RESEARCH ARTICLE

An assessment of optimizing biofuel yield percentage using K-fold integrated machine learning models for a sustainable future

Krishnamoorthy Ramalingam[1,2,*], Mohd. Zulkifly Abdullah[1*], Perumal Venkatesan Elumalai[3], Allam Sangeetha[4], Xu Yong[5*], Nasim Hasan[6*], Wei Shangzhi[7]

1 School of Mechanical Engineering, Engineering campus, Universiti Sains Malaysia, Nibong Tebal, Penang, Malaysia, **2** Department of Sustainable Energy Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Science, Chennai, India, **3** Department of Mechanical Engineering, Aditya University, Surampalem, India, **4** Department of Computer Science and Engineering, MLR Institute of Technology, Hyderabad, Telangana, India, **5** School of Artificial Intelligence and Smart Manufacturing, Hechi University, Yizhou, China, **6** Mettu University, Metu, Ethophia **7** Faculty of Education, Shinawatra University, Bang Toei, Thailand,

* nasimhasan78@gmail.com (NH); kskrishnamech@gmail.com (KR); mezul@usm.my (MZA); xuyong@hcnu.edu.cn (XY)

Abstract

Accelerating population and modernization has triggered a steady rise in energy demand and a significant rise in household waste, particularly municipal solid waste. In this context, waste-to-energy conversion has emerged as a sustainable solution. This study aims to maximize biofuel production yield using biomass-based banana peel catalyst waste by optimizing process parameters through machine learning models integrated with k-fold cross-validation. The models employed include Polynomial Regression (PR), Decision Tree (DT), Random Forest (RF), and Linear Regression (LR). The three key input variables including reaction temperature (RT), catalyst concentration (CC), and methanol-to-oil molar ratio (MOR) were used to train and test the models, with biodiesel yield as the measured output. Among the models, PR emerged as the best-performing one for predicting biofuel yield, demonstrated by its high R^2 value of 0.956 and low error metrics (RMSE = 1.54 MSE = 2.39 MAE = 1.43). The best model was determined through balancing bias and variance across k-fold validation iterations, where PR exhibited the highest average R^2 value of 0.868. Furthermore, the optimized process parameters predicted by PR for maximum biofuel yield were a RT of 59°C, CC of 2.96%, and a MOR of 9.21, resulting in a yield of 95.38%. These findings contribute to advancing large-scale machine learning-driven biofuel optimization, supporting industrial waste-to-energy applications, and fostering sustainable energy development.

OPEN ACCESS

Citation: Ramalingam K, Abdullah MZ, Elumalai PV, Sangeetha A, Yong X, Hasan N, et al. (2025) An assessment of optimizing biofuel yield percentage using K-fold integrated machine learning models for a sustainable future. PLoS One 20(8): e0328880. https://doi.org/10.1371/journal.pone.0328880

Editor: Bahram Hosseinzadeh Samani, Shahrekord University, IRAN, ISLAMIC REPUBLIC OF

Received: May 26, 2025

Accepted: July 8, 2025

Published: August 14, 2025

Copyright: © 2025 Ramalingam et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The datasets used and/or analysed during the current study are fully available within the manuscript.

Funding: This research was funded by the Guangxi Science and Technology Program

under Grant No: AA24010001 awarded to X.Y., This research was financially supported by First-class Discipline Construction Project of Hechi University, Guangxi Colleges awarded to X.Y., and Universities Key Laboratory of AI and Information Processing (Hechi University), Education Department of Guangxi Zhuang Autonomous Region awarded to X.Y.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: ML, Machine Learning; PR,Polynomial Regression; DT, Decision Tree; RF, Random Forest; LR, Linear Regression; NaOH, Sodium Hydroxide; KOH, Potassium Hydroxide; RMSE, Root Mean Square Error; MSE Mean square Error; MAE, Mean Absolute Error; SHAP, SHapley Additive exPlanations; R2, Coefficent of Determination; RT, Reaction Temperature; CC, Catalyst concentration; MOR, Methanol to oil molar ratio.

Introduction

Waste-to-energy conversion has emerged as a widely recognized strategy for promoting sustainability to addressing the dual challenges of biomass disposal and fossil fuel dependence. Surging global population growth combined with modernization has intensified global energy consumption, causing several countries to rely extensively on fossil fuels for transportation and electricity generation. However, fossil fuels are non-renewable, hydrocarbon-based energy sources that pose environmental and health risks while becoming progressively costlier due to depletion [1–3]. In response, researchers worldwide are exploring alternative energy solutions, including electric vehicles (EVs), hydrogen-based fuels, and biofuels derived from agricultural and biomass sources [2,4].

Among these renewable energy options, biofuels have emerged as a promising alternative, offering scalable solutions for sustainable energy production while addressing the limitations of EVs and hydrogen fuels. EV adoption is hindered by inadequate charging infrastructure, battery reliability concerns, and e-waste management challenges. Meanwhile, hydrogen-based fuels, particularly blue hydrogen, present environmental risks and storage complications, reducing their practicality in widespread energy applications. Given these constraints, biomass- and agriculture-based biofuels have gained prominence due to their efficient conversion processes and compatibility with existing fuel systems [5–8].

A critical factor in biofuel production is the selection of suitable feedstocks, which must meet criteria such as availability, cost-effectiveness, renewability, and non-edibility [9,10]. Seed-based feedstocks are commonly chosen for their high biodiesel yield; however, their elevated viscosity and fatty acid content pose challenges for combustion efficiency. To address these limitations, fuel upgrading processes have been developed, with transesterification standing out as one of the most effective techniques [11,12].

Transesterification has demonstrated superior thermal properties compared to other fuel upgrading techniques such as water emulsion, blending, and pyrolysis. However, biofuel yield is highly dependent on key reaction parameters namely RT, CC, MOR, and reaction duration, which directly affect conversion efficiency [13–15]. Researchers have investigated diverse catalysts utilized to facilitate transesterification, including homogeneous, heterogeneous, enzymatic, and supercritical alcohol-based catalysts. Among them, alkaline catalysts such as NaOH and KOH are commonly used for their high reactivity and fast conversion rates, particularly with low free fatty acid feedstocks. Acid catalysts perform better for feedstocks with high FFA content despite their slower reaction kinetics, while heterogeneous catalysts offer sustainability benefits through reusability and improved separation processes. Supercritical transesterification provides an alternative method, achieving high biofuel yields without catalyst contamination [16–18].

Many researchers have favored heterogeneous catalysts in biodiesel synthesis for their reusability, sustainability, and ease of separation from the final product. Unlike homogeneous catalysts, which require additional purification steps, heterogeneous

catalysts enable multiple reaction cycles, reducing waste and improving process efficiency [19,20]. Among the widely studied heterogeneous catalysts, metal oxides such as CaO, MgO, ZnO, TiO_2 and etc have demonstrated high catalytic activity and stability. Table 1 provides an extensive survey on the application of heterogeneous catalysts in biodiesel synthesis, highlighting their effectiveness across different feedstocks and optimization techniques. Such catalysts persist as key enablers in the evolution of biofuel research, contributing to sustainable energy advancements.

Traditionally, process optimization in biodiesel production has relied on design of experiments (DoE) techniques such as Taguchi, ANOVA, and Response Surface Methodology (RSM) to identify ideal parameter combinations. Although these statistical methods provide valuable insights, they often struggle to capture nonlinear interactions among multiple variables, resulting in prediction errors ranging from 4–5%. To enhance predictive accuracy and optimize process parameters more effectively, machine learning (ML) techniques have recently gained prominence [30–32].

Machine learning models including ANN, PR, SVR, DT, RF, and KNN offer refined approaches for biodiesel yield prediction, outperforming conventional statistical methods in terms of reliability and accuracy [33,34]. To further improve prediction efficiency, K-fold cross-validation is integrated into ML models, ensuring robust evaluation and generalization. By segmenting the dataset into distinct subsets in a systematic manner and iteratively training the model, K-fold validation reduces bias and mitigates overfitting, making the selected model more reliable for process optimization [35,36].

Based on the literature survey, minimal research has discovered the application of machine learning models in predicting biodiesel yield using previously published datasets, particularly in the context of biomass-based heterogeneous catalysts. Therefore, this study employs four machine learning models Polynomial Regression (PR), Linear Regression (LR), Decision Tree (DT), and Random Forest (RF) to analyze and optimize process parameters for biodiesel production. These models are integrated with K-fold cross-validation to improve prediction accuracy and enhance model reliability. Additionally, SHAP factor analysis and heatmap visualization are conducted to identify the most influential parameters affecting biodiesel yield, providing deeper insights into process optimization based on existing experimental data.

Table 1. Summary of Heterogeneous Catalysts Used in Biodiesel Production Along with Influential Process Parameter.

S.No	Feedstock	Heterogeneous Catalyst	Optimum Catalyst Concentration (wt%)	Optimum Reaction Temperature (°C)	Optimum Methanol-to-Oil Molar Ratio	Optimum Reaction Duration (min)	Optimum Biodiesel Yield (%)	Most Influential Parameter
1	Waste Cooking Oil [21]	CaO	5	65	12:1	120	96.47	Catalyst Concentration
2	Jatropha Oil [22]	MgO	7.5	60	10:1	50	92.3	Reaction Temperature
3	Soybean Oil [23]	ZnO based composite catalyst	4	65	9:1	15	97%	Methanol-to-Oil Ratio
4	Palm Oil [24]	Calcareous egg shell based catalyst	7	150	12:1	180	94.5	Catalyst Concentration
5	Microalgae Oil [25]	Fe3o4@SiO_2	5	60	13:1	55	97%	Methanol-to-Oil Ratio
6	Sunflower Oil [26]	K_2CO_3	4	120	15:1	50	89	Reaction Temperature
7	Castor Oil [27]	Zn/Mg	3	55	9:1	40	87	Catalyst Concentration
8	Cottonseed Oil [28]	KoH	6	55	6:1	60	96	Methanol-to-Oil Ratio
9	Waste chicken fat [29]	eggshell	2	65	10:1	55	90.2	Catalyst Concentration

https://doi.org/10.1371/journal.pone.0328880.t001

Material and methodology

Catalyst preparation

Biomass-based banana peel-derived catalyst was prepared using agro-waste collected from local fruit vendors, restaurants, and vegetable markets. Initially, the gathered banana peels underwent thorough washing using distilled water to ensure the removal of surface contaminants and residual dirt. Any residual fruit material was manually separated and discarded. The cleaned peels were then sundried to reduce surface moisture and cut into smaller pieces for uniform drying. Subsequently, the chopped peels were oven subjected to drying at 60°C for two hours to remove remaining moisture content. The oven-dried biomass was subsequently calcined at 900°C to induce complete decomposition of organic matter and conversion into a carbon-rich calcinated catalyst. After calcination, the solid product was finely milled and preserved in a sealed container to prevent moisture absorption.

