Automotive Experiences
Vol. 6 No. 1 (2023) pp. 4-13
p-ISSN: 2615-6202 e-ISSN: 2615-6636

Research Paper

Comparison of Various Prediction Model for Biodiesel Cetane Number using Cascade-Forward Neural Network

Sri Mumpuni Ngesti Rahaju[1], **April Lia Hananto**[2], **Permana Andi Paristiawan**[3], **Abdullahi Tanko Mohammed**[4,5], **Anthony Chukwunonso Opia**[6,7], **Muhammad Idris**[8]

[1]Department of Mechanical Engineering, Faculty of Engineering, Universitas Bung Karno, Jakarta 10320, Indonesia
[2]Faculty of Engineering and Computer Science, Universitas Buana Perjuangan Karawang, Karawang 41361, Indonesia
[3]Research Center for Metallurgy, National Research and Innovation Agency, South Tangerang 15314, Indonesia
[4]Department of Mechanical Engineering, Waziri Umaru Federal Polytechnic Birnin Kebbi, Nigeria
[5]Automotive Development Centre, Institute for Vehicle Systems and Engineering, Universiti Teknologi Malaysia, 81310, Johor Bahru, Malaysia
[6]Department of Marine Engineering, Niger Delta University, Wilberforce Island, Nigeria
[7]Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
[8]School of Environmental Science, University of Indonesia, Jakarta 10430, Indonesia

aprilia@ubpkarawang.ac.id

https://doi.org/10.31603/ae.7050

Check for updates

Published by Automotive Laboratory of Universitas Muhammadiyah Magelang collaboration with Association of Indonesian Vocational Educators (AIVE)

Article Info
Submitted:
05/05/2022
Revised:
12/06/2022
Accepted:
30/06/2022
Online first:
01/01/2023

Abstract

Cetane number (CN) is one of the important fuel properties of diesel fuels. It is a measurement of the ignition quality of diesel fuel. Numerous studies have been published to predict the CN of biodiesels. More recently, the utilization of soft computing methods such as artificial neural networks (ANN) has received considerable attention as a prediction tool. However, most studies in the use of ANN for estimating the CN of biodiesels have only used one algorithm to train a small number of datasets. This study aims to predict the CN of 63 biodiesels based on the fatty acid methyl esters (FAME) composition by developing an ANN model that was trained with 10 different algorithms. To the best of our knowledge, this is the first study to predict the CN of biodiesels using numerous ANN training algorithms utilizing sizeable datasets. Results revealed that the ANN model trained with Levenberg-Marquardt gave the highest prediction accuracy. LM algorithm successfully predicted the CN of biodiesels with the highest correlation and determination coefficient (R = 0.9615, R2 = 0.9245) as well as the lowest errors (MAD = 2.0804, RMSE = 3.1541, and MAPE = 4.2971). Hence, the Cascade neural network trained with the LM algorithm could be considered a promising alternative to the empirical correlations for predicting biodiesel's CN.

Keywords: Biodiesel; Cetane number; Cascade neural network; Artificial neural network; Fuel properties; FAME

1. Introduction

Despite the emerging trend of electric vehicles [1], [2], the internal combustion engine (ICE) remains the most dominant power engine due to its mature technology [3]–[6]. However, it is important to note that ICE needs fuel from fossil hydrocarbons that are rare and not available in large amounts, making it costly to produce [7]. Petroleum-based fuel is commonly known for its derivation products such as gasoline, diesel, and jet fuel [8]. Regrettably, fossil fuel is still dominating the transportation sectors worldwide as the primary energy source since it is suitable for most engine applications [9].

Furthermore, the increasing population, expanding urbanization, and better living

standards have raised global energy consumption [10]. On the contrary, fossil fuel reserves have been diminishing more and more in the last two decades. Environment concerns and anxiety over the decreasing fossil fuel have thus made alternative fuel to be a feasible option to meet future energy demands [11]–[13].

The esters from vegetable oil can be used as the energy source for diesel engines and, due to its natural ingredient, is often referred to as biodiesel [14]. Biodiesel is among the alternative fuel that has gained popularity globally [15]. The use of such fuel is being encouraged by many countries to replace conventional diesel fuel with fossil. The main objective is to decrease dependency on foreign oil as well as to reduce the greenhouse effect. Biodiesel is believed to be a non-toxic biodegradable fuel that can be used in diesel engines without major modification. In addition, fuel derived from biodiesel processes has emerged as renewable and non-toxic compared to conventional diesel fuel. Neither net carbon dioxide nor sulfur is released from such fuel, emitting fewer pollutants emissions to the environment.

During the late 90s, soybean oil was primarily used in the US, being the only oil that is sufficiently available to the national demand. Yet, the production cost to produce this edible vegetable oil was expensive. As a result, it was only used in cases of the serious scarcity of petroleum diesel fuel. To be commercially visible, it is important to reduce the cost of the feedstock. One possible way to make it more affordable is to use feedstocks that are less expensive. Used oils, greases, and animal fats are good sources of economical biodiesel.

Many, however, are concerned with the issue of increasing food and water prices if biodiesel is produced from edible sources. Hence, second-generation biofuels are developed, being made from waste materials. Waste cooking oil, for example, is abundant in most parts of the world. In fact, biodiesel from used cooking oil has been produced commercially in some countries. Non-edible vegetable oils are gaining popularity as promising substitutions for conventional edible food crops. With the enormous demand for edible oil as a food source, the deploying of such non-edible plant oils plays an important role to solve the food vs fuel dilemma.

