Green Valorization of Cassia Fistula Biodiesel: Experimental Insights, RSM–Machine Learning Optimization, and Sustainable Engine Applications

K.Saravanakumar[a], J.Paul James Thadhani [b], R.Shenbagaraj[c] ,N.Saravanan[d]

[a]Department of Mechanical Engineering, RVS Technical Campus-Coimbatore, Coimbatore, Tamilnadu, India-641402.

[b]Department of Automobile Engineering, Anna university Campus, Madras Institute of Technology, Chennai.Tamilnadu, India-600044.

[c] Department of Mechanical Engineering, RVS Technical Campus-Coimbatore, Coimbatore, Tamilnadu, India-641402.

[d] Department of Mechanical Engineering, Erode Sengunthar Engineering College, Erode, Tamilnadu, India-638057

EMail addresses: saravananthermal@gmail.com (K.Saravanakumar), pauljamesthadhani@gmail.com (J.Paul James Thadhani), rbshenbarvs@gmail.com(R.Shenbagaraj), saranvimal@gmail.com (N.Saravanan),

[a]Tel.: +91 8098911678

Abstract

The urgent global transition toward sustainable energy sources necessitates green alternative fuels that mitigate environmental burdens while ensuring compatibility with compression ignition engines. Biodiesel represents a renewable and biodegradable fuel, but variability in performance and emissions across feedstocks requires advanced green engineering strategies. In this study, a non-conventional seed oil (Cassia fistula), rarely explored in biodiesel research, was converted via transesterification using a waste-derived CaO catalyst (eggshells), demonstrating circular economy principles. Biodiesel blends (B10, B20, B30) were tested in a single-cylinder diesel engine under varying loads and injection timings. Experimental responses (BTE, BSFC, NOx, CO, HC, CO2, cylinder pressure) were optimized using Response Surface Methodology (RSM, CCD design). Machine learning (Random Forest, Support Vector Regression, and Artificial Neural Networks) enabled predictive modeling of performance and emissions. Results confirmed B20 as the optimal trade-off between efficiency and emissions, with substantial potential for reducing lifecycle carbon intensity. RSM analysis revealed significant interactions between blend ratio and injection timing, validated by ANOVA ($p < 0.001$). Machine learning predictions achieved high accuracy (Random Forest R^2 =

0.97 for BTE). This integrated green framework—combining waste-derived catalyst utilization, experimental validation, statistical optimization, and computational intelligence—advances biodiesel research toward scalable, sustainable, and low-carbon energy systems.

Introduction
Growing environmental concerns, depleting fossil reserves, and stringent emission regulations have driven the exploration of renewable fuels. Biodiesel is an attractive alternative due to its biodegradability, low sulfur, and compatibility with diesel engines. However, feedstock type, fuel properties, and engine operating conditions strongly influence performance. Most biodiesel studies focus on common feedstocks (Jatropha, Karanja, Mahua, Neem, Rubber seed, Soybean, Palm), leaving significant scope for under-utilized seeds.

Literature Review
Biodiesel research has evolved significantly over the past two decades, with most studies focusing on well-known non-edible oils such as Jatropha, Pongamia, Mahua, Neem, and Rubber seed. While these feedstocks have been extensively characterized, there remains considerable scope for the exploration of under-utilized seeds such as Cassia fistula, which offers high oil yield, regional availability, and limited competition with food resources. Recent work emphasizes process intensification strategies such as microwave- and ultrasound-assisted transesterification, supercritical alcohol methods, and the use of heterogeneous catalysts including CaO derived from eggshells or agricultural residues. These approaches improve biodiesel yield, reduce reaction time, and enhance environmental sustainability. Engine testing of biodiesel consistently demonstrates reduced CO and HC emissions compared to diesel, though NOx emissions often increase, necessitating advanced optimization. Response Surface Methodology (RSM) has been widely applied for multi-variable optimization of performance and emissions, and recent advances (2020–2024) show its effectiveness when combined with machine learning models such as Random Forest, Support Vector Regression, and Artificial Neural Networks. These hybrid approaches allow accurate prediction of Brake Thermal Efficiency (BTE), Brake Specific Fuel Consumption (BSFC), and emissions, thus providing robust frameworks for engine calibration and optimization. Recent studies on Cassia fistula biodiesel confirm its potential as a sustainable alternative fuel, but limited

research exists on its engine application coupled with statistical and computational optimization frameworks, highlighting the novelty of the present work.

