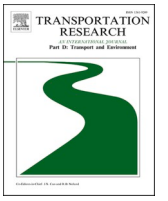
Transportation Research Part D 121 (2023) 103801

Contents lists available at ScienceDirect

Transportation Research Part D

journal homepage: www.elsevier.com/locate/trd



Comparison of deterministic and ensemble weather forecasts on ship sailing speed optimization

Xi Luo [a], Ran Yan [b,*], Shuaian Wang [a]

[a] Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Kowloon, Hong Kong Special Administrative Region
[b] School of Civil and Environmental Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore

ARTICLE INFO

Keywords:
Green shipping management
Ship energy efficiency
Ship speed optimization
Deterministic weather forecasts
Ensemble weather forecasts
Ship sailing and weather data fusion

ABSTRACT

Ship sailing speed optimization models are constructed based on prediction of ship fuel consumption, whose accuracy is highly influenced by the quality of sea and weather information. In this study, we develop two fusion methods for combining external meteorological data with ship noon report data, including the rhumb line based fusion method and the direct fusion method, and compare them in terms of accuracy in providing meteorological data. Next, we propose a framework based on the better data fusion strategy for comparing the impacts of deterministic and ensemble weather forecasts on ship speed optimization performance, enabling the evaluation of ship fuel consumptions under different speed plans based on weather forecast data available before departure. Results show that speed optimization based on ensemble weather forecasts has greater potential than that based on deterministic weather forecasts to diminish ship fuel consumption and thus to reduce greenhouse gas emissions.

1. Introduction

Maritime transport is extremely important in international trade and globalization. There are more than 80 % of the global merchandise trade that was carried by ships, with a volume of 10.9 billion metric tons in 2020 (UNCTAD, 2020; Wu et al., 2022; Tian et al., 2023b). Although ships can cost-effectively transfer large quantities of goods over long distances, over 50 % of total operational costs of ships are attributed to the cost of ship fuel consumption (Wang et al., 2018a), and thus it is a topic of great interest to shipping companies, especially when the fuel price is high or the shipping market is in recession (Wang et al., 2013; Yan et al., 2020; Yan et al., 2023). In addition, according to the report by the International Maritime Organization (IMO), approximately 3 % of global greenhouse gas (GHG) emissions are attributed to the international maritime industry (IMO, 2020). Therefore, the IMO has implemented several maritime regulations that aim to boost the energy efficiency of ships and thus diminish GHG emissions, such as Ship Energy Efficiency Management Plan (SEEMP) (IMO, 2009). Because ship companies wish to conform to these stringent emissions rules and increase revenue and competitiveness, they are eager to take advantage of practical methods to increase ship energy efficiency (Fan et al., 2022; Wang et al., 2023).

Such methods can be classified into the technical methods and the operational methods (Yan et al., 2021; Yang et al., 2019b, 2020; Du et al., 2022; Tian et al., 2023a). For existing ships, adopting technical methods, such as vessel size optimization and propeller upgrade, requires a large investment. In contrast, the operational methods recommended by SEEMP for existing ships, such as speed

* Corresponding author.
E-mail address: ran.yan@ntu.edu.sg (R. Yan).

https://doi.org/10.1016/j.trd.2023.103801
Received 10 February 2023; Received in revised form 18 April 2023; Accepted 30 May 2023
Available online 15 June 2023

optimization and weather routing, have significant energy saving potential and require relatively low investment from shipping companies (Yang et al., 2020). Among these methods, speed optimization can be easily implemented at a low cost and has great potential for improving ship energy efficiency (IMO, 2009; Yang et al., 2020). Because of these benefits, shipping companies favor speed optimization as a method for improving ship energy efficiency, which eventually reduces ship fuel consumption and GHG emissions (Beşikçi et al., 2016; Yang et al., 2020; Yan et al., 2020).

Ship speed optimization requires the establishment of an accurate fuel consumption prediction model (FCPM) in various scenarios (Yuan et al., 2020; Yang et al., 2020). Although ship fuel consumption is mainly determined by the sailing speed (Bialystocki and Konovessis, 2016; Ryder and Chappell, 1980), other factors also have an impact, such as trim condition, draught conditions, and sea and weather information (Sourtzi, 2019; Medina et al., 2020; Yu et al., 2021). In literature, ship noon reports are widely used to construct FCPMs (Gkerekos et al., 2018; Tarelko and Rudzki, 2020; Yuan and Nian, 2018). However, as the noon report is manually recorded by crews and there is only one record for each day (usually at noon according to the local time), the quality as well as the quantity of noon reports can be low, especially given that sea and weather conditions are often volatile (Li et al., 2022). Thus, incorporating sea and weather information from the external meteorological database can improve data quality as well as the data quantity, eventually improving the predicting performance of FCPMs (Yan et al., 2020; Gkerekos & Lazakis, 2020; Linh & Ngoc, 2021).

After a FCPM is constructed, a ship sailing speed optimization model is developed from the predicted fuel consumption rates. Therefore, sea and weather information also affect the reliability and feasibility of the speed optimization results generated by sailing speed optimization models. In previous studies, sea and weather conditions were usually obtained from a deterministic weather forecast, which is a single-valued estimation of the future atmospheric and oceanic state (Yan et al., 2020; Lee et al., 2018). To be specific, a deterministic weather forecast is generated by estimating the current atmospheric and oceanic state with meteorological observations collected by weather stations and satellites, and then calculating how the atmospheric and oceanic state will change based on a numerical weather prediction (NWP) model (Garcia-Moya et al., 2016; Rodwell et al., 2006). Despite the increasing skill of the NWP model forecasts, the accuracy of forecasting weather is still affected by several factors. First, due to the limitations in observational data and measurement errors, the uncertainty is thus introduced in the initial conditions (Garcia-Moya et al., 2016). Because of the chaos in the atmosphere and ocean, very tiny errors in the initial state will be amplified so that large errors will occur even in the forecast for the near future (Lorenz, 1965). In other words, when two slightly different initial conditions are input to the same NWP model, the predicted weather in a few days' time can be significantly different. In addition, numerous simplifications and approximations are made in the process of representing the complex behavior of the atmosphere and ocean, such as assumptions about the shape of the earth or the composition of the atmosphere, which can also lead to inaccuracies in the weather forecasts (Orrell et al., 2001).

To assess the impacts of these minor differences in initial conditions on the forecast results, a system of the ensemble weather forecast is applied to generate multiple forecasts. The NWP model is executed several times with slightly different initial conditions (World Meteorological Organization, 2012). The ensemble forecast can be regarded as a collection of deterministic forecasts that represent the possibilities of future weather, giving us a better idea of what weather events may happen at a specified time in the future. Thus, ship fuel consumption prediction results based on ensemble weather forecasts are more reliable than prediction results based on deterministic weather forecasts. Consequently, speed optimization based on the former prediction results can provide a better ship speed profile over a voyage, leading to a greater reduction in ship fuel consumption.

It is also notable that previous studies related to ship fuel consumption prediction rarely describe fusion methods that combine external meteorological data and the noon report data in detail, or compare the impacts of different fusion methods on the prediction accuracy of FCPMs. Thus, to improve the prediction accuracy of FCPMs, it is worth studying how to combine external meteorological data and noon report data to obtain a new dataset, which can be used to train a FCPM with higher prediction accuracy.

In this work, two data fusion methods are presented and their impacts on ship fuel consumption prediction accuracy are investigated. The better fusion method is selected to prepare the essential dataset for subsequent speed optimization. Then, the impacts of deterministic and ensemble weather forecasts on ship speed optimization based on the speed optimization model (Yan et al., 2020) are compared. The contributions of this paper are twofold:

First, it describes two fusion methods for combining external meteorological data and noon report data and thus provides new insights into ways of remedying the data quality and quantity issues of noon reports. The prediction accuracies of the two FCPMs trained by the datasets generated by the two fusion methods are compared, and the results suggest that a better data fusion strategy can improve prediction model performance. This provides a new research perspective for improving the performance of FCPMs.

Second, this study provides a framework for comparing the overall fuel consumption of a voyage based on speed profiles optimized under the deterministic versus the ensemble weather forecast settings when the actual fuel consumption under different conditions is not available. This framework demonstrates that speed optimization based on ensemble weather forecasts can reduce ship fuel consumption by 1 % compared to that based on deterministic weather forecasts. According to the IMO, total maritime fuel consumption reached 213 million tons in 2019 (IMO, 2019). If ensemble weather forecasts had been used in the speed optimization, around 2 million tons of fuel would have been saved in 2019. In addition, applying our findings allows maritime companies to operate their fleets more efficiently.

2. Literature review

Recently, due to the accessibility of large quantities of ship energy efficiency data and increases in computation resources, researchers have attempted to develop FCPMs based on machine learning algorithms (Yan et al., 2021). Several common machine learning algorithms include the tree-based model (Li et al., 2022; Soner et al., 2018), artificial neural networks (Jeon et al., 2018;

Petersen et al., 2012a,b; Kim et al., 2021; Zhu et al., 2021), regularized linear regression (Soner et al., 2019; Wang et al., 2018b), k-nearest neighbors models (Peng et al., 2020; Uyanık et al., 2020), and Gaussian process (Petersen et al., 2012a; Yuan and Nian, 2018).

Based on the ship fuel consumption predictions, a ship sailing speed optimization model that minimizes the overall fuel consumption of a voyage can be then proposed; this whole framework is termed a two-phase ship sailing optimization model (Du et al., 2019; Yan et al., 2020). For example, Wang and Meng (2012) first use the linear regression method to calibrate the relationship between the ship fuel consumption per day and the ship sailing speed using historical ship operation data; then, the authors propose an optimization model which optimizes the ship sailing speeds for a given route. Du et al. (2019) optimize the speed and trim for the container ship by developing a two-phase optimization method. In phase one, the authors propose a ship fuel consumption model based on artificial neural networks. In phase two, this research conducts several methods related to speed and trim optimization to cut down fuel consumption. In addition, Yan et al. (2020) propose a two-phase speed optimization model. In phase one, this study constructs a random forest model to estimate ship fuel consumption under various scenarios. In phase two, by leveraging such the prediction model, the authors establish a mixed integer linear programming model for optimizing sailing speeds, with constraints that the estimated arrival time must be less than the expected time at the destination port.

Most papers on ship fuel consumption prediction usually consider meteorological information, such as ocean currents, wind waves and swells, and wind force and direction, when predicting ship fuel consumption (Psaraftis and Kontovas, 2014; Sourtzi, 2019; Yan et al., 2021); some obtain sea and weather information based on public meteorological data (Lee et al., 2018; Li et al., 2022; Du et al., 2022a; Du et al., 2022b). However, few studies describe in detail a fusion method to combine two or more data sources, or discuss the impact of different fusion methods on the prediction accuracy of FCPMs. In addition, the ship fuel consumption predictions rely on information about sea and weather conditions obtained from deterministic weather forecasts in phase two of these two-phase methods. However, due to chaos in the atmosphere and ocean and the imperfect depiction of their initial state, the forecast accuracy of deterministic weather forecasts decreases as the forecast period increases (Yousuf et al., 2019). Thus, the speed optimization results become less reliable due to the decreasing accuracy of the FCPM.

To bridge these gaps, we first implement two different data fusion methods to combine noon reports and meteorological information obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF), generating two different datasets for training and testing a FCPM. By comparing the prediction performance on the test data, we use the better fusion method to prepare the essential data for the downstream speed optimization model.

Second, we compare the impacts of deterministic and ensemble weather forecasts on ship speed optimization to see which weather forecast data are better for voyage planning. To make such a comparison, a FCPM, which is assumed to calculate the actual fuel consumption, needs to be established. In this work, the model trained by the dataset based on the better fusion method is used as the actual FCPM. However, in reality, when optimizing ship sailing speed at the departure stage, the actual fuel consumption data for different situations are not available. Therefore, an alternative FCPM with prediction errors needs to be built, and the speed optimization model is constructed based on this alternative model. Then, speed optimization is implemented under two different settings. Under the first setting, the alternative FCPM predicts the fuel consumption under various sailing speed profiles, ship operational conditions, and meteorological information derived from deterministic weather forecasts. A speed optimization model based on the prediction results is adopted to minimize the overall fuel consumption of a voyage, generating the optimized speed profile under the deterministic weather forecast setting. Under the second setting, a similar process is implemented, the only difference being that the meteorological data are obtained from ensemble weather forecasts. The optimized speed profile under the ensemble weather forecast setting is also generated. Finally, based on the actual FCPM, the actual fuel consumption over the voyage for these two speed profiles is calculated and compared, revealing that ensemble weather forecasts are better for voyage planning.

