Review Article

A Review on Machine Learning Application in Biodiesel Production Studies

Yuanzhi Xing [ID],[1] **Zile Zheng** [ID],[1] **Yike Sun** [ID],[1] **and Masoome Agha Alikhani** [ID][2]

[1]*School of Management, Shanghai University, Shanghai 201800, China*
[2]*Fouman Faculty of Engineering, College of Engineering, University of Tehran, Fouman, Iran*

Correspondence should be addressed to Zile Zheng; bryce_zl@shu.edu.cn and Masoome Agha Alikhani; m.aghaalikhani@ut.ac.ir

The consumption of fossil fuels has exponentially increased in recent decades, despite significant air pollution, environmental deterioration challenges, health problems, and limited resources. Biofuel can be used instead of fossil fuel due to environmental benefits and availability to produce various energy sorts like electricity, power, and heating or to sustain transportation fuels. Biodiesel production is an intricate process that requires identifying unknown nonlinear relationships between the system input and output data; therefore, accurate and swift modeling instruments like machine learning (ML) or artificial intelligence (AI) are necessary to design, handle, control, optimize, and monitor the system. Among the biodiesel production modeling methods, machine learning provides better predictions with the highest accuracy, inspired by the brain's autolearning and self-improving capability to solve the study's complicated questions; therefore, it is beneficial for modeling (trans) esterification processes, physicochemical properties, and monitoring biodiesel systems in real-time. Machine learning applications in the production phase include quality optimization and estimation, process conditions, and quantity. Emissions composition and temperature estimation and motor performance analysis investigate in the consumption phase. Fatty methyl acid ester stands as the output parameter, and the input parameters include oil and catalyst type, methanol-to-oil ratio, catalyst concentration, reaction time, domain, and frequency. This paper will present a review and discuss various ML technology advantages, disadvantages, and applications in biodiesel production, mainly focused on recently published articles from 2010 to 2021, to make decisions and optimize, model, control, monitor, and forecast biodiesel production.

1. Introduction

Fossil fuel, the most popular fuel with an essential role in developing economy and politics in both established and developing countries, has been a common industrial energy source for several decades because of its perfect properties combination like easy transportability, versatility, accessibility, and costly prices [1–3]. Although many undiscovered oil reserves remain in geological structures and rich unconventional oil reservoirs like tar sands, heavy oil, and oil shale indicate a suitable possibility of commercially viable resources, they are nonrenewable and limited. The world energy demand is assumed to reach a 56% growth between 2010 and 2040; hence there is a dire need for a sustainable alternative energy resource [4–6]. In addition to resources limitation, fossil fuel consumption for economic and industrial activities causes many challenges like air pollution, global warming, environmental deterioration, health problems, global climate change issues, and emitting greenhouse gas (GHG) in the entire world [7]. The energy crisis followed by high dependence on fossil fuels, increasing resource fluctuation, and environmental challenges exacerbated the resources ending-up concern and leading the world towards eco-friendly energy resources to assure a sustainable energy supply and meeting the escalating energy requirements from a renewable source [8–12]. Fossil fuel production will not suddenly stop and remains a universal energy resource, but scientists try to obtain low carbon

footprint energy [13]. Biofuels, hydrogen, compressed natural gas, liquefied petroleum gas, and alcohol have enough potential to become alternative energy sources [14–16].

A study was performed on renewables to choose the best alternative energy where bioenergy, today's largest renewable energy resource, presented great potential in addressing climate change and global energy issues [17]. Biofuel includes biodiesel, bioethanol, and biogas, obtained from biomass resources, which can be applied instead of fossil fuels due to integrating enhanced energy security, environmental benefits, availability, renewability, and sustainability to produce various energy sorts like electricity, power, and heating or to sustain transportation fuels [6, 16, 18–22]. Figure 1 illustrates the research trend in the biofuels field. The number of published documents has increased sharply from 2002 till 2020. Since 2016 a decrease in the growing number of articles was observed; however, it is still progressing.

Among all sustainable alternatives to fossil fuel, biodiesel is a suitable choice for diesel engines due to lower engine emissions (41% less greenhouse emission), physical and chemical properties advantages, and no need for significant modifications [23–26]. Biodiesel and petrodiesel are miscible in any ratio, which leads to the use of their combination rather than pure biodiesel, not only in developed countries such as The United States, France, Italy, and Germany, but also in developing countries such as Malaysia, Brazil, Indonesia, and Argentina [7, 27–29]. Biodiesel production capacity is an attractive growing trend; the automotive biofuels market is growing dramatically; it has engaged many scientists and researchers to satisfy the ever-rising energy supply demands by producing alternative fuels [25, 30]. As shown in Figure 2, the share of renewable energy in generating power is expected to have a 23% increase by 2030.

The challenge is to identify the biofuel production process outputs relationship as a function of process parameters, then maintaining and optimizing effective parameters in an optimum range to ensure high quality and productivity [13, 32]. Various transesterification associated raw materials parameters and reaction conditions like temperature, oil and catalyst type, reaction duration, oil to alcohol molar ratio, and catalyst concentration can affect productivity and production process response features, estimated through physical experiments [13, 33–37]. Despite the necessity of experiments, the prediction of factors effect is not successful due to the underlying nonlinear relations between the responses and parameters and also plenty of process parameters; therefore, high accurate experimental modeling methods like machine learning-based prediction and artificial intelligence (AI) techniques are beneficial to overcome experiment methods limitations and traditional computing techniques challenges [13, 38–40].

They provide mathematical models or independent modeling approaches according to the nature of the process to prevent waste of time and money and, furthermore, to study a wide range of physical and chemical process parameters separately and generate experimentally inaccessible details [10, 12, 41–44].
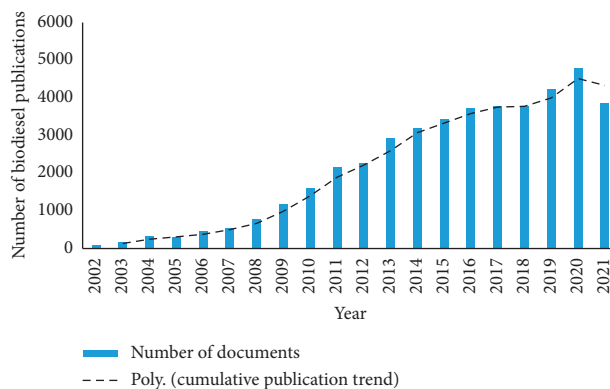

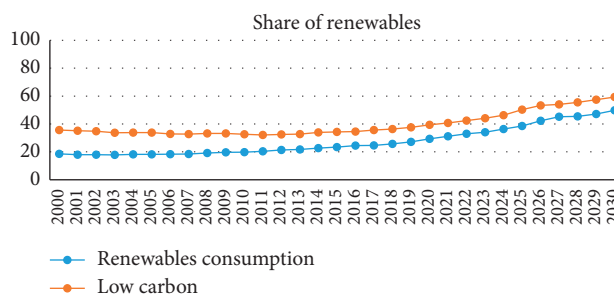Figure 1: The number of biodiesel research publications (data extracted from http://www.sciencedirect.com).

Figure 2: Share of renewables in power generation in the Sustainable Development Scenario, 2000–2030 [31].

2. An Introduction to AI and ML

AI is the ability of machines to simulate the human brain activities, applied through different computer science techniques, like heuristic algorithms, machine learning, and fuzzy logic [45–47]. It is chiefly employed to predict biomass and biofuel properties, bioenergy end-use systems performance, conversion process performance, supply chain modeling, and optimization. Recommended optimization methods are response surface methodology (RSM), genetic algorithm, and Taguchi method; in the meantime, artificial neural network (ANN), regression, and analytical methods are trending modeling methods in internal combustion engine research [48–51].

ML algorithms evolved with deep learning, reinforcement learning, transfer learning, and extreme learning are utilized in industrial processes to optimize, monitor, and control the systems, forecast maintenance, diagnose mistakes, and notify process attacks [2, 52–56]. Linear regression, Principal Component Analysis (PCA), Decision Trees (DT), Genetic Algorithms (GA), K-nearest Neighbor Classifier (KNN), Random Forests regression (RF), Artificial Neural Networks (ANN), and Support Vector Machines (SVM) are some powerful machine learning algorithms [57]. Machine learning refers to a programmed process using consecutive iterations based on inputs of external variants, gradually updating problem-solving capability and self-improvement to solve the study's complicated questions [57, 58].