Low-cost sources such as oils from restaurants have been widely used recently with regard to edible feedstocks. Nevertheless, producing produce fuel-grade biodiesel from this low-cost oil is far more difficult as they encompass a high level of free fatty acids. Various studies have indicated that the use of waste cooking oil biodiesel reduces a considerable amount of particulate matter, CO, and total hydrocarbon emissions with no efficiency loss in several types of diesel engines in comparison to conventional diesel fuel. Some studies have also shown that NOx emissions rise slightly. Note that sharp rises in both NOx and PM emissions may occur when cooking oil biodiesel fuel is used.

Many researchers have pointed out that biodiesel's properties bear a close resemblance to those of petroleum-based diesel fuel. The studies of biodiesel in optimization models have been analyzed and discussed by numerous researchers [16]–[18]. As a result, biodiesel can be widely used in diesel engines with little or no alteration. Some considerations, however, need to be addressed. Take flash point properties for example. The flash point is a substantial safety measure of fuel flammability. Biodiesel tends to have flash point that is two times higher than conventional diesel fuel. This, fortunately, can be solved by adding residual alcohol to lower its flash point. Another fuel property to be considered is viscosity. It affects the volume flow and injection spray characteristics of an engine. Biodiesel kinematic viscosity is known to be higher than that of diesel fuel. Consequently, biodiesel at low temperatures can be very viscous or even solidified, compromising the mechanical integrity of the injection pump drive system. Nevertheless, biodiesel still provides superior advantages, namely, relatively higher CN, no aromatics, and contains 10% to 11% oxygen by weight. These, consequently, decrease the emissions of carbon monoxide (CO), hydrocarbons (HC), and particulate matter (PM).

CN, in particular, is an indicator of the ignition quality of diesel fuel. The greater the number, the better the fuel burns. The CN is comparable to the octane number for gasoline fuel as a measurement of combustion quality. The major difference between cetane and octane numbers is that the octane number measures the ability of a fuel to resist unwanted initial ignition caused by compression so that the fuel only ignites from a

spark plug. By contrast, the CN measures the ignition delay of a fuel. Therefore, a higher-performance gasoline engine needs fuel with a greater octane number, whereas a high-performance diesel vehicle needs fuel with a greater CN to enable fast ignition and avoid incomplete combustion.

2. Novelty of the Study

Regardless Regardless of the feedstock sources, it is important to remember that fuel properties of biodiesel such as density, kinematic viscosity, and CN should meet certain standards like ASTM. As most fuel properties are significantly reliant on the compositions of the feedstock, it is therefore different feedstock will have different fuel properties.

Many previous studies have been published to estimate biodiesel CN. More recently, the utilization of soft computing methods such as artificial neural networks (ANN) has received considerable attention as a prediction tool [19]–[21]. However, most studies in the use of ANN for estimating the CN of biodiesels have only used less algorithms to train small number of datasets. Hence, by greater number of data, it would give a better prediction of CN. In addition, this study proposed 10 different techniques (algorithm) to predict the CN.

The 10 proposed techniques are not the new approach in ANN. However, to compare many techniques (more than 5) have not been considered yet previously in the literatures. Additionally, there was a lack of information on the best techniques. To find the best technique, many techniques shall be tested and compared.

This study aims to predict the CN of 63 biodiesels based on the fatty acid methyl esters (FAME) composition by developing an ANN model that was trained with 10 different algorithms as shown in **Figure 1**. To the best of our knowledge, this is the first study to predict the CN of biodiesel using numerous ANN training algorithms utilizing sizeable datasets.

3. Method

3.1. Data Gathering

Data from 63 biodiesel CNs along with their fatty acid methyl esters (FAME) composition have been gathered from previous studies [22]–[26] and are presented in **Table 1**. The dataset has values of CN ranging from 20.4 to 86.9.

3.2. Cascade Neural Network

Previous studies have noted the importance of artificial neural networks [27], [28]. However, numerous machine learnings have been studied for prediction models or other purposes, such as response surface method (RSM), surrogate model, and metamodelling.

