WALAILAK JOURNAL
http://wjst.wu.ac.th Engineering and Physical Sciences

Prediction of Fuel Properties of Biodiesel Using Two-Layer Artificial Neural Network

Solomon Olasunkanmi ODEYALE[*] and Adeleye Oladiran EGUNLETI

Olabisi Onabanjo University, College of Engineering and Environmental Studies, Ibogun, Ogun State, Nigeria

([*]Corresponding author's e-mail: solomon.odeyale@gmail.com)

Received: 6 May 2014, Revised: 3 October 2014, Accepted: 8 November 2014

Abstract

Biodiesel is an alternative fuel produced from a renewable source (biological). Biodiesel has properties similar to diesel, produced from fossil fuels, and this makes it a good substitute as fuel used in diesel engines. The experimental determination of various properties of biodiesel is costly, time consuming and also a tedious process. In order to reduce these problems, researchers have identified that the fatty acid composition of biodiesel determines its fuel properties. The percentage composition of the fatty acid content of each biodiesel plays a significant role, and it is the sole determinant of the fuel properties. Furthermore, artificial neural networks have been considered to be tools helpful in estimating these properties from the fatty acid composition of the fuel. In this study 4 properties (cetane number, flash point, kinematic viscosity, and density) have been modeled, using artificial neural network (ANN). These fuel properties were predicted using 5 fatty acids as input parameters. A 2-layer neural network was used with logsig and purelin in the hidden layers; the fatty acids considered as input parameters were; palmitic acid, stearic acid, oleic acid, linoleic acid and linolenic acid. Both cetane number and flash point used 6 neurons in the hidden layers, and the network had a determining factor (R^2) of 0.93488 and 0.9814 respectively. The network of the kinematic viscosity used 7 neurons in the hidden layer, and had a determining factor (R^2) of 0.83238, while the density network used 5 neurons and had a determining factor (R^2) of 0.819. The results obtained from this study closely agree with previous studies.

Keywords: Biodiesel, artificial neural network, fuel properties, alternative fuel, fatty acid

Introduction

The major sources of energy used in the world are from petrochemical sources, coal, natural gases, hydroelectricity and nuclear energy [1]. With the exception of hydroelectricity and nuclear energy, all other energy sources are finite. Few regions in the world have petroleum, but usage of petroleum fuel, which is a fossil-based fuel, is affected with the gradual depletion of petroleum reserve [2], increases in price, emission, and increase in global warming. Diesel fuel has significant and greater importance in the industrial, transportation and power generating sectors of the economy of developed and developing countries. The demand for diesel is rapidly increasing worldwide. Biodiesel has been considered as one of the alternative fuels to be used as a good replacement for diesel fuel in compression ignition engines for several reasons, among which is its production from a renewable source, such as vegetable oils and animal fats [3,4]. Some of the advantages of biodiesel usage in compression ignition engines, compared to petrodiesel, include; better lubricating property, higher flash point (which makes it safer), biodegradability, non toxicity, lower exhaust emission, and very low or no sulfur content [1,5].

Biodiesel

Biodiesel is monoalkyl esters of long chain fatty acids, derived from renewable feedstocks, such as vegetable oil or animal fats, for use in compression ignition engines [6]. In simple terms, it is the product gotten when vegetable oil or animal fat is chemically reacted with an alcohol to produce a new compound that is known as a fatty acid alkyl ester. It is diesel produced from a biological product. Biodiesel can be used alone or blended with a certain percentage of petrodiesel. It can also be used as a low carbon alternative to heating oil. Biodiesel is produced by a process known as transesterification. This is basically the conversion of feedstock that contains oil into methyl or ethyl esters. Biodiesel is mainly produced by the chemical reaction of 3 compounds, which are the triglyceride (oil) from the feedstock to be used, an alcohol, which can be methanol or ethanol, and a catalyst (acidic, alkaline or enzymic catalyst). In a transesterification or alcoholysis reaction, one mole of triglyceride reacts with 3 moles of alcohol (molar ratio of methanol to vegetable oil of 3:1) to form one mole of glycerol and 3 moles of the respective fatty acid alkyl esters [1,7].

Fuel properties

The amount of fatty acids present in the molecules of the triglyceride, contained in the feedstocks used in producing biodiesel, determines the fuel properties of biodiesel. Chain length and the number of double bonds determine the physical characteristics of both fatty acids and triglycerides [1]. Hence, the fatty acid compositions in vegetables oils result in the properties of biodiesels, such as viscosity, flash point, high-heating values, etc. [8,9]. Various vegetable oils and animal fats have been used in the production of biodiesel with different fatty acid compositions, which have led to the variation in their fuel properties. Standards have been set by different countries regarding the fuel properties of biodiesel that must be met before it can be used as fuel. Some of these fuel properties are highlighted below.

Cetane number (CN)

The cetane number is widely used as a diesel fuel quality parameter, related to the ignition delay time and combustion quality. The shorter the ignition delay time, the higher the CN, and vice versa. CN affects a number of engine performance parameters, like combustion, stability, drivability, white smoke, noise, and emissions of carbon monoxide and hydrocarbon. It is well known that the CN of a biodiesel depends on the feedstock used for its production. The longer the fatty acid carbon chains, and the more saturated the molecules, the higher the CN [1,5,10]. High CNs were observed from esters of saturated fatty acids, such as palmitic (C16:0) and stearic (C18:0) acids, and low CNs have been associated with unsaturated components, such as the esters of linoleic (C18:2) and linolenic (C18:3) acids. Biodiesel has a higher CN than conventional diesel fuel, which results in higher combustion efficiency [6]. Standards have been established worldwide for CN determination, for example, ASTM D613 in the United States, and internationally, ISO 5165.