AI applications to bioenergy systems are limited; however, studies indicate its great potential in addressing bioenergy development obstacles. Former reviews have separately focused on either a single AI approach or a part of bioenergy systems [2, 44, 48, 49, 59]. Due to the wide variety of AI techniques, conversion technologies, bioenergy products, biomass types, and supply chain design, a comprehensive review of AI applications throughout biomass agriculture to the consumption phase is necessary. This review intends to recommend advanced statistical methods and current popular machine learning algorithms conflux to obtain overall pragmatic models as an experiential agreement.

3. An Introduction to Biodiesel Production

Biodiesel is a clean, aromatic, biodegradable fatty acid methyl ester derived from waste oils, edible and nonedible vegetables oil, and animal fat (i.e., chicken and mutton tallow) as an alternative fuel source for diesel engines to reduce engine emissions, becoming a global mainstream for transportation [34, 45, 51, 60–62]. In addition to alternative transport fuel, biodiesel has other potential usages such as heating oil, plasticizers, power production, high boiling absorbents for cleaning gaseous industrial emissions, lubricants, and various solvent applications. Biodiesel has similar properties to diesel fuel, for instance, cetane number, viscosity, energy content, and phase variations. Biofuels can provide a new business for agricultural products and revitalizing rural areas [63].

3.1. Advantages

(i) Sulfur-free

(ii) Releasing fewer emissions

(iii) Profitable Physicochemical properties such as density, cetane number, flash point, viscosity, and lubrication

(iv) More complete combustion because it is highly oxygenated

(v) Promoting energy sufficiency [42].

3.2. Disadvantages

(i) Less energy content

(ii) Releasing more nitrogen oxides

(iii) Higher maintenance cost

(iv) High cost of establishment

(v) Separation and purification stage for product

(vi) Undesirable side reactions [51, 64]

Easy production from available renewable feedstock makes it more attractive. Nonedible tree seed oil resources are easily found everywhere, even in nonappropriate food crops land. Pure biodiesel, or a mixture of commercial diesel and biofuel, can be used in unmodified diesel engines due to the environmental sustainability advantages [51, 65]. Several

countries command to add biodiesel into all diesel fuels to encourage people to use biodiesel [63, 66].

The most common reaction in the biodiesel production process is transesterification, which uses heterogeneous or homogeneous acid and base catalysts to improve transesterification under mild reaction conditions. Sodium hydroxide and potassium hydroxide (NaOH, KOH) are regular alkaline catalysts that can provide higher biodiesel yield [67–70]. The transesterification reactions among the oil (i.e., canola oil, *Simarouba glauca* oil, soybean oil, sunflower seed oil, *Thevetia peruviana* seed oil, palm oil, etc.) and alcohols (i.e., methanol, ethanol) produce biodiesel [62, 71–76]. It is a costly energy-consuming production process which results from product purification and separation, requiring a pretreatment step to reduce water and free fatty acids over a long period [2]. Low esterification efficiency arose from undesired side reactions. Figure 3 illustrates the transesterification reaction for biodiesel production and input and output variables.

Various transesterification associated parameters and reaction conditions like temperature, oil and catalyst type, reaction duration, oil to alcohol molar ratio, and catalyst concentration affect productivity, and production process response features significantly affect transesterification reaction [37, 78]. Statistical tools and many physical experiments are necessary to predict reaction responses and interactions to each parameter due to optimizing transesterification [36, 42].

4. ML Methods Application in Biodiesel Life Cycle

Producing biodiesel from renewables includes the following steps: extracting oil, pretreating feedstock, transesterification reaction, separating products, recovering unreacted alcohol, neutralizing glycerin, washing, and purification of biodiesel [70, 79]. In this section, we attempted to categorize and review ML technology applications in 5 crucial steps of biodiesel production, including soil, feedstock, production, consumption, and emissions [57, 80].

ML technology can be beneficial in all five stages to enhance the quality of estimations. There are plenty of research reviews on applications of machine learning technology in modeling biodiesel-fueled engines and combustion approaches; therefore, this study mainly focuses on the first three stages. Figure 4 shows an overview of the biodiesel production trend, inspired by Aghbashlo et al. [79] and Ahmad et al. [57].

4.1. ML Applications in Soil Stage. Numerous studies on the plot and tree cases applying ML have been reported in the soil stage of the biofuels' life cycle. The most common ML methods in the soil stage are Random Forest (RF), Gaussian Process Model (GPM), and Support Vector Machines (SVM).

Sorghum crop is beneficial for producing health-promoting food from seeds, fodder, and biofuels from aboveground biomass [81]. To predict future trends in sorghum bicolor yield,
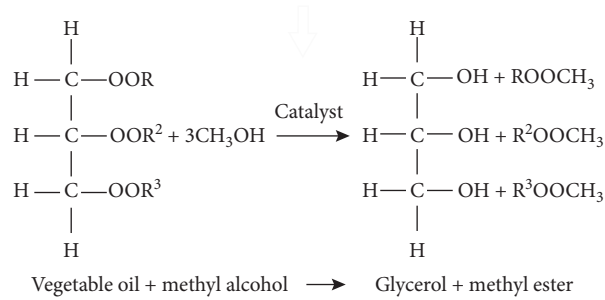
Figure 3: Biodiesel production procedure, transesterification reaction outline [77].
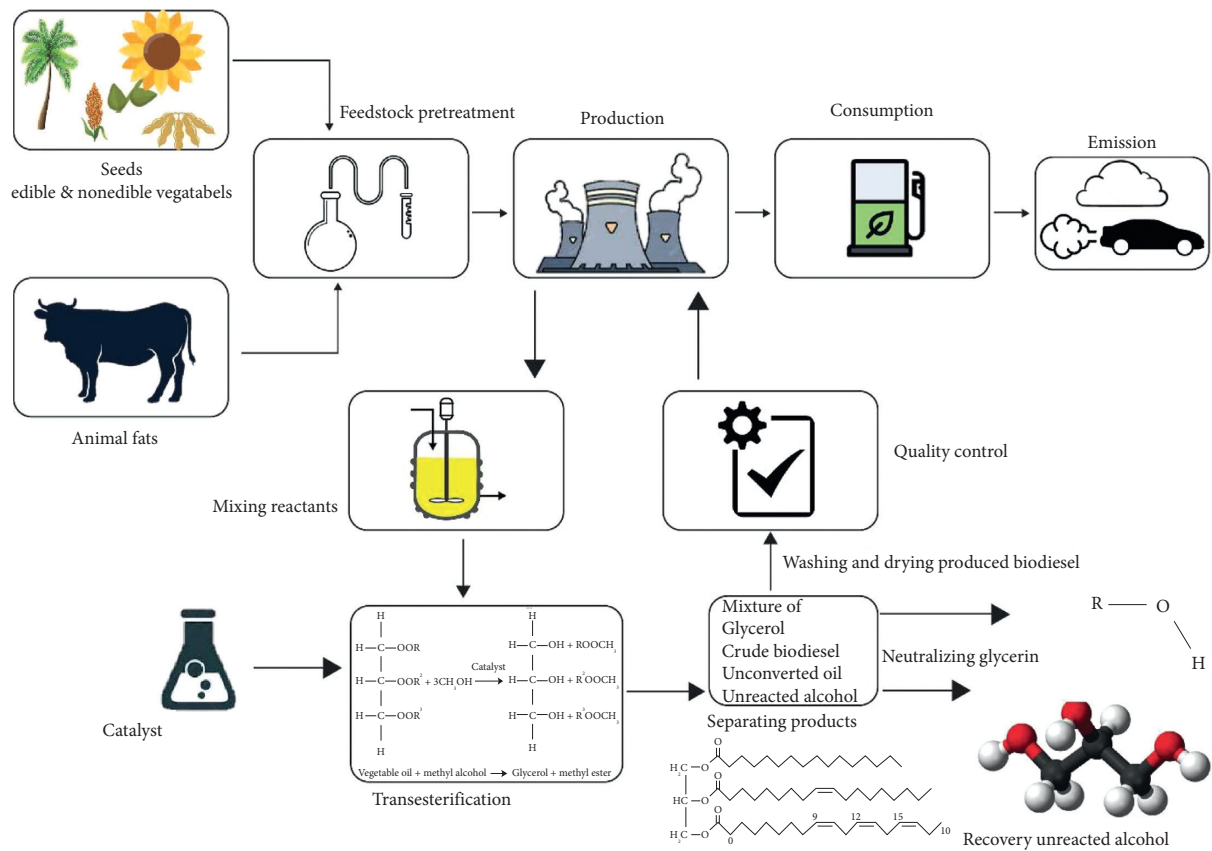

Figure 4: Critical stages of biodiesel life cycle (inspired from [57, 79]).