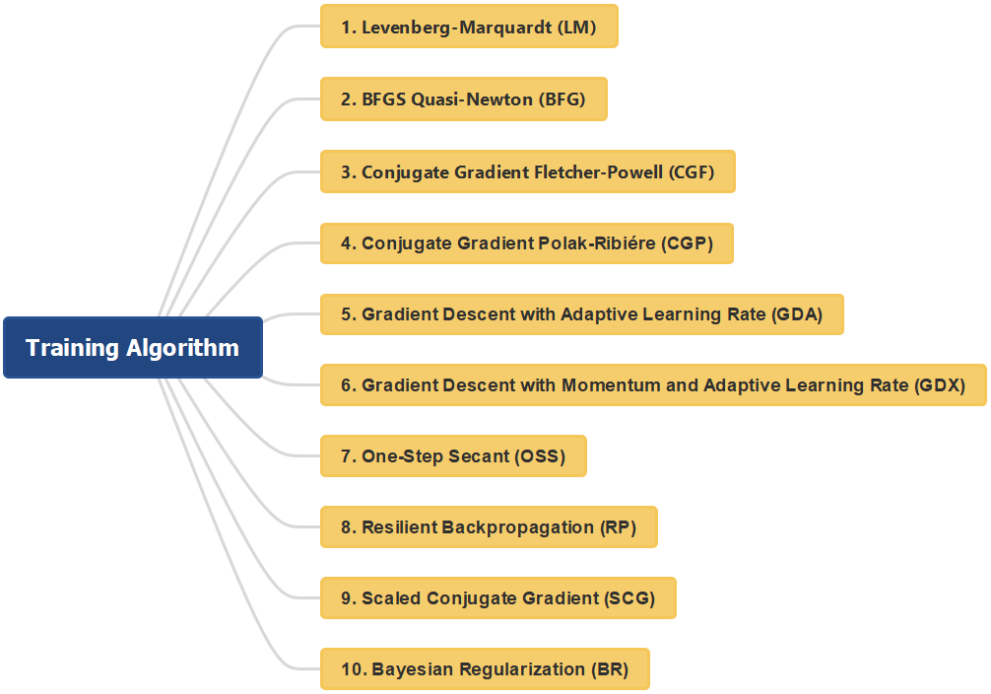

Figure 1. Ten algorithms used in this study for cascade neural network

Table 1. FAME composition of different biodiesel sources

Biodiesel	12:00	14:00	16:00	16:01	18:00	18:01	18:02	18:03	20:01	22:01	CN
Lauric	100	0	0	0	0	0	0	0	0	0	61.4
Myristic	0	100	0	0	0	0	0	0	0	0	66.2
Palmitic	0	0	100	0	0	0	0	0	0	0	74.5
Stearic	0	0	0	0	100	0	0	0	0	0	86.9
Oleic	0	0	0	0	0	100	0	0	0	0	55
Palmitoleic	0	0	0	100	0	0	0	0	0	0	51
Linoleic	0	0	0	0	0	0	100	0	0	0	42.2
Erucic	0	0	0	0	0	0	0	0	0	100	76
Eicosanoic	0	0	0	0	0	0	0	0	100	0	64.8
Linolenic	0	0	0	0	0	0	0	100	0	0	20.4
Soybean	0	0.1	10.5	0.1	3.7	23.2	48.9	1.2	0.3	0.1	47.2
Inedible tallow	0.1	2.1	23.9	2.8	19.5	38.5	6.4	0.3	0.5	0.1	61.7
Thevetia peruviana M.	0	0	15.6	0	10.5	60.9	5.2	7.4	0	0	57.5
Moringa oleifera Lam	0	0	9.1	2.1	2.7	79.4	0.7	0.2	0	0	56.7
Pongamia pinnata P.	0	0	10.6	0	6.8	49.4	19	0	2.4	0	55.8
Holoptelia integrifolia	0	3.5	35.1	1.9	4.5	53.3	0	0	0	0	61.2
Vallaris solanacea K.	0	0	7.2	0	14.4	35.3	40.4	0	0	0	50.3
Aleurites moluccana	0	0	5.5	0	6.7	10.5	48.5	28.5	0	0	34.2
Euphorbia helioscopia L	2.8	5.5	9.9	0	1.1	15.8	22.1	42.7	0	0	34.2
Garnicia morella D.	0	0	0.7	0	46.4	49.5	0.9	0	0	0	63.5
Saturega hortensis Linn	0	0	0.4	0	0.4	12	18	62	0	0	25.5
Actinodaphne angust.	87.9	1.9	0.5	0	5.4	0	0	0	0	0	63.2
Litsea glutinosa Robins	96.3	0	0	0	0	2.3	0	0	0	0	64.8
Neolitsea cassia Linn	85.9	3.8	0	0	0	4	3.3	0	0	0	64
Swietenia mahagoni J.	0	0	9.5	0	18.4	56	0	16.1	0	0	52.3
Argemone mexicana	0	0.8	14.5	0	3.8	18.5	61.4	0	0	0	44.5
Salvadora persica Linn	19.6	54.5	19.6	0	0	5.4	0	0	0	0	67.5
Madhuca butyracea M.	0	0	66	0	3.5	27.5	3	0	0	0	65.3
Rhus succedanea Linn	0	0	25.4	0	0	46.8	27.8	0	0	0	52.2
Basella rubra Linn	0	0.4	19.7	0.4	6.5	50.3	21.6	1.1	0	0	54
Corylus avellana	0	3.2	3.1	0	2.6	88	2.9	0	0	0	54.5
Jatropa curcas Linn	0	1.4	15.6	0	9.7	40.8	32.1	0	0	0	52.3
Croton tiglium Linn	0	11	1.2	0	0.5	56	29	0	0	0	49.9
Princeptia utilis Royle	0	1.8	15.2	0	4.5	32.6	43.6	0	0	0	48.9
Vernonia cinerea Less	0	8	23	0	8	32	22	0	0	0	57.5
Joannesia princeps V.	0	2.4	5.4	0	0	45.8	46.4	0	0	0	45.2
Garcinia combogia D.	0	0	2.3	0	38.3	57.9	0.8	0.4	0.3	0	61.5
Garcinia indica Choisy	0	0	2.5	0	56.4	39.4	1.7	0	0	0	65.2
Illicium verum Hook	0	4.4	0	0	7.9	63.2	24.4	0	0	0	50.7
Melia azadirach Linn	0	0.1	8.1	1.5	1.2	20.8	67.7	0	0	0	41.4
Myristica malabarica L	0	39.2	13.3	0	2.4	44.1	1	0	0	0	61.8
Urtica dioica Linn	0	0	9	0	0	14.6	73.7	2.7	0	0	38.7
Tectona grandis Linn	0	0.2	11	0	10.2	29.5	46.4	0.4	0	0	48.3
Canola	0	0.1	5.2	0.2	2.5	58.1	28.1	0.4	1.6	0.4	55
Lard	0.1	1.9	24.5	2.8	14.4	38.3	13.4	0.3	0.7	0.1	63.6
Yellow grease	0	1.1	17.3	2.2	9.5	45.3	14.5	1.3	1.3	0	52.9
Rape	0	0	4.8	0	1.6	33	20.4	7.9	9.3	23	55
Linseed	0	0	5	0	2	20	18	55	0	0	52
Wild mustard	0	0.1	2.6	0.2	0.9	7.8	14.2	13	5.4	45.7	61.1
Waste palm oil	0	1	39	0.2	4.3	43.7	10.5	0.2	0.2	0	60.4
Balanites roxburhii	0	0	17	4.3	7.8	32.4	31.3	7.2	0	0	50.5
Garnicia echinocarpa	0	0	3.7	0	43.7	52.6	0	0	0	0	63.1
Neolitsea umbrosa G.	59.1	11.5	0	0	0	21	6.7	0	0	0	60.8
Anamirta cocculus	0	0	6.1	0	47.5	46.4	0	0	0	0	64.3
Broussonetia p. Vent.	0	0	4	0	6.1	14.8	71	1	0	0	41.2
Salvadora oleoiles D.	35.6	50.7	4.5	0	0	8.3	0.1	0	0	0	66.1
Nephelium L.	0	0	0.2	0	13.8	45.3	0	0	4.2	0	64.9
Ziziphus maurit. L.	0	0	10.4	0	5.5	64.4	12.4	0	2.6	1.7	55.4
Jojoba	0	0	1.2	0	0	10.7	0.1	0.4	59.5	12.3	69
Rape	0	0	4.9	0	1.6	33	20.4	7.8	9.3	23	55
Peanut	0	0.1	8	0	1.8	53.3	28.4	0.3	2.4	0	53
Grape	0	0.1	6.9	0.1	4	19	69.1	0.3	0	0	48
Sunflower	0	0	6	0.1	2.9	17	74	0	0	0	49