Flash point (FP)

The flash point of a fuel is the temperature at which it will ignite when exposed to a flame or spark. This temperature correlates with its volatility, which is an important fuel feature for an engine's starting and warming. The combination of high viscosity and low volatility of a fuel causes bad cold engine start up, misfire, and ignition delay. In spite of the fact that the FPs of feedstocks are reduced through transesterification, they are still higher than those of diesel fuel, regardless of whether the biodiesel is from high-quality vegetable oils or from low-cost feedstocks. Since biodiesel has a higher FP than diesel, it is a safer fuel for transport purposes than diesel.

Density

The ASTM D941 test method can be used to measure the density of the biodiesel, diesel fuel and their blends. The densities of diesel fuels are slightly lower than those of biodiesels. The density of biodiesel using the ASTM D6751 standard usually varies between 0.86 and 0.90 g/cm^3. In many studies, it was observed that the biodiesel's density has not changed a lot, because the densities of methanol and oil are close to the density of the produced biodiesel [11].

Kinematic viscosity (KV)

Viscosity is an important property of any fuel, as it is an indication of the ability of a material to flow. Viscosity affects the atomization of a fuel upon injection into the combustion chamber, and thereby causes the formation of engine deposits. The higher the viscosity, the greater the tendency of the fuel to cause such problems [12]. The KV of biodiesel increases with chain length (number of carbon atoms), types of bonding, double bond configuration, and increasing degree of saturation [13]. ASTM D 445 provides a test method for determining the KV and the calculation method to determine the dynamic viscosity [14].

Other fuel properties

Several other biodiesel fuel properties are provided for in the standards, such as ASTM D6751, EN 14214 etc. These properties include: heat of combustion, cold point, pour point, lubricity, acid value, oxidative stability, phosphorous content, total sulfur content, ester content, iodine value, ash content, water content etc.

Methods of determining biodiesel fuel properties

Artificial neural network method

The ANN method involves the use of certain parameters of biodiesel, which have been previously known or obtained via experimental results, to predict future biodiesel properties. ANN is modeled to predict properties of biodiesel from one feedstock or blends of biodiesel in different proportions. A.S. Ramadhas *et al*. [15] successfully developed a multi layer feed forward ANN to predict the CN of biodiesel. ANN was employed to estimate the density of pure palm oil-based methyl ester biodiesel [16]. Kumar and Bansal examined 7 neural network (NN) architectures and 3 training algorithms, along with 10 different sets of weight and biases, to predict 4 properties of diesel-biodiesel blends [17]. The properties predicted were flash point, fire point, density and viscosity. The results showed that the NN gave the best estimate for diesel-biodiesel blends. Kraipat used a statistical model and ANN to predict kinematic viscosity and CN of biodiesels from their fatty acid compositions [8]. It was concluded by Kraipat that the ANN model predicted the properties accurately, and were closer to the experimental results than the statistical model.

Other methods include;

- Experimental and analytical method;
- Mathematical model;
- Statistical method.

Materials and methods

Data collections

The collected data include the Fatty Acid Methyl Ester (FAME) composition of biodiesels, or fatty acid composition of vegetable oil, and their corresponding fuel properties that are of interest to this work. It has been established that the fuel properties of biodiesel is not affected by the mode of preparation.

Selection of data

The fatty acid composition of oils (vegetable and animal oils) and fuel properties were obtained from experimental results, which satisfied established biodiesel standards as given by EN14214 and EN14213. The 5 predominant constituents of most biodiesels were selected for use. These were the palmitic (C16:0), stearic (C18:0), oleic (C18:1), linoleic (C18:2) and linolenic (C18:3) acids. The chain lengths of carbon for all these acids were 18, except for palmitic acid, which was 16. Both the palmitic and stearic acids are saturated acids, while the other 3 were unsaturated acids; oleic is monounsaturated with a single bond, linoleic acid is di-unsaturated, i.e., with a double bond, and linolenic is tri-unsaturated, with a triple bond. The fatty acid composition of biodiesel has been reported by many researchers as the main determinant of biodiesel properties upon which this present work is be developed. Some feedstocks in the data gathered from the literature have fatty acids with less than 16 carbon chains and fatty acids

with more than 18 carbon chains. **Table 1** shows the list of feedstocks whose fatty acid composition was modified for this present work. In order to consider these fatty acids, the percent values of all saturated acids with a carbon chain of less than 16 were added to the amount of palmitic acid (C16:0) present, while the saturated acids with a carbon chain of more than 18 were added to the stearic acid (C18:0). Also, acids with single bonds (a carbon chain of more than 18) were added to oleic acid (C18:1), those with double bonds were added to linoleic acid (C18:2), while those with triple bonds were added to linolenic acid (C18:3). Some feedstocks used in this work have been reported to have acids with a level of unsaturation above triple bond. They were also taken into consideration by adding those with 4 bonds to oleic acid (C18:1), those with 5 bonds added to linoleic acid (C18:2), and those with 6 bonds to linolenic acid (C18:3).

