



ELSEVIER

Contents lists available at [ScienceDirect](http://ScienceDirect)

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

## Data on artificial neural network and response surface methodology analysis of biodiesel production

A.A. Ayoola<sup>a,\*</sup>, F.K. Hymore<sup>b</sup>, C.A. Omonhinmin<sup>c</sup>, P.O. Babalola<sup>d</sup>,  
E.O. Bolujo<sup>e</sup>, G.A. Adeyemi<sup>e</sup>, R. Babalola<sup>f</sup>, O.A. Olafadehan<sup>g</sup><sup>a</sup> Chemical Engineering Department, Covenant University, Ota, Ogun State, Nigeria<sup>b</sup> Regent University College of Science and Technology, Accra, Ghana<sup>c</sup> Biological Sciences Department, Covenant University, Ota, Ogun State, Nigeria<sup>d</sup> Mechanical Engineering Department, Covenant University, Ota, Ogun State, Nigeria<sup>e</sup> Petroleum Engineering Department, Covenant University, Ota, Ogun State, Nigeria<sup>f</sup> Chemical/Petrochemical Engineering Department, Akwa Ibom State University, Nigeria<sup>g</sup> Chemical and Petroleum Engineering Department, University of Lagos, Nigeria

## ARTICLE INFO

## Article history:

Received 18 March 2020

Revised 13 May 2020

Accepted 13 May 2020

Available online 20 May 2020

## Keywords:

ANN

Biodiesel

KOH

NaOH

RSM

Waste soybean oil

## ABSTRACT

The biodiesel production from waste soybean oil (using NaOH and KOH catalysts independently) was investigated in this study. The use of optimization tools (artificial neural network, ANN, and response surface methodology, RSM) for the modelling of the relationship between biodiesel yield and process parameters was carried out. The variables employed in the experimental design of biodiesel yields were methanol-oil mole ratio (6 – 12), catalyst concentration (0.7 – 1.7 wt/wt%), reaction temperature (48 – 62°C) and reaction time (50 – 90 min). Also, the usefulness of both the RSM and ANN tools in the accurate prediction of the regression models were revealed, with values of R-sq being 0.93 and 0.98 for RSM and ANN respectively.

© 2020 Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

\* Corresponding author.

E-mail address: [ayodeji.ayoola@covenantuniversity.edu.ng](mailto:ayodeji.ayoola@covenantuniversity.edu.ng) (A.A. Ayoola).

## Specifications table

Subject area	Energy, Engineering, Environmental Science
More specific subject area	Renewable Energy, Sustainability and the Environment, Engineering, Environmental Engineering
Type of data	Table, graph, figure
How data were acquired	Experimental design on transesterification reaction, the use of ANN algorithm, use of Minitab 17 software, model
Data format	Raw, analysed
Parameters for data collection	The variables employed in the generation of biodiesel yield data were methanol-oil mole ratio (6 – 12), catalyst concentration (0.7 – 1.7 wt/wt%), reaction temperature (48 – 62°C) and reaction time (50 – 90 minutes). Box-Benkehn BB(4) design in Minitab 17 environment was used for the design of experiments, ANN algorithm and RSM software were used for optimization studies. The esterification process of free fatty acid (FFA) removal involved the addition of 40 mL of a mixture of 25 mL isopropyl alcohol and 15 mL benzene solution to waste soybean oil (heated to 55°C), as well as the addition of 2 drops of phenolphthalein and the mixture was then titrated with 0.1 M KOH. Removal of water was done by heating the esterified oil at 110°C for 20 min.
Description of data collection	RSM analysis was done through the adoption of three (3) continuous factors: one (1) number of categorical factor, one (1) number of block and one (1) number of replicate for the plots of biodiesel yields and prediction of statistical models, analysis of <i>Sum of Errors coefficients</i> (SE coefficients), <i>p values</i> and <i>F values</i> . ANN algorithms were sectioned into data division (random), training (Levenberg Marquardt) and performance (mean squared errors). The 27 samples from the Box-Benkehn BB(4) design were distributed into training, validation and testing in the proportion of 19 samples (70%), 4 samples (15%) and 4 samples (15%) respectively. Plots of regression (training, validation, test and overall), mean squared error and error histogram were made using ANN algorithm.
Data source location	Department of Chemical Engineering, Covenant University, Ota, Nigeria.
Data accessibility	With the article

