International Journal on Advanced Science Engineering Information Technology

Vol.15 (2025) No. 1
ISSN: 2088-5334

Machine Learning-Driven Insights for Optimizing Ship Fuel Consumption: Predictive Modeling and Operational Efficiency

Phuoc Quy Phong Nguyen [a], Duy Tan Nguyen [b], Nhi Ha Thi Yen [c], Quang Le [d], Nhat Thang Nguyen [d], Nguyen Dang Khoa Pham [b,*]

[a] *Maritime College II, Ho Chi Minh City, Vietnam*
[b] *Institute of Maritime, Ho Chi Minh City University of Transport, Ho Chi Minh City, Vietnam*
[c] *Faculty of Foreign Languages, Dong Nai Technology University, Bien Hoa City, Vietnam*
[d] *Faculty of Engineering, Dong Nai Technology University, Bien Hoa City, Vietnam*
*Corresponding author: *khoapnd@ut.edu.vn*

Abstract*—Specific fuel consumption significantly impacts the shipping industry's operational costs and environmental footprint. Therefore, the development of modeling for predicting ship fuel consumption aiming to maximize the operation efficiency and minimize the pollutant emissions from ship operation is very necessary. By the use of predictive models such as Artificial Neural Networks (ANN) and eXtreme Gradient Boosting (XGBoost), this work optimizes ship fuel usage using machine learning approaches, in which we developed and tested both models using a 149-point dataset split between 70% training and 30% testing data. With a high R-value of 0.9851, MSE of 129.8261, and a low MAPE of 0.2541%, ANN showed good training performance. XGBoost topped ANN in training with a flawless R^2 of 1.000, MSE of 0.01, and MAPE of 0.21%. With a MAPE of 0.2121%, ANN showed great proportional accuracy on the test set; nevertheless, its high MSE of 26,300 indicated strong sensitivity to outliers. XGBoost showed strong outliers handling with a lower MSE of 259.45 but had a larger MAPE of 16.97%, thereby suggesting lowered proportional accuracy. The findings highlight both models' trade-offs and merits. XGBoost shows higher performance in controlling severe deviations, ANN shines in proportional accuracy. These realizations emphasize the need for model selection depending on application needs as they help predictive modeling for enhancing operational efficiency and reducing fuel consumption in the shipping industry.

***Keywords*— Ship fuel consumption; ship operation efficiency; maritime technology; machine learning; XGBoost; neural networks.**

Manuscript received 18 Jul. 2024; revised 12 Nov. 2024; accepted 7 Dec. 2024. Date of publication 28 Feb. 2025.

I. INTRODUCTION

In its 2020 report on greenhouse gas emissions, the International Maritime Organization (IMO) revealed that the global shipping industry contributed about 2.89% of the planet's total emissions in 2018. This meant around 1.076 billion tons of greenhouse gases were dumped into the environment. Unbelievably high predictions of a 50% to 250% increase in carbon dioxide emissions from shipping without intervention highlighted the immediate necessity of sustainable practices in the marine industry [1]–[4]. Marine fuel costs, however, have been steadily high all over. Now, a major and rising percentage of fleet running costs are fuel expenditures, which put great financial strain on the shipping sector [5]–[7]. Since most ships worldwide run on marine diesel engines, improving their efficiency is essential to maintain operating profitability. Fuel optimization is a two-edged need as any inefficiencies in fuel usage not only result in high prices but also aggravate environmental effects [8]–[11].

In the maritime industry, operational efficiency is exactly correlated with fuel usage. Finding trends of inefficiency and places for development depends on an analysis of fuel use. Accurate fuel consumption statistics let operators create plans that maximize routes, reduce waste, and guarantee ships run within their most effective performance range [12], [13]. These steps help reduce running expenses and improve market competitiveness. These approaches will also help in coinciding with world carbon-reduction targets [14], [15]. Policies include the IMO's Energy Efficiency Existing Ship Index, and the Carbon Intensity Indicator demand complete awareness of a vessel's energy use profile. Employing continuous fuel consumption monitoring and analysis, operators may make informed decisions to reduce emissions and ensure long-term profitability [16]–[19].

Investments in technologies like alternative propulsion systems, voyage planning software, and advanced fuel monitoring systems might greatly improve operational efficiency. By enabling accurate modifications throughout journeys, these technologies ensure ships burn the least amount of fuel required and preserve the safety and performance objectives [20]. These developments not only lower fuel costs but also support more general environmental goals, therefore putting the maritime sector as a major participant in the worldwide battle against climate change [21]. Given the critical deadlines for carbon reduction and the growing costs of marine fuels, fuel consumption monitoring has become a foundation of operational strategy in the maritime sector. In a regulatory environment becoming increasingly demanding, it represents a vital beginning towards achieving both environmental responsibility and financial sustainability [22].

In recent years, specific fuel consumption (SFC) prediction for ships has attracted more and more active studies. Using shipping data, a growing corpus of creative works has investigated the application of machine learning (ML) models to predict SFC at sea [23]–[26]. Notwithstanding these developments, numerous important aspects make precisely forecasting SFC difficult. The main difficulty is SFC forecast accuracy depending on shipping data quality and volume [27]–[29]. Nonetheless, ship-related data is intrinsically private as it relates to the transportation secrecy of shippers as well as shipping firms. Many times, regulatory restrictions limit enterprises engaged in shared shipping from exchanging SFC-related data. These constraints complicate access to enough high-quality data for strong model training and validation. Furthermore, greatly affected by a variety of climatic elements in different and erratic navigation conditions is SFC performance [30], [31]. Variations in meteorological conditions such as wind speed, wave height, and ocean currents that fluctuate dynamically across various areas and periods provide significant errors in SFC forecasts. Each ship and navigation situation affects SFC differently from these elements, hence careful feature selection is necessary to guarantee correct forecasts [25], [32], [33]. Improving prediction dependability depends on choosing the most relevant variables from a complicated and changing collection of inputs. Accurate SFC prediction is hampered overall by two main obstacles: restricted availability of high-quality shipping data resulting from privacy concerns and the diversity of environmental elements in navigation [34]. Dealing with data security and adjusting to changing operating situations while keeping forecast accuracy calls for creative ideas that help to overcome these challenges [35].