The ANN model analysis

Steps/training procedures

- Fatty acid composition as the input and the fuel property as the target vector were written in a suitable format in the MATLAB workspace.
- A feed forward back propagation network with 2 layers was created in the NN toolbar of MATLAB. This is because back propagation uses a gradient descent method, and it allowed the NN to train with higher degree of efficiency.
- Trainlim and sigmod were chosen as the training and transfer functions, respectively. A sigmod transfer function can learn nonlinearity between inputs and target very fast and accurately.
- The numbers of neurons for the hidden layer was selected based on the input parameters for each network trained.
- The input and target vectors were introduced to the created network, and the weight initialized.
- Training parameters, such as epoch, maximum failure, error goal etc. were adjusted.
- The network was gradually trained. The process was completed when the defined error was reached.

The optimum number of hidden layer neurons was selected and determined during the training and leaning process by trial and error.

Properties to be modeled

Input parameters

For all the networks to be modeled, the input layer consisted of 5 vector elements, which were the palmitic acid, stearic acid, oleic acid, linoleic acid, and the linolenic acid, respectively. Except for the density network, in which 6 input parameters were used, the temperature in °C at which the density of the selected biodiesel was reported was used as the sixth parameter, apart from the other 5 parameters.

Output parameters

Kinematic viscosity

Values of KV at different temperatures from numerous FAMEs have been reported; the values selected were those of KV at 40 °C, and found to be within the limit of the biodiesel standards of ASTM D6751 and EN 14214 (ASTM D445 or EN ISO 3104). The average and modified value of the FAME composition and corresponding KV at 40 °C is shown in **Table 3**. For this work, several networks with different transfer functions, such as logsig, tansig and purelin, were used, with variation in the network parameters, such as neuron, epoch, maximum failure, hidden layer etc, so as to determine the best network using the selected data.

Cetane number

A back propagation learning algorithm was used to train the network. Purelin transfer function was used in the output layer, and the output vector element was the predicted CN. At the start, twelve networks were trained using 5, 6, 7 and 8 neurons, with logsig, tansig and purelin used respectively with these neurons for accuracy comparison. The entire networks were trained for each selected number of

neurons with different network parameter values, such as maximum failure, epoch, gradient etc. The network considered to be the best was the one with the least network output error.

Flash point

The fatty acid composition of feedstocks was used as the input vector and the FP as the target vector. A feed forward back propagation network was designed to model the NN. Six-network models, with 4, 5 and 6 neurons, were used, with tansig and logsig transfer functions as the transfer function in the first layer for prediction accuracy.

Density

The fatty acid composition of feedstocks at the selected temperature was used as the input parameter, while the density of these feedstocks at that temperature was used as the target output. Only feedstocks whose densities fall within the density standards for biodiesel were considered. Several NN with different numbers of neurons were trained in order to determine which had the best prediction accuracy. During the training, trial and error means were employed in the selection of the training parameters, such as epoch, gradient, maximum failure, etc. Likewise, both logsig and Tansig transfer functions were compared with different numbers of neurons.

Results and discussion

General training result

After several trainings with different transfer functions and variation of each set of training parameters, it was found that logsig as a transfer function yielded the best result for all networks trained for this study. Also, it was found that the same neurons did not give the best output for the various properties that were considered in this work. Though 4, 5, 6 and 7 neurons were considered for the density network, 5 neurons gave the best result. Neurons, ranging from 4 to 8 in number, were used in training both networks for flash point and cetane number, but it was found that 6 neurons gave the best acceptable result for both FP and CN networks. Kinematic viscosity was the only property whose prediction required the use of 7 neurons as the best number of neurons required, even though neurons between 4 and 8 were used to train the networks. Some parameters that were used in training NNs were; show, epoch, time, max_fail mu_reduc, min_grad, mu, mu_dec, mu_inc and mu_max. Others that were very important and cannot be overlooked were: the type of network, training function, adaptation learning function, performance function, number of layers, and properties of layers. The same network type was used for all the training, that is, the feedback propagation network. The training function used was Train LM (Levenberg-Marquardt back propagation), Learn GDM (Gradient descent with momentum weight and bias learning functions) was the adaptation leaning function used, and MSE (mean-squared error) was the performance function. All networks employed 2 layers, with logsig in the first layer and purelin in the second layer. This paper used the squared value (R^2) of the regression coefficient to justify the prediction of the outputs. Other metrics, such as mean absolute error (MAE) and mean square error (MSE), could also be used to justify prediction results [18-20].

NN of cetane number

CN is a very important diesel fuel property, similar to the octane number of petrol fuel. CN is the measure of the ignition quality of a diesel engine. The limit of CN given by United States (ASTM D 613) and Europe (EN ISO 5165) is at a minimum of 47 and 51 respectively. The data collected, as shown in **Table 1**, are those that satisfy both standards. Thirty five different biodiesel feedstocks, with 5 pure fatty acids, were used in this work to train the neural network of cetane number. **Table 1** shows the feedstocks used and their actual cetane numbers.

Table 1 Modified fatty acid composition and actual cetane number of different biodiesel.