## Value of the data

- The data aided the determination of the quality level and efficiency of the two processes investigated in biodiesel production.
- The data are useful to authors in the field of renewable energy (at the global level). This is because the data revealed the processing conditions for the production of high yield and high-quality WSO-biodiesel.
- The ANN and RSM data could be used to predict reliable models that comparatively connect yields, processing conditions and the quality of yields (in terms of statistical tools).
- The ANN and RSM data could be used to improve the biodiesel production process by focusing on the operating conditions that gave high-quality value of biodiesel.
- The data showed the comparative assessment of KOH and NaOH catalysed processes (under the same operating conditions) during the WSO biodiesel process. This is a piece of important information in the consideration of a more suitable catalyst for WSO biodiesel process.

## 1. Data Description

Fig. 1 depicts the ANN Architecture in which the 27 samples (from the experimental design) were distributed into training, validation and testing in the proportion of 19 samples (70%), 4 samples (15%) and 4 samples (15%) respectively. The experimental design (using Box-Behnken BB(4)) for WSO biodiesel production at three levels and four process parameters ( $X_1$ ,  $X_2$ ,  $X_3$  and  $X_4$ ) is presented in Table 1. However, when ANN technique was applied, these experimental

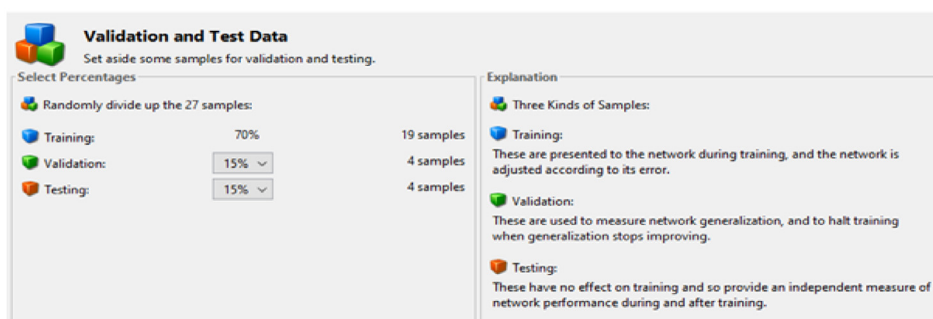
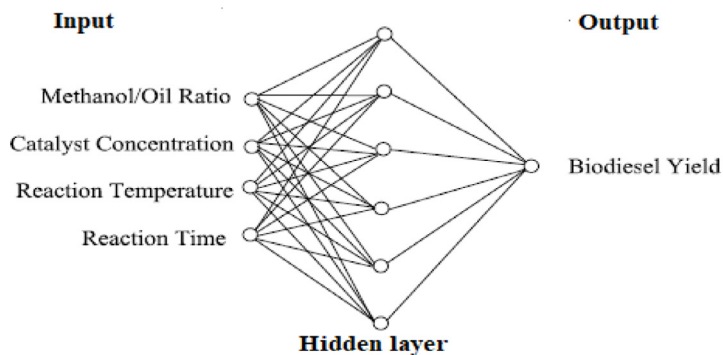


Fig. 1. ANN Architecture

Table 1

Experimental design for WSO biodiesel production using Box-Behnken BB(4)

Process parameter	Notation	Levels		
		-1	0	+1
Methanol to oil mole ratio	$X_1$	6:1	9:1	12:1
Concentration of catalyst, wt/wt% oil	$X_2$	0.7	1.2	1.7
Temperature of reaction, °C	$X_3$	48	55	62
Time of reaction, minutes	$X_4$	50	70	90

input parameters were re-assigned  $\alpha_m$ ,  $\alpha_c$ ,  $\alpha_T$  and  $\alpha_t$  to represent the mole ratio of methanol to oil, concentration of catalyst, temperature of reaction and time of reaction respectively. For ease of application of ANN algorithm, these four input parameters and the yields were computed by dividing each column of the actual process variables by the highest value in the column so as to have a range of zero (0) to one (1) values.

Table 2 shows the process parameters and the WSO biodiesel yield obtained from RSM and ANN tools.

