Article

Artificial Neural Network–Guided Optimization of Microwave-Assisted Biodiesel Production Using a Zn–Ce-Loaded Agro-Waste Activated Carbon Bifunctional Catalyst

Mudher M. Ali[1]*, Thaer A. Abdulla[1], Mohammed T. Yaseen[1], Sabah M. Hassan[1], Saad R. Ahmed[2], Omar A. Habib[3], Nalan Türköz Karakullukçu[4]

[1] Chemical Engineering Department, College of Engineering, Tikrit University, Iraq.
[2] Mechanical Engineering Department, College of Engineering, Tikrit University, Iraq.
[3] North Refineries Company (NRC), Baiji, Salahuldeen, Ministry of Oil, Iraq.
[4] Advanced Technology Research and Application Center, Ondokuz Mayis University, Samson, Turkey.

Emails:
Mudher M. Ali: Mr.mudher.m.ali@tu.edu.iq, Thaer A. Abdulla: Adnan.thaer@tu.edu.iq,
Mohammed T. Yaseen: mohammed.t.yaseen42949@st.tu.edu.iq, Sabah M. Hassan: sabahmohamed@tu.edu.iq ,
Saad R. Ahmed: saadramadhan82@tu.edu.iq , Omar A. Habib: omaraljbory44.oh@gmail.com , Nalan Türköz Karakullukçu: nturkoz@omu.edu.tr

Abstract:

Catalyst limits and the intricate, nonlinear interaction of operating parameters make it difficult to produce biodiesel from low-cost, high-free fatty acid feedstocks in an efficient and sustainable manner. By combining artificial intelligence-based process optimization with a bifunctional Zn–Ce-loaded activated carbon catalyst made from agricultural waste, this study seeks to increase the efficiency of biodiesel synthesis. Artificial neural network (ANN) models were developed and assessed using experimental data from 52 runs of the microwave-assisted production of biodiesel from waste cooking oil. The best feedforward neural network was combined with a genetic algorithm (GA) to estimate biodiesel yield and determine ideal operating parameters after a number of ANN topologies were investigated. With a correlation coefficient of roughly 0.996 and a mean square error of roughly 5.7×10^{-2}, the chosen ANN model with three hidden layers showed outstanding prediction ability. With a maximum anticipated biodiesel yield of 98.57%, ANN–GA optimization found ideal parameters of 0.3 weight percent catalyst loading, an oil-to-methanol ratio of 1:7, a reaction duration of 4.3 minutes, and a microwave power of 97.5%. These findings show that coupling ANN-GA optimization with a sustainable bifunctional catalyst offers a precise and economical method for enhancing microwave-assisted biodiesel production, providing a promising route for the effective use of feedstocks obtained from trash.

Keywords:
Artificial Neural Network (ANN); Microwave; Biodiesel; Genetic Algorithm (GA).

Highlights:

- Zn–Ce bifunctional catalyst from agro-waste optimizes sustainable biodiesel yield.
- Microwave-assisted synthesis enables rapid conversion in under 5 minutes.
- Hybrid ANN-GA modeling predicts biodiesel production with 99.6% accuracy.
- Optimized parameters achieve a maximum biodiesel yield of 98.57% from waste oil.
- Deep and shallow ANN architectures compared for nonlinear process optimization.

Corresponding Author*
Mudher M. Ali
Chemical Engineering Department, College of Engineering, Tikrit University, Iraq.
Email: Mr.mudher.m.ali@tu.edu.iq

1.INTRODUCTION

The expansion of transportation networks, industrialization, and population growth have all contributed to a steady rise in the world's energy demand. Fossil fuels continue to supply the majority of this expanding demand, but their limited supply, unstable prices, and negative effects on the environment have made the hunt for sustainable and renewable energy sources more intense. Because they may be easily incorporated into current energy infrastructures, liquid biofuels have drawn a lot of attention among these options [1]. Among the potential alternatives, biodiesel stands out as one of the most promising renewable fuels. Fatty acid methyl esters (FAMEs) make up biodiesel, which is seen as a promising renewable fuel because of its good combustion properties, non-toxicity, sulfur-free composition, and biodegradability. It can be made from a variety of feedstocks, such as animal fats, non-edible oils, and waste cooking oils (WCOs), allowing for waste valorization and cost savings. Biodiesel combustion is compatible with standard compression ignition engines and produces reduced emissions of carbon monoxide, unburned hydrocarbons, and particulate matter when compared to petroleum diesel [2]. Furthermore, biodiesel works with existing compression ignition (CI) engines with little or no modification, making it a drop-in fuel. Another benefit of biodiesel is its feedstock versatility. It may be made from edible oils (soybeans, rapeseed, palm), non-edible oils (Jatropha, Pongamia, castor), animal fats, microalgae, and, most crucially, waste cooking oils (WCOs). The usage of WCOs is particularly appealing since it saves raw material prices (which account for up to 70-80% of biodiesel production costs), minimizes environmental contamination caused by inappropriate oil disposal, and promotes circular economy concepts [3]. Despite its benefits, catalyst-related issues continue to limit the manufacturing of biodiesel on a broad scale. Although homogeneous base catalysts have great activity, they are very sensitive to water and free fatty acids (FFAs), which can result in soap production, lower yield, and complicated downstream separation. Acid catalysts have sluggish reaction rates and corrosion problems, but they can withstand high FFA content [4]. Although many solid catalysts show limited activity under mild circumstances and poor stability in the presence of FFAs and water, heterogeneous catalysts have been designed to circumvent these shortcomings by permitting catalyst recovery and reuse [5]. These restrictions are especially troublesome when processing inexpensive feedstocks like WCOs, which frequently have high FFA levels [6]. To address these restrictions, researchers have resorted to heterogeneous catalysts, which allow for faster recovery, reusability, and less downstream treatment. Metal oxides (CaO, MgO, ZnO), mixed oxides, zeolites, ion-exchange resins, and supported catalysts are among the solid catalysts under investigation. Although heterogeneous catalysts address many of the problems associated with separation and waste treatment, they still face limitations such as catalyst leaching, modest activity at low temperatures, and poor stability when exposed to FFAs and water [7, 8]. By allowing the simultaneous esterification of FFAs and transesterification of triglycerides in a single step, bifunctional catalysts with both basic and acidic active sites provide an efficient solution. Concurrently, the use of activated carbon made from agricultural waste as a catalyst support has drawn interest because of its large surface area, adjustable porosity, chemical stability, affordability, and capacity to promote waste valorization. Such supports can improve catalyst dispersion, stability, and resistance to deactivation when appropriately treated with metal oxides [9]. Along with advancements in catalyst design, sustainable material sourcing has emerged as a key research concern. Activated carbon (AC) generated from agricultural leftovers has received widespread interest as a catalytic support due to its huge surface area, adjustable porosity, chemical stability, and low cost. Rice husk, coconut shells, date palm fronds, maize cobs, sugarcane bagasse, and olive stones have been effectively transformed into activated carbon using pyrolysis and chemical activation [10, 11]. Numerous research have shown the efficacy of agro-waste-derived AC supports in biodiesel catalysis. For instance [7] discovered that palm shell AC supported CaO produced significant biodiesel yields while being reusable. [8]. found that K_2CO_3-impregnated coconut shell AC improved basicity and stability while converting waste oils. Olive stone AC modified with metal oxides accomplished simultaneous esterification and transesterification, allowing for the use of high-FFA oils [10]. These findings demonstrate that agricultural waste not only serves as a sustainable raw material for activated carbon manufacturing but also helps to reduce the cost and environmental impact of biodiesel production. Zinc provides basic active sites that are advantageous for transesterification reactions, while cerium has garnered attention among different metal modifiers because of its capacity to enhance surface acidity, oxygen mobility, and catalyst stability. While Zn- and Ce-based catalysts have shown encouraging results on their own, their combined application on activated carbon produced from agricultural waste as a bifunctional system for the synthesis of biodiesel has hardly been investigated. This gap offers a chance to create a sustainable catalyst that can effectively process high-FFA feedstocks in mild environments. Beyond catalyst development, process optimization is crucial for increasing biodiesel output. Several interdependent elements affect biodiesel synthesis, including reaction time, temperature, catalyst loading, methanol-to-oil ratio, and power input. Conventional optimization techniques, such as Response Surface Methodology (RSM), have been widely used to establish the ideal operating parameters. However, these statistical approaches frequently presume linear or quadratic correlations, which may fail to capture complicated nonlinear interactions. A number of critical operating parameters, such as reaction time, catalyst loading, alcohol-to-oil ratio, and power input, have a significant impact on biodiesel yield in addition to catalyst design. Because of the extreme nonlinearity of these linkages, traditional statistical optimization techniques like response