No.	Biodiesel	Palmitic acid (C16:0)	Stearic acid (C18:0)	Oleic acid (C18:1)	Linoleic acid (C18:2)	Linolenic acid (C18:3)	Cetane number
1	Peanut	8	7.5	55.7	28.4	0.3	53
2	Palm	37.5	7.2	46.4	8.6	0.3	61
3	Crude palm oil	46.16	4.44	40.19	8.94	0.27	62.4
4	Distilled palm oil	44.3	4.05	40.77	10.25	0.25	58.3
5	Sunflower oil	5.5	6.1	21.4	66.2	0.8	55.6
6	Sunflower	6.5	4.6	25.6	63.1	0.2	50
7	High-oleic sunflower	4.8	4.7	62.9	27.5	0.1	53
8	Rapeseed	4.9	1.6	65.3	20.4	7.9	55
9	Soybean oil	11.3	4	25.6	53	6.1	51.1
10	Jatropha oil	14.2	8.3	43.1	34.4	0.5	57.1
11	Guindilla	9.2	8.3	73.1	7.7	0.8	59
12	Olive	11.6	4	76	7.8	0.6	57
13	Grape	7	4.3	19.1	69.1	0.3	48
14	Almond	10.4	3.2	78	7.6	0.8	57
15	Corn	6.5	1.5	66.4	25.2	0.1	53
16	Okra oil	32.53	5.12	30.44	30.05	0.38	52.2
17	Karanja oil	10.6	18.6	51.8	19	0	55.1
18	G. abyssinica	9.2	10.1	9	71.7	0	57
19	P pinnata oil	9.8	6.2	72.2	11.8	0	55.1
20	Cotton oil	24.9	2.92	18.93	53.14	0	54.13
21	Soap stocks	17.2	4.4	15.7	55.6	7.1	51.3
22	Scum	58	15.76	19.21	0.48	0.26	60
23	Rice bran oil	18.8	3.1	43.1	33.2	0.6	63.8
24	Soya bean	14	4	24	52	0	45
25	Tallow	26	25	43	3	0	58.8
26	Conola oil	3.5	0.9	64.4	22.3	8.2	55
27	Safflower oil	6.85	2.11	14.2	75.98	0	52.32
28	Pongami(karanja)	9.8	6.2	72.2	11.8	0	55.1
29	Camelina oil	6.9	6.8	32.1	20.9	33.4	52.8
30	TME	20.3	9.2	7.3	15.7	2.8	57.2
31	Tobacco	10.69	3.34	14.74	69.49	0.69	51
32	Yellow grease	25.67	12.96	48.11	6.97	0.67	62.6
33	Mellon bug	33.3	3.5	57.3	3.9	0	55
34	Sorghum bug	12.3	7.3	41.9	34.5	0	55
35	S birra	14.3	8.8	67.5	0	0	62
36	Pure palmitic acid	100	0	0	0	0	74.3
37	Pure stearic acid	0	100	0	0	0	75.6
38	Pure stearic acid	0	0	100	0	0	57.2
39	Pure linoleic	0	0	0	100	0	42
40	Purelinolenic acid	0	0	0	0	100	22.7

Multilayers of neurons with nonlinear transfer function allowed the network to learn both the linear and nonlinear relationships that may exist between the input and the target vectors. In this study, the input layer consisted of 5 fatty acid compositions, palmitic, stearic, oleic, linoleic and linolenic acids, which were arranged in the sequence mentioned. The target vector was the cetane number that corresponds to the fatty acid of each row, as tabulated in **Table 2**. In this network, 2 hidden layers, with logsig in the first layer and purelin in the second layer, were used. The input element was the fatty acid composition, while the output element was the predicted cetane number. All the data to train, validate and test the network performance were provided to the network. Twenty four sets of data, which equaled 60 % of the total data, was used in the training set, 15 % of the total sets, which accounted for 8 sets of data, were used in

the validation, with the remaining 15 % in the testing. The selection of data used at different operation was done randomly during the training. At the start, the network was trained with the selected number of neurons i.e. 6 neurons, and the ignition values of the network parameters provided by matlab were used before being adjusted. The time of training was set to infinity, maximum failure was adjusted until 8 gave the best result, the goal was left at zero, mem_reduc was left at one, min_grad was adjusted to 10^{-10}, mu at 0.01, mu_dec at 0.1, mu_inc was given as 10, and mu-max at 10 billion.

The training was repeated several times until the predicted output and actual values had a very close relation. **Figures 1 - 4** show the plot regression of the data used, the training, the validation and the test data. The graphs below show the training, validation and test date, with R= 0.98083, 0.97806 and 0.99274 respectively. R when all predicted and target are compared is 0.96689. For the trained network, R^2 is equal to 0.93488. The actual and predicted outputs are listed in **Table 2** below. From this table the error difference can be seen as been calculated to be the difference between the target and the output predicted.

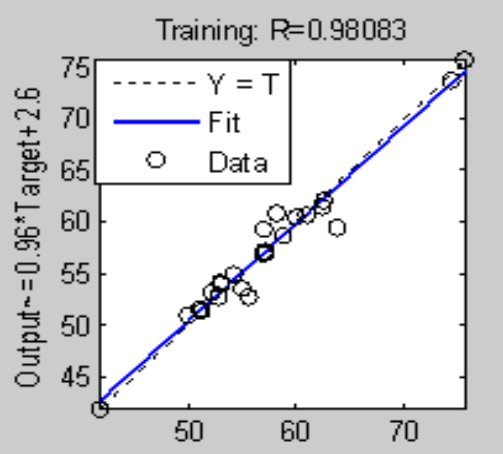

Figure 1 Regression of CN training data.

