The 8th International Symposium on Ship Operations, Management & Economics (SOME)
7-8 March 2023, Athens GR

From STEAM to Machine: Emissions control in the shipping 4.0 era

Dimitrios Kaklis[1,2,4], Takis J. Varelas[1], Iraklis Varlamis[2] , Pavlos Eirinakis[3] , George Giannakopoulos[4], Constantine V. Spyropoulos[4]
1. Danaos Research Center, 2. Harokopio University of Athens, Department of Informatics and Telematics
3. Department of Industrial Management & Technology University of Piraeus, Piraeus, Greece
4. Institute of Informatics & Telecommunications, NCSR Demokritos, Athens, Greece

The maritime sector is required to adhere to the IMO 2020 - mandated reduction of emissions. This reduction can be conducted by either using a compliant fuel with lower sulfur content, an alternative fuel (e.g. LNG, methanol), or clean its exhaust gasses with a "scrubber" technology to reduce the output of CO_2, NO_x and SO_x emissions. The objective of this paper is to present a holistic approach to continuously monitor and estimate the emissions of a vessel as well as to assess and improve the efficiency of scrubbers. Furthermore the deployment of a cutting-edge, integrated framework, incorporating the latest technological advances, that can offer the ability to capture, process and analyze vessels' operational data in order to improve efficiency, sustainability, and rule compliance is presented. Particularly the conceptualization and materialization of a big data application suite that exploits the IoT (Internet of Things) and AI (Artificial Intelligence) advancements and technologies, to employ a "digital replica" of the en-route vessel is demonstrated. By collecting a multitude of features from on-board sensor installments, we present how we can effectively utilize these features, harvested in real time, in order to accurately assess and estimate the environmental footprint of the vessel by employing robust Fuel Oil Consumption (FOC) predictors. Then we describe in detail the streamlined procedure from data acquisition to model deployment, utilizing the proposed big data framework, in order to assess and estimate the emissions during the operational state of the vessel. Finally, we demonstrate experimental results by deploying comparative analysis utilizing operational data from one containership-centric Living Lab (LL) in order to validate and confirm our approaches in terms of accuracy and performance in a real world setting.

KEY WORDS: Digital Twin; IoT; Deep Learning; Admiralty coefficient; Fuel Oil Consumption estimation; Emission Control; STEAM methodology

INTRODUCTION

State of the art efficient and comprehensive data management is an indispensable component for modern pertinent scientific research as well as for cross-sectoral multi-disciplinary applications and technological advancements. The maritime sector has witnessed an exponential growth in data availability over the past few years, a fact that renders mandatory the utilization of state of the art interoperable frameworks responsible for aggregating – processing and analyzing this vast amount of data. Cutting-edge technologies are linked to the so-called "Fourth Industrial Revolution", often known as "**Industry 4.0**" in general literature. Representative examples are Artificial Intelligence (**AI**) Big Data Analytics (**BDA**), **Cloud Computing** and Internet of Things (**IoT**) applications that are already influencing the maritime industry, which is gradually and steadily transitioning into a new operational blueprint, often termed as "shipping in the era of **digitalization**". Shipping companies promote digitalization as the future of the maritime industry and their efforts to set up relevant strategies are already in progress. Examining these initiatives in relation to digitalization would provide stakeholders with a better understanding towards the way the maritime industry is heading.

RELATED WORK

Traditionally, due to the lack of actual measurements, the vessel emissions are calculated using generic (i.e., non-vessel specific) mathematical models that do not take into consideration vessels' operational data. Although the predictions constitute a fair basis to extract some preliminary conclusions and insights regarding the carbon footprint of the vessel during a voyage, they obviously deviate from the ground truth. A well-known model is the generic STEAM (Ship Traffic Emissions Assessment Method) (Jalkanen et al. 2012), that is based on an open-source database of vessel particulars (IHS fairplay database - *fairplay.ihs.com*). The STEAM model assesses the power consumption, load of the engine and the fuel-oil consumption of the ship. Based on these values, the model is used to evaluate the emissions of NO_x, SO_x, CO, CO_2, as a function of time and location. In the current version of the model, the engine's loads during voyages can be determined with improved accuracy based on the ratio of the vessel's speed over the calculated resistance that the ship is required to overcome at a specified speed (Hollenbach 1998).

In the context of the EU project EMERGE (Moussiopoulos et al. 2020), the aforementioned method was further advanced and expanded in order to include environmental conditions in the emissions model. More specifically, the STEAM model was extended with a capacity to include the impact of various ambient contributions on ship's fuel consumption and emissions. External factors were included in the modeling like wind, sea-waves, -ice and -currents. This extension enables more realistic emission modeling, but can also be used to estimate the magnitude of the effect of different features to ships' fuel consumption. Although this method refined the initial STEAM approach it still lacks the approximation capabilities of an IoT based data-driven model continuously refined and adapted by real time sensory measurements.

Most of FOC theoretical calculations, found in pertinent naval engineering literature, are based on the **Admiralty coefficient** which is extensively used by marine practitioners and engineers in the estimation of the power that is required for a new build design to attain the required speed, and is given by the formula:

$$k = (\sqrt[3]{\Delta^2} * V^3)/ehp$$

where Δ is the displacement (tn) of the vessel, V is the desired speed and *ehp* is the effective horse power (kWh)

The techniques employed in the literature for estimating the FOC of the vessel, based on vessel characteristics and/or environmental conditions, can be grouped into the following categories:

- White box models, where analytical equations and approximation methods (e.g. Computational Fluid Dynamic equation - CFD), are exploiting a variety of vessel specific variables and hydrodynamic principles to model the added resistance of the hull of a specific vessel. The admiralty constant as well as the resistance constant ($RL/\Delta V^2$), where R is the total resistance and L is the length overall of the vessel, proposed by (Telfer 1963) has been utilized in the past to replicate the hydrodynamic behavior of new ship designs by using only the design point values of the corresponding parameters (speed (V), displacement (Δ), length (L), breadth (B), draught (T)), rather than the whole spectrum of the operational domain. Thus it can be easily inferred that the purpose of the Admiralty constant was to provide an initial baseline to compare the hydrodynamic performance of different ships in their respective design conditions, rather than monitor their operational state. This makes evident that the Admiralty constant was neither intended nor demonstrated to be a suitable operational hydrodynamic performance indicator for a ship.

- Data-oriented approaches that combine vessel-trajectory data, gathered from sensors, satellites (AIS data) or Noon Reports, with Machine and Deep-Learning algorithms. These techniques are ranging from simple Regression analysis, using stand-alone models like Support Vector Regression (SVR), Lasso Regression (LR), Polynomial Regression, to ensemble non-parametric schemes like Random Forest regression (RF), Decision Trees or AdaBoost where the approximation power of each model is combined appropriately in order to infer the underlying function.