The previous techniques provide a better result in small dataset. However, design of experiment (DOE) is required for those techniques. If the techniques is applied with random dataset

(scattered data), less accuracy will be achieved. Therefore, ANN give better result for random dataset.

Cascade neural network, in particular, is a neural network class that is analogous to feed-forward networks. Yet, the cascade neural network has a connection from the input as well as each preceding layer to following layers. In a network with three layers, for example, the output layer is directly connected with the input and hidden layer. Similar to feed-forward networks, a two or extra layer cascade neural network can learn the input-output relationship randomly well using sufficient hidden neurons. Cascade neural network can be utilised for any type of mapping from input to output. The benefit of this technique is that it contains the nonlinear input-output relationship by not disregarding such linear relationship. **Figure 2** shows the cascade neural networks architecture. The equation of cascade neural network is following Eq. 1.

Here, f^I and W_i^I are the activation function and weight from the input to the output layer, respectively. When there is an additional bias, Eq. 1 becomes Eq. 2.

$$y = \sum_{i=1}^{n} f^i W_i^i x_i + f^o\left(\sum_{j=1}^{k} W_j^o f_j^h\left(\sum_{i=1}^{n} W_{ji}^h x_i\right)\right) \quad (1)$$

$$y = \sum_{i=1}^{n} f^i W_i^i x_i + f^o\left(W^b + \sum_{j=1}^{k} W_j^o f_j^h\left(W_j^b + \sum_{i=1}^{n} W_{ji}^h x_i\right)\right) \quad (2)$$

In the present study, the data were distributed randomly by 70%, 15% and 15% for training, test and validation. Log-sigmoid (logsig) was set as the hidden layer transfer functions, whereas the linear (purelin) was set as the output layer. The flowchart of cascade neural network used in the present study is given in **Figure 3**.

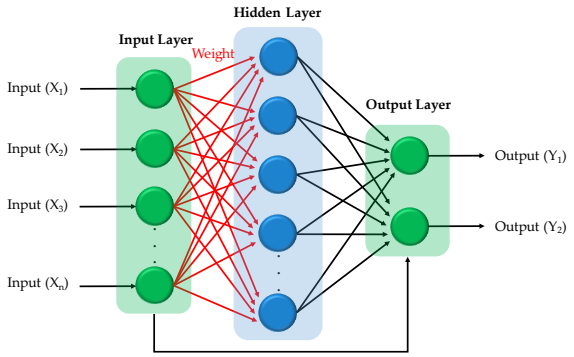

Figure 2. The general architecture of cascade neural networks

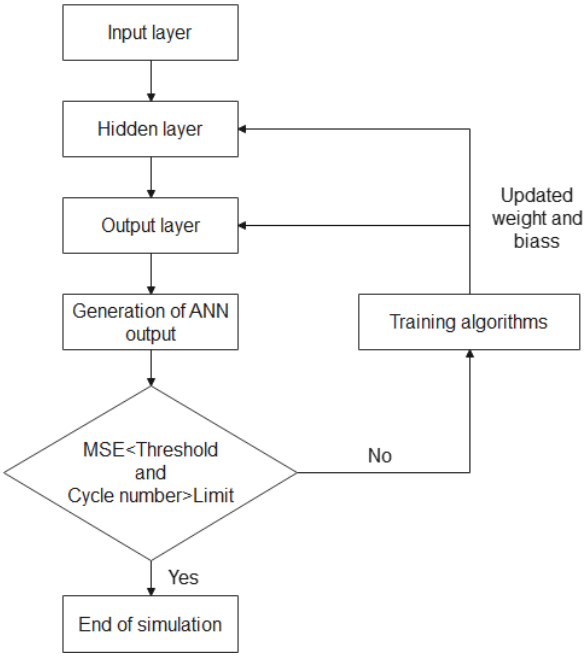

Figure 3. Flowchart of cascade network algorithm

3.3. *Performance Criteria*

Five different criteria were used; R, R^2, MAD, RMSE and MAPE to evaluate the prediction accuracy of each model using the following Eq. 3-Eq. 7.

