

Modelling of biodiesel production from transesterification process of sandbox (*Hura crepitans* L.) seed oil: performance comparison of artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS)

Olajide Olukayode Ajala (✉ ajalaolajide84@gmail.com)

Obafemi Awolowo University <https://orcid.org/0000-0003-2410-6603>

Joel Olatunbosun Oyelade

Obafemi Awolowo University

Emmanuel Olusola Oke

Michael Okpara University of Agriculture

Oluwole Oluwatoyin Oniya

Ladoke Akintola University of Technology

Research Article

Keywords: Modelling , Biodiesel , Sandbox seed oil, Artificial intelligent , Artificial neural networks , Adaptive neuro-fuzzy inference system

Posted Date: January 30th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-1696942/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Biodiesel has been seen as an alternative to diesel (fossil) fuel as a result of its favourable properties, energy security reasons and environmental benefits. In this research, transesterification of sandbox seed oil with ethanol to form biodiesel has been modelled using artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) techniques. Temperature ($^{\circ}\text{C}$), time (min.), catalyst concentration (w/w) and catalyst type (g) were used as input variables while biodiesel yield was used as output variable for modelling the efficiency of biodiesel production from sandbox seed oil. The results showed that ANN model gave R^2 of 0.925, RSME of 2.99255, MAE of 0.62196, SEP of 0.03689 and AD of 0.03194 while ANFIS model gave R^2 of 0.961, RSME of 1.97379, MAE of 0.0001, SEP of 0.02433 and AD of 0.000005136. The results prove that ANFIS model is more reliable in predicting biodiesel yield from sandbox seed oil than ANN model.

1 Introduction

The desires for the establishment of national energy self-reliance and development of alternative fuels to fossil fuels have given birth to bioenergy which is based on the use of renewable agricultural-based (low-cost) materials as feedstock (Haas et al. 2006). Over the last decades, a tremendous amount of studies/researches have been carried out by using low-costs feedstock to produce alternative fuels to petroleum/natural gas whose prices are soaring in the world market coupled with their rapid depletion (Sai et al. 2018). Furthermore, the challenges of global warming and emission of carbon monoxide (CO) and carbon dioxide (CO_2) are driving immense growth in the global production of bioenergy (Oyelade et al. 2017). It is important to note that the emissions of CO and CO_2 are partly or majorly due to the type of fuel used, quality of fuel used, incomplete combustion of the fuel and the state of the engine (Oniya, 2012). Therefore, biodiesel is an optimistic substitute to fossil diesel fuel due to similarity in their properties (Adewuyi et al., 2014). According to Sai et al. (2018) and Oyelade et al. (2017), biodiesel is non-toxic, eco-friendly, renewable, biodegradable and comparatively a clean fuel used in internal combustion engines.

Biodiesel, a mixture of mono-alkyl esters of both saturated and unsaturated long chain-fatty acids, is produced by transesterification of waste frying oil, oilseeds, algae or animal fats with either ethanol or methanol in the presence of an alkaline catalyst (Kassem et al. 2018). The transesterification process for biodiesel production can be carried out using both homogenous (acid or base) and heterogeneous (acid, base or enzymatic) catalysts. In the last decades, homogenous catalysts especially sodium hydroxide (NaOH) or potassium hydroxide (KOH) have attracted attention in alkaline-based biodiesel production (Phan and Phan, 2008). However, they are costly, and their removal is very difficult because of large amount of waste water produced (Omotoso and Akinsanya, 2015).

Daily increase in human population brings about food-fuel crisis due to the use of edible oils such as canola oil, soyabean oil, cotton oil, palm kernel oil and rapeseed oil for biodiesel production. Adewuyi et al. (2014) reported that over 95% of biodiesel produced globally is from edible oils which consequently leads to food shortage, and as a result, attention has been shifted to the use of non-edible oils for biodiesel production. They stated further that, non-edible oils contain higher free fatty acid (FFA) than edible oils. Therefore, pretreatment is needed so as to lower the FFA content present in non-edible oil to a workable/normal level. The major cost incurred in biodiesel production is feedstock cost which contributes about 70-85% to the total production cost (Oraegbunam et al. 2022; Adewuyi et al. 2014). One of the best way to reduce biodiesel production cost is the use of non-edible oils

as feedstock for biodiesel production (Samuel et al. 2021; Adewuyi et al. 2014). Therefore, it has become imperative to search for underutilized non-edible oils as feedstock for biodiesel production.

Among the non-edible oils that seems appealing for biodiesel production is *Hura crepitans* (also known as sandbox, possum wood or jabillo) seed oil due to its oil content range (36.4-72.2%) and presence of unsaturated fatty acid, linoleic (50%) (Oraegbunam et al. 2022). *Hura crepitans* is a dicotyledon plant which belongs to the family of *Euphorbiaceae*. It is usually grown in tropical region of North and South America in Amazon Rainforest, and its tree can be as tall as 100ft (about 30m) with its leaves about 2ft (Oniya et al. 2014). It has dark mark, pointed spines and smooth brown bark. Nigerians (most especially rural dwellers) underutilize *Hura crepitans* plant as it is being planted purposely for shelter. Some researchers (Adewuyi et al. 2014; Oniya et al. 2014) have reported transesterification of *Hura crepitans* seed oil for biodiesel production while optimal conditions for biodiesel production from *Hura crepitans* L. seed oil using RSM were determined by Oyelade et al. (2017). However, in an open literature, few researchers have worked on using artificial intelligence (AI) modelling techniques to predict biodiesel yield from *Hura crepitans* seed oil. Among the artificial intelligence modelling techniques, both artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) which gives hybrid algorithm are preferred in this study due to their ability to handle non-linear and complex stochastic dataset.

2 Literature Review

ANN and ANFIS techniques help to model complex systems without a need for mathematical equations, and give high prediction accuracy (Najafi et al. 2018). These techniques have been successfully used for modelling and optimization of biodiesel production from different feedstock. Yue et al. (2018) applied ANFIS technique for modelling transesterification of castor oil for biodiesel production. Najafi et al. (2018) used RSM, ANFIS and ANN to predict and optimize biodiesel yield from waste cooking oil. Kassem et al. (2018) used both the ANN and ANFIS for the prediction of density and kinematic viscosity of biodiesel blends at various temperatures and volume fraction of biodiesel. Samuel et al. (2021) carried out modelling of transesterification process of biodiesel production from tobacco seed oil using ANN and ANFIS. The results from all these studies showed that both the ANN and ANFIS models have good predictive ability. To the best of our knowledge as researchers, integrating ANN with the ANFIS as a hybrid tool for modelling transesterification process of *Hura crepitans* seed oil for biodiesel production has not been properly explored. Though, Oraegbunam et al. (2022) modelled and optimized transesterification process of *Hura crepitans* seed oil with ANN and genetic algorithm (GA) with good results while Omilakin et al. (2020) assessed ANFIS and RSM in the modelling of oil extraction process from *Hura crepitans* seed oil. But, this study focussed on evaluation efficiency of ANN and ANFIS models in predicting biogas yield from *Hura crepitans* seed oil.