Huntington et al. [82] used the RF approach under four greenhouse gas (GHG) emission scenarios and two different watering regimes. The most valuable sorghum productivity predictors were vapor pressure deficit, time, and irrigation practices. The RF model obtained a rational prediction accuracy by uniquely training and classifying data samples by year and country. Habyarimana et al. [81] performed a study based on sorghum fields satellite imaging to predict sorghum biomass yield using various ML methods like radial basis kernel (SVM-R), nonlinear kernel (SVM-G), PCA discriminant analysis (PCA-DA), PLS discriminant analysis (PLS-DA), SVM with linear classifier (SVM), radial basis kernel with polynomial basis kernel (SVM-P), simple linear model, RF, ANN, eXtreme Gradient Boosting-XgbLinear method (GBL), eXtreme

Gradient Boosting-xgbDART method (GBD), and eXtreme Gradient Boosting-xgbtree method (GBT), where the eXtreme Gradient Boosting-xgbtree method performed better results.

Gleason et al. [83] compared the Linear Mixed-effects Regression (LME), Cubist, Support Vector Regression (SVR), and Random Forest (RF) methods to predict biomass in a moderately dense forest with 40 to 60% canopy closure where SVR performed the most accurate biomass model. Lee et al. [84] conducted a four-scenario context emissions-based study using Boosted Regression Tree (BRT) model to estimate corn production environmental impacts from 2022 to 2100, where the BRT model achieved a 0.82 estimating eutrophication impacts correlation coefficient and 0.78 in global warming. Yang et al. [85] applied Gaussian Process

Model (GPM), a Bayesian inference method in a two-stage machine learning method, to achieve more accurate estimations. First of all, GPM crops yield downscaling and then an RF model estimated yield. Soil characteristics, solar radiation, average precipitation, wind speed, and temperature are usually input parameters, and the output parameters are future life cycle environmental impact and biomass yield. Table 1 provides a summary of soil phase studies in order to understand the effective method in each study. To form this table, various research papers have been used to extract data [81–85].

4.2. ML Applications in Feedstock. According to machine learning applications in the feedstock phase studies, ANN, multiple linear regression, statistical regression, and multiple nonlinear regression models are the most popular methods. Blend composition, temperature, mixing speed, and mixing time are typical input variables, and the output variables are viscosity, flash point, oxidation stability, density, methane fraction, higher heating values, and cetane number. Mairizal et al. [86] examined biodiesels generated from various resources such as walnut oil, sunflower oil, peanut oil, rapeseed oil, hydrogenated coconut oil, hydrogenated copra oil, and beef tallow to predict higher heating value, viscosity, flashpoint, biodiesel's oxidative stability, and density by using multiple linear regressions. Results showed that prediction performance increases by adding PU/MU (mono- and polyunsaturated fatty acids balance) as an independent parameter. Model inputs were feedstock polyunsaturated fatty acids content, iodine value, and saponification value. In another study on various biodiesels generated from fatty acid, the ANN method was applied to estimate cetane number, density, kinematic viscosity, and flashpoint [87]. Average absolute deviation and model's estimation accuracy are showed in following values, respectively: cetane number (1.637%; 96.6%), flash point (0.997%; 99.07%), kinematic viscosity (1.638%; 95.80%), and density (0.101%; 99.40%). Tchameni et al. [88] used multiple ANN and nonlinear regression (MNLR) to forecast waste vegetable oil rheological properties. Results presented ANN model superiority over the MNLR method performance. Using single linear regressions and multiple linear regressions to estimate methane yield in biomass structural components revealed a quite considerable correlation between methane biomass' potentials and chemical composition. Table 2 provides a summary of feedstock phase studies [86–88] to classify the efficient method and study purposes.

4.3. ML Applications in Production. In the production stage, choosing a proper ML method depends on produced biofuel type (i.e., biodiesel, biogas, and biohydrogen). Based on studies, machine learning applications in biodiesel study can be organized into four sections: both quality and yield optimization, estimating quality, estimating yield, estimating, and optimizing process conditions and efficiency [57].

4.3.1. Quality Prediction. The prevailing ML method for quality prediction is ANN developed by the regression model, using reaction temperature, reaction time,

calcination temperature, pressure, and flow rate as input variables and FAME (fatty acid methyl ester) content, viscosity, composition, quantity, cetane number, and density stand as output variables.

Soltani et al. [89] used an artificial neural network (ANN) to model various reaction parameter effects, i.e., calcination temperature, metal ratio, reaction time, and reaction temperature in a palm fatty acid (PFAD) to esters distillation, using sulfonated mesoporous zinc oxide SO_3HZnO catalyst. Assessed optimum conditions for predicting a 56.41 nm $SO_3H–ZnO$ nanocrystalline catalyst size were 160°C reaction temperature, 700 calcine temperature, and 0.004 mole of Zn concentration during 18 min reaction time. Zinc concentration and the reaction time are recognized as the most and least effective parameters, respectively.

Ahmad et al. [90] used an ensemble learning method like Least Squares Boosting (LSBoost) integrated with the polynomial chaos expansion method (PCE) to predict quantity, quality, flow rate, the cetane number of fatty acid methyl esters (FAME), and composition in the vegetable oil-based biodiesel production process. Predicted values showed 1% uncertainty in all process parameters using mean absolute deviation percent (MADP), showing high accuracy of the proposed model in outcomes prediction and quantification uncertainty effect in the process. During the biodiesel production process from vegetable oil, the PCA method was applied to estimate relative density, viscosity, and percentage of vegetable oil conversion to methyl esters. Using PCA is an effective technique to differentiate and discriminate between pure biodiesel, pure diesel, waste oil, and their mixture.

Sarve et al. [91] used artificial neural network (ANN) and response surface methodology (RSM) based on a central composite design (CCD) to predict fatty acid methyl ester (FAME) content in biodiesel production from sesame oil, using barium hydroxide as a basic catalyst. The best possible combination of optimum condition values is methanol-to-oil molar ratio (6.69 : 1), reaction time (40.30 min), catalyst concentration (1.79 wt.%), and (31.92°C) temperature, which resulted in 98.6% of FAME content. The study revealed that catalyst concentration has the main influence on the FAME contents in the final product. ANN has a better capability in predicting the FAME content due to better correlation coefficient, root mean square error (R^2), standard error of prediction (SEP), and relative percent deviation (RPD) values compared to RSM.

4.3.2. Yield Estimation. Several studies concentrated on ML methods application in predicting biodiesel synthesis from nonedible oils like anaerobic sludge, castor oil, and jatropha-algae.

Kumar et al. [92] trained an ANN model with Levenberg–Marquardt (LM) algorithm and backpropagation learning algorithm to predict biodiesel yield in the transesterification process, using jatropha-algae oil blends as inputs. The R-square value of 0.9976 compared with the experimental results confirmed the competency of the ANN technique.

TABLE 1: Various ML applications in the soil phase of biodiesel production outline.

Reference	Applied models	Field	Results
[81]	GBL, GBD, GBT, ANN, RF, SVR, SVM, SVM-P, SVM-R, SVM-G, PCA-DA, PLS-DA	Predict sorghum crop yield	GBT
[82]	RF	Predict sorghum crop yield	RF
[83]	LME, SVR, RF	Predict biomass yield in forest	SVR
[84]	BRT	Estimate corn production environmental impacts	BRT
[85]	GPM, RF	Land productivity	GPM

TABLE 2: Various ML applications in the feedstock phase of biodiesel production outline.