3. Data description

The ship noon report records the sailing conditions of a ship and the sea and weather information around the ship; reports are usually submitted by the captain to shipping companies and shore management parties for the purpose of archiving and planning (Yan et al., 2021). The noon report data usually include ship sailing conditions, such as draught, geographic coordinates, loading conditions, and the average sailing speed since the last report. It also records snapshot information about the surrounding weather and sea conditions. For example, the wind speed information is acquired from a single reading of the anemometer by the deck officer when filling out the noon report (Li et al., 2022). The period of the noon report used in this study spans from February 13, 2019 to August 3, 2020, with a total of 323 records. To compare the impacts of deterministic and ensemble weather forecasts on ship speed optimization, we first randomly select one 9-day laden voyage and one 9-day ballast voyage from the whole dataset for downstream comparison. For the remaining 305 records, we randomly select 20 % of them (61 records) to form the test set and 80 % (244 records) to form the training set.

The major disadvantage of the ship noon report is that the weather and sea information is recorded manually by crews once a day according to their readings of onboard instruments or subjective visual inspections. Consequently, the meteorological data in the noon report are usually neither adequate nor accurate. To address this issue, more reliable meteorological data should be collected to complement the noon report (Yan et al., 2021). In this study, ECMWF Reanalysis v5 (ERA5) is used to improve the quality of the sea and weather data. ERA5 is a reanalysis dataset provided by the ECMWF; it is regarded as the most complete picture of past weather and climate (ECMWF, 2018). It provides hourly estimates of numerous meteorological variables from 1979 to the present, with a resolution of 0.25° (longitude) × 0.25° (latitude) × 1 h (time) for the atmosphere and 0.5° (longitude) × 0.5° (latitude) × 1 h (time) for ocean waves. Therefore, ERA5 data are regarded as the ground truth of meteorological data, which can be combined with the noon report data for predicting the fuel consumption.

Table 1
Features in each data source.

Data source	Features	Description	Feature processing
Ship noon report	Hourly fuel consumption rate of the main engine (ME)	The average fuel consumption per hour of the ship's main engine in unit mt/hour.	The hourly fuel consumption rate is obtained by dividing the overall fuel consumption by the total sailing time.
	Average sailing speed	The average sailing speed since the last report in unit knot.	N.A.
	Sailing condition	Ballast voyage, if the ship is not loaded; laden voyage, if the ship is loaded.	"1" is used to present "laden voyage" and "0" is used to present "ballast voyage".
	Draught	The vertical distance between the waterline and the bottom of the hull (keel) in unit meter.	N.A.
ERA5 reanalysis data	10 m wind direction (WD)	This parameter measures the wind direction with respect to the ship's true heading in unit degrees at the height of 10 m.	The absolute wind direction is transformed to a relative direction to the ship's heading degree from AIS data.
	10 m wind speed (WS)	Wind speed at the height of 10 m in unit m/s.	N.A.
	Mean period of wind waves (MPWW)	This parameter is the average time in seconds which is required for two successive wind-sea wave crests to pass a given point on the surface of the sea.	N.A.
	Mean direction of wind waves (MDWW)	This parameter is the mean true direction of wind-sea waves with different heights, lengths, and directions in unit degree.	The absolute direction is transformed to a relative direction to the ship's heading degree from AIS data.
	Significant height of wind waves (SHWW)	This parameter represents the perpendicular distance from the wind-sea wave trough to the wind-sea wave crest in unit meter.	N.A.
	Mean period of total swell (MPTS)	This parameter is the average time in seconds which is required for two successive total swell crests to pass a given point on the surface of the sea.	N.A.
	Significant height of total swell (SHTS)	This parameter represents the perpendicular distance from the total swell trough to the total swell wave crest in unit meter.	N.A.
	Mean direction of total swell (MDTS)	This parameter is the mean true direction of swells with different heights, lengths, and directions in unit degree.	N.A.
	Significant height of combined wind waves and swell (SHCWWS)	This parameter measures the perpendicular distance in meters from the wave trough and the wave crest. Both wind-sea waves and swells are considered.	N.A.
	Mean wave direction (MWP)	This parameter is the average time in seconds which is required for two successive wave crests to pass a given point on the surface of the sea. Both wind-sea waves and swells are considered.	The absolute direction is transformed to a relative direction to the ship's heading degree from AIS data.
	Forecast surface roughness (FSR)	This parameter measures the surface resistance depending on the waves over the ocean. It is the aerodynamic roughness length in unit meter.	N.A.
	2 m temperature in degrees Celsius (2mT)	This parameter represents the highest temperature of the air at two meters above the surface of land, sea, or inland water in unit degrees Celsius (°C).	Temperatures measured in kelvin are converted to °C by subtracting 273.15.
GEFS ensemble forecast	10 m wind direction	This parameter measures the wind direction with respect to the ship's true heading in unit degrees at the height of 10 m.	The absolute direction is converted to a relative direction to the ship's heading degree from AIS data.
	10 m wind speed	This parameter measures the wind speed at the height of 10 m in unit m/s.	N.A.
	2 m temperature	This parameter represents the highest temperature of the air at two meters above the surface of land, sea, or inland water in unit °C.	Temperatures measured in kelvin are converted to °C by subtracting 273.15.

To improve the ship operational efficiency, weather information is used for voyage planning before the voyage starts (Zis et al., 2020). Due to the inaccessibility of the actual weather data, which are referred to ERA5 data in this study, weather forecast data are used for planning the ship voyage. Then, the impacts of deterministic and ensemble weather forecasts on ship speed optimization are compared to determine which weather forecasts are better for voyage planning. In this study, all weather forecasts are generated by the Global Ensemble Forecast System (GEFS), which is a comprehensive weather forecast model published by the National Oceanic Atmospheric Administration (NOAA) (NOAA, 2017). The weather forecasts generated by GEFS are given as a set of 21 forecast members, where one forecast obtained from the unperturbed initial weather value is referred to as a control member and is exempted from perturbances; perturbances to the initial weather values are added to the other 20 forecasts, which are thus called perturbed members. In our study, similar to Hinnenthal and Clauss (2010) and Skoglund et al. (2015), the control member is regarded as a deterministic weather forecast, and the 20 perturbed weather forecasts are used as the ensemble weather forecasts. The GEFS ensemble forecast data provide temperature, the u-component and v-component of wind, total precipitation, and other variables at different levels with a resolution at 0.5° (longitude) × 0.5° (latitude) × 6 h (time), and a forecast range of 16 days with four runs per day.

The last dataset is obtained from the automatic identification system (AIS). The AIS is widely adopted in the shipping industry to exchange sailing information between AIS-equipped terminals, such as between equipped ships and base stations. The AIS typically

Table 2
Original ship sailing information for voyage i.

Day	Sailing hours (h)	Total fuel (tons)	Speed (knots)	Hourly fuel consumption (tons/h)	Draught (m)	Actual meteorological data (from ERA5)										
						Relative wind direction (degree)	Wind speed (m/s)	Mean period of wind waves (s)	Relative mean direction of wind waves (degree)	Significant height of wind waves (m)	Mean period of total swell (s)	Significant height of total swell (m)	Relative mean directionof total swell (degree)	Significant height of combined wind waves and swell (m)	Mean wave direction (s)	Forecast surface roughness (10^{-2} m)	2 m temperature (°C)
1	24	32.2992	12.63	1.3458	8.6	57.0676	8.1503	4.9802	128.4499	1.2963	7.1365	1.0833	150.3852	1.8079	6.1582	0.0181	28.0580
2	24	32.5008	12.58	1.3542	8.6	4.2611	3.4006	2.8358	65.9841	0.1886	5.5691	0.5065	60.6949	0.5677	4.9930	0.0045	28.3143
3	24	32.5008	12.92	1.3542	8.6	8.7405	1.4705	2.5480	124.8815	0.0679	3.7525	0.3291	90.0941	0.3420	3.6393	0.0032	27.6437
4	24	31.9992	13.29	1.3333	8.6	126.4456	0.4662	3.4569	108.0065	0.0145	6.3865	0.2916	77.4096	0.2942	6.3555	0.0048	28.0852
5	24	32.2008	13.83	1.3417	8.6	157.3829	0.6602	3.1445	78.7957	0.0539	4.6846	0.3968	5.9885	0.4067	4.5522	0.0036	27.5683
6	24	32.2992	13.50	1.3458	8.6	101.5281	3.3529	2.5259	98.8893	0.1518	6.9266	0.6298	64.7416	0.6615	6.6298	8.0360	26.5783
7	24	32.5008	13.29	1.3542	8.6	174.5065	3.5061	2.5661	34.8911	0.1387	11.8568	1.6078	48.6969	1.6163	11.7638	0.0025	26.3400
8	24	32.8008	11.67	1.3667	8.6	157.8617	5.5250	2.9507	21.5785	0.4718	10.9085	1.9007	57.3629	1.9633	10.4231	0.0060	25.6933
9	24	21.9096	10.17	0.9129	8.6	133.6145	4.0771	2.1372	45.6302	0.2151	9.2389	1.5157	63.7959	1.5309	9.0986	0.0024	25.1473

Table 3
Original ship sailing information for voyage ii.

Day	Sailing hours (h)	Total fuel (tons)	Speed (knots)	Hourly fuel consum-ption (tons/h)	Drau-ght (m)	Actual meteorological data (from ERA5)											
						Relative wind direction (degree)	Wind speed (m/s)	Mean period of wind waves (s)	Relative mean direction of wind waves (degree)	Signif-icant height of wind waves (m)	Mean period of total swell (s)	Signif-icant height of total swell (m)	Relative mean directionof total swell (degree)	Significant height of combined wind waves and swell (m)	Mean wave direction (s)	Forecast surface roughness (10^{-2} m)	2 m temper-ature (°C)
1	24	21.3000	9.17	0.8875	9.2	140.7620	5.9578	3.1118	27.3673	0.4983	6.6657	1.2247	78.0075	1.3434	6.1678	0.0067	28.1749
2	24	20.7000	9.96	0.8625	8.6	114.6310	4.0419	2.4882	46.9734	0.1895	4.4306	0.4333	76.0421	0.4941	4.0285	0.0033	27.8245
3	24	21.1704	10.96	0.8821	8.6	147.9591	1.9898	3.2340	81.6656	0.1023	4.4670	0.4586	19.7456	0.4899	4.3052	0.0036	28.1829
4	24	26.3496	12.00	1.0979	8.6	164.6344	3.8953	2.0693	6.2955	0.1642	4.5724	0.4431	9.4587	0.4790	4.2506	0.0028	27.3150
5	24	40.8000	11.20	1.7000	8.6	48.3739	1.5031	2.9114	145.5869	0.0487	4.3055	0.3326	132.5486	0.3409	4.2229	0.0029	27.8522
6	24	32.5800	13.35	1.3575	8.6	126.7840	3.8328	3.0632	78.5747	0.3314	5.4641	0.5352	43.1780	0.6656	4.7684	0.0053	27.7590
7	24	31.3008	12.83	1.3042	8.6	97.6062	4.3871	2.9039	93.2445	0.3256	9.1373	1.0104	26.2728	1.0841	8.5744	0.115821	25.4205
8	24	32.5704	12.71	1.3571	8.6	98.6909	7.7959	4.1743	80.0752	1.0302	10.5849	1.8770	10.8017	2.1676	9.2449	0.000116	26.6168
9	24	32.4000	12.00	1.3500	8.6	84.9367	5.6870	2.9649	94.7588	0.5626	12.0735	1.8855	52.5861	2.0130	11.3822	0.000065	25.6271

Note: ERA5 meteorological data are retrieved using the data fusion method introduced in Section 4.2. Relative wind direction, relative mean direction of wind waves, and relative mean direction of total swell present the wind direction, mean direction of wind wave, and mean direction of total swells relative to the ship's heading, respectively.

