J. Emerg. Sci. Eng. 2024, 2(2), e21

Contents list available at CBIORE journal website

Journal of Emerging Science and Engineering

Journal homepage: https://journal.cbiore.id/index.php/jese/index

Research Article

Soft computing-based modelling and optimization of NOx emission from a variable compression ratio diesel engine

Prabhu Paramasivam[a,*], Khatir Naima[b], Marek Dzida[c]

[a] *Centre of Research Impact and Outcome, Chitkara University, Rajpura- 140401, Punjab, India*

[b] *Institute of Technology, University centre of Naama, Naama, Algeria*

[c] *Institute of Naval Architecture, Gdansk University of Technology, Gdansk, Poland.*

Abstract. Machine learning method and statistical method used for model prediction and optimization of third generation biodiesel-diesel blend powered variable compression engine High R2 values of 0.9998 and 0.9994 were observed in the training and testing phase of the model, respectively, indicating that The results confirm the robustness of the forecasting system. It was shown that the model accuracy means squared errors remained low at 0.0002 and 0.0014. These results were then confirmed by desirability-based optimization, which succeeded in achieving the values of the set parameters It should be noted that the compression ratio (CR), fuel injection pressure, and engine load were optimized to meet the defined parameters, resulting in a NOx emissions reduction as 222.8 ppm. The research illustrates the efficacy of desirability-based optimization in attaining targeted performance targets across important engine parameters whilst also reducing the impact on the environment.

Keywords: Machine learning; Alterative fuel; XGBoost; Desirability; Response surface methodology

Received: 12th March 2024; Revised: 5th April 2024; Accepted: 20th April 2024; Available online: 24th April 2024

1. Introduction

A concerning situation has arisen all across the world as a consequence of the fast depletion of conventional fuels and the subsequent damage to the environment that has resulted from this decline. It has virtually become a requirement to find a substitute fuel in order to bridge the gap between the demand for energy and energy availability (K. Ramalingam *et al.*, 2023). Sustainable development goal 7 categorically needs all-out efforts towards achieving a better effort towards mankind with economical and universal availability of energy at its core. The problem becomes more severe in the case of fossil fuel-powered compression ignition engines like diesel engines. The emission from these engines is being held responsible for global warming caused by greenhouse gases emitted by them. Herein, alternative fuels like biodiesel are attractive options (Mayer *et al.*, 2020; S. Ramalingam *et al.*, 2018).

The alternative fuel needs to have the capability of lowering the overall emissions that are brought up by the extensive utilization of liquid fuels that are derived from petroleum. Biodiesel, which is produced from sustainable biomass-based feedstocks such as vegetable oils and animal fats, is gaining popularity because it is kind to the environment and has fuel properties that are equivalent to those of fossil-based diesel (Arias *et al.*, 2021; Naik *et al.*, 2010; Pullagura *et al.*, 2024).

Animal fats, edible oils, non-edible oils, and microalgal oils may all be converted into biodiesel by the process of transesterification. However, because edible oils are consumed by humans, they are not regarded to be a potential feedstock for the production of biodiesel in India. Because local production is not adequate to meet the need for edible oil, more than 80 percent of its requirements are still imported from other countries of the world. In recent years, there has been a shift in focus towards non-edible oils, which continue to struggle with low productivity and sluggish growth rates, making them an impractical choice for the manufacture of biodiesel (Khan *et al.*, 2022; A. Sharma *et al.*, 2019; P. Sharma & Sharma, 2022).

When it comes to diesel engines, biodiesel has been employed in a great number of studies that have been published in the academic literature to explore engine performance, combustion features, and contaminants. Uyumaz *et al.* (Uyumaz *et al.*, 2020) researched to investigate the impact that opium poppy-derived methyl ester has on combustion, engine performance, emissions of nitrogen oxides, carbon monoxide, and soot at varying engine loads and speeds. As a result of their higher viscosity and density, biodiesel blends were found to have decreased in-cylinder exhaust temperature, increased in-cylinder pressure, and decreased in-cylinder exhaust pressure. It is important to note that opium poppy oil biodiesel significantly reduces emissions of carbon monoxide and soot, while marginally increasing emissions of nitrogen oxides. Solmaz *et al.* (Solmaz *et al.*, 2023) employed Response Surface Methodology (RSM) in order to enhance the features of diesel engines by using diesel, biodiesel, and multi wall carbon nanotubes (MWCNT) mixes. The purpose of this study was to explore the potential effects of MWCNT concentration and engine load on performance as well as exhaust pollutants. As the findings showed, increasing the load on the engine resulted in a rise in NOx

* Corresponding author
Email: lptprabhu@gmail.com (P. Paramasivama)

https://doi.org/10.61435/jese.2024.e20
ISSN: p-ISSN: 3026-0817; e-ISSN: 3026-0183.The Author(s). Published by CBIORE

emissions; however, an increase in the concentration of MWCNT results in a decrease in these emissions. The employment of MWCNTs results in a slight increase in BTE while simultaneously reducing BSFC.

Nanthagopal *et al.* (Nanthagopal *et al.*, 2017) in an investigation of a twin-cylinder diesel engine, the impact of adding nanoparticles of zinc oxide and titanium dioxide to biodiesel made from *Calophyllum inophyllum* was investigated. Through the use of an ultrasonication method, the nanofluids were created by combining distilled water with amounts of titanium dioxide and zinc oxide that were 50 and 100 parts per million, respectively. Using a mechanical stirrer, four distinct Calophyllum inophyllum nanoemulsions were created. The proportions of these nano emulsions were as follows: 93% of Calophyllum inophyllum biodiesel, 5% of nanofluids of zinc oxide and titanium dioxide, and 2% of span 80. The emission of nitrogen oxides was lower for all CIME nano emulsions than it was for pure *Calophyllum inophyllum* biodiesel, although it was somewhat higher than the emission of standard diesel fuel. The smoke emission of diesel engines was significantly decreased for all CIME nanoemulsions when compared to pure diesel and biodiesel fuels during all engine loads. This was the case regardless of the volume of the engine.