3 Methodologies

This section describes data gathering process as well as theories of ANN and ANFIS modelling techniques for the prediction of biodiesel yield.

3.1 Experimental data gathering

This research work is exclusively computational, and the dataset gotten from experimental research reported in the work of Oyelade et al. (2017) was used. The training, testing and validation processes for both ANN and ANFIS modelling were performed using this dataset. The input (temperature, time, catalyst concentration and catalyst type) and output (experimental biodiesel yield) dataset is presented in Table 1.

3.2 Development of artificial neural network (ANN) model

In this study, artificial neural network was employed as a tool for modelling biodiesel yield from transesterification process of sandbox seeds. This was done using ANN Toolbox of MATLAB 2013 software which is an in-built tool in MATLAB. It helps in modelling robust, complex and nonlinear systems that are difficult to model by conventional methods. The flowchart for ANN model development is shown in Fig. 1. ANN model was developed by using the input-output experimental data in Table 1 gotten from biodiesel production. The input sets used were residence time (min), temperature (°C), catalyst concentration (%) and catalyst type while the target was biodiesel yield (%). The multi-layered feed forward network was used to design ANN model while TRAINLM (Levenberg-Marquardt) back-propagation algorithm was used to train the network. Twenty four (24) samples of dataset were fed into ANN toolbox with 70% for training, 15% for validation and 15% for testing. The system architecture as shown in Fig. 2 has four (4) input layers with nodes indicating input variables to the problem, ten (10) hidden layers with nodes to help in capturing nonlinearity in the data, and one (1) output layer with node indicating a variable that is being modelled. Different numbers of hidden neurons inside the hidden layers were applied so as to make the best decision, while 10 hidden layers as shown in Fig. 3 give accurate decision boundary.

3.3 Development of adaptive neuro-fuzzy inference system (ANFIS) model

ANFIS as a hybrid modelling technique combines the strength of ANN and Fuzzy logic (Najafi et al. 2018). The input-output data obtained from the experiment was used for ANFIS modelling. The model was trained with part of the input and output data generated. Furthermore, the training data was divided into training dataset and checking dataset (checking dataset was employed to avoid over-fitting the model to the training dataset). The architecture of the proposed ANFIS model showing the input and output variables is presented in Fig. 3. The input and output crisp values were transformed into linguistic values which led to sectioning of membership function plot for easy and accurate analysis. The developed ANFIS model for the prediction of biodiesel yield using MATLAB is based on residence time (60–120 min.), temperature (55–65°C), catalyst concentration (1–3 w/w %) and catalyst type (KOH or CSS) as input variables. The data training was carried out with gaussian membership function (gaussmf). A fuzzy if-then rule based on a first-order Takagi-Sugeno fuzzy model was employed to provide the ANFIS structure, and the if-then rule is presented as follows:

If m is A and n is B , then $h = f(m, n)$

A and B represent fuzzy sets while $h = f(m, n)$ is a function of resulting from A and B .

3.4 Assessment of the developed ANN and ANFIS models

The goodness of the ANN and ANFIS models for predicting biodiesel production was investigated using statistical indices such as coefficient of determination (R^2), mean square error (MSE)/root mean square error (RMSE),

standard error prediction (SEP), mean average error (MAE) and average deviation (AD), and they are calculated as follows as applied by Samuel et al. (2021) and Okwu et al. (2020):

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_{i,pred} - Y_{i,exp})^2}{\sum_{i=1}^n (Y_{i,pred} - Y_{exp,ave})^2}$$

1

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_{i,exp} - Y_{i,pred})^2}{n}}$$

2

$$SEP = \frac{RMSE}{Y_{exp,ave}}$$

3

$$MAE = \sum_{i=1}^n \frac{(Y_{i,exp} - Y_{i,pred})}{n}$$

4

$$AD = \frac{100}{n} \sum_{i=1}^n \frac{(Y_{i,exp} - Y_{i,pred})}{Y_{i,exp}}$$

5

where $Y_{i,pred}$; $Y_{i,exp}$; $Y_{exp,ave}$ represent predicted output, experimental output and average experimental output respectively while n is the number of experimental data.

4 Results And Discussion

4.1 Biodiesel yield from sandbox (Huran crepitans) oil

The experimental and predicted biodiesel yield for each of the experimental runs using either KOH or CSS as catalyst is presented in Table 1.

Table 1
Results of actual and predicted biodiesel yield from sandbox oil.

Run	Temp (°C)	Time (min.)	Catalyst conc. (w/w)	Catalyst type (g)	Experimental biodiesel yield (%)	ANN predicted biodiesel yield (%)	ANFIS predicted biodiesel yield (%)
1	65.00	120.00	2.00	KOH	82	82.0106	81.9999
2	55.00	60.00	1.00	KOH	88	83.72714	87.9999
3	55.00	120.00	2.00	CSS	93	93.14045	92.9999
4	65.00	60.00	3.00	KOH	80	79.57857	77.4999
5	65.00	60.00	2.00	CSS	73	72.62402	72.9999
6	55.00	120.00	3.00	KOH	96	94.12281	93.9999
7	55.00	120.00	1.00	KOH	89	89.45493	88.9999
8	65.00	60.00	3.00	KOH	75	79.57857	77.4999
9	55.00	60.00	1.00	CSS	65	57.95468	64.9999
10	60.00	60.00	3.00	CSS	67	64.04127	65.5000
11	60.00	60.00	3.00	CSS	64	64.04127	65.5000
12	55.00	90.00	3.00	CSS	90	86.63227	87.9999
13	65.00	60.00	1.00	CSS	65	64.72773	65.0000
14	55.00	90.00	3.00	CSS	86	86.63227	87.9999
15	60.00	90.00	2.00	KOH	85	85.05496	84.9998
16	65.00	120.00	3.00	CSS	78	84.47732	83.4999
17	55.00	120.00	1.00	CSS	87	87.26429	86.9999
18	55.00	60.00	2.00	KOH	87	89.45724	86.9999
19	60.00	75.00	1.50	CSS	71	65.89702	70.9998
20	65.00	60.00	1.00	CSS	65	64.72773	65.0000
21	65.00	120.00	3.00	CSS	89	84.47732	83.4999
22	62.50	105.00	2.00	CSS	88	86.49664	87.9997
23	65.00	120.00	1.00	CSS	92	91.83102	91.9999
24	55.00	120.00	3.00	KOH	92	94.12281	93.9999