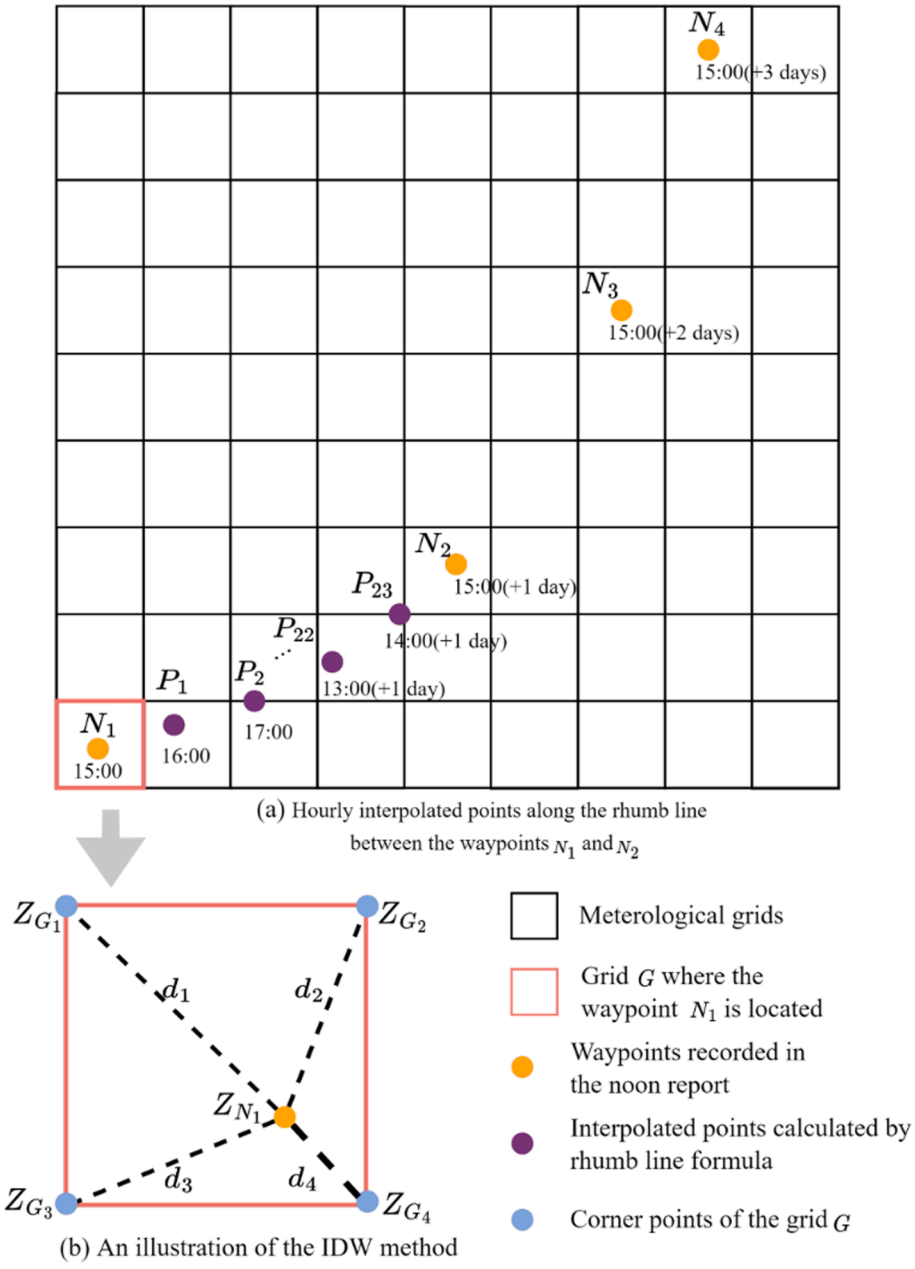


Fig. 1. Graphic explanation of the fusion method based on the rhumb line.

provides static, dynamic, and voyage-related information (Yang et al., 2019a). Because the direction of wind/waves in the meteorological data is absolute, we use the ship's heading degrees provided by the AIS data to calculate the direction of wind/waves with respect to the ship sailing direction.

The source and description of features used in predicting ship fuel consumption are summarized in Table 1. The column "feature processing" describes the conversion methods used to process each feature.

4. Data fusion methods

A voyage in the noon report contains several waypoints that represent the locations traveled through by the ship, and these waypoints can be used to represent a voyage. For example, Fig. 1(a) shows a voyage containing four waypoints (shown by yellow dots), and the voyage can be represented by (N_1, N_2, N_3, N_4). As shown in Fig. 1(a), the meteorological data are presented in regular latitude–longitude grids, which are called meteorological grids in this study, and only the corner points of a meteorological grid have

available meteorological values.

4.1. Fusion method of raw noon report and meteorological data

External meteorological data, which are highly accurate, are incorporated with the original noon report data in this section. Specifically, the first step is to locate the meteorological grid in the external meteorological data in which the waypoint falls, using the geographical coordinates recorded in the noon report. Due to the limited spatial resolution of the meteorological data, there might be no exact meteorological value for the location of the waypoint according to the noon report. To address this issue, the second step is to use the inverse distance weighted (IDW) interpolation method to calculate the meteorological values of the specific voyage point recorded in the noon report (Chen & Liu, 2012). The IDW interpolation method is a deterministic spatial interpolation approach to estimating an unknown value by calculating the weighted average of values of the neighborhood measured points. The weight of a measured point is the inverse distance of the measured point to the unmeasured point. As shown in Fig. 1(a), the sea/weather conditions of waypoint N_1 are unknown. As waypoint N_1 is located in the meteorological grid G and the meteorological values of the corner points of G are known, the meteorological values of N_1 can be calculated based on the corner points. Taking the calculation of wind speed at N_1 as an example, the formulas of the IDW method are as follows:

$$Z_{N_1} = \sum_{i=1}^{4} w_i \times Z_{G_i} \tag{1}$$

$$w_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{4} \frac{1}{d_i}}, \tag{2}$$

where Z_{N_1} is the unknown wind speed at waypoint N_1; w_i is the weight; Z_{G_i} is the wind speed of measured point G_i, i.e., the corner point of the meteorological grid; and d_i is the distance of N_1 to the measured point $G_i (i = 1, 2, 3, 4)$.

4.2. Fusion method of noon report and meteorological data based on the rhumb line

In the fusion method presented in Section 4.1, the weather data acquired at a time point cannot accurately represent the weather condition during the whole day. To provide more accurate weather information, it is necessary to interpolate more points traveled on ship's course between two consecutive waypoints of the voyage according to the temporal resolution of the meteorological data. As shown in Fig. 1(a), 23 points $(P_1, P_2, \ldots, P_{23})$ between waypoints N_1 and N_2 are interpolated points, and their geographic positions can be estimated by the rhumb line route. Then, the average sea/weather data along the interpolated points can be used as a substitute for the original sea/weather conditions in the noon report. Specifically, the first key step is to calculate the number of interpolated points according to the temporal resolution of the external meteorological data source and estimate their geographic positions by interpolation. For example, if the noon report is combined with the ERA5 data at a temporal resolution of 1 h, we can interpolate 23 points (as the time difference between two consecutive noon reports is 24 h); if the noon report is combined with the GEFS ensemble forecast data at a temporal resolution of 6 h, we can interpolate 3 points. In this study, the whole voyage is divided into a number of segments, on which the ship sails following a rhumb line route (Weintrit and Kopacz, 2011). For example, the voyage in Fig. 1(a) is segmented into three parts, and the ship moves along the rhumb line on the $N_1 \rightarrow N_2$, $N_2 \rightarrow N_3$, and $N_3 \rightarrow N_4$ segments. The start points of the three segments are N_1, N_2, and N_3, respectively. The latitude and longitude of each interpolated point can be calculated according to the formulas shown below (Williams, 2013):

$$\delta = \frac{d}{R} \tag{3}$$

$$\varphi_2 = \varphi_1 + \delta \times \cos\theta \tag{4}$$

$$\Delta\psi = \ln\left(\frac{\tan\left(\frac{\pi}{4} + \frac{\varphi_2}{2}\right)}{\tan\left(\frac{\pi}{4} + \frac{\varphi_1}{2}\right)}\right) \tag{5}$$

$$q = \begin{cases} \cos\varphi, & \text{the ship sails on constant latitude course } \varphi \\ \frac{|\varphi_2 - \varphi_1|}{\Delta\psi}, & \text{otherwise} \end{cases} \tag{6}$$

$$\Delta\lambda = \delta \times \frac{\sin\theta}{q} \tag{7}$$

$$\lambda_2 = \lambda_1 + \Delta\lambda, \tag{8}$$

where δ is the angular distance; d represents the distance between the start point of a segment and the interpolated point on the segment, and is calculated by multiplying the average sailing speed (knots) by sailing time (h); R is the earth's radius; λ_1 and φ_1 are the

longitude and latitude of the start point, respectively; λ_2 and φ_2 are the longitude and latitude of the interpolated point, respectively; ln is the natural log; θ is the ship's heading degree in the noon report; and $\Delta\lambda$ is the longitude difference. For example, when the ship sails on segment $N_1 \rightarrow N_2$, the coordinates of the hourly interpolated points $(P_1, P_2, \ldots, P_{23})$ between waypoints N_1 and N_2 can be obtained using equations (3)–(8), as shown in Fig. 1(a). Second, the sea and weather data at each interpolated point can be retrieved by the IDW method. Finally, we average the meteorological values of the interpolated points between the current and previous waypoint, and that of the current waypoint itself. This average is used to replace the original meteorological values of the current waypoint. For example, the wind speeds at the interpolated points are $(Z_{P_1}, Z_{P_2}, \ldots, Z_{P_{23}})$, and the wind speed at waypoint N_2 is Z_{N_2}. Then the recorded wind speed of N_2 in the noon report is replaced by the average wind speed $(Z_{P_1} + Z_{P_2} + \ldots + Z_{P_{23}} + Z_{N_2})/24$. Note that the meteorological values at the start point of the first segment in a voyage, i.e., waypoint N_1 of the voyage in Fig. 1(a), are calculated using the fusion method introduced in Section 4.1.

5. Development of tree-based models for ship fuel consumption prediction

5.1. Introduction of gradient boosted regression tree (GBRT) model

GBRT is an ensemble machine learning algorithm used to deal with regression tasks (Friedman, 2001). It aims to achieve a stronger learning effect by combining the predictions of individual weak learners, which are components of a GBRT and are usually decision trees (DTs). The basic idea of GBRT is to continually add new DTs to fit the residual of the previous round of prediction. This fitting process is repeated numerous times until a preset stopping criterion is reached, such as a specific number of DTs.

When the training process is completed, the final prediction results are generated by summing the prediction results of individual DTs. Let $D = \{(\mathbf{x}_i, y_i), i = 1, 2, \ldots, n\}$ represent the sample data, where $\mathbf{x}_i \in R^m$, and $y_i \in R$. In this study, $\mathbf{x}_i$ is the vector of the input variables, $\mathbf{x}_i = (v_i, \mathbf{c}_i, \mathbf{u}_i)$, where v_i is the sailing speed, $\mathbf{c}_i$ is the sea and weather information, and $\mathbf{u}_i$ is the vessel's loading condition and draught. y_i is the output variable, i.e., the hourly fuel consumption rate. The procedure of constructing a GBRT model is shown in Algorithm 1.

Algorithm 1 Construction of a GBRT model

Input: Training set D; the number of DTs T.
Output: A GBRT model $f^{GBRT}(\mathbf{x})$.

1: Initialize a DT tree which is denoted as $f_0(\mathbf{x}) = \underset{\rho}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, \rho)$, where $L(y_i, \rho) = \frac{1}{2}(y_i - \rho)^2$ is half of the squared loss. The value of ρ is estimated by minimizing the loss function and its optimal value is $\frac{\sum_{i=1}^{n} y_i}{n}$, i.e., $f_0(\mathbf{x}) = \frac{\sum_{i=1}^{n} y_i}{n}$.
2: **for** $t = 1, \ldots, T$ **do:**
3: Calculate the residual of each sample as predicted by the current prediction model by

$$r_{it} = -\left[\frac{\partial L(y_i, f(\mathbf{x}_i))}{\partial f(\mathbf{x}_i)}\right]_{f(x)=f_{t-1}(x)}, i = 1, \cdots, n. \quad (9)$$

4: Construct a new training set $D_m = \{(\mathbf{x}_i, r_{it}), i = 1, 2, \ldots, n\}$ to fit the t th DT whose leaf nodes are $R_{jt} (j = 1, 2, \cdots, J)$, where J represents the number of leaf nodes of the current DT. For each leaf node R_{jt}, the best fitting value is

$$\gamma_{jt} = \underset{\gamma}{\operatorname{argmin}} \sum_{\mathbf{x}_i \in R_{jt}} L(y_i, f_{t-1}(\mathbf{x}_i) + \gamma). \quad (10)$$

5: Update the model by $f_t(\mathbf{x}) = f_{t-1}(\mathbf{x}) + \sum_{j=1}^{J} \gamma_{jt} I(\mathbf{x} \in R_{jt})$, where $I(\mathbf{x} \in R_{jt})$ takes the value of 1 if the value of input $\mathbf{x}$ belongs to leaf node R_{jt}, and 0 otherwise.
6: **end for**
7: The final model $f^{GBRT}(\mathbf{x})$ can be presented by

$$f^{GBRT}(\mathbf{x}) = f_0(\mathbf{x}) + \sum_{t=1}^{T} \sum_{j=1}^{J} \gamma_{jt} I(\mathbf{x} \in R_{jt}). \quad (11)$$

8: **return** $f^{GBRT}(\mathbf{x})$.

5.2. Introduction of eXtreme gradient boosting (XGBoost) model

The XGBoost model, which is proposed by Chen and Guestrin (2016), is a variant of GBRT. Therefore, XGBoost also involves iteratively adding new models to the ensemble and adjusting their weights based on the prediction errors of the previous models. In addition, XGBoost uses a regularized objective function that combines the loss function with a penalty term to prevent overfitting (Fan et al., 2018). Given sample data $D = \{(\mathbf{x}_i, y_i), i = 1, 2, \ldots, n\}$, where $\mathbf{x}_i \in R^m$, and $y_i \in R$, the procedure of constructing an XGBoost model is shown in Algorithm 2.

Algorithm 2 Construction of an XGBoost model

Input: Training set $D = \{(\mathbf{x}_i, y_i), i = 1, 2, \ldots, n\}$; the number of weak learners T; the complexity cost γ; the regularization hyperparameter λ.
Output: An XGBoost model $f^{GBX}(\mathbf{x})$.