Constraints in conventional forecast methods make SFC in the marine sector exact is still a challenging task. Often, conventional static models ignore the dynamic and complex nature of navigational environments or the many influences of climatic elements. Especially concerning building robust, adaptable, and accurate forecasting models for SFC, these shortcomings provide significant research opportunities. Filling up these gaps will help the maritime sector improve operational efficiency, economic resilience, and compliance with ever more environmental requirements. This study guarantees accurate SFC predictions employing a revolutionary method using artificial neural networks (ANN) and eXtreme Gradient Boosting (XGBoost), therefore transcending the limitations of traditional methods. Integrated with ANN, XGBoost—known for its ability to handle high-dimensional data and complex relationships—is examined and increases the overall prediction accuracy. The framework builds a powerful model capable of managing numerous marine operational circumstances by using XGBoost's efficacy in feature selection and gradient-based optimization as well as the flexibility of the ANN in capturing nonlinear patterns. The predictive powers of ANN and XGBoost are assessed in hybrid architecture wherein XGBoost enhances ANN by improving input feature significance and reducing prediction errors, hence assuring precise SFC forecasting throughout many operating scenarios. By filling up certain crucial research gaps, our study improves SFC prediction by state-of-the-art standards. Comparative study of ANN and XGBoost ensures scalability and robustness in addition to improving forecasting accuracy, thus matching with global environmental and economic aims. The maritime sector will find great use for the findings of this research. With rising environmental regulations and fuel costs, the proposed idea provides a relevant solution for operating expenditure reduction. It also provides participants with effective tools to reach environmental aims and increase economic resilience. By allowing smart fuel consumption and educated decision-making, this study helps to equip a future-ready maritime industry.

II. Materials and Method

A. Data Collection, Cleaning, and Analysis

A ship register was the source of the data for this research; it offers an array of data on vessel operations, fuel consumption, and related environmental variables. This dataset is the foundation for creating and confirming the proposed PSO-optimized ANN architecture. Data collection included collating information from several ships operating in various navigational and climatic environments. The selection criteria assured that the dataset had a wide range of vessel types, operational features, and fuel efficiency patterns. This variation determines the generalizability of the predictive model across various shipping situations. Raw data was carefully cleaned to guarantee its validity and quality. Typical issues such as outliers, missing data, and repeated entries were identified and fixed. Missing values were imputed using statistical techniques to preserve the integrity of the dataset; duplicate entries were removed to avoid biased conclusions. Particularly in fuel use and meteorological data, outliers were extensively investigated and either deleted or corrected based on domain knowledge and statistical criteria.

By the use of a robust data pipeline, this stage guarantees that the employed dataset for model development is both comprehensive and consistent. The basis of the proposed approach is the mix of high-quality data with innovative analytical techniques, which assures exact and useful estimates of SFC in the maritime industry.

B. Artificial Neural Networks

Artificial Neural Networks (ANNs) have become a fundamental tool for regression tasks due to their ability to model complex and nonlinear relationships between inputs and outputs [36], [37]. ANNs provide a flexible and efficient structure in ship fuel consumption prediction to examine the complex interaction of operational characteristics and environmental circumstances affecting fuel economy [32]. An ANN's three fundamental layers—the input layer, hidden layers, and output layer, form a sequence. Data from several operational and environmental factors, like ship speed and distance covered during the voyage comes into the input layer. These factors project particular SFC. By the use of a network of linked neurons, the hidden layers analyze this input by using activation functions to capture the nonlinear connections. The projected SFC values then are produced by the output layer [5], [38].

Using a training procedure, ANNs get trained on the patterns in the data. The network changes the weights and biases of its connections throughout training to reduce the error between expected and real values. Gradient descent and other iteratively updating the weights depending on the gradient of the error function helps to accomplish this [39], [40]. Advanced optimization methods such as Particle Swarm Optimization (PSO) improve the training process even further by identifying ideal starting weights and biases, hence lowering the probability of the network being caught in local minima [41]. ANNs' capacity to generalize across several datasets is one of its main benefits for SFC prediction. In the marine sector, where ship kinds, routes, and operational circumstances vary greatly, this capacity is especially crucial. Because ANNs are flexible enough to fit these variances, they provide accurate forecasts for many situations. Incorporating meteorological variables allows the network to additionally consider dynamic environmental factors, like wind speed and wave height, thus guaranteeing strong performance under many navigation circumstances [42], [43].

The ANNs have the potential to replicate both linear as well as nonlinear connections between control factors and outputs. For SFC prediction, this factor is very important as a confluence of factors interacting in complex ways influences fuel consumption. By using these characteristics, ANNs may uncover subtle patterns and correlations buried by traditional regression methods, hence generating more accurate and consistent predictions. Moreover, the modular design of ANNs allows one to quickly scale and modify [44]. Investigators may modify the architecture by changing the number of hidden layers, neurons, and activation functions to match specific purposes. This versatility ensures that the network might be tailored to meet the particular requirements of SFC prediction in the maritime sector. Figure 1 shows typical ANN architecture.

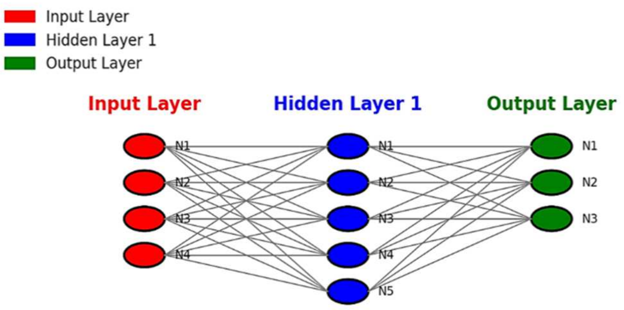

Fig. 1 ANN architecture

C. Extreme Gradient Boosting

eXtreme Gradient Boosting (XGBoost) is an advanced implementation of gradient-boosting algorithms. It is an efficient, flexible, and more precise version of gradient boosting for predictive modeling. The XGBoost regression develops an ensemble of decision trees in sequence. Herein, each tree is trained to minimize the residual errors of the preceding model [45], [46]. Guided by a gradient descent optimization approach, which ensures that the model always improves by focusing on reducing prediction errors, this process first produces predictions based on a constant value—usually the mean of the target variable—XGBoost begins its construction. Later additions to the mix include additional trees, each trained on the residual errors of the current model [47]. XGBoost measures the difference between anticipated and actual values using a loss function—such as mean squared error (MSE)—for regression during training. To direct the optimization process, it then determines the gradient of the loss function concerning the predictions [48]. Figure 2 shows a typical XGBoost architecture.

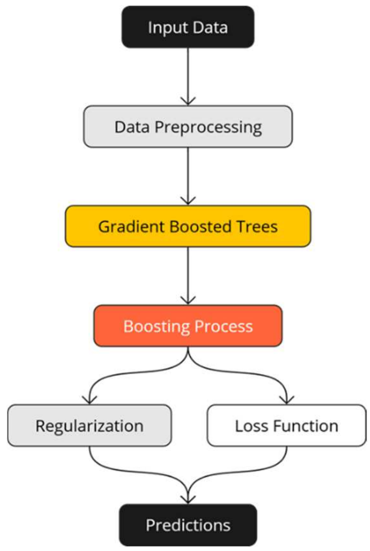

Fig. 2 Flow chart of XGBoost

XGBoost stands out mainly for its use of regularizing methods like L1 (Lasso) and L2 (Ridge), which penalize too complicated trees to avoid overfitting. The method also scales the contribution of every tree using a shrinkage parameter (learning rate), therefore guaranteeing smoother convergence [49]. XGBoost additionally employs column subsampling, in which case just a subset of characteristics is taken into account at every iteration hence improving generalization and processing efficiency. XGBoost's tree building is improved via a greedy split search wherein every node is split to maximize the loss reduction. Moreover, sparsity-aware techniques ensure efficient handling of missing data in XGBoost. The parallelization and support for distributed computing enable this method to analyze high-dimensional data with scalability, hence strengthening its robustness for use in many situations in regression analysis [50]–[52].