4.2 ANN model for prediction of biodiesel yield

In this study, MATLAB R2013a was employed to successfully implement the prediction of biodiesel yield using ANN. The highlights of the grouped available dataset are presented in Table 2. It is shown in Table 2 that the dataset was split into 70% for training, 15% for validation and 15% for testing. Five (5) iterations were performed

on the ANN toolbox with feasible solutions obtained at fifth iteration depicting accurate prediction of biodiesel yield. Furthermore, the results obtained show that ANN model was efficient with low values of MSE for training, validation and testing, and correlation coefficient (R^2) for training, validation and testing all close to unity. When R^2 is zero (0), it implies absence of a linear relationship, whereas unity (1) means a perfect relationship between variables. A good result is when R^2 value falls within 0.7-1.0 (Okwu et al. 2020).

Table 2
ANN result for MSE and R values

	% of separation	Number of sample	MSE	R
Training	70	16	4.47733	0.97852
Validation	15	4	11.77425	0.93756
Testing	15	4	24.04858	0.99624
All	100	24	-	0.96396

The regression plots of ANN model which show relationship between the target and output data is presented in Fig. 4. It is shown in the figure that the data is correlated by straight lines. Furthermore, Table 2 clearly shows MSE values for training (4.47733), validation (11.77425) and testing (24.04858) while the correlation coefficients (R) for training, validation and testing are respectively 0.97852, 0.93756 and 0.99624. According to these results, it shows that prediction by ANN model is accurate with overall R value of 0.96396. This implies that, minimal error is between the experimental biodiesel yield and predicted biodiesel yield by ANN model as presented in Table 1. Figure 5 presents the best validation performance of ANN model drawn from the curve of different epoch. The best validation performance is 11.7743 at epoch 2. The MSE least value obtained for validation performance confirms the best results, and ability of the ANN model to predict biodiesel yield effectively. In the error histogram plot presented in Fig. 6, training, validation and testing datasets have blue, green and red fonts respectively while the point where error tends to zero (0) is represented on the plot with a thin straight orange line.

4.3 ANFIS model for prediction of biodiesel yield

In this study, the fuzzy inference system (FIS) used for ANFIS modelling is shown in Fig. 7. The membership function for each of input and output variables was 3 (low-medium-high). The fuzzy logic rule viewer shown in Fig. 8 contains 81 rules with logical connector AND for each rule. The rulers at the left-hand side are the input variables which include residence time, temperature, catalyst concentration and catalyst type while ruler at the right-hand side of the rule viewer gave predicted biodiesel yield as output solution.

Graph of experimental biodiesel yield, ANN predicted biodiesel yield and ANFIS predicted biodiesel yield against experimental runs is shown in Fig. 9. ANFIS model predicted biodiesel yield that has almost same amount to actual biodiesel yield when compared with ANN model. Coefficient of determination (R^2) of 0.961 for ANFIS implies that 96.1% of the experimental data fit into ANFIS model while 92.5% ($R^2 = 0.925$) of the experimental data fit into ANN model. Other statistical indices also showed that ANFIS model performed better than ANN model.

4.4 Comparative assessment of ANN and ANFIS models

Statistical indices such as R^2 , RMSE, MSE, SEP, MAE and AD were used to determine the precision and performance of both ANN and ANFIS models. When the results of ANN and ANFIS models were compared as shown in Table 3, superiority in the performance of ANFIS model over ANN model was evident due to its high value of R^2 and low values of RMSE, MSE, SEP, MAE and AD.

Table 3
Comparative of statistical indices of
ANN and ANFIS

Variable	ANN	ANFIS
R^2	0.925	0.961
RSME	2.99255	1.97379
SEP	0.03689	0.02433
MAE	0.62196	0.0001
AD	0.03194	0.000005136

5 Conclusion

In this study, application of ANN and ANFIS in the modelling of transesterification process for biodiesel production from sandbox oil was established. The predictive ability of the ANN and ANFIS models were assessed and compared using statistical indices which include R^2 , RMSE, MSE, SEP, MAE and AD. The R^2 (0.925), RSME (2.99255), MAE (0.62196), SEP (0.03689) and AD (0.03194) of ANN model were compared with R^2 (0.961), RSME (1.97379), MAE (0.0001), SEP (0.02433) and AD (0.000005136) of ANFIS model. It was inferred from the results that ANFIS model has better predictive ability than ANN model for prediction of biodiesel yield from transesterification of sandbox oil. It is therefore recommended that techno-economic analysis of the process should be carried out in the future in order to commercialize the project among the populace with low income. Furthermore, exergy analysis of the biodiesel production from transesterification process of sandbox oil can also be explored in the future.

Declarations

Funding: No funding was received for this study

Availability of data and material: The dataset used in this study can be found using DOI: 10.1080/15567036.2017.1320691

Authors' contributions: All authors contributed to the idea of the article. Oyelade J.O developed the introductory part, Ajala O.O developed the models and discussed the results while Oke E.O and Oniya O.O critically revised the work.

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Competing interests: We declare that the authors have no affiliations or financial involvement with anyone or any organization with a financial interest or conflict with the subject or material discussed in this manuscript. We also confirm that this manuscript has not been published in any other journal and is not under consideration for publication elsewhere. The authors approved the manuscript and agreed to its submission in “Soft Computing”.