1: Initial a weak learner by $f_0(\mathbf{x}) = \underset{w}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, w)$, where $L(y_i, w) = \frac{1}{2}(y_i - w)^2$ is half of the squared loss. The initial prediction value is given by $\hat{y}_i^{(0)} = f_0(\mathbf{x}_i)$, $i = 1, 2, \ldots, n$.
2: **for** $t = 1, \ldots, T$ **do:**

(*continued on next page*)

(*continued*)

Algorithm 2 Construction of an XGBoost model

3: Calculate the first order and second order gradients of each sample by

$$g_{it} = \partial_{\hat{y}_i^{(t-1)}} L\left(y_i, \hat{y}_i^{(t-1)}\right) \quad (12)$$

$$h_{it} = \partial^2_{\hat{y}_i^{(t-1)}} L\left(y_i, \hat{y}_i^{(t-1)}\right). \quad (13)$$

4: Construct the tree structure ($q^t(\mathbf{x})$) of a new weak learner by using a greedy algorithm that recursively partitions the samples based on the values of the input features. $q^t(\mathbf{x})$ can divide samples into the corresponding leaf nodes. For each leaf node in the current tree, the greedy algorithm considers all possible splits of the samples and selects the split that maximizes the value of *Gain* according to the following the following equation

$$Gain = \frac{1}{2}\left[\frac{(G_{Lt})^2}{H_{Lt}+\lambda} + \frac{(G_{Rt})^2}{H_{Rt}+\lambda} - \frac{(G_{Lt}+G_{Rt})^2}{H_{Lt}+H_{Rt}+\lambda}\right] - \gamma, \quad (14)$$

where G_{Lt} and G_{Rt} (H_{Lt} and H_{Rt}) are the sum of the first (the second) order gradients of the samples in the left and right child leaf. (Note: different splits would result in different values for G_{Lt}, G_{Rt}, H_{Lt} and H_{Rt}).

5: Calculate the leaf weights of the weak learner. The weight of leaf node j is given by

$$w_{jt}^* = -\frac{\sum_{i\in I_j} g_{it}}{\sum_{i\in I_j} h_{it} + \lambda} \quad (15)$$

where I_j is the samples in the leaf node j. In addition, $\sum_{i\in I_j} g_{it}$ and $\sum_{i\in I_j} h_{it}$ are the sum of the first and the second order gradients of the samples contained in leaf node j, respectively. A new tree $f_t(\mathbf{x})$ with tree structure $q^t(\mathbf{x})$ and leaf weights w_{jt}^* is obtained.

6: Update the prediction value by $\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i), i = 1, 2, \ldots, n$.

7: **end for**

8: The final model can be represented by

$$f^{XGB}(\mathbf{x}) = \sum_{t=0}^{T} f_t(\mathbf{x}). \quad (16)$$

9: return $f^{GBX}(\mathbf{x})$.

This study employs XGBoost to predict the fuel consumption rate and the predictions are regarded as the ground truth. The XGBoost model is then used to compare the impacts of deterministic and ensemble weather forecasts on ship speed optimization. In reality, when optimizing sailing speeds for a ship at the departure stage, the actual ship fuel consumption in different situations is not available, and thus an alternative FCPM featured with prediction errors needs to be developed to provide the fuel consumption prediction results for the following speed optimization. The GBRT model is used as the alternative FCPM in this study.

6. Development and comparison of ship sailing speed optimization models

6.1. Development of ship sailing speed optimization models

This study adopts the speed optimization model developed by Yan et al. (2020), which is denoted as [M1]. The model is constructed based on the following background. A ship plans to sail along a fixed route from the port O to the port D. The ship's loading condition is predetermined and fixed, and the meteorological conditions during the voyage can be acquired in advance from weather forecasts. Because the sea conditions are highly dynamic, the whole route is divided into a number of segments according to the corresponding records in the noon report. We assume that the meteorological conditions and the international conventions the ship must comply with can be viewed as identical within each segment. The minimum and maximum allowable ship sailing speeds are also assumed to be the same in each segment. The departure time of the ship is set to 0, and the estimated time of arrival at port D should not exceed the latest allowable time of arrival. The optimization model can be solved to find optimal sailing speed of each segment, which can minimize the overall ship fuel consumption. In the [M1] model, n represents the total number of path segments; $i \in \{1, 2, \ldots, n\}$ is the index of a path segment. The main decision variable is the sailing speed v_i in segment i, and the auxiliary decision variable is the time of arrival t_i to the beginning of segment i. In addition, t_{n+1} presents the time of arrival at the end of segment n, i.e., port D. The model is formulated as follows.

[M1]

$$\min \sum_{i=1}^{n}\left(f(v_i, \mathbf{c}_i, \mathbf{u}_i) \times \frac{L_i}{v_i}\right) \quad (17)$$

subject to:

$$t_{i+1} = t_i + \frac{L_i}{v_i}, \forall i \in \{1, 2, 3, \ldots, n\} \quad (18)$$

$$t_1 = 0 \quad (19)$$

$$t_{n+1} \leqslant T_{\max} \quad (20)$$

$$v_{i,c_i,u_i}^{\min} \leqslant v_i \leqslant v_{i,c_i,u_i}^{\max}, \forall i \in \{1, 2, 3, \ldots, n\} \quad (21)$$

$$t_i \geqslant 0, \forall i \in \{1, 2, 3, \ldots, n\} \cup \{n+1\}. \quad (22)$$

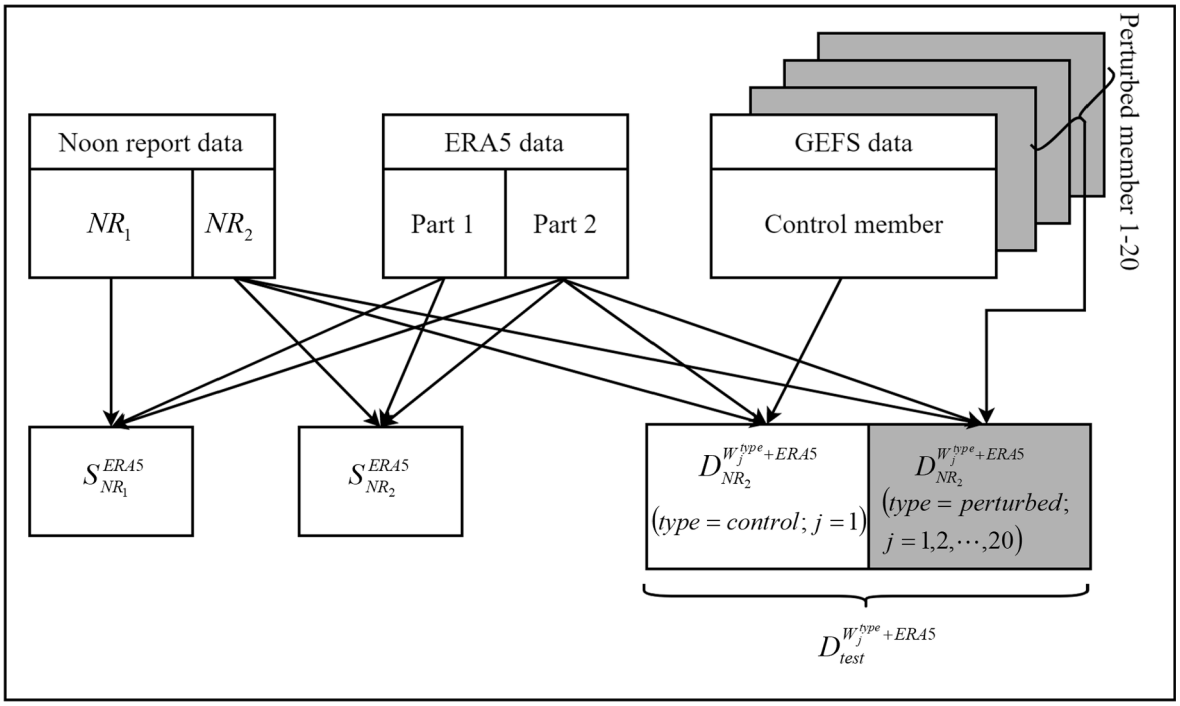


Fig. 2. The illustration of data fusion. Note: NR_1 contains 305 noon report records and NR_2 contains the noon reports of a 9-day laden voyage and a 9-day ballast voyage, i.e., a total of 18 noon report records as mentioned in Section 3; part 1 of ERA5 data includes MPWW, MDWW, SHWW, MPTS, SHTS, SHCWWS, MDTS, MWP, FSR; Part 2 of ERA5 data includes WS, WD and 2mT.

Objective function (17) aims to minimize the overall fuel consumption of a voyage, where $f(v_i, \mathbf{c}_i, \mathbf{u}_i)$ is the hourly fuel consumption of segment i when the sailing speed is v_i, the sea/weather information is $\mathbf{c}_i$, the loading condition and draught are $\mathbf{u}_i$, and L_i is the path length of segment i. The hourly fuel consumption is predicted by the FCPM implemented by the machine learning algorithm introduced in Section 5. Constraints (18) specify the relationship between the time of arrival at the beginning of segment i and at the beginning of segment $i+1$. In this set of constraints, t_i represents the time of arrival at the beginning of segment i, $i \in \{1, 2, 3, \ldots, n\}$. Therefore, as shown in constraints (19), $t_1 = 0$ represents the time of arrival at the beginning of segment 1. t_{n+1} represents the time of arrival at the end of segment n. Constraints (20) guarantee that the ship's time of arrival at port D does not exceed the allowable arrival time, where $T_{\max}$ represents the latest allowable time of arrival at the end of segment n according to the port's requirement or the contract between the shipper and carrier. Constraints (21) bound the sailing speed in the interval of $\left[v^{\min}_{i,\mathbf{c}_i,\mathbf{u}_i}, v^{\max}_{i,\mathbf{c}_i,\mathbf{u}_i}\right]$, where $v^{\min}_{i,\mathbf{c}_i,\mathbf{u}_i}$ and $v^{\max}_{i,\mathbf{c}_i,\mathbf{u}_i}$ are the minimum and maximum allowable sailing speeds in segment i with the draught and loading conditions $\mathbf{u}_i$ and sea/weather conditions $\mathbf{c}_i$. Constraints (22) ensure that the time of arrival at the beginning of every segment is nonnegative.

The model is then linearized so that the optimization software such as CPLEX can solve it. In each segment, the speed values are discretized at intervals of 0.1 knot considering the maximum and minimum allowable sailing speeds $v^{\max}_{i,\mathbf{c}_i,\mathbf{u}_i}$ and $v^{\min}_{i,\mathbf{c}_i,\mathbf{u}_i}$. Let m_i denote the number of candidate speed values in segment i. m_i is calculated by $\lfloor \frac{v^{\max}_{i,\mathbf{c}_i,\mathbf{u}_i} - v^{\min}_{i,\mathbf{c}_i,\mathbf{u}_i}}{0.1} \rfloor$. Therefore, the sailing speed after discretizing can be represented by $v_i^1 = v^{\min}_{i,\mathbf{c}_i,\mathbf{u}_i}$, $v_i^2 = v_i^1 + 0.1, \ldots$, $v_i^{m_i} = v_i^{m_i - 1} + 0.1$, and $v_i^{m_i} = v^{\max}_{i,\mathbf{c}_i,\mathbf{u}_i}$, and a specific sailing speed on segment i is $v_i^m = \{v_i^1, v_i^2, \ldots, v_i^{m_i}\}$. Furthermore, binary decision variable $y_i^m \in \{0, 1\}$ is introduced, which takes a value of 1 if $v_i = v_i^m$, and 0 otherwise. After linearization, the new main decision variable is $y_i^m (i \in \{1, 2, \ldots, n\}; m = 1, 2, \ldots, m_i)$, and the auxiliary decision variable is t_i, $i \in \{1, 2, \ldots, n\}$. According to these new decision variables, model [M1] can be converted to model [M2], which is a mixed-integer linear programming model, and model [M2] and model [M1] equivalent.