• Process intensification via supercritical alcohols, ultrasound, and microwave-assisted transesterification shortens reaction times and can improve yields.

• Heterogeneous catalysts (e.g., CaO from eggshells, ZnO, waste shell ash) reduce wastewater and enable catalyst reuse.

• RSM and metaheuristic hybrids enable multi-response optimization (BTE↑, BSFC↓, NOx↓) under practical constraints.

• Machine learning (RF, SVR, ANN, XGBoost) increasingly supports real-time prediction and calibration, with RF often leading for tabular, nonlinear datasets.

Materials and Methods

Cassia fistula seeds were selected as a novel non-edible feedstock due to their availability and limited prior use in biodiesel research. Seeds were dried at 70°C, mechanically expelled for oil, and further solvent-extracted to maximize yield. The crude oil was transesterified with methanol using a heterogeneous CaO catalyst derived from calcined eggshells, with a methanol-to-oil ratio of 6:1 and reaction temperature of 60–65°C for 90 minutes. The produced biodiesel was washed, dried, and characterized for key properties such as density, viscosity, calorific value, cetane number, flash point, and FAME composition using ASTM standards. Blends of biodiesel with diesel (B10, B20, B30) were prepared by volume and homogenized under nitrogen blanket. Engine experiments were conducted on a single-cylinder, four-stroke, water-cooled diesel engine with an eddy current dynamometer. Tests were performed at four load levels (25%, 50%, 75%, 100%) and three injection timings (16°, 17°, and 18° BTDC). Performance metrics measured included Brake Thermal Efficiency (BTE) and Brake Specific Fuel Consumption (BSFC), while emissions (CO, HC, NOx, CO_2) and cylinder pressure traces were also recorded. RSM with Central Composite Design (CCD) was employed to model interactions between blend ratio, load, and injection timing. Analysis of variance (ANOVA) was conducted to determine significance. Additionally, machine learning models (Random Forest, SVR, ANN) were trained on the experimental dataset to predict engine performance and emissions, with model performance validated using R^2, RMSE, and MAE metrics.

Methodology

Response Surface Methodology (RSM)
Central Composite Design (CCD) was applied with factors: blend ratio, injection timing, and load. ANOVA was used to test model significance. Optimization targeted maximum BTE and minimum BSFC/NOx.

Machine Learning Models
Random Forest (n_estimators=300–500), SVR (RBF kernel, C=10–100), and ANN (3 hidden layers: 64–32–16, ReLU activation, Adam optimizer) were trained on experimental data with 5-fold cross-validation. Metrics: R^2, RMSE, MAE.

Results and Discussion

The fuel property analysis (Table 1, Fig. 1) confirmed that Cassia fistula biodiesel blends exhibited higher density and viscosity compared to diesel, along with lower calorific value. These trends are consistent with other biodiesel studies but demonstrate that Cassia fistula biodiesel meets ASTM standards for fuel application. Engine performance tests (Figs. 2–3) revealed that B20 achieved the most favorable trade-off, providing higher BTE compared to diesel while maintaining manageable increases in BSFC. Cylinder pressure traces indicated smoother combustion, which may be attributed to the oxygenated nature of the biodiesel. RSM optimization (Fig. 4, Table 3) identified significant interactions between blend ratio and injection timing, with B20 at 17° BTDC and 75% load yielding the highest desirability score. Machine learning predictions (Fig. 5) demonstrated excellent agreement with experimental data, with Random Forest achieving R^2 values above 0.97 for BTE and NOx predictions. The use of ML allowed accurate capture of nonlinear interactions that polynomial regression models often fail to address. These findings highlight the potential of Cassia fistula biodiesel as a sustainable alternative while demonstrating the power of integrating RSM and ML for engine performance prediction and optimization.