Procedure for biodiesel production

The transesterification of the composite oil blend comprising 20% neem oil and 80% used cooking oil was performed using a heterogeneous base catalyst synthesized from calcinated banana peels, as outlined in Section 3.1. The biodiesel production process began with the pre-treatment to eliminate impurities and reduce moisture content. The composite oil was initially filtered and preheated to facilitate the removal of water, which can otherwise impede catalyst performance and encourage soap formation. Due to the elevated FFA content in the composite oil, an acid-catalyzed esterification step was carried out using sulfuric acid to lower the FFA concentration to acceptable levels. In this process, in this process, 20 vol% methanol and 1 wt% H_2SO_4 were added to the pretreated oil and stirred at 70°C under controlled conditions to convert FFAs into methyl esters while minimizing the risk of saponification. Upon completion of the reaction, the mixture was neutralized and Rinsed thoroughly using distilled water to remove any remaining acid and byproducts, preparing the composite oil for the subsequent transesterification step.

The transesterification was conducted in a three-necked round-bottom flask fitted with a mechanical stirrer and digital thermometer to maintain thermal stability. Transesterification experiments were carried out across a range of parameters: methanol-to-oil molar ratio (6–12:1), catalyst loading (1–3 wt% of oil), and reaction temperature (55–65°C), with maintained a fixed reaction duration of 60 minutes with continuous stirring at 600 rpm. At the beginning of each run, the required amount of catalyst was dispersed in methanol and stirred for 10 minutes to ensure uniform mixing before being introduced into the reactor containing the pre-heated composite oil. The reaction mixture was kept under constant agitation throughout to promote effective mass transfer and maintain homogeneity. Upon completion, the mixture was poured into a separating funnel and allowed to stand for 12 hours to enable phase separation under gravity. Two distinct layers formed: an upper biodiesel-rich layer and a lower glycerol-rich layer. The biodiesel phase was carefully decanted, washed repeatedly using heated distilled water to ensure complete removal of catalyst residues, soap, and methanol, and subsequently dried in a hot air oven at 105°C for two hours to ensure complete moisture removal. The end product was securely stored in closed containers for further characterization and analysis [37]. A schematic depiction of the biodiesel production steps is provided in Fig 1, highlighting both the esterification and transesterification stages.

Background of ML regression model

In machine learning applications for biodiesel production, researchers frequently utilize regression models, classifiers, and ANN to refine experimental parameters to achieve optimal outcomes. For this study, regression models were selected based on pragmatic factors such as data availability, computational efficiency, and their appropriateness for the intended research objectives. These models encompass a diverse range of algorithms designed to accommodate specific data structures and predictive requirements, ensuring robust and interpretable results [38,39].

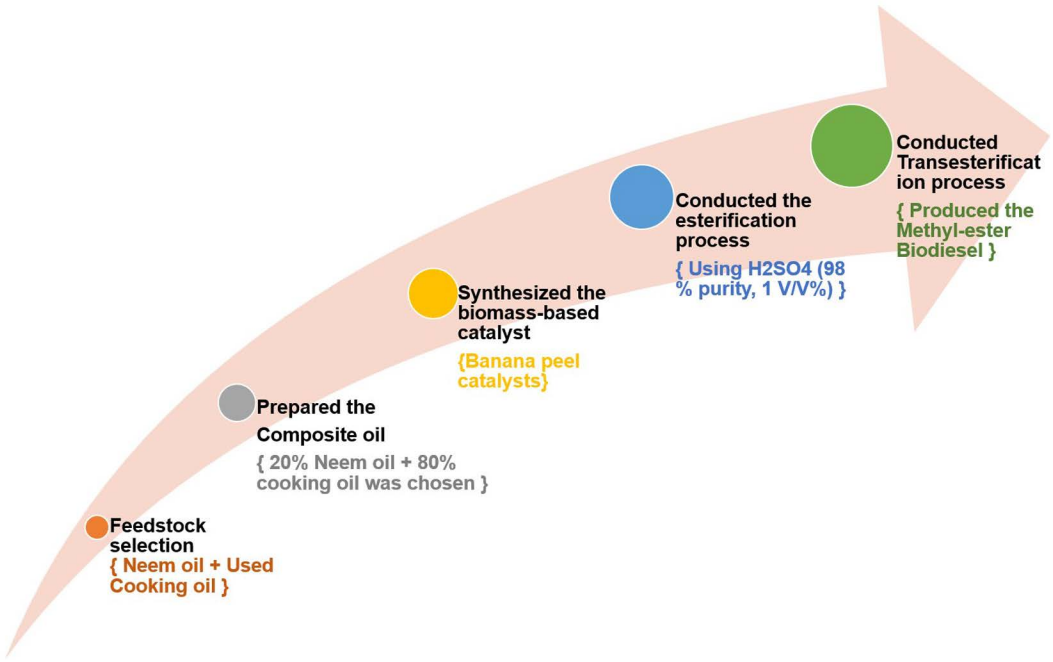


Fig 1. The schematic representation of steps involved in biodiesel production.

https://doi.org/10.1371/journal.pone.0328880.g001

To effectively predict and validate experimental results in biodiesel production, four regression algorithms were employed, each offering distinct advantages in handling diverse data patterns. Linear Regression (LR) is well-suited for capturing straightforward relationships between variables, making it ideal for modeling linear trends. In contrast, Polynomial Regression (PR) accommodates more complex, non-linear interactions, providing greater flexibility in identifying intricate patterns within the dataset. Decision Trees (DT) offer interpretable models for complex datasets, while Random Forests (RF) enhance prediction accuracy and mitigate overfitting by aggregating multiple decision trees [40,41]. The full machine learning workflow is presented in Fig 2.

Hyperparameters significantly influence how machine learning models learn and generalize to unseen data. To ensure learning accuracy, we have selected suitable hyper parameters of the models through a trial-and-error approach the selected hyperparameters were provided in Table 2. The intercept and normalization are key parameters in LR and degree of the polynomial is the key parameter in PR. Additionally, min_samples_leaf in DTs and n_estimators in RF are vital parameters to improve model stability and balance the underfitting and overfitting effectively. These adjustments ensure that each model learns meaningfully from the dataset while mitigating performance biases.

Data collection

The L27 orthogonal array design was employed as the Design of Experiments (DOE) methodology to systematically generate experimental data, as illustrated in Table 3. The dataset used in this study originates from a previously published work [37]. The dataset consists of 27 entries, partitioned into 80% for training and 20% for testing to support robust predictive modeling. The selected input parameters include reaction temperature (RT), catalyst concentration (CC), and methanol-to-oil molar ratio (MOR), while the output parameter being assessed is biodiesel yield.

The dataset spans a structured range, with RT varying between 55°C and 65°C, CC from 1 wt% to 3 wt%, and MOR extending from 6 to 12. These key variables are visually presented using contour plots in Figs 3a and 3b, facilitating a

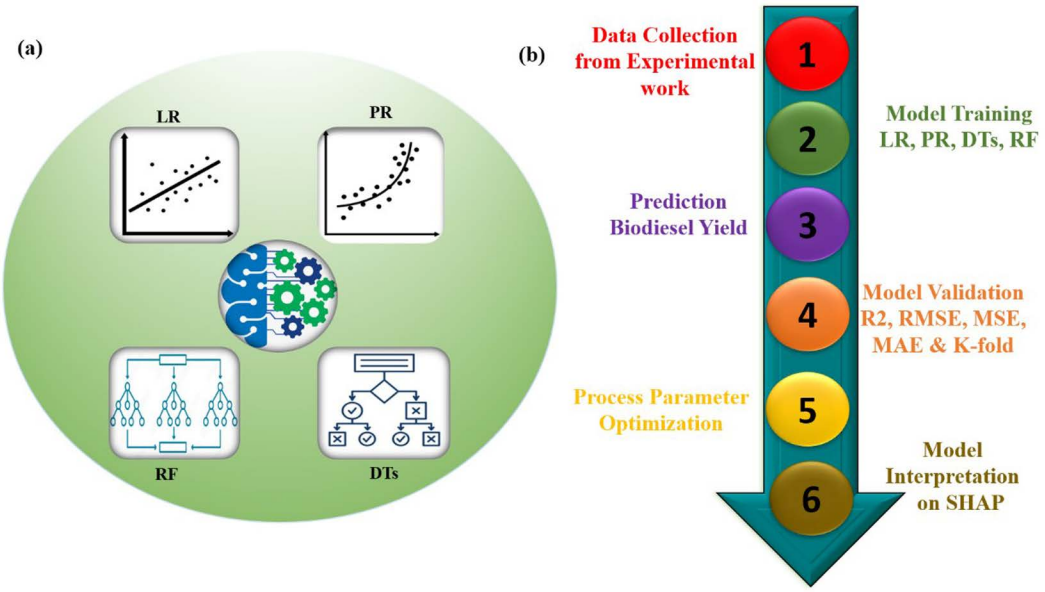


Fig 2. Overview of ML work flow a) ML models b) steps involve in work flow.

https://doi.org/10.1371/journal.pone.0328880.g002

Table 2. Hyper Parameters for all four ML learning models.

Sl. No	ML Models	Hyper Parameters
1	LR	Linear regression with intercept enabled and normalization turned off.
2	PR	PR of degree 2 without bias term.
3	DTs	No limit on tree depth, splitting requires 2 samples, each leaf must have 1 sample, and no feature limit.
4	RF	Random Forest with 10 trees, unlimited depth, minimum 2 samples to split, leaf size 1, feature selection is automatic, with bootstrapping and fixed seed of 42.

https://doi.org/10.1371/journal.pone.0328880.t002

comprehensive understanding of data distribution. Furthermore, Fig 3c provides a statistical summary, offering insights into overall trends and significant characteristics within the dataset. For ML prediction, Python Google Colab (Packaged as version 1.0.0 and compliant with Apache License, v2.0 terms) was used to build and train the model, utilizing the collected data from previous work along with Scikit-learn library.