Correlation Coefficient (R),

$$R = \sqrt{1 - \left\{\frac{\sum_{i=1}^{n}(M_i - P_i)^2}{\sum_{i=1}^{n} P_i^2}\right\}} \quad (3)$$

Determination Coefficient (R^2),

$$R^2 = 1 - \left\{\frac{\sum_{i=1}^{n}(M_i - P_i)^2}{\sum_{i=1}^{n} P_i^2}\right\} \quad (4)$$

Mean Absolute Deviation (MAD),

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|M_i - P_i| \quad (5)$$

Root Mean Squared Error (RMSE),

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - P_i)^2} \quad (6)$$

Mean Absolute Percentage Error (MAPE),

$$MAPE = \left\{\frac{100}{n}\sum_{i=1}^{n}\left|\left(\frac{M_i - P_i}{M_i}\right)\right|\right\}\% \quad (7)$$

4. Results and Discussion

Figure 4 shows the coefficient of correlation (R) and determination (R^2) for 10 different algorithms in cascade neural network examined in this study. R measures the association closeness of the points to a line of linear regression according to such points. Possible values of R varied from -1 to +1,

with -1 showing a perfectly linear inverse correlation and +1 showing that of positive correlation. Detailed R values of Cascade-LM algorithms for training, validation, test and all are presented in **Figure 5**. Moreover, the determination coefficient is a prediction criterion used to quantify how much variability of one particular factor can be triggered by its connection to another associated factor. The determination correlation is also known as the "goodness of fit," whose value ranging from 0 and 1, with value of 1 indicating a perfect fit, and is therefore a substantially reliable model for future predictions, while 0 indicating that the estimate does not accurately model the data at all.

The highest value of correlation and determination coefficient is given by Levenberg Marquardt algorithm. With an R value of 0.9615, LM algorithm signifies almost a perfectly linear positive correlation. Furthermore, the LM algorithm with a R^2 value of 0.9245 suggests that 92.45% of the dependent variable is predicted well by the independent variable. On the contrary, the lowest value of correlation and determination coefficient is given by Gradient Descent with Momentum and Adaptive Learning Rate (GDX) with an R value of 0.7885 and R^2 of 0.6217.