surface methodology (RSM) are unable to adequately capture intricate interactions. Without the need for predetermined functional correlations, artificial neural networks (ANNs) have become potent data-driven tools that can model nonlinear, multivariable systems. ANNs have mostly been used in biodiesel research to forecast yields or improve operating conditions with traditional catalysts [12, 13]. For example: created an ANN model to forecast biodiesel output from waste frying oil that outperformed experimental findings [14, 15]. examined ANN with RSM and found that ANN made more trustworthy predictions with lower error levels [16]. Used artificial neural networks (ANN) and genetic algorithms (GA) to optimize biodiesel yield, resulting in considerable improvements over conventional techniques [17]. Used artificial neural networks (ANN) to produce microwave-assisted biodiesel and showed that they outperformed RSM. Lately, researchers have been diving into hybrid ANN models that team up with Particle Swarm Optimization (PSO) and Differential Evolution (DE) to boost optimization performance even further. Nevertheless, the majority of current ANN-based biodiesel research does not combine machine learning with catalyst development and instead concentrates only on process modeling. Furthermore, ANN applications are usually restricted to traditional catalytic systems and single network designs. It has not yet been documented how ANN can be used to comprehensively assess various network configurations while concurrently optimizing operating parameters for the synthesis of biodiesel over a Zn−Ce-loaded agro-waste activated carbon bifunctional catalyst. Similarly, sophisticated machine learning techniques are rarely used in research on Ce-modified or agro-waste-derived catalysts to methodically optimize operating conditions. On the other hand, the current work presents a novel, integrated framework that combines a thorough ANN-based modeling and optimization strategy with the creation of a Zn−Ce-loaded agro-waste activated carbon bifunctional catalyst. To the best of the authors' knowledge, this is the first study to (i) produce biodiesel using Ce-based bifunctional catalysts derived from agricultural waste supports and (ii) systematically compare shallow and deep ANN architectures to optimize microwave-assisted biodiesel synthesis. This work goes beyond previous ANN-only or catalyst-only studies and offers a scalable, data-driven method for sustainable biodiesel production by concurrently optimizing catalyst performance and operating parameters.

2.ANN MODEL DEVELOPMENT

The development of ANN model for predicting biodiesel yield (data-driven model) consists of the following steps: Data Collecting, Data Preprocessing and Model Building [1,2].

2.1.Data Collecting

The ANN model was developed using experimental data from previously published research ([18], Table S.1) on the usage of a Zn−Ce-loaded activated carbon bifunctional catalyst to produce biodiesel from waste cooking oil. In order to speed up the esterification and transesterification reactions, microwave power was utilized as the main energy input in a batch microwave-assisted reactor. The feedstock was waste cooking oil, the alcohol was methanol, and the catalyst loading, reaction time, microwave power, and oil-to-methanol ratio were all systematically changed while the reaction temperature and stirring conditions remained constant. Each run's biodiesel yield was measured experimentally, and it served as the output variable for the creation of the ANN model.

2.2.Data Preprocessing

Before developing an ANN model, the experimental dataset was preprocessed to guarantee numerical consistency and steady training. The preprocessing procedures were as follows: (i) original experimental data; (ii) conversion to decimal format; (iii) data randomization; and (iv) data normalization. Figure 1 shows a schematic summary of these steps. Z-score standardization and Min−Max scaling were the two normalizing techniques that were assessed. The choice was based on ANN performance rather than broad recommendations. While Min−Max scaling consistently provided greater correlation coefficients and lower mean square errors, Z-score normalization yielded worse prediction accuracy and higher error for identical ANN topologies. For all subsequent ANN modeling, Min−Max normalization was therefore used. Finally, outliers and most influential variables have been detected using PLS within SIMCA Statistical software [1, 2].
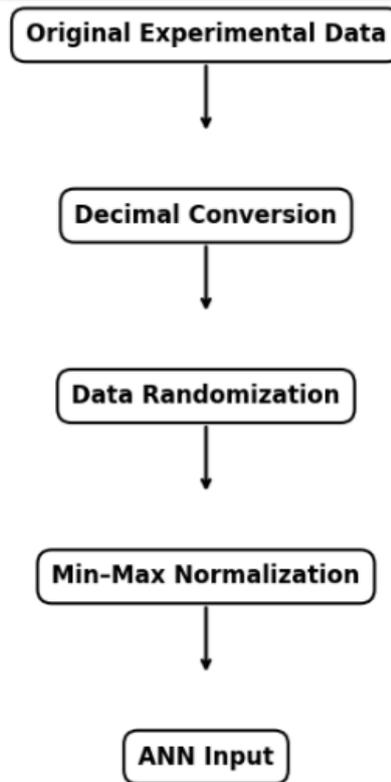
Original Experimental Data

Decimal Conversion

Data Randomization

Min-Max Normalization

ANN Input

Fig. 1 Summery of Steps of Data Preprocessing.

a. Modifying Dataset (dataset in decimal format):
Modified dataset (dataset in decimal format) is shown in Table S.2.

b. Randomizing dataset:
Dataset has been randomized using excel. Randomized dataset is shown in Table S.3.

c. Normalization of Dataset:
Normalization can be accomplished using several methods. The advantages of each method will be evaluated, and the most appropriate one for the data in the research will subsequently be selected [*].

➤ **Min-Max Scaling**
This method scales data to a predetermined range, which is typically (0, 1) or (-1, 1) [19]. It is the optimal solution when the original dataset relationships must be preserved while the data must be confined to a specific range. This method is represented by the following equation: [3]

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

Min-Max scaling is advantageous for algorithms that are susceptible to scale fluctuations, including SVMs [3]. Normalized Dataset using Min-Max Scaling is shown in Table 1.

➤ **Z-score Normalization (Standardization)**
This method rescales data by transforming it to have a mean of 0 and a standard deviation of 1 based on the mean and standard deviation. It is most effective when the data has a Gaussian (normal) distribution [20, 21]. This method is represented by the following equation: [3]

$$X_{norm} = \frac{(X - \mu)}{\sigma} \qquad (2)$$

where μ is the mean and σ is the standard deviation.
Standardization is extensively employed in statistical analysis and machine learning methods such as linear regression, logistic regression, and principal component analysis (PCA) [3].
Normalized dataset using Z-score Standardization is shown in Table 2.

Table 1 Normalized dataset using Min-Max Scaling Method.

ID	Time (min)	Power (%)	Catalyst (%wt)	oil-MEOH Frac	Yield observed (%)
1	0.75	0	0.105263	1	0.6296296
2	0	0.285714	0.105263	1	0.2222222
3	1	0.714286	0.105263	1	1
4	0.375	0.285714	0.105263	0	0.5185185
5	0.5	0	0.105263	1	0.5185185
6	0.5	0.714286	1	0	0.1481481
7	0.25	1	0.105263	0	1
8	1	0.285714	0.105263	1	0.8888889
9	0.5	0.714286	0.105263	1	1
10	0.75	0.714286	0.105263	1	0.8888889
11	0	1	0.105263	1	0.5925926
12	0.375	0.714286	0.105263	1	0.8518519
13	0.75	0.285714	0.105263	1	0.9259259
14	1	0.714286	0.105263	0	0.9259259
15	0.5	0	0.105263	0	0.3703704
16	1	0	0.105263	1	0.7777778
17	0.5	0.285714	0.105263	0	0.5925926
18	0	0	0.105263	1	0
19	0	1	0.105263	0	0.7407407
20	0.5	0.714286	0.473684	1	0.4444444
21	0	0.714286	0.105263	1	0.4444444
22	0.25	0.285714	0.105263	0	0.4074074
23	0	0.714286	0.105263	0	0.4444444
24	0.5	0.714286	0.210526	0	0.5925926
25	0.5	0.714286	0.736842	1	0.2592593
26	0.5	0.714286	0.105263	0	0.9259259
27	0.5	0.714286	0.210526	1	0.7037037
28	0.5	0.285714	0.105263	1	0.8148148
29	0.375	0.714286	0.105263	0	0.6296296
30	0	0.285714	0.105263	0	0.2222222
31	0.5	0.714286	0.105263	0	0.7037037
32	0.5	0.714286	0	0	0.7037037
33	0.5	0.714286	0.473684	0	0.4814815
34	0.375	0	0.105263	0	0.2962963
35	0.75	0.285714	0.105263	0	0.7037037
36	0.75	0	0.105263	0	0.6666667
37	0	0	0.105263	0	0.1481481
38	1	0.285714	0.105263	0	0.8148148
39	0.5	0.714286	0	1	0.6296296
40	0.25	0	0.105263	0	0.1851852
41	0.25	0.285714	0.105263	1	0.6666667
42	0.25	0.714286	0.105263	0	0.5925926
43	0.5	0.714286	1	1	0.1481481
44	0.25	0	0.105263	1	0.1111111
45	0.375	0	0.105263	1	0.4444444
46	0.75	0.714286	0.105263	0	0.7777778
47	0.5	0.714286	0.105263	1	0.9259259
48	1	0	0.105263	0	0.8518519
49	0.375	0.285714	0.105263	1	0.6296296
50	0.5	0.714286	0.736842	0	0.3333333
51	0.25	0.714286	0.105263	1	0.5185185
52	0.25	1	0.105263	1	0.9259259

Table 2 Normalized Dataset Using Z-Score Standardization.