D. Statistical Evaluation

The reliability and performance of machine learning (ML) models depend on quality statistical assessment. Important measures including Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and R-squared (R^2) were employed in the present work. While MSE and MAPE assess prediction errors, R^2 indicates explained variance, therefore gauging how well a model prediction matches the actual data. The following mathematical expressions were employed for the calculation of statistical metrics [53], [54]:

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

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

$$MAPE = \frac{1}{n}\sum_{i=1}\left|\frac{y_i - \hat{y}_i}{y_i}\right|^2 \times 100 \quad (3)$$

In this case, the y_i depicts the measured value of SFC, n is the count of samples used in the study, $\hat{y}_i$ depicts the model-predicted values of SFC, $\bar{y}_i$ depicts the mean of measured values for SFC.

III. Results and Discussion

A. Correlational Analysis of SFC Data

Figure 3 is a heatmap of data used in this study It depicts the correlation matrix among three variables: time (in hours), distance (in nautical miles), and fuel consumption (in gallons). Numerical values and color intensity indicate the strength of linear correlations between these variables. As might be predicted, a correlation coefficient of 1.00 along the diagonal denotes complete self-correlation. The off-diagonal values expose quite substantial positive correlations among the variables.

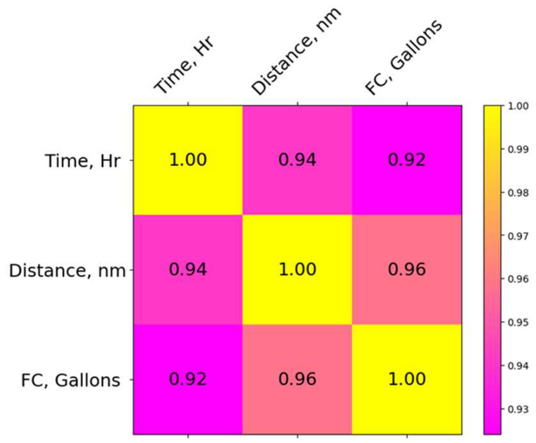

Fig. 3 Correlational heatmap

For example, the 0.96 correlation between distance and fuel consumption indicates that fuel consumption rises almost exactly with distance travel. In a similar vein, time and distance show a significant connection of 0.94, meaning that longer travel durations usually translate into more miles. Finally, time and fuel use exhibit a 0.92 connection, which reflects the link between journey length and fuel usage. These strong relationships validate their usage as inputs for predictive modelling as they imply that the factors are interconnected and thus affect ship fuel consumption [55], [56].

B. Model Development

1) Artificial neural network:

In this study, an Artificial Neural Network (ANN) model was developed to predict ship fuel consumption using key operational parameters. With 149 data rows, the dataset included goal variables related to the fuel consumption of the ship while input variables included the distance traversed and the time spent on journeys. Three subsets comprised the dataset: 70% of the data was used for training the model, 15% for validation to fine-tune hyperparameters and stop overfitting, and the remaining 15% for testing to assess the model's performance on unprocessed data. The careful design of the ANN architecture allowed it to capture the non-linear connections between the output variable and the input characteristics. Hidden layers used activation functions to improve the network's capacity to learn intricate data patterns. Optimization techniques were used during training to reduce the discrepancy between the expected and actual fuel consumption figures, therefore guaranteeing that the model fairly represented real-world situations. The validation method gave comments on the generalization capacity of the model, thereby allowing architectural changes and training parameter modifications for better performance. The ANN architecture used in this study had two inputs, one output, and ten neurons in a hidden layer as depicted in Figure 4.

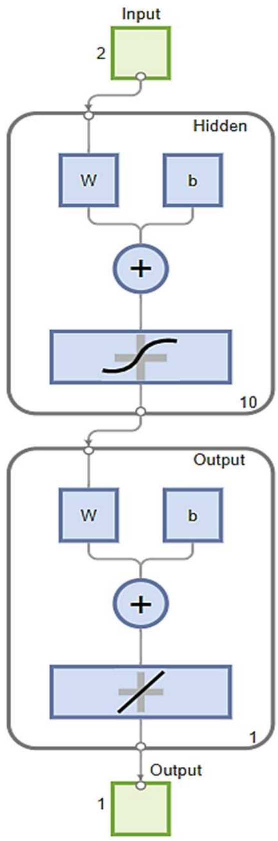

Fig. 4 ANN architecture used in this study

The training statistics of the ANN model used for ship fuel consumption prediction are shown in Table 1. Early ending conditions, including minimum validation error improvement, caused the training process to end after 13 epochs—far before the maximum aim of 1000 epochs. From an initial value of 2.33E+05 to 122, the performance metric—gauges error—probably MSE—improved considerably to show successful learning. Although the zero-goal performance value was not met, the attained value shows significant development. Indicating convergence, the gradient—which represents the rate of change in weights—dropped from 1.15E+06 to 223. Reaching the objective of 6, the validation checks imply no appreciable change in validation error during six consecutive epochs, so probably led to an early ending. From 0.001 to 10, the Mu value—a factor regulating the learning rate in Levenberg-Marquardt optimization—showcased signal model changes throughout training. Within the limits imposed by early stopping conditions, the model showed general good learning.

TABLE I
TRAINING PARAMETERS FOR ANN

Unit	Initial value	Stopped value	Target value
Epoch	0	13	1000
Elapsed Time	-	00:00:00	-
Performance	2.33E+05	122	0
Gradient	1.15E+06	223	1.00E-07
Mu	0.001	10	1.00E+10
Validation Checks	0	6	6

The ANN model demonstrated strong predictive capabilities across training, validation, and testing phases as listed in Table 2. With a correlation coefficient of 0.9851, the model obtained an MSE of 129.8261 during training, therefore demonstrating a very good match between the anticipated and actual data. With a mean absolute percentage error of 0.2541 percent, exact forecasts with little variation were shown. Showing strong generalizing ability, the MSE dropped to 25.57 in the validation phase using a correlation value of 0.9218 and a mean absolute percentage error of 3.258 percent. Strong prediction accuracy was shown in the test phase by the model obtaining a mean absolute percentage error of 0.2121 percent and a correlation value of 0.988. At 2.63E+04, the MSE was larger, however, indicating the likelihood of outliers or significant test data variance. The model showed consistent generalization to unprocessed data and did quite well in estimating ship fuel usage.