References

1. Adewuyi A Awolade P, Oderinde R (2014) Hura crepitans: An alternative feedstock for biodiesel production. *J. Fuels*, 1-8, <http://dx.doi.org/10.1155/2014/464590>.
2. Haas M, McAloon A, Yee W, Foglia T (2006) A process model to estimate biodiesel production costs. *Bioresource Technology*, 97:671–678, Doi:10.1016/j.biortech.2005.03.039.
3. Kassem Y, Çamur H, Bennur K (2018) Adaptive neuro-fuzzy inference system (ANFIS) and artificial neural networks (ANN) for predicting the kinematic viscosity and density of biodiesel-petroleum diesel blends. *American Journal of Computer Science and Technology* 1:8-18, Doi: 10.11648/j.ajcst.20180101.12.
4. Najafi B, Ardabili S, Shamshirband S, Chau K, Rabczuk T (2018) Application of ANNs, ANFIS and RSM to estimating and optimizing the parameters that affect the yield and cost of biodiesel production. *Engineering Applications of Computational Fluid Mechanics*, 12 (1):611-624, Doi: 10.1080/19942060.2018.1502688.
5. Okwu M, Samuel O, Otanocha O, Tartibu L, Omoregbee H, Mbachu, V (2020) Development of ternary models for prediction of biogas yield in a novel modular biodigester: a case of fuzzy Mamdani model (FMM), artificial neural network (ANN), and response surface methodology (RSM). *Biomass Conversion and Biorefinery*, <https://doi.org/10.1007/s13399-020-01113-1>.
6. Omilakin R, Ibrahim A, Sotunde B, Betiku E (2020) Process modelling of solvent extraction of oil from Hura crepitans seeds: adaptive neuro-fuzzy inference system versus response surface methodology. *Biomass Conversion and Biorefinery*, <https://doi.org/10.1007/s13399-020-01080-7>.
7. Omotoso MA, Akinsanya OA (2015) A review of biodiesel generation from non-edible seed oils crop using non-conventional heterogeneous catalyst. *J. Petroleum Technol. Altern. Fuel* 6:1–12.
8. Oniya OO, Akande FB, Adedeji AA, Olukayode OL (2014) Transesterification of Hura crepitans oil for biodiesel production. *J. Eng. Appl. Scientific Res.* 6 (1):91-97.
9. Oraegbunam J, Ishola N, Sotunde B, Betiku E (2022) Sandbox oil biodiesel production process modelling and optimization via artificial neural networks and genetic algorithm. *Proceedings of the OAU Faculty of Technology Conference*, (pp. 225-232). Ile-Ife.
10. Oyelade J, Idowu D, Oniya O, Ogunkunle O (2017) Optimization of biodiesel production from sandbox (Hura crepitans L.) seed oil using two different catalysts. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, DOI: 10.1080/15567036.2017.1320691.
11. Phan AN, Phan TM (2008) Biodiesel production from waste cooking oils. *Fuel* 87:3490–3496.
12. Sai B, Niju S, Meera B, Anantharaman N (2018) Optimization and modeling of biodiesel production using fluorite as a heterogeneous catalyst. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 1-17, Doi: 10.1080/15567036.2018.1549165.
13. Samuel O, Okwu M, Tartibu L, Giwa S, Sharifpur N, Jagun Z (2021) Modelling of Nicotiana Tabacum L. oil biodiesel production: comparison of ANN and ANFIS. *Frontiers in Energy Research*, 8:1-14, Doi: 10.3389/fenrg.2020.612165.

14. Yue X, Chen Y, Chang G (2018) Accurate modeling of biodiesel production from castor oil using ANFIS. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* , 1-7, DOI: 10.1080/15567036.2017.1422058.

Figures

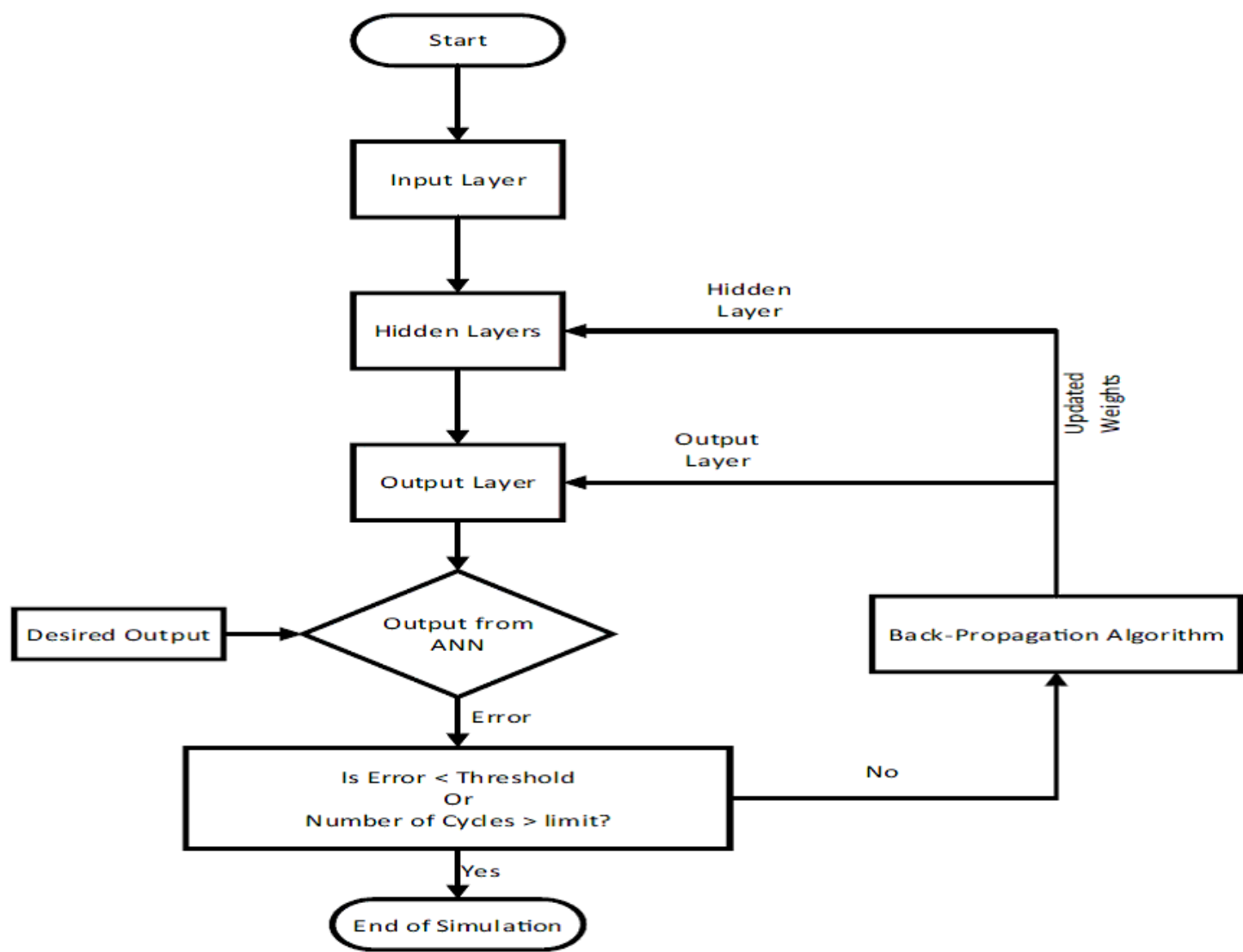


Figure 1

ANN flowchart for modelling biodiesel production.

