

Review

# Machine Learning Applications in Biofuels' Life Cycle: Soil, Feedstock, Production, Consumption, and Emissions

Iftikhar Ahmad <sup>1,\*</sup>, Manabu Kano <sup>2</sup>, Brenno C. Menezes <sup>3</sup>, Izzat Iqbal Cheema <sup>4,5</sup>, Adil Sana <sup>1</sup>, Junaid Shahzad <sup>1</sup>, Zahid Ullah <sup>1</sup>, Muzammil Khan <sup>1</sup>, and Asad Habib <sup>6</sup>

<sup>1</sup> Department of Chemical and Materials Engineering, National University of Sciences and Technology, Islamabad 44000, Pakistan

<sup>2</sup> Department of Systems Science, Kyoto University, Kyoto 606-8501, Japan

<sup>3</sup> Division of Engineering Management and Decision Sciences, College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar

<sup>4</sup> Department of Chemical, Polymer and Composite Materials Engineering, University of Engineering and Technology, Lahore (New Campus) 39021, Pakistan

<sup>5</sup> Center for Energy Research and Development, University of Engineering and Technology, Lahore (New Campus) 39021, Pakistan

<sup>6</sup> Institute of Computing, Kohat University of Science and Technology, Kohat 26000, Pakistan

\* Correspondence: [iftikhar.salarzai@scme.nust.edu.pk](mailto:iftikhar.salarzai@scme.nust.edu.pk)

**Abstract:** Machine learning (ML) is penetrating in all walks of life and is one of the major driving forces behind the fourth industrial revolution, typically known as Industry 4.0. This study reviews the state-of-the-art ML applications in the biofuels' life cycle stages, i.e., soil, feedstock, production, consumption, and emissions. A keyword search is performed to retrieve relevant articles from the databases of the Web of Science and Google Scholar. ML applications in the soil stage were mostly based on the use of satellite images of land for estimation of biofuels yield or suitability analysis of agricultural land. In the second stage of the life cycle, assessment of rheological properties of the feedstocks and their effect on the quality of biofuels were dominant studies reported in the literature. The production stage included estimation and optimization of quality, quantity, and process conditions. The fuel consumption and emissions stage included analysis of engine performance and estimation of emissions temperature and composition, such as NO<sub>x</sub>, CO, and CO<sub>2</sub>. This study identified the following trends: dominant ML method, the stage of life cycle getting more usage of ML, the type of data used for the development of the ML-based models, and the stage-wise frequently used input and output variables. The findings of this article are beneficial for academia and industry-related people involved in model development in different stages of biofuel's life cycle.

**Keywords:** Bio-energy; Artificial intelligence; Industry 4.0; Biodiesel; Biogas; Renewable energy; Supply Chain

## 1. Introduction

Machine learning (ML) is one of the major driving forces behind the fourth industrial revolution, typically known as Industry 4.0. This is fundamentally inspired by human brain learning. ML enables a computer system to solve knotty research questions through implicit and automated "learning", and additionally do self-improvement without being explicitly pre-programmed implementation for it. From an algorithmic point of view, the term machine refers to an automated process (computer program) that incrementally updates its problem-solving capability through successive iterations based on inputs of external variants. The concept of machine learning was introduced by Arthur Samuel, who is among the pioneers of modern Artificial Intelligence (AI) [1].

The early and dominant machine learning methods include Decision Trees (DT), Genetic Algorithms (GA), Artificial Neural Networks (ANN), Bayesian Networks, Random Forests (RF),

Support Vector Machines (SVM), Regression Analysis, K-Means clustering, Naive Bayes classifier, K-Nearest Neighbor Classifier (kNN), and Principal Component Analysis (PCA). These methods have evolved into deep learning, extreme learning, transfer learning, and reinforcement learning. ML has been applied across almost all types of industrial processes for various purposes, such as data mining [2], process automation [3], process-aware attacks on industrial control systems [4], predictive maintenance [5], optimization [6], process monitoring and fault diagnosis [7], and industrial tomography [8].

Ge et al. [2] categorized the industrial process ML techniques into supervised, unsupervised, and the semi-supervised ML algorithms. This review reported that the supervised learning algorithms focus on fault diagnosis and classification, operation mode localization, process monitoring, key performance index prediction, soft sensor modeling, and quality estimation, whereas the unsupervised learning algorithms are applied in features dimensionality optimization, data visualization, data clustering, density estimation, outlier detection, and process monitoring. The semi-supervised learning algorithms are used in hybrid tasks, such as data regression and classification. Extensive reviews have been performed to analyze the state-of-the-art of ML applications in process industries, such as chemical and petrochemical [9], petroleum industry [10], cement industry [11], pharmaceutical industry [12], and iron and steel industry [13]. The integration of ML in conventional industrial systems has brought about precision-based automation that subsequently resulted in the fabrication of high-quality products at a minimal cost in a shorter time.

Biofuel production is getting prominence in the manufacturing sector because of the renewable nature of its raw materials and environmentally friendly products, i.e., biofuels. ML is capable of playing a vital role in realizing the efficient and stable operation of biofuel production. ML application across all stages of the biofuels life cycle, i.e., soil, feedstock, production, consumption, and emission, has been reported as shown in Figure 1. The work from [14] and [15] present an overview of the ML and deep learning application in the biofuels. However, no review has been reported in the literature that exclusively discusses ML application in all stages of the biofuels life cycle. This article presents a systematic review in finding answers to the following questions: what are dominant ML methods?, which stage of the life cycle is getting more usage of ML?, what type of data is used for the development of the ML-based models? what are the stage-wise trends in input and output variables?. The findings of this article are beneficial for academia and industry related persons involved in model development in different stages of biofuel's life cycle.

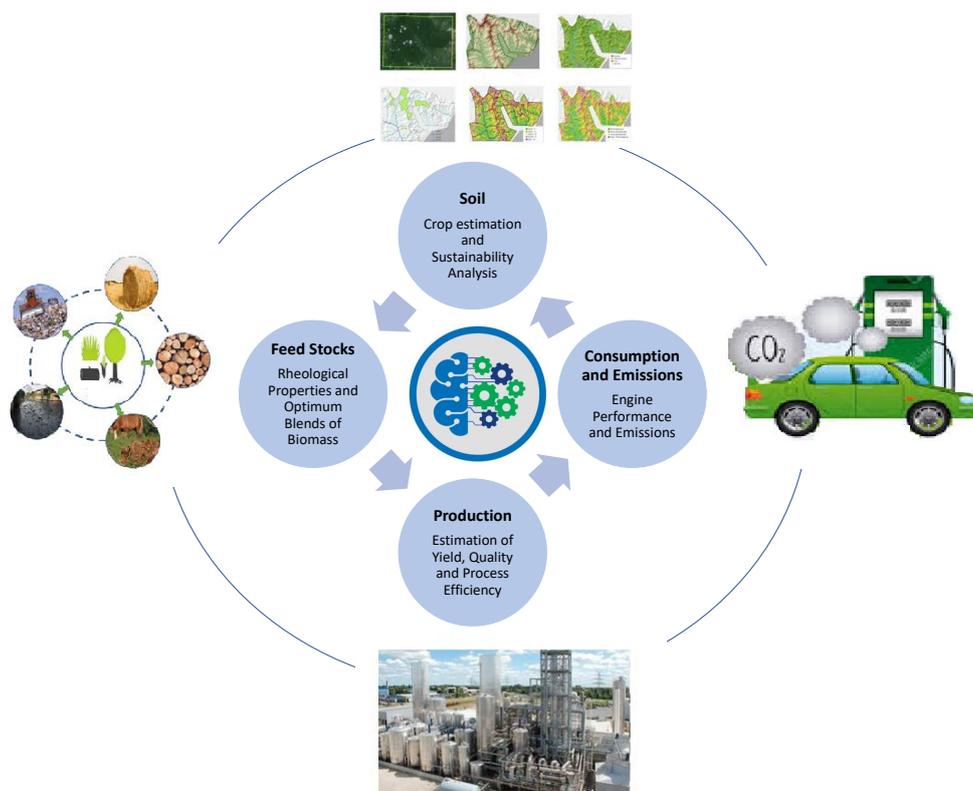
The article is structured as follows: in section 2, the methodology adopted in this work to perform the literature review and the approach used to obtain the citation data set are discussed followed by section 3 where the results and discussions of the review are presented. Conclusions and future prospects are discussed in section 4.

## 2. Methodology

For literature retrieval, the Web of Science and Google Scholar databases were used. To restrict the literature to a rational size for accommodation in a review paper, only research articles were reviewed. The dominant part of the literature was collected from the Web of Science database; however, Google Scholar was also used for collecting some gray literature regarding ML applications in soil stage. Keywords listed in Table 1 were used to search relevant literature in the Web of Science database.

The following two rules were applied while deciding the relevancy of the articles: 1) title of the article should contain at least one word from each of the first two keywords categories, i.e., machine learning and biofuels, of Table 1, 2) the title should also contain at least one word from any one of the next four categories, i.e., soil, feedstock, production, consumption, and emissions, of Table 1. The collected literature was divided into four stages of the life cycle on the basis of keywords found in step 2 of the above-mentioned procedure.

In the first round, the number of papers related to soil, feed stock, production, and consumption and emission stages were 10, 40, 90, and 75, respectively. The collected literature was studied to



**Figure 1.** Machine learning applications in biofuels life cycle

screen-out the irrelevant papers. Some papers were shifted across the stages. After the screening, the production stage had the largest number of papers (73) on ML application followed by consumption and emission with (49), soil (5) and feedstock (5).

**Table 1.** Keywords used in literature survey

Categories	Keywords used for search in the databases
Machine learning	artificial neural network   boosting   data-based   data-driven   decision tree   deep learning   dimensionality reduction algorithms   discriminant analysis   ensemble learning   estimation   extreme learning machine   genetic algorithm   inference   kNN   K-Means   least-squares   logistic regression   linear regression   machine learning   moving average   multi objective optimization   multi layered perception   Naive Bayes   neuro fuzzy   partial least squares   principal component analysis   prediction   random forest(s)   soft sensor   support vector machine   virtual sensor
Biofuels	bioalcohol   biodiesel   bioethers   biofuels   biogas   biohydrogen   dimethylfuran   green diesel
Soil	drone   land   land image   satellite   soil   surveillance
Feedstock	algae and aquatic biomass   biomass   biosolid(s)   corn   energy cane   feed   feedstock   forest thinning   high biomass sorghum   hybrid poplars   industrial waste gases   logging residues   lignocellulosic crops   lignocellulosic residues   manure slurries   micro algae   miscanthus   municipal waste   oil-based residues   oil crops   organic residues   plant   plastics   raw material(s)   shrub willows   sludge   starch crops   sugar crops   sweet sorghum   switch grass   vegetable oil   waste   waste food   waste gases
Production	catalytic synthesis   distillation   drying   fermentation   gas cleaning   gasification   operation   process   product   production   reactor   refining   unit   water gas shift   yield
Consumption & emissions	air pollution   carbon emission   emission   energy potential   engine   environment   exhaust gases   fuel consumption   fuel quality   fuel use   green house gases   greenhouse   mileage

### 3. Applications of ML methods in biofuels' life cycle

In this section, the stage-wise application of machine learning (ML) methods is discussed. ML application in soil, feedstock, production, and consumption, engine performance and emission stages are discussed in sections 3.1, 3.2, 3.3 and 3.4, respectively. At the end, applications summary is presented in section 3.5.

#### 3.1. Soil

Several studies on ML application in the soil stage of the biofuels' life cycle have been reported both at tree and plot level. For example, [16] compared the Linear Mixed-effects Regression, Support Vector Regression (SVR), RF, and Cubist for prediction of biomass in a 40 to 60 % canopy closure forest. SVR outperformed the other methods. [17] used RF to predict future trends in sorghum bicolor yield under two irrigation regimes and four Greenhouse gas (GHG) emission scenarios. The RF model trained on data samples uniquely identified by year and country achieved reasonable accuracy in the prediction. In another study [18], various ML methods were used for prediction of sorghum biomass yield based on satellite images of sorghum fields. This ML methods included PLS discriminant analysis (PLS-DA), PCA discriminant analysis (PCA-DA), ANN, RF, nonlinear kernel (SVM-G), radial basis kernel (SVM-R), radial basis kernel with polynomial basis kernel (SVM-P), SVM with linear classifier (SVM), eXtreme Gradient Boosting-xgbtree method (GBT), eXtreme Gradient Boosting-xgbLinear method (GBL), eXtreme Gradient Boosting-xgbDART method (GBD), and a simple linear model. The GBT outperformed the other methods. [19] used the Boosted Regression Tree (BRT) model to asses environmental impacts of corn production for a duration of years 2022 to 2100. The study was conducted in the context of four emissions scenarios where BRT model achieved correlation coefficients of 0.82 and 0.78, in estimating eutrophication impacts and global warming, respectively. [20] used two-step ML approach. Initially, Gaussian Process Model (GPM) was used for crop yield downscaling followed by yield estimation through RF model. The GPM which is a Bayesian inference method helped in realizing accurate estimations.