K-fold cross validation

Traditional regression models typically utilize an 80:20 holdout validation, where the dataset was split into training (80%) and testing (20%) subsets to assess model performance. Performance is assessed through R^2 and error metrics, with higher R^2 and lower error indicating better predictive accuracy. However, this single-split approach can introduce biases due to uneven data partitioning.

To enhance reliability, Five k-fold cross-validation was employed, partitioning the dataset into five equally sized subsets. The model undergoes five iterative cycles, systematically using each fold as a test set while training on the remaining folds. This ensures comprehensive data utilization and mitigates overfitting risks. Averaging the five R^2 values obtained across iterations provides a more robust and unbiased model evaluation, particularly beneficial for small datasets [42,43].

Table 3. Experimental Design Matrix Based on L27 Orthogonal Array for Process Parameter Optimization.

Sl. No	Reaction Temperature (RT)	Methanol to oil molar ratio (MOR)	Catalyst concentration (CC)	Experimental Yield
1	55	9	1	60.14
2	55	12	1	78.33
3	60	9	1	76.33
4	60	12	1	74.34
5	60	6	1	72.38
6	65	9	1	73.42
7	65	12	1	72.41
8	65	6	1	69.33
9	55	9	2	91.54
10	55	12	2	91.61
11	55	6	2	89.14
12	60	9	2	91.52
13	60	12	2	91.24
14	60	6	2	88.44
15	65	9	2	87.65
16	65	12	2	87.45
17	65	6	2	85.31
18	55	9	3	94.25
19	55	12	3	94.57
20	55	6	3	94.33
21	60	9	3	94.18
22	60	12	3	94.35
23	60	6	3	94.51
24	65	9	3	94.14
25	65	12	3	94.21
26	65	6	3	94.35
27	55	9	1	60.14

https://doi.org/10.1371/journal.pone.0328880.t003

The methodology of five k-fold cross-validation is schematically represented in Fig 4, demonstrating its role in improving process parameter predictions for optimizing biodiesel yield.

Results and interpretations

ML Learning model prediction on biodiesel yield

All four selected machine learning algorithms were utilized for forecasting biodiesel yield based on investigational data, with results illustrated in Fig 5. The comparison between predicted yield and experimental yield was visualized using fit lines, test points, and train points, enabling an assessment of each model's accuracy in replicating real-world trends. From this evaluation, Polynomial Regression (PR) and Random Forest (RF) exhibited the highest similarity to experimental outputs, demonstrating their effectiveness in capturing complex interactions within the dataset. Further comparison revealed that PR and RF exhibited strong alignment with the fit line and minimal deviation from test and train points, reinforcing their reliability in accurately predicting biodiesel yield across varying process parameters. This robust predictive capability highlights their potential for optimizing biodiesel production efficiency.

The selected machine learning models were further validated using critical evaluation indicators, including the R^2 and error values such as RMSE, MSE, and MAE. In general, an elevated R^2 combined with low error values signifies superior

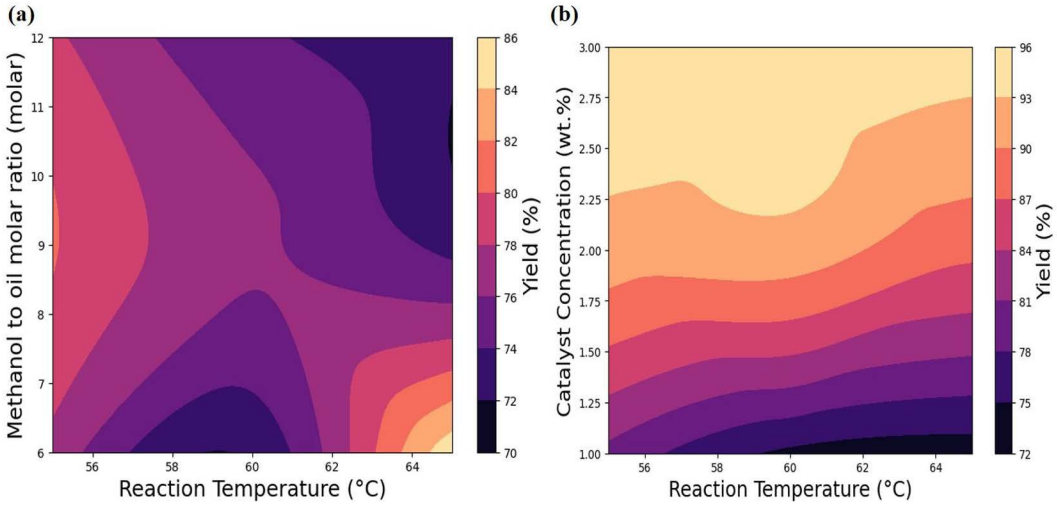


Fig 3. Visualization of Experimental data collection a) Reaction Temperature, Methanol to molar ratio verus biodiesel yield b) Reaction Temperature, Catalyst concentration versus biodiesel yield.

https://doi.org/10.1371/journal.pone.0328880.g003

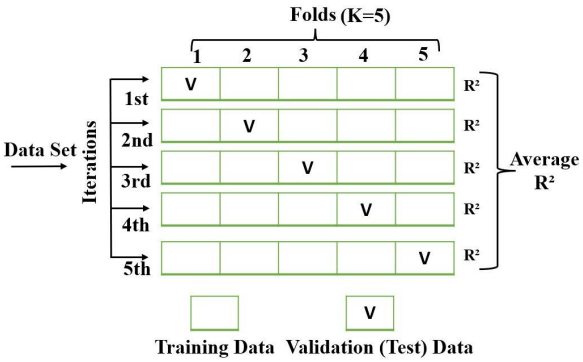


Fig 4. Basic concept of five k-fold cross-validation.

https://doi.org/10.1371/journal.pone.0328880.g004

predictive accuracy [44]. Thus, models demonstrating strong correlation between predicted and experimental outcomes were considered the most reliable for biodiesel yield prediction.

Fig 6 provides a detailed visual representation of these evaluation metrics across all models. The assessment confirmed that PR and RF unveiled the highest reliability, with R^2 values of 0.956 for PR and 0.911 for RF, indicating strong alignment with actual experimental results. Furthermore, the error values for PR and RF models were RMSE = 1.54, MSE = 2.39, MAE = 1.43 for PR, and RMSE = 2.21, MSE = 4.89, MAE = 1.92 for RF. These results highlight the robustness of PR and RF, reinforcing their effectiveness in accurately predicting biodiesel yield across various process parameters. Additionally, the reliability valuation emphasizes the significance of model selection and performance evaluation in optimizing biodiesel production processes.

Report on K-fold validation

K-fold cross-validation is a key technique for evaluating machine learning model performance via assessment of bias and variance. The interplay between bias and variance is graphically represented in Fig 7, classifies models into four distinct

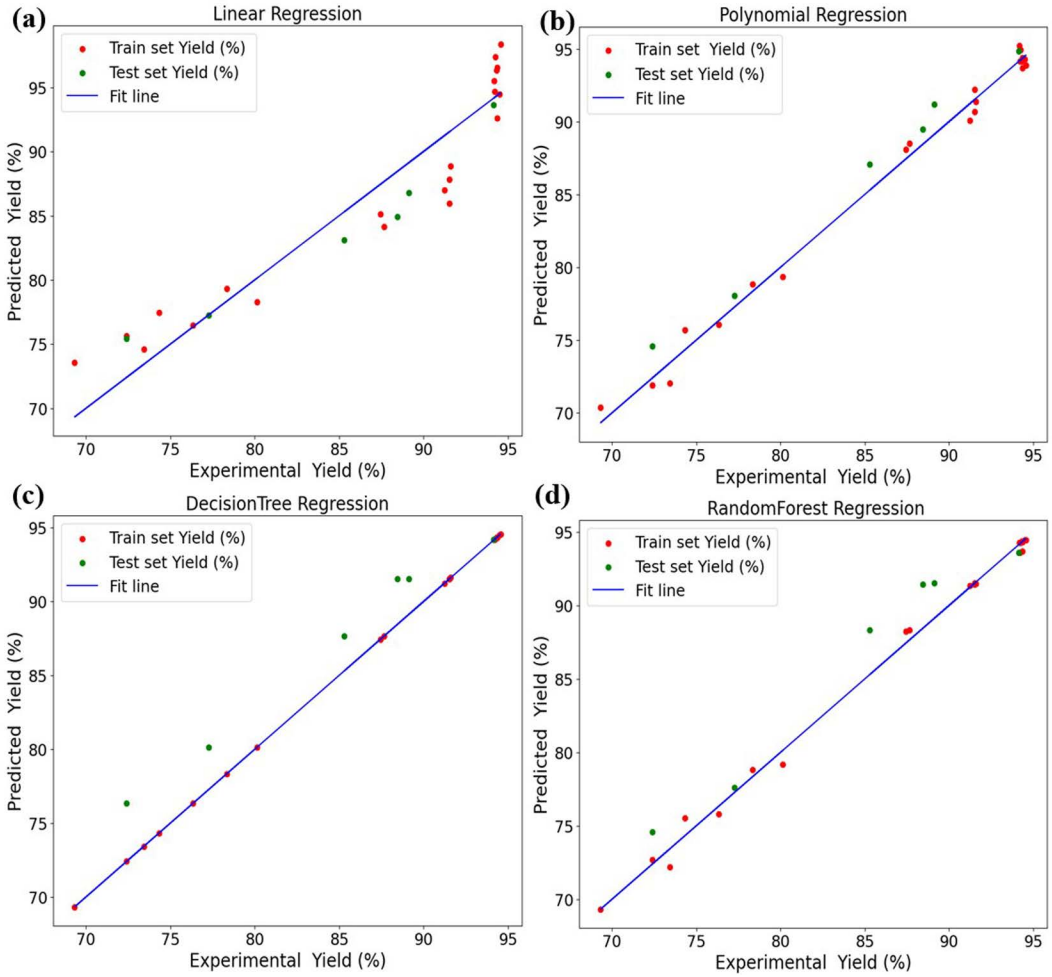


Fig 5. Comparison of experimental output verus prediction output by all four learning models a) LR, b) PR, c) DTs, d) RF.

https://doi.org/10.1371/journal.pone.0328880.g005

categories: optimal models, overly simple models, overly complex models, and unstable models. These distinctions are based on R^2 values, which serve as essential indicators of prediction accuracy. An overly simple model exhibits high bias but low variance, leading to generalization errors due to insufficient learning. Conversely, a high-bias, high-variance model struggles to identify patterns, making predictions unreliable. A low-bias, high-variance model captures trends but is highly sensitive to noise, resulting in inconsistent outputs. In contrast, an optimal model achieves low bias and low variance, ensuring stable and accurate predictions across diverse datasets. By systematically evaluating bias and variance, k-fold cross-validation enhances model reliability, minimizes errors, and optimizes parameter selection for better generalization in machine learning applications [45,46].