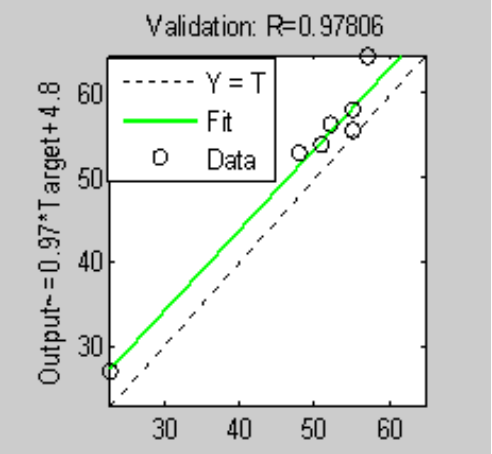

Figure 2 Regression of CN validation data.

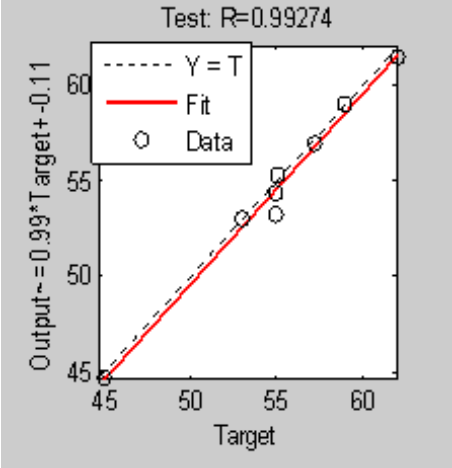

Figure 3 Regression of test sets for CN.

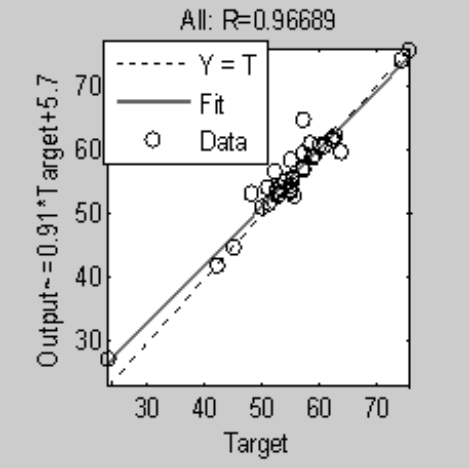

Figure 4 Regression of all data for CN.

Table 2 Actual cetane number against predicted cetane number.

S/N	Actual CN	Predicted CN	Error	% Difference
1	53	54.17814005	-1.17814005	-2.22291
2	61	60.55014193	0.449858074	0.737472
3	62.4	59.37393417	3.026065832	4.849464
4	58.3	58.61017797	-0.310177967	-0.53204
5	55.6	50.40250494	5.197495056	9.348013
6	50	49.74083011	0.259169889	0.51834
7	53	53.06509864	-0.065098639	-0.12283
8	55	52.93692049	2.063079509	3.751054
9	51.1	50.52711372	0.572886284	1.121108
10	57.1	56.30349915	0.796500845	1.394923
11	59	58.96773427	0.032265727	0.054688
12	57	56.4686657	0.531334298	0.932165
13	48	51.39534294	-3.395342943	-7.07363
14	57	56.23110642	0.768893576	1.348936
15	53	53.04128859	-0.04128859	-0.0779
16	52.2	52.69336448	-0.493364482	-0.94514
17	55.1	55.26089324	-0.160893239	-0.292
18	57	55.49463339	1.505366614	2.640994
19	55.1	55.39253605	-0.292536052	-0.53092
20	54.13	53.79097046	0.339029544	0.626325
21	51.3	51.26785466	0.032145339	0.062661
22	60	60.39752553	-0.397525525	-0.66254
23	63.8	56.10830325	7.691696748	12.05595
24	45	44.65126191	0.348738088	0.774974
25	58.8	63.50236653	-4.702366532	-7.99722
26	55	53.2705501	1.729449904	3.144454
27	52.32	52.50605328	-0.186053282	-0.35561
28	55.1	55.39253605	-0.292536052	-0.53092
29	52.8	52.74716039	0.052839606	0.100075
30	57.2	57.41058319	-0.210583189	-0.36815
31	51	52.4521459	-1.4521459	-2.84734
32	62.6	62.15063858	0.449361417	0.71783
33	55	57.5662845	-2.566284498	-4.66597
34	55	54.34198904	0.65801096	1.196384
35	62	61.39656386	0.603436145	0.973284
36	74.3	71.56570011	2.734299893	3.680081
37	75.6	73.45520963	2.144790366	2.837024
38	57.2	56.91268486	0.28731514	0.502299
39	42	41.85991735	0.14008265	0.33353
40	22.9	22.85010494	0.049895056	0.217882

It is seen from **Table 2** that the predicted output was in close agreement with the actual cetane number. The highest error deviation was 12.06 % of the target data, which was for the cetane number of rice bran oil with 63.8 and which was predicted to be 56.1083. Apart from the pure fatty acid used to support the network data, all other predicted results of the various biodiesel satisfied the standard of biodiesel, which is a minimum of 47 by the ASTM D613 standard. From the difference between the actual and predicted output, it was found out that error in prediction ranged from –4.702366532 to

7.691696748 %. This error witnessed could be as a result of the selected actual cetane number for a particular feedstock, since different researchers report different values of cetane number, which is sometimes in a range of ± 5 %.