#### 3.2. Feedstock

The feedstock stage of the biofuels also got attention in terms of ML application. Mahanty et al. [21] used ANN and statistical regression models to predict specific methane yield in the production of biogas from industrial sludges. The ANN model performed better than the statistical regression model. It was revealed that sludges from the chemical industry have a relatively higher impact on methane in the produced biogas. Mairizal et al. [22] used multiple linear regressions to predict viscosity, flash point, density, higher heating value, and oxidative stability of biodiesel produced from sunflower oil, peanut oil, hydrogenated coconut oil, hydrogenated copra oil, beef tallow, rapeseed oil, and walnut oil. Saponification value, iodine value, and the polyunsaturated fatty acids content of feedstock were used as inputs of the model. It was inferred from the results that the addition of PU/MU as an independent parameter increase prediction performance. Giwa et al. [23] used ANN to estimate cetane number, kinematic viscosity, flash point, and density of biodiesel produced from fatty acid. Accuracy of estimation and the average of absolute deviation of the model were as follows: CN (96.6%; 1.637%), FP (99.07%;0.997%), KV (95.80%; 1.638%), and density (99.40%; 0.101%). Tchameni et al. [24] used Multiple Non-Linear Regression (MNL) and ANN for prediction of rheological properties of waste vegetable oil for production of biodiesel. The ANN model had superior performance over the MNL method. Dahunsi et al. [25] used single and multiple linear regressions to estimate methane yield in biomass structural components. A fairly high correlation was found between the chemical composition and methane potentials of the biomass.

### 3.3. *Production*

Several studies have been reported for prediction and optimization of quality, yield (quantity), and process conditions, i.e., pressure, temperature, flow rate, etc., in the production stage of the biofuels' life cycle. These studies are categorized based on the type of biofuel produced, such as biodiesel, biogas, biohydrogen, and other miscellaneous cases. Trends in ML applications in the production stage are summarized in Table 2 and discussed in the following subsections.

**Table 2.** Summary of ML application in production phase

Production process	Purpose	ML method ranking	Input variables ranking	Output variables ranking	References
Biodiesel	Quality estimation	ANN (2), Least Squares Boosting (LSBoost) integrated with polynomial chaos expansion (PCE) method, regression models, principle component analysis	Reaction temperature (4), reaction time (2), metal ratio, and calcination temperature, flow rate, pressure, reactor residence time, reflux rate, oil fraction, , methanol-to-oil molar ratio, catalyst concentration	Viscosity (2), density (2), size of the nanocrystalline, composition, quantity, quality (cetane number), percentage conversion, and FAME content	[26–30]
	Yield estimation	ANN (11), ANFIS (2), linear regression (LR)	Temperature (11), methanol-to-oil molar ratio(10), catalyst concentration (9), reaction time (7), organic loading rate, influent–effluent pH, H <sub>2</sub> SO <sub>4</sub> concentration, total volatile fatty acid (VFA) of the effluent, xylose, influent–effluent alkalinity, initial pH, pressure, reactor diameter, liquid height and ultrasound intensity	Biodiesel yield (9), % fatty acid methyl ester (FAME) yield, biogas production rate, biohydrogen yield, and biodiesel production	[31–41]
	Quality and yield optimization	ANN-GA(9), Genetic Algorithm-based Support Vector Machines (4), ANFIS-GA (2), RSM (2), multi-objective Genetic Algorithm, multivariate regression analysis, multi-objective optimization with Orthogonal collocation on finite elements (OCFE) method	Methanol-to-oil molar ratio (14), reaction temperature (14), reaction time (12), stirring speed (6), catalyst concentration (6), catalyst weight (4), dosage of NaOH catalyst (3), humidity (3), impurities (3), mixing time (3), Free Fatty Acid (FFA) content, Sulfuric acid-to-rice bran ratio, methanol-to-rice bran ratio, FAME concentration, overall heat duty, initial acid value of vegetable oil, calcination temperature, reactor’s residence time, and pressure	Biodiesel yield (12), FAME yield (3), final acid value of oil (3), minimization of the overall heat duty of the R-S processes, and maximization of profit, viscosity turbidity, density, high heating value	[3,42–56]
	Estimation and optimization of process conditions and efficiency	ANN (3), RSM (2), multi-objective optimization program with Genetic Algorithm (2), ANFIS, ALFIMO (Artificial linear interdependent fuzzy multi-objective optimization), Multivariate Curve Resolution Alternative Least Square (MCR-ALS), GIP group interaction parameters with GA, and extreme learning machine with Lyapunov analysis	Reaction temperature (4), concentration (2), reaction time (3), methanol-to-oil molar ratio (2), water content, residence time, vapor pressure, heat capacity, liquid molar volume, and liquid viscosity, ethanol-to-oil molar ratio, heat of vaporization, initial CO <sub>2</sub> pressure, pH, metal ratio, heat of formation, calcination temperature, and phase equilibrium	Yield (2), pressure (2), functional exergy efficiency (FEE), specific energy consumption, FAME content, normalized exergy destruction (NED), conversion efficiency (CE), Universal Exergy Efficiency (UEE), maximum quality of material, size of the nanocrystalline, LLE, air-to-fuel ratio, critical temperature, critical pressure and acentric factor, temperature, reaction time, ethanol-to-oil molar ratio, initial CO <sub>2</sub> , concentration and type of component in the process, thermal efficiency, heat release in percentage, exergy efficiency of the enzymatic transesterification, and volume generated	[57–69]

Biogas	Quality estimation	ANN	Working days, influent ammonia, influent total phosphorus (TP), waste adding ratio (%), influent chemical oxygen demand (COD), influent pH, influent alkalinity, hydraulic retention time (HRT), and additive waste	Biogas performance	[69]
		ANN and ANFIS	total solids, fixed solids, volatile solids, Volatile fatty acids (VFAs), pH, and inflow rate	Biogas production (CH <sub>4</sub> , CO <sub>2</sub> and H <sub>2</sub> S) estimation	[70]
		Multiple regression model	COD, ammonia (NH <sub>4</sub> ), pH, total dissolved solids (TDS), total Kjeldahl nitrogen (TKN), alkalinity (Alk), chloride (Cl), conductivity (Cond), and total phosphorus (TP)	Biogas amount prediction	[71]
	Yield estimation	ANN (5), Multiple Non-Linear Regression (MNLN) models, and Partial Least Squares Regression (PLS-R)	pH (3), temperature (2), Total Volatile Solids (TVS) (2), volatile fatty acids (VFAs) (2), composition, time, moisture content (MC), and CH <sub>4</sub> , total Kjeldahl nitrogen, total chemical oxygen demand, total phosphorus, hydraulic retention time, fractions of different industrial sludges, ammonium, alkalinity, instrument measurements (FFT acoustic spectra), and Total Solids (TS)	Biogas Yield (4), specific methane yield, amount of dry matter in bioslurry	[72–77]
	Optimization of quality and yield	ANN-GA (4), ANFIS	Total Solids (TS) (2), pH (2), temperature (2), digestion time and C-to-N ratio (2), Total Volatile Solids (TVS), organic loading rate, retention time, mass of poultry droppings, cow dung, plantain peel, piggery waste, stirring intensity of substrates	Biogas yield (3), fraction of methane (% CH <sub>4</sub> ), Biogas production, effluent COD, TSS and VFA concentrations	[78–82]
	Estimation of process efficiency	ANN-GA	Paper waste, banana stem, cow dung, saw dust, and rice bran	Biogas performance index	[83]
Biohydrogen	Yield estimation	ANN (3), hybrid fuzzy clustering-ranking approach with ANN, gray model	Biomass concentrations (5), pH (4), substrate (2), temperature (2), time, agitation speed and flow rate	Biohydrogen yield (4), exergetic outputs	[84–88]
	Optimization	Gene Expression Programming (GEP)	pH, temperature, substrate concentration and inner column volume	COD removal, and H <sub>2</sub> yield	[89]
Bioethanol	Yield Estimation	Fuzzy Neural Network (FNN) and PSO	Temperature, glucose content, and fermentation time	Prediction of bioethanol production	[90]
Bisabolene	Yield, Quality, and Process Condition Estimation	Convolutional neural networks-based multi-objective optimization with hybrid stochastic search optimization algorithm: random search (RS), PSO and SA	Incident light intensity, recycling gas flow rate, number of holes, diameter of holes, and cardinal coordinates of sample	Biomass concentration, bisabolene production, and friction velocity	[91]

### 3.3.1. Biodiesel

The studies reported in the biodiesel production phase are further classified based on the output, i.e., quality, yield, and process efficiency, of the ML models.

#### Quality estimation

[26] used the ANN model to determine optimum conditions for realizing desired nanocrystalline size of mesoporous SO<sub>3</sub>HZnO catalyst. Optimized conditions were calcine temperature of 700 C, 160 C reaction temperature, 18 min reaction time, and 4 mmol of Zn concentration. ([27]) used Least Squares Boosting (LSBoost) integrated with polynomial chaos expansion method in the production of vegetable oil based biodiesel under uncertainty. The average mean absolute deviation percent (MADP) values in the predicted values of the target output were 0.84 in response to 1 % uncertainty in each input variable of the models. Gulum et al. [28] used Regression and ANN models to predict viscosity and density of ternary blends consisting of biodiesel, diesel, and vegetable oil. Exponential and rational models previously suggested [92–96] were compared with regression models and ANN approach. Tomazzoni et al. [29] used PCA to estimate viscosity, relative density, and percentage conversion of vegetable oil to methyl esters in the production of diesel from vegetable oil. Through the use of PCA, they were able to differentiate between pure samples waste oil, diesel, and biodiesel from their respective blends. Sarve et al. [30] used ANN and Central Composite Design (CCD) to predict Fatty Acid Methyl Ester (FAME) content in the production of biodiesel from sesame oil. The study revealed that catalyst concentration has the highest impact on the FAME contents in the final product. The ANN model showed better performance with a small number of data-set.

#### Yield estimation

Several studies based on biodiesel yield prediction through ML methods have been reported in the literature where the biodiesel was formed from jatropha-algae, castor oil, and anaerobic sludge [31–33]. [31] developed an ANN model to predict biodiesel yield using various jatropha-algae oil blends as inputs. [32] used ANN model for predicting the fractional formation of FAME. They also devised a kinetic model using the experimental and computed data. The experimental and the ANN based predicted data was used to estimate the rate constants of the akintic model. The ANN model was able to predict the % FAME yield within 8% deviation. Kanat and Saral [33] used ANN to estimate the production rate of biodiesel from anaerobic sludge in a thermophilic up-flow anaerobic sludge blanket reactor. Longer time periods for moving average showed higher correlation coefficient of 0.927.

In several studies, ANN was used together with Response Surface Methodology (RSM) for prediction of the yield of biodiesel from Jatropha-algae oil [34], goat tallow [35], and enterobacter species [36]. In [34], ANN outperformed the RSM. [35] used RSM and ANN for identifying optimal parametric values that resulted in maximum FA conversion while maintaining FAME yield that met the American Society for Testing and Materials (ASTM) biodiesel specifications. ANN and RSM had comparable predictability performance. [36] compared performance of RSM model and ANN model. The ANN model had much higher prediction abilities than the RSM model.

Kumar [37] used ANN and linear regression (LR) to predict soybean-based biodiesel yield where the ANN outperformed the LR. [38] used ANN and kinetic models to estimate the yield of biodiesel from Soybean oil. The ANN exhibited the capability of learning from experimental data and simple to apply in comparison to the classic kinetic modeling method.

Guo et al. [39] and Mostafa et al.[40] used Adaptive Neuro-Fuzzy Interference System (ANFIS) to estimate biodiesel yield from algae oil blend, vegetable oil, and waste cooking oil. [39] achieved a low absolute deviation between experimental data and ANFIS model based predicted data. It confirmed suitability of ANFIS for prediction of the biodiesel yield from vegetable oil. [40] used the ANFIS model together with RSM where the ANFIS model outperformed the RSM model in terms of robustness and

prediction accuracy. Maran et al. [41] used ANN and RSM to predict biodiesel yield from muskmelon oil. ANN model again outperformed the RSM model.

#### Quality and yield optimization

Several studies have been reported that focus on optimization of biodiesel yield and quality. Bobadilla et al. [42] used a GA-based SVM to estimate and optimize biodiesel yield of specific properties: higher heating value with decreased viscosity, density, and turbidity from waste cooking oil. To produce biodiesel of high quality, the optimum inputs were dosage of catalyst (NaOH) from 1.00 to 1.38 wt %, molar ratio from 6.0 wt % to 8.4, mixing speed from 500 to 999.99 rpm, time from 20.00 to 26.94 min, temperature from 28.75 to 37.5 °C, humidity from 0 to 2.31 wt %, and impurities from 0 to 2.99 wt %. Cheng et al. [43] used a GA-based evolutionary SVM to predict and optimize the final acid value of oil in the production of biodiesel from rice bran. They found GA-ESVM better than ANN-GA and SVM. Sivamani et al. [44] used RSM and ANN coupled with GA to predict the yield of Simarouba glauca biodiesel. For higher yield, the optimum values for oil-to-alcohol ratio, temperature, and duration were found to be 1:6.22, 677.25 °C, and 20 h, respectively. Ighose et al. [45] used an ANFIS coupled with RSM and GA to realize high yield of biodiesel from *Thevetia peruviana* seed oil via transesterification process. The ANFIS outperformed the RSM model. In addition, the use of GA resulted in higher yield at less time and catalyst loading than RSM.