From the five-fold cross-validation, a total of 20 R^2 values were obtained, with each learning algorithm contributing five high R^2 values. These values were averaged to determine the mean R^2 score for each model, as tabulated in Table 4. The tabulated results specify that the PR model attained the highest average R^2 value of 0.868, surpassing all other models. Similarly, the RF model recorded the second-highest average R^2 value of 0.836, demonstrating strong predictive reliability.

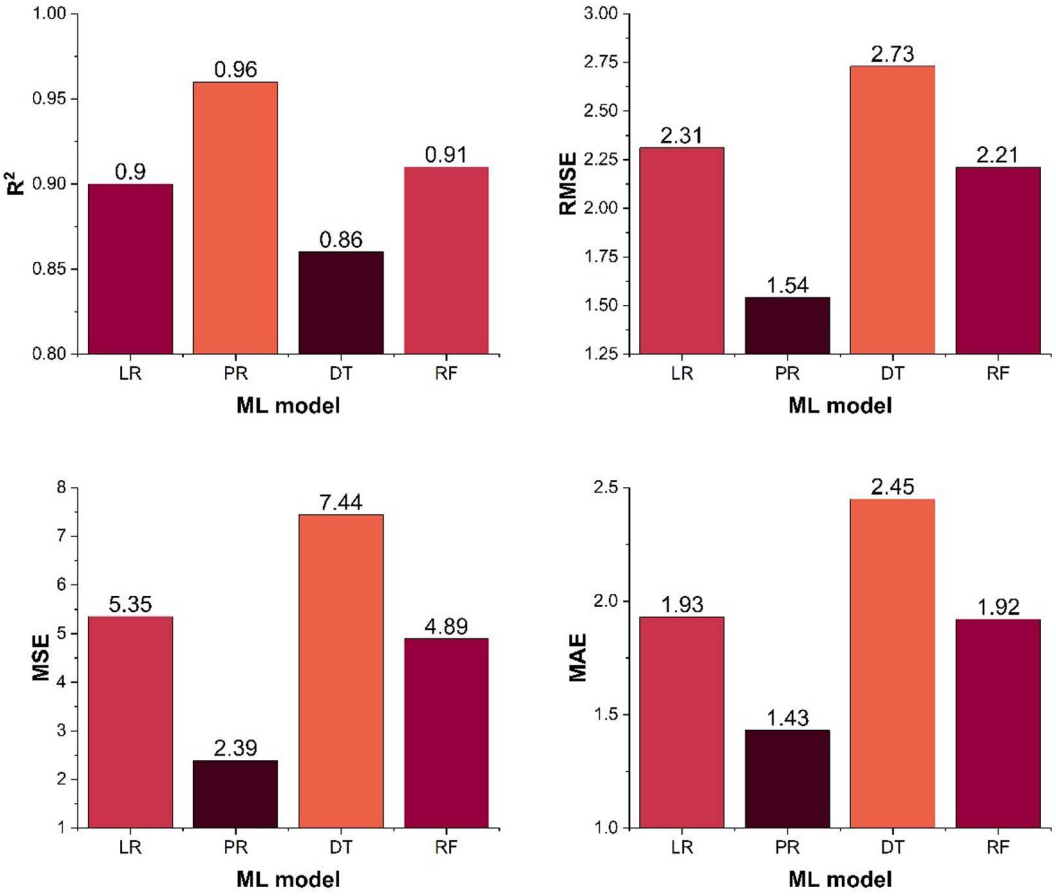


Fig 6. Validation metric for all four learning models a) R2, b) RMSE, c) MSE, d) MAE.

https://doi.org/10.1371/journal.pone.0328880.g006

Both PR and RF fall under the category of low bias and low variance, meaning they consistently yield predictions closer to the ideal value of 1 across all iterations. Their stable performance underscores their suitability for accurately modeling biodiesel yield prediction, reinforcing their efficacy in identifying nonlinear trends and interactions.

ML learning model interpretation

For model interpretation, SHAP values analysis, Feature importance, Partial dependence and Pearson correlation heat-map were employed to evaluate the best-performing learning algorithm, as determined via five k-fold cross-validation. The Polynomial Regression (PR) model demonstrated the highest prediction accuracy, closely aligning with experimental results. Consequently, the PR model underwent SHAP values analysis providing deeper insights into the influence of input variables on biodiesel yield predictions. The Feature importance and Partial dependence analysis provide quantities significant influence and trend of input variables. The Heatmap analysis used to provide liner corelation of input and out-put variables [47,48]. The Fig (8a) presents the relationship between SHAP values and feature inputs for the Polynomial Regression (PR) model, offering an in-depth interpretation of how individual parameters influence predictions.

The SHAP visualization follows a structured format, where wider sections highlight features that exert significant impact on the output response. In a SHAP summary plot, each dot denotes a prediction with its color, red for high, blue for low, and purple for medium values. The spread of dots for a given feature reflects its impact on predictions, with wider

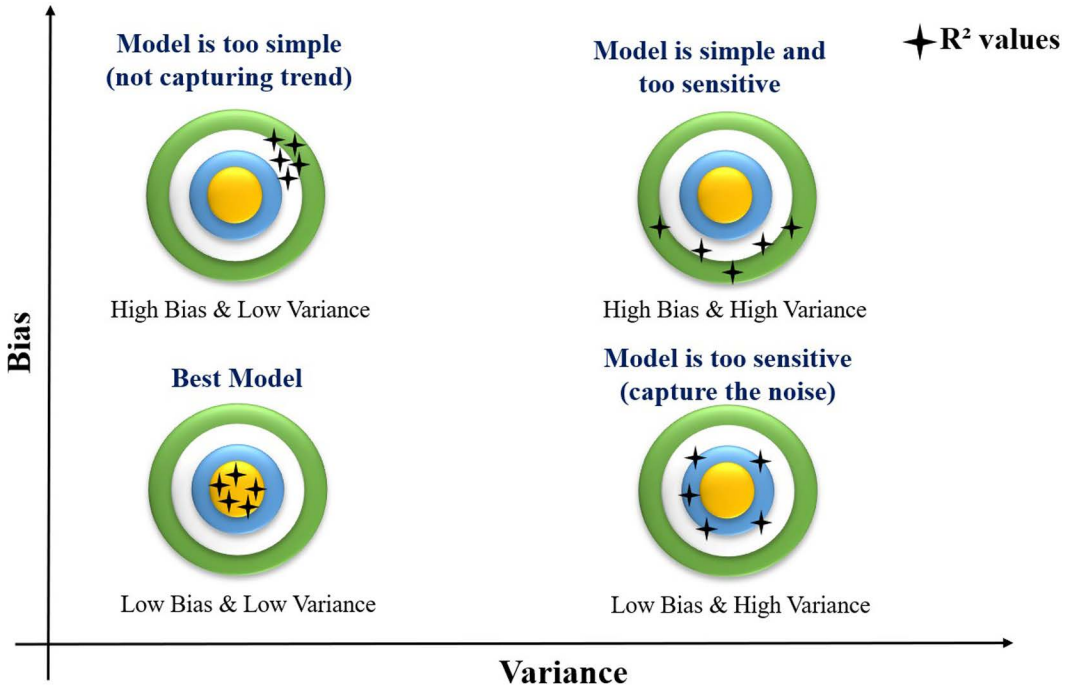


Fig 7. Concept of bias-variance tradeoff for all learning model.

https://doi.org/10.1371/journal.pone.0328880.g007

Table 4. K-fold cross validation on average R^2 results.

Iterations (k = 5)	R2 values for ML models			
	LR	PR	DTs	RF
Iteration 1	0.76	0.91	0.98	0.8
Iteration 2	0.91	0.91	0.97	0.99
Iteration 3	0.81	0.89	0.75	0.96
Iteration 4	−1.13	0.72	0.93	0.78
Iteration 5	0.87	0.91	0.54	0.65
Average R^2	0.444	0.868	0.834	0.836

https://doi.org/10.1371/journal.pone.0328880.t004

distributions signifying stronger influence. Features are ranked top to bottom in order of decreasing importance. An optimistic SHAP value signifies a feature's positive correlation with the output, A negative SHAP value implies that the feature contributes to lowering the predicted response. The SHAP value representation highlights catalyst concentration as a key determinant in biodiesel yield prediction, with its wider sections representing a substantial influence on the output. Furthermore, the feature value transition from blue to red suggests a positive correlation, implying that increasing the catalyst amount promotes greater biodiesel production within the range of 1–3% of catalyst.

The analysis presented in Fig (8b) indicates that catalyst concentration holds significant feature importance in biodiesel yield prediction, with a magnitude coefficient nearing 10. In contrast, RT and MOR exhibit considerably lower influence, with magnitude coefficients reaching only 0.5. This suggests that catalyst concentration accounts for approximately 95% of the impact on biodiesel yield, emphasizing its dominant role in process optimization.