Fig. 2 shows the comparative analyses of the WSO biodiesel yields from KOH and NaOH catalysed processes. Eqs. 1, 2 show the respective RSM models that relate the biodiesel yields and the four process variables investigated using KOH and NaOH catalysts. The fitness and reliability of the models were revealed by the high values of  $R$ -sq and  $R$ -sq (adj.), low values of probability value ( $p \leq 0.05$  at significant level), the low values of **Sum of Errors (SE) coefficient** and high values of  $F$ , as presented in Tables 3–4 for the statistical analyses of the RSM models for WSO biodiesel production using KOH and NaOH catalysts respectively.

Figs 3–5 reveal the respective plots of regression, mean squared error and error histogram for WSO biodiesel production using KOH catalyst. Figs. 6–8 reveal the respective plots of regression,

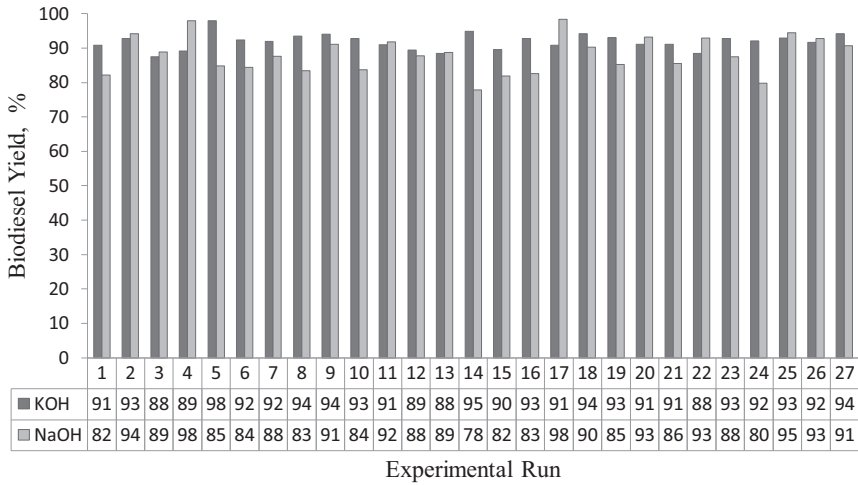


Fig. 2. WSO biodiesel yields from KOH and NaOH catalysed processes

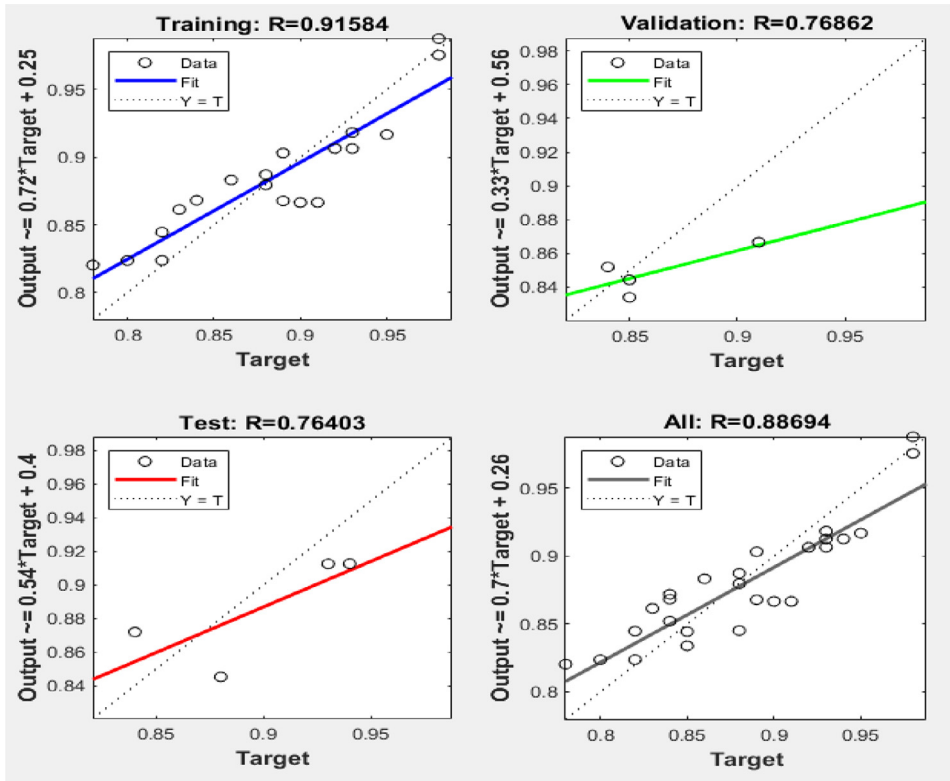


Fig. 3. Plots of regression: (i) training, (ii) validation (iii) test (iv) overall for WSO biodiesel production via KOH catalysed process

**Table 2**

Process variables and biodiesel yield (response) obtained from RSM and ANN techniques

S/N	Process variable				Yield			
	$\alpha_m$	$\alpha_c$	$\alpha_T$	$\alpha_t$	RSM (NaOH)	ANN (NaOH)	RSM KOH	ANN (KOH)
1	1.00	1.00	0.89	0.78	0.91	0.93	0.82	0.84
2	1.00	0.41	0.89	0.78	0.93	0.92	0.94	0.91
3	0.50	1.00	0.89	0.78	0.88	0.89	0.89	0.87
4	0.50	0.41	0.89	0.78	0.89	0.89	0.98	0.98
5	0.75	0.71	1.00	1.00	0.98	0.97	0.85	0.84
6	0.75	0.71	1.00	0.56	0.92	0.94	0.84	0.87
7	0.75	0.71	0.77	1.00	0.92	0.92	0.88	0.89