Dhingra et al. [46] used ANN and GA to predict and optimize yield in the production of polanga oil based biodiesel. They combined the ANN, RSM, and GA for predicting the optimized reaction conditions that resulted in a biodiesel yield of 91.08 % by weight significantly higher than 78.8 % obtained through RSM alone. Ishola et al. [47] used ANN, ANFIS, and GA to estimate and optimize biodiesel yield (methyl esters) from sorrel oil. ANFIS model outperformed the ANN model while RSM was the lowest performer in terms of prediction accuracy. In addition, GA also outperformed the RSM and obtained the highest biodiesel yield (methyl esters) of 99.71 wt % at the methanol-to-oil molar ratio, the catalyst weight and reaction time of 8:1, 1.23 wt %, and 43 min, respectively. Silitonga et al. [48] used ANN integrated with ant colony optimization to realize minimum acid value and maximum biodiesel yield from Cerberamanghas. For esterification, the optimum methanol-to-oil molar ratio was 10.5:1 while the best values for reaction time and reaction temperature were 71 min and 54.5 °C, respectively. The optimized catalyst concentration, reaction temperature, methanol-to-oil molar ratio, stirring speed, and reaction time for transesterification were 1.1 wt %, 55 °C, 10.9:1, 1100 rpm, and 72 min, respectively.

Chakraborty et al. [49] used multivariate regression analysis to predict optimum operating conditions for mustard oil (MO)-based biodiesel yield. Optimal values of methanol-to-MO molar ratio, calcination temperature, catalyst concentration, and stirrer speed, were 13.13:1, 950 °C, 3.44 wt %, and 890 rpm, respectively. Goharimanesh et al. [50] used multi-objective GA to obtain optimum reaction temperature for maximizing biodiesel production, amount of ester, and alcohol.

Oladipo et al. [51] used CCD and ANN to maximize FAME content in the production of biodiesel from crude neem, *Jatropha*, and waste cooking oils. They also found that safe reuse of the mesoporous catalyst can be carried out up-to five cycles. Rajendra et al. [52] used ANN, GA, and central compost in rotatable design to optimize the final acid value of oil in the production of biodiesel from *jatropha*, *simaruoba*, *mahua*, and rice bran oils. ANN-GA helped in determining optimum process conditions for realizing high quality yield. Corral et al. [53] used GA-SVM to predict optimum turbidity, high heating value, viscosity, density, and yield in the waste cooking oil based production of biodiesel. 1 wt % dosage of catalyst resulted in higher heating value, and lower viscosity, density, and turbidity.

Fahmi et al. [54] used ANN to predict virgin oil and methanol-based biodiesel yield. An optimization model was used to realize minimum net present sink in synthesis of biodiesel. The models for unit operation, thermodynamics, and mixing were replaced by the surrogate models to reduce computational load. Betiku et al. [55] used CCD, ANN, and RSM to realize high biodiesel yield from neem seed oil. Optimization of transesterification of pretreated NO using KOH as catalyst

resulted in 99 % yield of biodiesel. Zhang et al. [56] used LS-SVM with GA to estimate and optimize biodiesel yield from castor oil. Based on high accuracy in prediction, the use of the LS-SVM model was recommended for efficient prediction in the process. Optimum values of catalyst weight was 13g, MOR at 625, the temperature at 4060C, and time of 1240 min.

#### Estimation and optimization of process conditions and efficiency

Karimi et al. [57] used RSM and ANN to estimate FAME content and exergetic efficiency in waste cooking oil-based production of biodiesel. The method performed well in achieving high quality and exergetic efficiency of the process by optimizing the input variables. Reaction time, immobilized lipase, concentrations of water, and concentrations of methanol were the design variables. The catalyst concentration of 35%, the water content of 12%, methanol to WCO molar ratio of 6.7, and reaction time of 20h achieved FAME content and exergy efficiency of 86% and 80.1%, respectively. Aghbashlo et al. [58] used ANFIS with GA and linear interdependent fuzzy multi-objective optimization to predict functional exergy efficiency (FEE), normalized exergy destruction (NED), universal exergy efficiency (UEE), and conversion efficiency (CE) in production of biodiesel. Optimum values for transesterification temperature, residence time, and methanol-to-oil molar ratio were found to be 60 °C, 10 min, and 6.20, respectively. Patle et al. [59] used multi-objective GA optimization to estimate heat duty, profit, and organic waste in the palm waste cooking oil based biodiesel production. Waste cooking oil flow rate was the factor affecting heat duty, profit, and organic waste. Shukri et al. [60] used ANN to predict pressure in-cylinder in a bar, heat release in percentage, volume generated, and thermal efficiency in percentage in palm oil methyl ester blends-based biodiesel production. B10 was found to be more efficient due to the high heating value and cetane number. Sarve et al. [61] used ANN to estimate ethanol-to-oil molar ratio, temperature, reaction time, and initial CO<sub>2</sub> pressure in mahua oil based production of biodiesel. Sensitivity analysis of the ANN model was performed where the temperature was found the most effective variable followed by reaction time, ethanol-to-oil molar ratio, and initial CO<sub>2</sub> pressure. ANN outperformed the RSM model both in data fitting and prediction. Kuen et al. [62] applied an automatic tune control scheme consisting of Recursive Least Squares (RLS) and Internal Model Control (IMC) integration to get optimized values of the adaptive controller parameters for the biodiesel transesterification reactor. For introduced disturbance of 5 % rise in the reactor temperature and concentration loops from nominal values, in comparison to conventional PID controllers, adaptive controllers' response time was much faster, i.e. 370 s and 380 s less, respectively.

Rouchi et al. [63] used a Multivariate Curve Resolution Alternative Least Square (MCR-ALS) for interpretation and control of the reaction towards the desired route. For the said purpose, the number of components, concentration profiles, spectral, and reaction kinetics were evaluated for Soybean oil-based biodiesel with reagents consisting of methanol and NaOH. The correlation coefficient and standard deviation of residuals were 0.999 92 and 0.007 65, respectively, that showed the usefulness of MCR-ALS. López-Zapata et al. [64] used virtual sensors based on the Extended Kalman Filter to estimate concentrations of triglycerides, monoglycerides, methyl ester, diglycerides, glycerol and alcohol in jatropha oil-based production of biodiesel. The method has the potentials of real-time implementation because it needs only a few measured variables, such as temperature and pH. Nicola et al. [65] used multi-objective GA optimization to realize maximum purity of important compounds and minimum energy requirements in the production of vegetable oil-based biodiesel by two processes. The temperature of the water, reflux ratio, flash temperature, dryer temperature, the mass flow rate of water, and the number of trays were used as inputs of the process models. The specific energy consumptions for process schemes were 2.7 MJ/kg and 1.5 MJ/kg that met the required standards.

Fahmi and Cremaschi [66] used ANN as a surrogate model to identify the superstructure and operating conditions which minimized net present sink in the production of biodiesel. ANN was used as a surrogate model resulted in a less complex model with an efficient representation of the process synthesis. Soltaniet al. [67] used ANN to model nanocrystalline-sized mesoporous zinc oxide (SO<sub>3</sub>H-ZnO) catalyst for the efficient production of biodiesel from palm fatty acid distillate-based

production of biodiesel. The prediction error was in an acceptable range of 2.73 %. Noriega et al. [68] used group interaction parameters to predict Liquid-Liquid Equilibrium (LLE) in the vegetable oil based production of biodiesel. The most influential variable on LLE was the overall mass fraction followed by the length of alcohol chain. Wong et al. [69] used extreme learning machines with Lyapunov analysis to predict the air-to-fuel ratio in the production of biodiesel from biofuel blends. The proposed approach resulted in effectively regulating AFR to the desired level. The control strategy outperformed the engine built-in AFR controller and was highly recommended for dual-injection engines.

### 3.3.2. Biogas

#### Quality estimation

[97] used integrated ANN and Particle Swarm Optimization (PSO) model for realizing a robust control of production system of biogas. The integrated estimation and optimization framework increased production and quality of biogas, and boosted quantity of electricity production at the affiliated wastewater treatment facility. [98] used ANN for simulating and optimizing operating conditions of UASB reactors for biogas generation. It was observed that ANN can efficiently predict the biogas yield from a laboratory-scale upflow anaerobic sludge blanket (UASB) reactors. Asadi et al. [70] used ANN and ANFIS with subtractive clustering, Fuzzy C-Means Clustering (FCMC), grid partition for prediction of biogas production rate from an anaerobic digesters. Based on the results, the ANFIS-FCMC model outperformed the other sets of models. Akkaya et al. [71] used the Multiple Regression Model in the production of biogas from landfill leachates. The proposed method demonstrated sufficient prediction accuracy.

#### Yield estimation

Ghatak and Ghatak [72] used ANN to predict the yield of biogas from cattle dung, sugarcane bagasse, bamboo dust, and sawdust under mesophilic and thermophilic conditions. The capability of ANN modeling significantly reduced the processing time required for control of the process. Nair et al. [73] used ANN to evaluate biogas yield in an Anaerobic Bioreactor from the organic fraction of municipal solid consisting of vegetable waste, food waste, and yard trimming. It was inferred that an optimized CH<sub>4</sub> recovery can be realized at pH range between 6.6 and 7.1 with Total Volatile Solids (TVS) from 77 to 84 %. Antwi et al. [74] and [75] used different training algorithms for ANN models to estimate biogas and methane yield from chemical, industrial sludges of paper, automobile, petrochemical, and food industries. The conjugate gradient backpropagation and the Quasi-Newton method were the best among eleven training algorithms.

Ihunegbo et al. [76] used PLS to predict the yield of biogas from bioslurry. The results helped in establishing the fact that the acoustic chemometrics is a reliable PAT approach to monitor Total Solids (TS) in complex bioslurry and the same concept can be extended to other biomass processing industries as well. Qdais et al. [77] used ANN to predict biogas yield in the waste digester. The ANN model was effective in capturing the important features of the variables involved in biogas digester operation.

#### Optimization of quality and yield

[78] integrated ANN model with the GA for optimizing operational parameters that resulted in 6.9% increase in methane yield. [79] achieved the best performance of Upflow Anaerobic Contactor (UAC) with 87% COD removal, and hydraulic retention time of 16.67 days where 7.4% increase in biogas production. Barik and Murugan [80] used ANN and GA to estimate and optimize the yield of biogas from cattle dung and seed cake of Karanja in co-digestion. The product quality through the use of co-digestion of cake of Karanja and cattle dung mixture was higher than that of cattle dung samples for a mixing ratio of 1 cake of Karanja to 3 cattle dung. Oloko-Oba et al. [81] used ANN integrated with

GA to predict biogas production from cow dung, poultry droppings, and piggery waste. The optimal amount of poultry droppings, cow dung, plantain peel, and piggery waste were 0.7 kg, 0.0004 kg, 0.29 kg, and 0.61 kg, respectively. Zareei and Khodaei [82] used the ANFIS to estimate and optimize the production of anaerobic digestion-based biogas from cow manure and maize straw. The ANFIS model helped in optimizing the process conditions that resulted in 8 % rise in production. Kana et al. [83] used ANN and GA to predict the optimum combination of rice bran, paper waste, banana stem, sawdust, and concentration of cow dung that enhanced the yield of biogas.

### 3.3.3. Biohydrogen

[84] used the ANN model for the prediction of hydrogen production from different substrates. The initial pH, temperature, initial substrate, biomass concentrations, and time were used as inputs of the model. The ANN exhibited high capability in capturing the correlation among parameters and the process output. [85] predicted hydrogen yield through ANN and RSM. It was observed that ANN has greater accuracy than RSM. Ren et al. [86] used gray and ANN model for prediction of biohydrogen yield from feedstocks comprising of agricultural residues, paper wastes, and wood chips. The gray box model outperformed the ANN model in predicting the output in the context of uncertain data. Prakasham et al. [87] integrated ANN with GA for the prediction of biohydrogen yield from mixed anaerobic microbial consortia. The optimization strategy resulted in 16 % biohydrogen yield. Aghbashlo et al. [88] used a novel hybrid fuzzy clustering-ranking method with ANN to predict exergetic efficiencies in the production of hydrogen from photo-fermentation. Optimum values of flow rate of syngas and agitation speed were 13.68 ml/min and 348.62 rpm, respectively.

### 3.3.4. Miscellaneous (bioethanol, bisabolene)

Ezzatzadegan et al. [90] used Fuzzy Neural Network (FNN) and PSO to predict the yield of bioethanol from corn stover. The optimum fermentation time and required temperature were 69.39 h and 34.5 °C, respectively. Del et al. [91] used ANN-based multi-objective optimization with a hybrid stochastic search optimization in bisabolene production from microalgal biofuel.

## 3.4. Consumption, engine performance and emissions

ML application in consumption, engine performance, and emissions are mostly performed simultaneously in the studies reported in the literature. Hence, we combined all these aspects in this section. Dominant ML methods in these studies were ANN, ANFIS, Extreme learning, SVM, and PLS. The studies are predominantly based on biodiesel as a fuel hence classification is carried based on the ML methods as shown in Table 3 and discussed in the subsections 3.4.1 to 3.4.4.