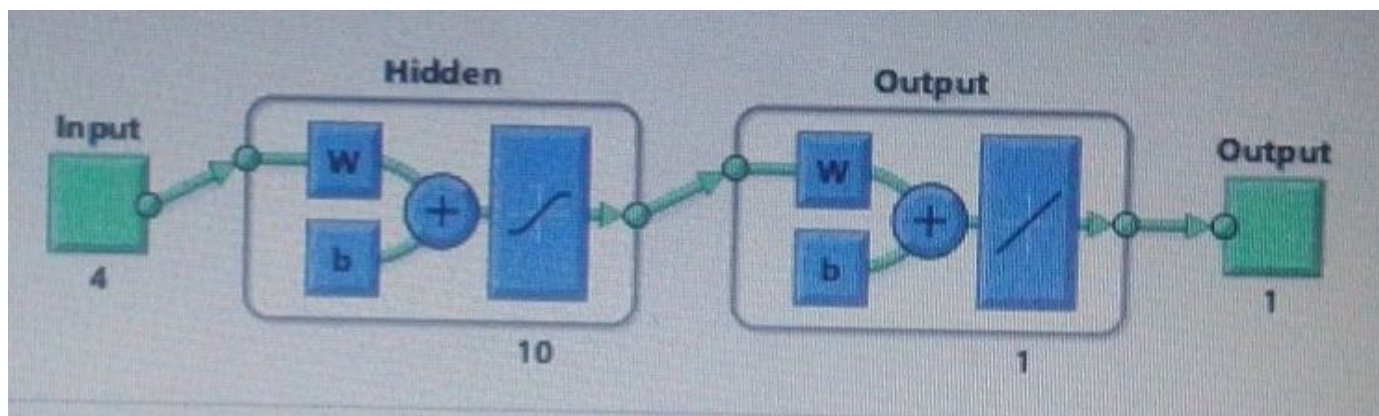


Figure 2

Architecture of ANN model for biodiesel production.

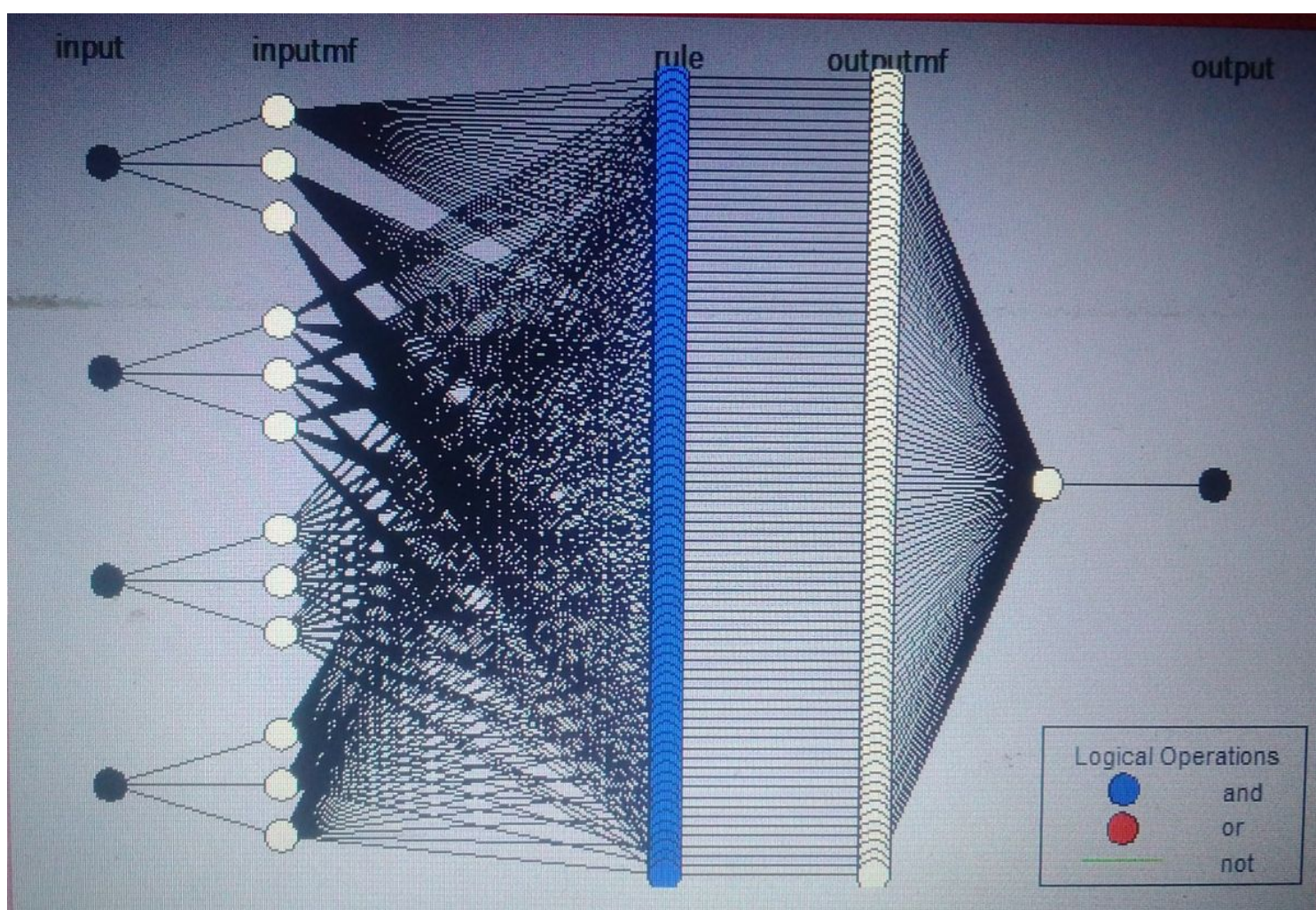


Figure 3

Architecture of ANFIS model for biodiesel production.