Reference	Applied models	Aim
[86]	Multiple linear regressions	To predict HHV, viscosity, FP, oxidative stability, density
[87]	ANN	To estimate C, density, kinematic viscosity, FP
[88]	ANN, MNLR, single and multiple linear regressions	To estimate oil rheological properties

Banerjee et al. [93] used the ANN and CCD model in castor oil and methanol transesterification using H_2SO_4 acid catalyst to predict the % fatty acid methyl ester content. They also devised a kinetic model using the experimental and computed data. Also using ANN-based predicted data and the experimental outputs, the rate constants of a kinetic model have been estimated. The temperature, catalyst concentration, and methanol-to-oil molar ratio are input parameters. The ANN model predicted a % fatty acid methyl ester yield with an 8% deviation.

Kanat et al. [94] used the ANN method and multilayer neural networks topology to model and estimate the anaerobe thermophilic upflow sludge blanket digester biodiesel and biogas production rate. Trained and tested experimental data were evaluated in both steady conditions and abnormal conditions; a high correlation coefficient showed ANN optimistic results for online monitoring of the thermophilic reactors. In a jatropha-algae oil blend study, ANN performed better than RSM [95].

A biodiesel synthesis process from waste goat tallow containing remarkable free fatty acids (FFAs) has been modeled by RSM and ANN to identify optimum parametric values that resulted in maximum FA conversion. Under optimal conditions, response surface methodology (RSM) and ANN presented similar predictability performance [96].

In another study, a linear regression (LR) and ANN model based on a Levenberg–Marquardt learning algorithm were developed for predicting soybean oil-based biodiesel transesterification yield, where the ANN performed better than LR [97]. Various conditions of soybean oil to biodiesel transesterification process have been studied to predict biodiesel yield [39]. In this study, the artificial neural network is applied with a multilayer feedforward neural network and kinetic models. The results showed the ANN model superiority, accuracy, and clarity over the kinetic modeling method. Guo et al. [98] used an adaptive neurofuzzy interference system (ANFIS) method, based on a statistical learning theory to estimate the biodiesel production yield as a function of methanol/oil ratio, pressure, reaction time, and temperature in the noncatalytic supercritical methanol (SCM) method. The high value of R-squared results indicates the ANFIS model's impact on biodiesel yield prediction. Mostafa et al. [35] compared adaptive neurofuzzy inference system (ANFIS) and response surface methodology (RSM) to predict and simulate the efficiency of these approaches in modeling the transesterification yield. Box-Behnken design of RSM and two ANFIS approaches (hybrid and backpropagation optimization methods) investigated independent variable's impact on the conversion of fatty acid methyl esters (FAME). The considerable R^2 value was 0.9669 for RSM compared with 0.9812 and 0.9808 for two ANFIS models indicating the ANFIS models superiority against the RSM model for modeling and optimizing. Maran et al. [49] compared artificial neural network (ANN) and response surface methodology (RSM) efficiencies to predict and simulate muskmelon oil-based biodiesel yield. Central composite rotatable design CCRD investigated the ANN model against the RSM model. Catalyst concentration, reaction time, reaction temperature, and methanol-to-oil molar ratio affect FAME conversion by Multilayer Perceptron (MLP) neural network and RSM. The R^2 value for RSM was 0.869, and it was 0.991 for ANN models, showing the ANN model superiority against the RSM to model and optimize FAME production.

4.3.3. Quality and Yield Estimation. Numerous studies have focused on biodiesel quality and yield optimization. Bobadilla et al. [77] used a set of Support Vector Machines (based on radial basic function kernel, linear kernel, and polynomial kernel) and linear regression methods to predict and improve biodiesel yield of particular properties like turbidity, higher heating value (HHV) with decreased viscosity, and density. Appling genetic algorithms to the regression models obtained more accurate biodiesel optimization

scenarios to identify the best combination of independent and dependent variables.

Cheng et al. [99] developed a GA-ESIM method which is the combination of Evolutionary Support Vector Machine Inference Model (ESIM) and K-means Chaotic Genetic Algorithm (KCGA) to predict precisely and optimize biodiesel mixture properties. They found GA-ESVM better than ANN-GA and SVM. Obtained results demonstrate that the GA-ESIM model performance in prediction is more accurate than other AI-based tools.

Sivamani et al. [100] used ANN-GA-based and RSM models to predict and optimize the biodiesel yield in *Simarouba glauca* transesterification. They used a gas chromatography-mass spectroscopic (GC-MS) analysis oil to observe free fatty acid (FFA) level, and alcohol ratio, reaction time, and reaction temperature were input variables.

Ighose et al. [101] focused on an RSM optimization tool alongside the ANFIS model to predict and optimize the biodiesel yield in the *Thevetia peruviana* seed oil transesterification process. In addition to ANFIS and RSM model, using GA resulted in higher *Thevetia peruviana* methyl esters yield (TPME) in less time. The results determined the priority of ANFIS prediction capability over the RSM model. Dhingra et al. [102] applied ANN and GA combination in polanga oil-based biodiesel production to predict and optimize reaction variables to maximize the transesterification process. The input variables are the ethanol-to-oil molar ratio, the reaction temperature, the catalyst concentration, the reaction time, and the stirring speed. Outputs were combined with GA to optimize reaction conditions resulting in 92% by weight biodiesel yield.

4.3.4. Estimation and Optimization of Process Conditions and Efficiency. Karimi et al. [103] implemented a multiobjective analysis, using RSM and ANN to estimate FAME content and exergetic efficiency in waste cooking oil transesterification (WCO) for biodiesel production. Water concentration, reaction time, immobile lipase, and methanol concentration have been optimized to achieve 95.7% predicted FAME content. Corresponded input variables are the 35% catalyst concentration, 12% water content, methanol-to-WCO molar ratio of 6.7, in 20 hours, produced 86% FAME content, and 80.1% exergy efficiency.

Patle et al. [104] used nondominated sorting GA-II (NSGA-II) multiobjective optimization to simulate and compare palm waste cooking oil esterification and transesterification reactions and optimizing heat duty, profit, and organic waste. As the heat duty increased, the profit improved, which increases the amount of organic waste. Rouchi et al. [105] used a Multivariate Curve Resolution Alternative Least Square (MCR-ALS) to process analysis and control the reaction parameters into the desired path. Multiple Scatter Correction preprocessing technique and MCR-ALS evaluate concentrations, the component's type, and spectra to obtain biodiesel production from the soybean process. The correlation coefficient and standard deviation of residuals demonstrated the suitability of the MCR-ALS

method. Shukri et al. [106] used ANN to optimize the engine performance, using a mixture of palm oil methyl ester and diesel as fuel in a diesel engine. Both experimental results and the ANN model showed better engine performance for the biodiesel 10 percent blend (B10) diesel fuel and palm oil blends due to the higher heating value and cetane number.

Aghbashlo et al. [107] developed an ANFIS model integrated with linear interdependent fuzzy multiobjective (ALIFMO) approaches and nondominated sorting genetic algorithm (NSGA-II) to optimize operating conditions as a function of inputs. Input parameters were reaction temperature, methanol/oil molar ratio, and residence time. Optimization minimized normalized exergy destruction (NED) and maximized functional exergy efficiency (FEE) and universal exergy efficiency (UEE) output parameters towards achieving the best conversion efficiency (CE), which is more than 96.5% of biodiesel content. Applied ANFIS models perfectly estimated the FEE, UEE, NED, CE parameters with an $R^2 \approx 1.0$.

Sarve et al. [108] compared ANN and RSM in biodiesel production optimization concerning their analysis sensitivity, predictivity and generalization capability, and parametric effects. 97.42% of fatty acid ethyl ester (FAEE) content have been obtained at optimized temperature, ethanol-to-oil molar ratio, initial CO_2 pressure, reaction time, and temperature, where the temperature was the most effective. ANN model performed better results than the RSM in mahua oil FAEE content predictions and data fitting.

In a biodiesel production process from vegetable oil, Nicola et al. [80] employed a multiobjective GA optimization to maximize important compounds' purification and minimize energy requirements by optimizing main parameters in the process. Input parameters to the process model are reflux ratio, the mass flow rate of water, the water temperature, flash temperature, the number of trays, and dryer temperature. Among all optimized configurations, the one which confirms the minimum specific energy consumption and meets the biodiesel quality required standards was detected. Noriega et al. [109] used group interaction parameters (GIP) to predict and validate all present two-phase equilibriums between liquids in the biodiesel production system, including glycerol, low molecular weight alcohols, water, fatty acids, and biodiesel. Results demonstrated that the amount of carbon, hydroxyl groups, and unsaturated bonds affect liquid-liquid equilibrium, and the most efficient parameter was distributed component overall mass fraction, afterward length of the alcohol chain.