**Table 3.** Summary of ML application in consumption and emission phases

Types of ML Methods	Input variables	Output variables	References
ANN	Biofuel blend (22), engine speed (15), load (11), cetane number (4), output torque (3), density of fuels (3), compression ratios (3), intake air temperature (2), EPS content (2), lower heating value (2), CO <sub>2</sub> , hydrogen flow rates, percent fuel for non-EGR engine, H <sub>2</sub> , brake power, environmental conditions, nano size, rpm, CNG flow rate, specific gravity, average molecular weight, net heat of combustion, kinematic viscosity, time, fuel temperature, C-to-H ratio, engine crank angle, performance of a compression ignition engine, compression ratio, injection timing, CH <sub>4</sub> ratio of the fuel, maximum cylinder pressure, pilot fuel and natural gas consumption, air-to-fuel ratio exhaust emissions, exhaust temperature values, smoke, HC, CO, fuel mass flow rate, injection pressure, and throttle position, Biodiesel volume, fuel type, and NO <sub>x</sub>	Emission characteristics ( NO <sub>x</sub> (14), CO (13), CO <sub>2</sub> (6), HC (5), unburned hydrocarbons UHC (2), HCO <sub>2</sub> , and smoke density (2)), BSFC (8), EGT (7), BTE (6), torque(5), power (4), hourly fuel consumption and specific fuel consumption (3), maximum pressure (P <sub>max</sub> ) (3), exergy efficiency (%) (3), thermal efficiency Performance (2), exergy of exhaust gas (2), fuel exergy (2), exergy transfer to cooling water (2), exergy destruction (2), sustainability index (2), exergy efficiency, maximum heat release rate, exergy transfer to ambient, volumetric efficiency (VE), HRR and cumulative HRR (CuHRR), values of flow rates, engine load, maximum injection pressure, brake power, maximum cylinder gas pressure, indicated mean effective pressure, pressure in cylinder, thermal efficiency in percentage, mechanical efficiency, exhaust gas temperature (EGT), location of maximum pressure, heat release in percentage, volume generated, combustion, noise, vibration level, sound level, mean effective pressure, mechanical efficiency, air-to-fuel ratio, power, engine load, engine speed, BSEC, and engine performance	[60,99–126]
ANFIS	Double bonds (3), blend (3), load (3), the average carbon numbers (2), injection timings, SV, IV, molar weight, speed, acceleration, RPM, VSP, passenger count, MAP, and temperature	Carbon monoxide (CO) (5), BTE (4), NO <sub>x</sub> (4), double bonds (3), blend (3), load (3), the average carbon numbers (2), EGT (2), carbon dioxide (CO <sub>2</sub> ) (2), smoke (2), cetane number (2), ignition delay (ID) (2), combustion delay (CD) (2) and maximum rate of pressure rise (MRPR) (2), injection timings, SV, oxides of nitrogen, molar weight, speed, acceleration, RPM, passenger count, MAP, temperature, biodiesel surface tension, VSP, HC, CO, Particulate Matter (PM), and hydrocarbon (HC)	[127–133]
Extreme learning method	Biodiesel ratio (5), engine speed (5), fuel consumption (5) Engine torque (3), concentrations of the emissions (3), idle air valve normal position (2), fuel injection time(2), ignition advance, throttle position, carbon monoxide, nitrogen oxide, smoke opacity, Brake thermal efficiency, combustion characteristics, exhaust emissions, performance, and the air-to-fuel ratio	Fuel consumption (5), the concentrations of the emissions(3) (HC, CO, CO <sub>2</sub> , NO <sub>x</sub> and PM), brake thermal efficiency , performance, exhaust emissions and combustion characteristics, and air to fuel ratio	[69, 134–139]
Support vector machine (SVM) algorithm and PLS	Composition	Yield	[140]
Nonlinear regression	Injection timing (I), power (P), and blend ratio (B)	BTE, HC, smoke and NO <sub>x</sub> emissions	[141]

### 3.4.1. ANN

ANN application for estimation of compositions of the emissions and performance of biofuels-based engines are categorized based on the feedstock of the biofuels, such as vegetable oil, waste cooking oil, and non-vegetable oil.

Ismail et al. [99] used ANN to predict CO, CO<sub>2</sub>, NO, unburned hydrocarbons, maximum heat release rate, the maximum pressure, location of maximum HRR, location of maximum pressure, and cumulative HRR (CuHRR) of an engine using the soybean and palm oil-based biodiesel. The ANN model demonstrated high prediction capability of engine combustion and emission behavior. Sharon et al. [100] used ANN to predict hydrocarbon (HC), brake thermal efficiency, brake specific fuel consumption, NO<sub>x</sub>, CO, and smoke density of biodiesel produced from vegetable fried oil and non-vegetable fried oil. The prediction accuracy for B15, B30, B60, and B90 was in an acceptable range. Javed et al. [101] used different training structures of ANN to predict BTE, BSFC, O<sub>2</sub>, CO, NO<sub>x</sub>, HC, CO<sub>2</sub>, and EGT of biodiesel produced from Jatropha Methyl Ester. Levenberge Marquardt as a training algorithm with 16 neurons had the best prediction performance. Canakci et al. [102] used ANN to predict emissions, flow rates, engine load, maximum injection pressure, thermal efficiency, and maximum cylinder gas pressure. The ANN performed well in terms of prediction of the output except emissions, such as NO<sub>x</sub>, CO, and UHC where mean error in prediction was higher. Ouguz et al. [103] used ANN to estimate hourly fuel consumption, power, moment, and specific fuel consumption of biodiesel. The ANN model was found suitable for the prediction of engine performance. Barma et al. [142] used ANN to predict mechanical efficiency, mean effective pressure, air-to-fuel ratio, fuel consumption, and torque of engine consuming biodiesel engine. The BPANN had adequate prediction accuracy for the different fuel blends. Celebi et al. [104] used ANN to estimate the noise and vibration level of biodiesel produced from a blend of sunflower, conventional diesel, and canola biodiesel. The ANN model outperformed the linear regression model.

Javed et al. [105] used ANN model for prediction of noise of engine operating on biodiesel with hydrogen dual-fueled zinc-oxide nanoparticle blend. The least noise was noticed for the H<sub>2</sub> flow rate of 1.51/min. Aydogan et al. [106] used ANN for estimation of engine performance, the engine torque, engine power, the specific fuel consumption (SFC), and EGT of engine using cotton and rapeseed oils biodiesel. The ANN accurately predicted the engine performance, the engine torque, SFC, and EGT. Shojaeefard et al. [107] used ANN to estimate performance and exhaust emissions, BSFC, brake power, and exhaust emissions of DI engine working on biodiesel developed from castor oil. The ANN performance was compared with a group method of data handling. The ANN model was better in terms of prediction accuracy but the group method of data handling models was superior in terms of simplicity. Abhishek et al. [108] used ANN for estimation of performance, BSFC, exhaust temperature, and exhaust composition of an engine using Polanga biodiesel. A very highly accurate prediction with a correlation coefficient closed to one was achieved. Omidvarborna et al. [109] used ANN to predict NO<sub>x</sub> emissions and concentration of NO<sub>x</sub> from EGR engines and non-EGR engines using soybean based biodiesel. The application of ANN was recommended for the estimation of NO<sub>x</sub> emissions from both EGR engines and non-EGR engines. Karthickeyan et al. [110] used ANN model for estimation of performance and emissions characteristics from engine using orange oil-based biodiesel. Orange oil Methyl Ester (OME) with the VCR engine demonstrated higher efficiency and lesser fuel consumption. Menon et al. [111] used ANN with GA to optimize emission characteristics and performance of a biodiesel engine. For realizing optimum biodiesel composition, the total saturated methyl ester contents were from 36 to 43 wt % and unsaturated contents were from 55 to 63 wt %, respectively.

Ghobadian et al. [112] used ANN to estimate fuel consumption, torque and emission of engine working on waste cooking oil-based biodiesel. The correlation coefficient and mean squared error (MSE) for torque, SFC, CO, and HC were close to 1 and 0.0004, respectively. Shivakumar et al. [113] used ANN to estimate emission characteristics and performance of a variable compression ratio CI-engine working on waste cooking oil based biodiesel. The mean error values of ANN were less than 8 %, which is acceptable. Muralidharan et al. [114] used ANN to predict emission and performance of

a four-stroke variable compression ratio engine and a single-cylinder using cooking oil-based biodiesel. A good agreement was found between predicted and experimental measurements. Najafi et al. [115] used ANN to predict energy and exergy efficiency, and exhaust temperature in the usage of waste cooking oil-based biodiesel. The ANN was more helpful than the first-principle models. Kanna et al. [116] used ANN for prediction of performance, torque, power, and specific fuel consumption of a biodiesel engine. The optimum values of injection timing and injection pressure were 25.50 bTDC and 280 bar, respectively. Jaliliantabar et al. [117] used ANN for prediction of emissions, load, and speed of an engine working on blend of biodiesel fuel derived from waste cooking oil in diesel. An optimum operation was achieved with the reduction for CO, CO<sub>2</sub>, HC, NO<sub>x</sub> and smoke emissions approximately 47.25 %, 48.23 %, 52.7 %, 94.55 % and 44.29 %, respectively. Aghbashlo et al. [118] used ELM with wavelet transform algorithm, ANN, and GA to determine the exergetic performance of a DI-engine and sustainability index of biodiesel produced from waste oil. Kurtgoz et al. [119] used ANN to predict biogas engine performance, BSFC, thermal efficiency (TE), and volumetric efficiency (VE) of biogas produced from bovine manure. It was concluded that ANN can accurately estimate TE, BSFC, and VE values.

Aydogan et al. [120] used ANN to predict NO<sub>x</sub>, SFC, and maximum cylinder inner pressure caused by the usage of various blends of biodiesel, bioethanol, and diesel. ANN exhibited high prediction accuracy with a correlation coefficient of 0.98. Ilangkumaran et al. [121] used ANN for estimation of the engine performance, HC, CO, CO<sub>2</sub>, NO<sub>x</sub>, BTE, and smoke from engine working on biodiesel from fish oil. The ANN model exhibited high prediction accuracy. Tosun et al. [122] used linear regression and ANN to predict torque and exhaust emissions (CO, NO<sub>x</sub>) of a naturally aspirated diesel engine running on biodiesel-alcohol mixtures. ANN had more accurate results than LR. Dharma et al. [123] used ANN to predict emission characteristics and performance of a single-cylinder DI-engine using mixed biodiesel-diesel fuel blends. The ANN model was able to accurately predict the outputs for different blends of the fuel. Najafi et al. [124] used ANN with GA to estimate exhaust emissions including NO<sub>x</sub>, PM, CO, and UHC of biodiesel blend of glycerol triacetate. With the use of biodiesel and additive, a reduction of emission of NO<sub>x</sub> and CO up to 63% and 42%, respectively, was realized. The PM was also substantially reduced by 27times. Ozgur et al. [125] used ANN to predict CO, CO<sub>2</sub>, NO<sub>x</sub> and NO<sub>2</sub> emissions from engine using soybean oil-based biodiesel. A close agreement was found between the predicted and experimental results.

#### 3.4.2. Neuro fuzzy inference system

ZareNezhad and Aminian [127] used ANFIS to predict surface tension of biodiesel prepared from Soybean, Rapeseed, Palm, and Sunflower. The ANFIS-based framework outperformed the reported work and the surface tensions values for ten different biodiesels. A high correlation was found between the model estimated values and the experimental data. Gopalakrishnan et al. [128] used ANFIS and the Dynamic Evolving NFIS (DENFIS) to predict emission from transit bus using real-world data of NO<sub>x</sub>, HC, CO, CO<sub>2</sub> and PM of biodiesel. The ANFIS outperformed the DENFIS in prediction of emissions. Mostafaei et al. [129] used ANFIS models to predict the cetane number of biodiesels from its FAME composition. The ANFIS models developed by Fuzzy C-Means (FCM) and grid partition FIS techniques had higher final desirability of 0.718, and 0.857, respectively. Sakthivel et al. [130] used Fuzzy logic and GA to predict emission, performance, and combustion parameters of CI-engine working on biodiesel from fish oil. For high engine performance and reduction in emissions, best blends were identified. The exact biodiesel proportion for no-load, 25, 50, 75, and 100 % loads were found out using Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) as 17%, 17%, 18%, 17%, and 20%, respectively. Sakthivel et al. [131] used fuzzy logic to predict BTE, HC, EGT, NO<sub>x</sub>, smoke, CO, CO<sub>2</sub>, Combustion Delay (CD), Ignition Delay (ID), and Maximum Rate of Pressure Rise (MRPR) of biodiesel produced from fish oil. The fuzzy approach had an edge over theoretical and empirical methods in terms of prediction accuracy. Debnath et al. [132] used GA to predict BTE, CO, NO<sub>x</sub> of engine butanol-based biodiesel. Less biodiesel and higher butanol percentage had good impact

on performance and emission. Blend with 10 % butanol, 10 % biodiesel, and 80 % diesel resulted in high heat release rate, cylinder pressure, BTE, and reduced  $\text{NO}_x$ . Ardabili et al. [133] used ANFIS model for estimation of cetane number of biodiesel. It was also investigated that rise in the carbon number of FAMES increases the viscosity, cetane number, and HHV but rise in the number of double bonds cause decrease in viscosity, cetane number, and HHV.

#### 3.4.3. Extreme learning method

Silitonga et al. [134] used ELM for prediction of BSFC and thermal efficiency of an engine running on biodiesel and bioethanol blends. The biodiesel-bioethanol-diesel blends had oxidation stability more than 20h that showed the potential of their commercialization. Silitonga et al. [135] used a Kernel-based ELM (K-ELM) to predict exhaust emissions, performance, and characteristics of combustion of biodiesel. K-ELM demonstrated the high potential of application in prediction and process optimization of biodiesel derived from a variety of feedstocks. Wong et al. [136] used K-ELM for prediction of the fuel consumption and emissions characteristic of engine working on biodiesel. With the K-ELM, cuckoo search (CS) was then used to determine the optimal biodiesel ratio. The CS optimization was compared with experimental results and PSO. In computational time, the CS and PSO were similar. However, in case of PSO, the time for tuning the parameters was more than CS because CS had lesser number of user-defined parameters than PSO. Wong et al. [137] used Bayesian ELM (B-ELM) and metaheuristic optimization to predict fuel consumption, the k-value, and emission characteristics from engine using gasoline and ethanol. The metaheuristic optimization methods was also applied for identifying optimal ECU setup.

