engineering proceedings

MDPI

Proceeding Paper

Modelling and Optimisation of Biodiesel Production from Margarine Waste Oil Using a Three-Dimensional Machine Learning Approach †

Pascal Mwenge *, Hilary Rutto and Tumisang Seodigeng

Clean Technology and Applied Materials Research Group, Department of Chemical and Metallurgical Engineering, Vaal University of Technology, Private Bag X021, Vanderbijlpark 1900, South Africa; tumisangs@vut.ac.za (T.S.)

* Correspondence: pascalm@vut.ac.za; Tel.: +27-16-950-7830

† Presented at the 3rd International Electronic Conference on Processes—Green and Sustainable Process Engineering and Process Systems Engineering (ECP 2024), 29–31 May 2024; Available online: https://sciforum.net/event/ECP2024.

Abstract: This work presents the use of three-dimensional machine learning approaches, namely the response surface methodology (RSM), the artificial neural network (ANN), and the adaptive neuro-fuzzy inference system (ANFIS), to optimise and model the biodiesel yield from waste margarine oil. The effect of the process parameters methanol-to-oil ratio (3–15 mole), catalyst ratio (0.3–1.5 wt. %), reaction time (30–90 min), and reaction temperature (30–70 °C) were studied. The performance metric results for the RSM, ANN, and ANFIS were 0.991, 996, and 0.998 for regression (R^2); 0.924, 0.566, and 0.324 for root mean square error (RMSE); 0.568, 0.267, and 0.202 for mean absolute error (MAE); 0.746, 0.333, and 0.226 for mean absolute percentage error (MAPE); 0.008, 0.004, and 0.003 for average relative error (ARE); and 4.503, 2.114, and 1.828 for mean percentage standard deviation (MPSD). The developed three-dimensional machine learning approach—the RSM, ANN, and ANFIS models—is a potential method for optimising and modelling biodiesel yield. The study results may be used to create sustainable, efficient, and economical solutions for recycling waste margarine oil.

Keywords: biodiesel; machine learning; artificial neural network; adaptive neuro-fuzzy inference system; response surface methodology; waste margarine oil

check for updates

Citation: Mwenge, P.; Rutto, H.; Seodigeng, T. Modelling and Optimisation of Biodiesel Production from Margarine Waste Oil Using a Three-Dimensional Machine Learning Approach. *Eng. Proc.* **2024**, *67*, 27. https://doi.org/10.3390/engproc2024067027

Academic Editor: Juan Francisco García Martín

Published: 31 August 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Many government policies have endorsed the use of biofuel to reduce the reliance on fossil fuels and the pursuit of energy security by partly substituting fossil fuels and reducing the risk of environmental pollution and global warming [1], which are mainly affected by population growth and current industrialisation. Therefore, new forms of energy sources are required. Biodiesel is an alternative that has considerably increased in consideration [2,3]. The transesterification reaction of oils and animal fats in the presence of a catalyst is the method that is mainly used to produce biodiesel. The catalysts used are either homogenous or heterogeneous. Potassium hydroxide and sodium hydroxide are the homogenous catalysts most used in biodiesel production [4–6].

The biodiesel production cost is mainly attributed to the cost of feedstock, for which 75% is attributed to the cost of oil. Therefore, there is a need to use alternatives such as non-edible oils, as virgin oil leads to a biodiesel production cost of 1.5 times more than fossil diesel. Waste oils such as waste cooking oil are about two to three times cheaper than virgin oil, decreasing the biodiesel production cost and making it competitive with petroleum diesel [7]. Close to five tons per month per factory of waste oils are generated in butter and margarine production facilities; unfortunately, some of these oils are flushed from production and disposed of by municipalities, by they can be used as low-cost feedstock

for biodiesel production [1]. Using waste oils from a butter plan can be a cost-effective way of producing biodiesel [8].