The partial dependence analysis illustrates the influence trends of input variables on the output. Fig (8c) shows that catalyst concentration has a continuously increasing effect on yield, with higher concentrations (within the range of 1% to

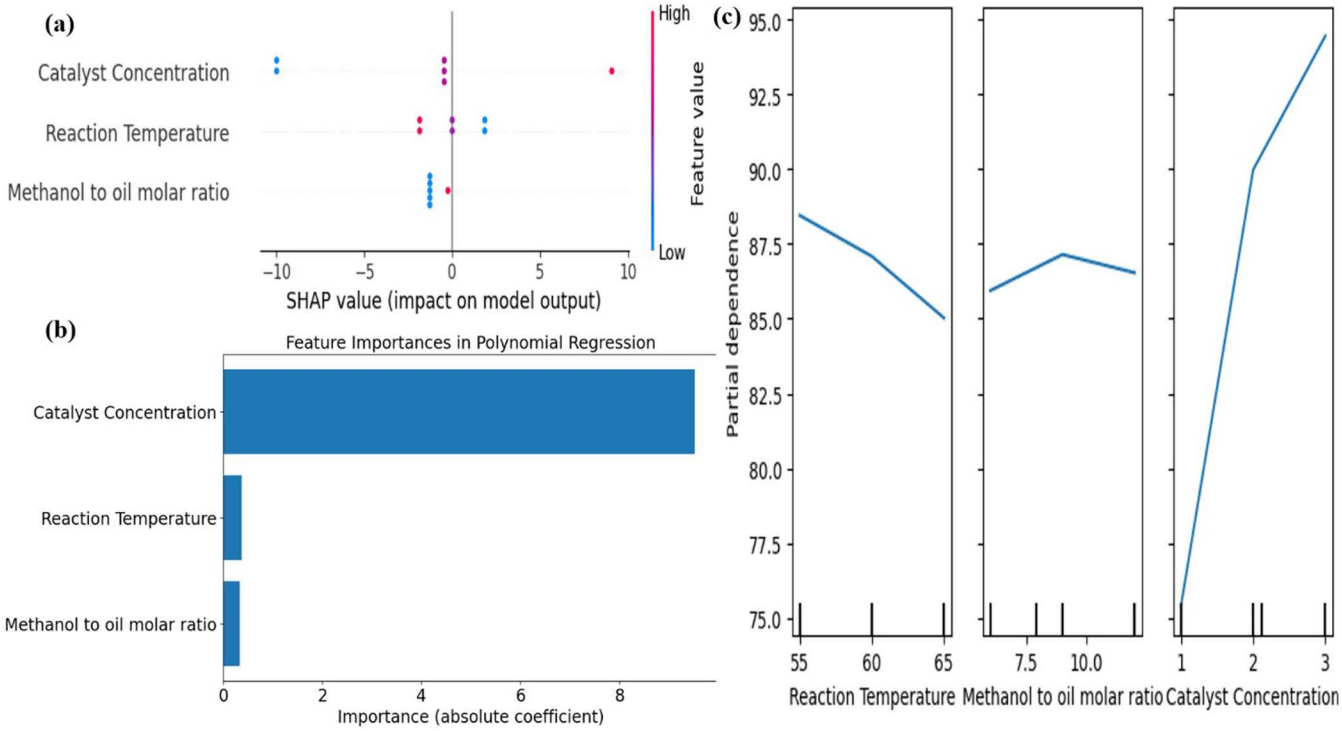


Fig 8. model interpretation a) SHAP b) Feature Importance c) Partial dependence.

https://doi.org/10.1371/journal.pone.0328880.g008

3%) leading to greater biodiesel yield. In contrast, the reaction temperature exhibits an inverse trend, where an increase in temperature (from 55°C to 65°C) results in a continuous decrease in yield. Meanwhile, the MOR initially enhances biodiesel yield as methanol concentration increases within a certain range. However, beyond this optimal range, the trend reverses, leading to a decline in yield.

The heatmap reveals three distinct relational patterns: positive, negative, and null correlations, with coefficient values ranging from –1 to +1. A coefficient nearing +1 denotes a direct relationship, where increasing input constraints correspond to a rise in output constraints. In contrast, values approaching –1 reflect an inverse relationship, indicating that higher input constraints lead to reduced output constraints. A coefficient near zero suggests no discernible linear association between the input and output variables [45,46]. Fig 9 highlights that catalyst concentration exhibits a strong relational pattern with biodiesel yield, as evidenced by a coefficient of 0.93. This recommends that greater catalyst concentrations lead to an enhanced biodiesel yield, reinforcing its significance in optimizing production efficiency

ML based optimization output by the best model

The optimization analyses were conducted by ML based PR learning model for finding the optimum process parameters for maximizing the biodiesel yield and it was visualized by two approaches namely 2D and the 3D by adapting PR model. In the 2D, one input variable was varied while the other two were kept constant, allowing for an isolated assessment of the impact of the selected variable on biodiesel yield. Conversely, in the 3D, two input variables were varied while keeping one constant, enabling a more comprehensive interaction analysis among multiple factors.

The optimization results obtained using the 2D are presented in Fig 10, illustrating the impact of different process parameters on biodiesel yield. In Fig 10a, the reaction temperature (RT) was varied while catalyst concentration (CC) and

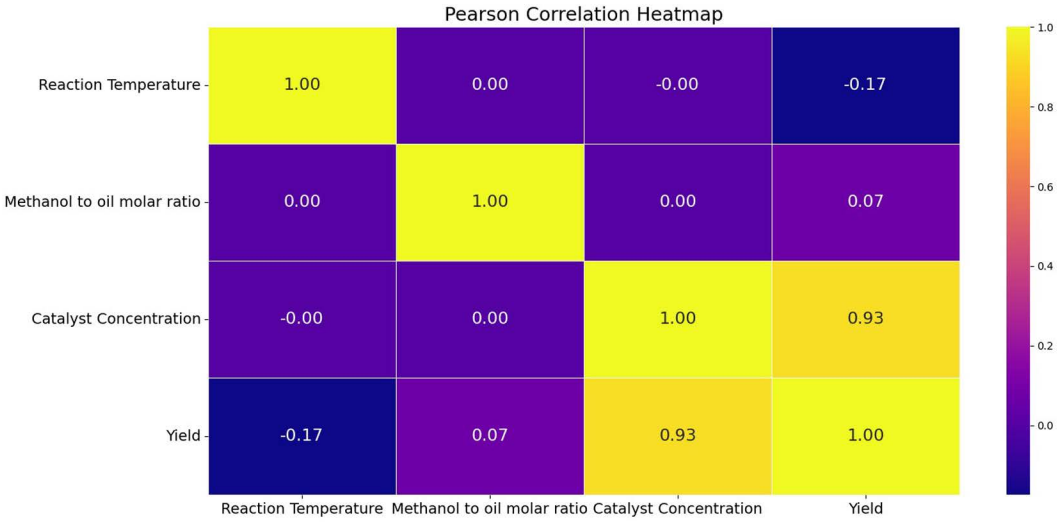


Fig 9. Model interpretation on Heatmap.

https://doi.org/10.1371/journal.pone.0328880.g009

methanol-to-oil molar ratio (MOR) remained constant. The analysis determined that the optimum reaction temperature for achieving maximum biodiesel yield was 59°C, leading to a biodiesel yield of 95.38%.

The results indicate a steady increase in biodiesel yield from 55°C, reaching an optimal level at 59°C. However, beyond this temperature, the yield begins to decline, suggesting an upper threshold where excessive thermal effects negatively impact reaction kinetics. This decline may be attributed to increased feedstock solubility at higher temperatures, potentially triggering side reactions that reduce overall efficiency. These findings align with previous research, which reports that beyond a certain temperature limit, increased feedstock solubility can lead to unintended reactions affecting biodiesel production efficiency [49,50].

Similarly, Fig 10b illustrates the optimization results for the MOR, keeping CC and RT constant. The study identified an optimum ratio of 9.21, achieving a maximum biodiesel yield of 95.38%. The results indicate a progressive increase in yield starting from a ratio of 6, but beyond 9.21, the yield declined. This decrease is likely due to excess methanol disrupting phase equilibrium, which negatively affects biodiesel conversion efficiency. Additionally, an excessive methanol ratio may promote unwanted by-product formation, complicate separation and lowering the overall biodiesel yield These findings align with previous research, which reports that exceeding an optimal methanol ratio can lead to reduced conversion efficiency and increased by-product formation [51,52].

Furthermore, Fig 10c presents the optimization analysis for CC, with MOR and RT held constant. The results confirmed that an optimum CC of 2.96% produced the highest biodiesel yield of 95.38%. The yield increased consistently from 1%, but beyond 2.96%, the reaction may reach a saturation point where additional catalyst no longer enhances conversion efficiency. Excess catalyst could lead to agglomeration or increased viscosity in the reaction mixture, reducing mass transfer efficiency and limiting further yield improvements. These findings align with previous research, which reports that exceeding the optimal catalyst concentration can result in diminished conversion rates due to viscosity changes and agglomeration effects, ultimately affecting biodiesel yield efficiency [53,54].

The optimization results obtained using the 3D surface plot in Fig 11a illustrate the interactive effects of RT and MOR on biodiesel yield, with CC held constant. The surface reveals a nonlinear synergistic relationship between these two parameters: increasing the RT enhances the transesterification rate by accelerating molecular interactions, with the optimum observed at 59°C. Concurrently, raising the MOR shifts the reaction equilibrium toward biodiesel formation, with the