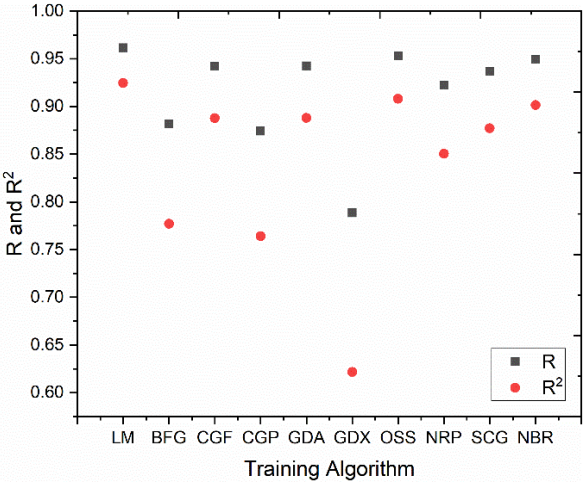

Figure 4. R and R^2 values for the ten investigated training algorithms

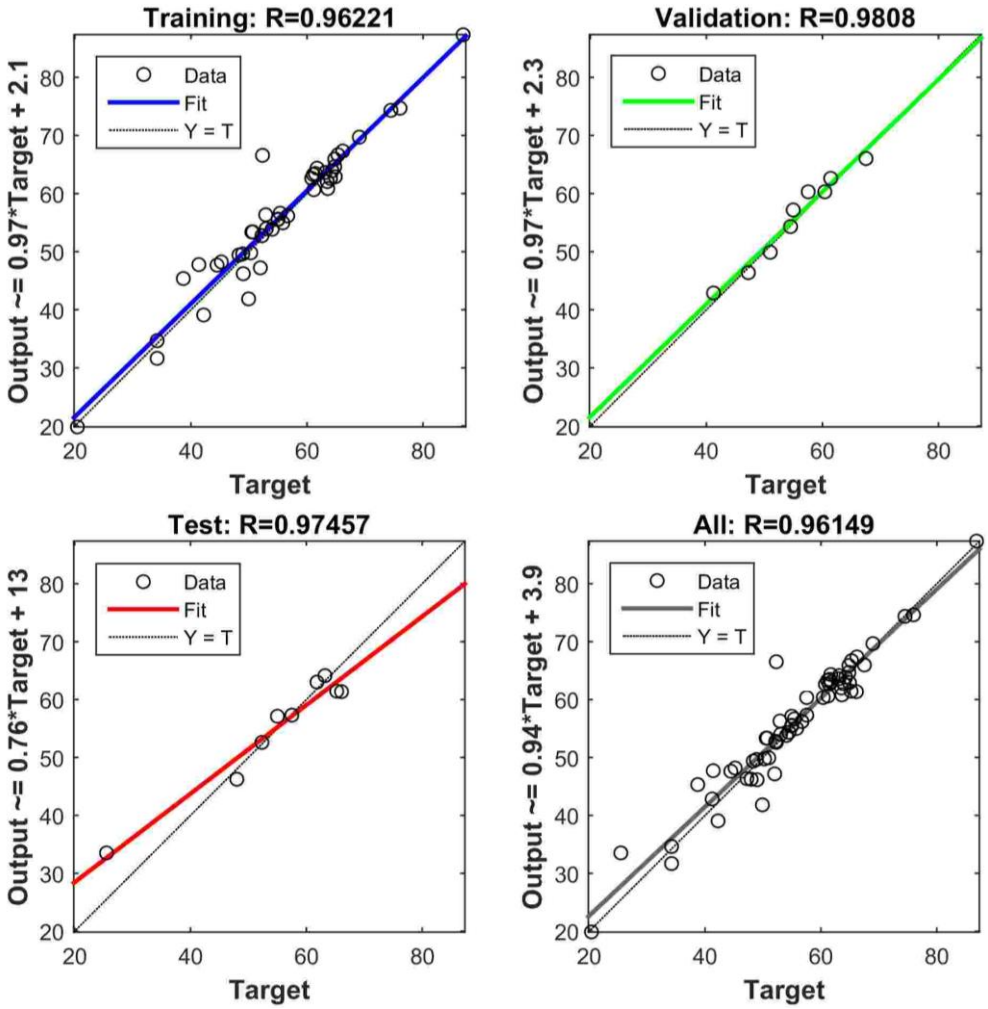

Figure 5. R values of Cascade-LM algorithms for training, validation, test and all

Figure 6 shows the value of MAD, RMSE and MAPE for all 10 training algorithms in cascade neural network examined in this study. MAD is a variability measure which denotes the average distance between examinations and their mean values. The term of the mean absolute deviation resonances to the standard deviation (SD). Both MAD and SD measure variability, but they have different calculations. Due to its relatively straightforward concept that better fits real life application, MAD is favoured to replace the standard deviation. Another method to evaluate how well a regression model fits a dataset is by calculating the RMSE. It is a metric that signify the average distance between the estimated values from the model and the real dataset values. The last prediction accuracy used in this study is MAPE. It measures a forecast system accuracy as a percentage. MAPE is considered as the most frequent measure used to predict error, working at its best when there are no extremes and no zeros.

The lowest value of MAD, RMSE and MAPE is given by Levenberg Marquardt algorithm, while that of the highest is given by Gradient Descent with Momentum and Adaptive Learning Rate (GDX). LM's lowest MAD value shows that data points bunching closer to its average value. Conversely, the highest MAD value of GDX indicates that the data points in Gradient Descent with Momentum and Adaptive Learning stretch further from the average values. LM's lowest RMSE value indicates the best prediction accuracy among of an Levenberg Marquardt among other investigated training algorithm to fit well a dataset. As far as the mean absolute percent error is concerned, the MAPE value of LM is 4.2971% meaning that the prediction is off by just over 4%. Note that since the MAPE is a percentage, it is far easier to comprehend its value compared to other accuracy measure.

Detailed results of R, R^2, MAD, RMSE, and MAPE are provided in **Table 2**. Furthermore, the best results provided by Levenberg Marquardt algorithm is comprehensively provided in **Figure 7** where actual and predicted CN for various test cases are presented. Looking at **Table 2** and **Figure 7**, it is apparent that the LM and GDX training algorithm shows the best and the worst prediction accuracy, respectively.

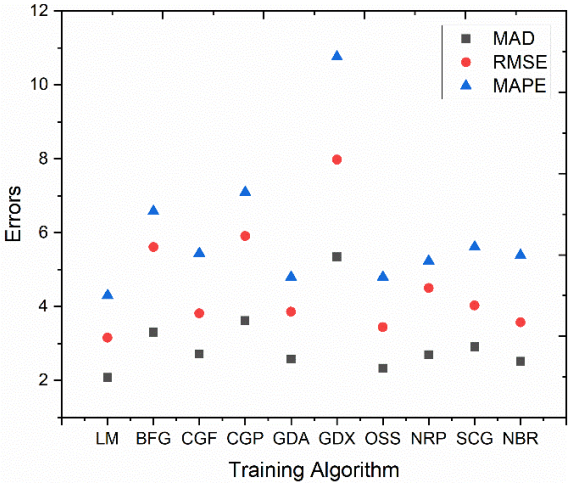

Figure 6. Errors for the ten investigated training algorithms

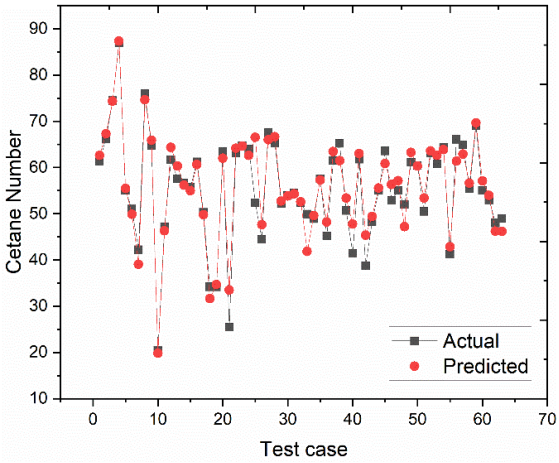

Figure 7. Actual and predicted CN for various test cases using Cascade Levenberg-Marquardt training algorithm

There are a number of explanations for the above results. Since there are two possible possibilities for the direction at every iteration, the LM is more robust. Levenberg Marquardt is also faster to converge than gradient descent. Moreover, LM is able to handle models with more than free parameters that are not exactly known. If the initial estimation is far-off from the mark, the LM algorithm is still able to discover the optimal solution. However, the Levenberg Marquardt may sometimes suffer from several drawbacks. For flat functions, the LM algorithm may become lost in parameter space. In several cases, such training algorithm can become exceptionally slow to converge. This is predominantly true when the model has more than ten parameters that needs the algorithm to move slowly along a closely well-defined crawl space.