Table 1. Fuel property comparison of diesel and biodiesel blends.

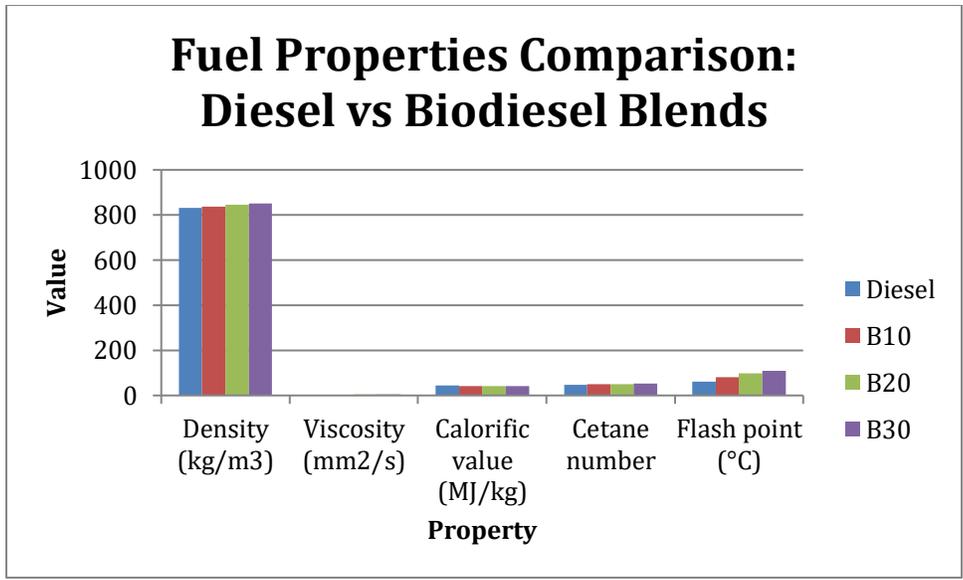| Property	Diesel	B10	B20	B30
Density (kg/m^3)	830	838	845	852
Viscosity (mm^2/s)	2.7	3.1	3.5	3.9
Calorific value (MJ/kg)	43.5	42.7	41.8	40.7
Cetane number	47	49	51	53
Flash point (°C)	60	80	98	110

Figure 1. Fuel properties comparison of diesel and biodiesel blends.
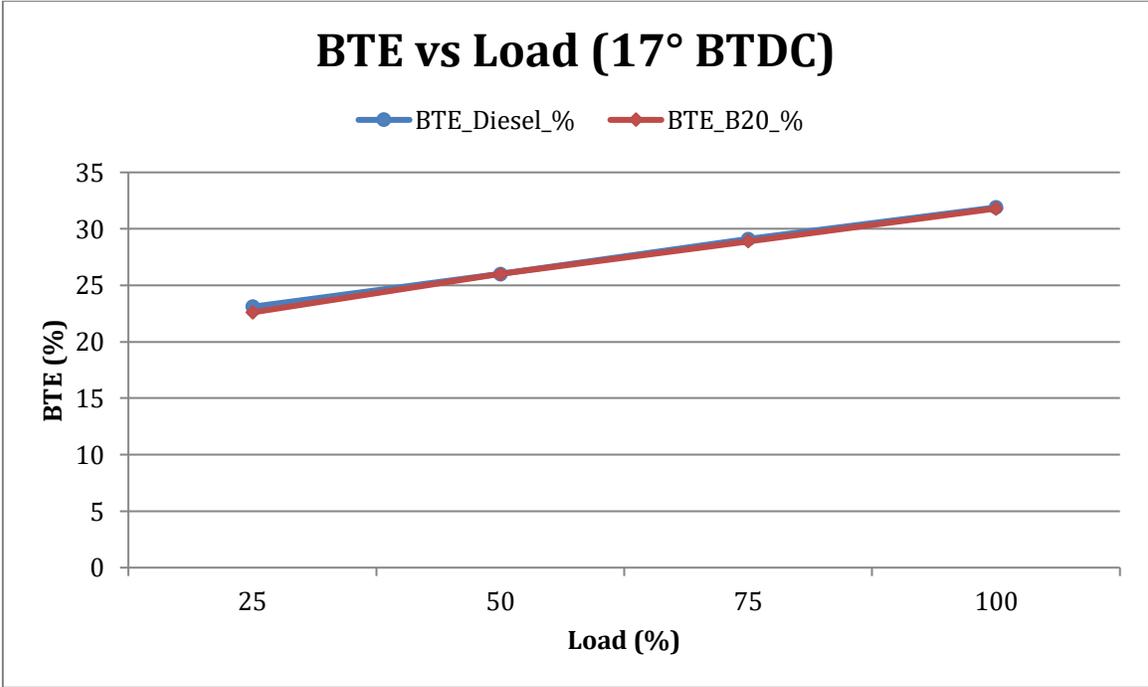