NN for kinematic viscosity

The high viscosity of pure vegetable oil has so far been a limitation to its use as a substitute fuel for diesel. Fuels with high viscosity have a greater tendency of forming deposits in the engine combustion chamber. The viscosity of transesterified oil is drastically reduced when compared to the parent oil. The viscosities of different reported biodiesel previously produced were used in the present work. 1.9 - 6.0 mm^2/s is the range of accepted biodiesel viscosity by ASTM D 6751 while a boundary of 3.5 - 5.0 mm^2/s is the standard by EN 14214. The accepted temperature at which viscosity is reported is 40 °C by standard.

For this work 46 feedstocks were chosen in training the viscosity network; this included 4 pure fatty acids. The input data were the palmitic, stearic, oleic, linoleic and linolenic acids respectively with the actual viscosity as the target vector. The list of data used in training this network is shown in **Table 3**. All the selected data were within the limit of accepted viscosity. After several trials, the network with a logsig transfer function in the first layer and purelin in the second layer was taken as the best network. Seven neurons with a logsig transfer function gave the best and closest relationship between the target and predicted output. The training parameters that were adjusted from the initial value were the max_fail, which was changed to 8, min_grad as 10^{-10}, mu, taken as 0.01, mu_dc, adjusted to 0.1, mu_inc, as 10 and mu_max, used as 10 billion.

Twenty eight data out of the 46 selected data were used for the training, and 9 for the validation, while the remaining 9 were used to test the performance. **Figures 5 - 7** shows the plot for the data chosen for training, validating and to test the performance, while **Figure 8** shows the plot of the total data for the network. 0.97663, 0.86094 and 0.89368 were the values of regression coefficient (R) for the training, validation, and testing, respectively. The overall target and predicted output has R equal 0.91235 and the overall square of R i.e. R^2 for this network is 0.83238. **Table 3** shows the comparison between the actual viscosity and predicted output; the error and the percentage difference are also included.

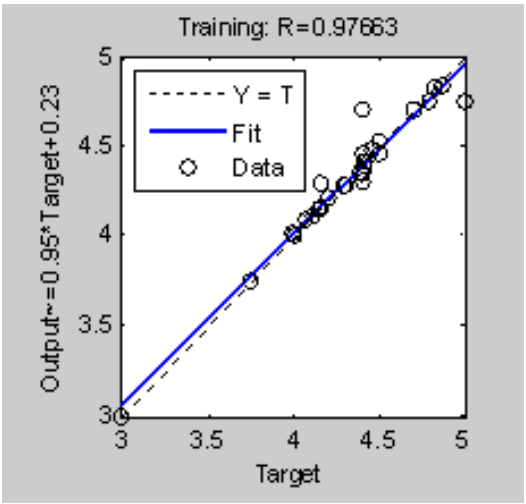

Figure 5 Regression of viscosity training data.

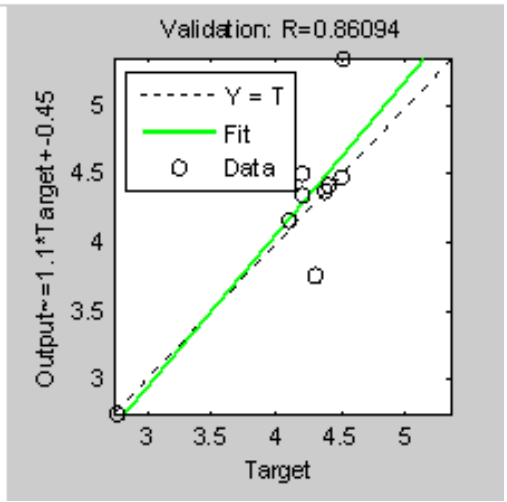

Figure 6 Regression of viscosity validation data.

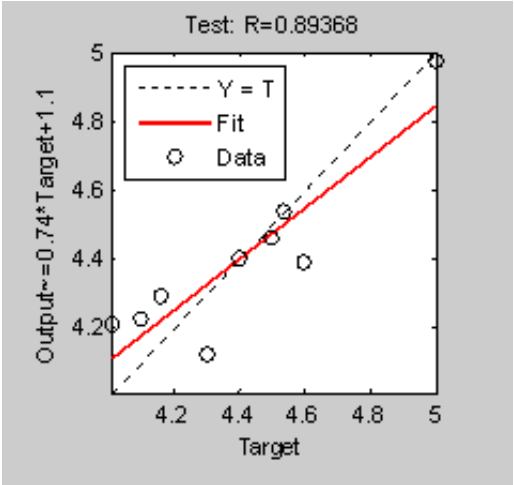

Figure 7 Regression of viscosity training data sets.