Table 2. Performance evaluation of ANN models

ANN training algorithm	R	R^2	MAD	MSE	RMSE	MAPE
Cascade-LM	0.9615	0.9245	2.0804	9.9486	3.1541	4.2971
Cascade-BFG	0.8815	0.7770	3.3011	31.4749	5.6103	6.5837
Cascade-CGF	0.9422	0.8877	2.7139	14.5417	3.8134	5.4417
Cascade-CGP	0.8741	0.7641	3.6203	34.9113	5.9086	7.0998
Cascade-GDA	0.9423	0.8880	2.5798	14.9114	3.8615	4.7941
Cascade-GDX	0.7885	0.6217	5.3485	63.6301	7.9768	10.7621
Cascade-OSS	0.9528	0.9079	2.3282	11.8651	3.4446	4.7966
Cascade-RP	0.9222	0.8504	2.6972	20.2798	4.5033	5.2346
Cascade-SCG	0.9365	0.8771	2.9094	16.2064	4.0257	5.6135
Cascade-BR	0.9495	0.9015	2.5191	12.7715	3.5737	5.3889

5. Conclusion

In this study, a dataset of 63 samples of biodiesel was used to build an ANN model. To predict the CN of biodiesels in terms of their fatty acid methyl esters composition, 10 different ANN training algorithms using a 10-10-1 network structure were compared and analyzed. Results revealed that the ANN trained with Levenberg-Marquardt gave the highest accuracy. This model successfully predicted the CN of biodiesels with the highest correlation and determination coefficient ($R = 0.9615$, $R^2 = 0.9245$) as well as the lowest errors (MAD = 2.0804, RMSE = 3.1541, and MAPE = 4.2971).

CN determination by means of experimental work is a relatively expensive and laborious process. For that reason, developing accurate models for predicting biodiesel's CN from its FAME composition for various feedstocks would be beneficial. In the present study, we have shown that the Cascade neural network trained with the Levenberg-Marquardt algorithm using a 10-10-1 network structure was considered a promising alternative to the expensive experimental or empirical correlations for predicting biodiesel's CN. This promising alternative has a potential approach for further research by continuing the prediction model of CN with larger data.

Author's Declaration

Authors' contributions and responsibilities

The authors made substantial contributions to the conception and design of the study. The authors took responsibility for data analysis, interpretation and discussion of results. The authors read and approved the final manuscript.

Funding

No funding information from the authors.

Availability of data and materials

All data are available from the authors.

Competing interests

The authors declare no competing interest.

Additional information

No additional information from the authors.

References

[1] I. Veza, M. F. Roslan, M. F. M. Said, and Z. A. Latiff, "Potential of range extender electric vehicles (REEVS)," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 884, no. 1, p. 12093, doi: 10.1088/1757-899X/884/1/012093.

[2] I. Veza *et al.*, "Electric Vehicles in Malaysia and Indonesia: Opportunities and Challenges," *Energies*, vol. 15, no. 7, p. 2564, 2022, doi: 10.3390/en15072564.

[3] A. Katijan, M. F. A. Latif, Q. F. Zahmani, S. Zaman, K. A. Kadir, and I. Veza, "An experimental study for emission of four stroke carbureted and fuel injection motorcycle engine," *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, vol. 62, no. 2, pp. 256–264, 2019.

[4] M. F. Roslan, I. Veza, and M. F. M. Said, "Predictive simulation of single cylinder n-butanol HCCI engine," in *IOP conference series: Materials science and engineering*, 2020, vol. 884, no. 1, p. 12099, doi: 10.31224/osf.io/y53bn.

[5] I. Veza *et al.*, "Strategies to Form Homogeneous Mixture and Methods to Control Auto-Ignition of HCCI Engine," *International Journal of Automotive and Mechanical Engineering*, vol. 18, no. 4, pp. 9253–9270, 2021, doi: 10.15282/ijame.18.4.2021.09.0712.

[6] I. Veza, M. F. M. Said, Z. A. Latiff, M. F. Hasan, R. I. A. Jalal, and N. M. I. N. Ibrahim,

"Simulation of predictive kinetic combustion of single cylinder HCCI engine," in *AIP Conference Proceedings*, 2019, vol. 2059, no. 1, p. 20017, doi: 10.1063/1.5085960.

[7] M. Q. Rusli *et al.*, "Performance and emission measurement of a single cylinder diesel engine fueled with palm oil biodiesel fuel blends," in *IOP Conference Series: Materials Science and Engineering*, 2021, vol. 1068, no. 1, p. 12020, doi: 10.1088/1757-899X/1068/1/012020.

[8] I. Veza, M. F. Roslan, M. F. M. Said, Z. A. Latiff, and M. A. Abas, "Physico-chemical properties of Acetone-Butanol-Ethanol (ABE)-diesel blends: Blending strategies and mathematical correlations," *Fuel*, vol. 286, p. 119467, 2021, doi: 10.1016/j.fuel.2020.119467.

[9] I. Veza, V. Muhammad, R. Oktavian, D. W. Djamari, and M. F. M. Said, "Effect of COVID-19 on biodiesel industry: A case study in Indonesia and Malaysia," *International Journal of Automotive and Mechanical Engineering*, vol. 18, no. 2, pp. 8637–8646, 2021, doi: 10.15282/ijame.18.2.2021.01.0657.