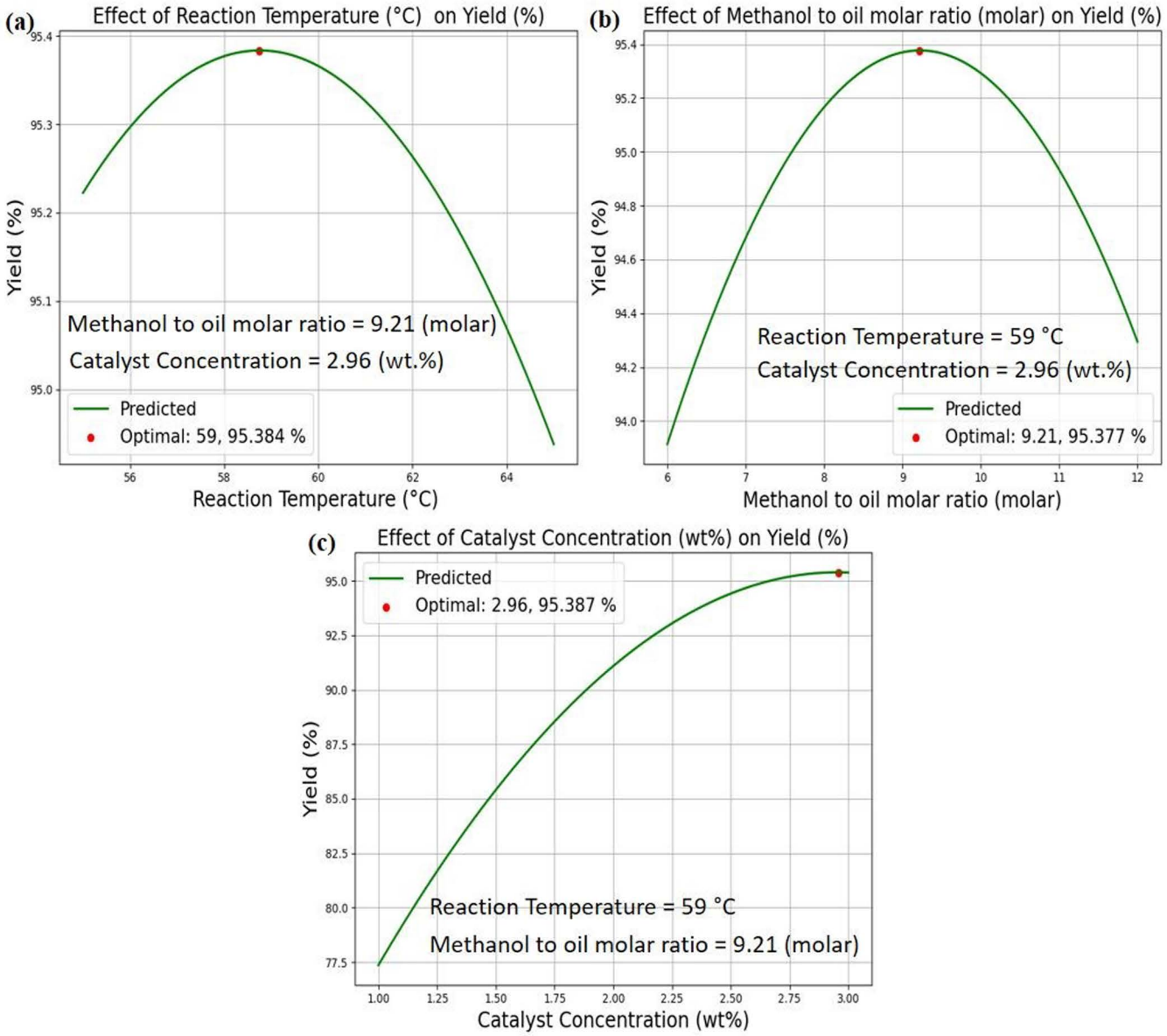


Fig 10. Optimum process parameters by best model a) Reaction Temperature (RT) b) Methanol to oil molar ratio (MOR), c) catalyst concentration (CC).

https://doi.org/10.1371/journal.pone.0328880.g010

optimum at 9.21. However, beyond these thresholds, the biodiesel yield begins to decline slightly, which can be attributed to excess methanol lowering the reactant concentration per unit volume (dilution effect), causing inefficient mixing and phase separation, and elevated temperatures accelerating methanol evaporation, which reduces the actual methanol available for the reaction. The curved response surface suggests that neither parameter alone maximizes yield; rather, their combined balance leads to the peak biodiesel yield of 95.38%. These findings align with previous research, which reports that exceeding certain thresholds such as RT above 60°C and MOR beyond approximately 9:1 can lead to unintended side reactions and phase imbalance due to increased feedstock solubility and reagent dilution, ultimately compromising biodiesel production efficiency [55–57].

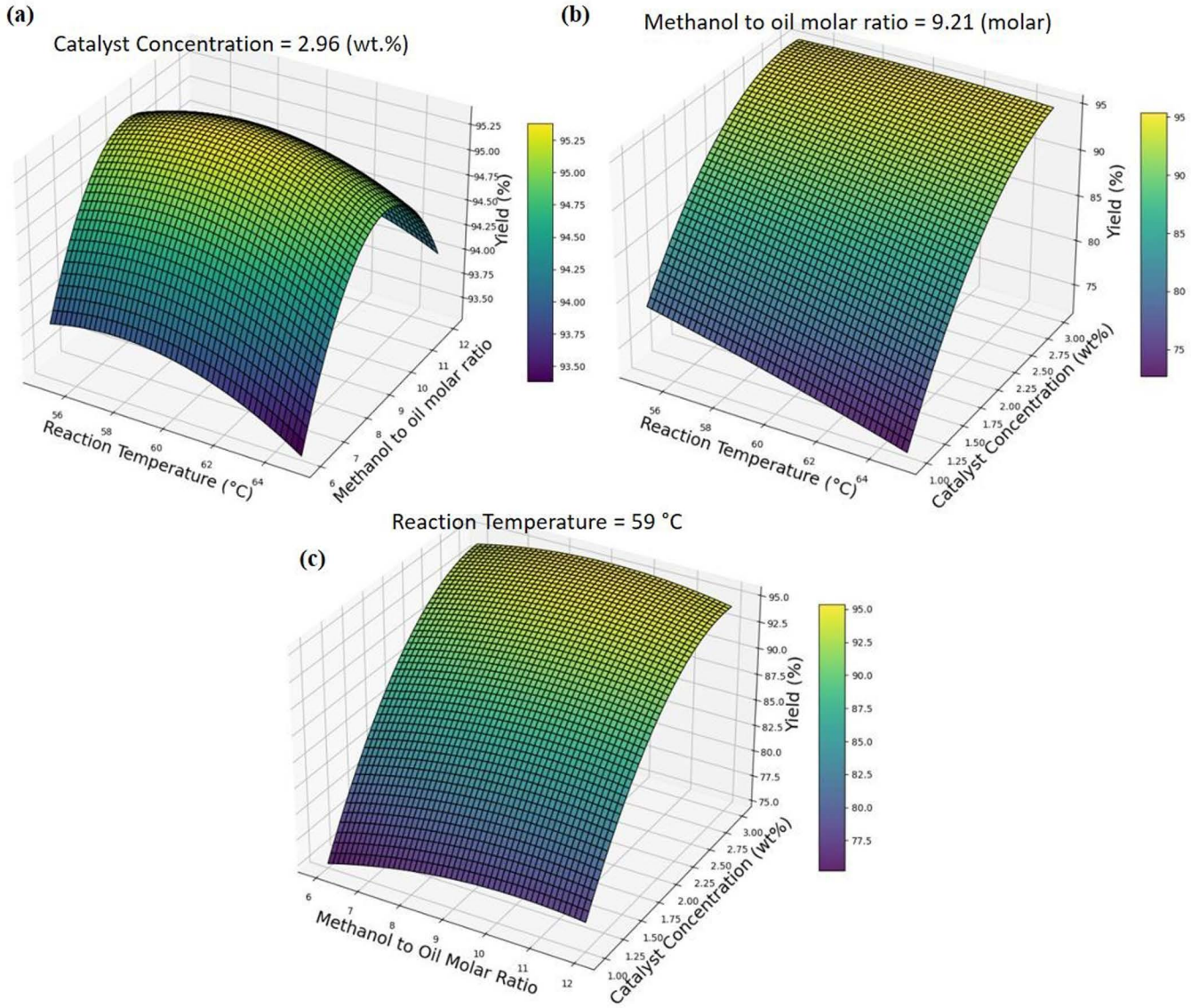


Fig 11. 3D illustration of optimum process parameters on maximum yield a) Reaction temperature (RT) versus Methanol to oil molar ratio (MOR), b) Reaction temperature (RT) versus catalyst concentration (CC), c) Methanol to oil molar ratio (MOR) versus catalyst concentration (CC).

https://doi.org/10.1371/journal.pone.0328880.g011

The Fig 11b illustrate the interactive effects of RT and CC on biodiesel yield, with the MOR held constant. The surface reveals a nonlinear coupled relationship between these two parameters: increasing the reaction temperature enhances transesterification kinetics by lowering activation energy barriers and promoting more rapid molecular collisions, with the optimal yield observed at 59°C. Simultaneously, increasing the CC provides more active sites for the reaction, effectively accelerating the conversion of triglycerides to methyl esters. However, beyond the optimal catalyst loading of 2.96%, the biodiesel production exhibits a downward trend. This drop is likely due to catalyst agglomeration or excessive base availability, which can promote saponification reactions and result in soap formation, thereby hindering mass transfer, increasing viscosity, and complicating product separation. The curvature of the response surface demonstrates that a balanced combination of both elevated temperature and adequate catalyst concentration is required to achieve the highest

yield, rather than the independent maximization of either factor. These findings are consistent with prior studies that report excessive catalyst dosages and high thermal input can introduce undesired side reactions and separation difficulties, ultimately reducing conversion efficiency and final product purity [53,58].

The Fig 11c illustrate the interactive effects of MOR and CC on biodiesel yield, with the RT held constant. The response surface reveals a nonlinear coupled relationship between these two parameters: increasing the MOR initially drives the transesterification reaction forward enhancing biodiesel conversion. Similarly, increasing the catalyst concentration introduces additional active sites, accelerating the reaction rate and supporting more efficient breakdown of triglycerides into methyl esters. However, beyond the optimal point identified near 9.21 molar ratio and 2.96% catalyst concentration the yield begins to plateau or decline. This behavior is attributed to methanol oversaturation, which can dilute reactant concentration, interfere with phase separation, and reduce the mass transfer rate. Concurrently, excessive catalyst can increase system viscosity or promote particle agglomeration and soap formation, all of which hinder effective mixing and limit reaction efficiency. The curvature of the response surface confirms that biodiesel yield is maximized only when both parameters are simultaneously optimized, rather than altered independently. These observations are consistent with prior research indicating that overloading either methanol or catalyst beyond their critical thresholds can lead to undesirable reaction conditions and reduced conversion performance [59–61]. From the optimizing analysis using PR model the optimised process parameters were founded using polynomial regression equation and it was mentioned in Table 5.

Polynomial regression equation. Yield = 60.2702 + (0.0000 * 1) + (0.2694 * Reaction Temperature) + (2.6898 * Methanol to oil molar) + (10.0242 * Catalyst Concentration) + (−0.0114 * Reaction Temperature^2) + (0.0072 * Reaction Temperature * Methanol to oil molar ratio) + (0.3353 * Reaction Temperature * Catalyst Concentration) + (−0.1406 * Methanol to oil molar ratio^2) + (−0.1714 * Methanol to oil molar ratio * Catalyst Concentration) + (−4.7206 * Catalyst Concentration^2)

Conclusion

The machine learning validation and optimization study successfully identified the most reliable predictive model and optimal process parameters for biodiesel production.