The literature reveals that the evaluation of a novel fuel as a substitute for fossil fuels may involve a large number of experiments, taking into consideration all of the features of the engine. The situation that results from taking into account both resources and time is inefficient. An increase in the effectiveness of the process of developing fuel maps has been achieved via the use of several software applications. Numerous simulations were carried out with an exceptional level of accuracy using a number of these computer programs, which resulted in a reduction in the total number of tests. One of these applications is called RSM which can help determine the ideal quantities of each factor and indicate which of the input variables that have been given are the most effective (P. Sharma & Sahoo, 2022; Sunil Kumar *et al.*, 2024). Besides this machine learning (ML) is an attractive approach that is often employed to assess the impact of every input variable upon the output variable. It makes use of data to understand underlying patterns in data to establish correlation among data. By using a hybrid strategy, it is possible to enhance the effectiveness of the process by concurrently employing ML model-based prediction and RSM to determine the best combination of any number of input variables (Silitonga *et al.*, 2018; Yang *et al.*, 2023; X. Zhang *et al.*, 2023).

Extreme Gradient Boosting (XGBoost), a sophisticated machine learning approach, was used in combination with Response Surface Methodology (RSM) in the present study in order to predict and optimize the amount of nitrogen oxides (NOx) emissions produced by a biodiesel-diesel powered engine. Researchers endeavor to improve the accuracy of emission forecasts and determine the ideal operating conditions. It is planned to combine the capabilities of RSM, which include optimization, with the capacity of XGBoost to handle complicated connections in data. This technique not only provides insights into the reduction of harmful emissions, but also demonstrates the synergy that exists between ML and statistical optimization approaches in the context of addressing environmental concerns within the realm of renewable energy.

2. Materials and methods

2.1 Test fuel and set up

In the present study, *Chlorella protothecoides* microalgal oil was employed for the production of biodiesel through the process of transesterification. A filter was employed for microalgal oil to eliminate the insoluble contaminants and then heated for 10 minutes maintaining the 100 °C temperature to reduce moisture from oil. A mechanical stirrer using the appropriate amounts of CH_3OH and KOH. The stated settings, which comprised a molar ratio of methanol vs oil as 7.12 (vol/vol), for 115 minutes, a reaction temperature of 64 °Celsius, and a concentration of catalyst as 1.019 (% wt/vol).

To complete the reaction, the mixture was heated for a predetermined amount of time at the specified temperature. Biodiesel generated using the aforementioned approach is utilized to power a computer-linked single-cylinder, 4-stroke, diesel engine (shown in **Fig. 1**). **Table 1** shows the exact specs for the test rig. An eddy current dynamometer was employed for engine loading and variation. The load sensors, mounted on a dynamometer, were used to transmit load indications to the crankshaft. Two piezometer sensors were used to quantify the combustion and fuel line pressures.

Table 1
Fuel characteristics of Chlorella Protothecoides microalgal biodiesel

Characteristics	*Chlorella protothecoides* microalgal biodiesel	Diesel
Viscosity @ 40 °C	4.41 cSt	3.029 cSt
Density	0.859 gm/cm^3	0.837 gm/cm^3
Flash point	166 °C	75 °C