TABLE II
STATISTICAL EVALUATION OF THE ANN MODEL

	Observations	MSE	R	MAPE, %
Training	105	129.8261	0.9851	0.2541
Validation	22	25.57	0.9218	3.258
Test	22	2.63E+04	0.988	0.2121

This subfigure (Figure 5a) shows the gradient, mu, and validation checks along with the training progress for 13 epochs. Reflecting the convergence of the model toward an optimum solution, the gradient, which reflects the rate of change of the error function, drops greatly from the beginning value and settles at 223.0834 by the last epoch. The mu parameter, which controls the step size in Levenberg-Marquardt optimization, rises gradually to show changes to the learning rate throughout training to balance speed and stability. Early halting was set off after six consecutive epochs without an appreciable increase in validation performance, according to validation tests. .

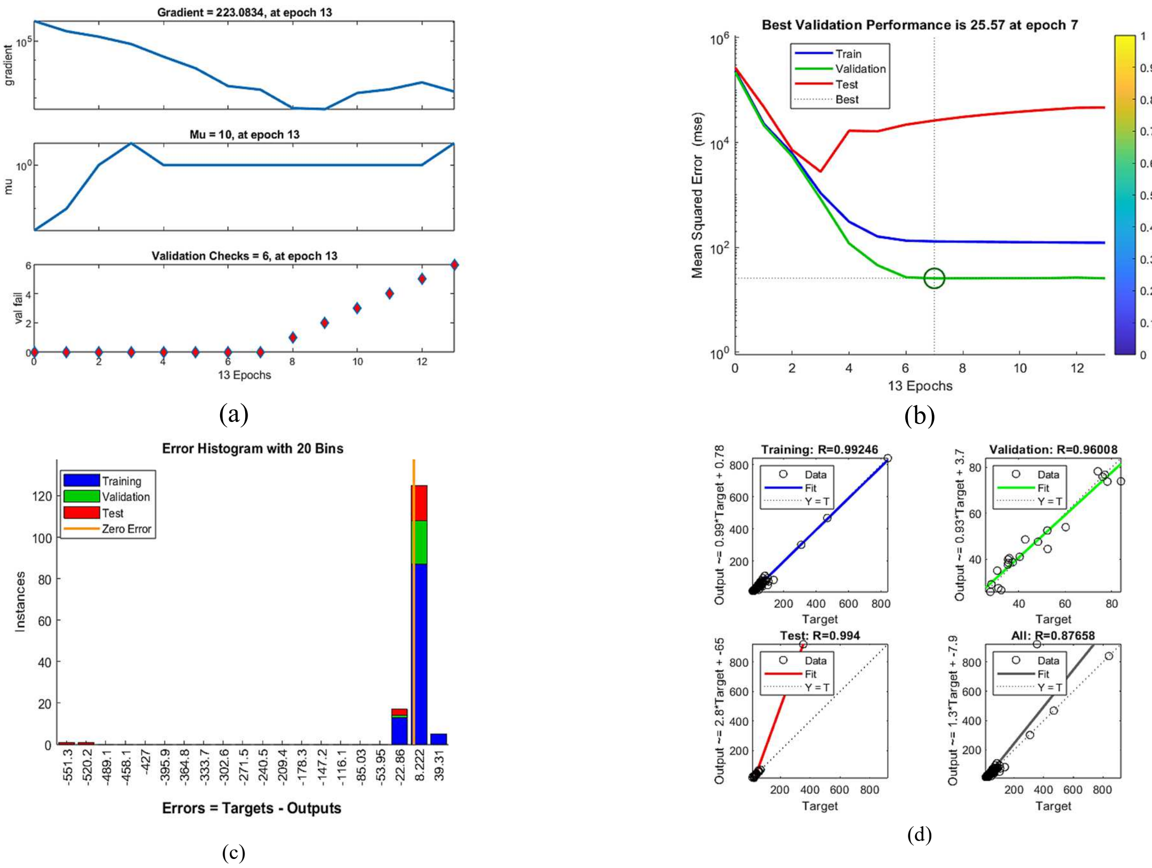

Fig. 5 ANN model (a) gradient, mu, and validation checks (b) training progress (c) error histogram (d) correlation results

Once the model begins to overfit the validation data, this method stops training, therefore preventing overfitting. Over 13 epochs, the subfigure (Figure 5b) the mean squared error (MSE) trends for the training, validation, and test sets. During the first epoch, the training and validation errors drop quickly; the validation error reaches its lowest level of 25.57 at epoch 7, therefore indicating the best validation performance. Past this point, the validation error starts to somewhat rise, suggesting possible overfitting. Though larger overall, the test error stabilizes and indicates that the model generally fits unobserved data.

For training, validation, and test datasets the histogram (Figure 5c) shows the distribution of errors—that is, variations between goals and outputs. Usually clustered close to zero, most mistakes indicate correct forecasts. A few outliers with more significant errors, especially in the test set, are evident, nevertheless, and may have affected the high test MSE. For training, validation, testing, and combined datasets the regression graphs (Figure 5d) show the link between expected and actual results. With a R = 0.99246 strong correlation value, the training plot exhibits a tight match. Strong but much less perfect fits are shown by the somewhat lower R values of 0.96008 and 0.994 respectively from the validation and test sets. With an R-value of 0.87658 overall, the model's dependability in estimating ship fuel consumption across all datasets is highlighted

2) XGboost-based SFC prediction model:

The XGBoost model was developed using a dataset of 149 data points containing key operational parameters for predicting SFC in gallons. To ensure that the training and assessment of the model were conducted on different datasets, the dataset was split into training and testing subsets in a 70:30 ratio. Comprising seventy percent of the data, the training sample helped to maximize the parameters of the model and equip it to spot trends and linkages within the dataset. The remaining thirty percent was set aside for testing to evaluate generalizing to unprocessed data and predicted the accuracy of the model. The graphic shows how well the XGBoost model forecasts ship fuel use. For both the training and testing sets, the graphic contrasts measured SFC values with expected ones. The training data are shown by the green stars, and the testing data by the red stars. Where anticipated and measured values are equal, the black dashed line—the ideal line of perfect prediction—is found. The green and orange dashed lines provide ±10% prediction boundaries, therefore offering a graphic assessment of prediction accuracy.

With an R^2 of 1, the model shows almost flawless performance for the training set, therefore demonstrating a perfect match between anticipated and actual values. At 0.01 the MSE is very low; the mean absolute percentage error (MAPE) is 0.21%, therefore verifying the model's accuracy throughout training. Most training data points exhibit great agreement by precisely on or very near the optimal prediction line. Though somewhat less than the training set, the model gets an R^2 of 0.9015 for the testing dataset, suggesting a significant connection between predicted and observed values. Higher than the training MSE, the test MSE is 259.45, indicating some aberrations in the model upon exposure to unprocessed data. Although it is larger than the training error, the test MAPE is 16.97%, still within reasonable bounds for many practical uses. A few testing data points, however, stray from the optimal line and lie outside the ±10% prediction limits, therefore suggesting the existence of prediction mistakes for certain data points. Based on high R^2 values and low error measurements, the XGBoost model shows overall great performance during training and good generalizing on the test data. The chart points out opportunities for improvement, including lowering prediction errors for certain outliers, and emphasizes the model's capacity to fairly estimate ship fuel usage. This assessment emphasizes the relevance of the concept in pragmatic marine fuel efficiency uses. Figure 6 presents actual vs predicted SFC values.

