

Regression Modelling Estimation of Marine Diesel Generator Fuel Consumption and Emissions

Onur Yuksel^a, Burak Koseoglu^b

This study aims to estimate the fuel consumption of marine diesel generators onboard. Objective technical specifications and operational data on the ship's power generating plants and port calls were collected from an oceangoing oil/chemical tanker and used to develop the mathematical model of the plant in the Python and MATLAB environment. The model consists of alternators, prime movers and load distributions of the ship's power generating plant and provides information on fuel consumption in metric tons calculated based on hours of operation and specific fuel consumption data. Regression models have helped predict future fuel consumption for the plant and the optimal model for the dataset was identified by comparing four

different algorithms. As the results have shown the Ordinary Least Squares Regression to be optimum, it was used to make one, five, and ten-year predictions. The predictions for one-year, five-year, and ten-year periods are 4,322,436, 10,684,860, and 18,615,472 t respectively. The selected model predicts fuel consumption with R² of 0.999, MAE of 3.932, and RMSE of 2.935. Fuel consumption predictions facilitated plant emission calculation.

KEY WORDS

- ~ Marine diesel engines
- ~ Mathematical modelling
- ~ Linear regression
- ~ Support vector regression
- ~ Artificial neural networks
- ~ Time series analysis
- ~ Ship emissions

a. Dokuz Eylul University, Graduate School of Natural and Applied Sciences, Izmir, Turkey

e-mail: onur.yuksel@deu.edu.tr

b. Dokuz Eylul University, Maritime Faculty, Maritime Transportation Engineering, Izmir, Turkey

doi: 10.7225/toms.v11.n01.w08

This work is licensed under 

Received on: Sep 10, 2021 / Revised on: Dec 12, 2021; Accepted on: Jan 6, 2022;

Published online: Feb 19, 2022

1. INTRODUCTION

Trustworthy ship system fuel consumption (FC) estimation is relevant in economic and environmental terms (Prpić-Oršić and Faltinsen, 2012). The understanding of fuel consumption dependencies increases the management efficiency of ships. FC systems can be optimized by using predictions. Future estimations are also helpful for the assessment of vessel and system conformity with the regulations of the International Maritime Organization (IMO). IMO regulations in maritime activities aim to lower shipping-related emissions since maritime transportation is a major contributor to air pollution (IMO, 2014). In addition to setting emission limits, the IMO aims to improve marine vessels efficiency by using methods such as Energy Efficiency Design Index (EEDI) and Ship Energy Efficiency Management Plan (SEEMP). Since emissions depend on the quantity of fossil fuels consumed by ship systems, FC optimization is becoming increasingly important. Ship-owners use fuel estimations to comply with regulations and improve the operational efficiency of their vessels (Eide et. al., 2011; Uyanik et. al., 2020).

Marine diesel generators power the ship by using energy obtained from fossil fuel combustion. They consist of a diesel engine as the prime mover and a synchronous alternator for three-phase electricity generation. Power generating plants may

have more than one marine diesel generator, depending on the ship's electrical load. The number of marine diesel generators in a plant can vary depending on the type of operation (McGeorge, 1995). Marine diesel generators are continuous fuel consumers that considerably contribute to air pollution. Ship electrical needs are particularly high during cargo transfer operations in ports, resulting in the release of pollutants (Styhre et al., 2017). Real time assessment of marine diesel generator FC can indicate emission quantities and encourage improvements.

Some studies conducted in the last decade have dealt with FC calculation and estimation onboard ships. For instance, Kesgin and Vardar (2001) computed emissions from ships in Istanbul and Canakkale straits using automated identification system (AIS) data. Miola and Ciuffo (2011) proposed an alternative approach to ship pollutant prediction and assessed the reliability of current techniques. Prpić-Oršić and Faltinsen (2012) proposed a method of estimation of ship speed loss and related CO₂ emissions for a container ship on the North Atlantic route. Winnes et al. (2015) computed GHG emissions using AIS and ship technical data from ships, then analyzed GHG emission reduction strategies for ships in port areas. They analyzed three scenarios, namely "Alternative Fuel", "Ship Design" and "Operation" at the Port of Gothenburg. Tichavska and Tovar (2015) built a model based on AIS data and the Ship Traffic Emission Assessment Model (STEAM). Their model calculated the emissions from cruisers and ferries in the Port of Las Palmas. Bialystocki and Konovessis (2016) proposed a statistical approach to predicting a ship's FC and speed curves. BalBesikci et al. (2016) constructed an artificial neural network (ANN) to analyze the relationship between engine revolutions per minute (rpm) and outside factors using the noon report data. Chang (2016) analyzed the relationship between carbon emission production and deadweight tonnage of shipping transportation. Leloup et al., (2016) built a ship propulsion system model involving kites and conducted a fuel estimation for the system. Styhre et al. (2017) developed a model that estimates GHG emissions from ship operations from different ports. Simonsen et al. (2018), constructed a model that predicts FC and energy usage of cruise vessels using AIS data and technical specifications of ships. Wang, et. al. (2018) estimated ship FC using Least Absolute Shrinkage and Selection Operator (LASSO) regression. Yang et al. (2019) built a genetic algorithm-based grey-box model to estimate a ship's FC using crude oil tanker's operational data. Le et al. (2020) presented an ANN model to predict the FC of container ships in Korea. Liu and Duru (2020) proposed a probabilistic Bayesian prediction algorithm that forecasts ship emissions based on ship movements gathered from AIS data. Le et al. (2020) conducted a study that predicted the FC of ocean-going container ships using a regression model. The data were acquired from a large container terminal in Korea. Farag and Ölçer (2020) used ANN and multi-regression methods to predict ship power and FC. Uyanik et. al. (2020) compared the performance of different regression

algorithms with respect to ship fuel consumption assessment. Reis et al. (2020) developed and tested two feature-oriented models to estimate shipping CO₂ emissions using an actual data set from a Ro-Pax ship. Li et. al. (2020) conducted a study that aimed to optimize the exhaust emissions from a marine dual-fuel engine using Response Surface Method. Zhu et. al. (2021) introduced a joint model to estimate the fuel consumption of a passenger ship. Kim et al. (2021) created an ANN model which can estimate the FC of a container ship using operational data. Moreira et al. (2021) used an ANN method to estimate FC and ship speed through the establishment of the correlation between propulsion, weather inputs, ship speed and FC. This overview of literature clearly shows that ANN is a popular strategy owing to its good performance with complex problems. The studies generally focus on and examine propulsion from an environmental standpoint. The objective of this study is to enable the estimation of fuel consumption and emissions of marine diesel generators used on a 29681 GT oil/chemical tanker. The ship's power generating plant is a constant emission producer that has not been studied in detail. The plant runs during both navigation and port operations. Particularly in port operations, the number of working generators increases due to higher load demands, resulting in higher air pollutant emissions. A detailed analysis of the plant can help determine exact pollutant quantities and consequently encourage plant innovations. This study provides an improved mathematical model that includes different plant operation modes and regression analysis based on the data obtained from the model's calculations. The model is based on real-time data and operations. The study is significant because the future estimations given focus on the plant's environmental impact. The remainder of the paper is divided into three sections. Section 2 explains the mathematical background, simulation modelling of a marine diesel generator plant and regression models. Section 3 gives the results, findings and comparison of these models, while Section 4 contains conclusions, discussions, and recommendations.

2. MATHEMATICAL BACKGROUND AND MODEL DESCRIPTION

This section gives the equations used in the ship power plant model, describes the logic and provides validation of the model. Evaluation metrics, regression algorithms and preferred choices are also mentioned in this section.

2.1. Ship Power Plant Simulation

Technical specifications, operational data, electrical load in different operating modes and cargo transfer information have been collected from a ship. Position and port of call data have been obtained from AIS. The data were collected between 06/12/2019

and 07/06/2021, and used to develop a mathematical model of the plant, while the plant's fuel consumption was established in Python and MATLAB in this same period. Various regression models were compared to find the best prediction algorithm for the data set.