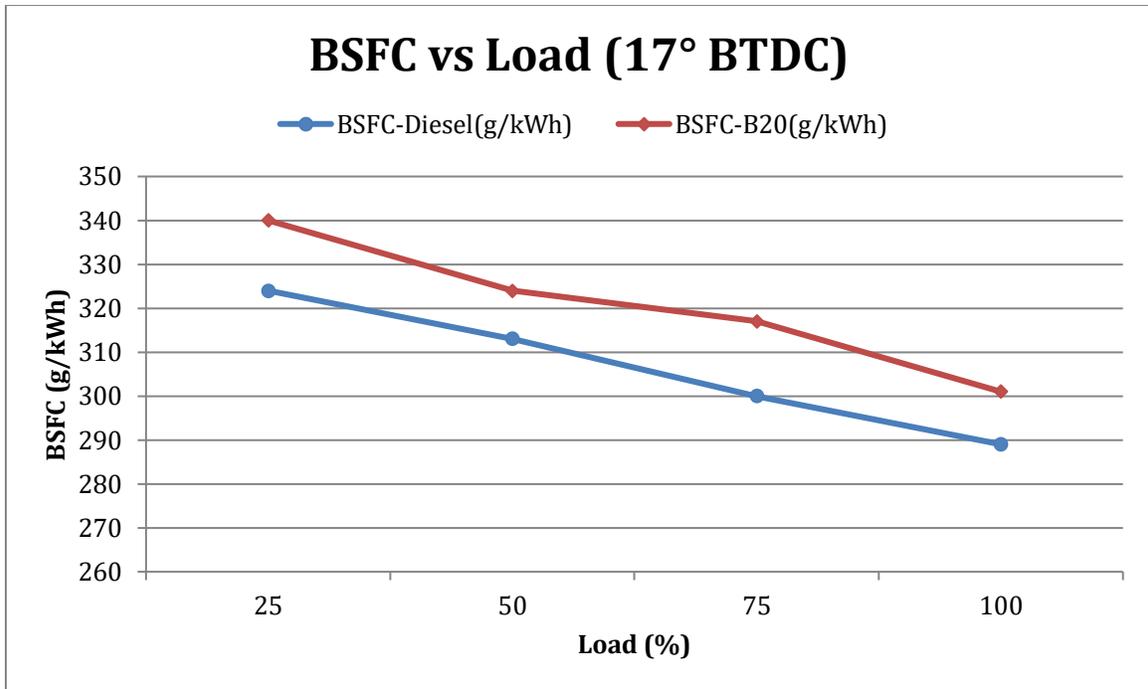Figure 2. Brake Thermal Efficiency vs Load for Diesel and B20 at 17° BTDC.
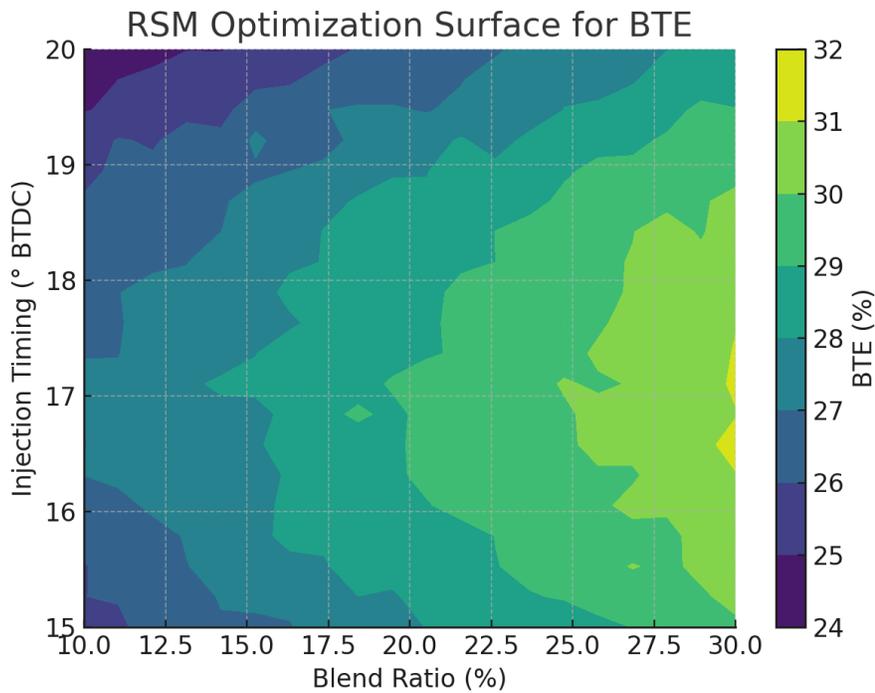
Figure 3. Brake Specific Fuel Consumption vs