- Polynomial Regression (PR) demonstrated the highest prediction accuracy, exhibiting strong alignment between experimental and predicted biodiesel yield ($R^2 = 0.956$). Its reliability was further validated through k-fold cross-validation, yielding a mean R^2 score of 0.868, ensuring robust and stable predictions across varying conditions. These findings underscore PR's effectiveness in capturing complex interactions within biodiesel production systems.
- Catalyst concentration identified as the dominant contributing factor, contributing roughly 95% of the impact on biodiesel yield, as confirmed by SHAP values and feature importance analysis. The Pearson correlation heatmap further reinforced this observation, showing a strong positive correlation (0.93) with biodiesel yield, highlighting the necessity of precise catalyst optimization to maximize conversion efficiency
- Optimization analysis using the PR model successfully identified the optimal process conditions for achieving maximum biodiesel yield (95.38%), with a RT of 59°C, a MOR of 9.21, and a CC of 2.96%. These results emphasize the critical role of fine-tuning process parameters to ensure efficient and stable biodiesel production
- Overall, this study validates the efficacy of PR-based optimization in enlightening biodiesel yield prediction and refining key process conditions, offering a reliable approach for enhancing production efficiency and sustainability

Table 5. Optimum process parameters for maximum biodiesel yield.

Reaction Temperature (RT)	Methanal to oil Molar ratio (MOR)	Catalyst concentration (CC)	Yield
59°C	9.21	2.96	95.38%

https://doi.org/10.1371/journal.pone.0328880.t005

- Although the findings are encouraging, several constraints remain, offering valuable directions for future research and methodological enhancement. The dataset, restricted to 27 experimental data points, may limit predictive generalizability, necessitating expansion to improve model robustness. Additionally, reliance on regression models constrains predictive flexibility, making ensemble learning methods for enhancing predictive stability. Furthermore, incorporating additional input features such as mixing speed, feedstock composition, and reaction time could further refine biodiesel yield optimization. These advancements will improve process efficiency and expand the applications of machine learning in sustainable biodiesel production.

Author contributions

Conceptualization: Wei Shangzhi.

Data curation: Mohd. Zulkifly Abdullah.

Formal analysis: P. V. Elumalai.

Methodology: Allam Sangeetha.

Supervision: Xu Yong.

Writing – original draft: Krishnamoorthy Ramalingam.

Writing – review & editing: Nasim Hasan.

References

1. Vellaiyan S, Kandasamy M, Nagappan B, Gupta S, Ramalingam K, Devarajan Y. Optimization study for efficient and cleaner production of waste-derived biodiesel through fuel modification and its validation. Process Integration and Optimization for Sustainability. 2024;8(3):939–52.
2. Ramalingam K, Vellaiyan S, Gupta M S, Nagappan B, Faujdar PK, Chandran D, et al. An experimental and ANN analysis of ammonia energy integration in biofuel powered low-temperature combustion engine to enhance cleaner combustion. Case Studies in Thermal Engineering. 2024;63:105284. https://doi.org/10.1016/j.csite.2024.105284
3. Lee J, Lin K-YA, Jung S, Kwon EE. Hybrid renewable energy systems involving thermochemical conversion process for waste-to-energy strategy. Chemical Engineering Journal. 2023;452:139218. https://doi.org/10.1016/j.cej.2022.139218
4. Sandaka BP, Kumar J. Alternative vehicular fuels for environmental decarbonization: A critical review of challenges in using electricity, hydrogen, and biofuels as a sustainable vehicular fuel. Chemical Engineering Journal Advances. 2023;14:100442. https://doi.org/10.1016/j.ceja.2022.100442
5. Shafiei E, Davidsdottir B, Leaver J, Stefansson H, Asgeirsson EI. Comparative analysis of hydrogen, biofuels and electricity transitional pathways to sustainable transport in a renewable-based energy system. Energy. 2015;83:614–27. https://doi.org/10.1016/j.energy.2015.02.071
6. Narayanan R D, N V, Rajkumar S, Thangaraja J, M S, Devarajan Y, et al. Techno-economic review assessment of hydrogen utilization in processing the natural gas and biofuels. International Journal of Hydrogen Energy. 2023;48(55):21294–312. https://doi.org/10.1016/j.ijhydene.2022.09.101
7. Ijaz Malik MA, Mujtaba MA, Kalam MA, Silitonga AS, Ikram A. Recent advances in hydrogen supplementation to promote biomass fuels for reducing greenhouse gases. International Journal of Hydrogen Energy. 2024;49:463–87. https://doi.org/10.1016/j.ijhydene.2023.09.154
8. Alonso DM, Bond JQ, Dumesic JA. Catalytic conversion of biomass to biofuels. Green Chem. 2010;12(9):1493. https://doi.org/10.1039/c004654j
9. Anwar M. Biodiesel feedstocks selection strategies based on economic, technical, and sustainable aspects. Fuel. 2021;283:119204. https://doi.org/10.1016/j.fuel.2020.119204
10. Lin C-Y. The Influences of Promising Feedstock Variability on Advanced Biofuel Production: A Review. Journal of Marine Science and Technology. 2022;29(6):714–30. https://doi.org/10.51400/2709-6998.2552
11. Al-Bawwat AK, Gomaa MR, Cano A, Jurado F, Alsbou EM. Extraction and characterization of Cucumis melon seeds (Muskmelon seed oil) biodiesel and studying its blends impact on performance, combustion, and emission characteristics in an internal combustion engine. Energy Conversion and Management: X. 2024;23:100637. https://doi.org/10.1016/j.ecmx.2024.100637
12. Devarajan Y, Beemkumar N, Ganesan S, Arunkumar T. An experimental study on the influence of an oxygenated additive in diesel engine fuelled with neat papaya seed biodiesel/diesel blends. Fuel. 2020;268:117254. https://doi.org/10.1016/j.fuel.2020.117254
13. Hundie KB, Shumi LD, Bullo TA. Investigation of biodiesel production parameters by transesterification of watermelon waste oil using definitive screening design and produced biodiesel characterization. South African Journal of Chemical Engineering. 2022;41:140–9. https://doi.org/10.1016/j.sajce.2022.06.002
14. Krishnamurthy KN, Sridhara SN, Ananda Kumar CS. Optimization and kinetic study of biodiesel production from Hydnocarpus wightiana oil and dairy waste scum using snail shell CaO nano catalyst. Renewable Energy. 2020;146:280–96. https://doi.org/10.1016/j.renene.2019.06.161

15. Elkelawy M, Bastawissi HA-E, Esmaeil KK, Radwan AM, Panchal H, Sadasivuni KK, et al. Maximization of biodiesel production from sunflower and soybean oils and prediction of diesel engine performance and emission characteristics through response surface methodology. Fuel. 2020;266:117072. https://doi.org/10.1016/j.fuel.2020.117072
16. Ma X, Liu F, Helian Y, Li C, Wu Z, Li H, et al. Current application of MOFs based heterogeneous catalysts in catalyzing transesterification/esterification for biodiesel production: A review. Energy Conversion and Management. 2021;229:113760. https://doi.org/10.1016/j.enconman.2020.113760
17. Shan R, Lu L, Shi Y, Yuan H, Shi J. Catalysts from renewable resources for biodiesel production. Energy Conversion and Management. 2018;178:277–89. https://doi.org/10.1016/j.enconman.2018.10.032
18. Jayakumar M, Karmegam N, Gundupalli MP, Bizuneh Gebeyehu K, Tessema Asfaw B, Chang SW, et al. Heterogeneous base catalysts: Synthesis and application for biodiesel production - A review. Bioresour Technol. 2021;331:125054. https://doi.org/10.1016/j.biortech.2021.125054 PMID: 33832828
19. Mukhtar A, Saqib S, Lin H, Hassan Shah MU, Ullah S, Younas M, et al. Current status and challenges in the heterogeneous catalysis for biodiesel production. Renewable and Sustainable Energy Reviews. 2022;157:112012. https://doi.org/10.1016/j.rser.2021.112012
20. Mandari V, Devarai SK. Biodiesel Production Using Homogeneous, Heterogeneous, and Enzyme Catalysts via Transesterification and Esterification Reactions: a Critical Review. Bioenergy Res. 2022;15(2):935–61. https://doi.org/10.1007/s12155-021-10333-w PMID: 34603592
21. Rabie AM, Shaban M, Abukhadra MR, Hosny R, Ahmed SA, Negm NA. Diatomite supported by CaO/MgO nanocomposite as heterogeneous catalyst for biodiesel production from waste cooking oil. Journal of Molecular Liquids. 2019;279:224–31. https://doi.org/10.1016/j.molliq.2019.01.096
22. Santoso A, Sukarianingsih D, Sari RM. Optimization of Synthesis of Biodiesel from Jatropha curcas L. with Heterogeneous Catalyst of CaO and MgO by Transesterification Reaction Using Microwave. J Phys: Conf Ser. 2018;1093:012047. https://doi.org/10.1088/1742-6596/1093/1/012047
23. Rodríguez-Ramírez R, Sosa-Rodríguez FS, Vazquez-Arenas J. Zinc oxide-co-sodium zirconate: A fast heterogeneous catalyst for biodiesel production from soybean oil. Journal of Environmental Chemical Engineering. 2022;10(4):108191. https://doi.org/10.1016/j.jece.2022.108191
24. Rahman WU, Khan RIA, Ahmad S, Yahya SM, Khan ZA, Rokhum SL, et al. Valorizing waste palm oil towards biodiesel production using calcareous eggshell based heterogeneous catalyst. Bioresource Technology Reports. 2023;23:101584. https://doi.org/10.1016/j.biteb.2023.101584
25. Chiang Y-D, Dutta S, Chen C-T, Huang Y-T, Lin K-S, Wu JCS, et al. Functionalized Fe3O4@silica core-shell nanoparticles as microalgae harvester and catalyst for biodiesel production. ChemSusChem. 2015;8(5):789–94. https://doi.org/10.1002/cssc.201402996 PMID: 25477296
26. Lukić I, Krstić J, Jovanović D, Skala D. Alumina/silica supported K2CO3 as a catalyst for biodiesel synthesis from sunflower oil. Bioresour Technol. 2009;100(20):4690–6. https://doi.org/10.1016/j.biortech.2009.04.057 PMID: 19477122
27. Navas MB, Ruggera JF, Lick ID, Casella ML. A sustainable process for biodiesel production using Zn/Mg oxidic species as active, selective and reusable heterogeneous catalysts. Bioresour Bioprocess. 2020;7(1). https://doi.org/10.1186/s40643-019-0291-3
28. Onukwuli DO, Emembolu LN, Ude CN, Aliozo SO, Menkiti MC. Optimization of biodiesel production from refined cotton seed oil and its characterization. Egyptian Journal of Petroleum. 2017;26(1):103–10. https://doi.org/10.1016/j.ejpe.2016.02.001
29. Odetoye TE, Agu JO, Ajala EO. Biodiesel production from poultry wastes: Waste chicken fat and eggshell. Journal of Environmental Chemical Engineering. 2021;9(4):105654. https://doi.org/10.1016/j.jece.2021.105654
30. Aghbashlo M, Peng W, Tabatabaei M, Kalogirou S, Soltanian S, Hosseinzadeh-Bandbafha H, et al. Machine learning technology in biodiesel research: A review. Progress in Energy and Combustion Science. 2021. https://doi.org/10.1016/J.PECS.2021.100904
31. Ishola NB, Epelle EI, Betiku E. Machine learning approaches to modeling and optimization of biodiesel production systems: State of art and future outlook. Energy Conversion and Management: X. 2024;23:100669. https://doi.org/10.1016/j.ecmx.2024.100669
32. Sultana N, Hossain SMZ, Abusaad M, Alanbar N, Senan Y, Razzak SA. Prediction of biodiesel production from microalgal oil using Bayesian optimization algorithm-based machine learning approaches. Fuel. 2022;309:122184. https://doi.org/10.1016/j.fuel.2021.122184
33. Sumayli A, Alshahrani SM. Modeling and prediction of biodiesel production by using different artificial intelligence methods: Multi-layer perceptron (MLP), Gradient boosting (GB), and Gaussian process regression (GPR). Arabian Journal of Chemistry. 2023;16(7):104801. https://doi.org/10.1016/j.arabjc.2023.104801
34. Kamal Abdelbasset W, Elkholi SM, Jade Catalan Opulencia M, Diana T, Su C-H, Alashwal M, et al. Development of multiple machine-learning computational techniques for optimization of heterogenous catalytic biodiesel production from waste vegetable oil. Arabian Journal of Chemistry. 2022;15(6):103843. https://doi.org/10.1016/j.arabjc.2022.103843
35. Wong T-T, Yeh P-Y. Reliable Accuracy Estimates from k-Fold Cross Validation. IEEE Trans Knowl Data Eng. 2020;32(8):1586–94. https://doi.org/10.1109/tkde.2019.2912815
36. Jiang G, Wang W. Error estimation based on variance analysis of k -fold cross-validation. Pattern Recognition. 2017;69:94–106. https://doi.org/10.1016/j.patcog.2017.03.025
37. Balamurugan S, Manoj Prabhakar BS, Kulkarni SS, Kalos PS, Dongre G, Karthikeyan C. Biodiesel synthesis from composite oil utilizing banana peel (Musa paradisiaca) derived catalyst and process parameter optimization using particle swarm method. Journal of the Indian Chemical Society. 2025;102(5):101664. https://doi.org/10.1016/j.jics.2025.101664
38. Xing Y, Zheng Z, Sun Y, Agha Alikhani M. A Review on Machine Learning Application in Biodiesel Production Studies. International Journal of Chemical Engineering. 2021;2021:1–12. https://doi.org/10.1155/2021/2154258