[10] I. Veza, M. F. M. Said, and Z. A. Latiff, "Improved performance, combustion and emissions of SI engine fuelled with butanol: A review," *International Journal of Automotive and Mechanical Engineering*, vol. 17, no. 1, pp. 7648–7666, 2020, doi: 10.15282/ijame.17.1.2020.13.0568.

[11] M. N. Atique *et al.*, "Hydraulic characterization of Diesel, B50 and B100 using momentum flux," *Alexandria Engineering Journal*, vol. 61, no. 6, pp. 4371–4388, 2022, doi: 10.1016/j.aej.2021.09.064.

[12] H. M. Khan, T. Iqbal, M. A. Mujtaba, M. E. M. Soudagar, I. Veza, and I. M. R. Fattah, "Microwave assisted biodiesel production using heterogeneous catalysts," *Energies*, vol. 14, no. 23, p. 8135, 2021, doi: 10.3390/en14238135.

[13] U. Rajak, Ü. Ağbulut, I. Veza, A. Dasore, S. Sarıdemir, and T. N. Verma, "Numerical and experimental investigation of CI engine behaviours supported by zinc oxide nanomaterial along with diesel fuel," *Energy*, vol. 239, p. 122424, 2022, doi: 10.1016/j.energy.2021.122424.

[14] A. T. Mohammed *et al.*, "Soil fertility enrichment potential of Jatropha curcas for sustainable agricultural production: A case study of Birnin Kebbi, Nigeria," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 4, pp. 21061–21073, 2021, doi: 10.1016/j.fuproc.2019.106179.

[15] I. Veza *et al.*, "Multi-objective optimization of diesel engine performance and emission using grasshopper optimization algorithm," *Fuel*, vol. 323, p. 124303, 2022, doi: 10.1016/j.fuel.2022.124303.

[16] A. Kolakoti and G. Satish, "Biodiesel production from low-grade oil using heterogeneous catalyst: an optimisation and ANN modelling," *Australian Journal of Mechanical Engineering*, pp. 1–13, 2020, doi: 10.1080/14484846.2020.1842298.

[17] A. Kolakoti, M. Setiyo, and B. Waluyo, "Biodiesel Production from Waste Cooking Oil: Characterization, Modeling and Optimization," *Mechanical Engineering for Society and Industry*, vol. 1, no. 1, pp. 22–30, 2021, doi: 10.31603/mesi.5320.

[18] S. Satya, A. Kolakoti, and R. Rao, "Optimization of palm methyl ester and its effect on fatty acid compositions and cetane number," *Mathematical Models in Engineering*, vol. 5, no. 1, pp. 25–34, 2019, doi: 10.21595/mme.2019.20469.

[19] I. Veza *et al.*, "Review of artificial neural networks for gasoline, diesel and homogeneous charge compression ignition engine," *Alexandria Engineering Journal*, vol. 61, no. 11, pp. 8363–8391, 2022, doi: 10.1016/j.aej.2022.01.072.

[20] S. Suryanarayanan, V. M. Janakiraman, J. Sekar, G. Lakshmi, and N. Rao, "Prediction of cetane number of a biodiesel based on physical properties and a study of their influence on cetane number," in *2007 Fuels and Emissions Conference*, 2007, no. 2007-01–0077, doi: 10.4271/2007-01-0077.

[21] R. Piloto-Rodríguez, Y. Sánchez-Borroto, M. Lapuerta, L. Goyos-Pérez, and S. Verhelst, "Prediction of the cetane number of biodiesel using artificial neural networks and multiple linear regression," *Energy Conversion and Management*, vol. 65, pp. 255–261, 2013, doi: 10.1016/j.enconman.2012.07.023.

[22] G. Knothe, "Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters," *Fuel processing technology*, vol. 86, no. 10, pp. 1059–1070, 2005, doi: 10.1016/j.fuproc.2004.11.002.

[23] M. M. Azam, A. Waris, and N. M. Nahar, "Prospects and potential of fatty acid methyl esters of some non-traditional seed oils for use as biodiesel in India," *Biomass and bioenergy*, vol. 29, no. 4, pp. 293–302, 2005, doi: 10.1016/j.biombioe.2005.05.001.

[24] D. Tong, C. Hu, K. Jiang, and Y. Li, "Cetane number prediction of biodiesel from the composition of the fatty acid methyl esters," *Journal of the American Oil Chemists' Society*, vol. 88, no. 3, pp. 415–423, 2011, doi: 10.1007/s11746-010-1672-0.

[25] A. K. Agarwal, "Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines," *Progress in energy and combustion science*, vol. 33, no. 3, pp. 233–271, 2007, doi: 10.1016/j.pecs.2006.08.003.

[26] M. J. Ramos, C. M. Fernández, A. Casas, L. Rodríguez, and Á. Pérez, "Influence of fatty acid composition of raw materials on biodiesel properties," *Bioresource technology*, vol. 100, no. 1, pp. 261–268, 2009, doi: 10.1016/j.biortech.2008.06.039.

[27] I. Veza, M. F. Roslan, M. F. Muhamad Said, Z. Abdul Latiff, and M. A. Abas, "Cetane index prediction of ABE-diesel blends using empirical and artificial neural network models," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, pp. 1–18, 2020, doi: 10.1080/15567036.2020.1814906.

[28] I. Veza, M. F. Muhamad Said, Z. Abdul Latiff, and M. A. Abas, "Application of Elman and Cascade neural network (ENN and CNN) in comparison with adaptive neuro fuzzy inference system (ANFIS) to predict key fuel properties of ABE-diesel blends," *International Journal of Green Energy*, vol. 18, no. 14, pp. 1510–1522, 2021, doi: 10.1080/15435075.2021.1911807.