Wong et al. [138] used ELM, LS-SVM, and RB-FNN to predict performance and the concentrations of  $\text{NO}_x$ , CO, HC,  $\text{CO}_2$ , and PM in emissions from engine working on biodiesel. The ELM performed better than LS-SVM and RB-FNN. Wong et al. [139] compared the prediction accuracy of SB-ELM with conventional ELM, B-ELM, and ANN to estimate performance parameters of engine, fuel consumption, and the concentrations of CO, HC, and  $\text{CO}_2$  emissions. SB-ELM outperformed the other methods.

#### 3.4.4. Support vector machine and least square methods

Alves et al. [140] used an SVM and PLS to estimate the biodiesel content in fuel blend. A comparison of SVM and PLS models was made, where SVM outperformed the former in terms of accuracy in prediction. Maheshwari et al. [141] used nonlinear regression to estimate the performance and emission characteristics, smoke, BTE, HC, and  $\text{NO}_x$  emissions from engine using biodiesel derived from Karanja. 13 % biodiesel-diesel blend was found optimum for emissions and efficiency. Shamshirband et al. [126] used SVM Wavelet Transform (SVM-WT), SVM-RBF, SVM Firefly Algorithm (SVM-FFA), and SVM based on quantum particle swarm optimization (SVM-QPSO) and ANN to predict exergetic parameter of a diesel engine and exhaust hot gas, exergy transfer rate to the cooling water, fuel exergy rate, and sustainability index of waste oil based biodiesel. The SVMWT approach was more efficient in prediction of exergetic efficiency and identification of best fuel composition.

#### 3.5. Application summary

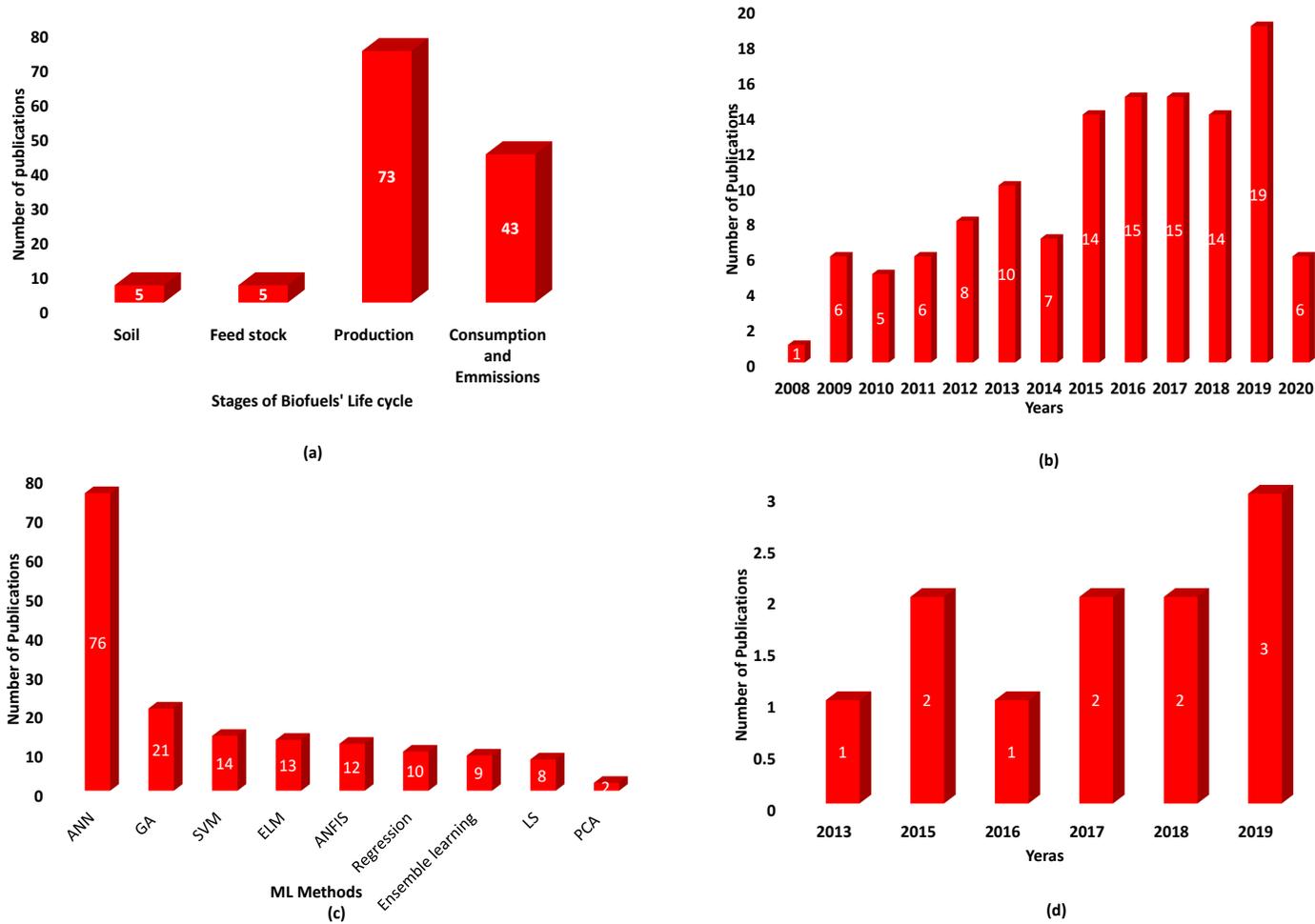
In the soil stage, RF, SVM, and GPM were the commonly used ML methods. The input variable for the ML application were soil characteristics, average precipitation, temperature, solar radiation, and wind speed. The output variable for the ML application was biomass yield and future life cycle environmental impacts. In the feedstock stage, ANN, statistical regression, multiple linear regression, and multiple nonlinear regression models were used. The input variables for the ML application were blend composition, mixing speed, mixing time, temperature. The output variables were viscosity, density, flash point, higher heating values, oxidating stability, cetane number, and methane fraction.

The production phase was further divided with regard to the type of biofuels produced, such as biodiesel, biogas, biohydrogen, etc. The biodiesel category was further divided into four categories based on the nature of the application, such as 1) quality estimation, 2) yield estimation, 3) quality,

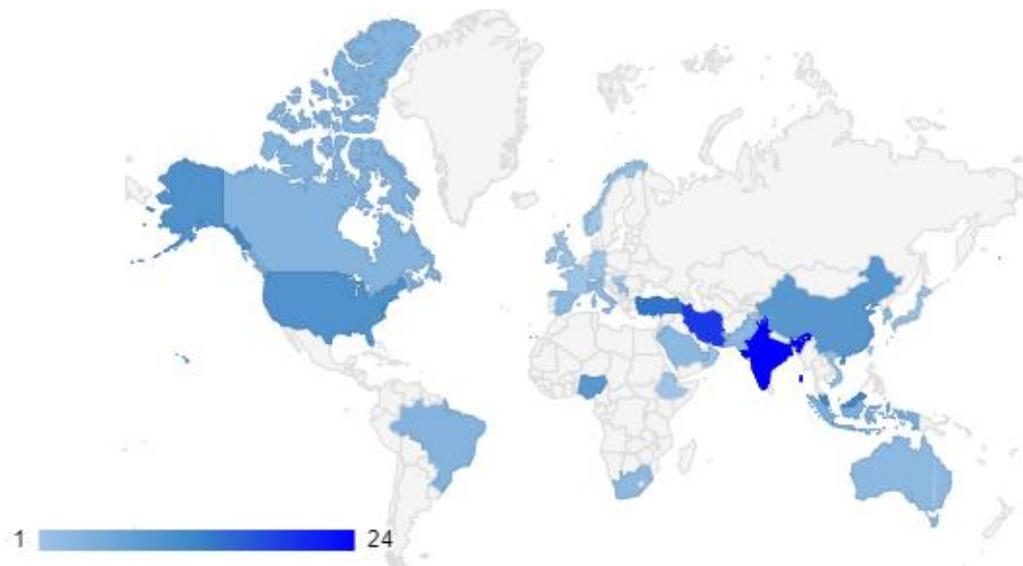
and yield optimization, and 4) estimation and optimization of process conditions, and efficiency. In the quality estimation, the dominant ML method was ANN followed by the regression model. The top five most commonly used input variables were reaction time, reaction temperature, calcination temperature, flow rate, and pressure. The top five most commonly used output variables were composition, quantity, and quality (cetane number), viscosity, density, and FAME content. In the yield estimation, the dominant ML method was ANN followed by ANFIS. The top five most commonly used input variables were catalyst concentration, reaction time, temperature, methanol-to-oil molar ratio, and total volatile fatty acid (VFA) of the effluent. The top five most commonly used output variables were biodiesel yield, % FAME yield, biogas production rate estimation, biohydrogen yield, and biodiesel production. In the quality and yield optimization section, ANN was the dominant ML method followed by GA-based SVM and ANFIS. The top five most commonly used input variables were methanol-to-oil molar ratio, reaction temperature, stirring speed, reaction time, and catalyst concentration. The top five most commonly used output variables were biodiesel yield, final acid value of oil, FAME yield, density, and high heating value. In the process efficiency and optimization section, ANN was the dominant ML method followed by ANFIS. The top five most commonly used input variables were concentration, water content, reaction time, temperature, and methanol-to-oil molar ratio. The top five most commonly used output variables were conversion efficiency (CE), biodiesel yield, FAME content, functional exergy efficiency (FEE), and universal exergy efficiency (UEE). The biogas category was further categorized based on the nature of applications, such as 1) quality estimation, 2) yield estimation, and 3) optimization of quality and yield. In the quality estimation, ANN was the dominant ML method followed by ANFIS and multiple regression models. The topmost commonly used input variables were volatile fatty acids (VFAs), total solids, fixed solids, volatile solids, and pH. The topmost commonly used output variable was biogas composition. In the yield estimation, ANN was the dominant ML method followed by MNLR models and PLS. The top five most commonly used input variables were temperature, pH, TVS, volatile fatty acids (VFAs), and composition. The topmost commonly used output variables were biogas yield, specific methane yield, and amount of dry matter. In the quality and yield optimization section, ANN was the dominant ML method followed by ANFIS. The top five most commonly used input variables were TS, TVS, pH, temperature, and carbon-to-nitrogen ratio. The top five most commonly used output variables were the fraction of methane (% CH<sub>4</sub>), effluent COD, TSS, and VFA concentration and biogas yield. For the biohydrogen, ANN was the dominant ML method. The top five most commonly used input variables were pH, substrate, biomass concentrations, temperature, and time. The topmost commonly used output variables were biohydrogen yield, exergetic outputs, and COD removal. In the miscellaneous category comprised of methane and bioethanol, ANN was the dominant ML method followed by single and multiple linear regressions. The top five most commonly used input variables were fractions of different industrial sludges, mechanical pretreatments, temperature, glucose content, and fermentation time. The top five most commonly used output variables were specific methane yield, biomass concentration, friction velocity, bisabolene production, and bioethanol production.

The consumption, engine performance, and emission cases were studied simultaneously in most of the related papers, hence they were reviewed in a single subsection. Biodiesel was the dominant type of biofuels in this stage so the literature was rather classified based on the type of ML methods, such as ANN, ANFIS, ELM, and SVM. In the ANN application, the top five input variables were biofuel blend, engine speed, load, cetane number, and output torque. The top five output variables were emission characteristics NO<sub>x</sub>, CO, CO<sub>2</sub>, BSFC and temperature. In the ANFIS application, the top five input variables were double bonds, blend, load, average carbon numbers, and temperature. The top five output variables were BTE, NO<sub>x</sub>, CO, smoke, and CO<sub>2</sub>, respectively. In the ELM application, the top five input variables were biodiesel ratio, engine speed, engine torque, fuel injection time, and idle air valve normal position. The top five output variables were fuel consumption, brake thermal efficiency, performance, exhaust emissions, and fuel concentrations. In the SVM and LS methods, the input and output variable were composition and yield, respectively.

The overall trends observed in the ML application in the biofuels' life cycle are shown in Figure 2. The phase-wise application of ML is shown in Figure 2 (a). The number of publications in the subject area is consistently increasing except for the years 2014 and 2018 as shown in Figure 2 (b). The number of publications related to the ML methods is shown in Figure 2 (c) where the use of ANN was reported 76, followed by GA at 21, SVM at 14, ELM at 13, and ANFIS at 12. The emergence of the ELM that belongs to the second generation of ML is shown in Figure 2 (d). Contribution in terms of studies conducted in the subject area is reported from across the globe as shown in Figure 3. The leading country was India with 24 publications followed by Iran, Turkey, Malaysia, and the USA with 19, 14, 12, and 10 publications.