The prime mover uses heavy fuel oil (HFO) during navigation and marine gas oil (MGO) during harbor and port operations. Fuel specifications were obtained from ship fuel records. The manufacturer of the engine provided look-up tables that include brake-specific FC data for the prime mover in the numerical model. The torque and power transmitted to the synchronous generator were calculated using the manufacturer's stroke, bore and break mean effective pressure (bmep) data. Eq. 1 is the displacement volume (V_d) calculation, where b is the bore and s piston stroke. Eq. 2 gives the transmitted power (P) calculation. Reference generators had six cylinders (n), while the number of firing strokes (N) was calculated by dividing the number of revolutions by 120. η_t is transmission efficiency, set at 0.99 (Altosole et al., 2019; Altosole and Figari, 2011; Pulkrabeck, 2004).

$$V_d = \pi \cdot \left(\frac{b^2}{4} \right) \cdot s \quad (1)$$

$$P = (V_d \cdot n \cdot bmep \cdot N/1000) \cdot \eta_t \quad (2)$$

Synchronous alternator onboard test trial provided the data for the simulation. Eq. 3 demonstrates the phase voltage calculation using armature current (I_A), armature resistance (R_A), synchronous reactance (X_s) and induced voltage (E_A).

$$V_\phi = E_A - jX_s I_A - R_A I_A \quad (3)$$

Terminal voltage (V_T) equals $\sqrt{3V_\phi}$ for Y-connected phases. Eq. 4 is the calculation of alternator input power (P_{in}). γ is the angle between E_A and I_A . Eq. 5 indicates synchronous generator output power that can be calculated using line or phase variables, where I_L is line voltage. Eq. 6 is the voltage drop (VD) percentage calculation, using no-load voltage (V_{nl}) and full-load voltage (V_{fl}). The higher the synchronous generator load, the lower the rpm of the prime mover. Eq. 7 is the calculation of the prime mover's speed drop (SD) percentage. The generator's no-load speed is n_{nl} , and full-load speed n_{fl} .

$$P_{in} = 3 E_A I_A \cos\gamma \quad (4)$$

$$P_{out} = \sqrt{3} V_T I_L \cos\theta = 3 V_\phi I_A \cos\theta \quad (5)$$

$$VD = (V_{nl} - V_{fl}) / V_{fl} \cdot 100 \quad (6)$$

$$SD = (n_{nl} - n_{fl}) / n_{fl} \cdot 100 \quad (7)$$

Eq. 8 explains the relationship between generator power and frequency. The generator's no-load frequency is f_{nl} , the system's operating frequency is f_{sys} , and the slope of the speed-power curve is s_p expressed in MW/Hz.

$$P_{out} = s_p (f_{nl} - f_{sys}) \quad (8)$$

In a system with three equivalent generators functioning in parallel, the total power is equal to the sum of their power. The total power (P_{tot}) of the system is shown in Eq. 9 (Chapman, 2005; Krause, et al., 2002; Hansen and Michalke, 2008).

$$P_{tot} = P_{load} = P_{out1} + P_{out2} + P_{out3} \quad (9)$$

The ship power plant simulation consists of three main parts. The prime mover, the synchronous generator and the required electrical load. Table 1 provides the essential specifications of these parts. Load test, governor test, parallel running test, open circuit and short-circuit test results have been obtained from alternator manufacturer's manuals and ship electrical equipment tests conducted before installation. FC and other required data for the prime mover model have been obtained from instructions and manuals of the diesel engine manufacturer. Data on electrical load in different operation modes have been obtained from electrical load tests conducted during ship trials. Figure 1 is a basic sketch of the ship's electrical distribution system.

Table 1.

General information on marine diesel generators and electrical load.

Prime Mover		Alternator	
Engine/Generator MCR	960/900 kW	Output Capacity	1312.5 KVA
Rpm	900	Voltage	450 V
Cylinder Bore/ Piston Stroke (mm)	210/320	Frequency	60 Hz
Swept Volume per Cylinder (dm ³)	11.1	Poles	8
Piston Mean Speed (m/s)	9.6	Voltage Variation	2.50 %
Compression Ratio	17:1	Voltage Adjust	5 %
Mean Effective Pressure (bar)	24.1	Power Factor	0.8 Lagging
Electrical Load			
Operational Mode	Number	Load (kW)	Current (A)
Sea Going	1	850.6	1364.152
Sea Going with Tank Heating	2	1195.7	1917.607
At Port	3	1122	1799.410
At Cargo Handling	4	1781.5	2857.085
At Harbor	5	571.3	916.224
At Harbor with Tank Heating	6	863	1384.038

The model compares the active power of generators, computed in the synchronous generator section, with load and current to determine the number of generators required for the given operating mode. Then, it ensures the operation of the ship's synchronous generators in a predetermined sequence, depending on the number of generators. Generator operation sequence can be determined at the beginning of the simulation and is initially set as 3, 1, 2. The generator sequence is updated in maintenance intervals obtained from the operation manual. In other words, the model is adjusted to ensure the overhauled generator is the last in the sequence. If the generator is in operation, the model fills the binary array with value 1, otherwise with 0. The program next examines the control array to calculate the frequency, speed drop and slope of the running generators' power curve. Then it computes system frequency, power, load, and current requirements for each generator, utilizing the control array and a number of generator data. The computed load of each generator, power of the prime mover, armature currents, internally generated voltages, line voltages, voltage regulation percentages and synchronous power output help determine

brake-specific FC under changing load conditions. FC and emissions produced by the generator, both in kg per hour, are calculated using brake-specific FC. Operation types and times have been established based on port call and operational records obtained from the ship. Therefore, the simulation computes total FC and emissions in metric tons. Figure 2 illustrates the logic of the marine diesel generator's mathematical model.

Emission calculations are based on fuel consumption computed by the model and emission factors taken over from Kuzu et al. and 2021; Trozzi, 2010. Eq.10 is the emission calculation and Table 2 shows marine gas oil (MGO) and heavy fuel oil (HFO) emission factors.

$$E_i = \sum_k FC \cdot EF \quad (10)$$

where E is total emission, EF emission factor, k operation mode and i pollutant (Kuzu et al., 2021).

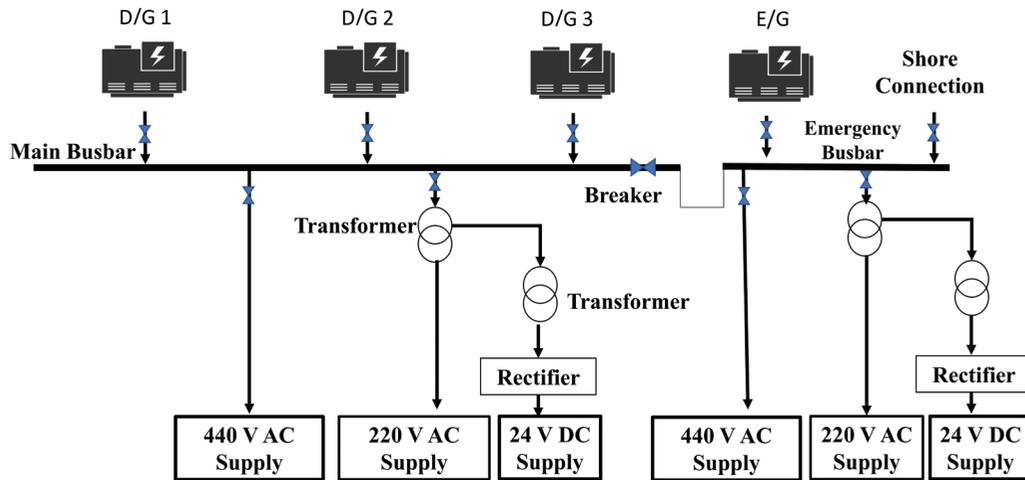


Figure 1. Ship power distribution system (Adapted from Cahyagi and Koenhardono, 2018, Oways, 2019).