López-Zapata et al. [110] used an Extended Kalman Filter (EKF) and virtual sensors to measure and estimate operating conditions variables, control performance, and monitor the reaction. Performance analysis used alcohol, triglycerides (TG), methyl ester, diglycerides (DG), glycerol (GL), and monoglycerides (MG) concentrations to evaluate jatropha oil-based biodiesel due to a minor number of measurable variables, like PH and temperature. Fahmi and Cremaschi [111] developed an ANN superstructure model to recognize the optimum biodiesel production plant and best operation conditions. The ANN model was an effective alternative for thermodynamics, unit operation, and mixing

models, presenting a less complicated model for the synthesis process. As mentioned before, Soltani et al. [89] used ANN to model various reaction parameter effects, using SO_3HZnO catalyst. Assessed optimum conditions were 160°C reaction temperature, 700 calcine temperature, and 0.004 moles of Zn concentration through 18-minute reaction time. Zinc concentration and the reaction time were the most and the least effective parameter, respectively.

5. Conclusions

According to the machine learning applications in this study, the most common ML methods in the soil stage are Random Forest, Gaussian Process Model, and Support Vector Machines. In the feedstock phase studies, ANN, multiple linear regression, statistical regression, and multiple nonlinear regression models are the most popular methods. Blend composition, temperature, mixing speed, and mixing time are typical input variables, and the output variables are viscosity, flash point, oxidation stability, density, methane fraction, higher heating values, and cetane number. The prevailing ML method for quality prediction is ANN developed by the regression model, using reaction temperature, reaction time, calcination temperature, pressure, and flow rate as input variables, and FAME content, viscosity, composition, quantity, cetane number, and density stand as output variables. The prevailing ML method for yield estimation is ANN accompanied by ANFIS, using methanol-to-oil molar ratio, reaction time, catalyst concentration, total volatile fatty acid of the effluent, and temperature, while % FAME yield, biogas production rate estimation, biodiesel yield, and biodiesel production are regular output variables. The prevailing ML method in optimizing yield and quality section is ANN accompanied by GA-based ANFIS and SVM. The top five main frequently used input variables are methanol-to-oil molar ratio, stirring speed, catalyst concentration, reaction time, and reaction temperature. The most common output variables are FAME yield, biodiesel yield, high heating value density, and oil's final acid value. The dominant ML method in the process efficiency and optimization portion is ANN accompanied by ANFIS. Frequently used input variables are reaction time, concentration, water content, methanol-to-oil molar division, and temperature, while CE, universal exergy efficiency (UEE), FAME content, biodiesel yield, and functional exergy efficiency are output variables. ANN, ANFIS, ELM, and SVM Machine Learning methods were employed to study consumption, engine performance, and emission.

Nomenclature

ALIFMO: Artificial linear interdependent fuzzy multiobjective optimization
AI: Artificial intelligence
ANFIS: Adaptive neurofuzzy interference system
ANN: Artificial neural networks
ALS: Alternative least square
B10: Biodiesel 10 percent blend

BRT: Boosted regression tree
CCD: Central composite design
CE: Conversion efficiency
CN: Cetane number
DA: Discriminant analysis
ELM: Extreme learning machine
FAME: Fatty acid methyl ester
FAs: Fatty acids
FEE: Functional exergy efficiency
FP: Flash point
GA: Genetic algorithm
GBD: eXtreme Gradient Boosting-xgbDART
GBL: eXtreme Gradient Boosting-xgbLinear
GBP: eXtreme Gradient Boosting-xgbtree
GBT: Gene expression programming
GHC: Greenhouse gas
GIP: Group interaction parameters
GPM: Gaussian process model
HC: Hydrocarbon
IAV: Initial acid value of vegetable oil
K-ELM: Kernel-based extreme learning machine
KV: Kinematic viscosity
LLE: Liquid-liquid equilibrium
LME: Linear mixed-effects
LR: Linear regression
LS: Least square
MAPE: Mean absolute percentage error
MCR: Multivariate curve resolution
ML: Machine learning
MNLR: Multiple nonlinear regression
MO: Mustard oil
MSE: Mean squared error
PU/MU: Mono- and polyunsaturated fatty acids balance
NED: Normalized exergy destruction
PAT: Process analytical technologies
PCA: Principal component analysis
PLS: Partial least square
RB-FNN: Radial basis function neural network
RF: Random forest
RFM: Random forest model
RLS: Recursive least squares
RSM: Response surface methodology
SVM: Support Vector Machines
SVR: Support vector regression
UEE: Universal exergy efficiency
UHC: Unburned hydrocarbons
VCR: Variable compression ratio.

Data Availability

The data used to support the findings of this study are provided within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] K. I. Wong and P. K. Wong, "Optimal calibration of variable biofuel blend dual-injection engines using sparse Bayesian extreme learning machine and metaheuristic optimization," *Energy Conversion and Management*, vol. 148, pp. 1170–1178, 2017.

[2] K. I. Wong, P. K. Wong, C. S. Cheung, and C. M. Vong, "Modeling and optimization of biodiesel engine performance using advanced machine learning methods," *Energy*, vol. 55, pp. 519–528, 2013.

[3] F. Zhang, J. Wang, S. Liu, S. Zhang, and J. W. Sutherland, "Integrating GIS with optimization method for a biofuel feedstock supply chain," *Biomass and Bioenergy*, vol. 98, pp. 194–205, 2017.

[4] N. A. Owen, O. R. Inderwildi, and D. A. King, "The status of conventional world oil reserves-hype or cause for concern?" *Energy Policy*, vol. 38, no. 8, pp. 4743–4749, 2010.

[5] A. Sieminski, "International energy outlook 2013," Report Number: DOE/EIA-0484, 2013, US Energy Information Administration (EIA), Washington, DC, USA, 2013.

[6] S. Ardabili, A. Mosavi, and A. R. Várkonyi-Kóczy, "Systematic review of deep learning and machine learning models in biofuels research," in *Proceedings of the International Conference on Global Research and Education*, Springer, Balatonfüred, Hungary, September 2019.

[7] A. K. Agarwal, "Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines," *Progress in Energy and Combustion Science*, vol. 33, no. 3, pp. 233–271, 2007.

[8] P. Zelenka, W. Cartellieri, and P. Herzog, "Worldwide diesel emission standards, current experiences and future needs," *Applied Catalysis B: Environmental*, vol. 10, no. 1, pp. 3–28, 1996.

[9] M. Ben Jebli and S. Ben Youssef, "The role of renewable energy and agriculture in reducing CO 2 emissions: evidence for North Africa countries," *Ecological Indicators*, vol. 74, pp. 295–301, 2017.

[10] R. Concu, M. N. D. S. Cordeiro, C. R. Munteanu, and H. González-Díaz, "PTML model of enzyme subclasses for mining the proteome of biofuel producing microorganisms," *Journal of Proteome Research*, vol. 18, no. 7, pp. 2735–2746, 2019.

[11] A. L. De Bortoli and F. N. Pereira, "Obtaining a reduced kinetic mechanism for methyl butanoate," *Journal of Mathematical Chemistry*, vol. 57, no. 3, pp. 812–833, 2019.

[12] E. A. Del Rio-Chanona, J. L. Wagner, H. Ali et al., "Deep learning-based surrogate modeling and optimization for microalgal biofuel production and photobioreactor design," *AIChE Journal*, vol. 65, no. 3, pp. 915–923, 2019.

[13] K. K. Gupta, K. Kalita, R. K. Ghadai, M. Ramachandran, and X.-Z. Gao, "Machine learning-based predictive modelling of biodiesel production-a comparative perspective," *Energies*, vol. 14, no. 4, p. 1122, 2021.

[14] I. M. Yusri, R. Mamat, G. Najafi et al., "Alcohol based automotive fuels from first four alcohol family in compression and spark ignition engine: a review on engine performance and exhaust emissions," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 169–181, 2017.