- Approaches where machine learning (ML) methods are used in conjunction with analytical models in order to increase the prediction accuracy. The former are also known as black-box models (**BBM**), and the latter are known as white-box models (**WBM**) and comprise equations of motion of a freely floating body moving with constant forward speed. The proposed models are known as grey-box models (**GBM**) (Corradu et al. 2017) (Kaklis et al. 2019).

ANNs (Artificial Neural Networks) have been in the center of attention lately in many research areas. As far as FOC of the vessel is concerned not many studies utilize the computational power of ANN's to approximate FOC mainly due to the problem of missing historical data. The studies found in pertinent literature dealing with FOC estimation from a deep learning perspective are presented briefly below. Some studies experimented with baseline sequential ANN's by applying a dropout in the weights in order to achieve better generalization error (Gkerekos and Lazakis 2020) or by tuning a number of hyper parameters (learning rate, number of neurons, number of layers, activation function) utilizing brute force methods like randomized grid search (Papandreou and Ziakopoulos 2020), (Savitha and Abdullah 2017) (Miyeon et al. 2018). Yongjie, Zuo, and Li (2020) employed a Recurrent NN in order to estimate FOC, but without further research as far as the architecture of the neural is concerned.

The methods presented above are usually tested for a single vessel and therefore lack the generalization capabilities of models evaluated in a variety of ships that are able to adjust and adapt to the underlying function that describes the relationship between FOC and each specific vessel, continuously, by exploiting the vast amount of data collected by IoT installments. Frameworks and technological advancements regarding the continuous monitoring of the vessel are inextricably linked with the emerging concept of the so-called **Digital Twin** in the shipping industry, as they employ a digital replica of the en-route vessel that is able to simulate-project and validate in real time the majority of the operational procedures. The term Digital Twin, first introduced in (Grieves, 2014), constitutes a virtual holistic representation of the vessel that spans its life-cycle and is updated from near to real-time data, utilizing simulation, machine learning and reasoning to help in decision-making, sensing and control actuation.

In the following paragraphs, we present a prototype of a Digital-Twin(ning) (DT) framework attempting to exploit the abundance of features collected from sensor installments on-board, by coupling the existing data acquisition network with a novel fault-tolerant adaptive streaming pipeline responsible for data processing and model deployment.

Utilizing this framework, we also introduce a novel deep learning architecture that combines the approximation capabilities of a data driven model, with theoretical naval engineering methods in a consolidated physics-informed approach, to enhance our predictions when estimating FOC and therefore the emissions of a fleet.

PROTOTYPE OF A DIGITAL-TWINNING FRAMEWORK

Kaklis et al. (2022b) demonstrate a Big Data Analytics - multipurpose - toolkit adapted to the needs of the maritime sector. The proposed framework incorporates a variety of state of the art streaming tools for real-time analysis of vessel data as well as tools for continuous integration/deployment (CI/CD) of ML/DL models regarding operational optimization, causal analysis and event recognition. By utilizing the company's existing in-house infrastructure concerning Edge-Headquarter (EDGE-HQ) communication between the vessel and the office, we can incorporate the aforementioned pipeline in a broader data acquisition network in order to aggregate, synchronize and process data coming from the vessel in real-time. The resulting platform (Fig. 1) constitutes a prototype version of a DT framework enabling sensing and control actuation on the vessel that aims to assist shipowners to achieve efficiency in fleet management with tangible benefits in terms of emission reduction, environmental compliance and protection of crew safety onboard.

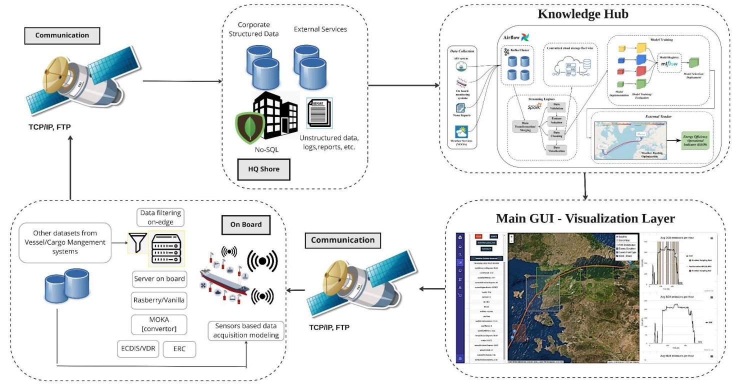

Figure 1 : The envisaged DT framework for the maritime sector

Mainly, the proposed DT framework consists of the following components:

- **IoT backbone suite -** *Data acquisition layer*
- **Knowledge Hub -** *Processing - Orchestration - Computing - Deployment layer*
- **Main GUI -** *Visualization layer*
- **Edge Computing -** *Sensing & Control actuation layer - Requirements & Refinements elicitation*

In order to collect a feature set that can be utilized in the context of a data-driven FOC approximation model, we aim to utilize the core functionality of the **Knowledge Hub** to assemble features that correspond mainly to:

- Speed Though Water of the vessel that is based on the GPS speed.
- The Mid Draft of the vessel that corresponds to the mean value of the Forward and the Aft draft of the vessel.
- The weather state at a specific time-instance and location. Weather state is described by a multitude of features like (Current Speed/Direction, Wind Speed/Direction, Waves Height/ Direction, Swell Wave Height/ Direction, Humidity, Visibility, Temperature etc)
- The Fuel Oil Consumption (FOC) at a specific time instance corresponding to the aforementioned features.

In the following subsections we will demonstrate a streamlined procedure, incorporating a set of state of the art algorithms, that aims to exploit, analyze and process the vast amount of data acquired utilizing the proposed DT-framework, in order to extract useful patterns and insights concerning the task of FOC approximation.

Knowledge Hub

In this section we briefly describe a core module of the proposed DT framework, as depicted in Figure 1, the Knowledge Hub (KH). KH incorporates a variety of multi-disciplinary approaches regarding data provision, re-usability and curation as

well as state of the art frameworks for model versioning and deployment. It constitutes a holistic approach that aims to create an adaptive versatile observatory for the shipping industry that comprises structured methodologies for inter-connecting each use case with the appropriate data, processing algorithms and simulation models. All these are joined together adequately, facilitating towards the decarbonization of the maritime sector. Figure 2 illustrates a multi-modal streamlined procedure, stored in KH, adapted to the task of FOC estimation.
The Knowledge Hub aims to largely simplify and standardize the way the various tools and services provided by the DT's ecosystem are operating and communicating with each other, following the standards of an ICT (Information Communication Technology) framework. The general streamlined procedure is based on

- Data Processing and feature selection
- Data Curation from bias and noise
- Model Deployment

Data processing focuses on determining the most important features for FOC estimation and data curation accounts for removing the bias (outliers, faulty measurements) from the bulk of data collected in real time from IoT installments. As a post-processing step the calculation of correlation between the most important features results in the selection of an ideal feature set that combines importance and independence. The resulting feature set is utilized accordingly in the training process of a data driven FOC estimation model.

