**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

# PREDICTION OF BIODIESEL YIELD FROM THE TRANSESTERIFICATION OF RUBBER SEED OIL WITH MACHINE LEARNING

**Ifeanyichukwu Edeh[1] and Nkwelle Daniel Arinze[1]**
[1]Department of Chemical Engineering, University of Port Harcourt, Nigeria

**Abstract:** Biodiesel has emerged as a promising renewable alternative to conventional diesel fuel, offering environmental advantages and supporting global sustainability efforts. This study applied machine learning techniques to predict biodiesel yield from rubber seed oil, a non-edible oil and locally available feedstock in Nigeria. A dataset of 20 experimental runs was used, with methanol-to-oil molar ratio, catalyst weight, temperature, and reaction time as input variables. Three models, Decision Tree Regressor (DTR), Random Forest Regressor (RFR), and Gradient Boosting Regressor (GBR) were developed and evaluated using R-squared and Mean Squared Error (MSE). The results obtained show that the square of the coefficient of regression ($R^2$) for the DTR, RFR, and GBR were 0.7900, 0.8612, and 0.9937, respectively. The MSE for the DTR, RFR, and GBR were 13.64, 8.97, and 0.40, respectively. The Gradient Boosting Regressor performed best, showing the highest predictive accuracy for the biodiesel yield of 75.32 % at the optimum conditions of temperature (30 °C), time (75 min), catalyst loading (0.5g), methanol-to-oil ratio (4:1), and weight of methanol (10 g). The results revealed that the methanol-to-oil molar ratio had the most significant influence on biodiesel yield, with yields increasing as the ratio improved within optimal limits. The results demonstrate that machine learning can offer a cost-effective and time-saving alternative to labor-intensive experimental methods, thereby improving the ability to predict and optimize biodiesel production. The ensemble-based learning models, particularly GBR, have the potential to be reliable tools for biodiesel yield prediction and support their integration into renewable energy system development frameworks.

**Keywords:** Biodiesel, Machine Learning, Decision Tree Regressor, Random Forest Regressor, Gradient Boosting Regressor, Model Evaluation Metrics

## 1.0 Introduction

The global shift towards cleaner and more sustainable energy sources has intensified interest in biodiesel as a promising alternative to fossil-based diesel fuels. Biodiesel is renewable, biodegradable, and environmentally friendly, offering significant potential to reduce greenhouse gas emissions and enhance energy security (Demirbas, 2009; Atabani *et al*., 2012). Among various non-edible feedstocks explored for biodiesel production, rubber seed oil has gained attention due to its high oil yield, local abundance in tropical regions such as Nigeria's Niger Delta, and its non-competition with food crops (Odetoye *et al.,* 2020). Despite these advantages, the production of biodiesel from rubber seed oil is influenced by a set of complex process parameters including reaction temperature, catalyst concentration, methanol-to-oil molar ratio, and reaction time which must be carefully optimized to ensure high yields and operational efficiency (Mekhilef *et al.,* 2011). Conventional optimization methods such as one-factor-at-a-time experiments or response surface methodology (RSM) are often labor-intensive, time-consuming, and costly, particularly when experimental trials are needed for multiple process combinations (Farobie *et al.,* 2015). This challenge is exacerbated in developing regions where laboratory infrastructure and research funding are limited. To address these limitations, recent studies have explored data-driven approaches using machine learning (ML) techniques to model and predict

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

biodiesel yields under varying process conditions (Chicco *et al.,* 2021). These models offer faster, scalable alternatives to empirical optimization by learning complex patterns from data. However, the effectiveness of such models depends on multiple factors, including the quality and quantity of available data, model interpretability, and their adaptability to regional feedstocks and conditions. In Nigeria, the potential of rubber seed oil for biodiesel production remains underexploited, in part due to several interrelated challenges. These include the scarcity of comprehensive regional datasets that capture the physicochemical properties of rubber seed oil and associated biodiesel yields (Adewole *et al*., 2018), uncertainty over the most suitable ML models due to varying performance and interpretability (Sahu & Agarwal, 2020), difficulties in validating predictive models under variable environmental conditions (Mohanty *et al.,* 2019), and the relatively high cost of rubber seed oil which undermines economic viability (Oladosu *et al.,* 2017). Furthermore, the majority of existing biodiesel modeling studies focus on feedstocks such as soybean, palm, or jatropha oils, which are either edible or not indigenous to regions like Nigeria (Atabani *et al.,* 2013; Odetoye *et al*., 2020). While techniques like artificial neural networks (ANN), support vector machines (SVM), and genetic algorithms have demonstrated strong predictive power, they often require large datasets and lack transparency in their decision-making process, which limits their usability in data-constrained environments (Mohanty *et al.,* 2019; Sahu & Agarwal, 2020). Tree-based ensemble models such as Random Forest, Gradient Boosting, and Decision Trees have emerged as promising alternatives due to their interpretability, lower data requirements, and strong performance in non-linear regression tasks (Chicco *et al.,* 2021; Breiman, 2001; Friedman, 2001). Thus, this study investigates the use of interpretable machine learning models such as Random Forest Regressor, Gradient Boosting Regressor, and Decision Tree Regressor to predict and optimize biodiesel yield from rubber seed oil using experimental data. The performance of the models was evaluated using both statistical metrics ($R^2$, MSE) and visual analysis.

## 2.0 Materials and Methods

### 2.1. Materials

#### 2.1.1 Data collection

The dataset was obtained from the optimization of the transesterification of rubber seed oil to produce biodiesel using rubber seed oil as a feedstock conducted by Edeh (2025). A total of 20 experimental data points were obtained through Design of Experiment using Central Composite Design (see Table 1).