[15] L. Das, R. Gulati, and P. K. Gupta, "A comparative evaluation of the performance characteristics of a spark ignition engine using hydrogen and compressed natural gas as alternative fuels," *International Journal of Hydrogen Energy*, vol. 25, no. 8, pp. 783–793, 2000.

[16] A. Demirbas, "Progress and recent trends in biofuels," *Progress in Energy and Combustion Science*, vol. 33, no. 1, pp. 1–18, 2007.

[17] Z. Abdmouleh, R. A. M. Alammari, and A. Gastli, "Review of policies encouraging renewable energy integration & best practices," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 249–262, 2015.

[18] S. Khishtandar, M. Zandieh, and B. Dorri, "A multi criteria decision making framework for sustainability assessment of bioenergy production technologies with hesitant fuzzy linguistic term sets: the case of Iran," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 1130–1145, 2017.

[19] N. E. Mougari, J. F. Largeau, N. Himrane, M. Hachemi, and M. Tazerout, "Application of artificial neural network and kinetic modeling for the prediction of biogas and methane production in anaerobic digestion of several organic wastes," *International Journal of Green Energy*, pp. 1–14, 2021.

[20] A. Mosavi, M. Salimi, S. Faizollahzadeh Ardabili, T. Rabczuk, S. Shamshirband, and A. Varkonyi-Koczy, "State of the art of machine learning models in energy systems, a systematic review," *Energies*, vol. 12, no. 7, p. 1301, 2019.

[21] H. M. Mahmudul, F. Y. Hagos, R. Mamat, A. A. Adam, W. F. W. Ishak, and R. Alenezi, "Production, characterization and performance of biodiesel as an alternative fuel in diesel engines-a review," *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 497–509, 2017.

[22] M. M. Hasan and M. M. Rahman, "Performance and emission characteristics of biodiesel-diesel blend and environmental and economic impacts of biodiesel production: a review," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 938–948, 2017.

[23] E. Jiaqiang, M. Pham, D. Zhao et al., "Effect of different technologies on combustion and emissions of the diesel engine fueled with biodiesel: a review," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 620–647, 2017.

[24] G. Knothe and L. F. Razon, "Biodiesel fuels," *Progress in Energy and Combustion Science*, vol. 58, pp. 36–59, 2017.

[25] S. Dharma, H. C. Ong, H. H. Masjuki, A. H. Sebayang, and A. S. Silitonga, "An overview of engine durability and compatibility using biodiesel-bioethanol-diesel blends in compression-ignition engines," *Energy Conversion and Management*, vol. 128, pp. 66–81, 2016.

[26] K. Sudhakar, M. Rajesh, and M. Premalatha, "Carbon mitigation potential of Jatropha biodiesel in Indian context," *Energy Procedia*, vol. 14, pp. 1421–1426, 2012.

[27] M. Jain, U. Chandrakant, V. Orsat, and V. Raghavan, "A review on assessment of biodiesel production methodologies from *Calophyllum inophyllum* seed oil," *Industrial Crops and Products*, vol. 114, pp. 28–44, 2018.

[28] M. Mofijur, H. H. Masjuki, M. A. Kalam, S. M. Ashrafur Rahman, and H. M. Mahmudul, "Energy scenario and biofuel policies and targets in ASEAN countries," *Renewable and Sustainable Energy Reviews*, vol. 46, pp. 51–61, 2015.

[29] F. Kusumo, A. S. Silitonga, H. H. Masjuki, H. C. Ong, J. Siswantoro, and T. M. I. Mahlia, "Optimization of transesterification process for Ceiba pentandra oil: a comparative study between kernel-based extreme learning

machine and artificial neural networks," *Energy*, vol. 134, pp. 24–34, 2017.

[30] F. Xue, X. Zhang, H. Luo, and T. Tan, "A new method for preparing raw material for biodiesel production," *Process Biochemistry*, vol. 41, no. 7, pp. 1699–1702, 2006.

[31] I. E. Agency, *Share of Renewables in Power Generation in the Sustainable Development Scenario, 2000-2030*IEA, Paris, France, 2000, https://www.iea.org/data-and-statistics/charts/share-of-renewables-in-power-generation-in-the-sustainable-development-scenario-2000-2030.

[32] M. F. Ortega, D. Donoso, H. Bousbaa et al., "Optimized production of fatty acid ethyl esters (FAEE) from waste frying oil by response surface methodology," *Waste and Biomass Valorization*, vol. 12, no. 5, pp. 2303–2310, 2021.

[33] C. Muthukumaran, R. Praniesh, P. Navamani, R. Swathi, G. Sharmila, and N. Manoj Kumar, "Process optimization and kinetic modeling of biodiesel production using non-edible Madhuca indica oil," *Fuel*, vol. 195, pp. 217–225, 2017.

[34] B. Aghel, M. Mohadesi, S. Sahraei, and M. Shariatifar, "New heterogeneous process for continuous biodiesel production in microreactors," *The Canadian Journal of Chemical Engineering*, vol. 95, no. 7, pp. 1280–1287, 2017.

[35] M. Mostafaei, H. Javadikia, and L. Naderloo, "Modeling the effects of ultrasound power and reactor dimension on the biodiesel production yield: comparison of prediction abilities between response surface methodology (RSM) and adaptive neuro-fuzzy inference system (ANFIS)," *Energy*, vol. 115, pp. 626–636, 2016.

[36] H. Moayedi, B. Aghel, L. K. Foong, and D. T. Bui, "Feature validity during machine learning paradigms for predicting biodiesel purity," *Fuel*, vol. 262, p. 116498, 2020.

[37] M. Guo, W. Jiang, C. Chen et al., "Process optimization of biodiesel production from waste cooking oil by esterification of free fatty acids using La^{3+}/ZnO-TiO_2 photocatalyst," *Energy Conversion and Management*, vol. 229, p. 113745, 2021.

[38] Y. Amini, M. B. Gerdroodbary, M. R. Pishvaie, R. Moradi, and S. M. Monfared, "Optimal control of batch cooling crystallizers by using genetic algorithm," *Case Studies in Thermal Engineering*, vol. 8, pp. 300–310, 2016.

[39] G. R. Moradi, S. Dehghani, F. Khosravian, and A. Arjmandzadeh, "The optimized operational conditions for biodiesel production from soybean oil and application of artificial neural networks for estimation of the biodiesel yield," *Renewable Energy*, vol. 50, pp. 915–920, 2013.

[40] H. Ghaderi, M. S. Pishvaee, and A. Moini, "Biomass supply chain network design: an optimization-oriented review and analysis," *Industrial Crops and Products*, vol. 94, pp. 972–1000, 2016.

[41] Y. Shi, H.-W. Ge, and R. D. Reitz, *Computational Optimization of Internal Combustion Engines*, Springer Science & Business Media, Berlin, Germany, 2011.

[42] S. Salam, T. Choudhary, A. Pugazhendhi, T. N. Verma, and A. Sharma, "A review on recent progress in computational and empirical studies of compression ignition internal combustion engine," *Fuel*, vol. 279, p. 118469, 2020.

[43] R. Anderson, D. Keshwani, A. Guru, H. Yang, S. Irmak, and J. Subbiah, "An integrated modeling framework for crop and biofuel systems using the DSSAT and GREET models," *Environmental Modelling & Software*, vol. 108, pp. 40–50, 2018.

[44] T. Kessler, E. R. Sacia, A. T. Bell, and J. H. Mack, "Artificial neural network based predictions of cetane number for furanic biofuel additives," *Fuel*, vol. 206, pp. 171–179, 2017.

[45] S. Russell and P. Norvig, *Artificial Intelligence-A Modern Approach*, M. Hirsch, Ed., Pearson Education, Inc, Hoboken, NJ, USA, 3rd edition, 2010.

[46] J. Mohd Ali, M. A. Hussain, M. O. Tade, and J. Zhang, "Artificial Intelligence techniques applied as estimator in chemical process systems-a literature survey," *Expert Systems with Applications*, vol. 42, no. 14, pp. 5915–5931, 2015.

[47] M. Liao and Y. Yao, "Applications of artificial intelligence-based modeling for bioenergy systems: a review," *GCB Bioenergy*, vol. 13, no. 5, pp. 774–802, 2021.