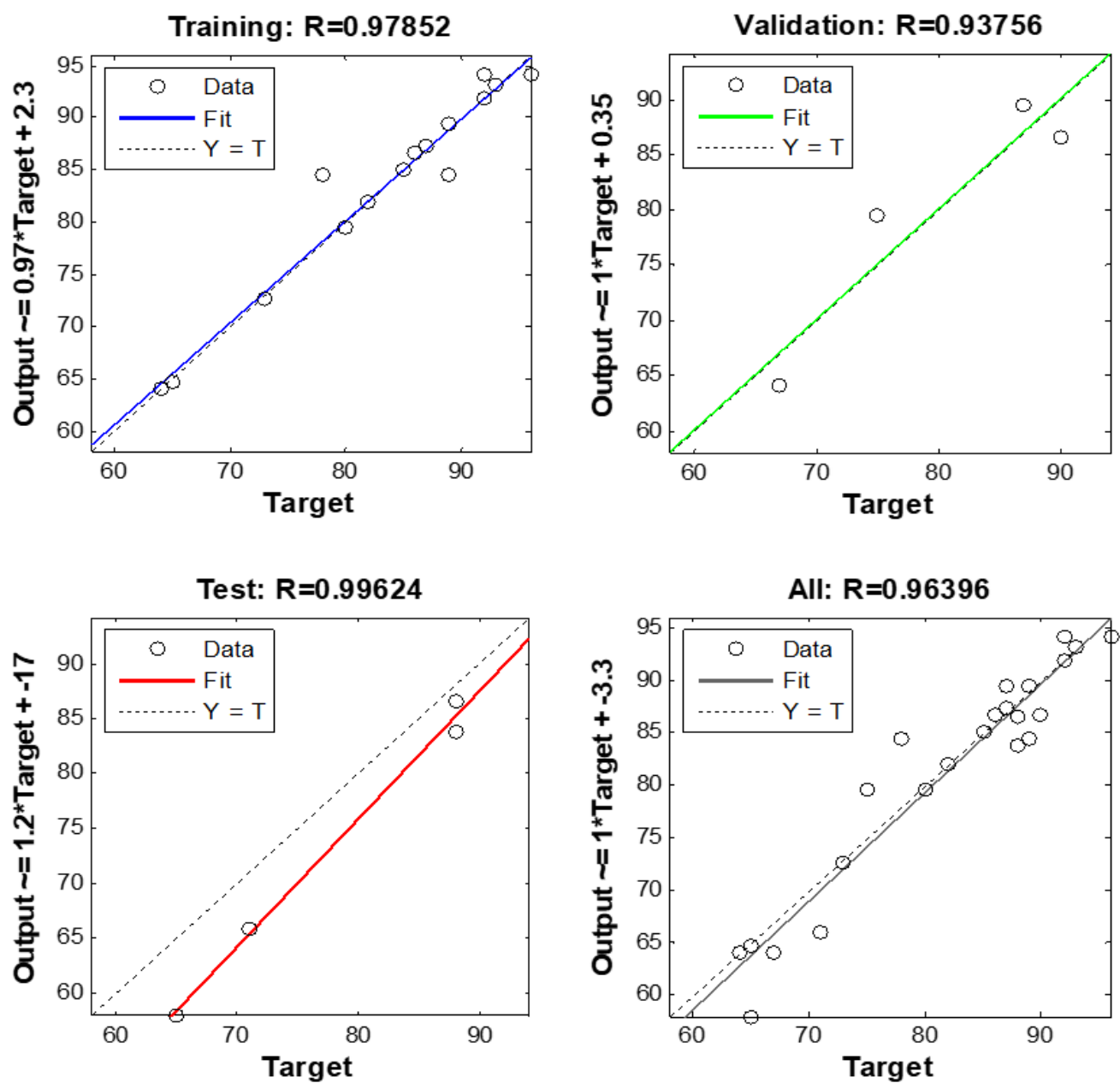


Figure 4

Regression plots of the ANN model for prediction of biodiesel yield.

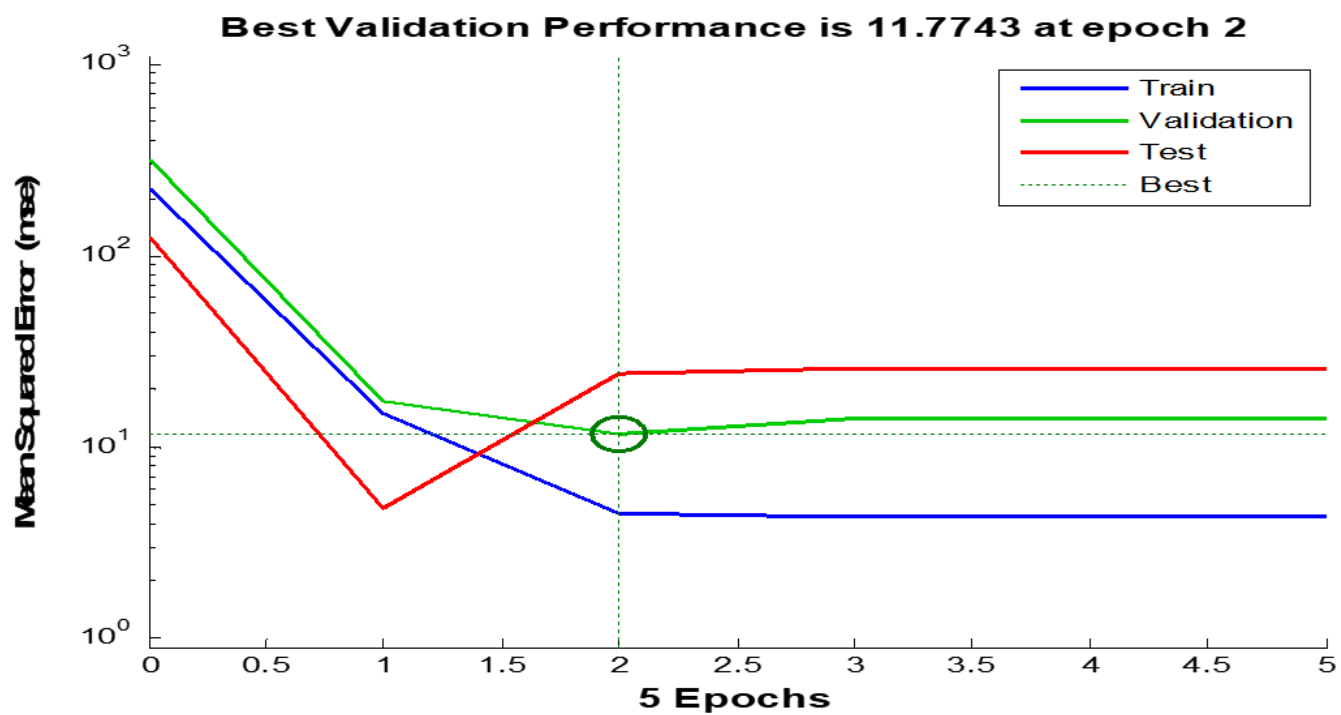


Figure 5

Performance plot of ANN model.