**Table 1. Optimization of biodiesel yield using Response Surface Methodology (Edeh, 2025)**

| *Run* | Temp (°C) | Time (min) | Catalyst loading (g) | Methanol/ Oil ratio | Biodiesel yield (%) |
|---|---|---|---|---|---|
| *1* | 30 | 120 | 0.5 | 6/1 | 72.04 |
| *2* | 45 | 120 | 0.5 | 4/1 | 66.1 |
| *3* | 45 | 75 | 0.5 | 4/1 | 51.3 |
| *4* | 45 | 75 | 0.5 | 4/1 | 44.2 |
| *5* | 60 | 120 | 0.5 | 6/1 | 72.9 |
| *6* | 45 | 75 | 0.5 | 4/1 | 67 |
| *7* | 30 | 75 | 0.5 | 4/1 | 76.6 |
| *8* | 30 | 75 | 0.5 | 6/1 | 58.2 |

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

| | | | | | |
|---|---|---|---|---|---|
| *9* | 60 | 30 | 0.5 | 6/1 | 52.7 |
| *10* | 30 | 75 | 0.5 | 4/1 | 57.7 |
| *11* | 45 | 30 | 0.5 | 2/1 | 61.5 |
| *12* | 30 | 120 | 0.5 | 2/1 | 71 |
| *13* | 45 | 75 | 0.5 | 4/1 | 60.4 |
| *14* | 45 | 30 | 0.5 | 2/1 | 55.4 |
| *15* | 60 | 120 | 2.0 | 16/1 | 54.7 |
| *16* | 45 | 75 | 24.0 | 24/1 | 60.05 |
| *17* | 60 | 120 | 20.0 | 20/1 | 70.12 |
| *18* | 45 | 75 | 20.0 | 20/1 | 64.2 |
| *19* | 60 | 30 | 8.0 | 8/1 | 60.83 |
| *20* | 45 | 75 | 24.0 | 24/1 | 58.46 |

**2.2 Methods**

2.2.1 Data Preprocessing

The data preprocessing steps were performed to prepare the data for model development. These included cleaning for missing values and inconsistencies, followed by standardization (mean = 0, standard deviation = 1) to ensure that each input feature contributed proportionally to the learning algorithms. Outliers were identified and removed to reduce the risk of biasing the model training phase (Han & Pei, 2012). The variables used included:

(1). X1 = X_1 =X1= Temperature (°C)
(2). X2 = X_2 =X2= Reaction time (min)
(3). X3 = X_3 =X3= Catalyst loading (g)
(4). X4 = X_4 =X4= Methanol-to-oil molar ratio
(5). X5 = X_5 =X5= Methanol weight (g)
(6). X6 = X_6 =X6= Oil weight (g)

The output variable was Y = Y = Y = biodiesel yield (wt.%). This dataset was originally developed from experimental runs performed under controlled laboratory conditions.

**2.2.2 Modeling**

The modeling was implemented in Python using the scikit-learn library. The libraries for implementing the models, including Decision Tree Regressor (DTR), Random Forest Regressor (RFR), Gradient Boosting Regressor (GBR), mean_squared_error, and r2_score were imported from the sklearn library. The dataset involving the biodiesel production process information was loaded using pandas and separated into features (X) and the target variable (y). This was followed by initializing the models with specified hyperparameters and training on the training data. The trained models (Equations 1 - 2) were then utilized in making predictions on the testing data, allowing for performance evaluation against actual biodiesel yield.

1. Random forest prediction (Y) = $\frac{1}{n}\sum_{i\text{-}1}^{n} Ti(x)$ (1)

2. Gradient Boosting prediction $F_m(x) = F_{m\text{-}1}(x)+y_m h_m(x)$ (2)

where: Y = predicted output (biodiesel yield); n = total number of trees in the forest; $Ti(x)$ = the prediction from the i-th tree for input x; $F_m(x)$ = model prediction after mmm-th iterations; $h_m(x)$ = current tree's prediction on the residuals ; $y_m$ = the learning rate, controlling how much each tree's prediction influences the overall model.

**2.2.3. Model Evaluation**

(1.) Coefficient of determination ($R^2$)

This was used to evaluate the performance of the models, including Decision Tree Regressor (DTR), Random Forest Regressor (RFR), and Gradient Boosting Regressor

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

(GBR). After making predictions on the X_test data, the $R^2$ score was computed using an inbuilt command in Jupyter Notebook. The following code snippet was used to calculate and print the $R^2$ value for the model:

```
from sklearn.metrics import r^2_score
r^2 = r^2_score(y_test, y_pred)
print(f'R^2: {r^2}')
```

The underpinning equation for computing the coefficient of determination ($R^2$) is presented in Equation (3).

$$R^2 = 1 - \frac{\sum_{I=1}^{N}(y_i - \bar{y}_i)^2}{\sum_{I=1}^{N}(y_i - \bar{y}_i)^2} \quad (3)$$

(2). Mean Squared Error (MSE)

After making predictions on the X_test data, the MSE can be computed using an inbuilt command in Jupyter Notebook. The following code was used to calculate and print the MSE for the model:

```
from sklearn.metrics import mean_squared_error
# Calculate Mean Squared Error
mse = mean_squared_error(y_test, y_pred)
print(f'MSE: {mse}')
```

The underpinning equation for determining the mean squared error (MSE) is presented in Equation 4.

$$\text{MSE} = \frac{\sum_{I=1}^{N}(y_i - \bar{y}_i)^2}{n} \quad (4)$$

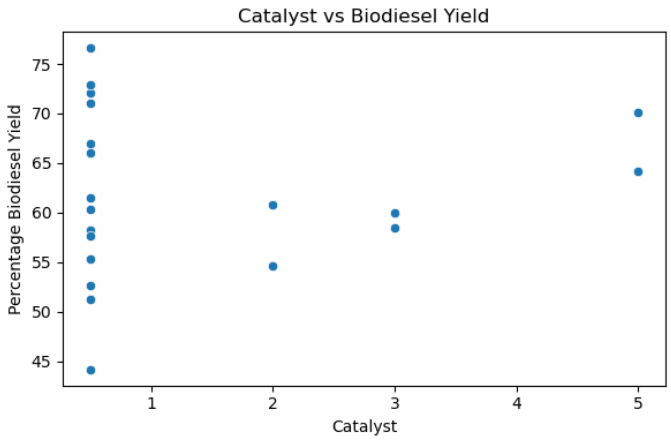
where n = number of runs; $\bar{y}_i$ = mean of the data set; $y_i$ = actual value of the biodiesel yield and $\bar{y}_i$ = predicated value of the biodiesel yield.