As showcased in Figure 2, the general procedure can be adapted to the needs of a specific use case by selecting the appropriate algorithms and models (shown on the side of each customized flow) and applying them into practice.

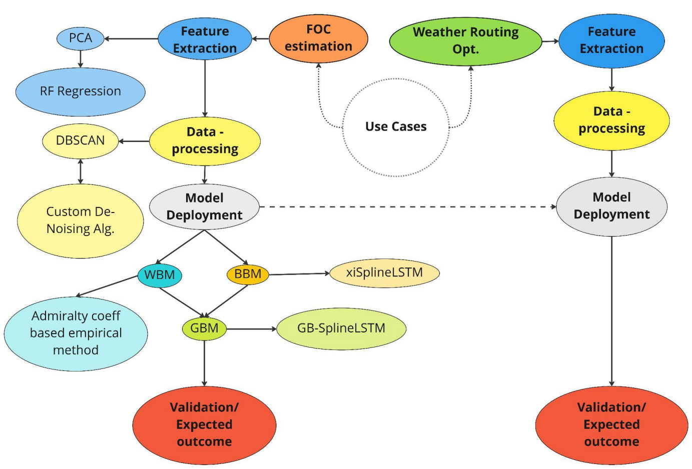

Figure 2: The streamline procedure adapted for the FOC estimation use case.

In the paragraphs that follow, we provide more details on each step from the perspective of its application to the FOC use case.

Feature Selection

Data coming from on-board sensor installments concern different compartments of the vessel (Bridge, Engine rooms, Deck) and consists of approximately $\simeq$ 500 features. However, FOC is highly impacted by the total resistance of the vessel as it moves forward, so this hydrodynamic force opposing the movement of the ship along its longitudinal axis, which is known as total resistance, is the most critical feature to estimate in order to correctly predict FOC.

The total resistance of the vessel consists of several components: frictional resistance, viscous-pressure resistance, wave resistance, added resistance and air resistance. **Frictional resistance** is due to shear stress and it depends on the size of the wetted area of the vessel. It usually represents about 70-90% of the ship's total resistance for low-speed ships (bulk carriers and tankers) and accordingly less than 40% for high-speed ships (containers and passenger ships). Viscous-pressure resistance depends on normal stresses and usually ranges from 5% to 10% of the frictional resistance. **Wave resistance** stems from the non-viscous pressure created by the wave system created by the ship moving steadily in calm waters and may rise up to 30% of the calm water resistance for moderate to high Froude numbers[1]. Furthermore, **Residual resistance** is defined as the total resistance minus the frictional resistance and as a result consists of the wave resistance and the viscous pressure resistance due to the form or curvature of the hull. **Added Resistance** measures the effect of waves and may rise up to 30% of the calm-water resistance. The characteristics of waves like their amplitude and wave length are determined from the ocean-wave spectra along the voyage path. These spectra are modeled on the basis of the strength and duration of the wind and the geographical area. Wave characteristics (height, length) are used as forcing terms in the system of hydrodynamic equations in order to model the wave-body interaction of ships moving through wind-generated waves. Another component that constitutes the total resistance is **Air resistance** that normally represents about 2% of the total resistance, however, for loaded container ships in head wind, the air resistance can be as much as 10%.

Based on the above, through Feature Selection we create a subset of the original set of ~500 features that consists of features which heavily affect total resistance, such as:

- Features that correspond to the frictional resistance and can be utilized in the context of a FOC estimation scheme like Speed Through Water (STW), Draft and Displacement.
- Features that describe the wave resistance components (Wave height/Direction, Wave Period, Swell Wave Height/Direction, Swell Period).
- Features that model the air resistance component (Wind Speed/Direction, Combined Wind Wave Height/Direction, Current Speed/Direction).

In order to further reduce the dimensionality of our dataset, we conduct PCA (Principal Component Analysis) to determine the

[1] The Froude number Fr=V/sqrt(gL) stems from the dimensional analysis of inertia vs gravitational forces.

principal components (expressed as appropriate linear combinations of the initial set of features), that can model the dataset in its entirety, in terms of explainable variance, in the best way possible. As depicted in the following graph (Fig. 3), the variance entailed in the dataset can be attributed, as a whole, to the first 10 components extracted from PCA.

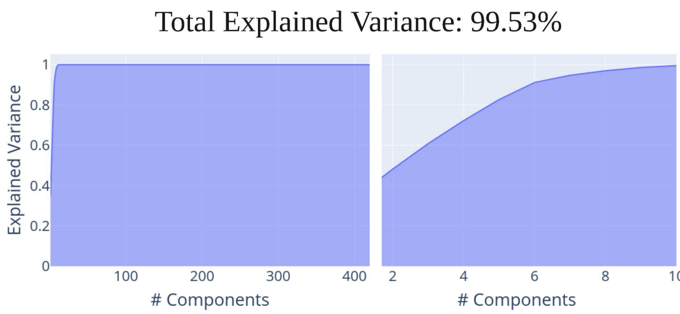

Figure 3: Dimensional reduction via PCA

In Kaklis et al. 2022b we demonstrated the applicability of a consolidated approach that combines *Random Forest* and *Spearman Correlation* to extract the most important and independent features to estimate FOC. By applying this algorithm and taking into account the PCA presented above, we conclude with the most important independent features to be exploited in the next sections in order to approximate FOC via data driven methods.

In Table 1 we depict the experimental results from conducting multiple regression analysis utilizing the aforementioned algorithm. Detailed description of the features and their abbreviation as well as their measurement unit is also presented.

Table 1: Feature ranking utilizing RF algorithm

	Algorithm		
	RF		
Ranking	**Feature**	Abbrv.	Meas. Unit
1°	Power	P	kW
2°	Speed Through Water	STW	kn
3°	Displacement	Δ	tn
4°	Draft	Dr	m
5°	Combined Wave Height	CWH	m
6°	Swell Wave Height	SWH	m
7°	Current Speed	CS	kn
8°	Current Period	CP	sec
9°	Swell Wave Period	SWP	sec
10°	Current Direction	CD	°
11°	Swell Wave Direction	SWD	°

Data Cleaning

Raw data, collected from the sensors of the vessel, are in time-series (minutely) form and tend to be "noisy" (high variance, high standard deviation from the mean) and in some cases even erroneous. In order to remove noise, Kaklis et al. (2022b) employ a fit & filter technique that effectively "cleans" the data, but at the same time keeps the bulk of information needed for training robust predictive models. The raw vessel's speed and corresponding FOC collected from the sensors versus quasi steady filtered data utilizing the algorithm from (Kaklis et al. 2022b) is depicted in Figure 4.