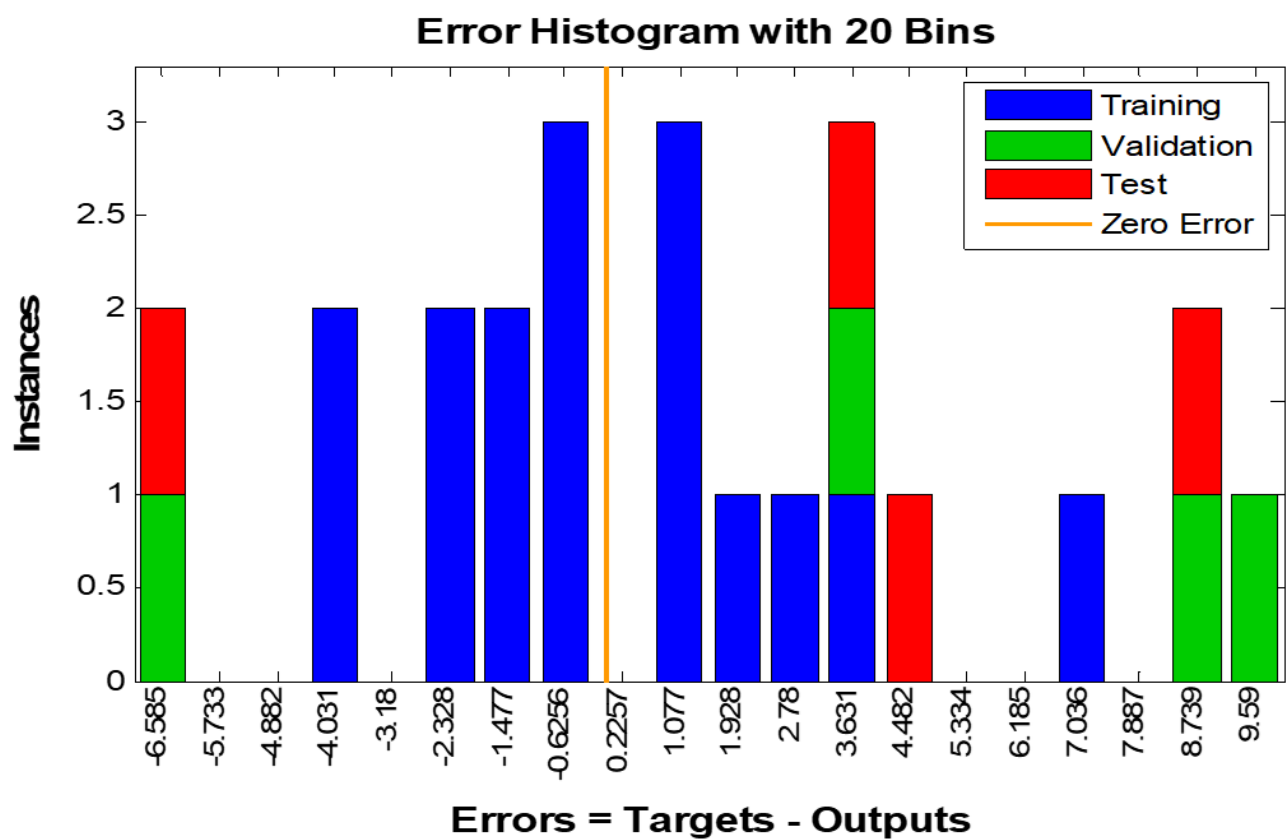


Figure 6

Error histogram plot of ANN model.

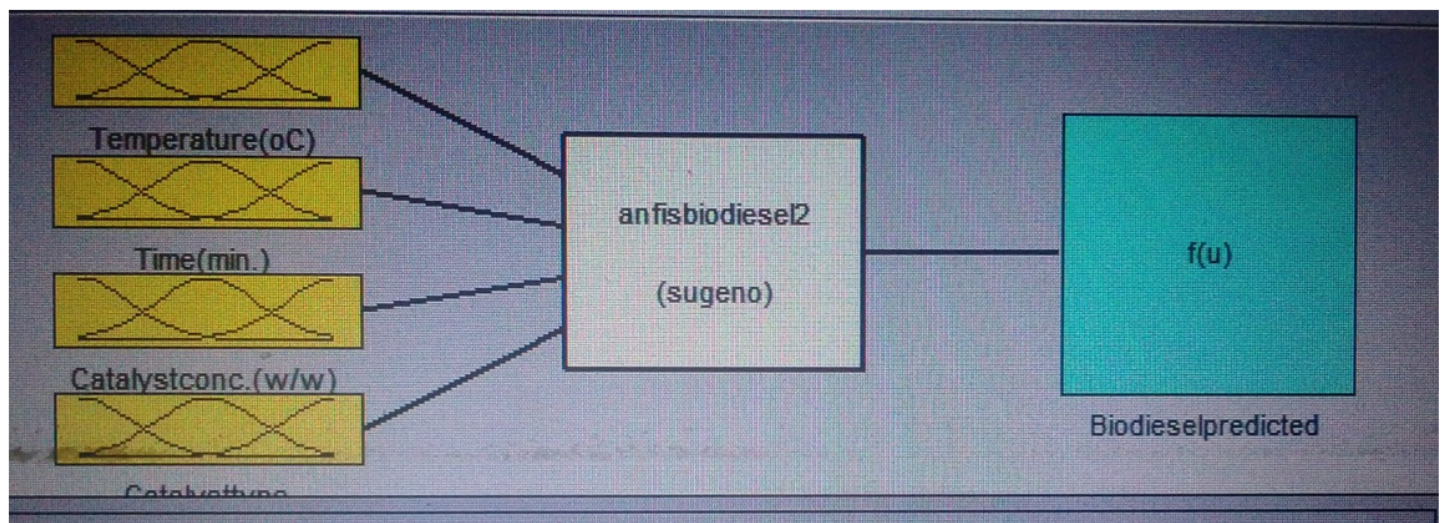


Figure 7

ANFIS model structure.