ID	Time (min)	Power (%)	Catalyst (%wt)	oil-MEOH Frac	Yield observed (%)
1	1.001709	-1.42424	-0.33327	0.9903379	0.142778
2	-1.56004	-0.56307	-0.33327	0.9903379	-1.36961
3	1.855625	0.728681	-0.33327	0.9903379	1.517675
4	-0.27916	-0.56307	-0.33327	-0.9903379	-0.26969
5	0.147793	-1.42424	-0.33327	0.9903379	-0.26969
6	0.147793	0.728681	3.758552	-0.9903379	-1.64459
7	-0.70612	1.589849	-0.33327	-0.9903379	1.517675
8	1.855625	-0.56307	-0.33327	0.9903379	1.105206
9	0.147793	0.728681	-0.33327	0.9903379	1.517675
10	1.001709	0.728681	-0.33327	0.9903379	1.105206
11	-1.56004	1.589849	-0.33327	0.9903379	0.005288
12	-0.27916	0.728681	-0.33327	0.9903379	0.967716
13	1.001709	-0.56307	-0.33327	0.9903379	1.242696
14	1.855625	0.728681	-0.33327	-0.9903379	1.242696
15	0.147793	-1.42424	-0.33327	-0.9903379	-0.81965
16	1.855625	-1.42424	-0.33327	0.9903379	0.692737
17	0.147793	-0.56307	-0.33327	-0.9903379	0.005288
18	-1.56004	-1.42424	-0.33327	0.9903379	-2.19455
19	-1.56004	1.589849	-0.33327	-0.9903379	0.555247
20	0.147793	0.728681	1.351597	0.9903379	-0.54467
21	-1.56004	0.728681	-0.33327	0.9903379	-0.54467
22	-0.70612	-0.56307	-0.33327	-0.9903379	-0.68216
23	-1.56004	0.728681	-0.33327	-0.9903379	-0.54467
24	0.147793	0.728681	0.14812	-0.9903379	0.005288
25	0.147793	0.728681	2.555075	0.9903379	-1.23212
26	0.147793	0.728681	-0.33327	-0.9903379	1.242696
27	0.147793	0.728681	0.14812	0.9903379	0.417757
28	0.147793	-0.56307	-0.33327	0.9903379	0.830227
29	-0.27916	0.728681	-0.33327	-0.9903379	0.142778
30	-1.56004	-0.56307	-0.33327	-0.9903379	-1.36961
31	0.147793	0.728681	-0.33327	-0.9903379	0.417757
32	0.147793	0.728681	-0.81466	-0.9903379	0.417757
33	0.147793	0.728681	1.351597	-0.9903379	-0.40718
34	-0.27916	-1.42424	-0.33327	-0.9903379	-1.09463
35	1.001709	-0.56307	-0.33327	-0.9903379	0.417757
36	1.001709	-1.42424	-0.33327	-0.9903379	0.280268
37	-1.56004	-1.42424	-0.33327	-0.9903379	-1.64459
38	1.855625	-0.56307	-0.33327	-0.9903379	0.830227
39	0.147793	0.728681	-0.81466	0.9903379	0.142778
40	-0.70612	-1.42424	-0.33327	-0.9903379	-1.5071
41	-0.70612	-0.56307	-0.33327	0.9903379	0.280268
42	-0.70612	0.728681	-0.33327	-0.9903379	0.005288
43	0.147793	0.728681	3.758552	0.9903379	-1.64459
44	-0.70612	-1.42424	-0.33327	0.9903379	-1.78208
45	-0.27916	-1.42424	-0.33327	0.9903379	-0.54467
46	1.001709	0.728681	-0.33327	-0.9903379	0.692737
47	0.147793	0.728681	-0.33327	0.9903379	1.242696
48	1.855625	-1.42424	-0.33327	-0.9903379	0.967716
49	-0.27916	-0.56307	-0.33327	0.9903379	0.142778
50	0.147793	0.728681	2.555075	-0.9903379	-0.95714
51	-0.70612	0.728681	-0.33327	0.9903379	-0.26969
52	-0.70612	1.589849	-0.33327	0.9903379	1.242696

At first, Z-score Normalization (Standardization) was used, and a group of neurons with no less than 25 neural networks was created. However, the results were not satisfactory, which led to the decision to change the technique to Min-Max Scaling (Normalization). This technique yielded excellent results.

d. Outlier Detection:

Observations that demonstrate a significant divergence from standard data variability, suggest that they arise from an unknown cause, are termed outliers. The reasons of these issues often include inaccurate observation, incorrect recording, erroneous data input, faulty sensors, and transposing results during measurements [4, 5]. The score plot from the outlier detection analysis applied to the experimental dataset is shown in Fig. 2 using SIMCA (Soft Independent Modeling of Class Analogy) statistical software. An individual experimental run projected onto the principal component space is represented by each point in the plot. The absence of anomalous observations and consistent behavior are indicated by the majority of the data points being clustered within the confidence boundary. There are no points outside the statistical threshold or far from the main cluster, indicating that there are no notable outliers in the dataset that could skew the ANN training procedure. This demonstrates that all experimental runs were kept for later modeling and optimization and that aberrant or inaccurate data has no effect on the ANN model's predicted performance. Analyzing this plot helps identify distinct groups of data, reveal outliers that deviate from clusters (Samples that are significantly different from others will appear as points far from the main clusters, indicating potential outliers), and understand the separation between different classes or categories in your data [6]. The result of this analysis is shown in Fig. 2:
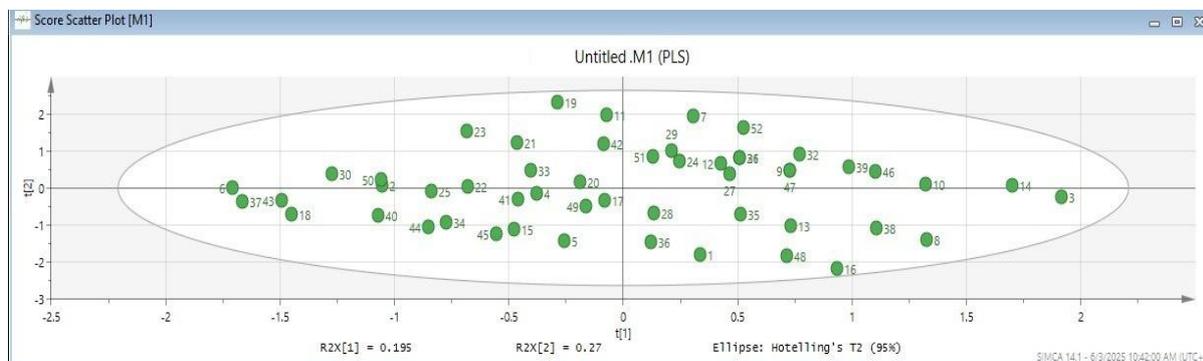
Fig. 2 Score Plot of the First Two Principal Components of a PLS Study for 52 Observations.

According to this analysis, outliers have not been detected with the dataset of this work.

e. Determination the most influential variables:

Thorough measurements of all process variables are necessary for building data-driven models, as are extra measurements for some additional variables. This produces metrics with a strong connection that are data-rich yet information-poor. This characteristic of the data is referred to as "collinearity" [7]. There are three distinct methods by which correlation can appear in linear regression problems: Little Variability, Large Correlation and Exact Linearity [8]. Several methodologies are implemented to identify the variables that exert the greatest influence on the process. The analysis of Variable Importance in Projection (VIP) scores procedure using SIMCA statistical software has been implemented in this research. In SIMCA, Variable Importance in Projection (VIP) scores identify the most influential variables for a Partial Least Squares (PLS) model by measuring a variable's overall contribution to both the X and Y blocks of the data. Variables with VIP scores greater than 1.0 are generally considered influential, with higher scores indicating greater importance [6]. The variable importance in projection (VIP) analysis offers useful advice for process optimization and sheds light on the relative impact of each operating parameter on biodiesel yield. As is frequently advised in multivariate analysis to find factors with above-average contribution to the model response, variables with VIP values greater than 1 were deemed highly influential. Reaction time and microwave power had the highest VIP scores in the current study, suggesting that these factors should be given priority during process optimization since they have a significant impact on biodiesel yield. The oil-to-methanol ratio, on the other hand, displayed a lower VIP value within the examined range, indicating a much smaller impact on yield whenever adequate surplus methanol is supplied. This result suggests that rather than raising the methanol percentage further, optimization efforts should concentrate mostly on energy input and reaction duration. Consistency with ANN–GA optimization results, where parameters with higher VIP values correlated to those identified as crucial for generating maximum biodiesel yield, justified the VIP threshold selection. The result of this analysis is shown in Fig. 3.
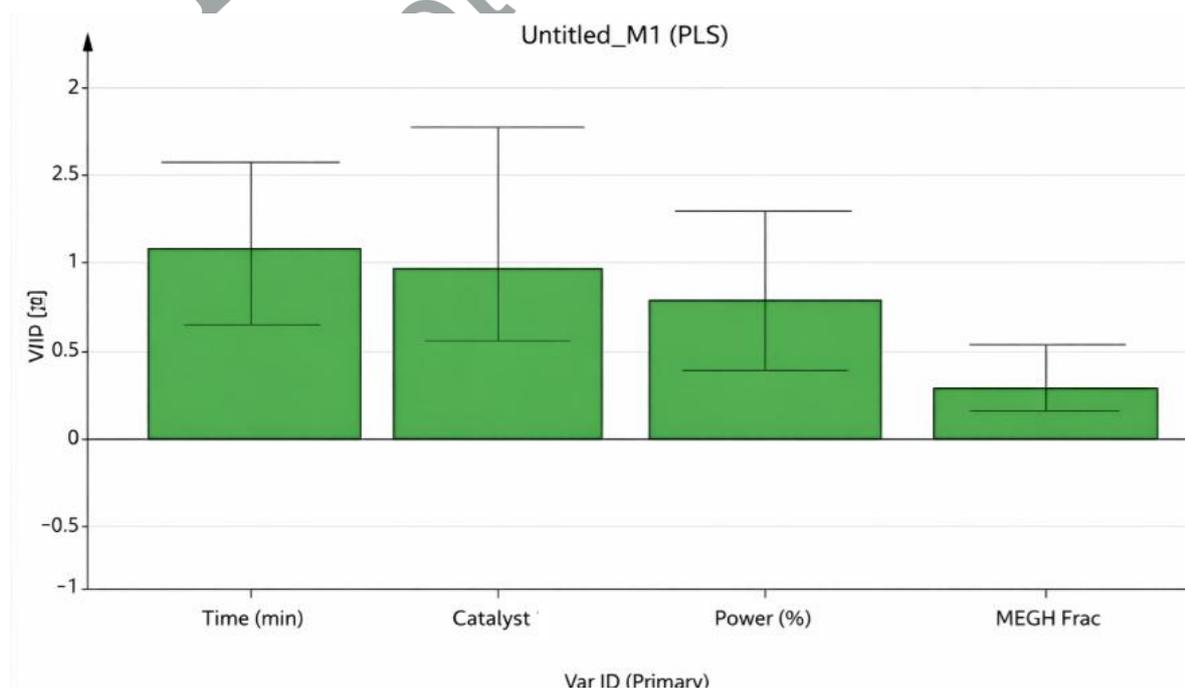

Fig. 3 PLS-VIP of the Input Variables.

According this analysis, time is the most influential variables while the oil-methanol ratio is the least influential variables.

7

3.ANN MODEL BUILDING

In order to estimate and predict biodiesel production as a function of four operating variables—reaction time, catalyst loading, microwave power, and oil-to-methanol ratio—artificial neural networks (ANNs) were created. The focus was on feedforward neural networks (FNNs), which are ideal for nonlinear regression issues with limited data, because of the comparatively short dataset size (52 experimental samples). The normalized preprocessed data has been splitting into three groups (training, testing and validation). Many Neural Networks have been developed (trained, tested and validated) using different configurations (neural network type, no. of layers, no. of neurons in each layer, transfer function of each layer, learning rate, data splitting fraction, and optimization algorithm) to reach the best configuration of neural network. This best configuration is determined based on the performance of neural network [mean square error (MSE), lowest value] and the value of R correlation coefficient (highest R value). The optimum ANN structure that has the minimum root mean square error (RMSE) and highest R value during the ANN building step is selected for developing ANN model for predicting biodiesel yield. ANN models have been developed using MATLAB software. Two methods have been used to create ANN models using MATLAB: first method [Fitting tool (nftool) in nnstart app] and second method (custom codes). First method, nnstart app within MATLAB has been used for developing ANN models. This app consists of four tools: Fitting (nftool), Pattern Recognition (nprtool), Clustering (nctool) and Time Series (ntstool). These tools are intended for training shallow neural networks. Fitting tool (nftool) is used to design neural networks for data fitting or performing regression and curve-fitting operations. It is used to estimate the relationships between inputs and outputs. This type of tools is the type adopted in our current research [10]. Fitting (nftool)) provides the ability to create neural networks of the type Feedforward Neural Network. The network is designed to learn non-linear relationships between inputs and outputs [22]. There is no need to write codes (in many cases); you can enter data and modify network settings through a graphical interface. It includes tools for visualizing data and inputs/outputs. This tool can deal with multiple types of data, such as entering data manually or importing it from external files. This tool provides the ability to easily customize neural network settings, such as the number of nodes, activation functions, number of layers, and network training methods. The tool provides a set of graphs that illustrate inputs, outputs, network performance, error rate, and other graphs. However, this tool does not support all types of neural networks, learning methods, and activation functions (transfer functions). It only supports a specific group of them [10]. Second method, a set of MATLAB codes has been developed to build the artificial network model to utilize additional neural network types, learning methods and activation functions, etc. These developed MATLAB codes have been utilized to build other type of neural networks or Feedforward Neural Network with different configurations that are not exist with first method (nftool).

3.1.Neural Network Type:

In this study, many ANN were built and developed, utilizing four types of Neural Networks, which are:

a) Feedforward Neural Networks (FNN)
b) Radial Basis Function Networks (RBF)
c) Convolutional Neural Networks (CNN)
d) Recurrent Neural Networks (RNN)

But the focus was mainly on the first type (Feedforward Neural Networks (FNN)) because it is the closest in use to our current data.

3.2.Learning Method & Training Optimization Algorithms (Solvers):

In our current study, ANNs were trained and taught using the Supervised Learning method. Unsupervised learning was not used because the data available is clear, fully understood, and does not require exploration, with well-defined, targeted outcomes. This process was carried out using many training optimization algorithms (solvers) supported by MATLAB. The choice of training optimization algorithm (solver) depends on factors such as the network architecture, the size and nature of the dataset, and computational resources. List of Optimization Algorithms (Solvers) that are used to build these neural networks are shown in Tables 3 and 4 [9].

Table 3 Training Optimization Algorithms for (Feedforward, Radial Basis Function, Elman, Layer Recurrent, Time Delay and Distributed Delay Neural Networks).

Training Optimization Algorithms	MATLAB Function
Levenberg-Marquardt	Trainlm
Bayesian Regularization	Trainbr
Scaled Conjugate Gradient	Trainscg
Resilient Backpropagation	Trainrp
Gradient Descent	Traingd
Gradient Descent with Momentum	Traingdm
Variable Learning Rate Gradient Descent	Traingdx
One Step Secant	Trainoss
Polak-Ribiére Conjugate Gradient	Traincgp
BFGS Quasi-Newton	Trainbfg

Table 4 Training Optimization Algorithms for (Deep Learning Neural Networks: Recurrent and Convolutional Neural Networks).

Training Optimization Algorithms (Solvers)	MATLAB Function
Stochastic Gradient Descent with Momentum	Sgdm
Root Mean Square Propagation	Rmsprop
Adaptive Moment Estimation	Adam
Limited-memory BFGS	Lbfgs
Levenberg-Marquardt	Lm

Transfer (Activation) Functions:

Careful selection of activation functions is crucial due to their impact on the neural network performance. Not all functions can work together without specific controls [23]. Activation functions are not universally applicable to all networks and datasets. Some activation functions may work well for a specific network but not for another. Additionally, two activation functions may not be compatible when used together. Therefore, it is important to understand the function of each activation function, its compatibility with the second function, and its suitability for the data environment in which it operates [9]. List of Transfer (Activation) Functions that are used to build these neural networks are shown in Table 5 [10].

Table 5 Transfer (Activation) Functions for (Feedforward, Radial Basis Function, Elman, Layer Recurrent, Time Delay and Distributed Delay Neural Networks).

Name	Input/Output Relation	MATLAB Function
Linear	$a = n$	purelin
Saturating Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$	Satlin
Log-Sigmoid	$a = 1/(1 + e^{\wedge}(-n))$	Logsig
Hyperbolic Tangent Sigmoid	$a = (e^{\wedge}n - e^{\wedge}(-n))/(e^{\wedge}n + e^{\wedge}(-n))$	Tansig
Positive Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$	poslin

Table 6 Transfer (Activation) Functions for (Deep Learning Neural Networks: Recurrent and Convolutional Neural Networks).