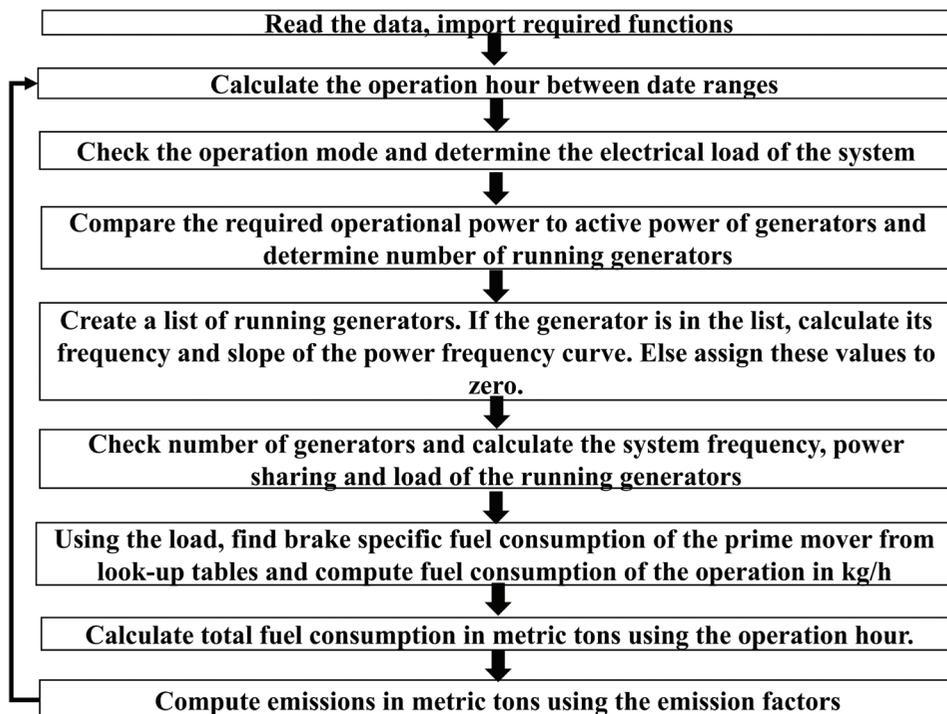


Figure 2. The algorithm scheme of the mathematical model.

Table 2.

EF of MGO (g emission/ g fuel) (Kuzu et al., 2021; Trozzi, 2010; IMO, 2014).

Pollutant	CO ₂	CH ₄	N ₂ O	NO _x	NM VOC	CO	PM ₁₀	SO ₂
EF MGO	3.206	0.00006	0.00015	0.0961	0.00308	0.00277	0.00097	0.01
EF HFO	3.114	0.00006	0.00015	0.0903	0.00308	0.00277	0.00278	0.025

2.2. Performance Evaluation Model

Error rate is the correlation between the actual value and the model's output that can be used to assess the model's performance. The validity and performance of the marine diesel generator model have been assessed using mean absolute error (MAE) as a relevant, understandable and reliable method (Chai & Drexler, 2014; Willmott & Matsuura, 2005). MAE is a metric that evaluates the discrepancy between real and anticipated values without taking direction into account. Therefore, lower MAE indicates better prediction performance (Chai & Drexler, 2014). Regression algorithms have been evaluated using MAE, correlation score (R^2) and root mean squared error (RMSE). The comparison of three evaluation metrics improves assessment reliability to avoid overfitting. R^2 is a well-known performance measurement metric that evaluates the strength of a relation in regression. It ranges from 0 to 1, with higher R^2 being indicative of better performance (Kasuya, 2019). RMSE is used to validate MAE for regression performance measurement purposes. Eq. 11, Eq. 12, and Eq.13 are the equations for MAE, RMSE, and R^2 , respectively.

$$MAE = \frac{\sum_{i=1}^n |y_p - y_t|}{n} \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_p - y_t)^2}{n}} \quad (12)$$

$$R^2 = \frac{\sum_{i=1}^n (y_t - \bar{y})^2}{\sum_{i=1}^n (y_p - y_t)^2} \quad (13)$$

where n is the total number of data, y_p calculated or predicted value, y_t real or true value and \bar{y} mean of the data. (Chai & Drexler, 2014; Kasuya, 2019).

2.3. Ship Power Plant Simulation Validation

Various performance measures can be used to assess model validation. These performance criteria have been derived both from real data obtained from a ship, and determined using the mathematical model. As a result, performance was tested by comparing calculated and real values. Figures 3. a, 3. b, and 3. c illustrate power-sharing based on total system load, as determined by parallel running tests on each generator. Generator 1 (Fig 3.a), Generator 2 (Fig 3.b), and Generator 3 (Fig 3.c) have MAEs of 0.0168, 0.0095, 0.0102 respectively for this indicator, all of which are small error rates. For generators 1, 2 and 3, Fig 4. a, Fig 4. b and Figure 4. c show the comparison between frequency decrease due to increasing load measured by governor tests and the frequency computed by the model. The MAEs of frequency calculations for generators 1, 2 and 3 are 0.0034, 0.0039, and 0.004, which is within acceptable limits for this assessment. Fig 4. d illustrates system frequency decrease depending on varying total system load. The parallel running test in ship trials starts with a stable system frequency of 60 Hz. Since the model initially ignores stabilization mechanisms, the system frequency computation error rate is slightly higher, with MAE of 0.0099. Fig 5. compares real and calculated power of synchronous generators as armature current increases. This benchmarking has the MAE of 0.000325, which is an acceptable error rate. Fig 5. b is a comparison between calculated and actual line voltages in the function of increasing armature current. As MAE in this analogy is 0.0035, model line voltage prediction can be assumed to be acceptable.

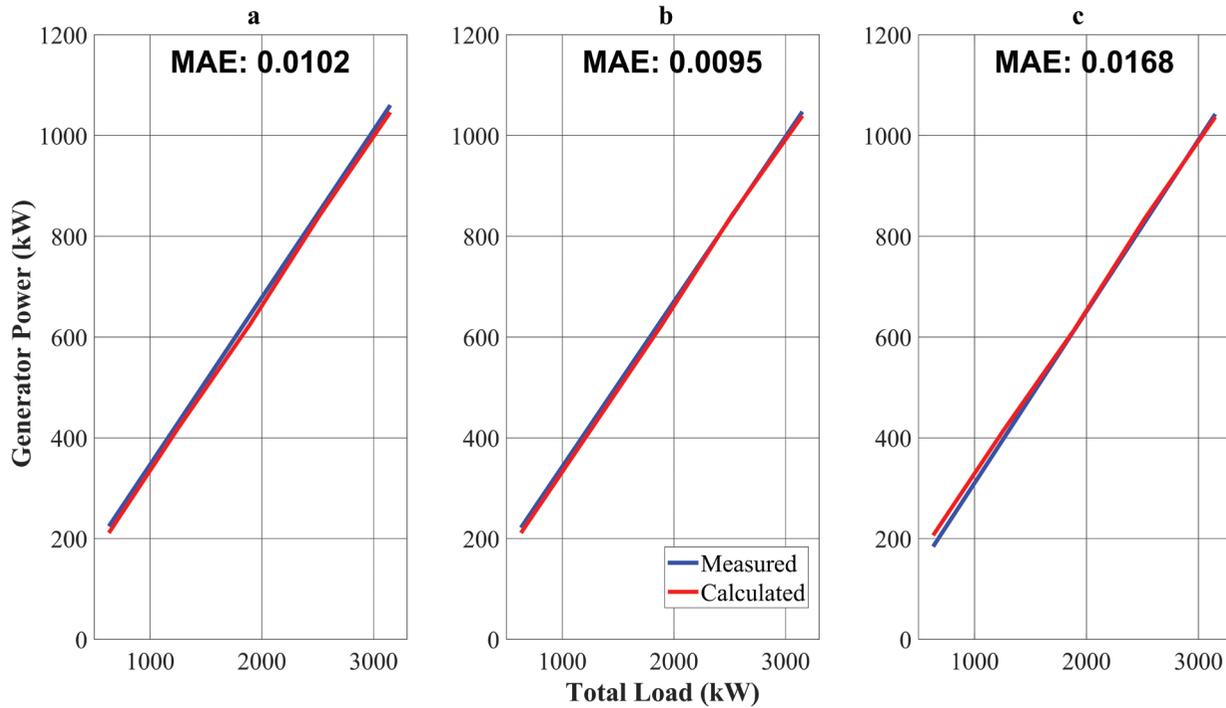


Figure 3.