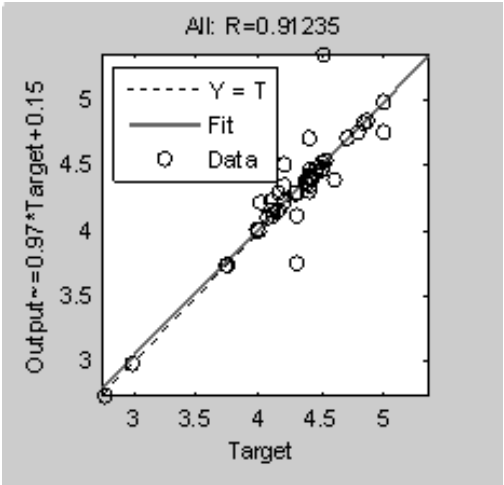

Figure 8 Regression of all viscosity data sets.

Table 3 Actual kinematic viscosity against predicted kinematic viscosity.

No.		Actual	Predicted	Error	% Difference
1	Peanut	4.6	4.388897	0.211103	4.589196
2	Palm oil	4.5	4.34332	0.15668	3.481778
3	Crude palm oil	4.502	4.509982	-0.00798	-0.1773
4	Distilled palm oil	4.415	4.42768	-0.01268	-0.2872
5	Sunflower oil	4.1	4.096929	0.003071	0.074902
6	Sunflower oil	4.2	4.243239	-0.04324	-1.0295
7	High oleic sunflower	4.4	4.438002	-0.038	-0.86368
8	Rapeseed	4.4	4.907185	-0.50719	-11.5269
9	Soybean oil	4.2	4.423509	-0.22351	-5.32164
10	Jatropha oil	4.4	4.399346	0.000654	0.014864
11	Guindilla	4.867	4.537548	0.329452	6.769098
12	Olive oil	4.5	4.307079	0.192921	4.287133
13	Grape	4.1	4.224253	-0.12425	-3.03056
14	Almond	4.2	4.330425	-0.13043	-3.10536
15	Corn	4.4	4.319368	0.080632	1.832545
16	Okra oil	4.01	4.211222	-0.20122	-5.018
17	Karanja oil	4.16	4.269347	-0.10935	-2.62853
18	G. Abyssinca	4.3	4.17668	0.12332	2.867907
19	p. pinnata oil	4.16	4.292133	-0.13213	-3.17627
20	Cotton oil	4.07	4.025117	0.044883	1.102776
21	Soap stocks	4.3	4.243017	0.056983	1.325186
22	Scum	3.75	3.567221	0.182779	4.874107
23	Soya bean	4.5	4.459895	0.040105	0.891222
24	Tallow	5	4.973701	0.026299	0.52598
25	Conola oil	5	4.97051	0.02949	0.5898
26	Non erucic brassicca oil	4.83	4.811454	0.018546	0.383975
27	Croton megalocarpus	4.46	4.453818	0.006182	0.13861
28	Carmelia oleifera	4.54	4.536282	0.003718	0.081894

No.		Actual	Predicted	Error	% Difference
29	Waste cooking oil	4.7	4.686016	0.013984	0.297532
30	ZBMSO	4	3.920744	0.079256	1.9814
31	Yellow horn oil	4.4	4.499749	-0.09975	-2.26702
32	Used cooking oil	4.79	4.484095	0.305905	6.386326
33	Polanga	3.99	3.937604	0.052396	1.313183
34	Karanja	4.37	4.381937	-0.01194	-0.27316
35	Polycarpa fruit	4.12	4.171672	-0.05167	-1.25417
36	Pumpkin seed	4.41	4.581858	-0.17186	-3.89701
37	Safflower oil	4.29	4.23367	0.05633	1.313054
38	Pongami (karanja) oil	4.16	4.292033	-0.13203	-3.17387
39	Camelina oil	4.15	4.391109	-0.24111	-5.80986
40	TME	4.42	4.395753	0.024247	0.548575
41	Poultry fat	4.39	4.377128	0.012872	0.293212
42	Terminalia	4.3	4.121739	0.178261	4.145605
43	Pure palmitic acid	2.778	2.783591	-0.00559	-0.20126
44	Pure stearic acid	2.986	3.459677	-0.47368	-15.8633
45	Pure oleic acid	4.518	4.67974	-0.16174	-3.5799
46	Pure linoleic acid	3.745	3.734747	0.010253	0.273778

NN for flash point

The temperature at which the fuel will ignite or start to burn when it comes in contact with fire is called the flash point. For this study, flash points of biodiesel selected were experimentally determined using the standard test method, using either ASTM D 93 or EN ISO 3679. The minimum standard set by the United States is 130, 120 °C by European, Australian and South African standards, and 100 °C by Brazilian standards.

During the training of this network, among the several numbers of neurons chosen for this training, the network with 6 neurons yielded the best result from several training operation. In the hidden layers, logsig was used as the first layer transfer function, while purelin was used in the second layer. Other training parameters that need to be discussed, as they were adjusted for good training, include epoch, whose value was 100, time, set to infinity, goal, at 0, max_fail, adjusted to 12, mem_reduc, 1, min_grad, set to 10^{-10}, mu_dec, at 0.1, and mu_inc, set to 10, with mu_max being the last parameters for the training parameters, the value always being large; for this network, 10 billion was the value that was used. The input parameters used for this network were the major fatty acid composition observed in most biodiesel; these were palmitic acid, stearic acid, oleic acid, linoleic acid and the linolenic acid, which were arranged sequentially to train. The target vector was the corresponding flash point of each grouped fatty acid composition. The set of data used to train this network is given in **Table 4**, the performance regression plots for the selected data in the training sets, validation sets, and test sets are shown in **Figures 9 - 12**.