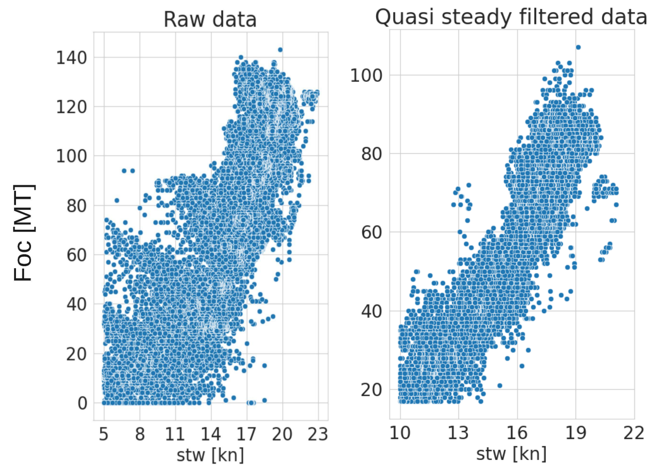

Figure 4: Measured STW vs FOC for a timespan of one year

An alternative method for removing noise from the dataset is to apply the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) clustering algorithm. DBSCAN is an unsupervised machine learning technique used to identify clusters of varying shape in a data set (Ester et al. 1996). It can identify clusters in large spatial datasets by looking at the local density of the data points. Its main advantage is its robustness against outliers and noise, which are removed from the clustering scheme.

Figure 5: "Cleansed" version of STW vs FOC with DBSCAN

DBSCAN can work without an expected number of clusters (such is the case with the popular K-Means clustering algorithm), and requires two parameters: *epsilon* and *minPoints* to define dense clusters of arbitrary shape. *Epsilon* is the radius of the circle to be created around each data point to check the density and *minPoints* is the minimum number of data points required inside that circle for that data point to be classified as a cluster.

Figure 5 depicts the resulting clusters, and corresponds to the points that remain after removing noise. What is even more important, when comparing the plots in Figures 4 (right) and 5 is that the DBSCAN based noise removal method is in agreement with the denoising procedure demonstrated in (Kaklis et al. 2022b). More specifically, if we apply the Kolmogorov-Smirnov (KS) test to the distributions of the two "cleaned" versions of the data, we validate that they are following the same pattern. KS is a non-parametric test (normality is not a prerequisite) that evaluates the maximum absolute difference between the cumulative distributions of the two groups as follows:

$$stat = sup_x|F_1(x) - F_2(x)| \tag{2}$$

where F_1, F_2 are the two cumulative distribution functions and x are the values of the underlying variable (here FOC).
We can visualize the value of the test statistic, by plotting the two cumulative distribution functions and the value of the test statistic as well as their histogram (Fig. 6).

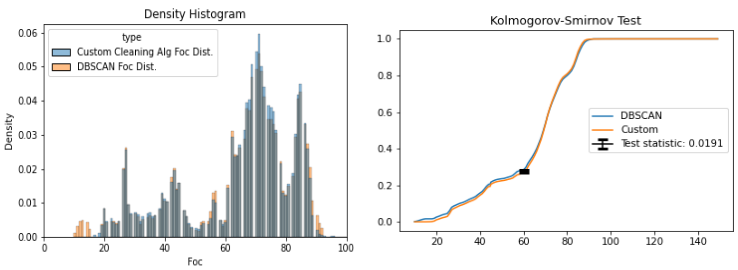

Figure 6: Histogram and KS-Test of the FOC values extracted after applying the Custom de-noising algorithm and DBSCAN

MODEL IMPLEMENTATION

The dynamic estimation of FOC based on vessel state and environmental conditions can be examined as a multivariate time-series prediction problem, that takes into account the actual values, as well as their recent history, and captures the information hidden in the values evolution over time. Based on the superiority of recurrent, Long short-term memory (LSTM) Neural Networks, over traditional time-series prediction methods (e.g., ARIMA in (Namini et al. 2018), LTSM NNs are chosen as the basis of our solution.
In a previous work (Kaklis et al. 2022b) we demonstrated the approximation capabilities of a novel recurrent neural network that was utilizing a reduced featured set (vessel, speed, basic weather features e.g wind speed and wind direction), in order to approximate FOC. To expand this limited set of features we employed a pre-training step that attempts to extract prior “knowledge” by training multiple Spline-based regression (Friedman 1991,) models in a parallel fashion and distilling this information gradually into the architecture of an LSTM neural architecture.

In what follows, we describe how this LSTM network is exploited and further adapted in an enriched feature set to approximate FOC and eventually CO_2 emissions, as well as how it can be utilized in conjunction with analytical models, calculating the total resistance on the basis of fluid dynamics and naval engineering.

White Box modeling

The admiralty coefficient is often used as a hydrodynamics performance indicator of the vessel and it summarizes the relationship between speed power and displacement (Gupta et al, 2021). In the context of preliminary ship design, it is necessary to approximate the power consumed by the propulsion system of the ship without resorting to model experiments. One method that has been widely used for this exact purpose utilizes the admiralty coefficient. This is based on the assumption that for small variations in speed the total resistance may be expressed in the form:

$$\Delta^{2/3}V^3/P \tag{3}$$

where V is the vessel speed, P is the power absorbed by the ship propulsion system and Δ is the displacement of the vessel. The coefficient was originally utilized to model the hydrodynamics performance for a variety of different ship designs. The speed (V) and power (P) were corresponding to the design point values of different ships rather than the whole operational domain. Nevertheless the idea that we can model calm water resistance of the vessel by employing a simple analytic function, connecting the displacement, speed and power is appealing if combined with a statistical approach, that exploits the vast amount of operational data collected on-board and thus ultimately unveiling the relationship between those variables and resistance in calm waters.

From a mathematical standpoint, the admiralty coefficient defines a log-linear relationship between the engine propulsion, the speed-through-water (V), shaft power (P) and vessel displacement. It also takes into account, indirectly, the vessel’s draft as the process of calculating displacement includes first determining the *mean draft* (averaging the port, starboard, sides forward, midships, and astern draft marks of the vessel). Since all the above are always subject to the vessel specifics, the proposed theoretical model can be applied in any vessel utilizing a multiplicative empirical exponential formula comprised of the top four features (Power, Speed, Draft and Displacement) as extracted from the preliminary data analysis conducted in previous section and is described as follows:

$$y(x_i, \ldots, x_n) = c_0 \prod_{i=1}^{n} x_i^{c_i} \tag{4}$$

where y is the FOC of the vessel, $x_1 \ldots x_n$ are the aforementioned variables and ci are constants that can be expressed as a function of characteristics of the hull below the waterline (shape and roughness of the hull).