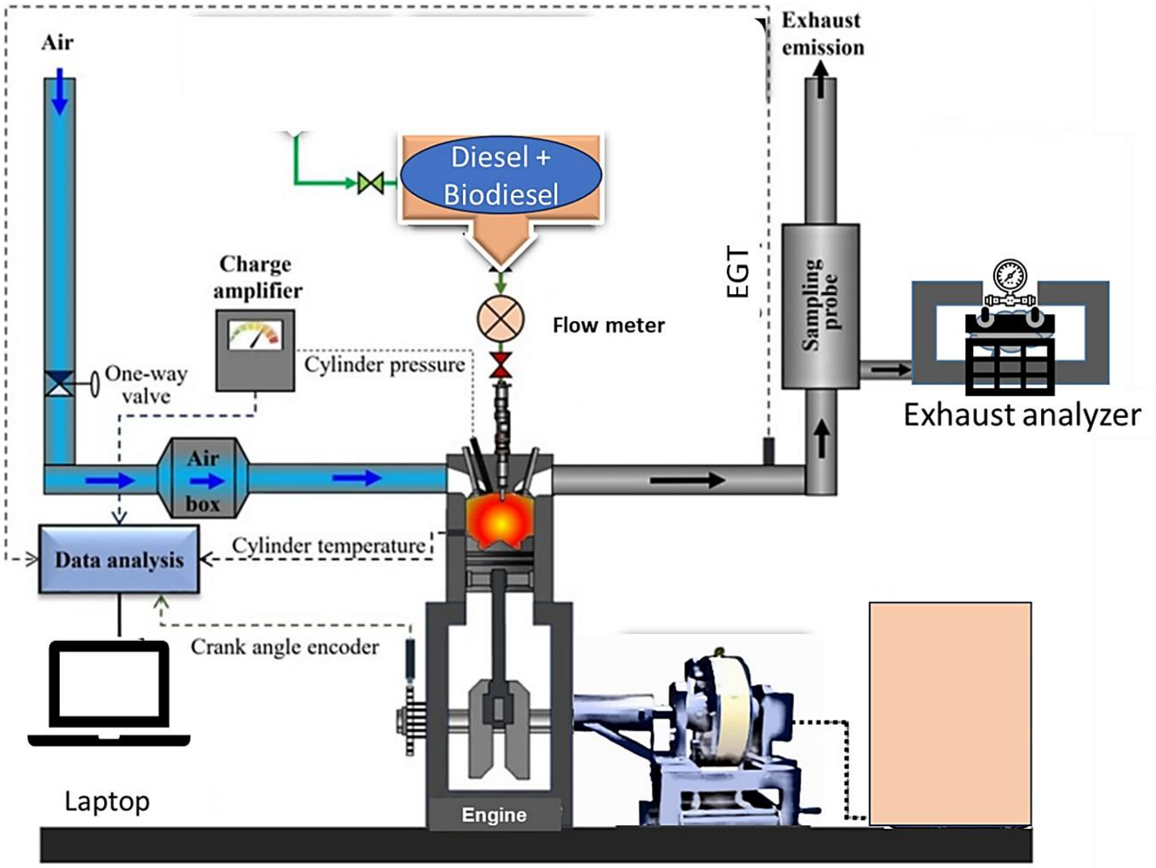


Fig. 1 Engine testing set up

2.2 Testing procedure

In this study, a variable compression engine was used so that different compression ratio (CR) can be used for fuel testing. The fuel injection pressure was also varied to investigate the effects on NOx emission. The engine load was varied with the help of eddy current dynamometer to explore the effects of engine load in NOx emission. The engine was antically run for 15 minutes to stabilize the engine cooling and lubrication temperature and the engine was run at different settings by varying load, CR and fuel injection pressure. The NOx data was recorded at different settings.

2.3 Soft computing

2.3.1. Response surface methodology

RSM is a collection of computational and mathematical methodologies that are applied to define and evaluate the relationships between a set of independent factors and a number of dependent variables (responses) in an environment (Nazarpour *et al.*, 2022). Numerous applications of the RSM approach may be found in the fields of experimental design, optimization of processes, and quality improvement. It is possible that RSM will be of particular use in the Analysis of Variance (ANOVA), and desirability-based optimization procedures when it relates to the context of engine emissions data (P. Sharma & Sharma, 2021, 2022).

Experiments that are aimed to systematically study the effects of different engine parameters (like fuel blends, injection pressure, and timing, air-fuel ratio, etc.) on emissions data (e.g. as NOx) may be planned with the use of RSM, which can be of assistance in the process of preparing successful experiments. When researchers make use of RSM, they are able to find the components that have the most influence as well as the optimal values for those factors, all while reducing the number of trials that are necessary. In the process of assessing the experimental data and establishing the significance of each component as well as the mechanisms by which they interact with other components in connection with engine emissions, the simultaneous use of RSM and ANOVA may be of great assistance. Because of this, researchers are able to identify the factors that are having a statistically significant impact on emissions, and they can then use this information to prioritize those components for future improvement (Pereira *et al.*, 2021; Said *et al.*, 2023). It is also possible to combine RSM with desirability functions with the goal of optimizing many engine parameters simultaneously while taking into account conflicting goals while doing so. It is the target values or variations that are provided that define the desirability score that is awarded to each response variable by desirability algorithms. This score is decided by the information that is provided. Desirability functions lead the way to optimization by discovering the optimal combination of parts that match the needed criteria. RSM is beneficial in the process of generating forecasting models for the responses, and desirability functions are also helpful in the process of optimizing (Harrington, 1965; Jain *et al.*, 2023; Lei *et al.*, 2019).

2.3.2. Extreme gradient boosting

Extreme Gradient Boosting (XGBoost) regression is a cutting-edge machine learning approach that has completely transformed the field of predictive analytics. XGBoost is an algorithm that has exceptional performance in a number of data science tasks and applications in the real world, providing speed and accuracy that have never been seen before. Both gradient boosting and ensemble learning are the driving forces behind it. XGBoost's superiority may be traced back to its forward-thinking approach to decision trees and boosting. The XGBoost algorithm makes use of a one-of-a-kind gradient boosting framework that places equal importance on mathematical precision and computing efficiency. Through the successive generation of multiple weak learners (decision trees) and the repeated revision of their predictions, XGBoost can construct a model that is both incredibly exact and resilient, surpassing the capabilities of a great number of previous methods (Chen & Guestrin, 2016; Qiu *et al.*, 2022).

The capability of XGBoost to manage large data sets that include a high level of complexity is one of its most notable strengths. The use of sophisticated normalization algorithms and concurrent processing enables XGBoost to handle millions of features and observations in a short amount of time. This makes it an excellent option for organizations that deal with enormous amounts of data, such as those in the fields of biology, e-commerce, and banking operations. The fact that XGBoost is completely transparent and easy to understand is yet another important aspect of this program. By providing a clear view of feature significance and commitment to the final forecast, XGBoost assists data analysts and domain experts in gaining important insights into the underlying properties of the data. This is accomplished by providing a clear perspective on the data. This interpretability is beneficial not just to the trustworthiness of the model but also to the design of features and decision-making that is driven by data (Amjad *et al.*, 2022; Pan *et al.*, 2022; P. Zhang *et al.*, 2022).

3. Results and discussion

3.1 Data analysis

The data gathered in the engine testing at different operational setting was used for model development. The data was also employed for plotting of correlational heatmap and data variation as depicted in **Figs. 2a** and **2b**, respectively. It can be observed that there is a strong positive correlation between engine load and NOx emission. The correlation between these two is strong that it makes other factors negligible.