Machine learning (ML) has been considered an effective method for transforming biodiesel production by enabling predictive modelling and process optimisation [9]. ML is a subdivision of artificial intelligence that uses actual values to train a model to manage complex tasks [10]. It has been reported that ML as a predictive model can predict responses with high precision. ML is inspired by the autolearning and self-improving ability of the brain to solve complex tasks. Thus, transesterification process modelling can be improved [11]. Various variables can affect biodiesel production, such as methanol-to-oil ratio, catalyst ratio, temperature, and time, for which experiments must be performed. Different statistical techniques are used to predict the effect of process variables on biodiesel yield. Response surface methodology has been mostly used for process optimisation, predicting output using fewer experiments [9]. Response surface methodology (RSM) is a tool that links one or more responses to process variables and combines statistics and mathematics to enhance and improve processes. Hence, it is used to develop prediction models.

Furthermore, the RSM, when used in transesterification processes, can boost the competitiveness of biodiesel over fossil fuels [12]. Artificial neural networks (ANNs) and adaptive neuro-fuzzy inference systems (ANFISs) are the most used ML predictive techniques. The ANN is a type of ML inspired by a biological brain's function, structure, and capacity for learning. ANNs provide solutions to issues that are shown to be difficult to solve by human or statistical standards. ANNs use complex datasets to model without needing knowledge of the governing system events. One of the disadvantages of the ANN is its inability to handle steady linguistic information or datasets. The ANFIS can overcome these flaws as it is built on a hybrid intelligence system that combines the computational power of the ANN and simple learning steps and uncertainties to illustrate the ability of fuzzy logic [13].

This work presents the optimisation and modelling of the transesterification of margarine waste oil using the catalyst ratio, time, methanol-to-oil ratio, and temperature as process variables and applying the three-dimensional machine learning approaches of the response surface methodology (RSM), the artificial neural network (ANN), and the adaptive neuro-fuzzy inference system (ANFIS). The performance of the ANN, ANFIS, and RSM is evaluated using error metrics to compare the robustness of ML.

2. Materials and Methods

Methanol (99.5%), potassium hydroxide (85%), and phenolphthalein indicators were purchased from ACE (Associate Chemical Enterprises). The waste margarine oil used as a triglyceride source was obtained from a local margarine production plant. The expected output (response) is biodiesel yield. The experimental setup, as shown in Figure 1, consists of a two-neck round-bottom flask as a reaction vessel, which was fitted with a reflux condenser to reflux methanol back to the reaction mixture; the heating source was a hot plate magnetic stirrer, which had an automatic timer and a temperature controller. Waste margarine oil was dried to remove moisture at a temperature of 110 °C for one hour, then let cool. The cooled oil was used to determine the free fatty acid (FFA) content of the oil, which was checked according to the method described by [14] and was found to be 1.79%, below 2%, suggesting that biodiesel could be produced from the margarine waste oil catalysed by potassium hydroxide (KOH) and avoiding saponification.

Biodiesel was produced by reacting approximately 100 g of waste margarine oil using process parameters, such as methanol-to-oil ratio (3–15 mol/mol), catalyst ratio (0.3–1.5 wt. %), time (30–90 min), and temperature (30–70 °C), as shown in Table 1. The methanol was mixed with the catalyst and transferred to the reaction vessel, where the oil was initially transferred. The reaction was carried out at a constant stirring speed of 450 RPM for a time, as per the range in Table 1. The product obtained was let settle and separated using a separation funnel, and the produced final biodiesel product obtained

was washed with distilled water and dried at 105 °C for one hour. The product yield was calculated using Equation (1).

Design Expert version 13 was used for experimental design, data analysis, optimisation, and predictive modelling in RSM using a central composite design. Design expert software was selected for its robustness capabilities in designing and optimising experiments. Twenty-one experimental runs were performed with methanol-to-oil ratio, catalyst ratio, time, and temperature as independent variables and biodiesel yield as a response. Neural Network Modular and Neuro-fuzzy were built with an NN toolbox using MATLAB 2021. The divider and function defaults in Matlab divided 70% of the data for training, 15% for testing, and 15% for validation. The network was trained using feed-forward propagation, and the Levenberg–Marquardt algorithm and linear (purelin) activation function were used to transfer data between layers. The ANFIS was generated using a grid partition and trained using a hybrid method; the Sugeno fuzzy inference system type was employed with the triangular membership function (trimf) used for the MF input type, and the output type MF used was constant. The ANFIS architecture, as shown in Figure 2, consisted of 5 layers with their respective colours: input (methanol-to-oil ratio, catalyst ratio, time, and temperature) (black) fuzzification (white), fuzzy rule base (blue), fuzzy inference (white), output defuzzification (white), and output (yield) (black).