In order to conclude with a linear combination of the variables in Eq. 4 we apply the natural logarithm ln on both sides of the equation, which is then transformed to:

$$ln(y(x_i, \ldots, x_n)) = ln(c_0) + \sum_{i=1}^{n} c_i ln(x_i) \tag{5}$$

Since data corresponding to vessel's specifics variables describing the geometry and the state of the hull are not available we attempt to give an initial approximation of the exponents $c_1, \ldots, c_n$, by training a baseline linear regression model $\mathcal{L}_r$ utilizing equation (5) with historical data, excluding the weather impact (calm water resistance) provided by on-board sensor installments.

In order to evaluate the proposed empirical formula with a well established model calculating the total resistance of the vessel we are employing the Guldhammer and Harvald method (Guldhammer and Harvald, 1965), an empirical method based on an extensive analysis of many published model tests. The method depends on relatively few parameters and is used for residual resistance prediction in the present analysis. The method presents curves for the total calm water resistance (C_T) as function of three parameters: i) The length-displacement ratio ($L/\Delta^{1/3}$), ii) the prismatic coefficient (C_p) and finally iii) the Froude number (F).
In Figure 7 we depict the calculated Power using Guldhammer-Harvald method, for a test vessel utilized in the current study, as well as the *quasi steady* mean M/E Power in calm water conditions acquired from on-board sensors installments for different speed ranges.

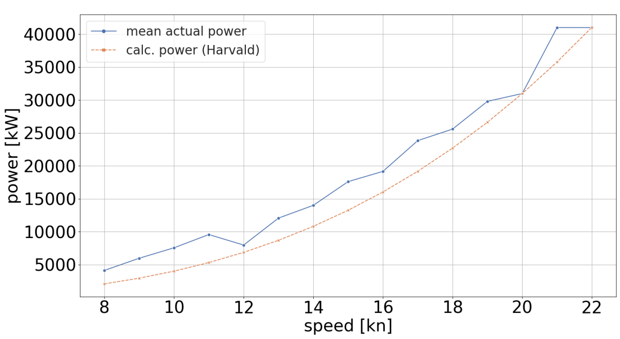

Figure 7: Power (P) vs Speed (V)

We can observe from the graph that the Speed-Power Guldhammer-Harvald curve follows the operational Power reported from the vessel, thus providing a strong baseline. It also allows us to extract useful insights and provide any reasoning or causality regarding deficiencies in the operation of the propulsion system, or the fouling-degradation of the hull. More specifically by evaluating the deviation between the reported and the calculated Power:

$$\Delta P = P_{data} - P_{emp} \tag{6}$$

we can estimate the fouling friction coefficient (ΔC_F) of the hull over time.

It is clear that the white box method that we propose for FOC approximation (Eq. 5) acts in a similar way by calculating the residuals between the reported FOC in calm water conditions and the approximated value.

Black Box modeling: SplineLSTM

In a previous work (Kaklis et al. 2022b) we demonstrated the approximation capabilities of spline-based regression models (Friedman 1991) and their ability to adapt to the linear and non-linear patterns that exist between dependent and independent variables, as those that describe the underlying function that approximates FOC. A spline regression of degree d partitions the input space in sub-domains separated by k knots. Each domain is approximated by different polynomials of degree up to d. Splines of order d have continuous $d-1$ derivatives, a property that balances the trade off between goodness-of-fit and smoothness of the spline interpolant, and results in a predictive scheme with good generalization capabilities.
An example of spline regression and the polynomials that are constructed in training time given a multivariate feature set: $x_1, \ldots, x_N$ and a target variable y is described as follows:

$$y = f(x) = \begin{cases} H_{1i}(x_{1i}, ..., x_{1N_i})b_{1i}, & x_{1i} \in [t_i, t_i + \Delta t] \\ \vdots & \\ H_{ki}(x_{ki}...x_{kN_i})b_{ki}, & x_{ki} \in [t_i + (k-1)\Delta t, t_i + k\Delta t], \\ 0 & otherwise \end{cases} \tag{7}$$

where:

- t_i is the corresponding time-step of observation x_i,
- k is the number of knots of the pre-trained SplineModel_i,
- $H(x)$ are piecewise continuous hinge functions of order $d \geq 1$ defined on subsequent time intervals, and
- b_{in} are the regression coefficients of the pre-trained spline models, where $i \in [1, k]$.

The proposed spline-based LSTM network uses the knowledge gained in the pre-spline-training step described above, for providing an extended input vector of size $k + N$ in each case, where k is the number of knots of the respective spline function. The Spline network evaluates each feature x_i on the corresponding hinge function H_i generated from Spline regression creating a k dimensional vector that quantifies the impact of each feature in FOC estimation.
The basic extension compared to a conventional LSTM is that by introducing this k-dimensional spline-informed vector the network is able to take into account not only the temporal but also the spatial structure of the features. This approach guides the network to form spatial-aware embeddings that help the

model learn a different set of functions for different sub-domains of interest.

With this transformation, the LSTM *looks back* m time steps to form the hidden state units h_{t-1}. The hidden state acts as the NN memory, for it holds information on data the network has *seen* before. The input vector is constructed by moving time windows that comprise:

- $(F_1, \ldots, F_N)$ values,
- k values generated from evaluating our feature set
- $(F_1, \ldots, F_N)$ values at each of the k knots of the pre-trained Spline model, and
- the corresponding FOC values.

In previous works (Kaklis et al. 2022a), (Kaklis et al. 2022b), it has been thoroughly demonstrated the efficiency and applicability of the proposed recurrent Neural scheme by evaluating its performance in a variety of settings. Ranging from training the model in a continuous manner (online) or utilizing its output as the main component of a cost function evaluating the next way-points in the context of a Weather Routing optimization algorithm, SplineLSTM showcased promising accuracy in terms of FOC estimation, when compared with baseline regression ML approaches.

Grey Box modeling: Integrating the theoretical FOC model with SplineLSTM: GB-SplineLSTM

Grey Box models (GBM), a term first introduced in (Bohlin 1991), are a combination of theoretical models (White Box Models (WBM) and data driven approaches (Black Box Models (BBM)). WBM's are usually simpler models in terms of computational complexity and they attempt to calculate the target variable of the approximation problem at hand from a theoretical standpoint by applying the contextually appropriate law of physics that governs each problem. In the context of this paper we will refer to a WBM as f_{WBM}.

In their simplest implementation GBMs are attempting to integrate prior knowledge extracted from a theoretical model into a BBM.
They do this by incorporating two approaches:

- a naive approach (N-GBM) where the output of the WBM is utilized as a new feature in the BBM training process.
- a more advanced approach (A-GBM) where a regularization term is introduced in the loss function of the BBM, that attempts to encompass the knowledge extracted from WBM in the weight vector w learned from the BBM.