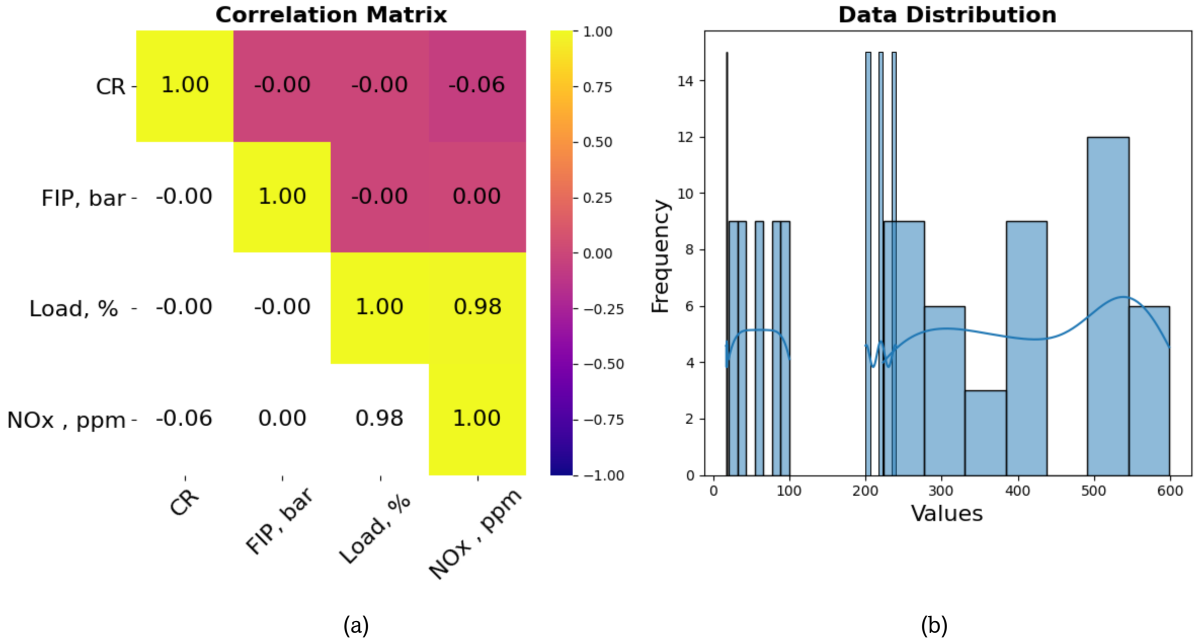


Fig. 2 NOx data (a) correlational heatmap; (b) distribution map

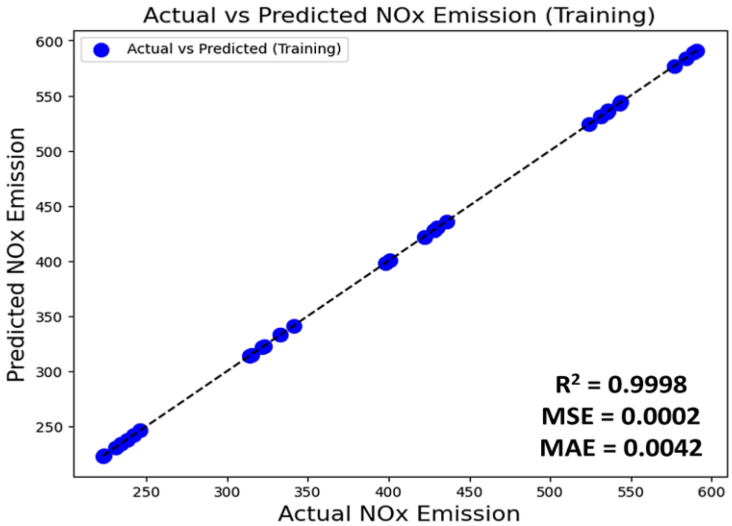
3.2 Model development

The python based open source libraries were employed for model development. The following pseudo code was used in this work

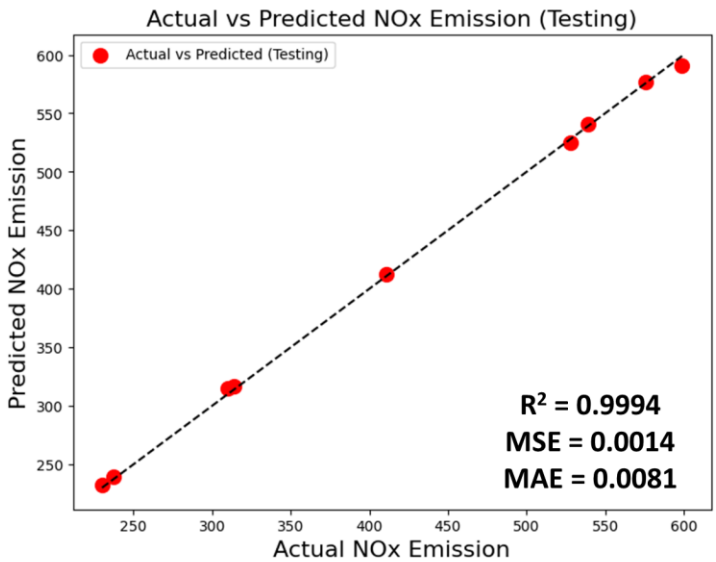
- Import necessary libraries: pandas, xgboost, train_test_split, r2_score, mean_squared_error, mean_absolute_error, and matplotlib.pyplot.
- Load data using pandas read_excel function and store it in a variable called data.
- Split the data into predictor variables (X) and response variable (y) by selecting appropriate columns from the data.
- Split the data into training and testing sets using train_test_split function from sklearn, with a test size of 0.2 and random state set to 42.

- Initialize an XGBoost model (XGBRegressor).
- Train the XGBoost model on the training data (X_train, y_train) using the fit method.
- Use the trained model to make predictions on both the training and testing data (X_train and X_test) using the predict method, and store the predictions in y_train_pred and y_test_pred respectively.
- Calculate performance metrics - R^2, MSE, and MAE - for both training and testing predictions using appropriate functions from sklearn.
- Define marker size, font size, and font color for the plots.
- Plot the actual vs. predicted values for the training set using scatter plot with blue color and label 'Actual vs Predicted (Training)', with annotations for performance metrics.
- Plot the 1:1 line on the plot to visually assess the fit of the model.
- Show the legend and display the plot.
- Plot the actual vs. predicted values for the testing set using scatter plot with red color and label 'Actual vs Predicted (Testing)', with annotations for performance metrics.
- Plot the 1:1 line on the plot to visually assess the fit of the model.
- Show the legend and display the plot.