**2.2.4 Model validation**

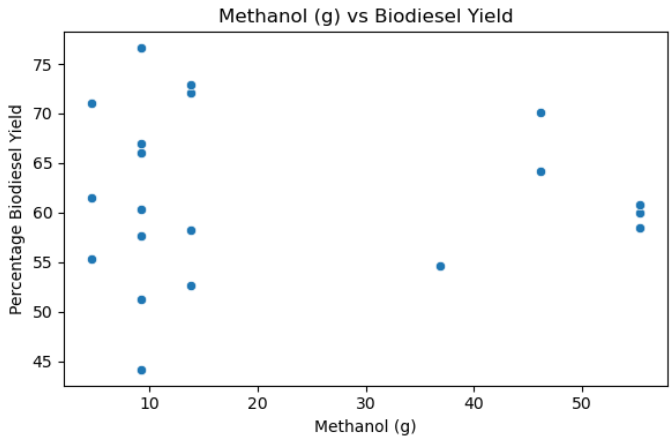
**3.0 Results and Discussion**

3.1 Effects of factors on yield from the production of biodiesel

A relationship was established between the various factors (input parameters) such as the weight of the catalyst, temperature, weight of the methanol, time, methanol/oil ratio, and weight of the oil, on the biodiesel yield. Figure 1 illustrates that the relationship is non-linear, a characteristic commonly observed in chemical processes, according to Ehinmowo et al. (2025).




(a)




(b)

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

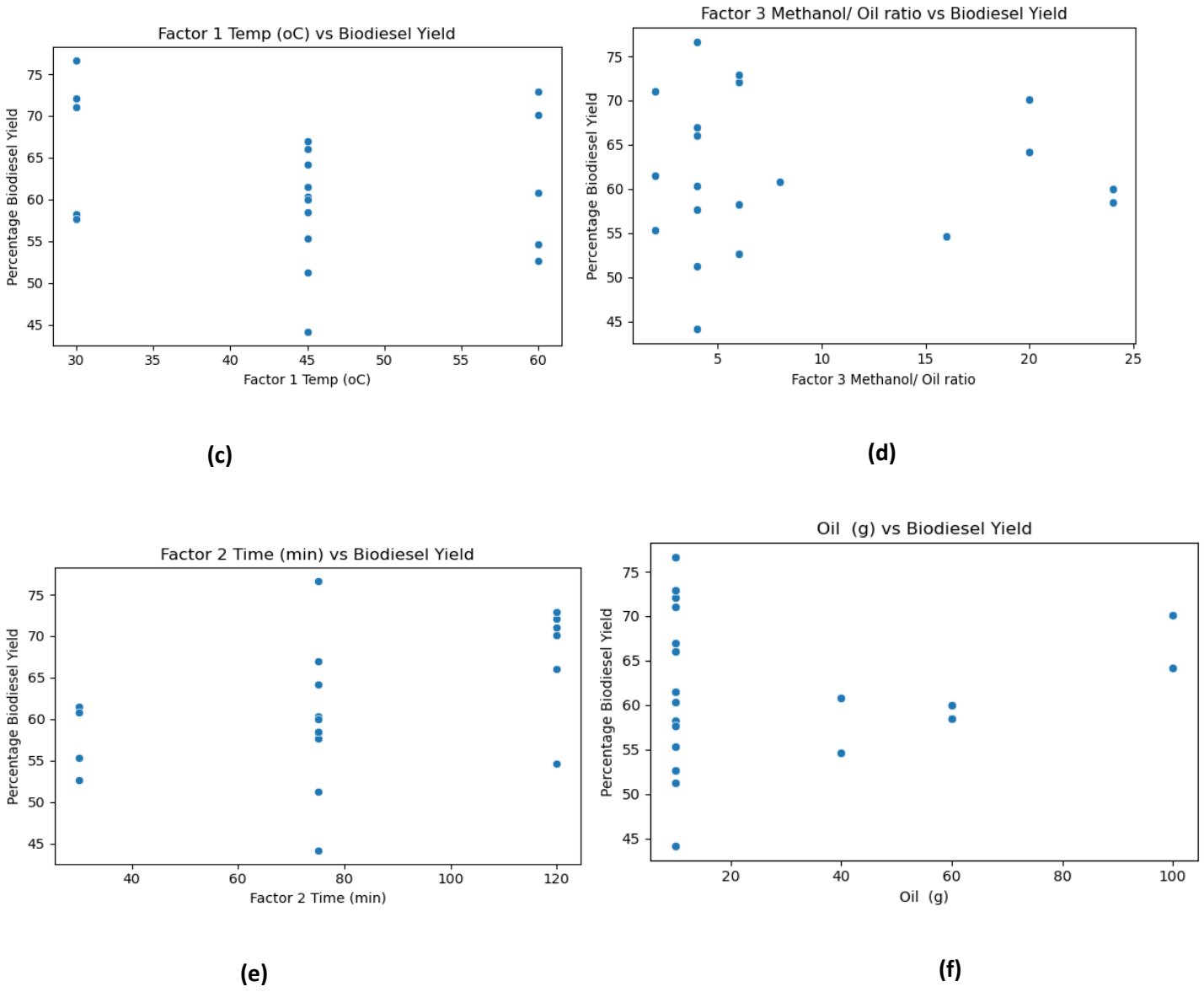



Figure 1: Effect of factors on the yield of biodiesel produced from the transesterification of rubber seed oil: (a). Effect of catalyst; (b). Effect of temperature; (c). Effect of the weight of methanol; (d). Effect of time; (e). Effect

Considering the correlation heatmap shown in Figure 2, the catalyst shows a strong correlation (1.0) to the weight of the oil.