Our aim in this work is to combine the BBM proposed in previous works with the theoretical backbone of Computational Fluid Dynamics (CFD) and develop a more efficient GBM that could better approximate the added resistance on the wetted area of the vessel. We test the N-GBM approach by incorporating in the training process of a BBM (*SplineLSTM*), the output of the theoretical baseline model introduced in section. The N-GBM approach allows the creation of a new dataset that takes the following form:

$$\mathcal{D}_n = \left\{ \left(\begin{bmatrix} x_1 \\ f_{WBM(x_1)} \end{bmatrix}, y_1 \right), \ldots, \left(\begin{bmatrix} x_n \\ f_{WBM(x_n)} \end{bmatrix}, y_n \right) \right\} \quad (8)$$

This new dataset can then be employed to generate a BBM in the form $f_{BBM}([x, f_{WBM}(x)])$ as presented in (Corradu et al. 2017). One way to include the prior information from the WBM is to replace the true label y_i in the dataset with the true label y_i minus the output of the $f_{WBM}(x_i)$ for the corresponding input. With this approach the dataset $\mathcal{D}_n$ takes the following form:

$$\mathcal{D}_n = (x_1, y_1 - f_{WBM}(x_1)), \ldots, (x_n, y_n - f_{WBM}(x_n)) \quad (9)$$

and the minimization process of the GBM algorithm is formulated as follows:

$$min \sum_{i=1}^{n} [w_i * f_{BBM}(x_i) - (y_i - f_{WBM}(x_i))]^2 \quad (10)$$

EXPERIMENTAL RESULTS

In this section we thoroughly describe the dataset utilized for our experiments and we demonstrate the results corresponding to different voyages conducted by a test vessel, comparing the approximation capabilities of the proposed FOC models described in previous sections versus the actual FOC acquired from the on-board flowmeter sensor.

Dataset

The experiments presented in the context of this work were conducted on one of Danaos container-ships in the context of Carbon Intensity Index (CII) calculation - projection. The ship is a 300-meter container vessel with a gross tonnage of 75200 tonnes, capable of transporting 6160 TEU (Twenty-foot Equivalent Unit - unit of cargo capacity used for container ships and terminals). The values collected correspond to a vast majority of different round-trip voyages at different periods and geographical locations (Fig. 7). As a whole, the extracted dataset covers a time span of one year (December 2020 - December 2021) comprising of approximately $4 * 10^5$ data points.

The ship was retrofitted with an open-loop exhaust gas cleaning system (i.e., scrubber). Along with the main scrubber unit, a dedicated software was also installed from the manufacturer to measure and monitor several parameters, such as PH, turbidity, SO_2/CO_2 ratio, inlet and outlet exhaust temperatures. The measurement of CO_2 and SO_2 emissions and the scrubber efficiency were evaluated during a nine-day voyage from Rotterdam to Kanakale.

During this period a team of environmental engineers, referred to as the EE-team in the sequel, boarded the vessel and conducted a preliminary research in the context of the EMERGE EU project regarding emission control and assessment. More specifically, they evaluated and compared the emissions calculated using STEAM methodology with the emissions acquired from on-board instruments. The results are further analyzed and assessed next, by comparing the emissions estimation, utilizing STEAM methodology, with the Grey Box model proposed in the context of this paper.

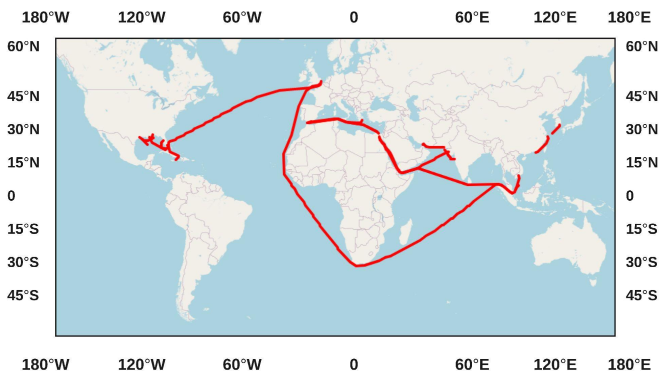

Figure 7: Route visualization

GB-SpLineSTM Performance in different voyages

In this section we demonstrate the approximation capabilities of GB-SplineSTM by evaluating its performance (mean absolute difference between the actual and the model's predicted FOC) on four different voyages conducted by DANAOS' test vessel. We also demonstrate and compare the BBM's (SplineLSTM) performance. In the graphs depicted below we can observe for different speed ranges ($\pm$ 0.5), the average deviation between the actual FOC as measured by the flowmeter (in red) and the predictions from SplineLSTM (in blue), GB-SplineLSTM (in green), as well as the number of observations found during a voyage, for a particular speed range (vertical bars - axis titled *Size*).

	Total Act FOC(MT/d)	Total Pred FOC(MT/d)	FOC Abs Diff(MT/d)	FOC Perc Diff %
SUEZ - ROTTERDAM	1100.62	1104.59	1.93	0.35
TANGER MED - SUEZ	411.7	413.42	1.64	0.41
MUNDRA JEDDAH	437.04	439.38	1.28	0.53
LE HAVRE TANGER MED	284.26	288.81	1.02	1.57
Total	**2233.62**	**2246.2**	**1.46**	**2.86**

Table 2: Computational performance of the FOC approximation model (GB-SplineLSTM)

STEAM vs GB-SplineLSTM emissions approximation comparison

The EE-team embarked on-board the aforementioned DANAOS' container vessel with a Horiba PG-350 portable multi-gas analyzer. Exhaust concentrations of CO, CO_2, NO_x, oxygen (O_2), and SO_x emissions were measured following the ISO 8178-2 protocol. The applied methodology is the generic STEAM (ship