$$Biodiesel\ Yield = \frac{Mass\ of\ dry\ biodiesel}{Mass\ of\ oil} \times 100 \quad (1)$$

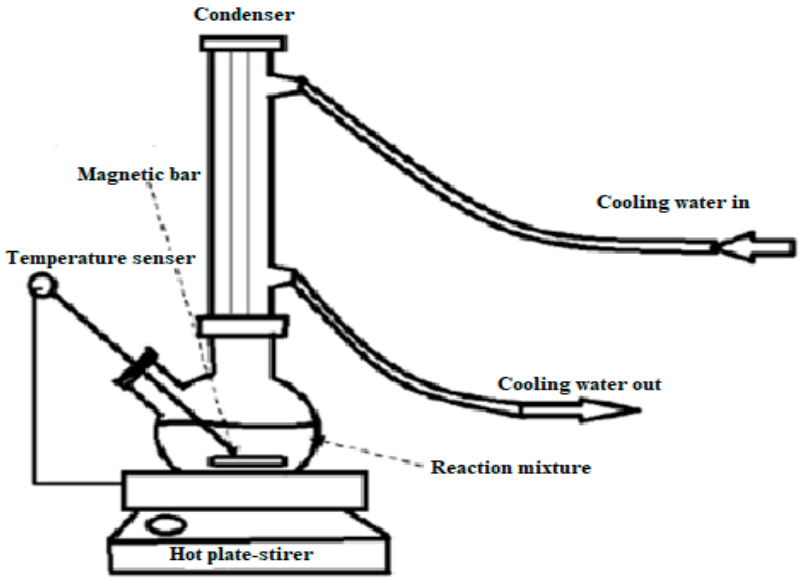

Figure 1. Experimental setup [14].

Table 1. Process variable range.

Input	Range	Output
Methanol/oil ratio (mol/mol)	3–15	Yield (%)
Catalyst ratio	0.3–1.5	
Time (minutes)	30–90	
Temperature (°C)	30–70	

The performance of the ML methods was evaluated using performance metrics such as regression coefficient (R^2), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), average relative error (ARE), and mean percentage standard deviation (MPSD).

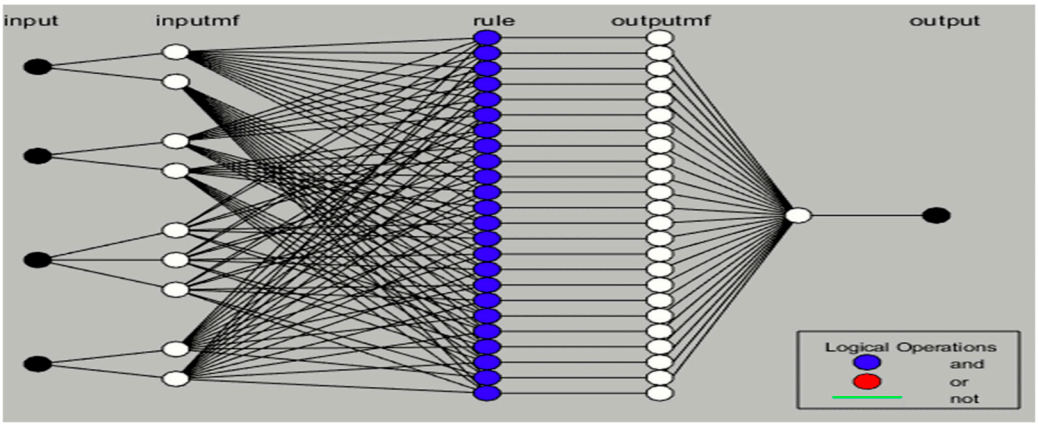

Figure 2. ANFIS model's architecture.