Comparison of measured and calculated power in correlation with varying system loads (a) generator 1 (b) generator 2 (c) generator 3.

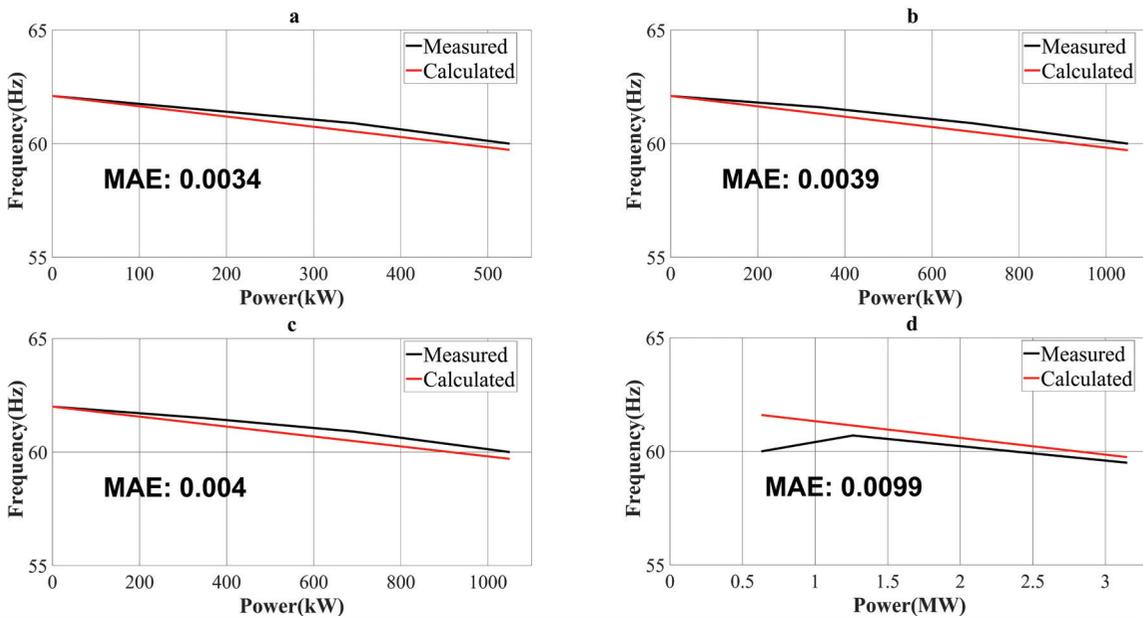


Figure 4.

Validation analogy between measured and calculated frequency in correlation with increasing generator power (a) generator 1 (b) generator 2 (c) generator 3 (d) system.

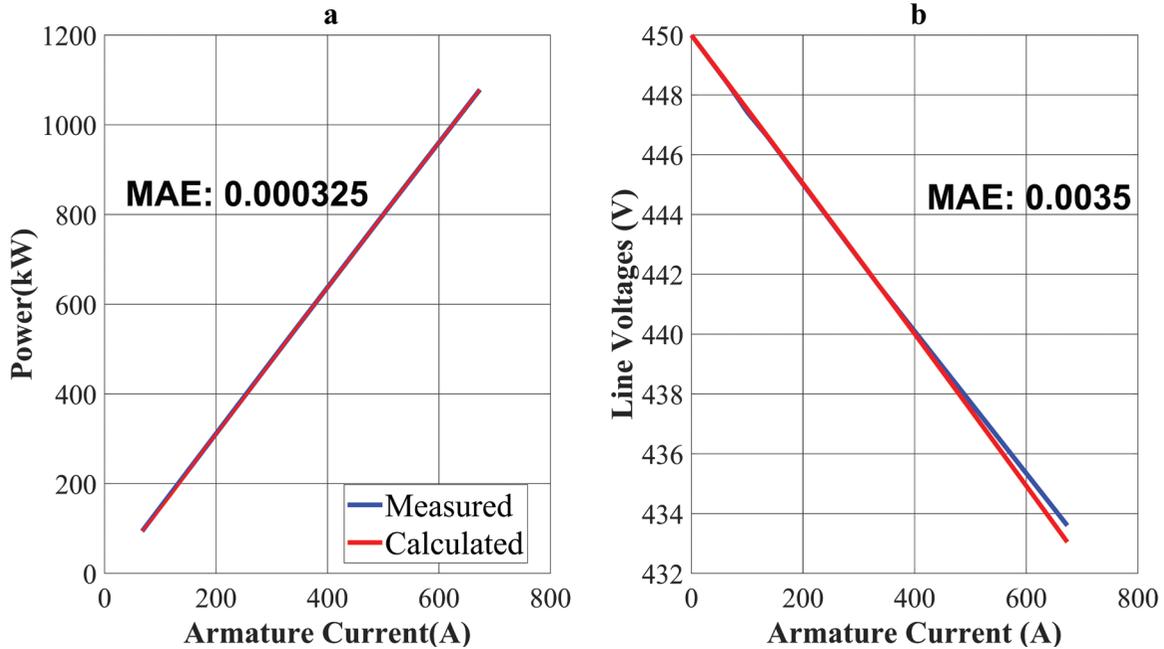


Figure 5. Validation analogy between measured and calculated (a) power and (b) line voltage as armature current increases.

2.4. Fuel Consumption Prediction

The mathematical model computed the plant's FC in the designated period. For future estimations, some regression techniques have been trained based on the data. However, the data vary depending on the working hours of the generator. Data standardization is ensured to improve performance and reduce computation time for complicated methods. Standardization is a scaling method that provides the mean of zero and the standard deviation of one by first subtracting each data point (x) from the mean of data (\bar{x}) and then dividing the standard deviation (σ). Eq. 14 explains the process (Trebuña et. al., 2014). Transformed and scaled data have 13268 rows and are split into 80 % train and 20 % test sets. The data split configuration is obtained from trials of various split combinations. According to results, the 80-20 split not only provides the optimum amount of data to test the model, but also a large quantity of training data. Regression models are trained with training data evaluated with test data using evaluation metrics. The optimum model is identified and future fuel consumption predictions ensured. Fig 6. illustrates the FC estimation process.

$$z = \frac{x - \bar{x}}{\sigma} \quad (14)$$

Techniques suitable for plant FC estimation are as follows:

- Ordinary least squares (OLS) regression
- Auto-regressive integrated moving average (ARIMA)
- Support vector regression (SVR)
- ANN

OLS is a type of generalized linear regression that can be used to model a single response variable that has been recorded on a scale of at least one gap (Craven and Islam, 2011). OLS deals with the correlation between dependent and independent variables. The value of the dependent parameter is obtained using the linear correlation between independent parameters plus error rate (e). Eq. 15 shows the OLS formula.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_k X_k + e \quad (15)$$

where β is the regression coefficient, X independent variable and y estimated value (Pohlman and Leitner 2003).

ARIMA is a well-known statistical approach to predicting time-related data. ARIMA(p, d, q) can be defined as a linear combination of past values of y_t and e . Eq. 16 shows the formulation of the ARIMA.