[M2]

$$\min \sum_{i=1}^{n} \sum_{m=1}^{m_i} \left[y_i^m \times f\left(v_i^m, \mathbf{c}_i, \mathbf{u}_i\right) \times \frac{L_i}{v_i^m} \right] \tag{23}$$

subject to:

$$t_{i+1} = t_i + \sum_{m=1}^{m_i} \left(\frac{L_i}{v_i^m} \times y_i^m \right), \forall i \in \{1, 2, 3, \ldots, n\} \tag{24}$$

$$t_1 = 0 \tag{25}$$

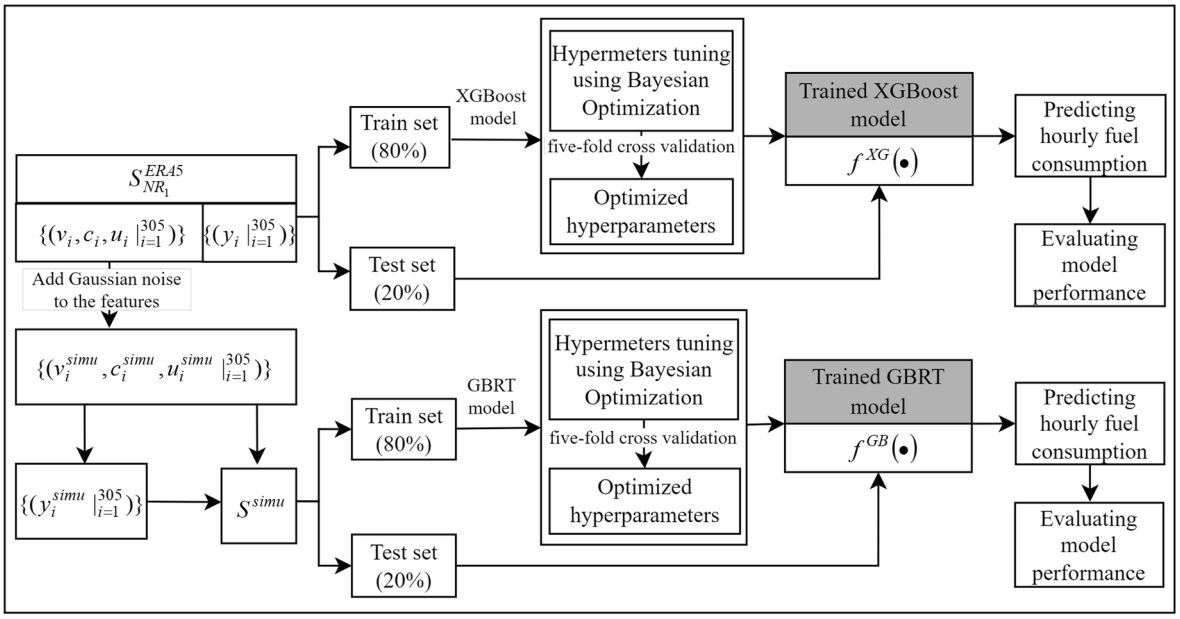


Fig. 3. The flowchart of the development of prediction model.

$$t_{n+1} \leqslant T_{\max} \tag{26}$$

$$y_i^m \in \{0,1\}, \forall i \in \{1,2,3,\ldots,n\}, m \in \{1,2,3,\ldots,m_i\} \tag{27}$$

$$\sum_{m=1}^{m_i} y_i^m = 1, \forall i \in \{1,2,3,\ldots,n\} \tag{28}$$

$$t_i \geqslant 0, \forall i \in \{1,2,3,\ldots,n\} \cup \{n+1\}. \tag{29}$$

6.2. *Comparison of the influences of deterministic and ensemble weather forecasts on ship speed optimization*

The procedure for comparing the impacts of deterministic and ensemble weather forecasts on ship speed optimization is described below.

(i) Data fusion: This step combines the noon report and meteorological data to produce three datasets for fuel consumption prediction and speed optimization. Fig. 2 shows the data fusion process to generate three datasets, where the first dataset $S_{NR_1}^{ERA5} = \left\{ (v_i, \mathbf{c}_i, \mathbf{u}_i, y_i)\,|_{i=1}^{305} \right\}$ is obtained by combining the original noon report data NR_1, which contains 305 noon reports, and ERA5 reanalysis data; the second dataset $S_{NR_2}^{ERA5} = \left\{ (v_i, \mathbf{c}_i, \mathbf{u}_i, y_i)\,|_{i=1}^{18} \right\}$ is obtained by combining the original noon report data NR_2, which contains 18 noon reports (9 from a laden voyage and 9 from a ballast voyage), and ERA5 reanalysis data. NR_1 and NR_2 are mutually exclusive to each other. Dataset $S_{NR_1}^{ERA5}$ is utilized for training and testing a FCPM, which in this study is assumed to be able to predict actual fuel consumption with 100 % accuracy, and dataset $S_{NR_2}^{ERA5}$ is used to compare the influences of deterministic and ensemble forecasts on optimizing ship speeds. In both datasets, v_i represents the average sailing speed, $\mathbf{c}_i$ represents the sea/weather conditions as introduced in Table 1, $\mathbf{u}_i$ represents the loading condition and draught, and y_i represents the hourly fuel consumption of ME. The third dataset is produced by combining the noon report data NR_2, ERA5 reanalysis data, and each member of GEFS ensemble forecast data W_j^{type} (if $type = perturbed$, $j = 1,2,\ldots,20$; if $type = control$, $j = 1$) separately, and is denoted as $D_{NR_2}^{W_j^{type}+ERA5} = \left\{ \left(v_i, \mathbf{c}_i^{W_j^{type}}, \mathbf{c}_i^{ERA5}, \mathbf{u}_i \right)\,|_{i=1}^{18} \right\}$, where $\mathbf{c}_i^{W_j^{type}}$ represents the sea/weather conditions, including WS, WD, and 2mT, and is derived from the j th perturbed or control member, and $\mathbf{c}_i^{ERA5}$ is the sea/weather conditions introduced in Table 1 (excluding WS, WD, and 2mT). Fig. 2 shows that $D_{NR_2}^{W_j^{type}+ERA5}$ equals $D_{NR_2}^{W_1^{control}+ERA5}$ if $type = control$; $D_{NR_2}^{W_j^{type}+ERA5}$ equals $D_{NR_2}^{W_j^{perturbed}+ERA5} (j = 1,2,\ldots,20)$ if $type = perturbed$. Dataset $D_{NR_2}^{W_1^{control}+ERA5}$ presents deterministic weather forecasts, and dataset $D_{NR_2}^{W_j^{perturbed}+ERA5} (j = 1,2,\ldots,20)$ presents ensemble weather forecasts.

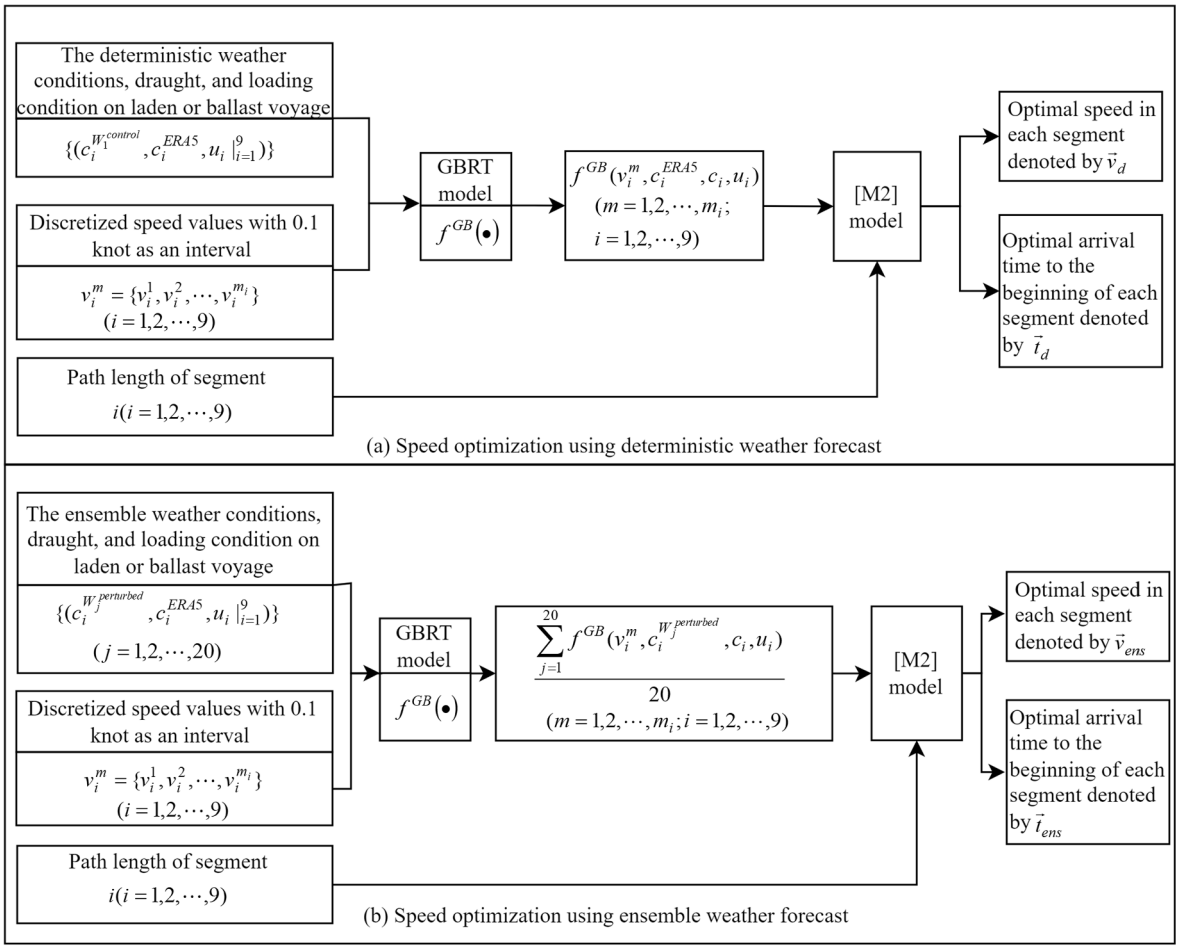


Fig. 4. The flowchart of developing speed optimization model [M2] using different datasets.

(ii) Model implementation: As shown in Fig. 3, this step develops an XGBoost model $f^{XG}(\cdot)$ and a GBRT model $f^{GB}(\cdot)$ to predict ship fuel consumption rates. Specifically, the XGBoost model is employed to compute the actual ship fuel consumption rates by inputting the sailing speeds, actual sea/weather conditions, and operational data. The GBRT model is used to predict the ship fuel consumption rate when the actual sea/weather conditions are unknown. Fig. 3 shows the whole training and testing process of model $f^{XG}(\cdot)$. First, the entire dataset $S_{NR_1}^{ERA5}$ is randomly divided according to an 8:2 ratio; the larger subset is used to tune model hyperparameters and train the model, while the generalization ability of models is evaluated by the smaller subset. Then, the optimal values of the hyperparameters in the XGBoost model are found by a Bayesian optimization method with five-fold cross validation on the whole training set. Finally, the performance of the XGBoost model is evaluated on the test set in terms of the mean squared error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). To train and test the GBRT model, a simulated dataset first needs to be generated. As shown in Fig. 3, the simulated input variables $\left\{\left(v_i^{simu}, \mathbf{c}_i^{simu}, \mathbf{u}_i^{simu}\right)|_{i=1}^{305}\right\}$ are generated by adding Gaussian noise to the features in the original dataset $S_{NR_1}^{ERA5}$, while the simulated output variable $\left\{y_i^{simu}|_{i=1}^{305}\right\}$ is calculated by inputting the simulated input variables to the XGBoost model $f^{XG}(\cdot)$. Then, the simulated dataset S^{simu} can be represented by $\left\{\left(v_i^{simu}, \mathbf{c}_i^{simu}, \mathbf{u}_i^{simu}, y_i^{simu}\right)|_{i=1}^{305}\right\}$, which is utilized for training and testing the GBRT model. The train/test set split ratio is set to 8:2. The training and testing procedures are the same as those for XGBoost model $f^{XG}(\cdot)$.