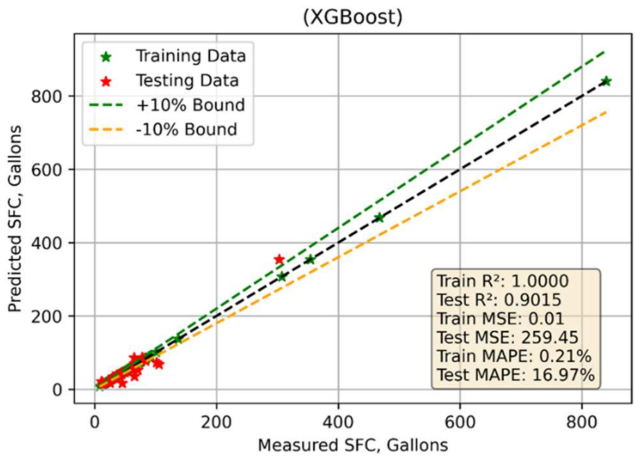

Fig. 6 Actual vs predicted SFC values.

Figure 7 evaluates the performance of the XGBoost model in predicting ship fuel consumption and is divided into four subplots, focusing on point-to-point comparisons and residual analysis for both training and test datasets. A detailed discussion of these subplots follows. For the training set, the subplot shown in Figure 7a contrasts the measured SFC (solid blue line) with the projected SFC (dashed green line). The high degree of overlap between the two lines indicates that the model faithfully reflects the trends and variances in the training data. The predictions closely reflect peaks and troughs in the observed data, suggesting the model learns intricate interactions between input factors and output very well. Small variations, however, might arise at certain places suggesting small prediction mistakes most likely related to overfitting to the training data. Consistent with the previous-recorded R^2 of 1, the almost perfect alignment shows outstanding performance on the training set overall. Plot Figure 7b shows for the test dataset the comparison between observed SFC (solid red line) and anticipated SFC (dashed red line). The alignment is less exact than in the training data, even though the expected values usually track the observed patterns. There are some variances, particularly at the highest levels.

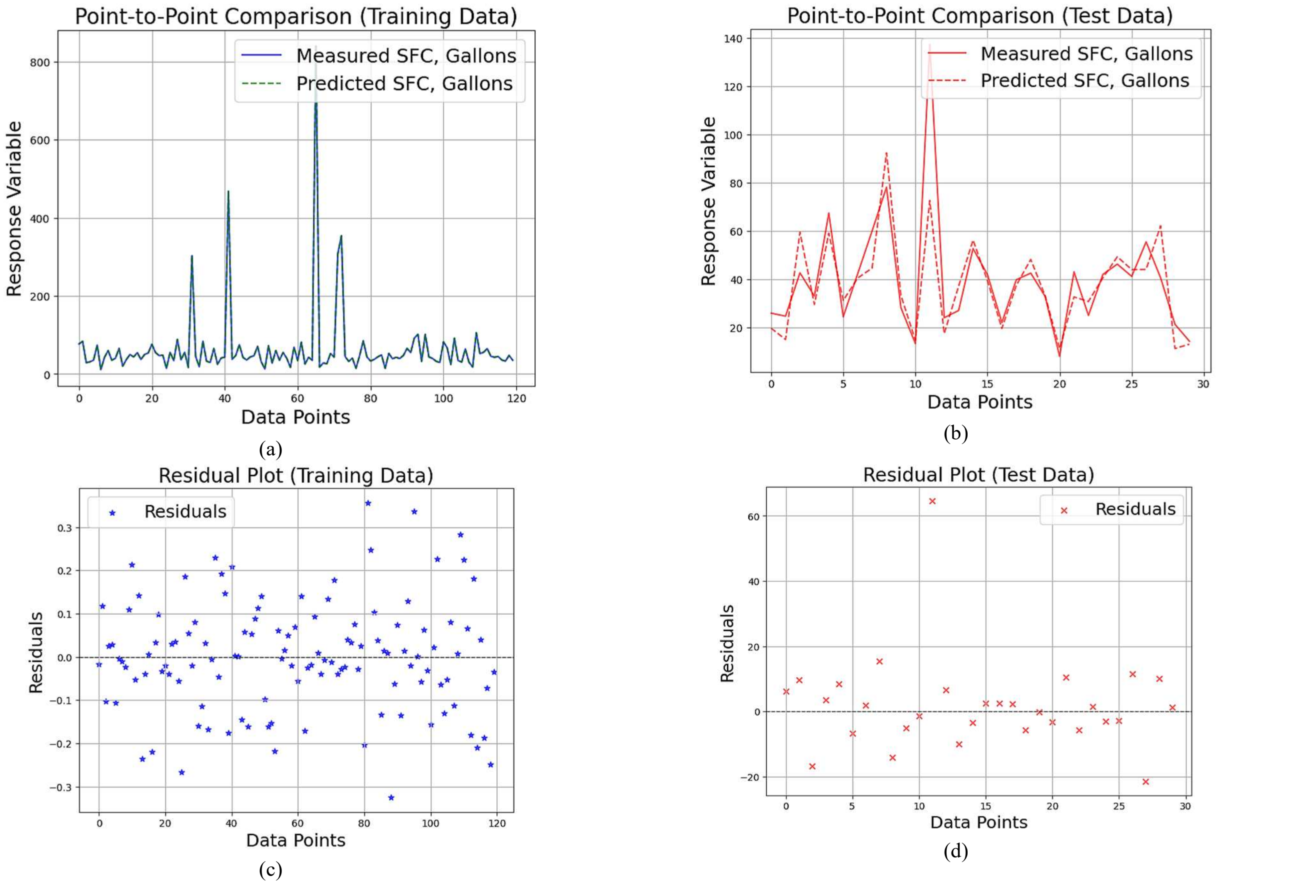

Fig. 7 XGBoost-based model's (a) Data point comparison during training (b) Data point comparison during test (c) Residual plot for training (d) Residual plot for test phase

These differences might be ascribed to noise, data variability, or inadequate training set scenario representation of certain events. Still, the model works reasonably on unknown data, and the general trend is well represented. Though less than for the training set, the findings line up with the test R^2 of 0.9015, which demonstrates strong predictive power. The deviations point out locations that extra hyperparameter adjustment or feature inclusion may be needed. Plotting residuals, that are the difference between observed and expected values—against data points for the training set is shown in Figure 7c. Residuals should ideally scatter randomly about 0 without any obvious pattern, meaning the model picks all systematic tendencies in the data. Confirming the validity of the model, the residuals in this subplot are zero and show no systematic bias. The few residuals support the low training MSE of 0.01, therefore verifying the correctness of the model in matching the training data. On the other hand, the grouping of certain elements refers to small local deviations, which in the whole framework are hardly significant. Figure 7d shows, like the training data, the residuals for the test dataset. Test residuals show more variation than training residuals; certain locations differ far from zero. These anomalies highlight certain test data points where the model has difficulty forecasting precisely. Such events might result from underrepresented patterns or noise in data attributes not sufficiently recorded during training. Still, most residuals fall within reasonable limits, therefore validating the test MAPE of 16.97%. The distribution indicates no systematic bias, meaning the model generalizes well and emphasizes the necessity of improving its performance even further in certain situations.