3. Results and Discussion

RSM was applied using Design Expert 13 (version 13.0.5.0) software; from the fit summary and lack-of-fit tests, a quadratic model was suggested over the linear model, a two-factor interaction (2FI) model, since the quadratic model had higher adjusted and predicted R^2 values and the cubic model was found aliased, as shown in Table 2. A quadratic mathematical model, shown in Equation (2), was obtained, which can predict the biodiesel yield using the four process parameters in this study. Figure 3 represents the experimental and predicted data of the RSM and shows an excellent fitting as the points are very close to the fit line. Numerical optimisation was performed in the RSM, and an optimum biodiesel yield of 89.09% was obtained with the following conditions: 60 min reaction time, 50 °C reaction temperature, 9 mol methanol-to-oil ratio, and 0.9 wt. % catalyst ratio.

$$Y = +89.06 + 7.91A + 7.16B + 3.64C + 7.09D + 3.08AB + 0.4762AC + 2.89AD - 1.09BC + 5.43BD + 1.34CD - 6.07A^2 - 4.35B^2 - 2.74C^2 - 3.91D^2 \quad (2)$$

Y is the yield, A is the methanol-to-oil ratio, B is the catalyst ratio, C is time, and D is temperature. The equation can be used to predict the biodiesel yield for the given ranges of factors and the respective units of the parameters. The factor coefficients can also be used to identify the relative impact of individual process variables by comparing the factor coefficients.

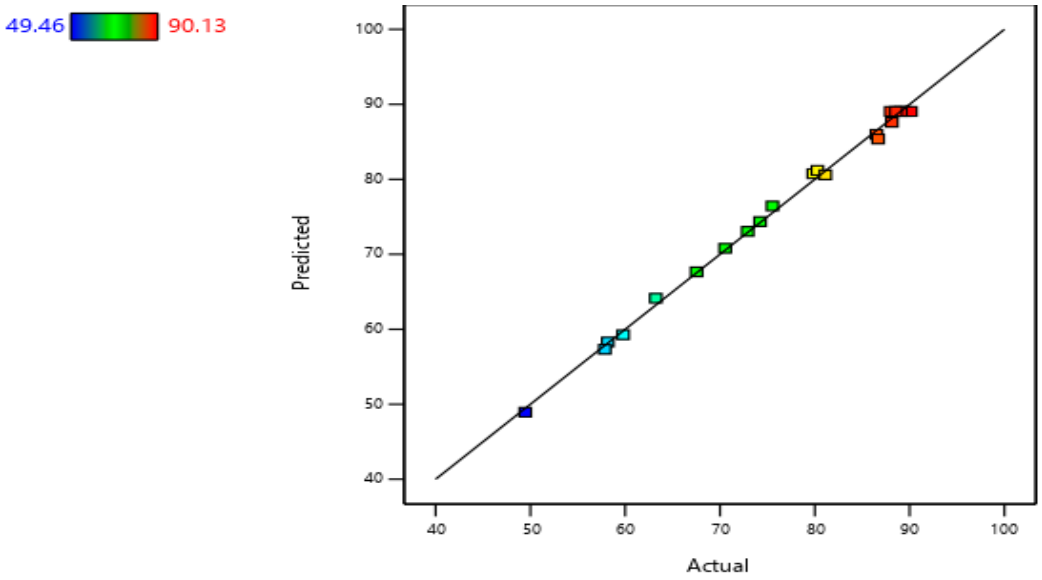

Figure 3. Actual vs. predicted yield from RSM (colour points are by value of yield).

Table 2. Fit Summary.

Source	Sequential *p*-Value	Lack-of-Fit *p*-Value	Adjusted R^2	Predicted R^2	
Linear	0.0140	<0.0001	0.4031	0.2790	
2FI	0.9363	<0.0001	0.1801	−2.4451	
Quadratic	<0.0001	0.0808	0.9900	0.7929	Suggested
Cubic	0.0808		0.9958		Aliased

Twenty-one experimental data points were randomly divided into 70% for training and 30% for validation and testing. The ANN architecture is shown in Figure 4, and the architecture consists of three layers: an input layer with four neurons, a hidden layer with 17 different nodes, and an output layer with one neuron. The network used for training was back-propagation. Where W, B and + in the hidden layer represent weights, bias and activation function, respectively.