(iii) Speed optimization: Based on the ship fuel consumption rates predicted by the GBRT model $f^{GB}(\cdot)$, the optimized sailing speed in each segment of a voyage can be obtained using model [M2], which is introduced in Section 6.1. To compare the impacts of deterministic and ensemble weather forecasts on ship speed optimization, these two different weather forecasts are input to the $f^{GB}(\cdot)$ model to obtain the predicted ship fuel consumption rates. Then, the predicted fuel consumption rates are input to model [M2] to obtain the optimal sailing speed in each segment, denoted as $\vec{v}_d \in \left\{v_i|_{i=1}^{9}\right\}$, and the optimal time of arrival at the

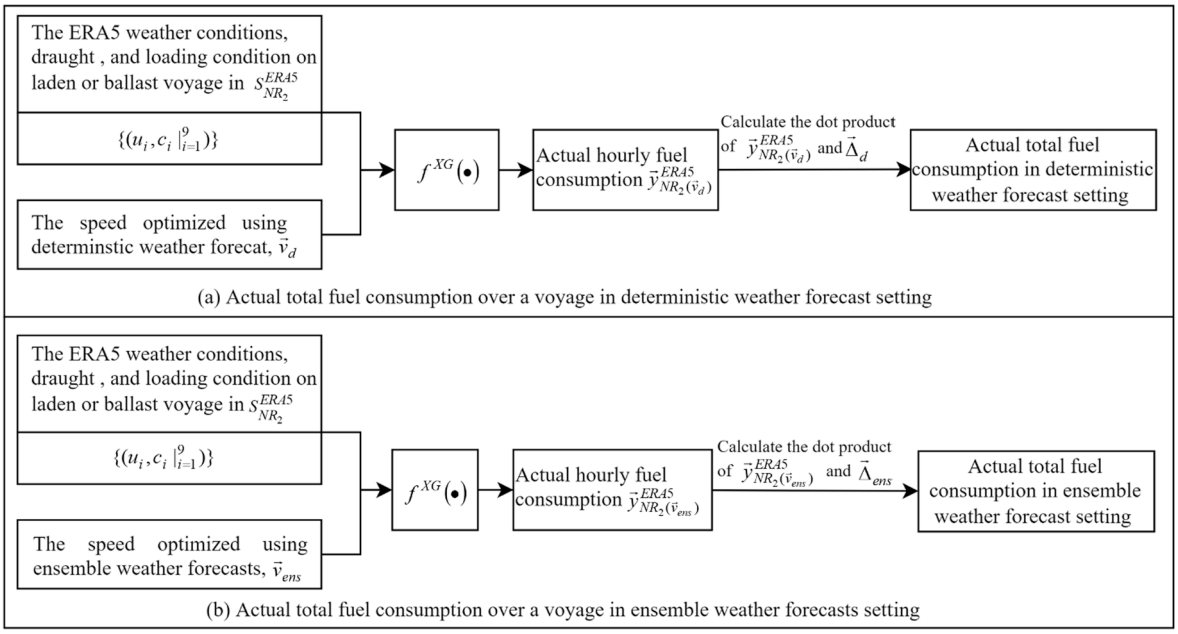


Fig. 5. The flowchart of comparing the actual total ship fuel consumption in deterministic and ensemble weather forecast settings.

beginning of each segment, denoted as $\overrightarrow{t}_d \in \left\{t_i|_{i=1}^{10}\right\}$ ($t_1 = 0$; t_{10} is the time of arrival at the end of the last segment). Specifically, according to the deterministic forecast data approach shown in Fig. 4(a), deterministic sea/weather conditions $\left\{\left(\mathbf{c}_i^{W_1^{control}}, \mathbf{c}_i^{ERA5}\right)|_{i=1}^{9}\right\}$, ship operational data $\left\{(\mathbf{u}_i)|_{i=1}^{9}\right\}$, and discretized speed values $v_i^m = \{v_i^1, v_i^2, ..., v_i^{m_i}\}$ for each segment of the laden or ballast voyage are input to $f^{GB}(\cdot)$, and we can obtain the hourly fuel consumption prediction for each segment. Then, the predicted hourly fuel consumption and sailing distance of each segment are input to the speed optimization model [M2] to find $\overrightarrow{v}_d$ and $\overrightarrow{t}_d$. In addition, the sailing time in each segment, denoted as $\overrightarrow{\Delta t}_d$, can be calculated based on $\overrightarrow{t}_d$. The objective function of [M2] can be defined as below.

$$\min \sum_{i=1}^{n} \sum_{m=1}^{m_i} \left[y_i^m \times f^{GB}\left(v_i^m, \mathbf{c}_i^{ERA5}, \mathbf{c}_i^{W_1^{control}}, \mathbf{u}_i\right) \times \frac{L_i}{v_i^m} \right]. \tag{30}$$

In the ensemble forecast data approach, the sea/weather conditions of each perturbed member $\left\{\left(\mathbf{c}_i^{W_j^{perturbed}}\right)|_{i=1}^{9}\right\} (j = 1, 2, ..., 20)$, sea/weather conditions of ERA5 $\left\{\left(\mathbf{c}_i^{ERA5}\right)|_{i=1}^{9}\right\}$, and ship operational data $\left\{\mathbf{u}_i|_{i=1}^{9}\right\}$ of dataset $D_{NR_2}^{W_j^{type}+ERA5}$ are input to $f^{GB}(\cdot)$, and the hourly fuel consumption $\left(\sum_{j=1}^{20} f^{GB}\left(v_i^m, c_i^{W_j^{perturbed}}, c_i^{ERA5}, u_i\right)\right) \Big/ 20$ on each segment can be calculated by averaging the 20 hourly fuel consumption predictions. Then, the average hourly fuel consumption and the sailing distance on each segment are input to the speed optimization model [M2] to determine the optimal sailing speed in each segment, denoted as $\overrightarrow{v}_{ens}$, and the optimal time of arrival at the beginning of each segment, denoted as $\overrightarrow{t}_{ens} \in \left\{t_i|_{i=1}^{10}\right\}$ ($t_1 = 0$; t_{10} is the time of arrival at the end of the last segment). In addition, the sailing time in each segment, denoted as $\overrightarrow{\Delta t}_{ens}$, can be calculated using $\overrightarrow{t}_{ens}$. The objective function of [M2] is mathematically represented as below.

$$\min \sum_{i=1}^{n} \sum_{m=1}^{m_i} \left[y_i^m \times \left(\frac{\sum_{j=1}^{20} f^{GB}\left(v_i^m, \mathbf{c}_i^{W_j^{perturbed}}, \mathbf{c}_i^{ERA5}, \mathbf{u}_i\right)}{20} \right) \times \frac{L_i}{v_i^m} \right]. \tag{31}$$

(iv) Comparison: The last step compares the actual overall fuel consumption of a voyage when the ship sails at speeds in $\overrightarrow{v}_d$ and $\overrightarrow{v}_{ens}$, respectively. As shown in Fig. 5(a), when the ship sails at speeds in $\overrightarrow{v}_d$, the vector of actual hourly fuel consumption $\overrightarrow{y}_{NR_2}^{ERA5}\left(\overrightarrow{v}_d\right)$ is predicted by inputting speeds $\overrightarrow{v}_d$, the actual sea/weather conditions $\left\{(\mathbf{c}_i)|_{i=1}^{9}\right\}$ of the dataset $S_{NR_2}^{ERA5}$, and the ship operational data $\left\{(\mathbf{u}_i)|_{i=1}^{9}\right\}$ of dataset $S_{NR_2}^{ERA5}$ to the XGBoost model $f^{XG}(\cdot)$. Then, the actual overall fuel consumption over a

Table 4
Hyperparameters in XGBoost models $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$.

Hyperparameter	Interpretation	Searching space	Selected value in ($f_{S1}^{XG}(\cdot)$)	Selected value in ($f_{S2}^{XG}(\cdot)$)
max_depth	Maximum depth of a tree.	[2,17]	15	6
n_estimators	The number of trees.	[10,300]	48	83
learning_rate	Learning rate shrinks the step size to prevent overfitting.	[0.00001,1]	0.45	0.16
min_child_weight	Minimum sum of instance weight (hessian) required in a child.	[0,10]	5.58	9.54
colsample_bytree	The subsample ratio of columns in the construction of each tree.	[0.1,1]	0.85	0.64
subsample	Subsample ratio of the training instances.	[0.4,1]	0.77	0.94
reg_alpha	L1 regularization term on weights.	[0,2]	0.20	0.23
reg_lambda	L2 regularization term on weights.	[0,2]	0.61	1.76
gamma	Minimum loss reduction required for further partitioning on a tree leaf node.	[0,2]	0.29	0.18

Table 5
Performance of the XGBoost model $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$ on the test set.

Model	MSE	MAE	MAPE	R^2
$f_{S1}^{XG}(\cdot)$	0.023	0.114	8.023	0.721
$f_{S2}^{XG}(\cdot)$	0.024	0.107	7.706	0.777

voyage in the deterministic weather forecast setting is obtained by calculating the dot product of the vector of the optimized sailing time $\overrightarrow{\Delta t}_d$ and the vector of actual hourly fuel consumption $\overrightarrow{y}_{NR_2}^{ERA5}\left(\overrightarrow{v}_d\right)$. As shown in Fig. 5(b), the vector of actual hourly fuel consumption $\overrightarrow{y}_{NR_2}^{ERA5}\left(\overrightarrow{v}_{ens}\right)$ and the actual overall fuel consumption in the ensemble weather forecast setting are calculated using the same method as shown in Fig. 5(a) when the ship sails at speeds $\overrightarrow{v}_{ens}$. Finally, we can compare the actual total ship fuel consumption of the voyage under the two different weather forecast settings to determine which weather forecast is better for ship speed optimization.

7. Computational experiments

We introduce two data fusion methods in Section 4, which can improve the accuracy of sea and weather data of the noon report. NR_1 combined with ERA5 meteorological data using fusion methods in Section 4.1 and Section 4.2 are denoted as $S1_{NR_1}^{ERA5}$ and $S2_{NR_1}^{ERA5}$, respectively. $S1_{NR_1}^{ERA5}$ and $S2_{NR_1}^{ERA5}$ are used to train and test FCPMs $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$. $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$ are developed using the same machine learning model, i.e., XGBoost model and parameter turning process, the only difference being the data fusion method used to obtain datasets $S1_{NR_1}^{ERA5}$ and $S2_{NR_1}^{ERA5}$. Therefore, by comparing the predictive performances of $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$ on the test data, we can conclude which fusion method can obtain more accurate meteorological data and then finally choose the corresponding XGBoost model to generate the simulated dataset. The simulated dataset S^{simu} is generated according to step (ii) of Section 6.2, and the standard deviations of the Gaussian noise are set to 5 % of the standard deviations of the features in the original dataset $S1_{NR_1}^{ERA5}$(or $S2_{NR_1}^{ERA5}$). Then, S^{simu} is used to train and test $f^{GB}(\cdot)$. Finally, step (iii) and step (iv) of Section 6.2 are performed to compare the impacts of deterministic and ensemble weather forecasts on ship speed optimization.

7.1. Implementation of fuel consumption prediction model

The XGBoost 1.6.1 library0F[1] is used to develop the XGBoost regressors $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$ to predict the actual ship fuel consumption using actual meteorological data, i.e., the ERA5 reanalysis data in this study. The parameters for the XGBoost regressor are founded by the Bayesian optimization method implemented via the hyperopt 0.2.2 library1F[2] (Hyperopt) with five-fold cross validation. Details of the hyperparameters tuned in XGBoost regressors $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$ are presented in Table 4. Aside from those hyperparameters, the default values adopted by the XGBoost 1.6.1 library are used. Table 5 lists the predictive performance of $f_{S1}^{XG}(\cdot)$ and $f_{S2}^{XG}(\cdot)$ on the test set. Although $f_{S1}^{XG}(\cdot)$ performs slightly better than $f_{S2}^{XG}(\cdot)$ in terms of MSE, $f_{S2}^{XG}(\cdot)$ is better than $f_{S1}^{XG}(\cdot)$ regarding MAE, MAPE, and R^2. We can conclude that the data fusion method proposed in Section 4.2 can obtain more accurate meteorological data as a supplement for the noon report data than the fusion method proposed in Section 4.1. Therefore, the fusion method proposed in 4.2 is used to combine the noon report data NR_2 and the meteorological data from weather forecasts for the downstream speed

[1] https://xgboost.readthedocs.io/.
[2] https://hyperopt.github.io/hyperopt/.

Table 6
Parameters in GBRT regressor $f^{GB}(\cdot)$.

Name	Interpretation	Searching space	Selected value
max_depth	Maximum depth of a single estimator.	[2,10]	3
min_samples_leaf	The minimum number of samples required per leaf node.	[1,20]	7
min_samples_split	The minimum number of samples required for splitting an internal node.	[2,20]	17
max_features	The number of features to be considered when searching for the best split.	[1,15]	15
n_estimators	The number of regression trees.	[10,300]	264
learning_rate	Learning rate diminishes the contribution of each tree.	[0.00001,1]	0.06
subsample	The fraction of samples used to fit each base learner.	[0.4,1]	0.87

Table 7
Performance of GBRT model $f^{GB}(\cdot)$ on the simulated test set.

Model	MSE	MAE	MAPE	R^2
$f^{GB}(\cdot)$	0.003	0.039	2.738	0.932

Table 8
The sailing information of the ship in voyage i after speed optimization using deterministic weather forecasts.

Day	Sailing time(h)	Sailing speed $\vec{v}_d$(knots)	Actual hourly fuel consumption (tons/h)	Actual total fuel consumption (tons)
1	22.7433	13.1	1.3438	30.5624
2	23.0810	13.1	1.3489	31.1340
3	23.3887	13.1	1.3419	31.3853
4	24.4178	13.0	1.3026	31.8067
5	31.2338	10.7	1.3035	40.7133
6	24.3290	13.1	1.3655	33.2213
7	29.0115	10.7	1.1108	32.2260
8	15.6583	18.0	2.1463	33.6074
9	18.6601	13.1	1.3730	25.6203
Total	212.5235	/	/	**290.2765**

optimization experiment.

In addition, the prediction performance of model $f_{S2}^{XG}(\cdot)$ is compared with several popular machine learning algorithms, including random forest (RF), decision tree (DT), least absolute shrinkage and selection operator (LASSO) regression, and artificial neural network (ANN). It is found that $f_{S2}^{XG}(\cdot)$ shows the best performance on the test set among five machine learning algorithms. The detailed performances of five machine learning models are described in Appendix A. To train and test the GBRT model $f^{GB}(\cdot)$, we use $f_{S2}^{XG}(\cdot)$ to generate the simulated dataset S^{simu}. In addition, $f_{S2}^{XG}(\cdot)$ is also used to calculate the actual hourly fuel consumption in step (iv) of Section 6.2.