Considering the correlation heatmap shown in Figure 2, the catalyst shows a strong correlation (1.0) to the weight of the oil.

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

of methanol/oil ratio and (f). Effect of the weight of the oil of biodiesel yield, respectively

### 3.2. Correlation Analysis

Considering the correlation heatmap shown in Figure 2, the weight of methanol shows a strong correlation (0.9) to the methanol/oil ratio, weight of the oil shows a strong correlation (0.88) to the methanol/oil ratio, while a positive strong correlation of 0.85 occurs between weight of the catalyst and methanol, and also between weight of the oil and methanol. The strong positive correlation is an indication that these variables strongly depend on each other resulting to a high yield of biodiesel. Explicitly, increase in one variable leads to a corresponding increase in the other variable and vice versa. Conversely, a strong negative correlation of -0.10 was recorded between temperature and time resulting to a low yield. This shows that the interaction of temperature and time had no appreciable effect on the yield of the biodiesel.

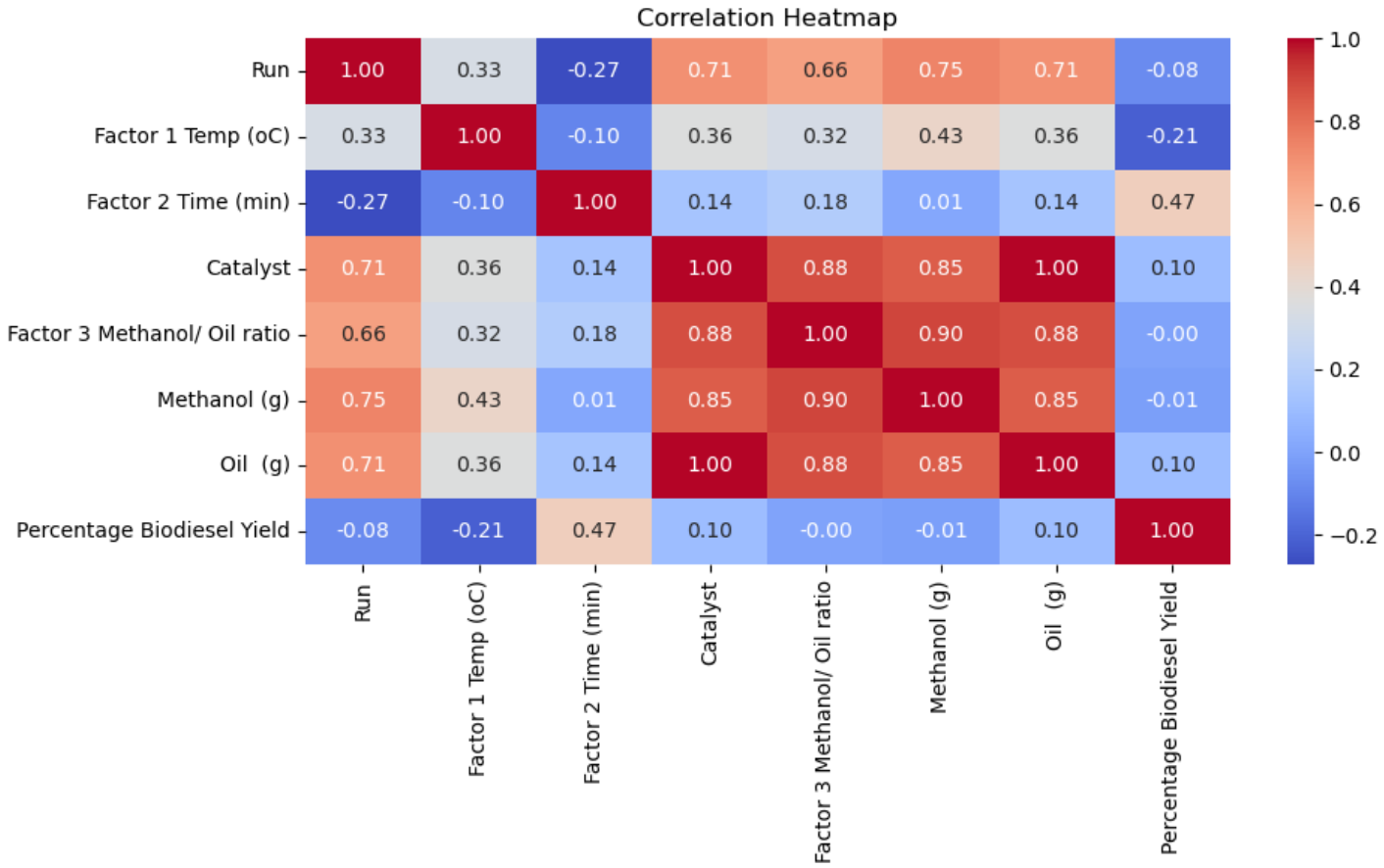



Figure 2. Correlation Heatmap of biodiesel yield

### 3.3 Features relationship

This was conducted to determine the effect of the factors such as temperature, contact time, catalyst loading, weight of methanol, weight of oil, and methanol to oil ratio on the yield of biodiesel. As shown in Figure 3, time was the most predominant factor resulting to high yield of biodiesel.