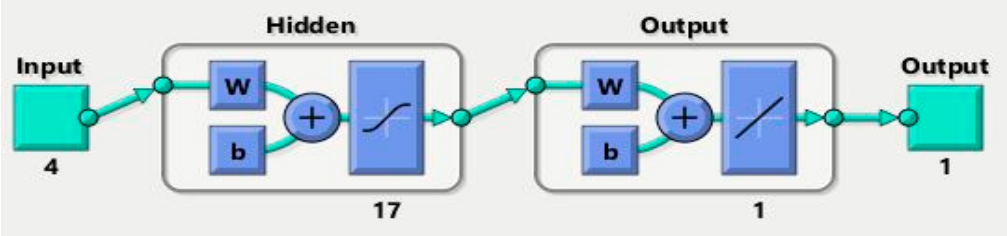

Figure 4. ANN architecture.

Figure 5 depicts how the network interacts with the training, testing, and validation data. The correlation coefficients for the training, testing, validation, and test data were found to be 0.999, 0.996, 0.999, and 0.998, respectively. The straight line demonstrates a linear correlation, implying that the targeted experimental data (yield) correlate adequately with the output (predicted data: yield).

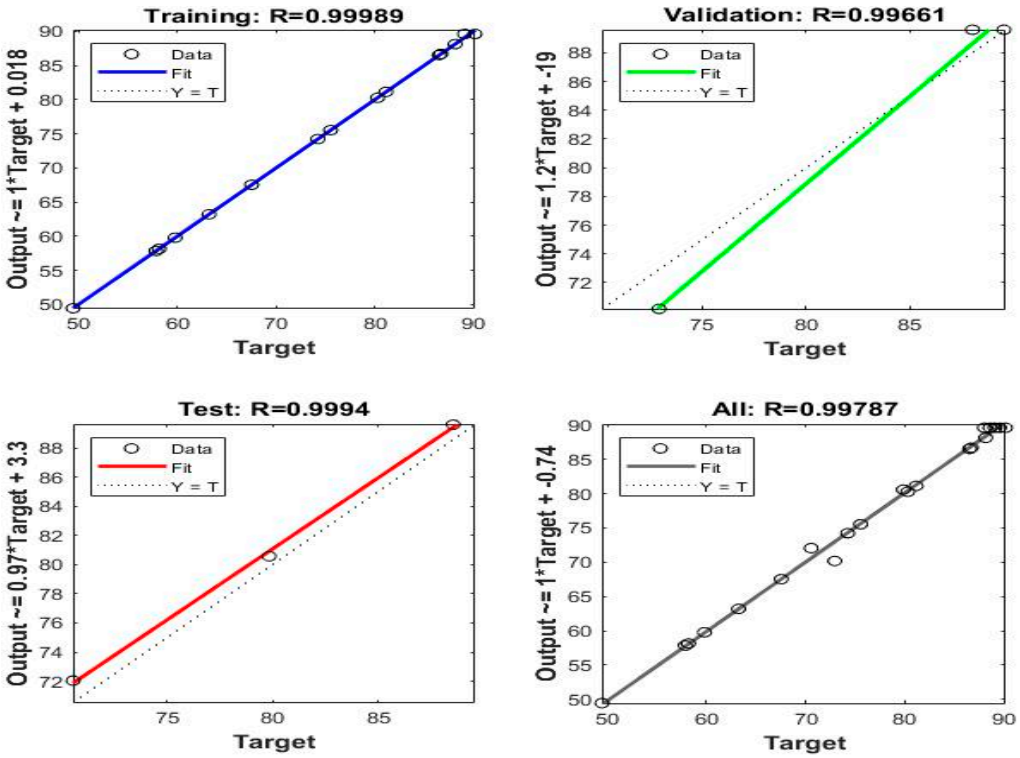

Figure 5. Training, validation, and testing for the ANN.