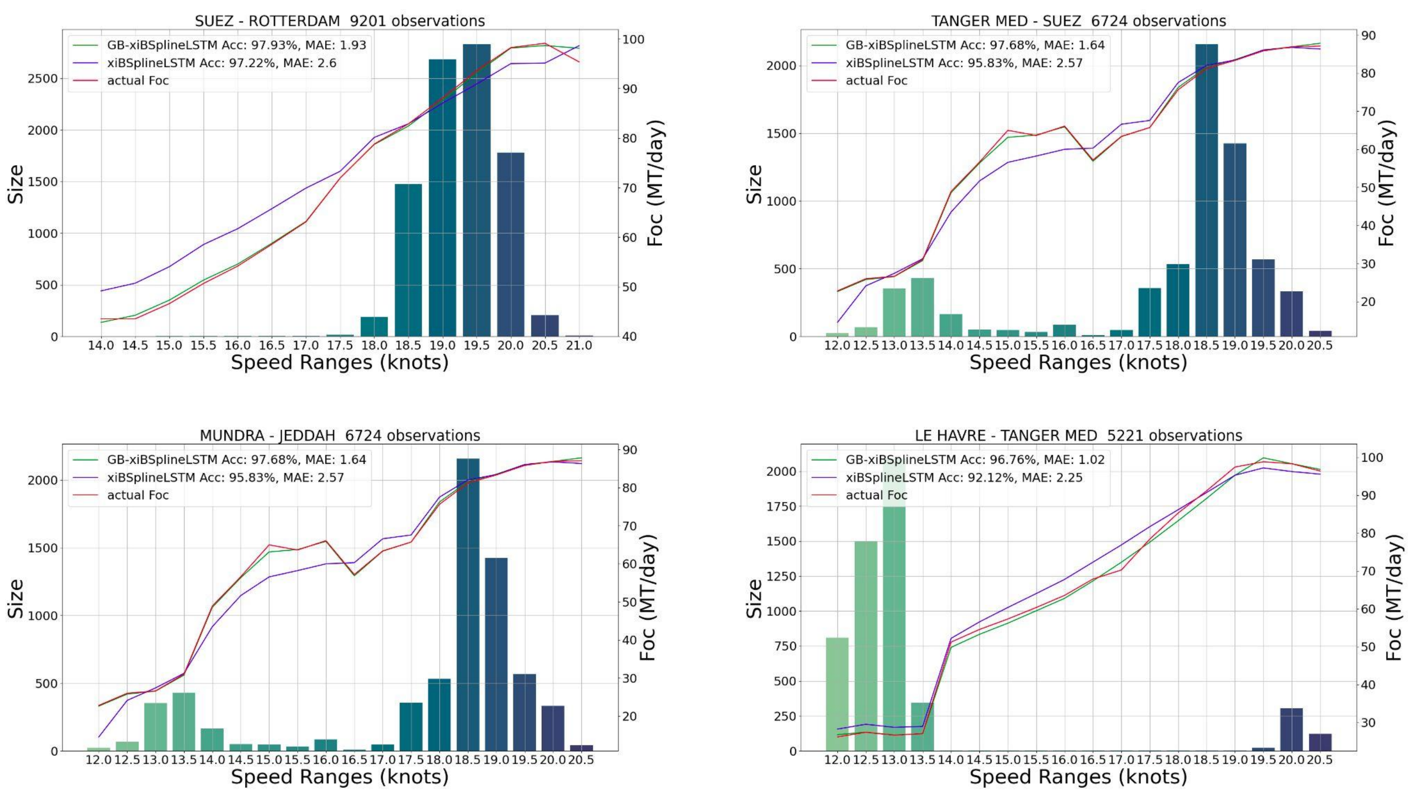

Figure 8: Actual FOC vs SplineLSTM vs GB-SplineLSTM predictions in 4 different voyages

traffic emissions assessment method). Based on the ships' particulars (taken from HIS fairplay database) the model evaluates initially the power consumption, load of the engine, and the fuel consumption of the ship. Based on these values, STEAM eventually evaluates the emissions of NO_x, SO_x, CO, CO_2, as a function of time and location. Towards this direction it was attempted to calculate FOC according to the STEAM model, where engine loads during voyages can be determined with reasonable accuracy based on the ratio of ship speed and the calculated resistance that the ship is required to overcome at a specified speed. In the graph below (Fig. 9), we depict the Specific Fuel Oil Consumption (SFOC) curve (*green line*) for the particular vessel that was utilized to approximate the total FOC and therefore the emissions on different engine loads and speed ranges. A second order polynomial (*blue line*) was "fitted" in order to be able to interpolate between arbitrary M/E Loads and the corresponding SFOC values.

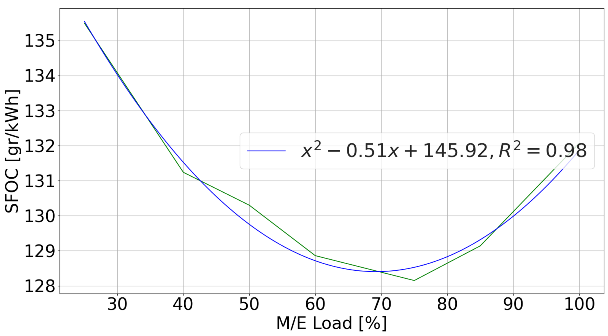

Figure 9: SFOC - M/E Load curve for the container-ship test vessel

The mass flow rates of NO_x, SO, CO_2, are calculated based on engine power and SFOC, accounting for the combustion stoichiometry and NO_x chemistry, as follows:

$$m = P * SFOC * EF = FOC * EF \quad (11)$$

where m is the emissions mass flow rate (in $grams/hour$), P (engine power) is provided in kW, SFOC is provided in gr/kWh and EF is the Emissions Factor presented in Table 3[2]. EF calculation is thoroughly described in Appendix A.

Table 3: Emissions factor (EF) used for calculating CO_2, NO_x and SO_x

CO_2	3.114 (tn CO_2/ tn fuel)
NO_x	0.092 (tn NO_x / tn fuel)
SO_x	2.023 x S mass fuel fraction in fuel (tn SO_x/ tn S in fuel)

In Figure 10 we depict the emissions CO_2 $(tn\, CO_2)/(tn\, Fuel)$ measured from the HORIBA instrument installed in the exhaust gas system of the vessel as well as the calculated emissions with STEAM and GB-SplineLSTM approaches, during a voyage. Mean Absolute Error (MAE) between the actual and the predicted emissions utilizing STEAM and GB-SplineLSTM approach is also demonstrated.

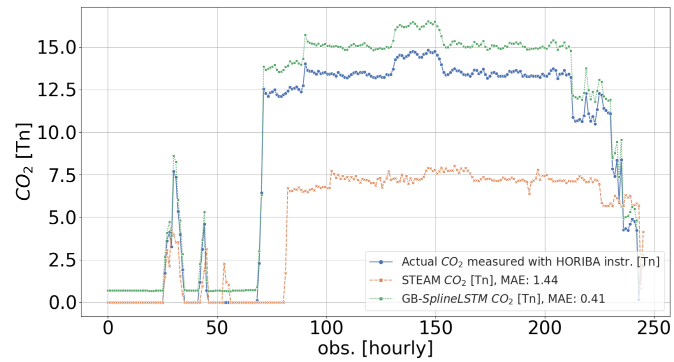

Figure 10: Actual vs STEAM vs GB-SplineLSTM CO_2 emissions

It is evident from the graph (Fig: 10) depicted above that the physics informed neural network (*GB-SplineLSTM*) trained on the tailor-made feature set, accounting for the operational as well as the weather state of the en-route vessel, is able to project CO_2 emissions in a much more accurate manner than the STEAM methodology (0.41 MAE on ~250 hourly observations (10 days), in contrast with 1.44 MAE for STEAM). Representative examples of the computational superiority of the *GB-SplineLSTM* versus the STEAM method are also illustrated in Table 2 where its performance is demonstrated through four different voyages involving 28k observations (during 20 days). *GB-SplineLSTM* achieved ~97% mean accuracy in FOC estimation, alternatively an overall 1.46 Metric Tonnes of fuel discrepancy on average between the predicted and the actual FOC measured by the flowmeter.