The results of models during model training as well as testing are deputed in **Fig. 3a** and **3b**, respectively. The graphical presentation shows a robust prediction framework in both cases. The R^2 values were 0.9998 and 0.9994 during the phases of model training and testing. In both cases the mean squared error was on lower side as 0.0002 and 0.0014 for model training and testing, correspondingly (Wang *et al.*, 2020; M. Zhang *et al.*, 2022).



(a)



(b)

Fig. 3 Model performance during (a) training and (b) testing phases

3.3 Response surface

The data acquired using engine testing was used for model development and optimization using design expert software-based ANOVA and desirability-based optimization. The ANOVA outcomes for the test data are listed in **Table 2**.

Table 2
ANOVA results

Source	SUM of Squares	df	'F' Value	'p' value (Prob > F)	
Model	712559.3	9	284.6111	< 0.0001	significant
CR 'C'	2218.8	1	7.976116	0.0078	
FIP 'F'	5.633333	1	0.020251	0.8877	
Load 'L'	700484.4	1	2518.093	< 0.0001	
CF	26.45	1	0.095082	0.7596	
CL	2912.067	1	10.46826	0.0027	
CF	0.016667	1	5.99E-05	0.9939	
C^2	1026.844	1	3.691288	0.0629	
F^2	780.2778	1	2.804933	0.1029	
L^2	5104.794	1	18.35065	0.0001	
error	9736.317	35			
Cor Total	722295.6	44			

The relevance of the model is shown by the Model F-value, which is 284.61. This figure suggests that there's only a 0.01% possibility that such a huge "Model F-Value" may be the result of noise. When the value of "Prob > F" is less than 0.0500, it indicates that the model terms are significant. For the sake of this discussion, the words A, C, AC, and C2 are regarded as being relevant. Values that are greater than 0.1000, on the other hand, indicate insignificance. When there are a large number of model terms that are not significant (with the exception of those that are required for hierarchy), decreasing the model may improve its effectiveness. The model has a high level of explanatory power, as seen by its standard deviation value of 16.68 and its R-squared value of 0.9865. The adjusted R-squared value of 0.9831 and the predicted R-squared value of 0.9782 are in reasonable agreement, which indicates that the model is valid. Having a signal-to-noise ratio that is sufficient for efficient exploration of the design space is shown by the fact that the Adeq. A precision of 49.638 is higher than the intended threshold of 4 (Dimitriou *et al.*, 2013; Srinidhi *et al.*, 2021; Thodda *et al.*, 2023).

Fig. 4 depicts the results of the RSM base analysis. **Fig. 4a** depicts the measured vs model-predicted values of NOx results. It can be observed that most of the data points lie on the best-fit line. **Figs. 4b** and **4c** depict the 2-D FIP vs CR contour plots and surface diagrams, respectively. **Fig. 4d** Similarly, the contour plot of CR vs engine load indicates how these factors affect NOx emissions. **Fig. 4e** the surface plot of CR against load provides a more thorough view of NOx emission patterns when both parameters fluctuate. **Fig. 4f** The contour plot of FIP vs engine load illustrates the interaction of these factors and their impact on NOx emissions. **Fig. 4g** Finally, the surface diagram of FIP against load offers a full perspective of NOx emission patterns influenced by both factors. Together, these visualizations provide light on the complicated interactions between engine operating parameters and NOx emissions, assisting in the optimization of combustion processes for decreased environmental effects.