The ANFIS membership parameters and rules were determined using a trial and error method to determine an appropriate model that predicts the input and output data. Different tasks were performed to aid in identifying the product layer. The system created "and" rules using the four process input parameters and membership functions, sending input through the same number output membership function to convert to output. Figure 6 shows the rule viewer of the ANFIS, which shows the ability of the ANFIS to predict the yield as the optimum condition of 9 mol methanol-to-oil ratio, 0.9 wt. % catalyst-to-oil ratio, 60 min, and 50 °C, giving a yield of 89.1%, close to the one obtained by RSM optimisation at 89.09%.

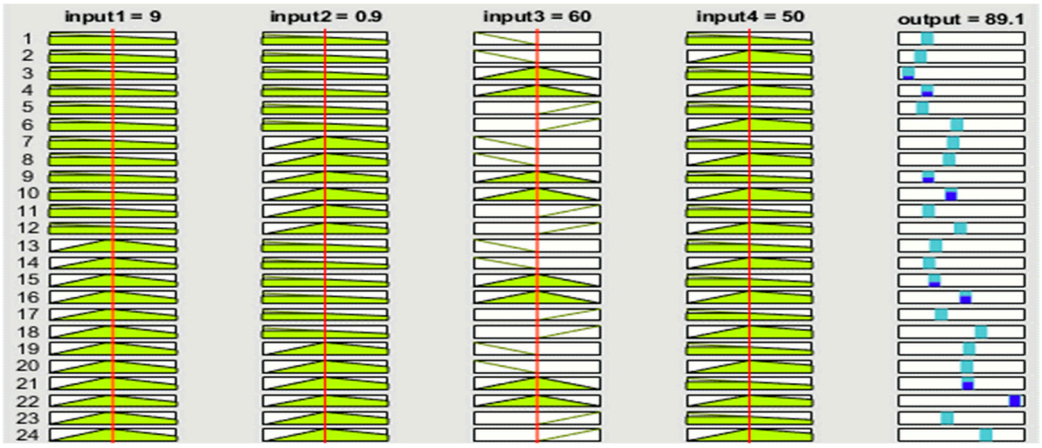

Figure 6. Rule viewer of the ANFIS model.

Figure 7 depicts the 3D response curves using the four process variables. As can be observed from Figure 7, increasing the methanol-to-oil ratio to 9 molar ratio results in an increase in biodiesel yield, but an increase beyond that resulted in a decrease in yield that can be attributed to the fact that transesterification, being a reversible reaction excess among the methanol, favours the forward reaction, but after a specific limit, biodiesel mixing with methanol causes the reverse reaction, and with the catalyst amount up to 0.9 wt. % beyond that, a decrease was observed. This behaviour can be attributed to the catalyst forming soap. A temperature of about 50 °C was enough to reach a high yield; a temperature above 60 °C could cause methanol evaporation, affecting the contact with oil and thus decreasing yield. A time of 60 min was enough to reach a high yield beyond that, which caused a slight decrease, which can be attributed to the fact that the reaction reached equilibrium, and above that, a reverse reaction can occur; the same behaviour was observed by [15].

The performance of the three ML methods used in this study was evaluated using error metrics, and the results are summarised in Table 3. From the results presented in the table, all three ML methods used presented adequate fittings with R^2, with all higher than 0.99: 0.991 for RSM, 0.996 for ANN, and 0.998 for ANFIS. ANFIS has shown slightly higher prediction with its highest R^2 and lower error functions, as shown in Table 3.

Table 3. The RSM, ANN, and ANFIS statistical error analysis.

