Coppini. 2022. Waves and surface ocean circulation for least-carbon ship routes. https://doi.org/10.5281/zenodo.7274349

[20] Juan Moreno-Gutiérrez, Fátima Calderay, Nieves Saborido, Maria Boile, Rafael Rodríguez Valero, and Vanesa Durán-Grados. 2015. Methodologies for estimating shipping emissions and energy consumption: A comparative analysis of current methods. *Energy* 86 (2015), 603–616.

[21] Juan Moreno-Gutiérrez, Emilio Pájaro-Velázquez, Yolanda Amado-Sánchez, Rubén Rodríguez-Moreno, Fátima Calderay-Cayetano, and Vanesa Durán-Grados. 2019. Comparative analysis between different methods for calculating on-board ship's emissions and energy consumption based on operational data. *Science of the Total Environment* 650 (2019), 575–584.

[22] International Maritime Organization. 2020. Fourth IMO Greenhouse Gas Study.

[23] Stian Glomvik Rakke. 2016. *Ship emissions calculation from AIS.* Master's thesis. NTNU.

[24] Daniel A Schwarzkopf, Ronny Petrik, Volker Matthias, Markus Quante, Elisa Majamäki, and Jukka-Pekka Jalkanen. 2021. A ship emission modeling system with scenario capabilities. *Atmospheric Environment: X* 12 (2021), 100132.

[25] Andreas Tritsarolis, George-Stylianos Theodoropoulos, and Yannis Theodoridis. 2021. Online discovery of co-movement patterns in mobility data. *International Journal of Geographical Information Science* 35, 4 (2021), 819–845.

[26] Kai Wang, Jianhang Wang, Lianzhong Huang, Yupeng Yuan, Guitao Wu, Hui Xing, Zhongyi Wang, Zhuang Wang, and Xiaoli Jiang. 2022. A comprehensive review on the prediction of ship energy consumption and pollution gas emissions. *Ocean Engineering* 266 (2022), 112826.

[27] Shengzheng Wang, Baoxian Ji, Jiansen Zhao, Wei Liu, and Tie Xu. 2018. Predicting ship fuel consumption based on LASSO regression. *Transportation Research Part D: Transport and Environment* 65 (2018), 817–824.

[28] Zhaokun Wei, Xinlian Xie, and Xiaoju Zhang. 2020. AIS trajectory simplification algorithm considering ship behaviours. *Ocean Engineering* 216 (2020), 108086.

[29] Ran Yan, Shuaian Wang, and Harilaos N Psaraftis. 2021. Data analytics for fuel consumption management in maritime transportation: Status and perspectives. *Transportation Research Part E: Logistics and Transportation Review* 155 (2021), 102489.

[30] Liqian Yang, Gang Chen, Niels Gorm Malỳ Rytter, Jinlou Zhao, and Dong Yang. 2019. A genetic algorithm-based grey-box model for ship fuel consumption prediction towards sustainable shipping. *Annals of Operations Research* (2019), 1–27.

[31] Jun Yuan and Victor Nian. 2018. Ship energy consumption prediction with Gaussian process metamodel. *Energy Procedia* 152 (2018), 655–660.

[32] Tianrui Zhou, Qinyou Hu, Zhihui Hu, and Rong Zhen. 2022. An adaptive hyper parameter tuning model for ship fuel consumption prediction under complex maritime environments. *Journal of Ocean Engineering and Science* 7, 3 (2022), 255–263.