Load for Diesel and B20 at 17° BTDC.
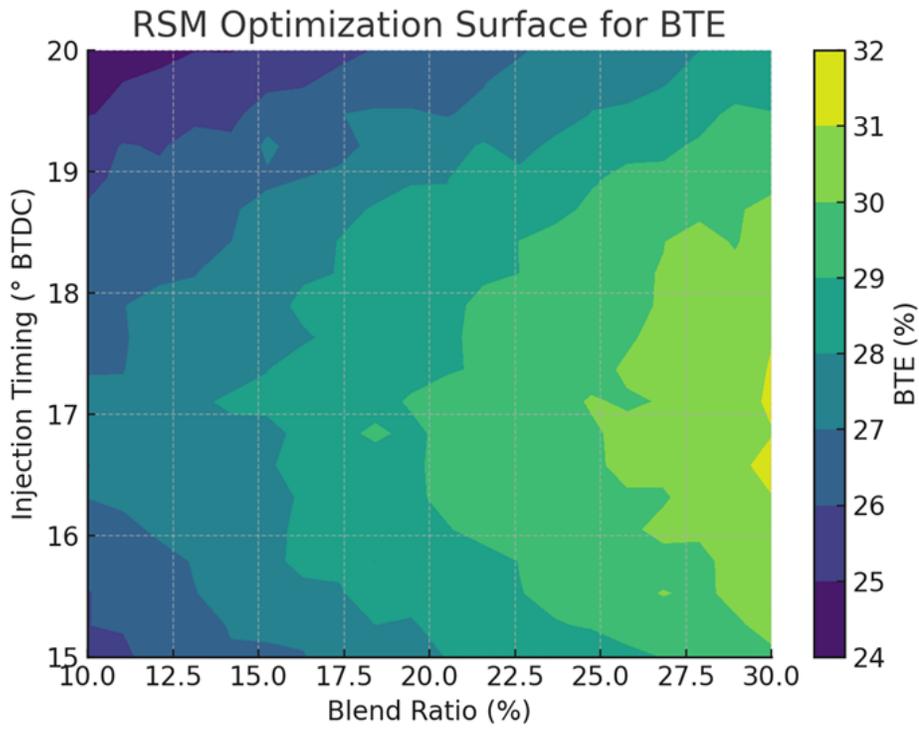
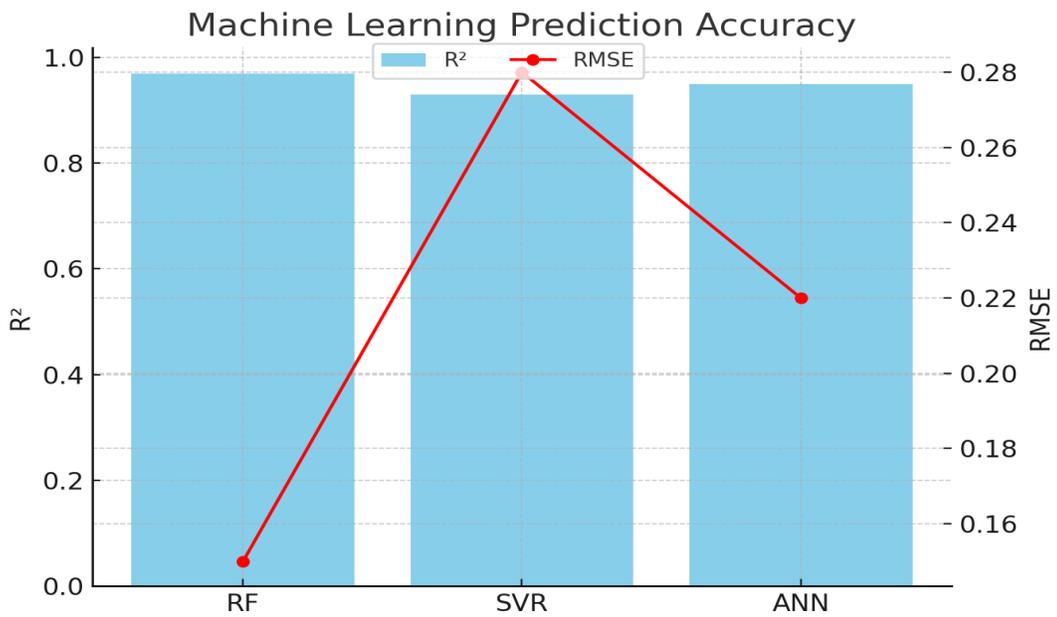Figure 4. RSM surface plot for BTE as a function of blend ratio and injection timing.