8	0.75	0.71	0.77	0.56	0.94	0.92	0.84	0.87
9	0.75	0.71	0.89	0.78	0.94	0.94	0.91	0.87
10	1.00	0.71	0.89	1.00	0.93	0.93	0.84	0.85
11	1.00	0.71	0.89	0.56	0.91	0.92	0.92	0.91
12	0.50	0.71	0.89	1.00	0.89	0.89	0.88	0.88
13	0.50	0.71	0.89	0.56	0.88	0.88	0.89	0.90
14	0.75	1.00	1.00	0.78	0.95	0.97	0.78	0.82
15	0.75	1.00	0.77	0.78	0.90	0.93	0.82	0.82
16	0.75	0.41	1.00	0.78	0.93	0.94	0.83	0.86
17	0.75	0.41	0.77	0.78	0.91	0.91	0.98	0.99
18	0.75	0.71	0.89	0.78	0.94	0.94	0.90	0.87
19	1.00	0.71	1.00	0.78	0.93	0.90	0.85	0.83
20	1.00	0.71	0.77	0.78	0.91	0.91	0.93	0.91
21	0.50	0.71	1.00	0.78	0.91	0.91	0.86	0.88
22	0.50	0.71	0.77	0.78	0.88	0.88	0.93	0.91
23	0.75	1.00	0.89	1.00	0.93	0.94	0.88	0.85
24	0.75	1.00	0.89	0.56	0.92	0.93	0.80	0.82
25	0.75	0.41	0.89	1.00	0.93	0.93	0.95	0.92
26	0.75	0.41	0.89	0.56	0.92	0.93	0.93	0.91
27	0.75	0.71	0.89	0.78	0.94	0.94	0.91	0.87

**Table 3**

Statistical analysis of the RSM model for WSO biodiesel production using KOH catalyst

Term	Coefficient	SE Coefficient	F	p
Constant	73.9378	6.11326	35.169	0.000000
Linear				
$X_1$	6.5672	0.55554	139.744	0.000000
$X_4$	-0.5117	0.15424	11.006	0.003831
Square				
$X_1X_1$	-0.3379	0.03065	121.552	0.000000
$X_2X_2$	-5.9667	1.00819	35.025	0.000013
$X_4X_4$	-0.0010	0.00069	2.132	0.161516
Interaction				
$X_2X_3$	0.2521	0.04434	32.314	0.000022
$X_3X_4$	0.0125	0.00232	29.144	0.000040

mean squared error and error histogram for WSO biodiesel production using NaOH catalyst. The best ANN topology could be observed at the least mean squared error value of approximately zero (0) and a correlation coefficient of one (1). The training of the network was terminated at the overshooting point.