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

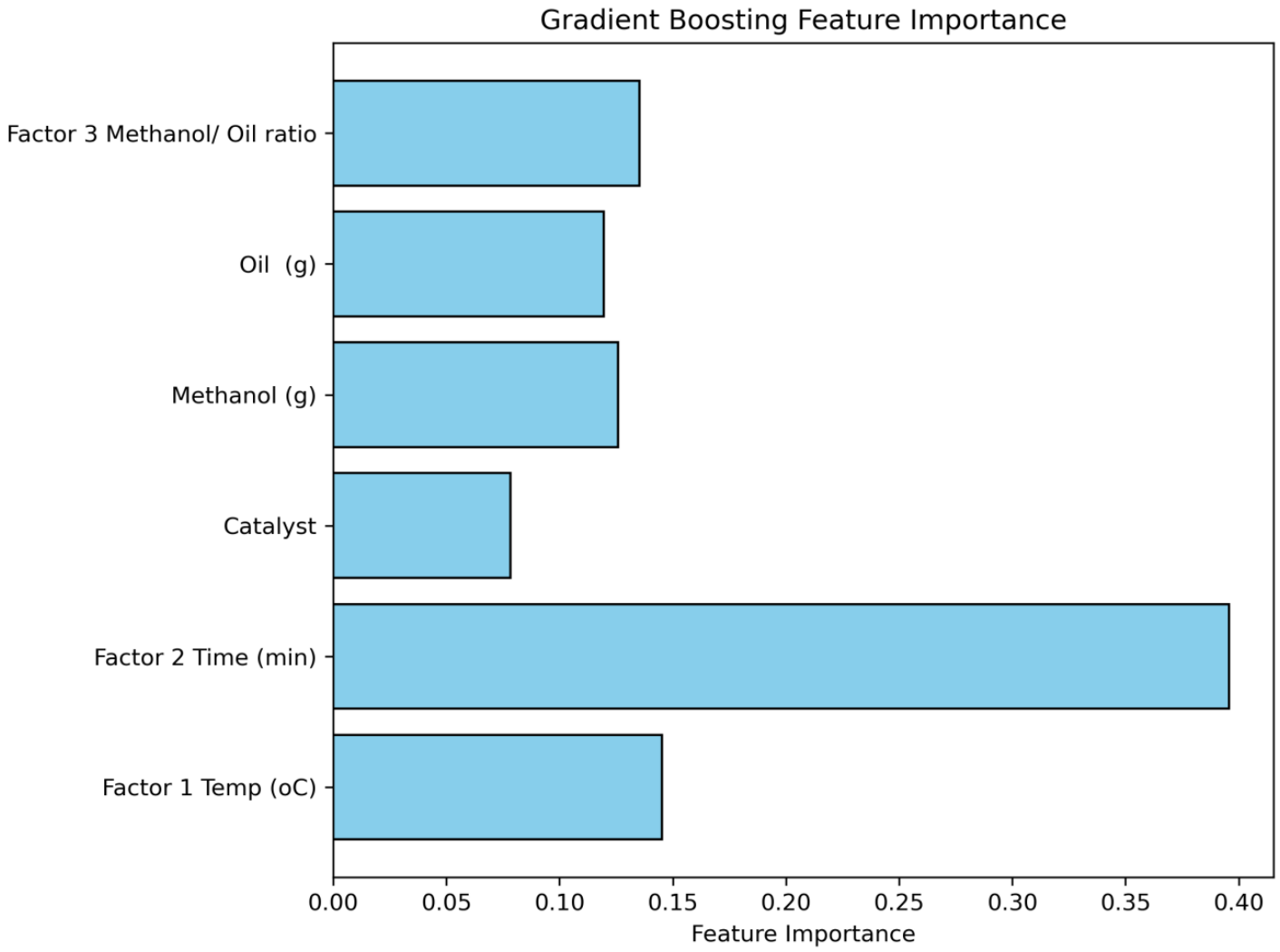
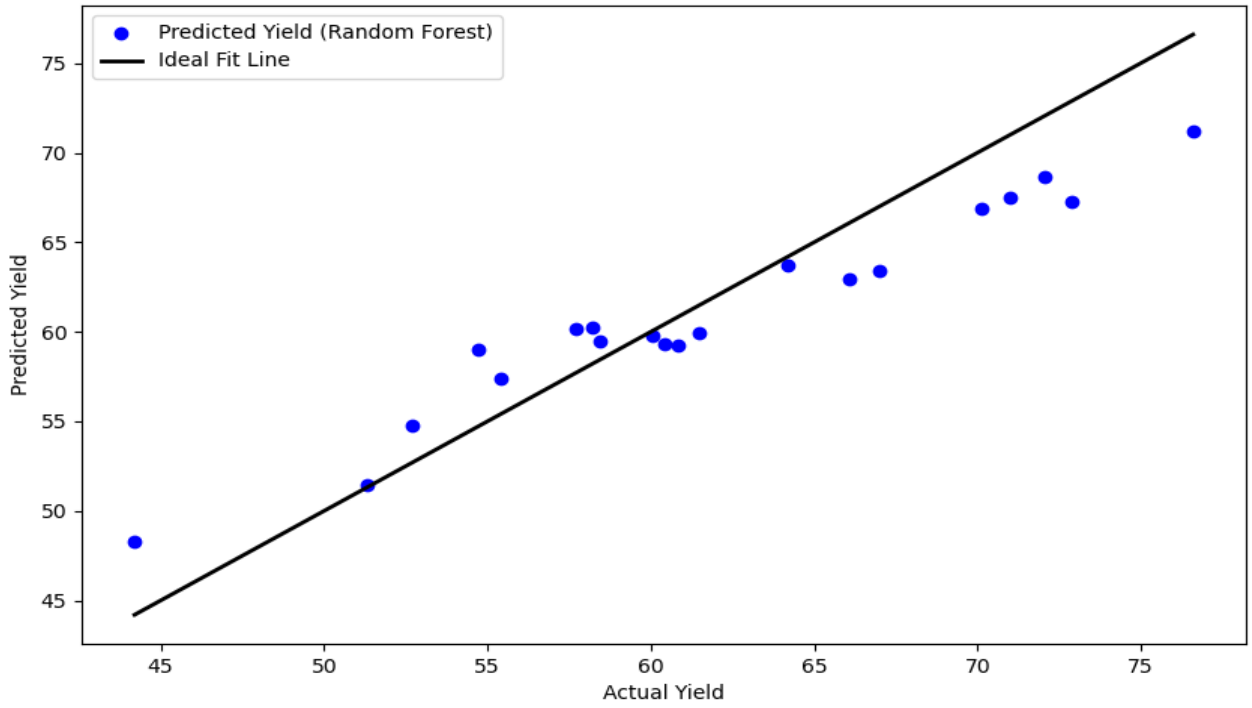