Both the point-to-point comparison and residual plot indicate that the XGBoost model works very well on the training dataset, precisely predicting SFC with minimum mistakes, so. This good result shows the capacity of the model to efficiently learn the fundamental connections in the data. Despite occasional outliers and somewhat higher errors, the model generalizes effectively on the test set to capture main patterns. Although the test performance is strong, especially in managing peak values and outliers, it still offers the opportunity for development. Residual analysis shows neither dataset has any consistent inclination. The modest residuals in the training set and good generalization-oriented tolerable deviations in the test set point to a well-trained model. Improving test performance might call for hyperparameter adjustment, more diverse training data, or investigating other characteristics to better capture complicated interactions. Dealing with test outliers might help the model's dependability even further. All things considered, Table 3 shows how well the XGBoost model predicts ship fuel usage, shows great training results and good generalizing to unknown data, and points out areas that need work.

TABLE III
STATISTICAL EVALUATION OF THE XGBOOST MODEL

Metric	Train	Test
R^2	1	0.9015
MSE	0.01	259.45
MAPE	0.21%	16.97%

3) Model comparison:

The comparative analysis of the ANN and XGBoost models reveals distinct strengths and weaknesses in terms of training and test performance metrics. With an R of 0.9851, MSE of 129.8261, and a rather low MAPE of 0.2541%, the ANN model shows great training performance. On training data, the XGBoost model beats by perfect R^2 of 1.000, a far smaller MSE of 0.01, and even a lower MAPE of 0.21%. This suggests that XGBoost shines in remarkably accurate learning of the patterns in the training data. The performance varies greatly on the test set. With a high R of 0.988, a modest MAPE of 0.2121%, and a disproportionately high MSE of 26,300—indicating the existence of significant outlier errors in its predictions—the ANN model performs. On the other hand, albeit having a somewhat lower R^2 at 0.9015, the XGBoost model has a considerably smaller MSE of 259.45 and a somewhat higher MAPE of 16.97%. This implies, based on the greater MAPE, XGBoost shows less consistent accuracy compared to the real values even if its total prediction errors are fewer. With a reduced MAPE and reflecting its capacity to preserve proportional accuracy, the ANN model shows greater test consistency. It's quite high test MSE, however, suggests difficulties in managing severe deviations. Conversely, XGBoost fails to maintain proportional accuracy throughout the whole range of test data but shows strong management of outliers indicated by its low MSE. Overall, ANN shows great proportional accuracy with low MAPE across both datasets but suffers from great outlier sensitivity. XGBoost achieves a low test MSE by handling outliers better and delivering outstanding training performance; but its greater MAPE suggests that it compromises proportional accuracy. The application's tolerance for outliers and proportionate accuracy needs will choose which of the models to use.

IV. Conclusion

This research provides a comparative analysis of ANN and XGBoost models for optimizing ship fuel consumption through predictive modelling. The results draw attention to the different strengths and constraints of every model in test performance measures as well as in training. Having a low MAPE of 0.2541% during training and keeping a decent MAPE of 0.2121% on the test set, ANN shows outstanding proportional accuracy. However, its high test MSE of 26,300 points to difficulties controlling severe deviations and outliers. Conversely, XGBoost shows its capacity to precisely learn training data patterns by delivering a flawless R^2 of 1.000, a much lower MSE of 0.01, and MAPE of 0.21%. With a substantially lower MSE of 259.45, XGBoost efficiently controls outliers for test data; nonetheless, its larger MAPE of 16.97% represents inferior proportional accuracy throughout the data range.

The operational priorities determine which of ANN and XGBoost to use. While XGBoost is useful for situations that give controlling outliers priority and guarantee solid overall prediction performance, ANN is better suited for applications needing constant proportional accuracy and low variance. Particularly for complicated systems like ship fuel usage, the findings highlight the crucial importance model choice plays in predictive analytics. This work helps to further develop machine learning applications for marine operational efficiency, therefore allowing sustainable and reasonably priced solutions in the shipping sector.

References

[1] IMO, "Fourth IMO GHG study 2020," 2020. [Online]. Available: https://safety4sea.com/wp-content/uploads/2020/08/MEPC-75-7-15-Fourth-IMO-GHG-Study-2020-Final-report-Secretariat.pdf.

[2] M. Zhang, N. Tsoulakos, P. Kujala, and S. Hirdaris, "A deep learning method for the prediction of ship fuel consumption in real operational conditions," *Eng. Appl. Artif. Intell.*, vol. 130, p. 107425, Apr. 2024, doi: 10.1016/j.engappai.2023.107425.

[3] H. P. Nguyen *et al.*, "The electric propulsion system as a green solution for management strategy of CO2 emission in ocean shipping: A comprehensive review," *Int. Trans. Electr. Energy Syst.*, vol. 31, no. 11, p. e12580, 2021, doi: 10.1002/2050-7038.12580.

[4] V. V. Pham and A. T. Hoang, "Technological perspective for reducing emissions from marine engines," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 9, no. 6, pp. 1989–2000, 2019, doi: 10.18517/ijaseit.9.6.10429.

[5] W. Tarelko and K. Rudzki, "Applying artificial neural networks for modelling ship speed and fuel consumption," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17379–17395, Dec. 2020, doi: 10.1007/s00521-020-05111-2.

[6] K. Rudzki, P. Gomulka, and A. T. Hoang, "Optimization Model to Manage Ship Fuel Consumption and Navigation Time," *Polish Marit. Res.*, vol. 29, no. 3, pp. 141–153, Sep. 2022, doi: 10.2478/pomr-2022-0034.

[7] V. G. Nguyen *et al.*, "Using Artificial Neural Networks for Predicting Ship Fuel Consumption," *Polish Marit. Res.*, vol. 30, no. 2, pp. 39–60, Jun. 2023, doi: 10.2478/pomr-2023-0020.

[8] T. Chu Van, J. Ramirez, T. Rainey, Z. Ristovski, and R. J. Brown, "Global impacts of recent IMO regulations on marine fuel oil refining processes and ship emissions," *Transp. Res. Part D Transp. Environ.*, vol. 70, pp. 123–134, May 2019, doi: 10.1016/j.trd.2019.04.001.

[9] IMO, "Initial IMO Strategy on reduction of GHG emissions from ships," 2018.

[10] P. Serra and G. Fancello, "Towards the IMO's GHG Goals: A Critical Overview of the Perspectives and Challenges of the Main Options for Decarbonizing International Shipping," *Sustainability*, vol. 12, no. 8, p. 3220, Apr. 2020, doi: 10.3390/su12083220.