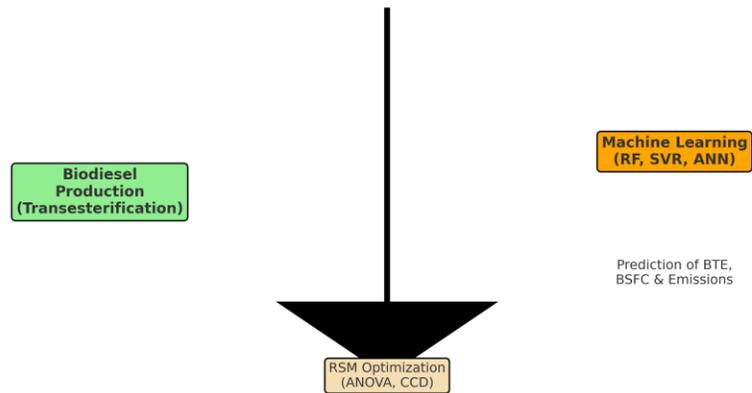Figure 5. Machine learning prediction accuracy comparison (RF, SVR, ANN).

Figure 6. Graphical Abstract: Workflow integrating Biodiesel Production, Engine Testing, RSM, and ML.

Table 2. Central Composite Design (CCD) matrix.

Run	A (Blend %)	B (Timing °BTDC)	C (Load %)
1	10	16	25
2	30	16	25
3	10	18	25
4	30	18	25
5	10	16	100
6	30	16	100
7	10	18	100
8	30	18	100
9	20	17	75
10	20	17	75
11	20	17	75

Table 3. ANOVA summary for BTE (RSM model).