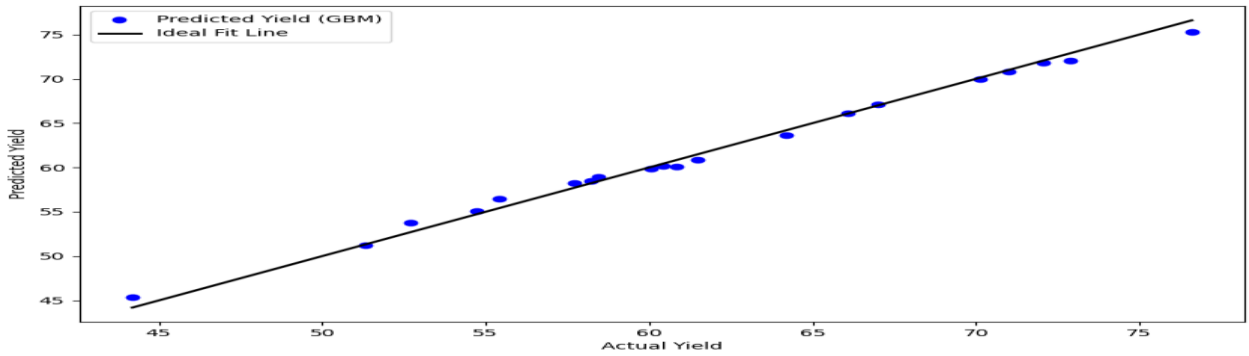

Figure 3. Features relationship analysis for biodiesel yield

**3.4 Model Development**

The models Random Forest Regressor, Gradient Boosting Regressors, and Decision Tree Regressor prediction of the biodiesel yield are presented in Figures 4 - 6. The Random Forest Regressor (Figure 2) shows a fairly good predictive accuracy compared to the Decision Tree Regressor as the actual and predicted biodiesel yield fail to align properly unlike the Gradient Boosting Regressor. The result shows that the Gradient boosting Regressor model performed better than the Random Forest Regressor and the Decision Tree Regressor suggesting high predictive accuracy and minimal error. The superior performance may be attributed to the boosting mechanism of GBR, in which weak learners are sequentially improved, reducing bias and variance simultaneously. The Figure 5 shows close clustering along the identity line, indicating a strong fit. The Decision Tree Regressor reveals broader dispersion of predicted values, indicating less reliability (Figure6).

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37



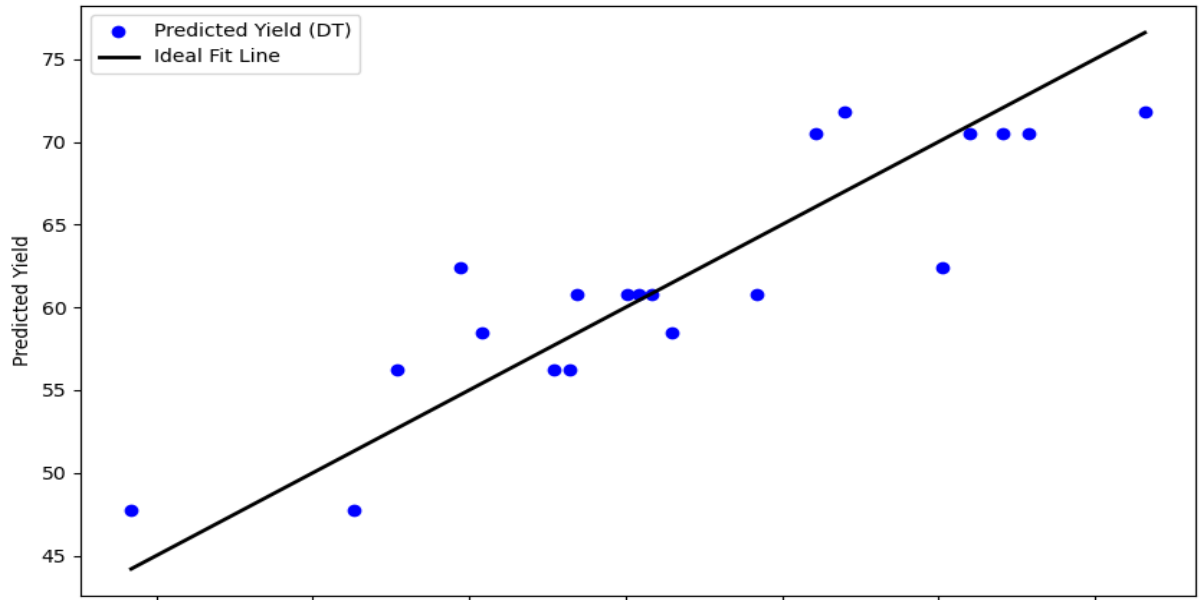



Figure 5. Actual and predicted percentage biodiesel yield using Gradient Boosting Regressor

Figure 6. Actual and predicted percentage biodiesel yield using Decision Tree Regressor

3.5 Model Evaluation Metrics

The models were evaluated using two key performance metrics, including the coefficient of determination ($R^2$) and mean squared error (MSE). These metrics were used to assess the performance of the gradient boost regressor (GBR), random forest regressor, and decision tree regressor. The result obtained is presented in Figure 7. The figure show that the lowest $R^2$ was 0.79 obtained from the decision tree regressor and the highest was 0.99 from gradient boost regressor. The lowest MSE was 0.4 obtained from gradient boost regressor and the highest was 13.64 from decision tree regressor. A low MSE means the predicted values were very close to the actual yield values, and the high $R^2$ confirms that the model explained over 99% of the variance in the dataset. The result suggests that gradient boost regressor predicted the biodiesel yield better than the random forest regressor and decision tree regressor, indicating that it is the most accurate and reliable model. This superior performance can be attributed to the boosting mechanism of GBR, where weak learners are sequentially improved, reducing bias and variance simultaneously.