Name	Input/Output Relation	MATLAB Function
Linear	$a = n$	Purelin
Log-Sigmoid	$a = 1/(1 + e^{\wedge}(-n))$	Sigmoid
Hyperbolic Tangent Sigmoid	$a = (e^{\wedge}n - e^{\wedge}(-n))/(e^{\wedge}n + e^{\wedge}(-n))$	Tanh
Rectified Linear Unit	$a = max(0,n)$	ReLU

3. RESULTS AND DISCUSSION

In order to estimate biodiesel production as a function of reaction time, catalyst loading, microwave power, and oil-to-methanol ratio, artificial neural networks (ANNs) were created. Feedforward neural networks (FNNs) were prioritized due to the small dataset size (52 experimental runs), whereas deep learning models (CNN and RNN) and radial basis function (RBF) were merely assessed for comparison. The correlation coefficient (R) and mean square error (MSE) were used to evaluate the model's performance on test and validation datasets. Feedforward neural networks (FNNs), radial basis function networks (RBFs), and a few deep learning models (CNNs and RNNs) were among the several ANN architectures that were first assessed. However, it was determined that it was superfluous and possibly distracting to include every tested setting in the main text. As a result, only representative models—one deep learning model for comparison, one RBF network, and the top-performing FNN—are covered here. The Supplementary Material contains comprehensive results for every network that was evaluated. The performance of these example models is summarized in Table 7. With the lowest mean square error (MSE = 5.7×10^{-2}) and the highest overall correlation coefficient (R ≈ 0.996), the FNN model designated as nn30new showed the best prediction accuracy among them. Simpler network topologies are more appropriate for the current dataset size, as demonstrated by the RBF and deep learning models' lower accuracy and higher prediction errors.

Table 7 Comparison of the Performance of Different Models to Prediction the Optimum Parameters of the Biodiesel Production.

Model type	Network ID	Architecture (hidden layers)	Validation/Test R	Validation/Test MSE
FNN (best)	nn30new	13–7–15	0.996	5.7×10^{-4}
RBF	nn215new	40 neurons (radbas)	0.95	7.0×10^{-3}
CNN	nn175new	2 conv layers + FC	0.97	4.4×10^{-3}
RNN	nn155new	4 layers, 10 neurons	0.95	6.7×10^{-3}

The performance of these example models is summarized in Table 7. With the lowest mean square error (MSE = 5.7×10^{-2}) and the highest overall correlation coefficient (R ≈ 0.996), the FNN model designated as nn30new showed the best prediction accuracy among them. Simpler network topologies are more appropriate for the current dataset size, as demonstrated by the RBF and deep learning models' lower accuracy and higher prediction errors. The ANN model's high generalization and the fact that the observed accuracy is not the consequence of training data reuse are demonstrated by the prediction errors shown in Table 8, which only match validation and test datasets. Based on three factors—highest prediction accuracy, consistent performance across training, validation, and test

datasets, and suitable model complexity—the nn30new model was chosen as the best ANN architecture. The network uses a linear activation function in the output layer and nonlinear activation functions in the three hidden layers, which have 13, 7, and 15 neurons, respectively. Beyond this structure, increasing network depth or complexity did not enhance prediction accuracy and, in certain situations, decreased generalization. This finding emphasizes how crucial it is to strike a balance between dataset size and model complexity in order to prevent overfitting.

Table 8 The Prediction Errors for the Cases Tested.

Case	Time (min)	Power (%)	Catalyst (wt%)	Oil:MeOH	Experimental yield (%)	Predicted yield (%)	Absolute error (%)
1	4	30	0.3	1:7	74.0	74.1	0.14
2	8	50	0.3	1:7	96.0	95.8	0.21
3	6	80	0.3	1:7	96.0	97.0	1.04
4	10	50	0.3	1:9	93.0	93.6	0.65
5	8	30	0.3	1:9	89.0	89.2	0.22
6	6	80	0.3	1:9	90.0	93.0	3.33
7	5	30	0.3	1:9	79.0	77.9	1.44
8	6	80	0.5	1:9	87.0	87.0	0.06

Among all developed neural networks, nn30new (FNN neural network type) has the best configurations giving highest performance (lowest MSE=0.0 and highest R=0.99596). This network has been selected to simulate and optimize the operating conditions to obtain the highest yield of biodiesel. The configuration of this network is shown in Fig. 4:
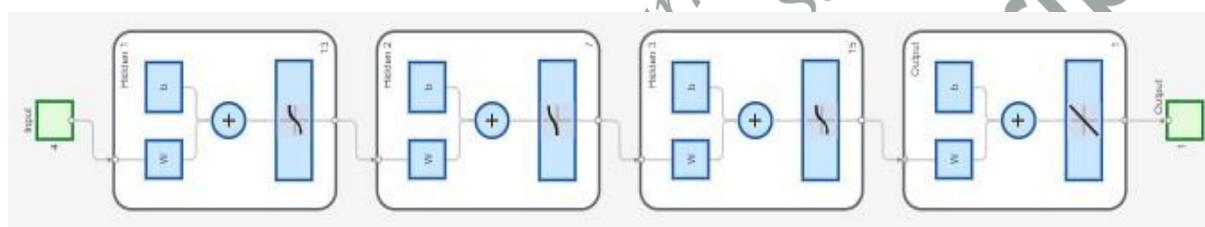

Fig. 4 Neural Network Configuration.

This figure shows that neural network (nn30new) consists of 4 layers (3 hidden layers and 1 output layer). Number of neurons is 13, 7 and 15 in the 1st, 2nd and 3rd hidden layers respectively. Activation functions are the following: tansig, losig, tansig and purelin in the 1st, 2nd, 3rd and 4th layers respectively. The training of this neural network has been made with the following settings:

    a) Data Splitting Ratio (Train 90%, Validation 5% and Test 5%).
    b) Goal=1e-20.
    c) Epochs=2000.
    d) Learning Rate=0.001.

The training results of this network are shown in Fig. 5:

Fig. 5 MATLAB Window of Training Results.

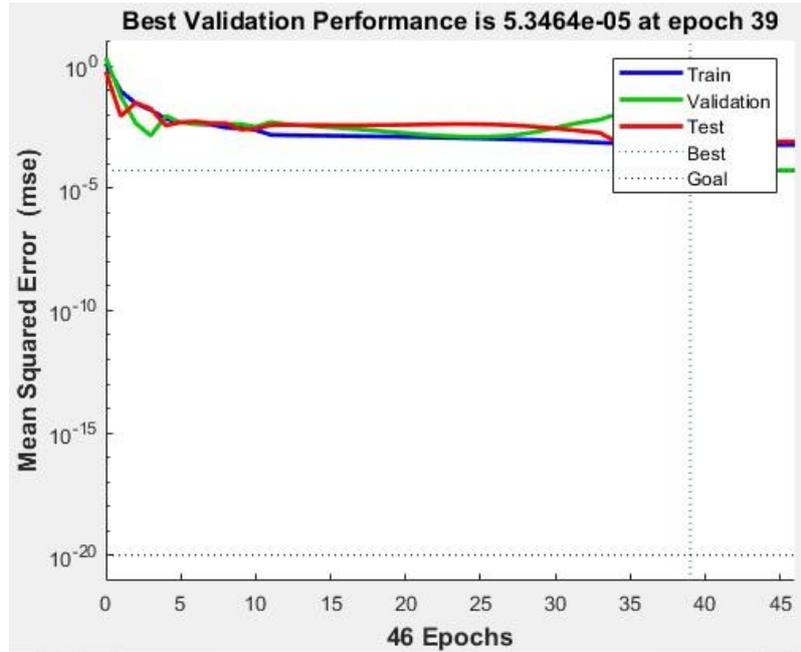The training plots of this network are shown in Fig. 6.

Fig. 6 Performance Plot.

Figure 6 shows the relationship between the number of epochs used for training and (MSE). It can be observed that the value of (MSE) is relatively high when the number of epochs is small (between 0 and 2), and then it starts to stabilize, reaching its lowest value when the number of epochs exceeds (5). The best validation performance (mse=5.3464e-05) has been noticed at epoch 39. Best performance (MSE) of this neural network can be seen in Table 9.

Table 9 Best Training Performance Results of Developed NN.