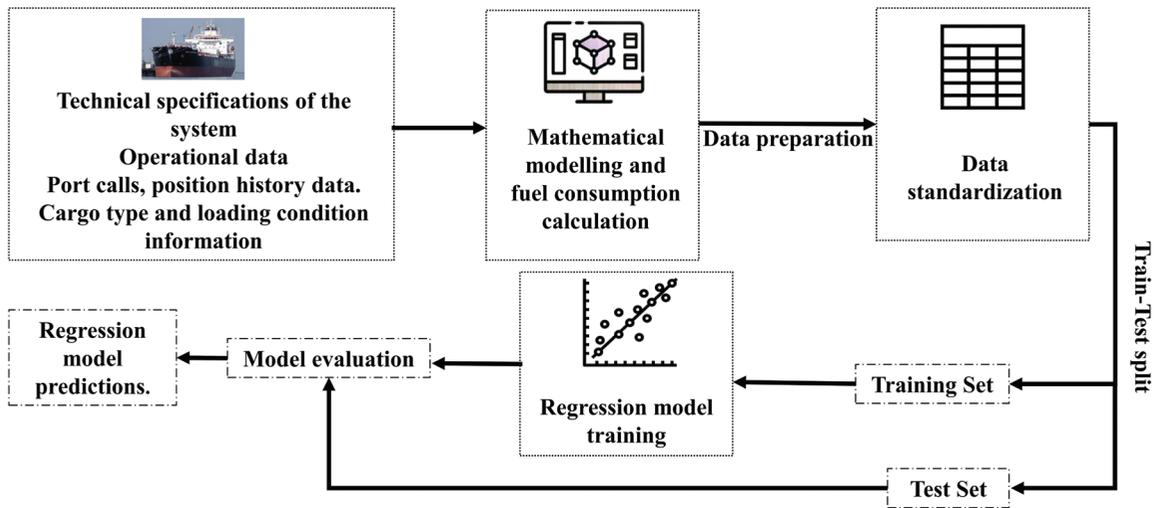


Figure 6.
FC estimation process.

$$y_t = \varphi_0 + a_1 \cdot y_{t-1} - m_1 \cdot e_{t-1} + a_2 \cdot y_{t-2} - m_2 \cdot e_{t-2} + \dots + a_p \cdot y_{t-p} - m_p \cdot e_{t-p} + e_t \quad (16)$$

$$\min_w \frac{1}{2} \|w\|^2 \quad (17)$$

where p , d , and q are ARIMA orders, a and m estimated regression weights, and φ_0 trend member (Mills, 2019). Optimal model orders for each data set are identified by computing Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) using an iterative algorithm. The algorithm iterated every combination of p , q , d within the specified limits. The p and q range was 1-5, while d range was 1-3. These limits depend on autocorrelation and partial autocorrelation plots. Algorithm results suggest that ARIMA (0,2,2) is the ideal model for the data.

SVR determines the acceptable error in the model, as well as the optimal data fitting line. The purpose of the objective function is to minimize coefficients and constraints. The approximation of magnitude of the normal vector ($\|w\|$) is given in Eq. 17 (Awad and Khanna, 2015)

ANN is one of the deep learning algorithms and is widely used in both classification and regression problems. The Python trial results helped identify the number of layers, neurons, activation function, batch size, epochs and optimizer. Rectified linear unit (Relu) performs best with the data compared to other activation which can be defined as $y = \max(0, x)$. Since the data are free of complex patterns or large bias, more basic activation functions like Relu can be applied and reduce the ANN training time (Dreyfus, 2005). The Adam optimizer is used for network creation due to its computational efficiency, easy implementation and suitability for large datasets (Kingma and Ba, 2014). Epochs are 50 and batch size is 10 in the model. Figure 7 depicts ANN structure.

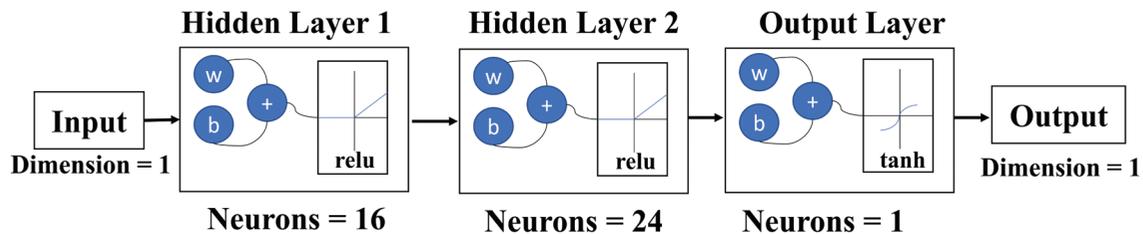


Figure 7.
ANN structure.

3. RESULTS AND FINDINGS

The mathematical model calculated the FC of the marine diesel generator plant in different ship operation modes between December 6, 2019 and June 7, 2021. Fig 8 illustrates total FC and FC by operation modes in metric tons. Air pollutants from the operations were calculated using FC and emission factors shown in Table 2. Table 3 gives emission production in each operation mode in metric tons. 65 % of plant operations occurred during

navigation and the rest were in harbor and port areas since the ship's routes in the period examined involved many oceangoing voyages. Even though cargo handling operations require the highest electrical load, they accounted for only 4 % of utilization time and had the lowest fuel consumption of all operations. Fig 9 illustrates the plant's utilization time distribution by mode of operation. The plant produces 8,592.38 t of CO₂, which is remarkable compared to other pollutants. NO_x and SO₂ are the second and third highest generated pollutants, respectively.

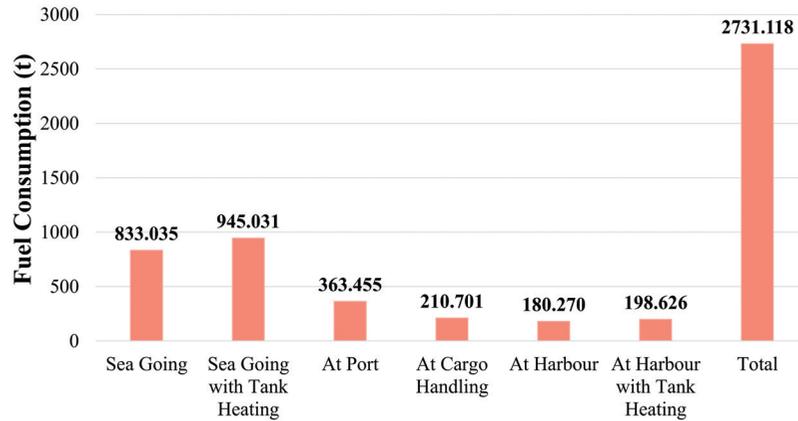


Figure 8.
Total generator FC by operation mode.

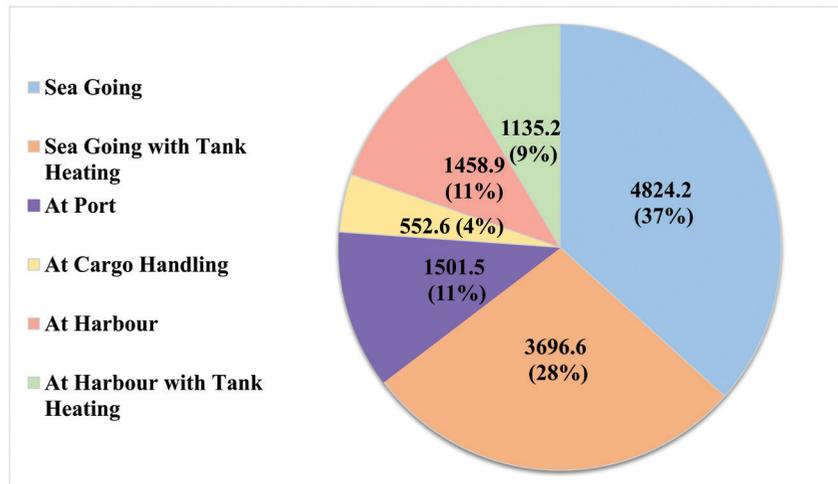


Figure 9.
Plant utilization time distribution by operation mode.

Table 3.