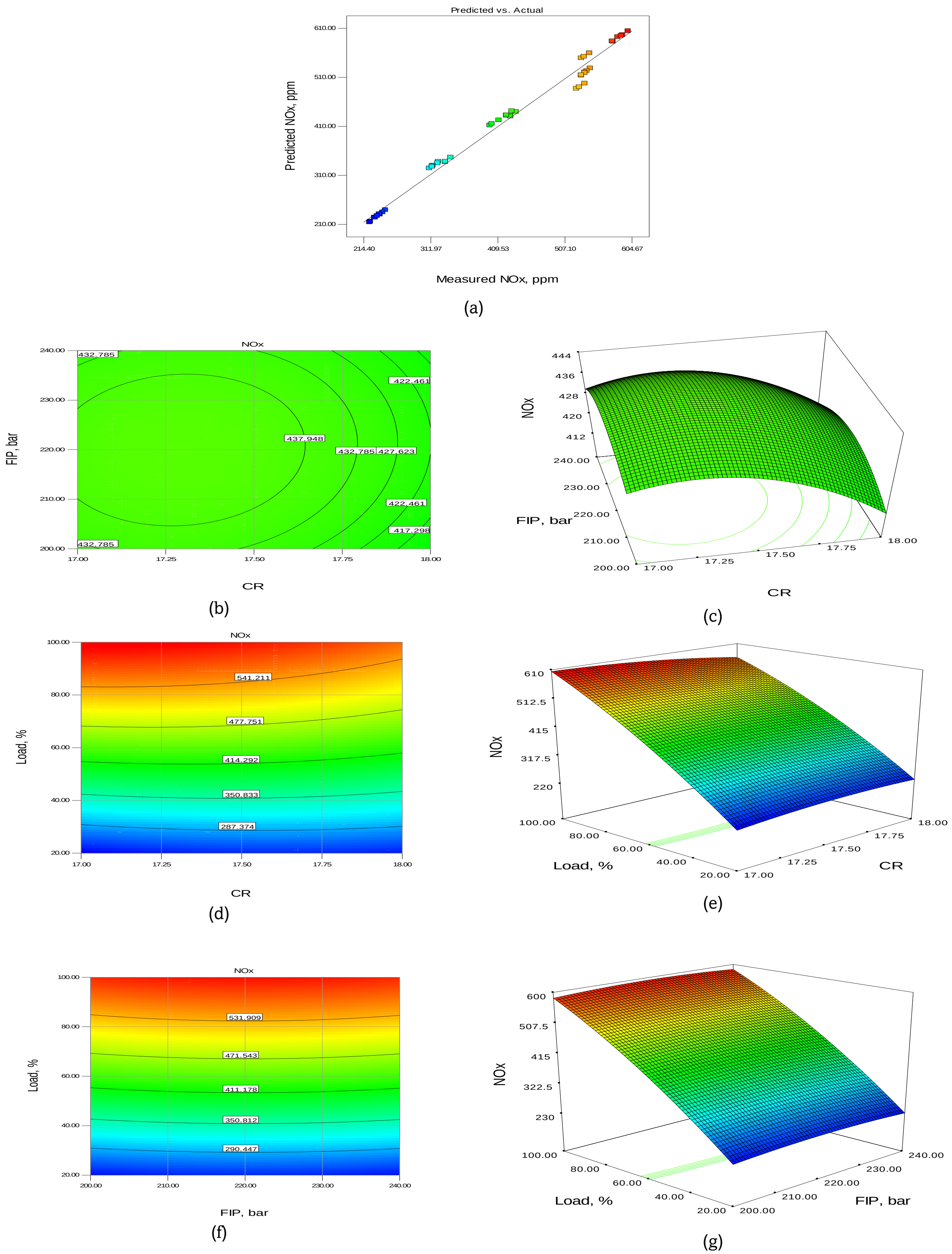


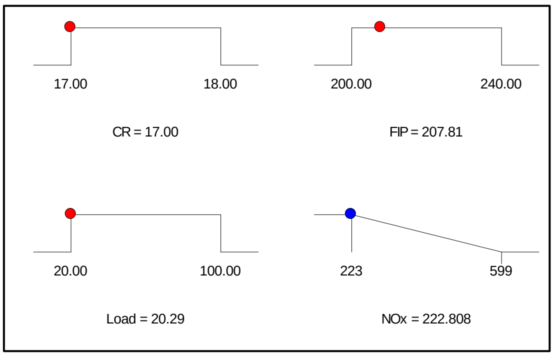
Fig. 4 RSM based NOx mdoel based (a) actual vs predictd; (b) CR vs FIP contour plot; (c) CR vs FIP surafce diagram; (d) CR vs load contour plot; (e) CR vs load surafce diagram; (f) FIP vs laod contour plot; (g) FIP vs load surface diagram

3.4 Desirability based optimization

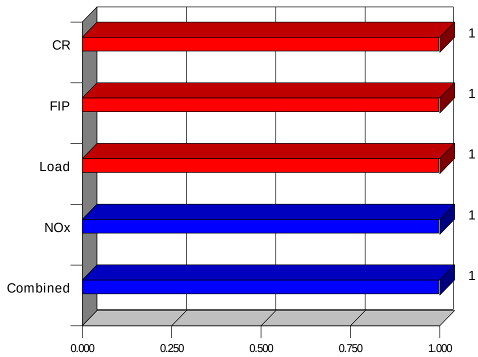
The outcomes of desirability-based optimization show that the objectives that were set were achieved within the parameter ranges that were established. This demonstrates that appropriate parameter tuning was performed in order to achieve improved performance. The Compression Ratio (CR) has been effectively kept within the required range of 17 to 18, which aligns with the assigned significance level of 3, as listed in **Table 3**. This is one of the metrics that has been optimized. Similar to the previous example, the Fuel Injection Pressure (FIP), which is essential for the efficiency of the engine, has been adjusted to 207.81 bar. This value falls within the planned range of 200 to 240 bar and has a high significance level of 3. The load on the engine, which is critical for maintaining operating stability, has been adjusted to 20.29 percent, which satisfies the stated range of 20 percent to 100 percent with equal meaning. Notably, the optimization procedure has also successfully reduced NOx emissions to 222.8 ppm, which is much more than the intended range of 223 to 599 ppm. This is of the highest significance. These findings provide further evidence that desirability-based optimization is an effective method for reaching desired performance objectives across crucial engine parameters while also reducing detrimental effects on the environment (Harrington, 1965; Kumar *et al.*, 2022; Lei *et al.*, 2019). The ramp diagram is depicted in **Fig. 5a** while the bar plot is shown in **Fig. 5b**.

Table 3
Optimized setting and results

Name	Goal	Lower limit	Upper limit	Importance	Optimized value
CR	is in range	17	18	3	17
FIP, bar	is in range	200	240	3	207.81
Load, %	is in range	20	100	3	20.29
NOx, ppm	minimize	223	599	1	222.8



(a)



(b)

Fig. 5 Desirability based (a) Ramp plot; (b) bar plot

4. Conclusion

In conclusion, employing a Chlorella Protothecoides microalgal biodiesel-diesel mix to power a VCR engine, together with the integration of XGBoost and RSM for model prediction and optimization, produced encouraging results. A robust prediction framework was developed with XGBoost. It achieved a R^2 values of 0.9998 and 0.9994 during training and testing, respectively. In addition, the MSE was low at 0.0002 and 0.0014 throughout the model training and testing stages, demonstrating the precision of models. RSM-based desirability approach for optimization confirmed these results by successfully achieving predefined targets within defined parameter ranges. Notably, Compression Ratio (CR) was successfully maintained within the needed range of 17 to 18, while Fuel Injection Pressure (FIP) and engine load were adjusted to 207.81 bar and 20.29%, respectively, in accordance with the prescribed ranges and importance levels. Most impressively, NOx emissions were decreased to 222.8 ppm, proving the efficiency of the optimization approach in reducing environmental impact. The ramp diagram and the bar plot give further information about the optimization process. Overall, our results demonstrate the effectiveness of desirability-based optimization in meeting targeted performance goals across crucial engine parameters while also addressing environmental issues.