	Performance (MSE)
Training	5.9641e-04
Validation	5.3464e-05
Test	8.1270e-04
All	5.7756e-04

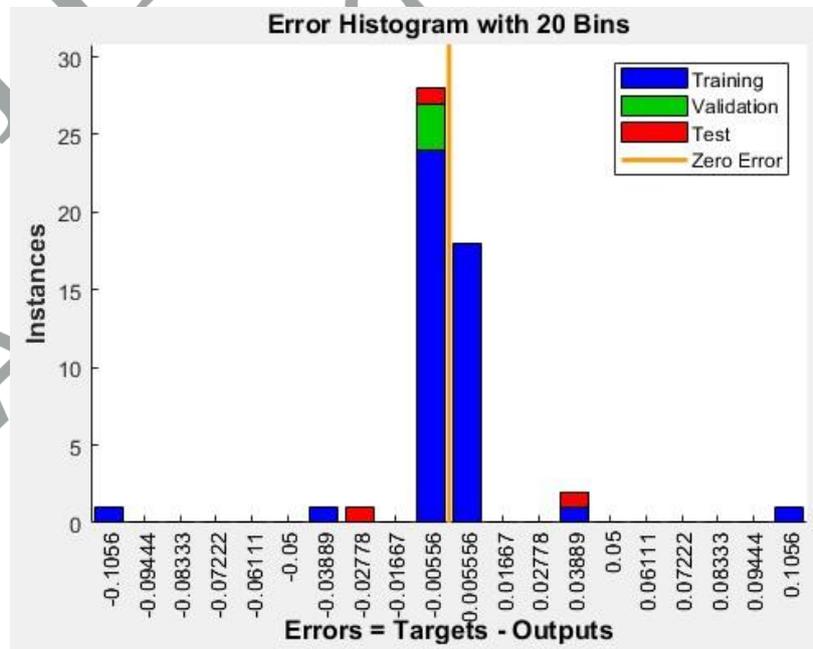Error Histogram Plot can be seen in Fig. 7:

Fig. 7 Error Histogram Plot

Figure 7 shows the error rate for each of the training, validation and test data. The (orange) line is an indicator that the error rate is zero and is considered an excellent indicator. From the figure, it can be observed that the error rate is high at the zero line, which reflects the excellent performance of the ANN.
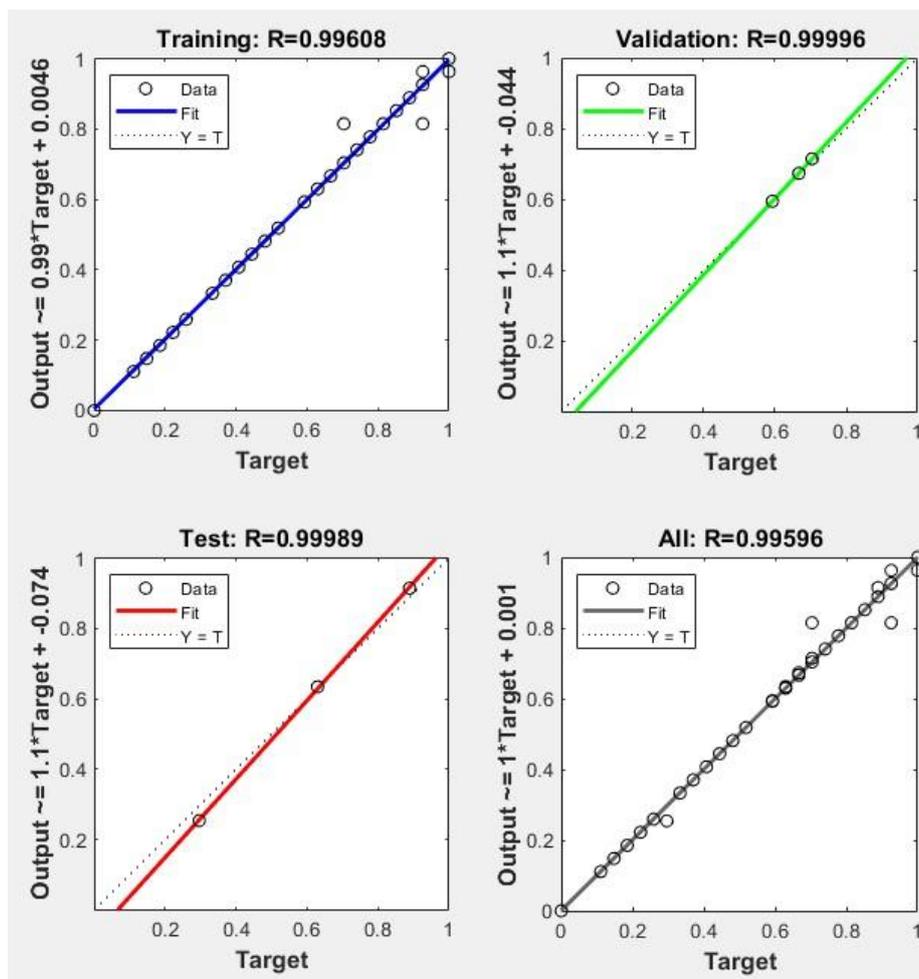
11

Fig. 8 Regression Plot (R-Value).

Figure 8 displays four graphs comparing model outputs to goal values for several data sets: training, validation, test, and all data. Here a breakdown:

1. Training data:
    a) Regression value ((R)): 0.99608
    b) Linear fit equation: output = 0.99 * Target + 0.0046

2. Validation data:
    • Regression value ((R)): 0.99996
    • Linear fit equation: output = 1.1 * Target + 0.044

3. Test data:
    • Regression value ((R)): 0.99989
    ✓ Linear fit equation: output = 1.1 * Target + 0.074

4. All data:
    ✓ Regression value ((R)): 0.99596
    ✓ Linear fit equation: output = 1 * Target + 0.001

Each graphics demonstrates a robust correlation between the model output and the target, signifying an excellent match across all data sets. With an overall correlation coefficient of roughly 0.996 and a mean square error of roughly 5.7×10^{-2}, the FNN architecture labeled as nn30new showed the best prediction ability among the assessed models. The RBF and deep learning models, in contrast, showed greater prediction errors and decreased accuracy. This demonstrates that for small experimental datasets with static input variables, fairly sophisticated FNN architectures are preferable to deep learning models. To find operating parameters that maximize the yield of biodiesel, a genetic algorithm (GA) was combined with the optimized ANN model. A maximum expected biodiesel production of 98.57% was projected by the ANN–GA framework under ideal conditions of 0.3 weight percent catalyst loading, an oil-to-methanol ratio of 1:7, a reaction duration of 4.3 minutes, and a microwave power of 97.5%. Table 10 shows the comparison between observed biodiesel yield (obtained experimentally from the transesterification and esterification of leftover cooking oil using a ZnO.Ce/AC bifunctional catalyst) and simulated biodiesel yield (using developed ANN model, nn30new) under different circumstances (time, power, catalyst wt.% and oil to methanol ratio). The expected and experimental biodiesel yields may be compared using this table. The absolute error percentage between the simulated and experimental findings might be as high as 3% in only 2 runs and it is close to 1 % in other 4 runs while the error percentage is almost zero in all other runs. This Table shows that the best operating conditions

(giving highest yield 98%) using oil to methanol ratio (1:7) are time=10min, power=80%, catalyst wt.%=0.3. Also, this table shows that increasing the methanol content leads to improve the yield slightly. The experimental biodiesel yield and the biodiesel yield derived using the nonlinear regression technique are compared in Fig. 9. These outcomes demonstrate the effectiveness of the plan that is put into practice. The comparison's findings are shown in the figures in the text.

Table 9 Comparison of the Biodiesel Production Yield (Simulated and Observed).