[48] E. Betiku, S. S. Okunsolawo, S. O. Ajala, and O. S. Odedele, "Performance evaluation of artificial neural network coupled with generic algorithm and response surface methodology in modeling and optimization of biodiesel production process parameters from shea tree (Vitellaria paradoxa) nut butter," *Renewable Energy*, vol. 76, pp. 408–417, 2015.

[49] J. Prakash Maran and B. Priya, "Modeling of ultrasound assisted intensification of biodiesel production from neem (Azadirachta indica) oil using response surface methodology and artificial neural network," *Fuel*, vol. 143, pp. 262–267, 2015.

[50] M. L. Traver, R. J. Atkinson, and C. M. Atkinson, "Neural network-based diesel engine emissions prediction using in-cylinder combustion pressure," *SAE Transactions*, vol. 108, pp. 1166–1180, 1999.

[51] H.-Y. Shin, S.-M. Lim, S. C. Kang, and S.-Y. Bae, "Statistical optimization for biodiesel production from rapeseed oil via transesterificaion in supercritical methanol," *Fuel Processing Technology*, vol. 98, pp. 1–5, 2012.

[52] S. Wu, Y. Wang, and S. Cheng, "Extreme learning machine based wind speed estimation and sensorless control for wind turbine power generation system," *Neurocomputing*, vol. 102, pp. 163–175, 2013.

[53] H. Guang-Bin, Z. Qin-Yu, and S. Chee-Kheong, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, Budapest, Hungary, July 2004.

[54] A. Keliris, H. Salehghaffari, B. Cairl et al., "Machine learning-based defense against process-aware attacks on industrial control systems," in *Proceedings of the 2016 IEEE International Test Conference (ITC)*, November 2016.

[55] D. Weichert, P. Link, A. Stoll et al., "A review of machine learning for the optimization of production processes," *The International Journal of Advanced Manufacturing Technology*, vol. 104, no. 5, pp. 1889–1902, 2019.

[56] H. Lahdhiri, M. Said, K. Ben Abdellafou, O. Taouali, and M. Faouzi Harkat, "Supervised process monitoring and fault diagnosis based on machine learning methods," *The International Journal of Advanced Manufacturing Technology*, vol. 102, no. 5, pp. 2321–2337, 2019.

[57] Z. Liu and A. Baghban, "Application of LSSVM for biodiesel production using supercritical ethanol solvent," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 39, no. 17, pp. 1869–1874, 2017.

[58] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210–229, 1959.

[59] K. Castillo-Villar, "Metaheuristic algorithms applied to bioenergy supply chain problems: theory, review, challenges, and future," *Energies*, vol. 7, no. 11, pp. 7640–7672, 2014.

[60] H. N. Bhatti, M. A. Hanif, M. Qasim, and H. Ata-ur-Rehman, "Biodiesel production from waste tallow," *Fuel*, vol. 87, no. 13, pp. 2961–2966, 2008.

[61] S. H. Teo, A. Islam, H. R. F. Masoumi et al., "Effective synthesis of biodiesel from Jatropha curcas oil using betaine assisted nanoparticle heterogeneous catalyst from eggshell of *Gallus domesticus*," *Renewable Energy*, vol. 111, pp. 892–905, 2017.

[62] R. R. C. Bastos, A. P. Da Luz Corrêa, P. T. S. Da Luz, G. N. Da Rocha Filho, J. R. Zamian, and L. R. V. Da Conceição, "Optimization of biodiesel production using sulfonated carbon-based catalyst from an amazon agro-industrial waste," *Energy Conversion and Management*, vol. 205, p. 112457, 2020.

[63] K. Sandesh and P. Ujwal, "Trends and perspectives of liquid biofuel-process and industrial viability," *Energy Conversion and Management: X*, vol. 10, p. 100075, 2021.

[64] K. Bunyakiat, S. Makmee, R. Sawangkeaw, and S. Ngamprasertsith, "Continuous production of biodiesel via transesterification from vegetable oils in supercritical methanol," *Energy & Fuels*, vol. 20, no. 2, pp. 812–817, 2006.

[65] M. J. Haas, A. J. McAloon, W. C. Yee, and T. A. Foglia, "A process model to estimate biodiesel production costs," *Bioresource Technology*, vol. 97, no. 4, pp. 671–678, 2006.

[66] L. F. P. Brandão, J. W. B. Braga, and P. A. Z. Suarez, "Determination of vegetable oils and fats adulterants in diesel oil by high performance liquid chromatography and multivariate methods," *Journal of Chromatography A*, vol. 1225, pp. 150–157, 2012.

[67] M. J. Montefrio, T. Xinwen, and J. P. Obbard, "Recovery and pre-treatment of fats, oil and grease from grease interceptors for biodiesel production," *Applied Energy*, vol. 87, no. 10, pp. 3155–3161, 2010.

[68] K. G. Harding, J. S. Dennis, H. von Blottnitz, and S. T. L. Harrison, "A life-cycle comparison between inorganic and biological catalysis for the production of biodiesel," *Journal of Cleaner Production*, vol. 16, no. 13, pp. 1368–1378, 2008.

[69] G. Vicente, M. Martinez, and J. Aracil, "Optimisation of integrated biodiesel production. part I. a study of the biodiesel purity and yield," *Bioresource Technology*, vol. 98, no. 9, pp. 1724–1733, 2007.

[70] I. M. Atadashi, M. K. Aroua, A. R. A. Aziz, and N. M. N. Sulaiman, "Refining technologies for the purification of crude biodiesel," *Applied Energy*, vol. 88, no. 12, pp. 4239–4251, 2011.

[71] F. Qiu, Y. Li, D. Yang, X. Li, and P. Sun, "Biodiesel production from mixed soybean oil and rapeseed oil," *Applied Energy*, vol. 88, no. 6, pp. 2050–2055, 2011.

[72] C. L. Peterson and T. Hustrulid, "Carbon cycle for rapeseed oil biodiesel fuels," *Biomass and Bioenergy*, vol. 14, no. 2, pp. 91–101, 1998.

[73] F. Ma and M. A. Hanna, "Biodiesel production: a review1 journal series #12109, agricultural research division, institute of agriculture and natural resources, University of Nebraska-Lincoln.1," *Bioresource Technology*, vol. 70, no. 1, pp. 1–15, 1999.

[74] W. A. G. Pessoa Junior, M. L. Takeno, F. X. Nobre et al., "Application of water treatment sludge as a low-cost and eco-friendly catalyst in the biodiesel production via fatty acids esterification: process optimization," *Energy*, vol. 213, p. 118824, 2020.

[75] D. Y. C. Leung and Y. Guo, "Transesterification of neat and used frying oil: optimization for biodiesel production," *Fuel Processing Technology*, vol. 87, no. 10, pp. 883–890, 2006.

[76] K. G. Georgogianni, M. G. Kontominas, P. J. Pomonis, D. Avlonitis, and V. Gergis, "Conventional and in situ transesterification of sunflower seed oil for the production of biodiesel," *Fuel Processing Technology*, vol. 89, no. 5, pp. 503–509, 2008.

[77] M. Corral Bobadilla, R. Fernández Martínez, R. Lostado Lorza, F. Somovilla Gómez, and E. Vergara González, "Optimizing biodiesel production from waste cooking oil using genetic algorithm-based support vector machines," *Energies*, vol. 11, no. 11, p. 2995, 2018.

[78] X. Meng, M. Jia, and T. Wang, "Neural network prediction of biodiesel kinematic viscosity at 313K," *Fuel*, vol. 121, pp. 133–140, 2014.

[79] M. Aghbashlo, W. Peng, M. Tabatabaei et al., "Machine learning technology in biodiesel research: a review," *Progress in Energy and Combustion Science*, vol. 85, p. 100904, 2021.

[80] G. Di Nicola, M. Moglie, M. Pacetti, and G. Santori, "Bio-energy II: modeling and multi-objective optimization of different biodiesel production processes," *International Journal of Chemical Reactor Engineering*, vol. 8, no. 1, 2010.

[81] E. Habyarimana, I. Piccard, M. Catellani, P. De Franceschi, and M. Dall'Agata, "Towards predictive modeling of sorghum biomass yields using fraction of absorbed photosynthetically active radiation derived from sentinel-2 satellite imagery and supervised machine learning techniques," *Agronomy*, vol. 9, no. 4, p. 203, 2019.