References

Amjad, M., Ahmad, I., Ahmad, M., Wróblewski, P., Kamiński, P., & Amjad, U. (2022). Prediction of Pile Bearing Capacity Using XGBoost Algorithm: Modeling and Performance Evaluation. *Applied Sciences*, *12*(4), 2126. https://doi.org/10.3390/app12042126

Arias, D. M., Ortíz-Sánchez, E., Okoye, P. U., Rodríguez-Rangel, H., Balbuena Ortega, A., Longoria, A., Domínguez-Espíndola, R., & Sebastian, P. J. (2021). A review on cyanobacteria cultivation for carbohydrate-based biofuels: Cultivation aspects, polysaccharides accumulation strategies, and biofuels production scenarios. *Science of The Total Environment*, *794*, 148636. https://doi.org/10.1016/j.scitotenv.2021.148636

Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Dimitriou, P., Peng, Z., Lemon, D., Gao, B., & Soumelidis, M. (2013, September 8). *Diesel Engine Combustion Optimization for Bio-Diesel Blends Using Taguchi and ANOVA Statistical Methods*. https://doi.org/10.4271/2013-24-0011

Harrington, E. C. (1965). The Desirability Function. *Industrial Quality Control*, *21*, 494–498.

Jain, A., Bora, B. J., Kumar, R., Sharma, P., Deepanraj, B., Irshad, K., & Ravikiran, C. (2023). Application of hybrid Taguchi L16 and desirability for model prediction and optimization in assessment of the performance of a novel Water Hyacinth biodiesel run diesel engine. *Fuel*, *339*, 127377. https://doi.org/10.1016/j.fuel.2022.127377

Khan, M. M., Kadian, A. K., & Sharma, R. P. (2022). Attempt to mitigate marine engine emissions with improved performance by the investigation of alcohol inclusion in sunflower biodiesel-sunflower oil-diesel blend. *Environmental Science and Pollution Research*, *30*(12), 33974–33991. https://doi.org/10.1007/s11356-022-24147-6

Kumar, A. N., Ashok, B., Nanthagopal, K., Ong, H. C., Geca, M. J., Victor, J., Vignesh, R., Jeevanantham, A. K., Kannan, C., & Kishore, P. S. (2022). Experimental analysis of higher alcohol–based ternary biodiesel blends in CI engine parameters through multivariate and desirability approaches. *Biomass Conversion and Biorefinery*, *12*(5), 1525–1540. https://doi.org/10.1007/s13399-020-01134-w

Lei, X., Shuang, J., Yang, P., & Liu, Y. (2019). Parametric study and optimization of dimpled tubes based on Response Surface Methodology and desirability approach. *International Journal of Heat and Mass Transfer*, *142*, 118453. https://doi.org/10.1016/j.ijheatmasstransfer.2019.118453

Mayer, F. D., Brondani, M., Vasquez Carrillo, M. C., Hoffmann, R., & Silva Lora, E. E. (2020). Revisiting energy efficiency, renewability, and sustainability indicators in biofuels life cycle: Analysis and standardization proposal. *Journal of Cleaner Production*, *252*, 119850. https://doi.org/10.1016/j.jclepro.2019.119850

Naik, S. N., Goud, V. V., Rout, P. K., & Dalai, A. K. (2010). Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, *14*(2), 578–597. https://doi.org/10.1016/j.rser.2009.10.003

Nanthagopal, K., Ashok, B., Tamilarasu, A., Johny, A., & Mohan, A. (2017). Influence on the effect of zinc oxide and titanium dioxide nanoparticles as an additive with Calophyllum inophyllum methyl ester in a CI engine. *Energy Conversion and Management*, *146*, 8–19. https://doi.org/10.1016/j.enconman.2017.05.021

Nazarpour, M., Taghizadeh-Alisaraei, A., Asghari, A., Abbaszadeh-Mayvan, A., & Tatari, A. (2022). Optimization of biohydrogen production from microalgae by response surface methodology (RSM). *Energy*, *253*, 124059. https://doi.org/10.1016/j.energy.2022.124059

Pan, S., Zheng, Z., Guo, Z., & Luo, H. (2022). An optimized XGBoost method for predicting reservoir porosity using petrophysical logs. *Journal of Petroleum Science and Engineering*, *208*, 109520. https://doi.org/10.1016/j.petrol.2021.109520

Pereira, L. M. S., Milan, T. M., & Tapia-Blácido, D. R. (2021). Using Response Surface Methodology (RSM) to optimize 2G bioethanol production: A review. *Biomass and Bioenergy*, *151*, 106166. https://doi.org/10.1016/j.biombioe.2021.106166

Pullagura, G., Vanthala, V. S. P., Vadapalli, S., Bikkavolu, J. R., Barik, D., Sharma, P., & Bora, B. J. (2024). Enhancing performance characteristics of biodiesel-alcohol/diesel blends with hydrogen and graphene nanoplatelets in a diesel engine. *International Journal of Hydrogen Energy*, *50*, 1020–1034. https://doi.org/10.1016/j.ijhydene.2023.09.313

Qiu, Y., Zhou, J., Khandelwal, M., Yang, H., Yang, P., & Li, C. (2022). Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. *Engineering with Computers*, *38*(S5), 4145–4162. https://doi.org/10.1007/s00366-021-01393-9

Ramalingam, K., Vellaiyan, S., Venkatesan, E. P., Khan, S. A., Mahmoud, Z., & Saleel, C. A. (2023). Challenges and Opportunities of Low Viscous Biofuel—A Prospective Review. *ACS Omega*, *8*(19), 16545–16560. https://doi.org/10.1021/acsomega.3c00387