The GBRT model $f^{GB}(\cdot)$ is used for predicting the fuel consumption in the downstream speed optimization experiment and developed using the scikit-learn python library 1.0.22F[3]. Hyperparameters in $f^{GB}(\cdot)$ are found using the Hyperopt tool with five-fold cross validation and are listed in Table 6. Aside from those hyperparameters, other hyperparameters are configured as the defaults in the scikit-learn library. Table 7 lists the predictive performance of $f^{GB}(\cdot)$ on the simulated test data. It can be seen that $f^{GB}(\cdot)$ has good performance on the simulated test set in all metrics.

7.2. Speed optimization based on deterministic and ensemble weather forecasts

We consider two voyages from the noon report that are not in the training and test sets, namely voyage i and voyage ii. Voyage i is a 9-day ballast voyage with a total steaming time of 216 h and total fuel consumption of 281.0112 tons. Voyage ii is a 9-day laden voyage with a total steaming time of 216 h and total fuel consumption of 259.1712 tons. Detailed information from voyage i and voyage ii, including ship operational data (i.e. sailing condition and draught) and actual meteorological data (as obtained using fusion method developed in Section 4.2), are presented in Table 2 and Table 3.

7.2.1. Optimizing sailing speed using deterministic weather forecasts

In the speed optimization experiment using deterministic weather forecasts, the hourly fuel consumption in model [M2] is predicted by $f^{GB}(\cdot)$ with discretized speed values, deterministic weather forecasts, draught, and sailing condition as the inputs. The optimized sailing speed and sailing time solved by the CPLEX optimizer in each segment of voyage i and voyage ii are presented in

[3] https://scikit-learn.org/stable/.

Table 9
The sailing information of the ship in voyage ii after speed optimization using deterministic weather forecasts.

Day	Sailing time(h)	Sailing speed $\vec{v}_d$(knots)	Actual hourly fuel consumption (tons/h)	Actual total fuel consumption (tons)
1	16.9857	13.0	1.4905	25.3172
2	22.2359	10.7	1.1460	25.4823
3	19.9880	13.1	1.3542	27.0678
4	22.1410	13.0	1.5634	34.6152
5	27.2183	10.7	1.3285	36.1595
6	28.4722	10.7	1.0838	30.8582
7	16.6841	13.1	1.3938	23.2543
8	27.0107	10.7	1.1805	31.8861
9	22.9783	13.1	1.3564	31.1678
Total	203.7142	/		**265.8084**

Table 10
The sailing information of the ship in voyage i after optimizing the speed using ensemble weather forecasts.

Day	Sailing time(h)	Sailing speed $\vec{v}_{ens}$(knots)	Actual hourly fuel consumption (tons/h)	Actual total fuel consumption(tons)
1	22.7433	13.1	1.3438	30.5625
2	23.0810	13.1	1.3489	31.1340
3	23.3887	13.1	1.3419	31.3853
4	24.4178	13.0	1.3026	31.8066
5	18.5668	18.0	2.0745	38.5168
6	24.3290	13.1	1.3655	33.2213
7	29.0115	10.7	1.1110	32.2318
8	26.3411	10.7	1.1083	29.1938
9	18.6601	13.1	1.3730	25.6203
Total	212.5393	/	/	**283.6723**

Table 11
The sailing information of the ship in voyage ii after optimizing the speed using ensemble weather forecasts.

Day	Sailing time(h)	Sailing speed $\vec{v}_{ens}$(knots)	Actual hourly fuel consumption (tons/h)	Actual total fuel consumption (tons)
1	16.9857	13.0	1.4905	25.3172
2	22.2359	10.7	1.1460	25.4832
3	19.9880	13.1	1.3542	27.0677
4	22.1410	13.0	1.5634	34.6152
5	27.2183	10.7	1.3285	36.1595
6	23.4348	13.0	1.3705	32.1174
7	16.6841	13.1	1.3938	23.2543
8	27.0107	10.7	1.1805	31.8861
9	22.9783	13.1	1.3564	31.1678
Total	173.9864	/	/	**267.0676**

Table 8 and Table 9, respectively. The actual hourly fuel consumption and actual total fuel consumption in Table 8 and Table 9 are calculated according to step (iv) in Section 6.2. Specifically, $f_{S2}^{XG}(\cdot)$ obtained in Section 6.1 is used to predict the actual hourly fuel consumption for the two voyages. The input features include the optimized sailing speed and the actual meteorological conditions and ship operation data shown in Table 2 and Table 3. The actual total fuel consumption data in Table 8 and Table 9 are calculated by multiplying the optimized sailing time by the actual hourly fuel consumption.

7.2.2. Optimizing sailing speed using ensemble weather forecasts

In the speed optimization experiment using ensemble weather forecasts, the hourly fuel consumption in model [M2] is predicted by $f^{GB}(\cdot)$ with discretized speed values, ensemble weather forecasts, and operational data as inputs. The optimized sailing speed and sailing time solved by the CPLEX optimizer in each segment of voyage i and voyage ii are presented in Table 10 and Table 11, respectively. The methods for calculating the actual hourly fuel consumption and total fuel consumption in Table 10 and Table 11 are the same as those in Section 7.2.1.

The actual overall fuel consumption after optimizing the sailing speeds and the overall fuel consumption prediction based on the original sailing speeds in the noon report for voyages i and ii are summarized in Table 12; both are obtained via $f_{S2}^{XG}(\cdot)$. As shown in Table 12, the ship consumes 556.0849 tons of fuel to complete voyage i and voyage ii in the deterministic weather forecast setting, and it consumes 550.7399 tons of fuel to complete the same voyages in the ensemble weather forecast setting. It can be concluded that, compared to using the deterministic weather forecasts, using the ensemble weather forecasts to optimize the sailing speed can reduce the overall fuel consumption of the voyage by more than 1 %. If we compare the fuel consumption after sailing speed optimization using ensemble weather forecasts and the actual total fuel consumption before sailing speed optimization, i.e., 596.5593 tons, we can

Table 12
Actual fuel consumption of two 9-day sailing voyage after speed optimization in deterministic and ensemble weather forecast settings.

Voyage	Actual total fuel consumption before speed optimization F_{before} (tons)	Actual total fuel consumption after speed optimization (tons)		Reduction ratio	
		Deterministic weather forecast setting F_d	Ensemble weather forecast setting F_{ens}	$\frac{F_{before}-F_{ens}}{F_{before}}\times 100\%$	$\frac{F_d-F_{ens}}{F_d}\times 100\%$
Voyage i	304.1203	290.2765	283.6723	6.7 %	2.28 %
Voyage ii	292.4390	265.8084	267.0676	8.7 %	−0.47 %
Total	596.5593	556.0849	550.7399	7.7 %	1.00 %

Note: The actual total fuel consumption before speed optimization is predicted by $f_{S2}^{XG}(\cdot)$ when the ship sails at the speeds in the noon report under the actual sea/weather conidtions derived from the ERA5 dataset.

conclude that fuel consumption can be reduced by more than 7.7 %. Compared to the use of deterministic weather forecasts, using ensemble weather forecasts to optimize the sailing speed can reduce the total fuel consumption to complete voyage i by over 2.27 % (a reduction from 290.2765 tons to 283.6723 tons, respectively). However, the total fuel consumption to complete voyage ii increases by over 0.47 % (from 265.8084 tons to 267.0676 tons, respectively) if we use the ensemble weather forecast to optimize the speed. As mentioned above, the actual total fuel consumption at certain speeds is predicted by $f_{S2}^{XG}(\cdot)$. However, as $f_{S2}^{XG}(\cdot)$ is not completely accurate, it will inevitably introduce errors into the prediction results. Therefore, the actual total fuel consumption predicted by $f_{S2}^{XG}(\cdot)$ will fluctuate to some extent.

8. Future work

The framework proposed in this study provides an alternative solution to compare the impacts of different weather forecasts on ship sailing speed optimization when the actual fuel consumption is not accessible. In this framework, we assume that the actual fuel consumptions are predicted by $f^{XG}(\cdot)$. $f^{XG}(\cdot)$ is trained based on historical data which combine the noon report data and the meteorological values obtained from the ERA5 dataset. However, $f^{XG}(\cdot)$ cannot accurately predict the actual fuel consumption with 100 % accuracy. In the future, we could consider building a simulation platform that simulates ship fuel consumption rates with all possible combinations of sailing speed profile and the surrounding sea/weather conditions. This would enable us to compare the impacts of different weather forecasts on sailing speed optimization in a fairer manner.

In addition, the sea and weather conditions used for optimizing sailing speeds are static, as they are obtained through the latest updated weather forecasts before the ship's departure. However, the static sea and weather conditions cannot reflect the continuously changing environment at sea. Therefore, the vessel sailing speed optimization based on the static sea and weather conditions may not generate the optimal speeds throughout the voyage. In the future, we can implement the rolling-horizon approach in the vessel sailing speed optimization decision problem. More specifically, the sailing speeds of the vessel can be continuously optimized for the remaining voyage based on the established speed optimization model once the latest updated weather forecast data are obtained. Therefore, the rolling-horizon method can offset the distortion of speed optimization results caused by the changing sea and weather conditions over time. Similarly, based on the rolling-horizon approach, we can also compare the impacts of deterministic and ensemble weather forecasts on ship sailing speed optimization.

Furthermore, the meteorological dataset used in this study has a limited spatial resolution, which may not provide meteorological values for the exact locations of waypoints and interpolated points. Therefore, the IDW is used to calculate the weighted average of the meteorological values at the four corner points of the grid in which the waypoint or interpolated point is located. This weighted average is then used as the meteorological value for the waypoint or the interpolated point. In the future, we can consider more grid points, such as corner points of eight adjacent grids around the grid where the waypoint is located, as measured points to estimate the meteorological values for the waypoint or interpolated point. Furthermore, we can investigate the impact of the number of measured points on the accuracy of sea and weather information derived for waypoints or interpolated points, such that we can determine the optimal number of measurement points to obtain more accurate sea and weather information, thereby improving the performance of the vessel sailing speed optimization model.

9. Conclusion

In this work, we investigate the impacts of two different data fusion methods on a FCPM of ship and compare the impacts of deterministic and ensemble weather forecasts on ship speed optimization. First, we develop two different data fusion methods to combine noon report data and external meteorological data to generate two different datasets for training and testing an actual FCPM. From the prediction accuracy on the test set, we find the rhumb line based fusion method produces more accurate meteorological data than the fusion method that directly connects meteorological data with raw noon report records. Therefore, the rhumb line based fusion method is employed to prepare the essential data for the two-phase speed optimization model.

Second, based on a two-phase speed optimization model, the impacts of deterministic weather forecasts and ensemble weather

Table 13
Performance of five models on test set.

Model/Metric	MSE	MAE	MAPE	R^2
XGBoost	**0.024**	**0.107**	**7.706**	**0.777**
RF	0.040	0.114	9.508	0.702
DT	0.045	0.111	9.213	0.658
LASSO	0.049	0.163	11.156	0.479
ANN	0.048	0.153	10.840	0.641

forecasts on speed optimization are investigated. Specifically, in phase one, a simulated dataset is generated by adding Gaussian noise to the features of the original fusion dataset, which is adopted for training and testing a FCPM based on the GBRT algorithm. In phase two, a speed optimization model established on the predicted fuel consumption is adopted to minimize the total fuel consumption over a voyage, generating optimized speed profiles under the deterministic and ensemble weather forecast settings, respectively. In numerical experiments on two 9-day sailing voyages, the total fuel consumption with the speed profile based on ensemble weather forecasts is 1 % lower than that based on deterministic weather forecasts. Based on global marine fuel consumption in 2019, speed optimization based on ensemble weather forecasts could save 2 million tons of fuel, demonstrating the potential of ensemble weather forecasts in speed optimization.

This study contributes to the field in two ways. First, the comparison of the two data fusion methods provides new insights into ways to remedy the quality and quantity issues in the sea and weather conditions of ship noon reports. Second, this study proposes a framework to compare the impacts of ensemble versus deterministic weather forecasts on ship sailing speed optimization when the actual ship fuel consumptions are unavailable, making it possible to evaluate the fuel consumption of multiple speed plans based on different weather forecast data before a voyage. Based on the evaluation results, we suggest that shipping companies should plan the sailing speed of their ships based on ensemble weather forecasts, which will reduce ship fuel consumption and GHG emissions to a larger extent. Therefore, the study will help both shipping firms and policy makers improve the environmental sustainability of maritime transport.

Declaration of Competing Interest

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

Acknowledgements

This work was supported by the National Natural Science Foundation of China [Grant Nos. 72071173, 71831008], AF Competitive Grants [Grant No. ZZQS], and the Research Grants Council of the Hong Kong Special Administrative Region, China [Project number HKSAR RGC TRS T32-707-22-N].