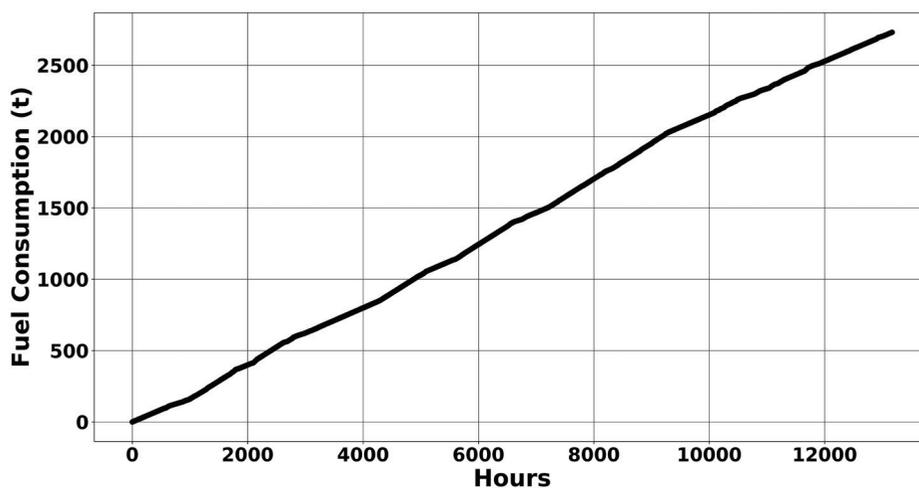
Emissions from marine diesel generators by operation mode.

Operation	CO ₂	CH ₄	N ₂ O	NO _x	NMVOC	CO	PM ₁₀	SO ₂
1	2,594.07	0.05	0.12	75.22	2.57	2.31	2.32	20.83
2	2,942.83	0.06	0.14	85.34	2.91	2.62	2.63	23.63
3	1,165.24	0.02	0.05	34.93	1.12	1.01	0.35	3.63
4	675.51	0.01	0.03	20.25	0.65	0.58	0.20	2.11
5	577.95	0.01	0.03	17.32	0.56	0.50	0.17	1.80
6	636.79	0.01	0.03	19.09	0.61	0.55	0.19	1.99
Total	8,592.38	0.16	0.41	252.15	8.41	7.57	5.87	53.98

The simulation computed FC by operation mode and total operating hours. The data were transformed hourly to prepare them for regression analysis. Fig 10 shows hourly FC increments of the plant over the period of 13,268 hours. The figure demonstrates that there are linear relationships between the data and that the models' R² scores are high as expected. Four regression models, from basic to complex, were adjusted to the data. An attempt was made to use other linear regression algorithms, such as Ridge, Lasso and Elastic Net, with their parameter optimization ending with OLS. Figure 11, 12, 13, and 14 are a comparison between actual (test data) and predicted values, evaluation

metrics and one-year predictions of each regression algorithm. As anticipated, R² scores were higher for each model, with OLS achieving the highest R² of 0.9992. OLS had the lowest MAE and RMSE metrics of 3.932 and 2.935 respectively.

Figure 11 shows that OLS has the best curve overlap of all the methods applied. Lower MAE and RMSE values also support this. However, in spite of OLS being the fastest and the most straightforward method, towards the end, the predicted value line is higher than the actual value line, i.e. predicted values obtained are higher than those obtained by other techniques.

**Figure 10.**

FC of the plant by working hours.

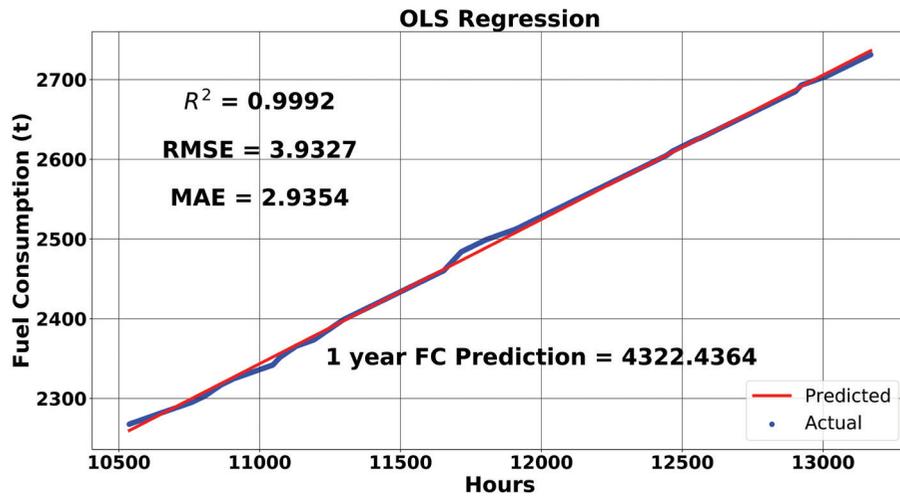


Figure 11.

Comparison of test data and predicted values, evaluation metrics, and one-year prediction using OLS.

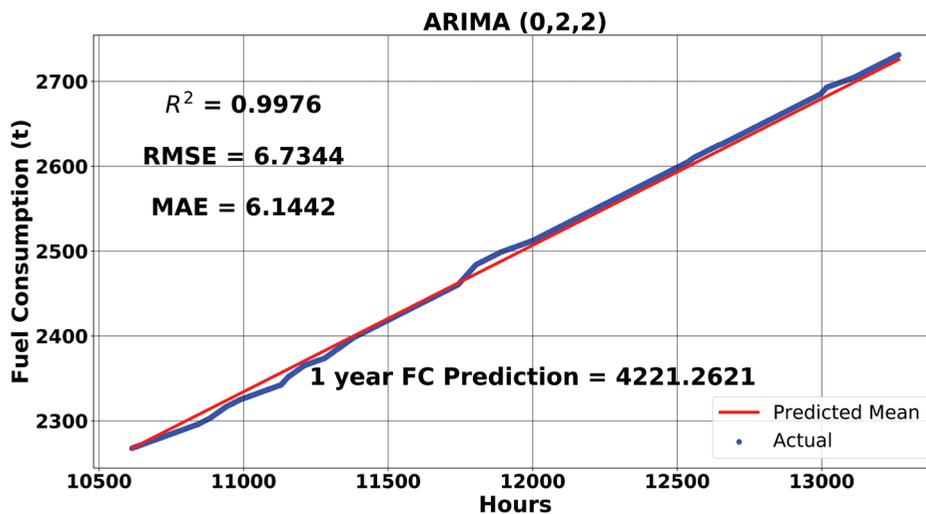


Figure 12.

Comparison of test data and predicted values, evaluation metrics, and one-year prediction using ARIMA (0, 2, 2).

A similar phenomenon can be seen in Figure 12 for ARIMA (0,2,2). In spite of having the highest MAE, successful curve overlap can be observed. Nevertheless, in the second part of the data, its predicted value curve is lower than the actual value curve and thus yields lower future forecasts. In addition, ARIMA's running time is the second-highest after ANN.

SVR's MAE and RMSE are particularly higher, especially at the beginning of the data set. In the second part of the data set,

SVR gives a perfect overlap which seemingly suggests better predictions, however, MAE and RMSE values obtained are slightly higher compared to other regression techniques.

Even though it is a complex methodology, the computation time of the SVR model dramatically drops with data standardization.

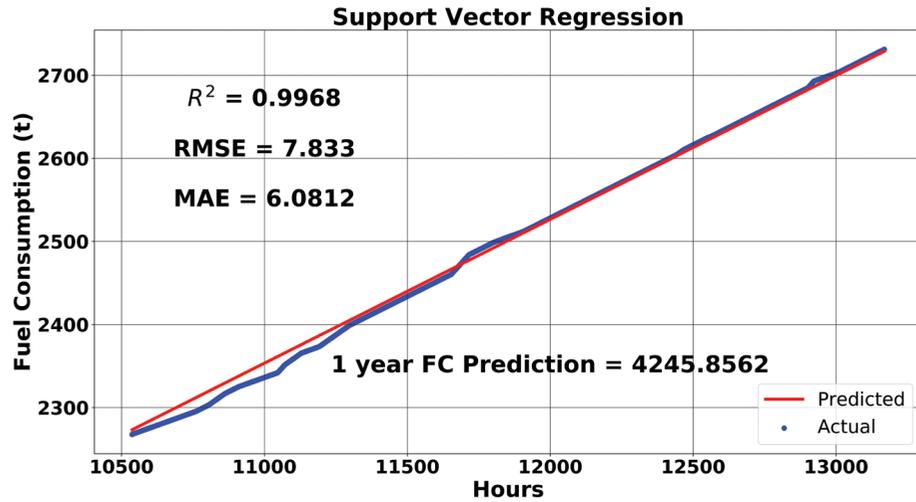


Figure 13.