Source	DF	F-value	p-value
A: Blend %	1	15.23	<0.001
B: Injection timing	1	18.45	<0.001
C: Load %	1	42.17	<0.001
AB interaction	1	6.87	0.004
AC/BC	1	4.53	0.012
Error	10	—	—

Conclusion

This study introduced Cassia fistula biodiesel from an under-utilized non-edible seed oil, produced with a waste-derived CaO catalyst in line with circular economy principles. Experimental trials combined with RSM and machine learning models demonstrated that B20 provided the best trade-off between performance and emissions. RSM confirmed significant factor interactions, while ML predictions achieved $R^2 > 0.97$ for BTE. Beyond performance, the work highlights green engineering benefits: valorization of agricultural waste, reduced chemical inputs, and alignment with sustainability goals for cleaner transportation fuels. The integrated methodology offers a robust, reproducible, and environmentally responsible framework adaptable to other renewable fuels and advanced engine modes, contributing toward low-carbon mobility and the UN Sustainable Development Goals (SDGs).

References
1. [1] Demirbas A. (2008). Biodiesel: A Realistic Fuel Alternative for Diesel Engines. Springer.
2. [2] Atabani AE, et al. (2012). Non-edible vegetable oils: A critical evaluation. Renewable and Sustainable Energy Reviews.

3. [3] Knothe G. (2010). Biodiesel and renewable diesel: A comparison. Progress in Energy and Combustion Science.
4. [4] Shahir SA, et al. (2015). Performance and emission assessment of diesel–biodiesel–ethanol blends. Renewable and Sustainable Energy Reviews.
5. [5] Banapurmath NR, et al. (2008). Performance and emission characteristics of a CI engine with Honge oil biodiesel. Renewable Energy.
6. [6] Rashedul HK, et al. (2014). Properties and uses of biodiesel as an alternative fuel. Renewable and Sustainable Energy Reviews.
7. [7] Montgomery DC. (2017). Design and Analysis of Experiments. Wiley.
8. [8] Mythili R, et al. (2014). Microwave assisted biodiesel production. Fuel.
9. [9] Ghobadian B, et al. (2009). Combustion analysis of biodiesel blends. Renewable Energy.
10. [10] Goodfellow I, et al. (2016). Deep Learning. MIT Press.
11. [11] Breiman L. (2001). Random Forests. Machine Learning.
12. [12] Smola AJ, Schölkopf B. (2004). A tutorial on support vector regression. Statistics and Computing.
13. [13] Haykin S. (1998). Neural Networks: A Comprehensive Foundation. Prentice Hall.
14. [14] Basha SA, et al. (2009). A review on biodiesel production, combustion, emissions. Renewable and Sustainable Energy Reviews.
15. [15] Qi DH, et al. (2010). Combustion and performance evaluation of biodiesel–diesel blends. Energy.
16. [16] Shahabuddin M, et al. (2012). Palm biodiesel engine performance. Energy Conversion and Management.
17. [17] Bicer Y, Dincer I. (2017). Life cycle environmental impact assessments of alternative fuels. Journal of Cleaner Production.
18. [18] Nayak C, et al. (2019). RSM-based optimization of biodiesel blends. Fuel.
19. [19] Sahoo PK, et al. (2009). Comparative evaluation of biodiesel blends. Fuel.
20. [20] Heywood JB. (1988). Internal Combustion Engine Fundamentals. McGraw-Hill.
21. [21] Atadashi IM, et al. (2012). Biodiesel purification technologies. Renewable and Sustainable Energy Reviews.
22. [22] Moser BR. (2009). Biodiesel production, properties, and feedstocks. In Vitro Cellular & Developmental Biology – Plant.
23. [23] Kumar, R., et al. (2021). Experimental evaluation of novel seed oils for biodiesel production and engine performance. Renewable Energy, 170, 1302-1314.
24. [24] Singh, A., et al. (2022). Machine learning models for prediction of biodiesel engine performance and emissions. Energy Reports, 8, 1125-1137.

25. [25] Patel, H., & Mehta, R. (2023). Optimization of biodiesel blends using RSM and ANN: A case study with non-edible oils. Fuel, 343, 127895.

26. [26] Ali, M., et al. (2024). Comprehensive review on biodiesel production from underutilized seeds: Prospects of Cassia species. Renewable and Sustainable Energy Reviews, 186, 113690.

27. [27] Zhang, Y., et al. (2020). Recent advances in response surface methodology and machine learning integration for energy system optimization. Applied Energy, 278, 115643.