[11] A. T. Hoang *et al.*, "Technological solutions for boosting hydrogen role in decarbonization strategies and net-zero goals of world shipping: Challenges and perspectives," *Renew. Sustain. Energy Rev.*, vol. 188, p. 113790, Dec. 2023, doi: 10.1016/j.rser.2023.113790.

[12] T. Cepowski, P. Chorab, and D. Łozowicka, "Application of an Artificial Neural Network and Multiple Nonlinear Regression to Estimate Container Ship Length Between Perpendiculars," *Polish Marit. Res.*, vol. 28, no. 2, pp. 36–45, 2021, doi: doi:10.2478/pomr-2021-0019.

[13] J. Herdzik, "Decarbonization of Marine Fuels—The Future of Shipping," *Energies*, vol. 14, no. 14, p. 4311, Jul. 2021, doi:10.3390/en14144311.

[14] L. Changxiong, Y. Hu, Z. Yang, and H. Guo, "Experimental Study of Fuel Combustion and Emission Characteristics of Marine Diesel Engines Using Advanced Fuels," *Polish Marit. Res.*, vol. 30, no. 3, pp. 48–58, Sep. 2023, doi: 10.2478/pomr-2023-0038.

[15] W. Zeńczak and A. K. Gromadzińska, "Preliminary Analysis of the Use of Solid Biofuels in a Ship's Power System," *Polish Marit. Res.*, vol. 27, no. 4, pp. 67–79, Dec. 2020, doi: 10.2478/pomr-2020-0067.

[16] V. N. Nguyen, N. Chung, G. N. Balaji, K. Rudzki, and A. T. Hoang, "Internet of things-driven approach integrated with explainable machine learning models for ship fuel consumption prediction," *Alexandria Eng. J.*, vol. 118, pp. 664–680, Apr. 2025, doi:10.1016/j.aej.2025.01.067.

[17] F.-M. Tzu and D.-T. Su, "Evaluation of carbon dioxide emission based on energy efficiency existing ship index during oceanographic navigation," *J. Oper. Oceanogr.*, vol. 17, no. 3, pp. 151–164, Sep. 2024, doi: 10.1080/1755876X.2023.2254133.

[18] V. Yalama, O. Yakovleva, V. Trandafilov, and M. Khmelniuk, "Future Sustainable Maritime Sector: Fishing Carriers and their Adoption to the Environmental Regulations. Part I," *Polish Marit. Res.*, vol. 29, no. 3, pp. 69–77, Sep. 2022, doi: 10.2478/pomr-2022-0027.

[19] A. T. Hoang *et al.*, "Energy-related approach for reduction of CO2 emissions: A critical strategy on the port-to-ship pathway," *J. Clean. Prod.*, vol. 355, p. 131772, Jun. 2022, doi:10.1016/j.jclepro.2022.131772.

[20] Q. Meng, Y. Du, and Y. Wang, "Shipping log data based container ship fuel efficiency modeling," *Transp. Res. Part B Methodol.*, 2016, doi: 10.1016/j.trb.2015.11.007.

[21] N. Gray, S. McDonagh, R. O'Shea, B. Smyth, and J. D. Murphy, "Decarbonising ships, planes and trucks: An analysis of suitable low-carbon fuels for the maritime, aviation and haulage sectors," *Adv. Appl. Energy*, vol. 1, p. 100008, Feb. 2021, doi: 10.1016/j.adapen.2021.100008.

[22] A. Fan, J. Yang, L. Yang, D. Wu, and N. Vladimir, "A review of ship fuel consumption models," *Ocean Eng.*, vol. 264, p. 112405, Nov. 2022, doi: 10.1016/j.oceaneng.2022.112405.

[23] Z. Hu, Y. Jin, Q. Hu, S. Sen, T. Zhou, and M. T. Osman, "Prediction of Fuel Consumption for Enroute Ship Based on Machine Learning," *IEEE Access*, vol. 7, pp. 119497–119505, 2019, doi:10.1109/access.2019.2933630.

[24] P. Karagiannidis, N. Themelis, G. Zaraphonitis, C. Spandonidis, and C. Giordamlis, "Ship fuel consumption prediction using artificial neural networks," in *Proceedings of the Annual meeting of marine technology conference proceedings, Athens, Greece*, 2019, pp. 46–51.

[25] S. Nguyen, X. Fu, D. Ogawa, and Q. Zheng, "An application-oriented testing regime and multi-ship predictive modeling for vessel fuel consumption prediction," *Transp. Res. Part E Logist. Transp. Rev.*, vol. 177, p. 103261, Sep. 2023, doi: 10.1016/j.tre.2023.103261.

[26] T. T. Le *et al.*, "Development of comprehensive models for precise prognostics of ship fuel consumption," *J. Mar. Eng. Technol.*, vol. 23, no. 6, pp. 451–465, Nov. 2024, doi: 10.1080/20464177.2024.2372888.

[27] T. Zhou, Q. Hu, Z. Hu, and R. Zhen, "An adaptive hyper parameter tuning model for ship fuel consumption prediction under complex maritime environments," *J. Ocean Eng. Sci.*, vol. 7, no. 3, pp. 255–263, Jun. 2022, doi: 10.1016/j.joes.2021.08.007.

[28] Y. Zhu, Y. Zuo, and T. Li, "Predicting Ship Fuel Consumption based on LSTM Neural Network," in *2020 7th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS)*, Nov. 2020, pp. 310–313, doi: 10.1109/ICCSS52145.2020.9336914.

[29] Z. Wang and S. Chen, "Real-time Forecast of Fuel Consumption of Ship Main Engine Based on LSTM Neural Network [J]," *J. Wuhan Univ. Technol. (Transportation Sci. Eng.*, vol. 44, no. 05, pp. 923–927, 2020.

[30] C. Sui, P. de Vos, D. Stapersma, K. Visser, and Y. Ding, "Fuel Consumption and Emissions of Ocean-Going Cargo Ship with Hybrid Propulsion and Different Fuels over Voyage," *J. Mar. Sci. Eng.*, vol. 8, no. 8, p. 588, Aug. 2020, doi: 10.3390/jmse8080588.

[31] T. Uyanık, Ç. Karatuğ, and Y. Arslanoğlu, "Machine learning approach to ship fuel consumption: A case of container vessel," *Transp. Res. Part D Transp. Environ.*, vol. 84, p. 102389, Jul. 2020, doi: 10.1016/j.trd.2020.102389.

[32] L. Moreira, R. Vettor, and C. Guedes Soares, "Neural Network Approach for Predicting Ship Speed and Fuel Consumption," *J. Mar. Sci. Eng.*, vol. 9, no. 2, p. 119, Jan. 2021, doi: 10.3390/jmse9020119.

[33] Z. Yuan, J. Liu, Y. Liu, Y. Yuan, Q. Zhang, and Z. Li, "Fitting Analysis of Inland Ship Fuel Consumption Considering Navigation Status and Environmental Factors," *IEEE Access*, vol. 8, pp. 187441–187454, 2020, doi: 10.1109/ACCESS.2020.3030614.