Comparison of test data and predicted values, evaluation metrics, and one-year prediction using SVR.

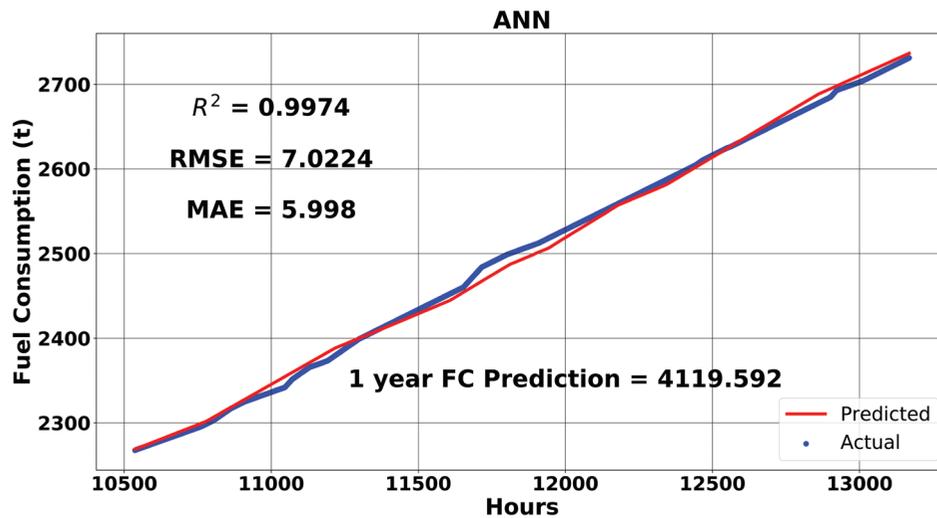


Figure 14.

Comparison of test data and predicted values, evaluation metrics, and one-year prediction using ANN.

ANN, on the other hand, has a high R^2 score, and its future predictions and other evaluation metrics are decent. Figure 14 shows the curve overlap, metrics and predictions obtained with ANN. ANN seems not to be the optimal choice from this data set. Its computation time is the highest by far and is better suited to more complex datasets with multiple inputs and uncertainty. For this particular dataset, OLS is preferable to other prediction methods owing to its superior performance metrics

and computation time. One year later, following the expiry of the period observed, the plant's total FC predicted by the OLS model is 4,322,436 t. Five-year and ten-year predictions are 10,684.86 and 18,615,472 t, respectively. Table 4 shows emission predictions for these periods based on estimated data to highlight the environmental effect of the plant. The emission calculation of the forecasted FC is based on the assumption that the route of the ship will remain the same.

Table 4.

OLS emission forecast.

Pollutant	1-year forecast (t)	5-year forecast (t)	10-year forecast (t)
CO ₂	13,599,249	33,616,707	58,567,998
CH ₄	0.259	0.641	1,117
N ₂ O	0.648	1.603	2,792
NO _x	399,091	986,533	1,718,767
NM VOC	13,313	32,909	57,336
CO	11,973	29,597	51,565
PM ₁₀	9,278	22,935	39,958
SO ₂	85,368	211,026	367,656

4. CONCLUSION AND DISCUSSION

The research examined the FC and emissions of a ship power plant through mathematical modelling. The numerical simulation calculated FC based on the data obtained from an oceangoing oil-chemical tanker. The study focused on marine diesel generators due to their major adverse environmental impact, especially in port and harbor areas. In addition, ships powered by two-stroke main engines use at least one diesel generator during navigation, resulting in continuous emissions by the plant. Between 6 December 2019 and 7 June 2021, model outputs have shown that the plant has produced 8,592.38 t of CO₂, 252.15 t of NO_x, and 53.98 t of SO₂. These values indicate substantial emissions even from a single ship plant. Even though IMO takes measures to reduce NO_x and SO₂, CO₂ reduction seems impossible without abandoning fossil fuels. To highlight the issue more efficiently, regression models were used to forecast future FC levels. Four different regression algorithms, suitable for the data, were trained and compared. The analogy identified OLS as the most suitable FC forecasting regression model for the current data. Although the other three algorithms performed successfully in terms of computation time, evaluation metrics and curve overlap quality, OLS is the optimal choice. One, five and ten-year forecasts obtained by OLS application to periods after 7 June 2021 showed that the plant's CO₂ production capacity was 58,567,998 t. This quantity of emissions has a significant impact on air pollution, especially given that they were generated by a single ship's auxiliary engines. Though promising, green ports supplying ships with electricity only prevent air pollution in the urban area, and their development is still an ongoing process. Short-term onboard solutions include plant hybridization using alternative energy sources or waste heat energy. Their installation and usage can be faster and would help reduce emissions from marine diesel generators. Long-term solutions include the usage

of alternative fuels and energy approaches onboard. However, they are still ongoing projects that need to be improved to be efficiently implemented on ships. Future studies will focus on battery additions to the plant and battery charging strategies due to their faster applicability.

ACKNOWLEDGMENT:

We thank Omurcan Yıldız for his expertise and assistance in all aspects of our study, as well as on his help with the collection of data for mathematical modelling.

CONFLICT OF 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.

REFERENCES

- Armstrong, V.N. & Banks, C., 2015. Integrated approach to vessel energy efficiency. *Ocean Engineering*, 110, pp.39–48. Available at: <http://dx.doi.org/10.1016/j.oceaneng.2015.10.024>.
- Altosole, M. et al., 2019. A Diesel Engine Modelling Approach for Ship Propulsion Real-Time Simulators. *Journal of Marine Science and Engineering*, 7(5), p.138. Available at: <http://dx.doi.org/10.3390/jmse7050138>.
- Altosole, M. & Figari, M., 2011. Effective simple methods for numerical modelling of marine engines in ship propulsion control systems design. *Journal of Naval Architecture and Marine Engineering*, 8(2), pp.129–147. Available at: <http://dx.doi.org/10.3329/jname.v8i2.7366>.
- Awad, M. & Khanna, R., 2015. Support Vector Regression. *Efficient Learning Machines*, pp.67–80. Available at: http://dx.doi.org/10.1007/978-1-4302-5990-9_4.
- Bal Beşikçi, E. et al., 2016. An artificial neural network based decision support system for energy efficient ship operations. *Computers & Operations Research*, 66, pp.393–401. Available at: <http://dx.doi.org/10.1016/j.cor.2015.04.004>.