**Academic Journal of Innovative Engineering and Technology**
Vol.06, No. 04; April-2026;
ISSN (2343–403X);
p –ISSN 3244–5621
Impact factor: 6.37

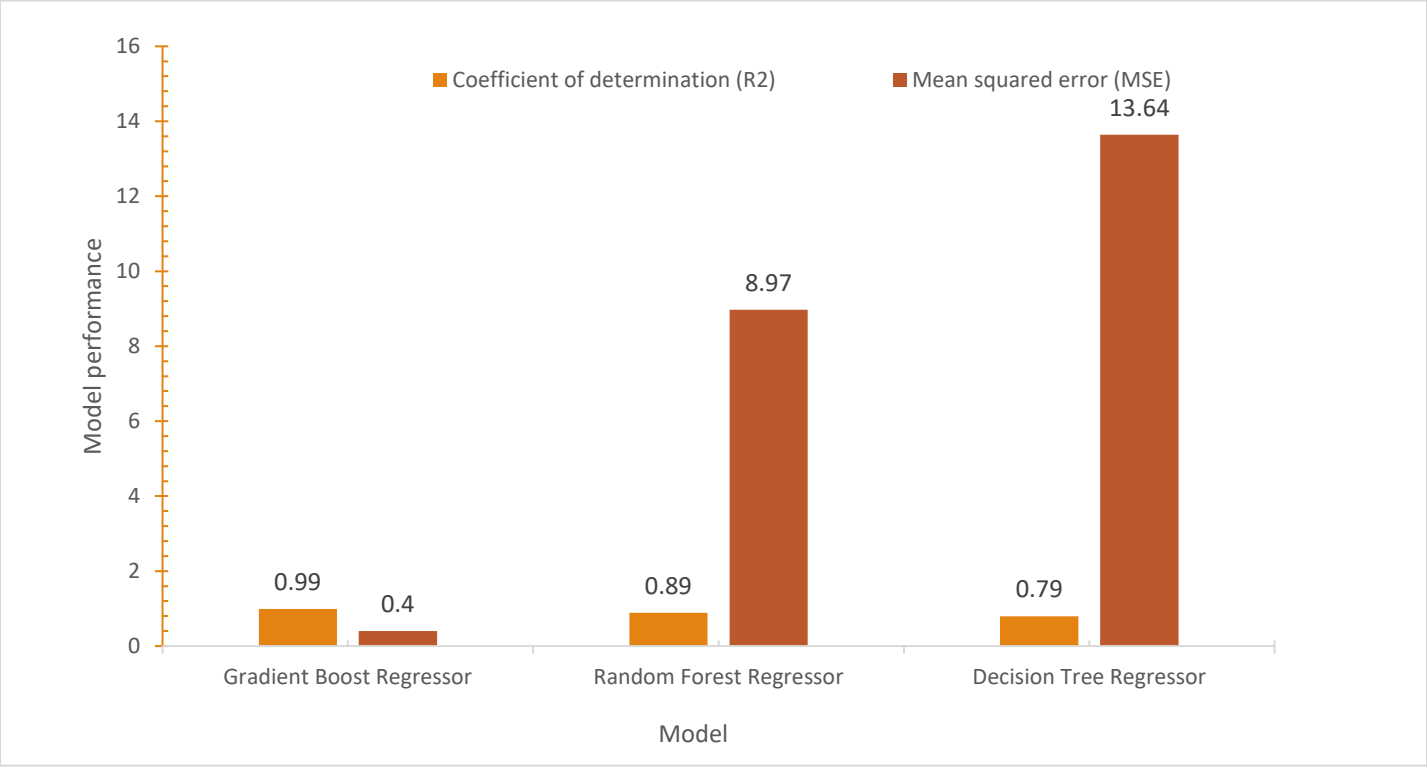



Figure 7. Model evaluation

Second to GBR in predicting the biodiesel yield was the random forest regressor (RFR) with an MSE of 8.9789 and $R^2$ of 0.8612. While this model also leveraged ensemble learning, it did so through averaging multiple decision trees, which helped to reduce overfitting compared to a single decision tree. However, it may have been less effective than GBR in predicting the biodiesel yield due to its inability to correct the errors of weak learners iteratively. The decision tree regressor (DTR) had the poorest performance, with the highest MSE of 13.6400 and the lowest $R^2$ of 0.7900. This indicates that the model struggled to generalize well to new data. DTRs tend to overfit training data, especially with small or noisy datasets, which likely contributed to its relatively poor predictive accuracy. The differences in performance can be attributed to the underlying learning mechanisms. The GBR, by sequentially focusing on the residuals of previous models, incrementally reduces prediction errors. In contrast, RFR averages predictions without accounting for sequential corrections, and DTR uses a single tree structure prone to overfitting. The findings of this study align with those of Rani et al. (2021), who reported that Gradient Boosting techniques outperformed both Random Forest and Decision Trees in predicting yield in a palm oil biodiesel production study. Similarly, Sharma and Kumar (2020) found that GBR achieved lower MSE values in modeling chemical process parameters, emphasizing its robustness for nonlinear relationships. In contrast, Mohammed et al. (2019) observed that Random Forest outperformed Gradient Boosting in a dataset with significant outliers. However, this discrepancy is likely due to the higher sensitivity of GBR to outliers, which was not a major issue in the rubber seed oil dataset used in this study.

Moreover, Ajiboye et al. (2022) in their work on biodiesel from jatropha seed oil found that Decision Tree Regressors showed erratic performance with large MSE values, which is consistent with this study's finding of DTR being the least reliable. The high performance of GBR in this study demonstrates the model's capacity to learn complex,

nonlinear interactions among transesterification process variables such as temperature, catalyst weight, reaction time, and methanol/oil ratio. This suggests that boosting-based ensemble methods are particularly well-suited for biodiesel yield prediction tasks and could be applied in similar contexts involving other feedstocks.