**Figure 2.** (a) Stage-wise share of publications, (b) Year-wise publications in the subject area, (c) ML method-wise publications, and (d) Year-wise publication from the second generation ML methods



**Figure 3.** Region-wise publications related on ML applications in biofuels' life cycle

#### 4. Conclusions and Future Work

In the biofuels' life cycle, out of four stages, the production stage got 58% share of the reported ML applications followed by consumption and emission, soil, and feedstocks stages with 34%, 4%, and 4%, shares, respectively. ANN has been consistently dominant from the first generation of ML methods. Interestingly, GA based optimization was the second-highly reported work after ANN. GA outperformed the conventionally used RSM approach when compared in several studies in realizing optimum process conditions. The ML methods in descending order in terms of application are ANN, GA, SVM, ELM, ANFIS, Regression (linear/non-linear), Ensemble Learning, LS, and PCA. The application of the second generation was reported for the first time in 2013. ELM was the dominant method from the second generation ML with several variations reported every year after 2013. Contribution in terms of studies conducted in the subject area was reported from across the globe, however, India, Iran, Turkey, Malaysia, and the USA collectively got authors' affiliation to 49% of the reported articles.

Although applications of ML in biofuels are found in the whole life cycle, there is no research applying ML to cover integrated supply chains in biofuels involving agriculture production, feedstock management, and quality control, bioprocessing to transform biomass into biofuels and the consumption and related emissions of the biofuels altogether. The challenge is to better determine the interplay over the decision-making among multiple players of multi-commodities of the value chain of biofuels. Making determinations on complex design, operation, and control of today's industry may count on novel capabilities of advanced analytics (AA) in engineering. As an application of AA, to model supply chain resilience (SCR) of transactions, logistics, operations, etc., of such complex representation of SC elements in the industry, machine learning approaches can be used to determine optimizable surrogate models to correlate independent variables (e.g., resistance and recovery of the SC resilience) to the dependent variable SCR. An ML methodology to quantify SCR based on continuous  $x$  and binary  $y$  variables of resistance (avoidance and containment) and recovery (stabilization and return) can consider ad hoc relationships of dependent and independent variables to be part of the SCR predictions in the machine learning method. Such SCR algebraic or analytical formulas obtained in constrained decision regression approach can be used in optimization and control problems to move from traditional independent networks in order to create more flexibility in fulfilling demand through the complementary behavior of heterogeneous resources. Such coupling of multi-layered networks paves the way for optimal resource exchange, efficient decision making, and knowledge

discovery through developing machine learning, control, and optimization techniques for large-scale interdependent decision making.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ALFIMO	Artificial linear interdependent fuzzy multi- objective optimization
AFR	Air-Fuel Ratio
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Interference System
ANN	Artificial Neural Networks
ALS	Alternative Least Square
ASTM	American Society for Testing and Materials
B10	Biodiesel 10 per cent blend
BRT	Boosted Regression Tree
BSFC	brake specific fuel consumption
BTE	Brake thermal efficiency
CCD	Central composite design
CE	Conversion efficiency
CN	Cetane number
COD	Chemical oxygen demand
CS	Cuckoo search
DA	Discriminant analysis
ECU	Engine Control Unit
EGR	Exhaust gas recirculation
EGT	exhaust gas temperature
ELM	Extreme Learning Machine
ERT	Extremely Randomized Trees
EU	Eutrophication
FAME	Fatty Acid Methyl Ester
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FAs	Fatty acids
FCM	Fuzzy C-Means
FCMC	Fuzzy C-Means Clustering
FEE	Functional exergy efficiency
FP	Flash point
GA	Genetic Algorithm
GBD	eXtreme Gradient Boosting- xgbDART
GBL	eXtreme Gradient Boosting-xgbLinear
GBT	eXtreme Gradient Boosting- xgbtree
GEP	Gene Expression Programming
GHG	greenhouse gas
GIP	group interaction parameters
GPM	Gaussian Process Model
HC	Hydrocarbon
HSME	biodiesel yield (methyl esters (HSME))
IAV	Initial acid value of vegetable oil
IMC	Internal Model Control
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 Non-Linear Regression
MO	Mustard oil
MSE	Mean squared error
PU/MU	Mono and poly-unsaturated fatty acids balance
NED	Normalized exergy destruction
OME	Orange oil Methyl Ester

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
RMSE	Root Mean Square Error
RSM	Response Surface Methodology
SC	Subtractive clustering
SFC	Specific fuel consumption
SVM	Support Vector Machines
SVM-FFA	Support Vector Machine Firefly Algorithm
SVM-QPSO	Support vector machine based on quantum particle swarm optimization
SVM-RBF	Support Vector Machine with the Radial Basis Function
SVM-WT	Support Vector Machine Wavelet Transform
SVR	Support Vector Regression
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
TPME	T. peruviana methyl esters
TS	Total Solids
TSS	Total suspended solids
TVS	Total Volatile Solids
UAC	Upflow Anaerobic Contactor
UASB	upflow anaerobic sludge blanket
UEE	Universal exergy efficiency
UHC	Unburned hydrocarbons
VCR	Variable compression ratio
VFA	Total volatile fatty acid

## References

- Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development* **1959**, *3*, 210–229.
- Ge, Z.; Song, Z.; Ding, S.X.; Huang, B. Data mining and analytics in the process industry: The role of machine learning. *Ieee Access* **2017**, *5*, 20590–20616.
- De, R.; Rajan, A.; Govindaraj, K.; Kinage, A.; Ramamurthy, R.K.; Schreder, J.; Peters, C. System and method for industrial process automation controller farm with flexible redundancy schema and dynamic resource management through machine learning, 2019. US Patent 10,416,630.
- Keliris, A.; Salehghaffari, H.; Cairl, B.; Krishnamurthy, P.; Maniatakos, M.; Khorrami, F. Machine learning-based defense against process-aware attacks on industrial control systems. 2016 IEEE International Test Conference (ITC). IEEE, 2016, pp. 1–10.
- Kanawaday, A.; Sane, A. Machine learning for predictive maintenance of industrial machines using IoT sensor data. 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS). IEEE, 2017, pp. 87–90.
- Weichert, D.; Link, P.; Stoll, A.; Rüping, S.; Ihlenfeldt, S.; Wrobel, S. A review of machine learning for the optimization of production processes. *The International Journal of Advanced Manufacturing Technology* **2019**, *104*, 1889–1902.
- Lahdhiri, H.; Said, M.; Abdellafou, K.B.; Taouali, O.; Harkat, M.F. Supervised process monitoring and fault diagnosis based on machine learning methods. *The International Journal of Advanced Manufacturing Technology* **2019**, *102*, 2321–2337.
- Rymarczyk, T.; Kozłowski, E.; Kłosowski, G.; Niderla, K. Logistic Regression for Machine Learning in Process Tomography. *Sensors* **2019**, *19*, 3400.
- Min, Q.; Lu, Y.; Liu, Z.; Su, C.; Wang, B. Machine learning based digital twin framework for production optimization in petrochemical industry. *International Journal of Information Management* **2019**, *49*, 502–519.
- Anifowose, F.A.; Labadin, J.; Abdulraheem, A. Ensemble machine learning: An untapped modeling paradigm for petroleum reservoir characterization. *Journal of Petroleum Science and Engineering* **2017**, *151*, 480–487.
- Oey, T.; Jones, S.; Bullard, J.W.; Sant, G. Machine learning can predict setting behavior and strength evolution of hydrating cement systems. *Journal of the American Ceramic Society* **2020**, *103*, 480–490.

12. Unnikrishnan, S.; Donovan, J.; Macpherson, R.; Tormey, D. Machine Learning for Automated Quality Evaluation in Pharmaceutical Manufacturing of Emulsions. *Journal of Pharmaceutical Innovation* **2019**, pp. 1–12.
13. Guo, S.; Yu, J.; Liu, X.; Wang, C.; Jiang, Q. A predicting model for properties of steel using the industrial big data based on machine learning. *Computational Materials Science* **2019**, *160*, 95–104.
14. Sewsynker-Sukai, Y.; Faloye, F.; Kana, E.B.G. Artificial neural networks: an efficient tool for modelling and optimization of biofuel production (a mini review). *Biotechnology & Biotechnological Equipment* **2017**, *31*, 221–235.
15. Ardabili, S.; Mosavi, A.; Várkonyi-Kóczy, A.R. Systematic review of deep learning and machine learning models in biofuels research. International Conference on Global Research and Education. Springer, 2019, pp. 19–32.
16. Gleason, C.J.; Im, J. Forest biomass estimation from airborne LiDAR data using machine learning approaches. *Remote Sensing of Environment* **2012**, *125*, 80–91.
17. Huntington, T.; Cui, X.; Mishra, U.; Scown, C.D. Machine learning to predict biomass sorghum yields under future climate scenarios. *Biofuels, Bioproducts and Biorefining* **2020**, *14*, 566–577.
18. Habyarimana, E.; Piccard, I.; Catellani, M.; De Franceschi, P.; Dall'Agata, M. 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* **2019**, *9*, 203.
19. Lee, E.K.; Zhang, W.J.; Zhang, X.; Adler, P.R.; Lin, S.; Feingold, B.J.; Khwaja, H.A.; Romeiko, X.X. Projecting life-cycle environmental impacts of corn production in the US Midwest under future climate scenarios using a machine learning approach. *Science of The Total Environment* **2020**, *714*, 136697.
20. Yang, P.; Zhao, Q.; Cai, X. Machine learning based estimation of land productivity in the Contiguous US using biophysical predictors. *Environmental Research Letters* **2020**.
21. Mahanty, B.; Zafar, M.; Park, H.S. Characterization of co-digestion of industrial sludges for biogas production by artificial neural network and statistical regression models. *Environmental technology* **2013**, *34*, 2145–2153.
22. Mairizal, A.Q.; Awad, S.; Priadi, C.R.; Hartono, D.M.; Moersidik, S.S.; Tazerout, M.; Andres, Y. Experimental study on the effects of feedstock on the properties of biodiesel using multiple linear regressions. *Renewable Energy* **2020**, *145*, 375–381.
23. Giwa, S.O.; Adekomaya, S.O.; Adama, K.O.; Mukaila, M.O. Prediction of selected biodiesel fuel properties using artificial neural network. *Frontiers in Energy* **2015**, *9*, 433–445.
24. Tchameni, A.P.; Zhao, L.; Ribeiro, J.X.; Li, T. Predicting the rheological properties of waste vegetable oil biodiesel-modified water-based mud using artificial neural network. *Geosystem Engineering* **2019**, *22*, 101–111.
25. Dahunsi, S. Mechanical pretreatment of lignocelluloses for enhanced biogas production: Methane yield prediction from biomass structural components. *Bioresource technology* **2019**, *280*, 18–26.
26. Soltani, S.; Rashid, U.; Roodbar Shojaei, T.; Nehdi, I.A.; Ibrahim, M. Modeling of the nanocrystalline-sized mesoporous zinc oxide catalyst using an artificial neural network for efficient biodiesel production. *Chemical Engineering Communications* **2019**, *206*, 33–47.
27. Ahmad, I.; Ayub, A.; Ibrahim, U.; Khattak, M.K.; Kano, M. Data-Based Sensing and Stochastic Analysis of Biodiesel Production Process. *Energies* **2019**, *12*, 63.
28. Gulum, M.; Onay, F.K.; Bilgin, A. Evaluation of predictive capabilities of regression models and artificial neural networks for density and viscosity measurements of different biodiesel-diesel-vegetable oil ternary blends. *Environmental and Climate Technologies* **2018**, *22*, 179–205.
29. Tomazzoni, G.; Meira, M.; Quintella, C.M.; Zagonel, G.F.; Costa, B.J.; de Oliveira, P.R.; Pepe, I.M.; da Costa Neto, P.R. Identification of vegetable oil or biodiesel added to diesel using fluorescence spectroscopy and principal component analysis. *Journal of the American Oil Chemists' Society* **2014**, *91*, 215–227.
30. Sarve, A.; Sonawane, S.S.; Varma, M.N. 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* **2015**, *26*, 218–228.