- Bialystocki, N. and Konovessis, D., 2016. On the estimation of ship's fuel consumption and speed curve: A statistical approach. *Journal of Ocean Engineering and Science*, 1(2), pp. 157-166. Available at: <http://dx.doi.org/10.1016/j.joes.2016.02.001>.
- Bratić, K. et al., 2021. Cruise Vessels Air Pollution Inventory for the Port of Kotor. *Transactions on Maritime Science*, 10(1). Available at: <http://dx.doi.org/10.7225/toms.v10.n01.016>.
- Cahyagi, D., 2018. Study of Shipboard Power Distribution System: Review on an Application of AC Zonal Distribution. *IPEK Journal of Engineering*, 4(1). Available at: <http://dx.doi.org/10.12962/joe.v4i1.3734>.
- Chai, T. & Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), pp.1247–1250. Available at: <http://dx.doi.org/10.5194/gmd-7-1247-2014>.
- Craven, B. D. and Islam, S. M., 2011. Ordinary least-squares regression. *The SAGE dictionary of quantitative management research*, pp.224-228. ISBN: 9781412935203.
- Dreyfus G., 2005. *Neural Networks: Methodology and Applications*. Berlin: Springer Science & Business Media. ISBN: 978-3-642-06187-5.
- Eide, M.S. et al., 2011. Future cost scenarios for reduction of ship CO₂ emissions. *Maritime Policy & Management*, 38(1), pp.11–37. Available at: <http://dx.doi.org/10.1080/03088839.2010.533711>.
- Farag, Y.B.A. & Ölçer, A.I., 2020. The development of a ship performance model in varying operating conditions based on ANN and regression techniques. *Ocean Engineering*, 198, p.106972. Available at: <http://dx.doi.org/10.1016/j.oceaneng.2020.106972>.
- Hansen, A.D. & Michalke, G., 2008. Modelling and control of variable-speed multi-pole permanent magnet synchronous generator wind turbine. *Wind Energy*, 11(5), pp.537–554. Available at: <http://dx.doi.org/10.1002/we.278>.
- IMO, 2014. *Third IMO GHG Study Executive Summary*. London, UK.
- Kasuya, E., 2018. On the use of r and r squared in correlation and regression. *Ecological Research*, 34(1), pp.235–236. Available at: <http://dx.doi.org/10.1111/1440-1703.1011>.
- Kesgin, U. & Vardar, N., 2001. A study on exhaust gas emissions from ships in Turkish Straits. *Atmospheric Environment*, 35(10), pp.1863–1870. Available at: [http://dx.doi.org/10.1016/s1352-2310\(00\)00487-8](http://dx.doi.org/10.1016/s1352-2310(00)00487-8).
- Kim, Y.-R., Jung, M. & Park, J.-B., 2021. Development of a Fuel Consumption Prediction Model Based on Machine Learning Using Ship In-Service Data. *Journal of Marine Science and Engineering*, 9(2), p.137. Available at: <http://dx.doi.org/10.3390/jmse9020137>.
- Kingma, D. P. and Ba, J., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Krause, P. C. et al., 2002. *Analysis of Electric Machinery and Drive Systems 2*. New York: IEEE press. ISBN: 978-1-118-02429-4.
- Kuzu, S.L., Bilgili, L. & Kiliç, A., 2020. Estimation and dispersion analysis of shipping emissions in Bandirma Port, Turkey. *Environment, Development and Sustainability*, 23(7), pp.10288–10308. Available at: <http://dx.doi.org/10.1007/s10668-020-01057-6>.
- Le, L.T. et al., 2020. Neural network-based fuel consumption estimation for container ships in Korea. *Maritime Policy & Management*, 47(5), pp.615–632. Available at: <http://dx.doi.org/10.1080/03088839.2020.1729437>.
- Le, L.T. et al., 2019. Voyage-based statistical fuel consumption models of ocean-going container ships in Korea. *Maritime Policy & Management*, 47(3), pp.304–331. Available at: t.
- Leloup, R. et al., 2016. A continuous and analytical modeling for kites as auxiliary propulsion devoted to merchant ships, including fuel saving estimation. *Renewable Energy*, 86, pp.483–496. Available at: <http://dx.doi.org/10.1016/j.renene.2015.08.036>.
- Liu, J. & Duru, O., 2020. Bayesian probabilistic forecasting for ship emissions. *Atmospheric Environment*, 231, p.117540. Available at: <http://dx.doi.org/10.1016/j.atmosenv.2020.117540>.
- McGeorge, H. D., 1998. *Marine Auxiliary Machinery*. 7th Ed. Elsevier. ISBN: 9781483193892.
- Mills, T. C., 2019. *Applied Time Series Analysis: A Practical Guide to Modelling and Forecasting*. Academic Press. ISBN: 978-0128131176.
- Miola, A. & Ciuffo, B., 2011. Estimating air emissions from ships: Meta-analysis of modelling approaches and available data sources. *Atmospheric Environment*, 45(13), pp.2242–2251. Available at: <http://dx.doi.org/10.1016/j.atmosenv.2011.01.046>.
- Moreira, L., Vettor, R. & Guedes Soares, C., 2021. Neural Network Approach for Predicting Ship Speed and Fuel Consumption. *Journal of Marine Science and Engineering*, 9(2), p.119. Available at: <http://dx.doi.org/10.3390/jmse9020119>.
- Oways, 2019. *Electrical Distribution*. Available at: <https://owaysonline.com/electrical-distribution/>.
- Pohlman, J. T. and Leitner, D. W., 2003. A comparison of ordinary least squares and logistic regression. *Ohio Journal of Science*, 103(5), pp. 118-125.
- Prpić-Oršić, J. & Faltinsen, O.M., 2012. Estimation of ship speed loss and associated CO₂ emissions in a seaway. *Ocean Engineering*, 44, pp.1–10. Available at: <http://dx.doi.org/10.1016/j.oceaneng.2012.01.028>.
- Pulkrabek, W. W., 2004. *Engineering Fundamentals of the Internal Combustion Engine*. ISBN: 9781292027296.
- Reis, M. S. et al., 2020. Predicting ship's CO₂ emissions using feature-oriented methods. *Applied Stochastic Models in Business and Industry*, 36(1), pp. 110-123. Available at: <https://doi.org/10.1002/asmb.2477>.
- Simonsen, M., Walnum, H. & Gössling, S., 2018. Model for Estimation of Fuel Consumption of Cruise Ships. *Energies*, 11(5), p.1059. Available at: <http://dx.doi.org/10.3390/en11051059>.
- Styhre, L. et al., 2017. Greenhouse gas emissions from ships in ports – Case studies in four continents. *Transportation Research Part D: Transport and Environment*, 54, pp.212–224. Available at: <http://dx.doi.org/10.1016/j.trd.2017.04.033>.
- Tichavska, M. & Tovar, B., 2015. Port-city exhaust emission model: An application to cruise and ferry operations in Las Palmas Port. *Transportation Research Part A: Policy and Practice*, 78, pp.347–360. Available at: <http://dx.doi.org/10.1016/j.tra.2015.05.021>.
- Trebuna, P. et al., 2014. The importance of normalization and standardization in the process of clustering. *2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*. Available at: <http://dx.doi.org/10.1109/sami.2014.6822444>.
- Trozzi, C., 2010. Emission estimate methodology for maritime navigation. In *19th International emission inventory conference*, 27–30 September, San Antonio, USA.
- Uyanık, T., Karatuğ, Ç. & Arslanoğlu, Y., 2020. Machine learning approach to ship fuel consumption: A case of container vessel. *Transportation Research Part D: Transport and Environment*, 84, p.102389. Available at: <http://dx.doi.org/10.1016/j.trd.2020.102389>.

Wang, S. et al., 2018. Predicting ship fuel consumption based on LASSO regression. *Transportation Research Part D: Transport and Environment*, 65, pp.817–824. Available at: <http://dx.doi.org/10.1016/j.trd.2017.09.014>.

Winnes, H., Styhre, L. & Fridell, E., 2015. Reducing GHG emissions from ships in port areas. *Research in Transportation Business & Management*, 17, pp.73–82. Available at: <http://dx.doi.org/10.1016/j.rtbm.2015.10.008>.

Willmott, C. & Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, pp.79–82. Available at: <http://dx.doi.org/10.3354/cr030079>.

Yang, L. et al., 2019. A genetic algorithm-based grey-box model for ship fuel consumption prediction towards sustainable shipping. *Annals of Operations Research*. Available at: <http://dx.doi.org/10.1007/s10479-019-03183-5>.

Zanne, M., Počuča, M. & Bajec, P., 2013. Environmental and Economic Benefits of Slow Steaming. *Transactions on Maritime Science*, 2(2), pp.123–127. Available at: <http://dx.doi.org/10.7225/toms.v02.n02.005>.

Zhu, Y., Zuo, Y. & Li, T., 2021. Modeling of Ship Fuel Consumption Based on Multisource and Heterogeneous Data: Case Study of Passenger Ship. *Journal of Marine Science and Engineering*, 9(3), p.273. Available at: <http://dx.doi.org/10.3390/jmse9030273>.