$$(\text{WSObiodieselyield})_{\text{KOH}}$$

$$= 73.9378 + 6.5672 X_1 - 0.5116 X_4 - 0.3379 X_1 X_1 - 5.9666 X_2 X_2 \\ + 0.2520 X_2 X_3 - 0.0089 X_3 X_3 + 0.0125 X_3 X_4 - 0.0010 X_4 X_4$$

$$R - \text{sq} = 93.99\%$$

$$R - \text{sq}(\text{adj}) = 91.31\%$$

(1)

**Table 4**  
Statistical analysis of the RSM model for WSO biodiesel production using NaOH catalyst

Term	Coefficient	SE Coefficient	F	p
Constant	-48.6299	3.0973	27.632	0.003400
<b>Linear</b>				
X <sub>2</sub>	-56.6067	0.7498	117.043	0.089000
X <sub>3</sub>	5.6742	0.2134	10.806	0.063291
X <sub>4</sub>	1.0185			
<b>Square</b>				
X <sub>1</sub> X <sub>1</sub>	0.1080	0.06885	91.791	0.03050
X <sub>3</sub> X <sub>3</sub>	-0.0647	2.0097	32.505	0.00496
X <sub>4</sub> X <sub>4</sub>	-0.0050	0.00092	1.094	0.76270
<b>Interaction</b>				
X <sub>1</sub> X <sub>4</sub>	0.0321	0.05304	30.147	0.00022
X <sub>2</sub> X <sub>3</sub>	0.8400	0.03772	26.375	0.00921

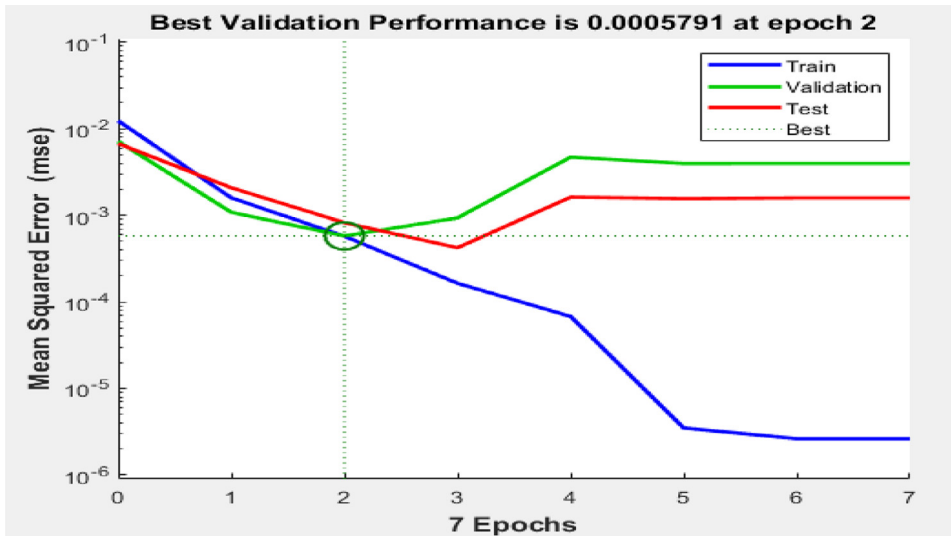


Fig. 4. Mean squared error for biodiesel production from WSO using KOH catalyst

$$\begin{aligned}
 &(\text{WSObiodieselyield})_{\text{NaOH}} \\
 &= -48.6299 - 56.6067 X_2 + 5.6742 X_3 + 1.0185 X_4 + 0.1080 X_1 X_1 - 0.0321 X_1 X_4 \\
 &\quad + 0.84 X_2 X_3 - 0.0647 X_3 X_3 - 0.0050 X_4 X_4 \\
 &R - sq = 80.18\% \qquad R - sq(\text{adj}) = 71.37\% \qquad (2)
 \end{aligned}$$

**2. Experimental Design, Materials, and Methods**

The pre-treatment of WSO involves the removal of impurities, free fatty acid (FFA) and water to prevent low yield of biodiesel (and even formation of other products) during transesterification reaction [1–6]. Hence, WSO was pre-treated before the transesterification process was carried out. Impurities were removed through filtration process (using the industrial sieve of 50 μm pore size), FFA was removed through the esterification process which involved mixing 25 mL isopropyl alcohol with 10 mL benzene. 40 mL of this solution was then added to WSO and