[34] R. Yan, S. Wang, and Y. Du, "Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship," *Transp. Res. Part E Logist. Transp. Rev.*, vol. 138, p. 101930, Jun. 2020, doi: 10.1016/j.tre.2020.101930.

[35] S. Feng *et al.*, "The Impact of Alternative Fuels on Ship Engine Emissions and Aftertreatment Systems: A Review," *Catalysts*, vol. 12, no. 2, p. 138, Jan. 2022, doi: 10.3390/catal12020138.

[36] T. Cong My, L. Dang Khanh, and P. Minh Thao, "An Artificial Neural Networks (ANN) Approach for 3 Degrees of Freedom Motion Controlling," *JOIV Int. J. Informatics Vis.*, vol. 7, no. 2, p. 301, May 2023, doi: 10.30630/joiv.7.2.1817.

[37] C. Halim, N. G. Eka Putra, N. A. Nugroho, and D. Suhartono, "Chest X-ray Image Classification to Identify Lung Diseases Using Convolutional Neural Network and Convolutional Block Attention Module," *JOIV Int. J. Informatics Vis.*, vol. 7, no. 3, pp. 651–658, Sep. 2023, doi: 10.30630/joiv.7.3.1136.

[38] J. Hadi, Z. Y. Tay, and D. Konovessis, "Ship Navigation and Fuel Profiling based on Noon Report using Neural Network Generative Modeling," *J. Phys. Conf. Ser.*, vol. 2311, no. 1, p. 012005, Jul. 2022, doi: 10.1088/1742-6596/2311/1/012005.

[39] P. K. Kanti, P. Sharma, B. Koneru, P. Banerjee, and K. D. Jayan, "Thermophysical profile of graphene oxide and MXene hybrid nanofluids for sustainable energy applications: Model prediction with a Bayesian optimized neural network with K-cross fold validation," *FlatChem*, vol. 39, p. 100501, May 2023, doi:10.1016/j.flatc.2023.100501.

[40] I. Veza *et al.*, "Review of artificial neural networks for gasoline, diesel and homogeneous charge compression ignition engine," *Alexandria Eng. J.*, vol. 61, no. 11, pp. 8363–8391, Nov. 2022, doi:10.1016/j.aej.2022.01.072.

[41] T. M. Shami, M. A. Summakieh, M. Alswaitti, M. A. Al Jahdhami, A. M. Sheikh, and A. A. El-Saleh, "TPPSO: A Novel Two-Phase Particle Swarm Optimization," *JOIV Int. J. Informatics Vis.*, vol. 7, no. 3–2, p. 2095, Nov. 2023, doi: 10.30630/joiv.7.3-2.2331.

[42] H. Adun, I. Wole-Osho, E. C. Okonkwo, O. Bamisile, M. Dagbasi, and S. Abbasoglu, "A neural network-based predictive model for the thermal conductivity of hybrid nanofluids," *Int. Commun. Heat Mass Transf.*, vol. 119, p. 104930, Dec. 2020, doi:10.1016/j.icheatmasstransfer.2020.104930.

[43] Z. Pang, F. Niu, and Z. O'Neill, "Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons," *Renew. Energy*, vol. 156, pp. 279–289, Aug. 2020, doi: 10.1016/j.renene.2020.04.042.

[44] A. Tuan Hoang *et al.*, "A review on application of artificial neural network (ANN) for performance and emission characteristics of diesel engine fueled with biodiesel-based fuels," *Sustain. Energy Technol. Assessments*, vol. 47, p. 101416, Oct. 2021, doi:10.1016/j.seta.2021.101416.

[45] P. Kumar K, M. Alruqi, H. A. Hanafi, P. Sharma, and V. V. Wanatasanappan, "Effect of particle size on second law of thermodynamics analysis of Al2O3 nanofluid: Application of XGBoost and gradient boosting regression for prognostic analysis," *Int. J. Therm. Sci.*, vol. 197, p. 108825, Mar. 2024, doi:10.1016/j.ijthermalsci.2023.108825.

[46] Z. Said, P. Sharma, B. J. Bora, and A. K. Pandey, "Sonication impact on thermal conductivity of f-MWCNT nanofluids using XGBoost and Gaussian process regression," *J. Taiwan Inst. Chem. Eng.*, vol. 145, p. 104818, Apr. 2023, doi: 10.1016/j.jtice.2023.104818.

[47] B. Akbar, H. Tayara, and K. T. Chong, "Unveiling dominant recombination loss in perovskite solar cells with a XGBoost-based machine learning approach," *iScience*, vol. 27, no. 3, p. 109200, Mar. 2024, doi: 10.1016/j.isci.2024.109200.

[48] J. Noh, T. Na, and S.-S. Kang, "Modeling Solar Irradiance in Tajikistan with XGBoost Algorithm," *J. Eng. Geol.*, vol. 33, no. 3, pp. 403–411, 2023.

[49] T. Chen and C. Guestrin, "XGBoost," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[50] J. Dong, Y. Chen, B. Yao, X. Zhang, and N. Zeng, "A neural network boosting regression model based on XGBoost," *Appl. Soft Comput.*, vol. 125, p. 109067, Aug. 2022, doi: 10.1016/j.asoc.2022.109067.

[51] Q. T. Phan, Y. K. Wu, and Q. D. Phan, "A Hybrid Wind Power Forecasting Model with XGBoost, Data Preprocessing Considering Different NWPs," *Appl. Sci.*, vol. 11, no. 3, p. 1100, Jan. 2021, doi:10.3390/app11031100.

[52] O. Sagi and L. Rokach, "Approximating XGBoost with an interpretable decision tree," *Inf. Sci. (Ny).*, vol. 572, pp. 522–542, Sep. 2021, doi: 10.1016/j.ins.2021.05.055.

[53] S. Roy, A. Ghosh, A. K. Das, and R. Banerjee, "Development and validation of a GEP model to predict the performance and exhaust emission parameters of a CRDI assisted single cylinder diesel engine coupled with EGR," *Appl. Energy*, vol. 140, pp. 52–64, Feb. 2015, doi:10.1016/j.apenergy.2014.11.065.

[54] P. Sharma, "Artificial intelligence-based model prediction of biodiesel-fueled engine performance and emission characteristics: A comparative evaluation of gene expression programming and artificial neural network," *Heat Transf.*, vol. 50, no. 6, pp. 5563–5587, Sep. 2021, doi: 10.1002/htj.22138.

[55] B. Kwon, F. Ejaz, and L. K. Hwang, "Machine learning for heat transfer correlations," *Int. Commun. Heat Mass Transf.*, vol. 116, p. 104694, Jul. 2020, doi: 10.1016/j.icheatmasstransfer.2020.104694.

[56] S. Pak and T. Oh, "Correlation and Simple Linear Regression," *J. Vet. Clin.*, vol. 27, no. 4, pp. 427–434, 2010.