Error Function	RSM	ANN	ANFIS
R^2	0.991	0.996	0.998
RMSE	0.924	0.566	0.324
MAE	0.568	0.267	0.202
MAPE	0.746	0.333	0.226
ARE	0.008	0.004	0.003
MPSD	4.503	2.114	1.828

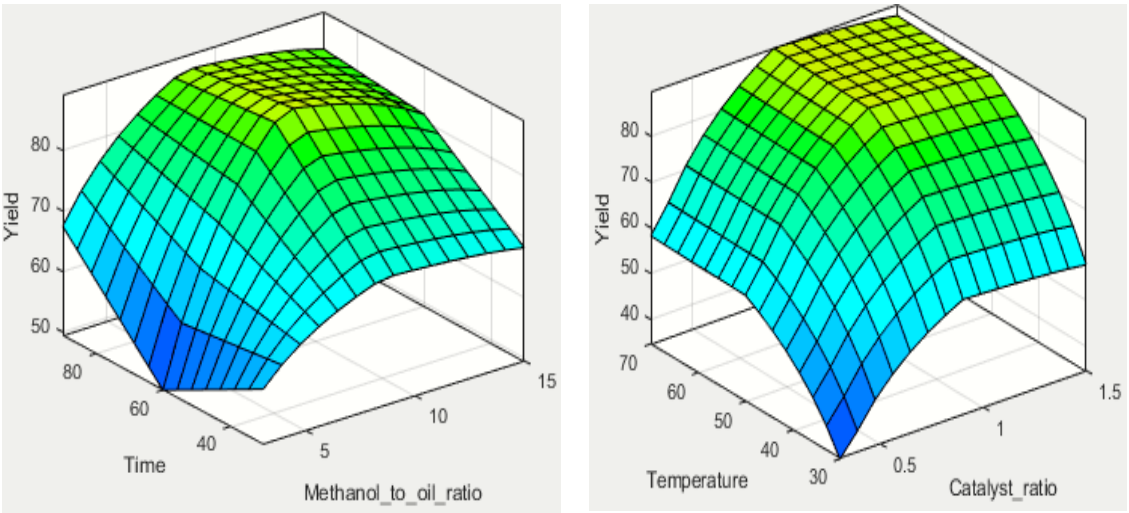

Figure 7. The 3D surface plots of biodiesel yield using the four process parameters (blue to yellow, low to high yield).

4. Conclusions

The response surface methodology (RSM), artificial neural network (ANN), and adaptive neuro-fuzzy inference system (ANFIS) were used for predictive modelling and optimisation with R^2 and ARE values of 0.991 and 0.008, 0.996 and 0.004, and 0.998 and 0.002 for RSM, ANN, and ANFIS, respectively. Four process parameters, methanol-to-oil ratio (3–15 mol), catalyst ratio (0.3–1.5 wt. %), reaction time (30–90 min), and reaction temperature (30–70 °C), were studied. The transesterification reaction catalysed by potassium hydroxide was optimised using a central composite design in the RSM, with the optimum yield obtained at a 9 mol methanol-to-oil ratio, 0.9 wt. % catalyst ratio, 60 min reaction time, and 50 °C reaction temperature, with 89.09% yield. According to the results, the developed three-dimensional machine learning approach—the RSM, ANN, and ANFIS models--is a potential method for optimising and modelling the production of biodiesel from waste margarine oil. The study results may be used to create sustainable, efficient, and economical solutions for recycling waste margarine oil.

Author Contributions: Conceptualisation, P.M.; methodology, P.M.; software, P.M.; validation, P.M., H.R., and T.S.; formal analysis, P.M., H.R., and T.S.; investigation, P.M., H.R., and T.S.; resources, P.M., H.R., and T.S.; data curation, T.S.; writing—original draft preparation, P.M.; writing—review and editing, P.M., H.R., and T.S.; visualisation, P.M., H.R., and T.S.; project administration, H.R.; funding acquisition, H.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank the Department of Chemical and Metallurgical Engineering of the Vaal University of Technology for providing research facilities.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mwenge, P.; Dlamini, N.; Rutto, H.; Seodigeng, T. Biodiesel production from waste margarine using homogeneous catalyst Potassium Hydroxide. In Proceedings of the 14th PARIS International Conference on Agriculture, Biological & Environmental Sciences (PABE-19), Paris, France, 5–7 September 2019. [CrossRef]
2. Osman, A.I.; Nasr, M.; Farghali, M.; Rashwan, A.K.; Abdelkader, A.; Al-Muhtaseb, A.H.; Ihara, I.; Rooney, D.W. Optimising biodiesel production from waste with computational chemistry, machine learning and policy insights: A review. *Environ. Chem. Lett.* **2024**, *22*, 1005–1071. [CrossRef]