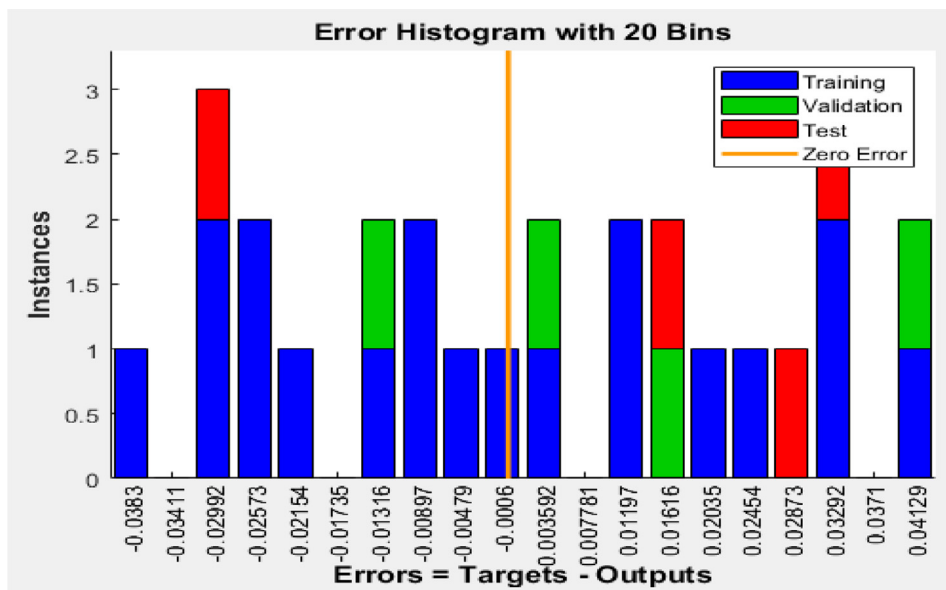


Fig. 5. Error histogram for biodiesel production from WSO using KOH catalyst

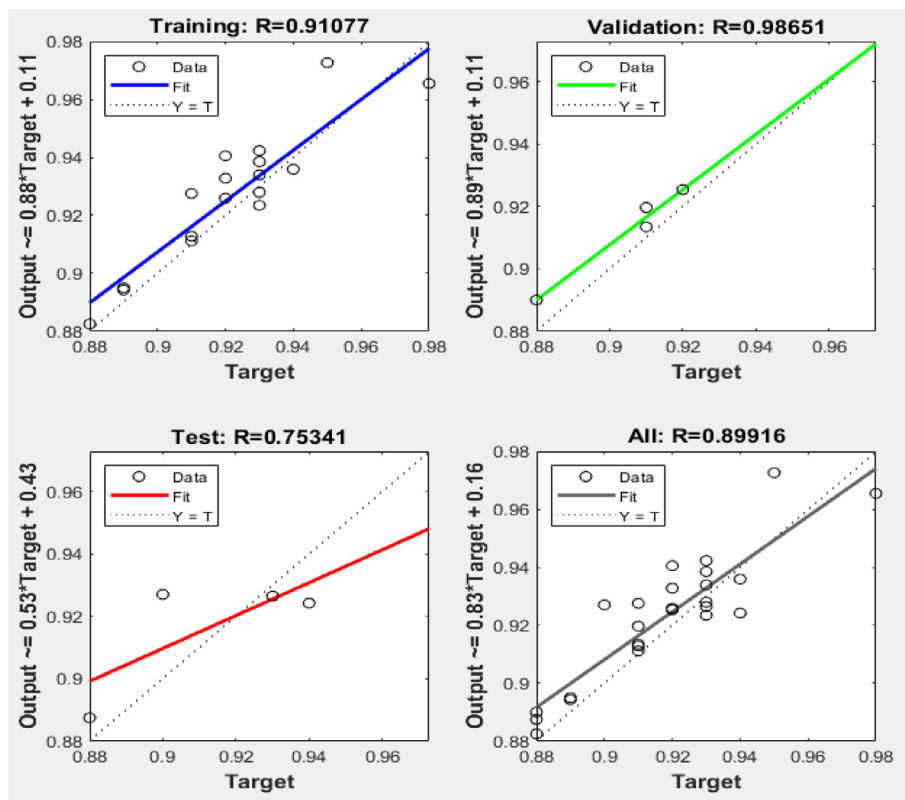


Fig. 6. Plots of regression: (i) training, (ii) validation (iii) test (iv) overall for NaOH catalysed WSO- biodiesel production