CONCLUSIONS

In the context of this work we demonstrated a prototype of a DT framework that aims to facilitate the transition of the maritime industry towards a sustainable, low emission and high efficiency ship operation. This vision is realized through innovative ways by encapsulating both Model Deployment and Decision Support Systems (DSS) in a shared, distributed, centralized data space KH (Knowledge Hub). The tools and services employed in the KH were utilized accordingly to offer a comprehensive approach to incorporate, analyze and further exploit the vast amount of data collected in real time from the vessels in order to build robust CO_2 predictors. These models assess the environmental footprint of the vessel offering to the shipowner an accurate insight regarding CII (Carbon Intensity Index), as well as the possibility to act in a preventive way by pertaining a tailor-made mitigation strategy to lower carbon dioxide emissions for a particular vessel.

[2] Emission factors for HSFO fuel type derived from the IMO Resolution MEPC 245 (66) 2014 'Guidelines on the method of calculation of the attained Energy Efficiency Index (EEDI) for new ships'

The approximation method introduced in this paper for FOC - CO_2 estimation constitutes a novel approach that incorporates standard marine engineering knowledge, into the architecture of a deep learning model. In the experimental section of the paper we showcased the approximation capabilities of this method (*GB-SplineLSTM)* in a series of different voyages conducted by the same container ship vessel. Finally, utilizing the well known and established STEAM methodology as a baseline we further validated and enhanced the predictions acquired from the proposed model, by comparing STEAM and *GB-SplineLSTM* projections with the actual emissions measured by a research team, on-board a container ship, in the context of an experiment conducted for the EU project EMERGE.

One of the main directions to expand this work is to employ a multi-objective WR (Weather Routing) algorithm utilizing the proposed digital ecosystem for simulation, reasoning and control-actuation purposes. Furthermore exploiting the distributed, streamlined architecture of the envisaged platform we aim to collect and eventually store a broader category of variables composed of vessel's particulars, building a physics informed library for different types of vessels. This dataset can then be employed in conjunction with incremental and transfer learning algorithms between different fleets of vessels, methods that aspire to tackle one of the main obstacles maritime industries are facing nowadays, that is the lack of historical data for many vessels.

APPENDIX

A. Emissions factors calculation

The emission factor, F^{CO_2} for carbon dioxide (CO_2), depends on the molecular weight of carbon dioxide m(CO_2), the molecular weight of carbon $m(C)$, the fuel carbon concentration fcc and is given with the following formula:

$$F^{CO_2} = \frac{m(CO_2) * fcc}{m(C)}$$

ACKNOWLEDGEMENTS

This Publication was supported partially by the program of Industrial Scholarships of Stavros Niarchos Foundation, and partially supported by the EMERGE project, an European Union's Horizon 2020 research and Innovation programme under the Grant Agreement No 874990 and by the DT4GS project an European Union's Horizon 2020 research and Innovation programme under the Grant Agreement No 101056799. Access to industrial data has been provided by Danaos Shipping Co.

REFERENCES

Bohlin, T. 1991. "Interactive System Identification: Prospects and Pitfalls." *Springer-Verlag*.

Corradu, Andrea, Luca Oneto, Frnasesco Baldi, and Davide Anguita. 2017. "Vessels fuel consumption forecast and trim optimisation: a data analytics perspective." *Ocean Engineering* 130:351--370.

Friedman, Jerome. 1991. "Multivariate adaptive regression splines." *The annals of statistics*, 1-67.

Gkerekos, Christos, and Iraklis Lazakis. 2020. "A novel, data-driven heuristic framework for vessel weather routing." *Ocean Engineering* 197:106887.

GuldHammer, and Harvald. 1965. "Ship Resistance."

Hollenbach, K.U. 1998. "Estimating resistance and propulsions for single screw and twin screw ships." *Ship Technology Research*, 72-76.

Jalkanen, J-P, Lasse Johansson, Kukkonen Jaakko, Brink Anders, Kalli Juha, and Stipa Tapani. 2012. "Extension of an assessment model of ship traffic exhaust emissions for particulate matter and carbon monoxide." *Atmospheric Chemistry and Physics* 12:2651-2659.

Kaklis, Dimitrios, Pavlos Eirinakis, George Giannakopoulos, Constantine Spyropoulos, Takis J. Varelas, and Iraklis Varlamis. 2022a. "A big data approach for Fuel Oil Consumption estimation in the maritime industry." *022 IEEE Eighth International Conference on Big Data Computing Service and Applications (BigDataService)*, 39-47.

Kaklis, Dimitrios, George Giannakopoulos, Iraklis Varlamis, Constantine Spyropoulos, and Takis J. Varelas. 2019.

"A data mining approach for predicting main-engine rotational speed from vessel-data measurements." *IDEAS '19: Proceedings of the 23rd International Database Applications \& Engineering Symposium* 23:1-10.

Kaklis, Dimitrios, George Giannakopoulos, Iraklis Varlamis, Constantine Spyropoulos, and Takis J. Varelas. 2022a. "Online Training for Fuel Oil Consumption Estimation: A Data Driven Approach." *23rd IEEE International Conference on Mobile Data Management (MDM)*, 12-22.

Miyeon, Jeon, Noh Yoojeong, Shin Yongwoo, Lim O-Kaung, Lee Inwon, and Cho Daeseung. 2018. "Prediction of ship fuel consumption by using an artificial neural network." *Journal of Mechanical Science and Technology* 32:5785-6796.

Moussiopoulos, Nicolas, Sokhi Ranjeet, George Tsegas, Evangelia Fragkou, Eleftherios Chourdakis, and Ioannis Pipilis. 2020. "Environmental impacts of shipping: from global to local scales." *12th International Conference on Air Quality, Science and Application*.

Namini, Siami, Sima Tavakoli, Neda Namin, and AKbar Siami. 2018. "A comparison of ARIMA and LSTM in forecasting time series." *2018 17th IEEE ICMLA*, 1294-1401.

Papandreou, Christos, and Apostolos Ziakopoulos. 2020. "Predicting VLCC fuel consumption with machine learning using operationally available sensor dat." *Ocean Engineering* 197:106887.

Savitha, R., and Al Mamun Abdullah. 2017. "Regional ocean wave height prediction using sequential learning neural networks." *Ocean Engineering* 129:605-612.

Shengzeng, Wang, Ji Baoxian, Zhao Jiansen, Liu Wei, and Xu Tie. 2018. "Predicting ship fuel consumption based on LASSO regression." *Transportation Research Part D: Transport and Environment* 65:817--824.

Telfer, E.V. 1963. "he design presentation of ship model resistance data."

Yongjie, Zhu, Yi Zuo, and Tieshan Li. 2020. "Predicting Ship Fuel Consumption based on LSTM Neural Network." *2020 7th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS* 32:310--313.