3. Elendu, C.C.; Liu, C.; Aleem, R.D.; Shan, Y.; Cao, C.; Ramzan, N.; Duan, P.-G. Optimization and Deep Learning Modeling of the Yield and Properties of Baobab-Derived Biodiesel Catalyzed by Waste Banana Bunch Stalk Biochar. *ACS Omega* **2024**, *9*, 12941–12955. [CrossRef] [PubMed]
4. Mwenge, P.; Rutto, H.; Enweremadu, C. Production of Biodiesel using Calcined Brine Sludge Waste from Chor-Alkali Industry as a Heterogeneous Catalyst. *Environ. Clim. Technol.* **2021**, *25*, 621–630. [CrossRef]
5. Mwenge, P.; Seodigeng, T. In Situ Real-Time Multivariate Analysis of Methanolysis Monitoring of Sunflower Oil Using FTIR. *Int. J. Chem. Mol. Eng.* **2020**, *14*, 136–148.
6. Cubas, A.L.V.; Machado, M.M.; Pinto, C.R.S.C.; Moecke, E.H.S.; Dutra, A.R.A. Biodiesel production using fatty acids from food industry waste using corona discharge plasma technology. *Waste Manag.* **2016**, *47*, 149–154. [CrossRef] [PubMed]
7. Linganiso, E.C.; Tlhaole, B.; Magagula, L.P.; Dziike, S.; Linganiso, L.Z.; Motaung, T.E.; Moloto, N.; Tetana, Z.N. Biodiesel Production from Waste Oils: A South African Outlook. *Sustainability* **2022**, *14*, 1983. [CrossRef]
8. Schabort, C.J.; Visser, H.P.; Hartzenberg, A.; Venter, R.; Marx, S. Waste butter and margarine as feedstock for biodiesel production. In Proceedings of the 20th European Biomass Conference and Exhibition, Milan, Italy, 18–22 June 2012.
9. Gupta, K.K.; Kalita, K.; Ghadai, R.K.; Ramachandran, M.; Gao, X.-Z. Machine Learning-Based Predictive Modelling of Biodiesel Production—A Comparative Perspective. *Energies* **2021**, *14*, 1122. [CrossRef]
10. Esmi, F.; Dalai, A.K.; Hu, Y. Comparison of various machine learning techniques for modeling the heterogeneous acid-catalysed alcoholysis process of biodiesel production from green seed canola oil. *Energy Rep.* **2024**, *12*, 321–328. [CrossRef]
11. Xing, Y.; Zheng, Z.; Sun, Y.; Agha Alikhani, M. A Review on Machine Learning Application in Biodiesel Production Studies. *Int. J. Chem. Eng.* **2021**, *2021*, 2154258. [CrossRef]
12. Asaad, S.M.; Inayat, A.; Ghenai, C.; Shanableh, A. Response Surface Methodology in Biodiesel Production and Engine Performance Assessment. *Int. J. Thermofluids* **2024**, *21*, 100551. [CrossRef]
13. Ajala, O.O.; Oke, E.O.; Odejobi, O.J.; Adeoye, B.K.; Oyelade, J.O. Artificial neuro-fuzzy intelligent prediction of techno-economic parameters of computer-aided scale-up for palm kernel oil based biodiesel production. *Clean Chem. Eng.* **2023**, *5*, 100098. [CrossRef]
14. Mwenge, P.; Pilusa, J.; Seodigeng, T. Optimisation of Biodiesel Production from Sunflower Oil Using Central Composite Design. *Int. J. Chem. Mol. Eng.* **2018**, *12*, 600–605.
15. Liu, Y.; Sayed, B.T.; Sivaraman, R.; Alshahrani, S.M.; Venkatesan, K.; Thajudeen, K.Y.; Al-Bahrani, M.; Hadrawi, S.K.; Yasin, G. Novel and robust machine learning model to optimise biodiesel production from algal oil using CaO and CaO/Al_2O_3 as catalyst: Sustainable green energy. *Environ. Technol. Innov.* **2023**, *30*, 103018. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.