[82] T. Huntington, X. Cui, U. Mishra, and C. D. Scown, "Machine learning to predict biomass sorghum yields under future climate scenarios," *Biofuels, Bioproducts and Biorefining*, vol. 14, no. 3, pp. 566–577, 2020.

[83] C. J. Gleason and J. Im, "Forest biomass estimation from airborne LiDAR data using machine learning approaches," *Remote Sensing of Environment*, vol. 125, pp. 80–91, 2012.

[84] E. K. Lee, W.-J. Zhang, X. Zhang et al., "Projecting life-cycle environmental impacts of corn production in the U.S. Midwest under future climate scenarios using a machine learning approach," *Science of The Total Environment*, vol. 714, p. 136697, 2020.

[85] P. Yang, Q. Zhao, and X. Cai, "Machine learning based estimation of land productivity in the contiguous US using biophysical predictors," *Environmental Research Letters*, vol. 15, no. 7, p. 074013, 2020.

[86] A. Q. Mairizal, S. Awad, C. R. Priadi et al., "Experimental study on the effects of feedstock on the properties of biodiesel using multiple linear regressions," *Renewable Energy*, vol. 145, pp. 375–381, 2020.

[87] S. O. Giwa, S. O. Adekomaya, K. O. Adama, and M. O. Mukaila, "Prediction of selected biodiesel fuel properties using artificial neural network," *Frontiers in Energy*, vol. 9, no. 4, pp. 433–445, 2015.

[88] A. P. Tchameni, L. Zhao, J. X. F. Ribeiro, and T. Li, "Predicting the rheological properties of waste vegetable oil biodiesel-modified water-based mud using artificial neural network," *Geosystem Engineering*, vol. 22, no. 2, pp. 101–111, 2019.

[89] S. Soltani, U. Rashid, T. Roodbar Shojaei, I. A. Nehdi, and M. Ibrahim, "Modeling of the nanocrystalline-sized mesoporous zinc oxide catalyst using an artificial neural network for efficient biodiesel production," *Chemical Engineering Communications*, vol. 206, pp. 1–15, 2018.

[90] I. Ahmad, A. Ayub, U. Ibrahim, M. Khattak, and M. Kano, "Data-based sensing and stochastic analysis of biodiesel production process," *Energies*, vol. 12, no. 1, p. 63, 2019.

[91] A. Sarve, S. S. Sonawane, and M. N. Varma, "Ultrasound assisted biodiesel production from sesame (*Sesamum indicum L.*) oil using barium hydroxide as a heterogeneous catalyst: comparative assessment of prediction abilities between response surface methodology (RSM) and artificial

neural network (ANN)," *Ultrasonics Sonochemistry*, vol. 26, pp. 218–228, 2015.

[92] S. Kumar, S. Jain, and H. Kumar, "Prediction of Jatropha-algae biodiesel blend oil yield with the application of artificial neural networks technique," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 41, no. 11, pp. 1285–1295, 2019.

[93] A. Banerjee, D. Varshney, S. Kumar, P. Chaudhary, and V. K. Gupta, "Biodiesel production from castor oil: ANN modeling and kinetic parameter estimation," *International Journal of Industrial Chemistry*, vol. 8, no. 3, pp. 253–262, 2017.

[94] G. Kanat and A. Saral, "Estimation of biogas production rate in a thermophilic UASB reactor using artificial neural networks," *Environmental Modeling & Assessment*, vol. 14, no. 5, pp. 607–614, 2009.

[95] S. Kumar, S. Jain, and H. Kumar, "Process parameter assessment of biodiesel production from a Jatropha-algae oil blend by response surface methodology and artificial neural network," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 39, no. 22, pp. 2119–2125, 2017.

[96] R. Chakraborty, S. Das, P. Pradhan, and P. Mukhopadhyay, "Prediction of optimal conditions in the methanolysis of mustard oil for biodiesel production using cost-effective mg-solid catalysts," *Industrial & Engineering Chemistry Research*, vol. 53, no. 51, pp. 19681–19689, 2014.

[97] S. Kumar, "Comparison of linear regression and artificial neural network technique for prediction of a soybean biodiesel yield," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 42, no. 12, pp. 1425–1435, 2020.

[98] J. Guo and A. Baghban, "Application of ANFIS strategy for prediction of biodiesel production using supercritical methanol," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 39, no. 17, pp. 1862–1868, 2017.

[99] M.-Y. Cheng, D. Prayogo, Y.-H. Ju, Y.-W. Wu, and S. Sutanto, "Optimizing mixture properties of biodiesel production using genetic algorithm-based evolutionary support vector machine," *International Journal of Green Energy*, vol. 13, no. 15, pp. 1599–1607, 2016.

[100] S. Sivamani, S. Selvakumar, K. Rajendran, and S. Muthusamy, "Artificial neural network-genetic algorithm-based optimization of biodiesel production from Simarouba glauca," *Biofuels*, vol. 10, no. 3, pp. 393–401, 2019.

[101] B. Ogaga Ighose, I. A. Adeleke, M. Damos, H. Adeola Junaid, K. Ernest Okpalaeke, and E. Betiku, "Optimization of biodiesel production from Thevetia peruviana seed oil by adaptive neuro-fuzzy inference system coupled with genetic algorithm and response surface methodology," *Energy Conversion and Management*, vol. 132, pp. 231–240, 2017.

[102] S. Dhingra, K. K. Dubey, and G. Bhushan, "A polymath approach for the prediction of optimized transesterification process variables of polanga biodiesel," *Journal of the American Oil Chemists' Society*, vol. 91, no. 4, pp. 641–653, 2014.

[103] M. Karimi, B. Jenkins, and P. Stroeve, "Multi-objective optimization of transesterification in biodiesel production catalyzed by immobilized lipase," *Biofuels, Bioproducts and Biorefining*, vol. 10, no. 6, pp. 804–818, 2016.

[104] D. S. Patle, S. Sharma, Z. Ahmad, and G. P. Rangaiah, "Multi-objective optimization of two alkali catalyzed processes for biodiesel from waste cooking oil," *Energy Conversion and Management*, vol. 85, pp. 361–372, 2014.

[105] M. Babaei Rouchi, M. Khanmohammadi Khorrami, A. B. Garmarudi, and M. De la Guardia, "Application of infrared spectroscopy as process analytics technology (PAT) approach in biodiesel production process utilizing multivariate curve resolution alternative least square (MCR-ALS)," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 213, pp. 347–353, 2019.

[106] M. R. Shukri, D. Ramasamy, K. Kadirgama, and R. ShukriM, "Artificial neural network optimization modeling on engine performance of diesel engine using biodiesel fuel," *International Journal of Automotive & Mechanical Engineering*, vol. 11, 2015.

[107] M. Aghbashlo, S. Hosseinpour, M. Tabatabaei, and M. Mojarab Soufiyan, "Multi-objective exergetic and technical optimization of a piezoelectric ultrasonic reactor applied to synthesize biodiesel from waste cooking oil (WCO) using soft computing techniques," *Fuel*, vol. 235, pp. 100–112, 2019.

[108] A. N. Sarve, M. N. Varma, and S. S. Sonawane, "Response surface optimization and artificial neural network modeling of biodiesel production from crude mahua (Madhuca indica) oil under supercritical ethanol conditions using CO_2 as co-solvent," *RSC Advances*, vol. 5, no. 85, pp. 69702–69713, 2015.

[109] M. A. Noriega and P. C. Narváez, "UNIFAC correlated parameters for liquid-liquid equilibrium prediction of ternary systems related to biodiesel production process," *Fuel*, vol. 249, pp. 365–378, 2019.

[110] B. Y. López-Zapata, M. Adam-Medina, P. E. Álvarez-Gutiérrez, J. P. Castillo-González, H. R. H.-D. León, and L. G. Vela-Valdés, "Virtual sensors for biodiesel production in a batch reactor," *Sustainability*, vol. 9, no. 3, p. 455, 2017.

[111] I. Fahmi and S. Cremaschi, "Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models," *Computers & Chemical Engineering*, vol. 46, pp. 105–123, 2012.