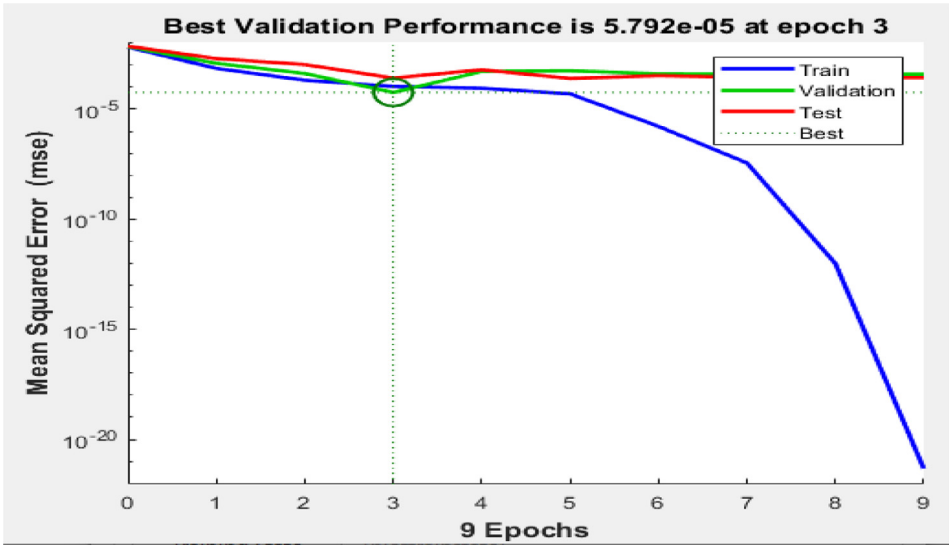


Fig. 7. Mean squared error for biodiesel production from WSO using NaOH catalyst

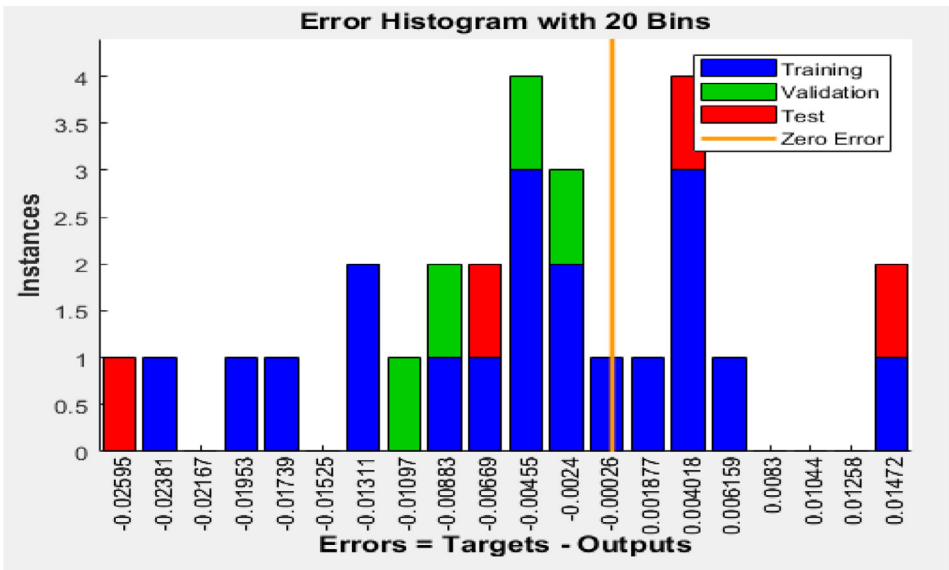


Fig. 8. Error histogram for biodiesel production from WSO using NaOH catalyst

heated to 55°C. 2 drops of phenolphthalein were added to the mixture and then titrated with 0.1 M KOH until a permanent pale pink colouration was observed. Removal of water was done by heating the esterified oil at 110°C for 20 min [7,8].