**4.0 Conclusion**

The study has demonstrated that accurate machine learning models can be developed for predicting biodiesel yield from rubber seed oil by acquiring a reliable dataset and applying Decision Tree Regressor, Random Forest Regressor, and Gradient Boosting Regressor. The models were trained using standardized data and evaluated through $R^2$ and MSE. Among the three, the Gradient Boosting Regressor demonstrated superior predictive performance with the highest $R^2$ (0.9937) and the lowest MSE (0.4058), indicating its effectiveness in modeling the nonlinear relationships inherent in biodiesel production processes. These findings validate the use of ensemble-based approaches like GBR for biodiesel yield prediction and show significant improvement over conventional regression methods reported in previous studies.

**Reference**

Abhishek, K., Agarwal, S., & Kumar, R. (2021). Application of SHAP for model interpretability in machine learning-based biodiesel yield prediction. *Energy Reports, 7*, 3502–3512.

Adewole, J. K., Ogunniyi, D. S., & Daramola, M. O. (2018). Rubber seed oil as a potential feedstock for biodiesel production: A review. *Renewable and Sustainable Energy Reviews, 92*, 817–828.

Aghbashlo, M., Tabatabaei, M., & Ghanavati, H. (2021). Machine learning applications in biofuel production: A comprehensive review. *Energy Reports, 7*, 2316–2340.

Atabani, A. E., Silitonga, A. S., Ong, H. C., Mahlia, T. M. I., Masjuki, H. H., Badruddin, I. A., & Fayaz, H. (2013). Non-edible vegetable oils: A critical evaluation of oil extraction, fatty acid compositions, biodiesel production, characteristics, engine performance and emissions production. *Renewable and Sustainable Energy Reviews, 18*, 211–245.

Atabani, A. E., Silitonga, A. S., Ong, H. C., Mahlia, T. M. I., Masjuki, H. H., Badruddin, I. A., & Fayaz, H. (2012). A comprehensive review on biodiesel as an alternative energy resource and its characteristics. *Renewable and Sustainable Energy Reviews, 16(4)*, 2070–2093.

Balat, M., & Balat, H. (2010). Progress in biodiesel processing. *Applied Energy, 87(6)*, 1815–1835.

Bali, V., & Singla, P. (2022). Comparative study of regression models for predicting biodiesel yield from jatropha oil. *Fuel, 314*, 122845.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Chauhan, R., Pal, A., & Saxena, R. (2020). Machine learning models for biodiesel production from Mahua oil. *Fuel, 280*, 118537.

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science, 7*, e623.

Demirbas, A. (2009). Progress and recent trends in biodiesel fuels. *Energy Conversion and Management, 50*(1), 14–34.

Ehinmowo, A. B., Nwaneri, B. I., Olaide, J. O. (2025). *Predictive modeling of hydrogen production and methane conversion from biomass-derived methane using machine learning and optimization techniques*. Next Energy 7, 1-17

Ejeh, B., & Suleiman, M. (2022). Techno-economic feasibility of biodiesel production from rubber seed oil in Nigeria. *Renewable Energy, 183*, 569–578.

Eze, S. O., Ozoani, H. A., & Oduola, M. K. (2021). Rubber seed oil as a sustainable biodiesel feedstock: A review of physicochemical properties and process optimization. *Journal of Renewable Energy and Environmental Sustainability, 8*, 1–10.

Farobie, O., Matsumura, Y., & Budiman, A. (2015). Application of response surface methodology for optimization of biodiesel production over biochar-supported CaO catalyst. *Bioresource Technology, 183*, 350–357.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232.

Issariyakul, T., & Dalai, A. K. (2014). Biodiesel from vegetable oils. *Renewable and Sustainable Energy Reviews, 29*, 678–695.

Jain, S., & Sharma, M. P. (2011). Prospects of biodiesel from Jatropha in India: A review. *Renewable and Sustainable Energy Reviews, 15(9)*, 4732–4741.

Jisieike, C. A., Chukwuma, O. C., & Uzochukwu, C. (2023). Application of ANFIS and Artificial Neural Network models for predicting biodiesel yield from non-edible oils. *Biofuels, Bioproducts and Biorefining, 17*(2), 255–267.

Kolakoti, N., Shah, N., & Rathi, R. (2020). Optimization of biodiesel production using hybrid genetic algorithm and response surface methodology. *Renewable Energy, 145*, 1900–1911.

Mekhilef, S., Siga, S., & Saidur, R. (2011). A review on palm oil biodiesel as a source of renewable fuel. *Renewable and Sustainable Energy Reviews, 15(4)*, 1937–1949.

Mohanty, S. K., Parthasarathy, R., & Kundu, K. (2019). Machine learning and regression analysis based hybrid modeling for biodiesel yield prediction. *Renewable Energy, 138*, 1064–1074.

Odetoye, T. E., Olatunji, O. M., & Afolabi, I. S. (2020). Comparative study of the biodiesel production from rubber seed oil and palm kernel oil. *Energy Reports, 6*, 1416–1424.

Oladosu, G. A., Salihu, A., & Abdulkadir, A. (2017). Cost analysis of biodiesel production from rubber seed oil. *International Journal of Energy Economics and Policy, 7*(1), 1–5.

Sahu, A. K., & Agarwal, A. K. (2020). A review on prediction of biodiesel properties using machine learning and data-driven approaches. *Renewable and Sustainable Energy Reviews, 128*, 109900.

Teo, S. H., Yusoff, M. H., & Yusup, S. (2022). Response surface methodology vs machine learning in process optimization: A case study on biodiesel synthesis. *Renewable and Sustainable Energy Reviews, 156*, 111994.

Wuest, T., Weimer, D., Irgens, C., & Thoben, K. D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research, 4*(1), 23–45.