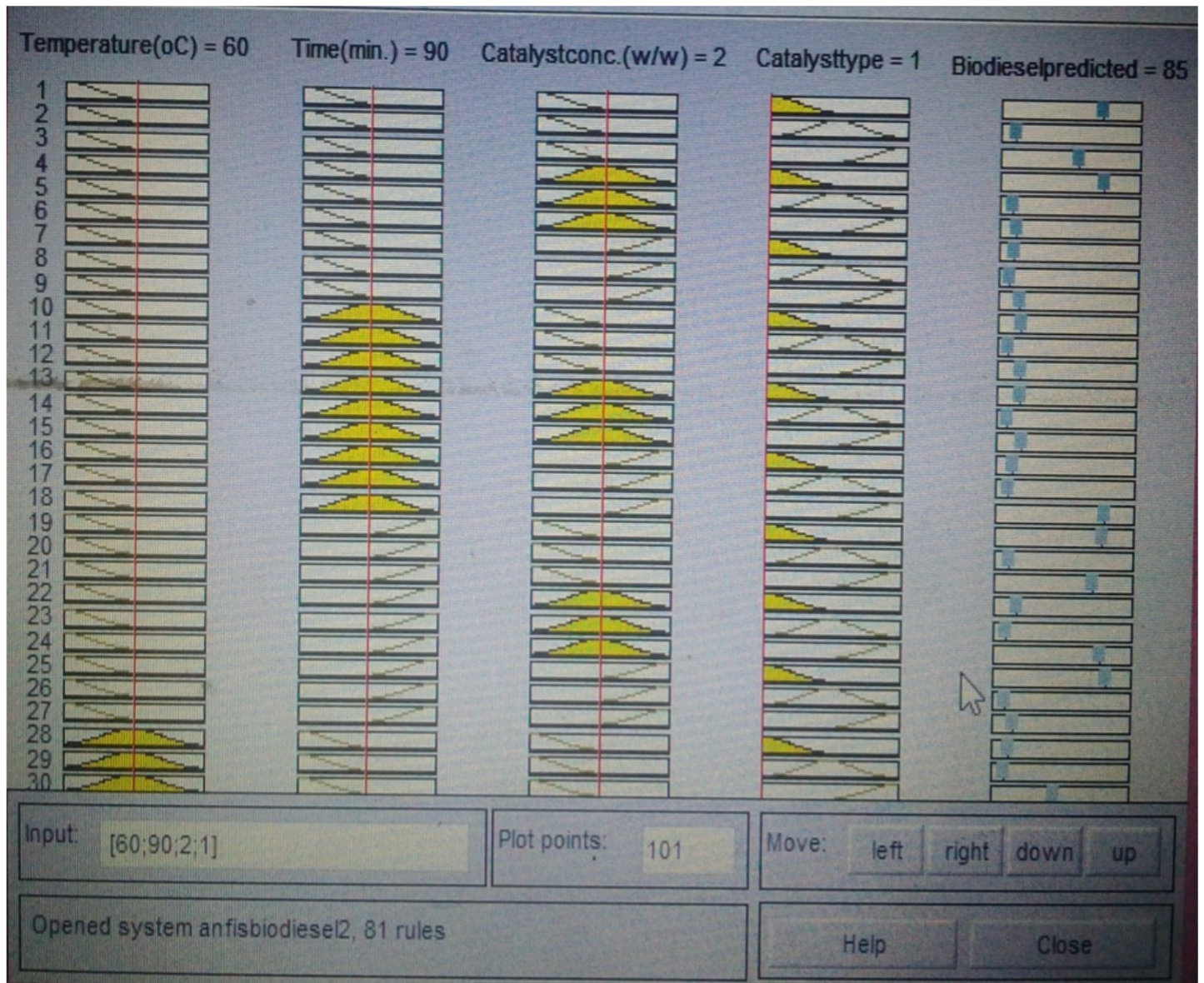


Figure 8

Fuzzy logic rule viewer.

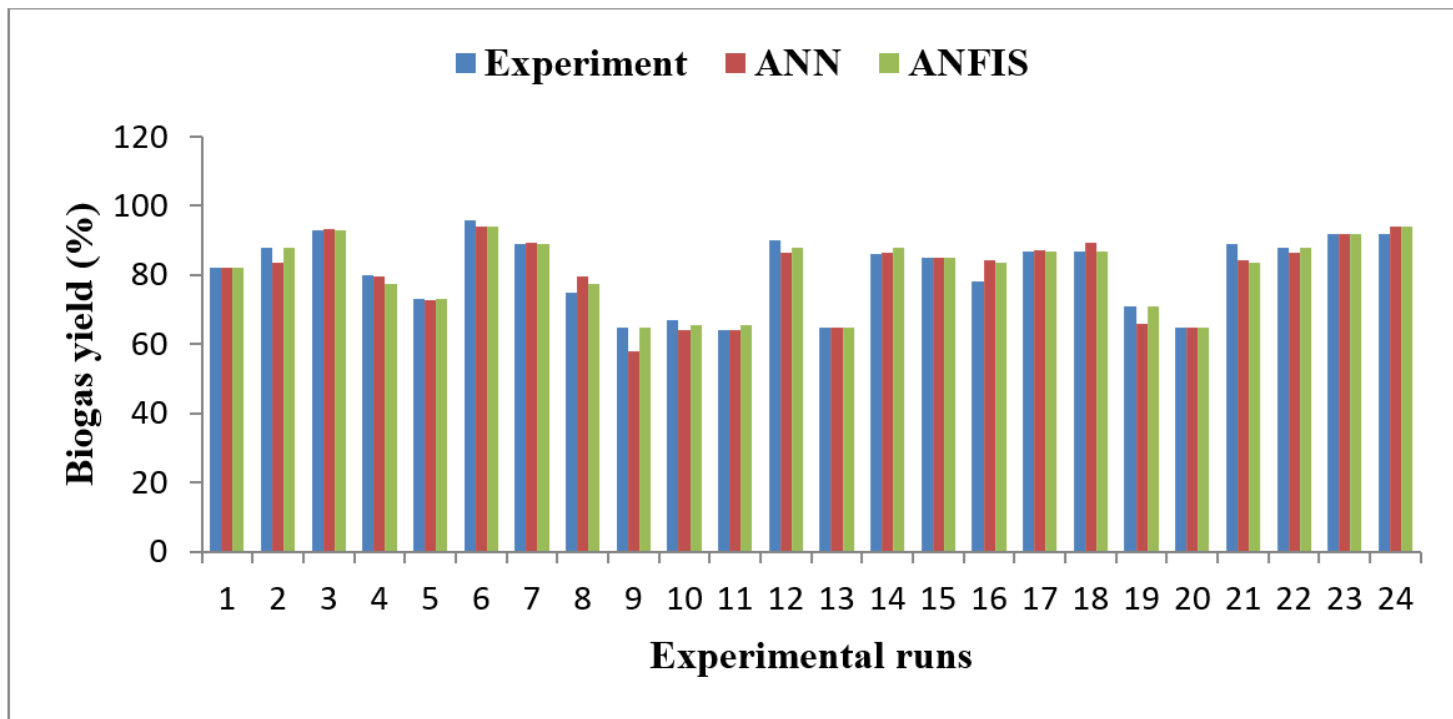


Figure 9

Experimental runs against experimental, ANN and ANFIS predicted biodiesel yield.