Run	Time (min)	Power (%)	Catalyst (%wt)	OIL-MEOH Ratio	Yield observed (%)	Yield predicted (%)	Percentage error (%)
1	2	30%	0.3	1:7	71	71.000044	0.000062
2	4	30%	0.3	1:7	74	74.000058	0.000078
3	6	30%	0.3	1:7	85	85.000049	0.000058
4	8	30%	0.3	1:7	88	88.122417	0.139110
5	10	30%	0.3	1:7	92	92.000021	0.000023
6	2	50%	0.3	1:7	77	77.000029	0.000038
7	4	50%	0.3	1:7	89	88.999977	0.000025
8	6	50%	0.3	1:7	93	92.999985	0.000017
9	8	50%	0.3	1:7	96	96.000005	0.000005
10	10	50%	0.3	1:7	95	95.684243	0.720255
11	2	80%	0.3	1:7	83	82.999978	0.000027
12	4	80%	0.3	1:7	85	85.000011	0.000013
13	6	80%	0.3	1:7	96	97.000012	1.041679
14	8	80%	0.3	1:7	95	94.999987	0.000013
15	10	80%	0.3	1:7	98	97.999997	0.000003
16	2	100%	0.3	1:7	87	86.999990	0.000012
17	4	100%	0.3	1:7	96	95.999998	0.000002
18	5	30%	0.3	1:7	83	83.000040	0.000048
19	5	50%	0.3	1:7	88	87.999983	0.000019
20	5	80%	0.3	1:7	94	94.000031	0.000033
21	6	80%	0.1	1:7	88	88.000007	0.000008
22	6	80%	0.3	1:7	98	97.000012	1.020396
23	6	80%	0.5	1:7	90	90.000013	0.000014
24	6	80%	1	1:7	83	82.999988	0.000014
25	6	80%	1.5	1:7	78	77.999977	0.000030
26	6	80%	2	1:7	75	74.999964	0.000048
27	2	30%	0.3	1:9	75	75.000009	0.000012
28	4	30%	0.3	1:9	76	76.000046	0.000061
29	6	30%	0.3	1:9	81	81.000102	0.000125
30	8	30%	0.3	1:9	89	89.180911	0.203271
31	10	30%	0.3	1:9	94	93.998548	0.001545
32	2	50%	0.3	1:9	77	77.000001	0.000002
33	4	50%	0.3	1:9	82	82.000015	0.000019
34	6	50%	0.3	1:9	87	87.000005	0.000006
35	8	50%	0.3	1:9	90	90.285987	0.317763
36	10	50%	0.3	1:9	93	92.998775	0.001318
37	2	80%	0.3	1:9	83	82.999982	0.000021
38	4	80%	0.3	1:9	87	86.999995	0.000006
39	6	80%	0.3	1:9	90	92.999989	3.333322
40	8	80%	0.3	1:9	92	91.999654	0.000376
41	10	80%	0.3	1:9	96	95.999556	0.000462
42	2	100%	0.3	1:9	91	90.999987	0.000015
43	4	100%	0.3	1:9	98	98.000002	0.000002
44	5	30%	0.3	1:9	79	77.862366	1.440043
45	5	50%	0.3	1:9	85	85.000009	0.000011
46	5	80%	0.3	1:9	88	88.000010	0.000011
47	6	80%	0.1	1:9	90	89.999951	0.000055
48	6	80%	0.3	1:9	96	92.999989	3.125011
49	6	80%	0.5	1:9	87	87.049078	0.056411
50	6	80%	1	1:9	84	83.999977	0.000027
51	6	80%	1.5	1:9	80	79.999985	0.000019
52	6	80%	2	1:9	75	74.999960	0.000053

13

predicted yield% = 0.9985*(experimental yield%) + 0.1357
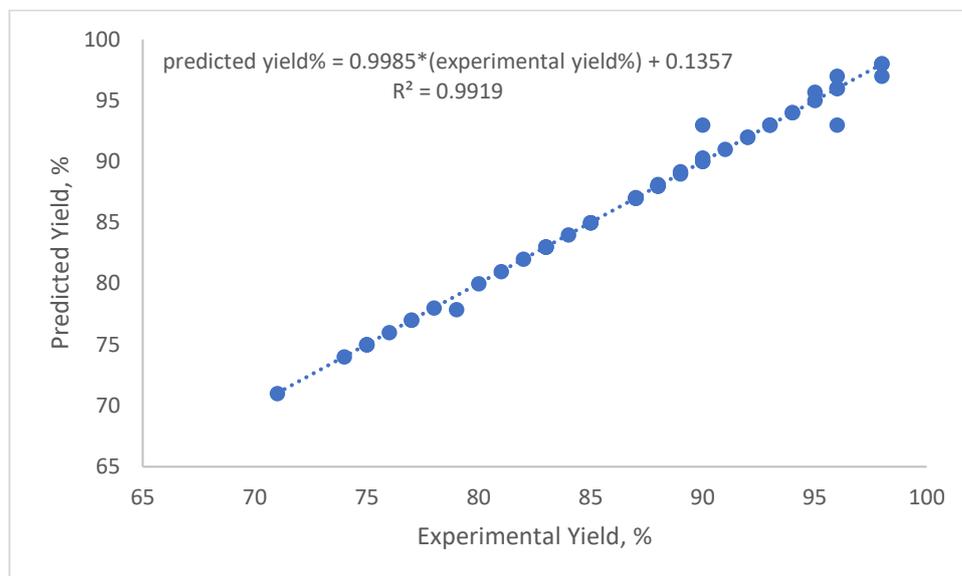R² = 0.9919

Fig. 9 Parity Chart of Biodiesel Yields under Various Reaction Conditions, Experimental and Predicted.

The developed neural network (nn30new) and MATLAB code of genetic algorithm (ga; MATLAB optimization function) have been used to optimize the experimental results of the present work (obtaining the optimal operating conditions of time and power using catalyst 0.3 wt.% and 1:7 oil-methanol ratio to achieve the highest yield of biodiesel). The optimum operating conditions to achieve highest yield of biodiesel (98.57%) are the following: catalyst 0.3 wt%, 1:7 oil-methanol ratio, time 4.3 min and power 97.5%. Also, the developed ANN and MATLAB code of ga have been used also to predict the operating conditions (time and power) to reach specific yield percentages of biodiesel via minimizing the square error between the actual (target) and predicted yield percentages. The predicted operating conditions at different percentages of biodiesel yield are presented in the following Table 10. The reaction mechanism and microwave-assisted process parameters are physically consistent with the ideal operating circumstances that the ANN–GA framework anticipated (0.3 weight percent catalyst, 1:7 oil-to-methanol ratio, 4.3 minutes reaction time, and 97.5% microwave power). For simultaneous esterification and transesterification, a moderate catalyst loading (0.3 weight percent) offers enough active acidic and basic sites without producing excessive viscosity or particle agglomeration, which might impede mass transfer. Diffusion restrictions and decreased efficient use of active sites result from raising catalyst loading above this threshold.

Table 10 Results of Predicted Operation Conditions (Time and Power) Using Catalyst 0.3 wt.% and 1:7 Oil-Methanol Ratio at different Yield Percentages of Biodiesel.

Yield % (Target)	Time (min) (Predicted)	Power (%) (Predicted)
96.65%	9.2	47.2
92.60%	3.5	100
88.55%	4.1	49.7
84.50%	4	39.9
80.45%	4.3	34.1
76.40%	3.1	35.3
72.35%	3.1	32.7

The study's findings show that using ANN-GA optimization in conjunction with a bifunctional Zn−Ce-loaded activated carbon catalyst is a successful strategy for enhancing biodiesel production in microwave-assisted environments. The primary findings are interpreted in terms of model performance, process behavior, and practical consequences in the discussion that follows. The feedforward neural network (nn30new) outperforms RBF, CNN, and RNN models, which emphasizes how crucial it is to choose model complexity that is suitable for the issue structure and dataset size. Deeper designs demonstrated less generalization and did not provide additional predictive potential with only 52 experimental samples and static input variables. This finding is in line with machine learning best practices, which state that when data is scarce, simpler designs frequently perform better than deep models. The chosen ANN successfully captured the nonlinear interactions between operating parameters and biodiesel yield, as seen by the strong correlation coefficient (R = 0.996) and low MSE ($\approx 5.7 \times 10^{-1}$). According to the ANN projections, the most important factors in the range under study are microwave power and reaction time. High microwave power increases volumetric heating, which raises the temperature quickly, improves molecular mobility, and increases mass transfer across the catalyst, methanol, and oil phases. Consequently, compared to traditional heating, reaction kinetics are greatly increased. The effectiveness of microwave-assisted transesterification, where equilibrium can be attained quickly and extended exposure yields diminishing returns, is reflected in the very short optimal reaction time. Another important factor is catalyst loading. A balance between offering enough

14

active acidic and basic sites for concurrent esterification and transesterification and avoiding an excessive catalyst concentration that might raise mixture viscosity, impede mass transfer, or encourage particle agglomeration is suggested by the ideal value of 0.3 weight percent. The observed decrease in yield with larger catalyst loadings can be explained by these factors. The equilibrium of the reaction is influenced by the ratio of oil to methanol. Without producing excessive dilution or making downstream separation more difficult, the determined ideal ratio of 1:7 offers sufficient extra methanol to propel the reversible transesterification reaction toward the synthesis of biodiesel. Beyond this point, increasing the methanol percentage only slightly improves the process, indicating that it becomes kinetically rather than equilibrium constrained. The fast exploration of the multidimensional operational space and the discovery of ideal circumstances that are in line with experimental trends while providing increased precision were made possible by the combination of ANN with a genetic algorithm. The ANN−GA framework found a smaller and more effective operating window with a shorter reaction time (4.3 min) and near-maximum power input in comparison to the empirically measured optimal region (about 0.3 weight percent catalyst, 1:7 ratio, 5–10 min, and 80–100% power). This illustrates how data-driven optimization can improve process conditions beyond trial-and-error experimentation. The ANN−GA approach captures complicated nonlinear interactions among variables, in contrast to traditional optimization techniques like response surface methodology, which presume simpler polynomial correlations. Furthermore, the current work merges ANN−GA optimization with a sustainable bifunctional catalyst system under microwave-assisted conditions, whereas prior ANN-based biodiesel research have mostly concentrated on prediction using traditional catalysts. This combination strategy goes beyond catalyst-only or model-only studies by improving both catalytic performance and process efficiency [24, 25]. Our ANN model demonstrated the efficacy of neural networks for capturing complicated correlations in biodiesel production data, in line with the results [26, 27].