31. Kumar, S.; Jain, S.; Kumar, H. 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* **2019**, *41*, 1285–1295.
32. Banerjee, A.; Varshney, D.; Kumar, S.; Chaudhary, P.; Gupta, V. Biodiesel production from castor oil: ANN modeling and kinetic parameter estimation. *International Journal of Industrial Chemistry* **2017**, *8*, 253–262.
33. Kanat, G.; Saral, A. Estimation of Biogas Production Rate in a Thermophilic UASB Reactor Using Artificial Neural Networks. *Environmental Modeling & Assessment* **2008**, *14*, 607–614.
34. Kumar, S.; Jain, S.; Kumar, H. 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* **2017**, *39*, 2119–2125.
35. Chakraborty, R.; Sahu, H. Intensification of biodiesel production from waste goat tallow using infrared radiation: Process evaluation through response surface methodology and artificial neural network. *Applied Energy* **2014**, *114*, 827 – 836.
36. Pandu, K.; Joseph, S.; Arun, N.; Sundaramoorthy, K. Optimization of biohydrogen production by Enterobacter species using artificial neural network and response surface methodology. *Journal of Renewable and Sustainable Energy* **2013**, *5*.
37. Kumar, S. 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* **2020**, *42*, 1425–1435.
38. Moradi, G.; Dehghani, S.; Khosravian, F.; Arjmandzadeh, A. The optimized operational conditions for biodiesel production from soybean oil and application of artificial neural networks for estimation of the biodiesel yield. *Renewable Energy* **2013**, *50*, 915 – 920.
39. Guo, J.; Baghban, A. Application of ANFIS strategy for prediction of biodiesel production using supercritical methanol. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* **2017**, *39*, 1862–1868.
40. Mostafaei, M.; Javadikia, H.; Naderloo, L. 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* **2016**, *115*, 626 – 636.
41. Maran, J.P.; Priya, B. Comparison of response surface methodology and artificial neural network approach towards efficient ultrasound-assisted biodiesel production from muskmelon oil. *Ultrasonics Sonochemistry* **2015**, *23*, 192–200.
42. Bobadilla, M.; Fernandez, R.; Lostado-Lorza, R.; Somovilla Gómez, F.; Vergara, E. Optimizing Biodiesel Production from Waste Cooking Oil Using Genetic Algorithm-Based Support Vector Machines. *Energies* **2018**, *11*, 2995. doi:10.3390/en11112995.
43. Cheng, M.Y.; Prayogo, D.; Ju, Y.H.; Wu, Y.W.; Sutanto, S. Optimizing mixture properties of biodiesel production using genetic algorithm-based evolutionary support vector machine. *International Journal of Green Energy* **2016**, *13*, 1599–1607.
44. Sivamani, S.; Selvakumar, S.; Rajendran, K.; Muthusamy, S. Artificial neural network–genetic algorithm-based optimization of biodiesel production from Simarouba glauca. *Biofuels* **2019**, *10*, 393–401.
45. Ighose, B.O.; Adeleke, I.A.; Damos, M.; Junaid, H.A.; Okpalaek, K.E.; Betiku, E. 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* **2017**, *132*, 231 – 240. doi:https://doi.org/10.1016/j.enconman.2016.11.030.
46. Dhingra, S.; Dubey, K.K.; Bhushan, G. A polymath approach for the prediction of optimized transesterification process variables of polanga biodiesel. *Journal of the American Oil Chemists' Society* **2014**, *91*, 641–653.
47. Ishola, N.B.; Okeleye, A.A.; Osunleke, A.S.; Betiku, E. Process modeling and optimization of sorrel biodiesel synthesis using barium hydroxide as a base heterogeneous catalyst: appraisal of response surface methodology, neural network and neuro-fuzzy system. *Neural Computing and Applications* **2019**, *31*, 4929–4943.
48. Silitonga, A.S.; Mahlia, T.M.I.; Shamsuddin, A.H.; Ong, H.C.; Milano, J.; Kusumo, F.; Sebayang, A.H.; Dharma, S.; Ibrahim, H.; Husin, H.; Mofijur, M.; Rahman, S.M.A. Optimization of Cerbera manghas

- Biodiesel Production Using Artificial Neural Networks Integrated with Ant Colony Optimization. *Energies* **2019**, *12*.
49. Chakraborty, R.; Das, S.; Pradhan, P.; Mukhopadhyay, P. Prediction of optimal conditions in the methanolysis of mustard oil for biodiesel production using cost-effective mg-solid catalysts. *Industrial & Engineering Chemistry Research* **2014**, *53*, 19681–19689.
  50. Goharimanesh, M.; Lashkaripour, A.; Akbari, A.A. Optimization of biodiesel production using multi-objective genetic algorithm **2016**. *19*, 117–124. doi:10.6180/jase.2016.19.2.02.
  51. Oladipo, A.S.; Ajayi, O.A.; Oladipo, A.A.; Azarmi, S.L.; Nurudeen, Y.; Atta, A.Y.; Ogunyemi, S.S. Magnetic recyclable eggshell-based mesoporous catalyst for biodiesel production from crude neem oil: Process optimization by central composite design and artificial neural network. *Comptes Rendus Chimie* **2018**, *21*, 684 – 695. doi:https://doi.org/10.1016/j.crci.2018.03.011.
  52. Rajendra, M.; Jena, P.C.; Rahman, H. Prediction of optimized pretreatment process parameters for biodiesel production using ANN and GA. *Fuel* **2009**, *88*, 868–875.
  53. Corral Bobadilla, M.; Fernández Martínez, R.; Lostado Lorza, R.; Somovilla Gómez, F.; Vergara González, E.P. Optimizing Biodiesel Production from Waste Cooking Oil Using Genetic Algorithm-Based Support Vector Machines. *Energies* **2018**, *11*, 2995.
  54. Fahmi, I.; Cremaschi, S. Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models. *Computers & Chemical Engineering* **2012**, *46*, 105–123.
  55. Betiku, E.; Omilakin, O.R.; Ajala, S.O.; Okeleye, A.A.; Taiwo, A.E.; Solomon, B.O. Mathematical modeling and process parameters optimization studies by artificial neural network and response surface methodology: A case of non-edible neem (*Azadirachta indica*) seed oil biodiesel synthesis. *Energy* **2014**, *72*, 266–273.
  56. Zhang, Y.; Niu, C. Toward estimation of biodiesel production from castor oil using ANN. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* **2018**, *40*, 1469–1476.
  57. Karimi, M.; Jenkins, B.; Stroeve, P. Multi-objective optimization of transesterification in biodiesel production catalyzed by immobilized lipase: Multi-objective optimization of biodiesel production. *Biofuels, Bioproducts and Biorefining* **2016**, *10*. doi:10.1002/bbb.1706.
  58. Aghbashlo, M.; Hosseinpour, S.; Tabatabaei, M.; Soufiyan, M.M. 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* **2019**, *235*, 100–112.
  59. Patle, D.S.; Sharma, S.; Ahmad, Z.; Rangaiyah, G. Multi-objective optimization of two alkali catalyzed processes for biodiesel from waste cooking oil. *Energy Conversion and Management* **2014**, *85*, 361 – 372.
  60. Shukri, M.R.; Rahman, M.; Ramasamy, D.; Kadirgama, K. ARTIFICIAL NEURAL NETWORK OPTIMIZATION MODELING ON ENGINE PERFORMANCE OF DIESEL ENGINE USING BIODIESEL FUEL. *International Journal of Automotive & Mechanical Engineering* **2015**, *11*.
  61. Sarve, A.N.; Varma, M.N.; Sonawane, S.S. Response surface optimization and artificial neural network modeling of biodiesel production from crude mahua (*Madhuca indica*) oil under supercritical ethanol conditions using CO<sub>2</sub> as co-solvent. *RSC Advances* **2015**, *5*, 69702–69713.
  62. Kuen, H.Y.; Mjalli, F.S.; Koon, Y.H. Recursive Least Squares-Based Adaptive Control of a Biodiesel Transesterification Reactor. *Industrial & Engineering Chemistry Research* **2010**, *49*, 11434–11442.
  63. Rouchi, M.B.; Khorrami, M.K.; Garmarudi, A.B.; de la Guardia, M. 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* **2019**, *213*, 347 – 353.
  64. López-Zapata, B.; Adam Medina, M.; Alvarez Gutierrez, P.; Castillo González, J.; Hernandez-De-Leon, H.; Valdés, L. Virtual Sensors for Biodiesel Production in a Batch Reactor. *Sustainability* **2017**, *9*, 1–11.
  65. Nicola, G.; Moglie, M.; Pacetti, M.; Santori, G. Bioenergy II: Modeling and Multi-Objective Optimization of Different Biodiesel Production Processes. *International Journal of Chemical Reactor Engineering* **2010**, *8*. doi:10.2202/1542-6580.1951.
  66. Fahmi, I.; Cremaschi, S. Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models. *Computers & Chemical Engineering* **2012**, *46*, 105 – 123.

67. Soltani, S.; Rashid, U.; Shojaei, T.R.; Nehdi, I.A.; Ibrahim, M. Modeling of the nanocrystalline-sized mesoporous zinc oxide catalyst using an artificial neural network for efficient biodiesel production. *Chemical Engineering Communications* **2019**, *206*, 33–47.
68. Noriega, M.A.; Narváez, P.C. UNIFAC correlated parameters for liquid-liquid equilibrium prediction of ternary systems related to biodiesel production process. *Fuel* **2019**, *249*, 365–378.
69. Wong, K.I.; Wong, P.K. Adaptive air-fuel ratio control of dual-injection engines under biofuel blends using extreme learning machine. *Energy conversion and management* **2018**, *165*, 66–75.
70. Asadi, M.; Guo, H.; McPhedran, K. Biogas production estimation using data-driven approaches for cold region municipal wastewater anaerobic digestion. *Journal of Environmental Management* **2020**, *253*, 109708. doi:https://doi.org/10.1016/j.jenvman.2019.109708.
71. Akkaya, E.; Demir, A.; Varank, G. Estimation of biogas generation from a uasb reactor via multiple regression model. *International journal of green energy* **2015**, *12*, 185–189.
72. Ghatak, M.D.; Ghatak, A. Artificial neural network model to predict behavior of biogas production curve from mixed lignocellulosic co-substrates. *Fuel* **2018**, *232*, 178 – 189.
73. Nair, V.V.; Dhar, H.; Kumar, S.; Thalla, A.K.; Mukherjee, S.; Wong, J.W. Artificial neural network based modeling to evaluate methane yield from biogas in a laboratory-scale anaerobic bioreactor. *Bioresource Technology* **2016**, *217*, 90 – 99. Special Issue on Bioenergy, Bioproducts and Environmental Sustainability.
74. Antwi, P.; Li, J.; Boadi, P.O.; Meng, J.; Shi, E.; Deng, K.; Bondinuba, F.K. Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network. *Bioresource Technology* **2017**, *228*, 106 – 115. doi:https://doi.org/10.1016/j.biortech.2016.12.045.
75. Mahanty, B.; Zafar, M.; Park, H.s. Characterization of co-digestion of industrial sludges for biogas production by artificial neural network and statistical regression models. *Environmental technology* **2013**, *34*, 2145–53.
76. Ihunegbo, F.N.; Madsen, M.; Esbensen, K.H.; Holm-Nielsen, J.B.; Halstensen, M. Acoustic chemometric prediction of total solids in bioslurry: A full-scale feasibility study for on-line biogas process monitoring. *Chemometrics and Intelligent Laboratory Systems* **2012**, *110*, 135–143.
77. Qdais, H.A.; Hani, K.B.; Shatnawi, N. Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm. *Resources, Conservation and Recycling* **2010**, *54*, 359–363.
78. Qdais, H.A.; Hani, K.B.; Shatnawi, N. Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm. *Resources, Conservation and Recycling* **2010**, *54*, 359 – 363.
79. Dibaba, O.; Lahiri, S.; T'Jonck, S.; Dutta, A. Experimental and Artificial Neural Network Modeling of a Upflow Anaerobic Contactor (UAC) for Biogas Production from Vinasse. *International Journal of Chemical Reactor Engineering* **2016**, *14*.
80. Barik, D.; Murugan, S. An Artificial Neural Network and Genetic Algorithm Optimized Model for Biogas Production from Co-digestion of Seed Cake of Karanja and Cattle Dung. *Waste and Biomass Valorization* **2015**, *6*.
81. Oloko-Oba, M.I.; Taiwo, A.E.; Ajala, S.O.; Solomon, B.O.; Betiku, E. Performance evaluation of three different-shaped bio-digesters for biogas production and optimization by artificial neural network integrated with genetic algorithm. *Sustainable Energy Technologies and Assessments* **2018**, *26*, 116 – 124. doi:https://doi.org/10.1016/j.seta.2017.10.006.
82. Zareei, S.; Khodaei, J. Modeling and optimization of biogas production from cow manure and maize straw using an adaptive neuro-fuzzy inference system. *Renewable Energy* **2017**, *114*, 423 – 427.
83. Kana, E.G.; Oloke, J.; Lateef, A.; Adesiyun, M. Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network and Genetic Algorithm. *Renewable Energy* **2012**, *46*, 276 – 281.
84. Nasr, N.; Hafez, H.; Naggat, M.H.E.; Nakhla, G. Application of artificial neural networks for modeling of biohydrogen production. *International Journal of Hydrogen Energy* **2013**, *38*, 3189 – 3195.
85. Whiteman, J.; Kana, E. Comparative Assessment of the Artificial Neural Network and Response Surface Modelling Efficiencies for Biohydrogen Production on Sugar Cane Molasses. *BioEnergy Research* **2014**, *7*.
86. Ren, J.; Gao, S.; Tan, S.; Dong, L. Prediction of the yield of biohydrogen under scanty data conditions based on GM (1, N). *International journal of hydrogen energy* **2013**, *38*, 13198–13203.