Appendix A. Comparison of different machine learning algorithms on predicting the hourly fuel consumption rate

Based on the dataset which combines the noon report data and meteorological data derived from the ERA5 dataset by the rhumb line based fusion method, the prediction performance of the XGBoost regressor is compared with several popular machine learning algorithms. Four popular regression models are selected for comparison: RF, DT, LASSO regression, and ANN. DT uses a tree structure to partition the training data into smaller subsets, and then predicts the output of each subset by calculating the mean of the outputs of the samples contained in the subset. Random forest builds multiple decision trees and combines their predictions to improve the accuracy and robustness of the model. ANN is a machine learning model inspired by the structure and function of the human brain, consisting of interconnected nodes that process and transmit information to make predictions. LASSO is a linear regression technique that uses regularization to prevent overfitting and improve the model's accuracy and interpretability by shrinking the coefficients of less important features towards zero. The datasets utilized to train and test the models are identical to those employed for the XGBoost regressor. In addition, we adopt the scikit-learn python library 1.0.2 to construct the RF, DT, LASSO, and ANN models. The hyperparameters for these models are tuned by the Hyperopt tool with five-fold cross validation. Table 13 summarizes the performance of the prediction models. It shows that the XGBoost model demonstrates the best prediction performance with the minimum MSE, MAE, and MAPE, as well as the highest R^2 among all the models.

References

Beşikçi, E., Kececi, T., Arslan, O., Turan, O., 2016. An application of fuzzy-AHP to ship operational energy efficiency measures. Ocean Eng. 121, 392–402.
Bialystocki, N., Konovessis, D., 2016. On the estimation of ship's fuel consumption and speed curve: A statistical approach. J. Ocean. Eng. Sci. 1 (2), 157–166.
Chen, T. and Guestrin, C., 2016, August. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

Chen, F.-W., Liu, C.-W., 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy Water Environ. 10 (3), 209–222.
Du, L., Gao, R., Suganthan, P.N., Wang, D.Z., 2022. Graph ensemble deep random vector functional link network for traffic forecasting. Appl. Soft Comput. 131, 109809.
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. B Methodol. 122, 88–114.
Du, Y., Chen, Y., Li, X., Schönborn, A., Sun, Z., 2022a. Data fusion and machine learning for ship fuel efficiency modeling: Part II-Voyage report data, AIS data and meteorological data. Commun. Transport. Res. 2, 100073.
Du, Y., Chen, Y., Li, X., Schönborn, A., Sun, Z., 2022b. Data fusion and machine learning for ship fuel efficiency modeling: Part III - Sensor data and meteorological data. Commun. Transport. Res. 2, 100072.
ECMWF, 2018. ERA5 hourly data on single levels from 1979 to present. Accessed 10 December 2022. https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview.
Fan, A., Yang, J., Yang, L., Liu, W., Vladimir, N., 2022. Joint optimisation for improving ship energy efficiency considering speed and trim control. Transp. Res. Part D: Transp. Environ. 113, 103527.
Fan, J., Yue, W., Wu, L., Zhang, F., Cai, H., Wang, X., Lu, X., Xiang, Y., 2018. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. Agric. For. Meteorol. 263, 225–241.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.
Garcia-Moya, J., Casado, J., Marco, I., Fernández-Peruchena, C., Gastón, M., 2016. Deterministic and probabilistic weather forecasting. Accessed 10 Dec 2022. https://hal.archives-ouvertes.fr/hal-02380127.
Gkerekos, C., Lazakis, I., Papageorgiou, S., 2018. Leveraging big data for fuel oil consumption modelling. In: Proceedings of the 17th conference on computer and IT application in the maritime industrial, pp. 144–152.
Gkerekos, C., Lazakis, I., 2020. A novel, data-driven heuristic framework for vessel weather routing. Ocean Eng. 197, 106887.
Hinnenthal, J., Clauss, G., 2010. Robust Pareto-optimum routing of ships utilising deterministic and ensemble weather forecasts. Ships Offshore Struct. 5 (2), 105–114.
IMO, 2009. Guidance for the development of a ship energy efficiency management plan (SEEMP). Accessed 10 December 2022. https://www.register-iri.com/wp-content/uploads/MEPC.1-Circ. 683.pdf.
IMO, 2019. Report of fuel oil consumption data submitted to the IMP Ship Fuel Oil Consumption Database in GISIS. Accessed 10 December 2022. https://www.imo.org/en/ourwork/envir onment/pages/data-collection-system.aspx.
IMO, 2020. Fourth IMO GHG Study 2020. Accessed 10 December 2020. https://www.imo.org/en/OurWork/Environment/Pages/Fourth-IMO-Greenhouse-Gas-Study-2020.aspx.
Jeon, M., Noh, Y., Shin, Y., Lim, O.-K., Lee, I., Cho, D., 2018. Prediction of ship fuel consumption by using an artificial neural network. J. Mech. Sci. Technol. 32 (12), 5785–5796.
Kim, Y., Jung, M., Park, J., 2021. Development of a fuel consumption prediction model based on machine learning using ship in-service data. J. Marine Sci. Eng. 9 (2), 1–25.
Lee, H., Aydin, N., Choi, Y., Lekhavat, S., Irani, Z., 2018. A decision support system for vessel speed decision in maritime logistics using weather archive big data. Comput. Oper. Res. 98, 330–342.
Li, X., Du, Y., Chen, Y., Nguyen, S., Zhang, W., Schönborn, A., Sun, Z., 2022. Data fusion and machine learning for ship fuel efficiency modeling: Part I - Voyage report data and meteorological data. Commun. Transport. Res. 2, 100074.
Linh, B., Ngoc, V., 2021. Utilization of a deep learning-based fuel consumption model in choosing a liner shipping route for container ships in Asia. Asian J. Ship. Log. 37 (1), 1–11.
Lorenz, E.N., 1965. A study of the predictability of a 28-variable atmospheric model. Tellus 17, 321–333.
Medina, J.R., Molines, J., González-Escrivá, J.A., Aguilar, J., 2020. Bunker consumption of containerships considering sailing speed and wind conditions. Transp. Res. Part D: Transp. Environ. 87, 102494.
NOAA, 2017. Global ensemble weather forecast. Accessed 10 December 2022. http://www.ncei.noaa.gov/products/weather-climate-models/global-ensemble-forecast.
Orrell, D., Smith, L., Barkmeijer, J., Palmer, T.N., 2001. Model error in weather forecasting. Nonlinear Process. Geophys. 8 (6), 357–371.
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. Clean. Prod. 264, 121564.
Petersen, J., Jacobsen, D., Winther, O., 2012a. Statistical modelling for ship propulsion efficiency. J. Mar. Sci. Technol. 17 (1), 30–39.
Petersen, J., Winther, O., Jacobsen, D., 2012b. A machine-learning approach to predict main energy consumption under realistic operational conditions. Ship Technol. Res. 59 (1), 64–72.
Psaraftis, H., Kontovas, C., 2014. Ship speed optimization: Concepts, models and combined speed-routing scenarios. Transport. Res. Part C: Emerg. Technol. 44, 52–69.
Rodwell, M., 2006. Comparing and combining deterministic and ensemble forecasts: How to predict rainfall occurrence better. ECMWF Newsletter 106, 17–23.
Ryder, S.C., Chappell, D., 1980. Optimal speed and ship size for the liner trades. Marit. Policy Manag. 7 (1), 55–57.
Skoglund, L., Kuttenkeuler, J., Rosén, A., Ovegård, E., 2015. A comparative study of deterministic and ensemble weather forecasts for weather routing. J. Mar. Sci. Technol. 20 (3), 429–441.
Soner, O., Akyuz, E., Celik, M., 2018. Use of tree based methods in ship performance monitoring under operating conditions. Ocean Eng. 166, 302–310.
Soner, O., Akyuz, E., Celik, M., 2019. Statistical modelling of ship operational performance monitoring problem. J. Mar. Sci. Technol. 24 (2), 543–552.
Sourtzi, V., 2019. Forecasting the fuel consumption on passenger vessels. University of Piraeus, Piraeus. Master's thesis.
Tarelko, W., Rudzki, K., 2020. Applying artificial neural networks for modelling ship speed and fuel consumption. Neural Comput. Applic. 32 (23), 17379–17395.
Tian, Q., Lin, Y.H., Wang, D.Z., 2023a. Joint scheduling and formation design for modular-vehicle transit service with time-dependent demand. Transp. Res. C: Emerg. Technol. 147, 103986.
Tian, X., Yan, R., Liu, Y., Wang, S., 2023b. A smart predict-then-optimize method for targeted and cost-effective maritime transportation. Transp. Res. B: Methodol. 172, 32–52.
UNCTAD, 2020. Review of Maritime Transportation 2020. Accessed 10 December 2022. https://unctad.org/system/files/official-document/rmt2020_en.pdf.
Uyanık, T., Karatuğ, Ç., Arslanoğlu, Y., 2020. Machine learning approach to ship fuel consumption: A case of container vessel. Transp. Res. Part D: Transp. Environ. 84, 102389.
Wang, S., Meng, Q., 2012. Sailing speed optimization for container ships in a liner shipping network. Transport. Res. Part E: Log. Transport. Rev. 48 (3), 701–714.
Wang, S., Meng, Q., Liu, Z., 2013. Bunker consumption optimization methods in shipping: A critical review and extensions. Transport. Res. Part E: Log. Transport. Rev. 53, 49–62.
Wang, S., Ji, B., Zhao, J., Liu, W., Xu, T., 2018b. Predicting ship fuel consumption based on LASSO regression. Transp. Res. Part D: Transp. Environ. 65, 817–824.
Wang, H., Yan, R., Au, M.H., Wang, S., Jin, Y.J., 2023. Federated learning for green shipping optimization and management. Adv. Eng. Inform. 56, 101994.
Wang, K., Yan, X., Yuan, Y., Jiang, X., Lin, X., Negenborn, R.R., 2018a. Dynamic optimization of ship energy efficiency considering time-varying environmental factors. Transp. Res. Part D: Transp. Environ. 62, 685–698.
Weintrit, A., Kopacz, P., 2011. A novel approach to loxodrome (rhumb line), orthodrome (great circle) and geodesic line in ECDIS and navigation in general. TransNav-Int. J. Marine Navigation Safety Sea Transport. 5 (4), 507–517.
Williams E., 2013. Aviation formulary V1.47. Accessed 10 December 2022. https://edwilliams.org/avform147.htm.
World Meteorological Organization, 2012. Guidelines on ensemble prediction systems and forecasting. World Meteorological Organization, Geneva, Switzerland.

Wu, L., Adulyasak, Y., Cordeau, J.F., Wang, S., 2022. Vessel service planning in seaports. Oper. Res. 70 (4), 2032–2053.
Yan, R., Wang, S., Du, Y., 2020. Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship. Transport. Res. Part E: Log. Transport. Rev. 138, 101930.
Yan, R., Wang, S., Psaraftis, H.N., 2021. Data analytics for fuel consumption management in maritime transportation: Status and perspectives. Transport. Res. Part E: Log. Transport. Rev. 155, 102489.
Yan, R., Wang, S., Zhen, L., 2023. An extended smart "predict, and optimize"(SPO) framework based on similar sets for ship inspection planning. Transp. Res. E: Logist. Transp. Rev. 173, 103109.
Yang, L., Chen, G., Rytter, N.G.M., Zhao, J., Yang, D., 2019b. A genetic algorithm-based grey-box model for ship fuel consumption prediction towards sustainable shipping. Ann. Oper. Res. https://doi.org/10.1007/s10479-019-03183-5.
Yang, L., Chen, G., Zhao, J., Rytter, N.G.M., 2020. Ship speed optimization considering ocean currents to enhance environmental sustainability in maritime shipping. Sustainability 12 (9), 1–24.
Yang, D., Wu, L., Wang, S., Jia, H., Li, K.X., 2019a. How big data enriches maritime research – a critical review of Automatic Identification System (AIS) data applications. Transp. Rev. 39 (6), 755–773.
Yousuf, M.U., Al-Bahadly, I., Avci, E., 2019. Current perspective on the accuracy of deterministic wind speed and power forecasting. IEEE Access 7, 159547–159564.
Yu, H., Fang, Z., Fu, X., Liu, J., Chen, J., 2021. Literature review on emission control-based ship voyage optimization. Transp. Res. Part D: Transp. Environ. 93, 102768.
Yuan, J., Nian, V., 2018. Ship energy consumption prediction with Gaussian process metamodel. Energy Proc. 152, 655–660.
Yuan, Y., Wang, K., Yin, Q., Yan, X., 2020. Review on ship speed optimization. J. Traffic Transport. Eng. 20 (6), 18–34.
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), 1–22.
Zis, T.P.V., Psaraftis, H.N., Ding, L.i., 2020. Ship weather routing: A taxonomy and survey. Ocean Eng. 213, 107697.