Ramalingam, S., Rajendran, S., Ganesan, P., & Govindasamy, M. (2018). Effect of operating parameters and antioxidant additives with biodiesels to improve the performance and reducing the emissions in a compression ignition engine – A review. *Renewable and Sustainable Energy Reviews*, *81*, 775–788. https://doi.org/10.1016/j.rser.2017.08.026

Said, Z., Sharma, P., Bora, B. J., Nguyen, V. N., Bui, T. A. E., Nguyen, D. T., Dinh, X. T., & Nguyen, X. P. (2023). Modeling-optimization of performance and emission characteristics of dual-fuel engine powered with pilot diesel and agricultural-food waste-derived biogas. *International Journal of Hydrogen Energy*, *48*(18), 6761–6777. https://doi.org/10.1016/j.ijhydene.2022.07.150

Sharma, A., Ansari, N. A., Pal, A., Singh, Y., & Lalhriatpuia, S. (2019). Effect of biogas on the performance and emissions of diesel engine fuelled with biodiesel-ethanol blends through response surface methodology approach. *Renewable Energy*, *141*, 657–668. https://doi.org/10.1016/j.renene.2019.04.031

Sharma, P., & Sahoo, B. B. (2022). An ANFIS-RSM based modeling and multi-objective optimization of syngas powered dual-fuel engine. *International Journal of Hydrogen Energy*, *47*(44), 19298–19318. https://doi.org/10.1016/j.ijhydene.2022.04.093

Sharma, P., & Sharma, A. K. (2021). AI-Based Prognostic Modeling and Performance Optimization of CI Engine Using Biodiesel-Diesel Blends. *International Journal of Renewable Energy Research*, *v11i2*. https://doi.org/10.20508/ijrer.v11i2.11854.g8191

Sharma, P., & Sharma, A. K. (2022). Statistical and Continuous Wavelet Transformation-Based Analysis of Combustion Instabilities in a Biodiesel-Fueled Compression Ignition Engine. *Journal of Energy Resources Technology*, *144*(3). https://doi.org/10.1115/1.4051340

Silitonga, A. S., Masjuki, H. H., Ong, H. C., Sebayang, A. H., Dharma, S., Kusumo, F., Siswantoro, J., Milano, J., Daud, K., Mahlia, T. M. I., Chen, W.-H., & Sugiyanto, B. (2018). Evaluation of the engine performance and exhaust emissions of biodiesel-bioethanol-diesel blends using kernel-based extreme learning machine. *Energy*, *159*, 1075–1087. https://doi.org/10.1016/j.energy.2018.06.202

Solmaz, H., Calam, A., Yılmaz, E., Şahin, F., Ardebili, S. M. S., & Aksoy, F. (2023). Evaluation of MWCNT as fuel additive to diesel–biodiesel blend in a direct injection diesel engine. *Biofuels*, *14*(2), 147–156. https://doi.org/10.1080/17597269.2022.2122154

Srinidhi, C., Madhusudhan, A., Channapattana, S. V., Gawali, S. V., & Aithal, K. (2021). RSM based parameter optimization of CI engine fuelled with nickel oxide dosed Azadirachta indica methyl ester. *Energy*, *234*, 121282. https://doi.org/10.1016/j.energy.2021.121282

Sunil Kumar, K., Surakasi, R., Patro, S. G. K., Govil, N., Ramis, M. K., Razak, A., Sharma, P., Alsubih, M., Islam, S., Khan, T. M. Y., Almakayeel, N., & Chintakindi, S. (2024). Performance, Combustion, and Emission analysis of diesel engine fuelled with pyrolysis oil blends and n-propyl alcohol-RSM optimization and ML modelling. *Journal of Cleaner Production*, *434*, 140354. https://doi.org/10.1016/j.jclepro.2023.140354

Thodda, G., Madhavan, V. R., & Thangavelu, L. (2023). Predictive Modelling and Optimization of Performance and Emissions of Acetylene Fuelled

CI Engine Using ANN and RSM. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, *45*(2), 3544–3562. https://doi.org/10.1080/15567036.2020.1829191

Uyumaz, A., Aydoğan, B., Yılmaz, E., Solmaz, H., Aksoy, F., Mutlu, İ., İpci, D., & Calam, A. (2020). Experimental investigation on the combustion, performance and exhaust emission characteristics of poppy oil biodiesel-diesel dual fuel combustion in a CI engine. *Fuel*, *280*, 118588. https://doi.org/10.1016/j.fuel.2020.118588

Wang, S., Liu, S., Zhang, J., Che, X., Yuan, Y., Wang, Z., & Kong, D. (2020). A new method of diesel fuel brands identification: SMOTE oversampling combined with XGBoost ensemble learning. *Fuel*, *282*, 118848. https://doi.org/10.1016/j.fuel.2020.118848

Yang, Y., Gao, L., Abbas, M., Elkamchouchi, D. H., Alkhalifah, T., Alturise, F., & Ponnore, J. J. (2023). Innovative composite machine learning approach for biodiesel production in public vehicles. *Advances in Engineering Software*, *184*, 103501. https://doi.org/10.1016/j.advengsoft.2023.103501

Zhang, M., Chen, W., Yin, J., & Feng, T. (2022). Lithium Battery Health Factor Extraction Based on Improved Douglas–Peucker Algorithm and SOH Prediction Based on XGboost. *Energies*, *15*(16), 5981. https://doi.org/10.3390/en15165981

Zhang, P., Jia, Y., & Shang, Y. (2022). Research and application of XGBoost in imbalanced data. *International Journal of Distributed Sensor Networks*, *18*(6), 155013292211069. https://doi.org/10.1177/15501329221106935

Zhang, X., Li, H., Sekar, M., Elgendi, M., Krishnamoorthy, N. R., Xia, C., & Priya Matharasi, D. (2023). Machine learning algorithms for a diesel engine fuelled with biodiesel blends and hydrogen using LSTM networks. *Fuel*, *333*, 126292. https://doi.org/10.1016/j.fuel.2022.126292