39. Aghbashlo M, Peng W, Tabatabaei M, Kalogirou SA, Soltanian S, Hosseinzadeh-Bandbafha H, et al. Machine learning technology in biodiesel research: A review. Prog Energy Combust Sci. 2021;85:100904.
40. Aslan V, Eryilmaz T. Polynomial regression method for optimization of biodiesel production from black mustard (Brassica nigra L.) seed oil using methanol, ethanol, NaOH, and KOH. Energy. 2020;209:118386. https://doi.org/10.1016/j.energy.2020.118386
41. Bukkarapu KR, Krishnasamy A. Investigations on the applicability of machine learning algorithms to optimize biodiesel composition for improved engine fuel properties. International Journal of Engine Research. 2024;25(7):1299–314. https://doi.org/10.1177/14680874241227540
42. Ghasemzadeh H, Hillman RE, Mehta DD. Toward Generalizable Machine Learning Models in Speech, Language, and Hearing Sciences: Estimating Sample Size and Reducing Overfitting. J Speech Lang Hear Res. 2024;67(3):753–81. https://doi.org/10.1044/2023_JSLHR-23-00273 PMID: 38386017
43. Bami Z, Behnampour A, Doosti H. A new flexible train-test split algorithm, an approach for choosing among the hold-out, k-fold cross-validation, and hold-out iteration. ArXiv. 2025. https://doi.org/10.48550/arXiv.2501.06492
44. Rainio O, Teuho J, Klén R. Evaluation metrics and statistical tests for machine learning. Sci Rep. 2024;14(1):6086. https://doi.org/10.1038/s41598-024-56706-x PMID: 38480847
45. Vu HL, Ng KTW, Richter A, An C. Analysis of input set characteristics and variances on k-fold cross validation for a Recurrent Neural Network model on waste disposal rate estimation. J Environ Manage. 2022;311:114869. https://doi.org/10.1016/j.jenvman.2022.114869 PMID: 35287077
46. Lumumba V, Kiprotich D, Mpaine M, Makena N, Kavita M. Comparative Analysis of Cross-Validation Techniques: LOOCV, K-folds Cross-Validation, and Repeated K-folds Cross-Validation in Machine Learning Models. AJTAS. 2024;13(5):127–37. https://doi.org/10.11648/j.ajtas.20241305.13
47. Vega García M, Aznarte JL. Shapley additive explanations for NO2 forecasting. Ecological Informatics. 2020;56:101039. https://doi.org/10.1016/j.ecoinf.2019.101039
48. Nordin N, Zainol Z, Mohd Noor MH, Chan LF. An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach. Asian Journal of Psychiatry. 2023;79:103316. https://doi.org/10.1016/j.ajp.2022.103316
49. Faisal F, Rasul MG, Chowdhury AA, Jahirul MI. Optimisation of Process Parameters to Maximise the Oil Yield from Pyrolysis of Mixed Waste Plastics. Sustainability. 2024;16(7):2619. https://doi.org/10.3390/su16072619
50. Bezergianni S, Dimitriadis A, Kalogianni A, Pilavachi PA. Hydrotreating of waste cooking oil for biodiesel production. Part I: Effect of temperature on product yields and heteroatom removal. Bioresour Technol. 2010;101(17):6651–6. https://doi.org/10.1016/j.biortech.2010.03.081 PMID: 20395136
51. Ghoreishi SM, Moein P. Biodiesel synthesis from waste vegetable oil via transesterification reaction in supercritical methanol. The Journal of Supercritical Fluids. 2013;76:24–31. https://doi.org/10.1016/j.supflu.2013.01.011
52. Khan E, Ozaltin K, Spagnuolo D, Bernal-Ballen A, Piskunov MV, Di Martino A. Biodiesel from Rapeseed and Sunflower Oil: Effect of the Transesterification Conditions and Oxidation Stability. Energies. 2023;16(2):657. https://doi.org/10.3390/en16020657
53. Salmasi MZ, Kazemeini M, Sadjadi S. Transesterification of sunflower oil to biodiesel fuel utilizing a novel K2CO3/Talc catalyst: Process optimizations and kinetics investigations. Industrial Crops and Products. 2020;156:112846. https://doi.org/10.1016/j.indcrop.2020.112846
54. Adepoju T, Ukanwa K, Eyibio U, Etim V, Eloka-Eboka A, Balogun T. Biodiesel production from renewable biosources ternary oil blends and its kinetic-thermodynamic parameters using Eyring Polanyi and Gibb's-Duhem equations. South African Journal of Chemical Engineering. 2023;44:103–12. https://doi.org/10.1016/j.sajce.2023.01.004
55. Elumalai PV. Graphene Oxide Nanoparticle Blended Tamanu Methyl Ester as a Promising Alternative Fuel for Unmodified Compression Ignition Engine. Int Res J multidiscip Technovation. 2025;151–64. https://doi.org/10.54392/irjmt25111
56. Fadhil AB, Ahmed AI. Production of mixed methyl/ethyl esters from waste fish oil through transesterification with mixed methanol/ethanol system. Chemical Engineering Communications. 2018;205(9):1157–66. https://doi.org/10.1080/00986445.2018.1434514
57. Tsaoulidis D, Garciadiego-Ortega E, Angeli P. Intensified biodiesel production from waste cooking oil and flow pattern evolution in small-scale reactors. Front Chem Eng. 2023;5. https://doi.org/10.3389/fceng.2023.1144009
58. Daramola MO, Mtshali K, Senokoane L, Fayemiwo OM. Influence of operating variables on the transesterification of waste cooking oil to biodiesel over sodium silicate catalyst: A statistical approach. Journal of Taibah University for Science. 2016;10(5):675–84. https://doi.org/10.1016/j.jtusci.2015.07.008
59. Mansoorsamaei Z, Mowla D, Esmaeilzadeh F, Dashtian K. Sustainable biodiesel production from waste cooking oil using banana peel biochar-Fe2O3/Fe2K6O5 magnetic catalyst. Fuel. 2024;357:129821. https://doi.org/10.1016/j.fuel.2023.129821
60. Erchamo YS, Mamo TT, Workneh GA, Mekonnen YS. Improved biodiesel production from waste cooking oil with mixed methanol-ethanol using enhanced eggshell-derived CaO nano-catalyst. Sci Rep. 2021;11(1):6708. https://doi.org/10.1038/s41598-021-86062-z PMID: 33758293
61. Munimathan A, Rajendran S, Tripathi AK, Jayabalan J, Palanivel V, Varma R, et al. ML techniques increasing the power factor of a compression ignition engine that is powered by Annona biodiesel using SATACOM. Sci Rep. 2025;15(1):11669. https://doi.org/10.1038/s41598-025-91162-1 PMID: 40188138