Box-Benken BB(4) experimental design (through the application of Minitab 17 software) was used for the design of experiments. The software was also used for statistical analysis, which involved the plots of experimental biodiesel yields and statistical modelling. The procedures used for the transesterification of the waste soybean oil using methanol and the two forms of catalysts (KOH and NaOH) separately were exhaustively outlined in Ayoola et al. [9,10].



The 27 samples from the experimental design were distributed into training, validation and testing in the proportion of 19 samples (70%), 4 samples (15%) and 4 samples (15%) respectively. To generate the R-sq values and least error conditions, the plots of regression (training, validation, test and overall), mean squared error and error histogram were made using the ANN algorithm.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

### Acknowledgement

Covenant University Centre for Research Innovation and Discovery (CUCRID) is acknowledged for financial involvement that makes the publication possible.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi: [10.1016/j.dib.2020.105726](https://doi.org/10.1016/j.dib.2020.105726).

### References

- [1] G.R. Moradi, S. Dehghani, F. Khosraviyan, A. Arjmandzadeh, The optimized operational conditions for biodiesel production from soybean oil and application of artificial neural networks for estimation of the biodiesel yield, *Renew. Energy* 50 (2013) 915–920 <https://doi.org/10.1016/j.renene.2012.08.070>.
- [2] B. Aghel, M. Rahimi, A. Sepahvand, M. Alitabar, H.R. Ghasempour, Using a wire coil insert for biodiesel production enhancement in a microreactor, *Energy Convers. Manag.* 84 (2014) 541–549 <https://doi.org/10.1016/j.enconman.2014.05.009>.
- [3] B. Aghel, M. Mohadesi, S. Sahraei, M. Shariatifar, New heterogeneous process for continuous biodiesel production in microreactors, *Can. J. Chem. Eng.* 95 (2017) 1280–1287 <https://doi.org/10.1002/cjce.22763>.
- [4] A.A. Ayoola, K.F. Hymore, C.A. Omonhinmin, Optimization of biodiesel production from selected waste oils using response surface methodology, *Biotech.* 16 (2017) 1–9 <https://doi.org/10.3923/biotech.2017.1.9>.
- [5] O. Agboola, S.E. Sanni, D.T. Oyekunle, A.O. Ayeni, B.A. Oni, A. Ayoola, P. Popoola, R. Sadiku, Assessment of the performance of osmotically driven polymeric membrane processes, *IOP Conference series* 1378 (032047) (2019).
- [6] V.E. Efevbokhan, J.D. Udonne, A.A. Ayoola, O.T. Shogbam, R. Babalola, A study of the effects of phenolic de-emulsifier solutions in xylene on the de-emulsification of a Nigerian crude oil emulsion, *J. Appl. Res. & Tech.* 15 (2) (2017) 110–121.
- [7] F.A.N. Fernandes, R.M. Lopes, M.P. Mercado, E.S. Siqueira, Production of soybean ethanol-based biodiesel using CaO heterogeneous catalysts promoted by Zn, K and Mg, *Int. J. Green Energy* 13 (2016) 417–423 <https://doi.org/10.1080/15435075.2014.977441>.
- [8] J. Goli, O. Sahu, Development of heterogeneous alkali catalyst from waste chicken eggshell for biodiesel production, *Renew. Energy.* 128 (2018) 142–154 <https://doi.org/10.1016/j.renene.2018.05.048>.
- [9] A.A. Ayoola, F.K. Hymore, C.A. Omonhinmin, O.C. Olawole, O.S.I. Fayomi, D. Babatunde, O. Fagbiele, Analysis of waste groundnut oil biodiesel production using response surface methodology and artificial neural network, *CDC* 22 (100238) (2019).
- [10] A.A. Ayoola, D.O. Adeniyi, S.E. Sanni, K.I. Osakwe, J.D. Jato, Investigating production parameters and impacts of potential emissions from soybean biodiesel stored under different conditions, *EER* 23 (1) (2017) 54–61.