87. Prakasham, R.; Sathish, T.; Brahmaiah, P. Imperative role of neural networks coupled genetic algorithm on optimization of biohydrogen yield. *International Journal of Hydrogen Energy* **2011**, *36*, 4332 – 4339. Emerging Materials Technology: Materials in Clean Power System.
88. Aghbashlo, M.; Shamshirband, S.; Tabatabaei, M.; Yee, P.L.; Larimi, Y.N. The use of ELM-WT (extreme learning machine with wavelet transform algorithm) to predict exergetic performance of a DI diesel engine running on diesel/biodiesel blends containing polymer waste. *Energy* **2016**, *94*, 443 – 456.
89. Akhbari, A.; Ibrahim, S.; Zinatizadeh, A.A.; Bonakdari, H.; Ebtehaj, I.; S. Khozani, Z.; Vafaeifard, M.; Gharabaghi, B. Evolutionary Prediction of Biohydrogen Production by Dark Fermentation. *CLEAN – Soil, Air, Water* **2019**, *47*, 1700494.
90. Ezzatzadegan, L.; Morad, N.A.; Yusof, R. PREDICTION AND OPTIMIZATION OF ETHANOL CONCENTRATION IN BIOFUEL PRODUCTION USING FUZZY NEURAL NETWORK. *Jurnal Teknologi* **2016**, *78*.
91. del Rio-Chanona, E.A.; Wagner, J.L.; Ali, H.; Fiorelli, F.; Zhang, D.; Hellgardt, K. Deep learning-based surrogate modeling and optimization for microalgal biofuel production and photobioreactor design. *AIChE Journal* **2019**, *65*, 915–923.
92. Gülüm, M.; Bilgin, A. Density, flash point and heating value variations of corn oil biodiesel–diesel fuel blends. *Fuel Processing Technology* **2015**, *134*, 456–464.
93. Gülüm, M.; Bilgin, A. Measurements and empirical correlations in predicting biodiesel-diesel blends' viscosity and density. *Fuel* **2017**, *199*, 567–577.
94. Gülüm, M.; Bilgin, A. A comprehensive study on measurement and prediction of viscosity of biodiesel-diesel-alcohol ternary blends. *Energy* **2018**, *148*, 341–361.
95. Nogueira Jr, C.A.; Carmo, F.R.; Santiago, D.F.; Nogueira, V.M.; Fernandes, F.A.; Aguiar, R.S.; de Sant'Ana, H.B. Viscosities and densities of ternary blends of diesel+ soybean biodiesel+ soybean oil. *Journal of Chemical & Engineering Data* **2012**, *57*, 3233–3241.
96. Baroutian, S.; Aroua, M.K.; Raman, A.A.; Sulaiman, N.M. Viscosities and densities of binary and ternary blends of palm oil+ palm biodiesel+ diesel fuel at different temperatures. *Journal of Chemical & Engineering Data* **2010**, *55*, 504–507.
97. Akbaş, H.; Bilgen, B.; Turhan, A.M. An integrated prediction and optimization model of biogas production system at a wastewater treatment facility. *Bioresource technology* **2015**, *196*, 566–576.
98. Tufaner, F.; Avşar, Y.; Gönüllü, M.T. Modeling of biogas production from cattle manure with co-digestion of different organic wastes using an artificial neural network. *Clean Technologies and Environmental Policy* **2017**, *19*, 2255–2264.
99. Ismail, H.M.; Ng, H.K.; Queck, C.W.; Gan, S. Artificial neural networks modelling of engine-out responses for a light-duty diesel engine fuelled with biodiesel blends. *Applied Energy* **2012**, *92*, 769 – 777.
100. Sharon, H.; Jayaprakash, R.; Sundaresan, A.; Karuppasamy, K.; others. Biodiesel production and prediction of engine performance using SIMULINK model of trained neural network. *Fuel* **2012**, *99*, 197–203.
101. Javed, S.; Murthy, Y.S.; Baig, R.U.; Rao, D.P. Development of ANN model for prediction of performance and emission characteristics of hydrogen dual fueled diesel engine with *Jatropha Methyl Ester* biodiesel blends. *Journal of Natural Gas Science and Engineering* **2015**, *26*, 549 – 557.
102. Canakci, M.; Ozsezen, A.N.; Arcaklioglu, E.; Erdil, A. Prediction of performance and exhaust emissions of a diesel engine fueled with biodiesel produced from waste frying palm oil. *Expert systems with Applications* **2009**, *36*, 9268–9280.
103. Oğuz, H.; Saritas, I.; Baydan, H.E. Prediction of diesel engine performance using biofuels with artificial neural network. *Expert Systems with Applications* **2010**, *37*, 6579–6586.
104. Çelebi, K.; Uludamar, E.; Tosun, E.; Yıldızhan, Ş.; Aydın, K.; Özcanlı, M. Experimental and artificial neural network approach of noise and vibration characteristic of an unmodified diesel engine fuelled with conventional diesel, and biodiesel blends with natural gas addition. *Fuel* **2017**, *197*, 159–173.
105. Javed, S.; Baig, R.U.; Murthy, Y.S. Study on noise in a hydrogen dual-fuelled zinc-oxide nanoparticle blended biodiesel engine and the development of an artificial neural network model. *Energy* **2018**, *160*, 774–782.
106. Aydogan, H.; Altun, A.A.; Ozcelik, A.E. Performance analysis of a turbocharged diesel engine using biodiesel with back propagation artificial neural network. *Energy Education Science and Technology Part A* **2011**, *28*, 459–468.

107. Shojaeefard, M.; Etghani, M.; Akbari, M.; Khalkhali, A.; Ghobadian, B. Artificial neural networks based prediction of performance and exhaust emissions in direct injection engine using castor oil biodiesel-diesel blends. *Journal of Renewable and Sustainable Energy* **2012**, *4*, 063130.
108. Sharma, A.; Sahoo, P.K.; Tripathi, R.; Meher, L.C. Artificial neural network-based prediction of performance and emission characteristics of CI engine using polanga as a biodiesel. *International journal of ambient energy* **2016**, *37*, 559–570.
109. Omidvarborna, H.; Kumar, A.; Kim, D. Artificial neural network prediction of NO<sub>x</sub> emissions from EGR and non-EGR engines running on soybean biodiesel fuel (B5) during cold idle mode. *Environmental Progress & Sustainable Energy* **2016**, *35*, 1537–1544.
110. Karthickeyan, V.; Balamurugan, P.; Rohith, G.; Senthil, R. Developing of ANN model for prediction of performance and emission characteristics of VCR engine with orange oil biodiesel blends. *Journal of the Brazilian Society of Mechanical Sciences and Engineering* **2017**, *39*, 2877–2888.
111. Menon, P.R.; Krishnasamy, A. A composition-based model to predict and optimize biodiesel-fuelled engine characteristics using artificial neural networks and genetic algorithms. *Energy & fuels* **2018**, *32*, 11607–11618.
112. Ghobadian, B.; Rahimi, H.; Nikbakht, A.; Najafi, G.; Yusaf, T. Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network. *Renewable energy* **2009**, *34*, 976–982.
113. Pai, P.S.; Rao, B.S.; others. Artificial neural network based prediction of performance and emission characteristics of a variable compression ratio CI engine using WCO as a biodiesel at different injection timings. *Applied Energy* **2011**, *88*, 2344–2354.
114. Muralidharan, K.; Vasudevan, D. Applications of artificial neural networks in prediction of performance, emission and combustion characteristics of variable compression ratio engine fuelled with waste cooking oil biodiesel. *Journal of the Brazilian Society of Mechanical Sciences and Engineering* **2015**, *37*, 915–928.
115. Najafi, B.; Faizollahzadeh Ardabili, S.; Mosavi, A.; Shamshirband, S.; Rabczuk, T. An intelligent artificial neural network-response surface methodology method for accessing the optimum biodiesel and diesel fuel blending conditions in a diesel engine from the viewpoint of exergy and energy analysis. *Energies* **2018**, *11*, 860.
116. Kannan, G.; Balasubramanian, K.; Anand, R. Artificial neural network approach to study the effect of injection pressure and timing on diesel engine performance fueled with biodiesel. *International Journal of Automotive Technology* **2013**, *14*, 507–519.
117. Jaliliantabar, F.; Ghobadian, B.; Najafi, G.; Yusaf, T. Artificial neural network modeling and sensitivity analysis of performance and emissions in a compression ignition engine using biodiesel fuel. *Energies* **2018**, *11*, 2410.
118. Aghbashlo, M.; Shamshirband, S.; Tabatabaei, M.; Yee, L.; Larimi, Y.N. The use of ELM-WT (extreme learning machine with wavelet transform algorithm) to predict exergetic performance of a DI diesel engine running on diesel/biodiesel blends containing polymer waste. *Energy* **2016**, *94*, 443–456.
119. Kurtgoz, Y.; Karagoz, M.; Deniz, E. Biogas engine performance estimation using ANN. *Engineering science and technology, an international journal* **2017**, *20*, 1563–1570.
120. Aydogan, H. Prediction of diesel engine performance, emissions and cylinder pressure obtained using Bioethanol-biodiesel-diesel fuel blends through an artificial neural network. *Journal of Energy in Southern Africa* **2015**, *26*, 74–83.
121. Ilangkumaran, M.; Sakthivel, G.; Nagarajan, G. Artificial neural network approach to predict the engine performance of fish oil biodiesel with diethyl ether using back propagation algorithm. *International Journal of Ambient Energy* **2016**, *37*, 446–455.
122. Tosun, E.; Aydin, K.; Bilgili, M. Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures. *Alexandria Engineering Journal* **2016**, *55*, 3081–3089.
123. Dharma, S.; Hassan, M.H.; Ong, H.C.; Sebayang, A.H.; Silitonga, A.S.; Kusumo, F.; Milano, J. Experimental study and prediction of the performance and exhaust emissions of mixed *Jatropha curcas*-*Ceiba pentandra* biodiesel blends in diesel engine using artificial neural networks. *Journal of cleaner production* **2017**, *164*, 618–633.
124. Najafi, B.; Akbarian, E.; Lashkarpour, S.M.; Aghbashlo, M.; Ghaziaskar, H.S.; Tabatabaei, M. Modeling of a dual fueled diesel engine operated by a novel fuel containing glycerol triacetate additive and biodiesel

- using artificial neural network tuned by genetic algorithm to reduce engine emissions. *Energy* **2019**, *168*, 1128–1137.
125. Ozgur, T.; Tuccar, G.; Ozcanli, M.; Aydin, K. Prediction of emissions of a diesel engine fueled with soybean biodiesel using artificial neural networks. *Energy Education Science and Technology Part A: Energy Science and Research* **2011**, *27*, 301–312.
  126. Shamshirband, S.; Tabatabaei, M.; Aghbashlo, M.; Yee, L.; Petković, D. Support vector machine-based exergetic modelling of a DI diesel engine running on biodiesel–diesel blends containing expanded polystyrene. *Applied Thermal Engineering* **2016**, *94*, 727–747.
  127. ZareNezhad, B.; Aminian, A. Accurate prediction of surface tension of biodiesel fuels at different operating conditions using a neuro-fuzzy model. *Journal of Molecular Liquids* **2015**, *207*, 206 – 210.
  128. Gopalakrishnan, K.; Mudgal, A.; Hallmark, S. Neuro-fuzzy approach to predictive modeling of emissions from biodiesel powered transit buses. *Transport* **2011**, *26*, 344–352.
  129. Mostafaei, M. ANFIS models for prediction of biodiesel fuels cetane number using desirability function. *Fuel* **2018**, *216*, 665–672.
  130. Sakthivel, G.; Sivaraja, C.; Ikua, B.W. Prediction OF CI engine performance, emission and combustion parameters using fish oil as a biodiesel by fuzzy-GA. *Energy* **2019**, *166*, 287–306.
  131. Sakthivel, G. Prediction of CI engine performance, emission and combustion characteristics using fish oil as a biodiesel at different injection timing using fuzzy logic. *Fuel* **2016**, *183*, 214–229.
  132. Debnath, R.; Sastry, G.R.K.; Rai, R.N. An experimental investigation of performance and emission of thumba biodiesel using butanol as an additive in an IDI CI engine and analysis of results using multi-objective fuzzy-based genetic algorithm. *Environmental Science and Pollution Research* **2019**, *26*, 2281–2296.
  133. ardabili, S.; Najafi, B.; Shamshirband Phd, D. Fuzzy Logic Method for the prediction of cetane number using carbon number, double bounds, iodine and saponification values of biodiesel fuels. *Environmental Progress & Sustainable Energy* **2018**.
  134. Silitonga, A.S.; Masjuki, H.H.; Ong, H.C.; Sebayang, A.H.; Dharma, S.; Kusumo, F.; Siswanto, J.; Milano, J.; Daud, K.; Mahlia, T.M.I.; others. Evaluation of the engine performance and exhaust emissions of biodiesel-bioethanol-diesel blends using kernel-based extreme learning machine. *Energy* **2018**, *159*, 1075–1087.
  135. Silitonga, A.S.; Hassan, M.H.; Ong, H.C.; Kusumo, F. palm oil methyl ester blends as biodiesel. *Environmental Science and Pollution Research* **2017**, *24*, 25383–25405.
  136. Wong, P.K.; Wong, K.I.; Vong, C.M.; Cheung, C.S. Modeling and optimization of biodiesel engine performance using kernel-based extreme learning machine and cuckoo search. *Renewable Energy* **2015**, *74*, 640–647.
  137. Wong, K.I.; Wong, P.K. Optimal calibration of variable biofuel blend dual-injection engines using sparse Bayesian extreme learning machine and metaheuristic optimization. *Energy conversion and management* **2017**, *148*, 1170–1178.
  138. Wong, K.I.; Wong, P.K.; Cheung, C.S.; Vong, C.M. Modeling and optimization of biodiesel engine performance using advanced machine learning methods. *Energy* **2013**, *55*, 519–528.
  139. Wong, K.I.; Vong, C.M.; Wong, P.K.; Luo, J. Sparse Bayesian extreme learning machine and its application to biofuel engine performance prediction. *Neurocomputing* **2015**, *149*, 397–404.
  140. Alves, J.C.L.; Poppi, R.J. Biodiesel content determination in diesel fuel blends using near infrared (NIR) spectroscopy and support vector machines (SVM). *Talanta* **2013**, *104*, 155–161.
  141. Maheshwari, N.; Balaji, C.; ; Ramesh, A. A nonlinear regression based multi-objective optimization of parameters based on experimental data from an IC engine fueled with biodiesel blends. *Biomass and bioenergy* **2011**, *35*, 2171–2183.
  142. Barma, S.; Das, B.; Giri, A.; Majumder, S.; Bose, a. Back propagation artificial neural network (BPANN) based performance analysis of diesel engine using biodiesel. *Journal of Renewable and Sustainable Energy* **2013**, *03*, 1–12.