4.CONCLUSIONS

This work showed that an efficient method for improving microwave-assisted biodiesel production from waste cooking oil is to combine ANN-GA optimization with a Zn−Ce-loaded agro-waste-derived activated carbon bifunctional catalyst. With a high correlation coefficient (R = 0.996) and a low mean square error (MSE ≈ 5.7×10^{-2}) on validation and test datasets, a feedforward neural network with three hidden layers (13−7−15 neurons) had the greatest prediction accuracy among the assessed models. With a maximum anticipated biodiesel output of 98.57%, the ANN−GA framework determined ideal operating conditions of 0.3 weight percent catalyst loading, an oil-to-methanol ratio of 1:7, a reaction duration of 4.3 minutes, and a microwave power of 97.5%. In addition to providing a more accurate and effective operating window, these settings are physically consistent with reaction kinetics, mass-transfer behavior, and the increased heating characteristics of microwave irradiation. They also closely match experimentally established optimal trends. The findings demonstrate that data-driven optimization can greatly improve operating conditions beyond experimental trial-and-error methods and that moderately complicated ANN structures perform better than deeper models when applied to short experimental datasets. The very tiny dataset and the limited working range, however, provide constraints to the conclusions. To improve the generalizability and practical usability of the model, more experimental validation with independent datasets and extended settings is advised. This study has some drawbacks, despite its great predictive performance. The ANN model may not be as generalizable outside of the examined operating range because it was developed on a comparatively limited dataset (52 experimental runs) [28]. Therefore, it is advised to confirm the robustness of the model through extra experimental validation utilizing separate datasets and other operating circumstances. To further enhance the applicability of the suggested ANN−GA system, future research should investigate bigger datasets, different machine learning techniques, and longer catalyst stability studies [29]. This accuracy not only streamlined the experimental process but also provided insights into optimal conditions for maximizing biodiesel production, demonstrating the model's practical utility in sustainable energy research.

REFERENCES:

[1] Knothe G. **Biodiesel and Renewable Diesel: A Comparison.** *Progress in Energy and Combustion Science* 2010; **36**(3): 364-373.

[2] Leung DY, Wu X, Leung MKH. **A Review on Biodiesel Production Using Catalyzed Transesterification.** *Applied Energy* 2010; **87**(4): 1083-1095.

[3] Gui MM, Lee KT, Bhatia S. **Feasibility of Edible Oil vs. Non-Edible Oil vs. Waste Edible Oil as Biodiesel Feedstock.** *Energy* 2008; **33**(11): 1646-1653.

[4] Kang J, Zhang L, Wang H, Liu Y, Chen X, Zhao Q. **Vernicia Montana Seed Shell-Derived Solid Acid for Heterogeneous Catalytic Transesterification from Waste Cooking Oil to Biodiesel: An Artificial Neural Network-Guided and Sustainable Approach.** *Molecular Catalysis* 2025; **585**: 115349.

[5] Melero JA, Iglesias J, Morales G. **Heterogeneous Acid Catalysts for Biodiesel Production: Current Status and Future Challenges.** *Green Chemistry* 2009; **11**(9): 1285-1308.

[6] Cordero-Ravelo V, Schallenberg-Rodriguez J. **Biodiesel Production as a Solution to Waste Cooking Oil (WCO) Disposal. Will Any Type of WCO Do for a Transesterification Process? A Quality Assessment.** *Journal of Environmental Management* 2018; **228**: 117-129.

[7] Xie F, Zhang TA, Dreisinger D, Doyle F. **A Critical Review on Solvent Extraction of Rare Earths from Aqueous Solutions.** *Minerals Engineering* 2014; **56**: 10-28.

[8] Nargotra P, Sharma V, Bajaj BK. **Application of Ionic Liquid and Alkali Pretreatment for Enhancing Saccharification of Sunflower Stalk Biomass for Potential Biofuel-Ethanol Production.** *Bioresource Technology* 2018; **267**: 560-568.

[9] Wilcox EG, Levy R, Futrell R, Levy RP. **On the Predictive Power of Neural Language Models for Human Real-Time Comprehension Behavior.** *arXiv Preprint* 2020.

[10] Verma S, Gupta S, Bandhiwal N, Kumar T, Bharadwaj C, Bhatia S. **High-Density Linkage Map Construction and Mapping of Seed Trait QTLs in Chickpea (Cicer Arietinum L.) Using Genotyping-by-Sequencing (GBS).** *Scientific Reports* 2015; **5**(1): 17512.

[11] Palitsakun S, et al., **Transesterification of Jatropha Oil to Biodiesel Using SrO Catalysts Modified with CaO from Waste Eggshell**. *Catalysis Communications* 2021; **149**: 106233.

[12] Abdulla TA, Osman M, Edgar TF. **Neural Network Based Soft Sensor for Pilot-Plant Distillation Column.** *2019 Spring Meeting and 15th Global Congress on Process Safety, American Institute of Chemical Engineers (AIChE)* 2019.

[13] Gülüm M, Bilgin A. **Measurements and Empirical Correlations in Predicting Biodiesel-Diesel Blends' Viscosity and Density.** *Fuel* 2017; **199**: 567-577.

[14] Singh D, Singh B. **Investigating the Impact of Data Normalization on Classification Performance.** *Applied Soft Computing* 2020; **97**: 105524.

[15] Chakraborty T, Hsu A, Manya D, Sheriff G. **Disproportionately Higher Exposure to Urban Heat in Lower-Income Neighborhoods: A Multi-City Perspective.** *Environmental Research Letters* 2019; **14**(10): 105003.

[16] Singh J, Singh J. **COVID-19 and Its Impact on Society.** *Electronic Research Journal of Social Sciences and Humanities* 2020; **2**.

[17] Cavayas YA, Yusuff H, Porter R. **Fungal Infections in Adult Patients on Extracorporeal Life Support.** *Critical Care* 2018; **22**(1): 98.

[18] Akream NS, Abdullah AC, Al-Qasmi RM, Hamad AF, Saleh MA. **High-Yield Activated Carbon Based ZnO-Ce Bifunctional Catalyst for Production of Biodiesel from Waste Cooking Oil.** *Energy Conversion and Management* 2024; **321**: 119054.

[19] Baltagi BH, Baltagi BH. **A Companion to Theoretical Econometrics.** *Wiley Online Library* 2001.

[20] Abusnina A. **Gaussian Process Adaptive Soft Sensors and Their Applications in Inferential Control Systems.** *University of York* 2014.

[21] Webster JG, Eren H. **Measurement, Instrumentation, and Sensors Handbook: Two-Volume Set (Electrical Engineering Handbook).** *CRC Press: Boca Raton, FL, USA* 2014.

[22] Demuth HB, Beale MH, De Jesus O, Hagan MT. **Neural Network Design.** *Martin Hagan* 2014.

[23] Demuth H, Beale M, Hagan M. **Neural Network Toolbox.** *For Use With MATLAB. The MathWorks Inc* 2000; **2000**: 415.

[24] 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.

[25] Yusuff AS, Gbadamosi AO, Popoola LT. **An Artificial Intelligence Approach to Model and Optimize Biodiesel Production from Used Cooking Oil Using CaO Incorporated Zeolite Catalyst.** *Energy Conversion and Management: X* 2023; **20**: 100452.

[26] Arunyanart P, Piyatamrong A, Srichat A, Chollacoop N, Sudaprasert K. **The Prediction of Biodiesel Production Yield from Transesterification of Vegetable Oils with Machine Learning.** *Results in Engineering* 2024; **24**: 103236.

[27] Arunyanart W, et al. **Developmental and Autism Screening: A Survey Across Six States.** *Infants & Young Children* 2012; **25**(3): 175-187.

[28] Yusuff AS, Olutoye MA, Akpan UG, Anyanwu EE. **Synthesis and Characterization of Coal Fly Ash Supported Zinc Oxide Catalyst for Biodiesel Production Using Used Cooking Oil as Feed.** *Renewable Energy* 2021; **170**: 302-314.

[29] Maleki B, Esmaeili H. **Upgrading of NiFe2O4 via Decoration of g-C3N4 Nanoparticles to a Magnetic Catalyst for the Ultrasound-Assisted Production of Biodiesel.** *Scientific Reports* 2025.