Regression coefficient (R) equaled 0.99675, 0.98895 and 0.98369 for the training, validation and test sets respectively, all the data has R = 0.99066 whose square (R^2) equals 0.9814. The sets of predicted outputs of the network and the actual flash point are given in **Table 4**. From these sets of data, the error difference between the actual and output is also included in this table.

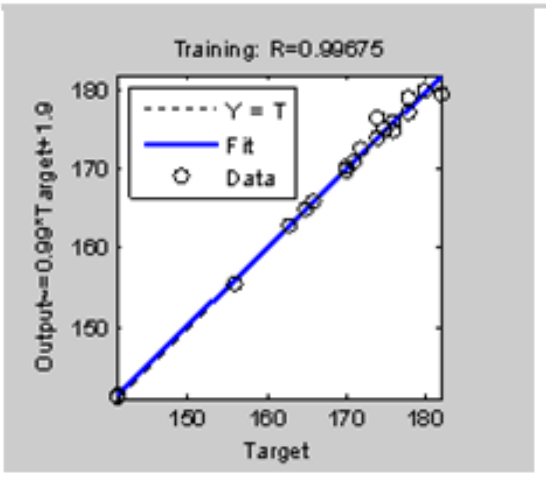

Figure 9 Regression of flash point training data.

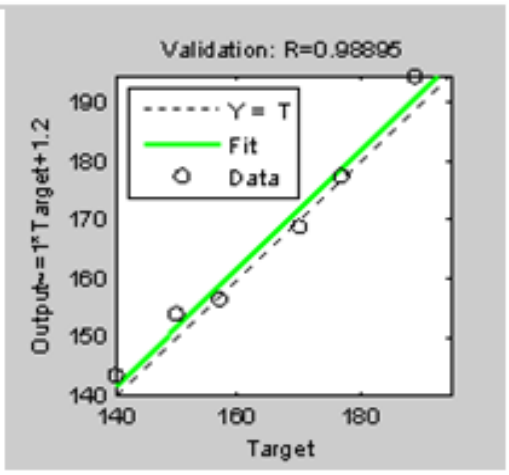

Figure 10 Regression of flash point validation data.

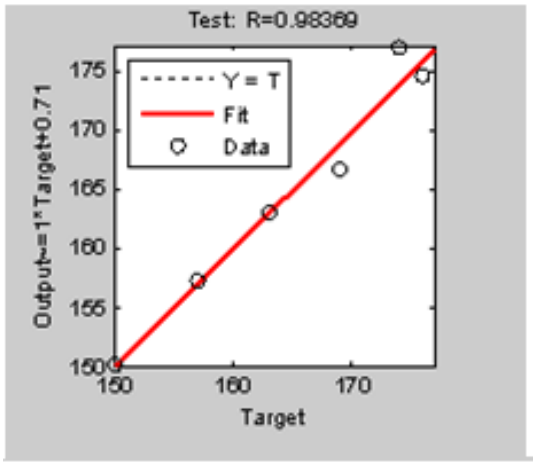

Figure 11 Regression of flash point training data sets.

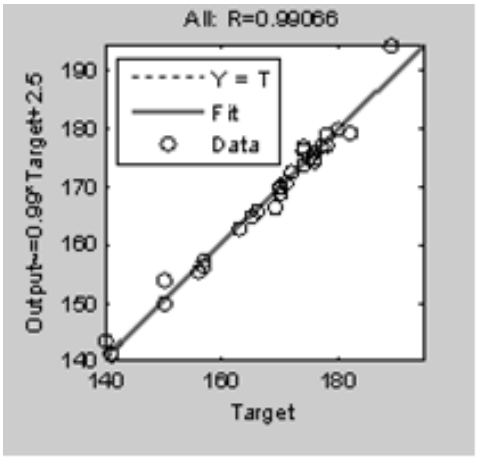

Figure 12 Regression of flash point training validation sets.

NN for density

The density of biodiesel was determined experimentally using the standard test methods by biodiesel standard. ASTM D1298 or EN ISO 3675/12185 are the accepted test methods, and the range of acceptance falls within 860 - 900 kg/m^3. The temperature at which the density is reported is at 15 °C. For this study, the data selected were those whose density fell within the above limit; although not all biodiesel densities used were reported at 15 °C, some were reported at temperature above this standard. This led to a consideration of temperature at which the biodiesel density was determined. The sets of data that were used for the training of density network are shown in **Table 4**. The input parameters were palmitic, stearic, oleic, linoleic, linolenic acids and the temperature at which density was reported, while the target vector was the density of each fatty acid composition. Thirteen sets of data were used for the training, 4 for the validation, and the last 4 to test the performance of the developed neural network for

density. The sets of data used in the training, validation, and training account to 60 % of the total available data for the training, 15 % for the validation and 15 % for the testing, respectively.

The trained network used 5 neurons in the first hidden layer with a logsig transfer function. The second layer used a purelin function. The training parameters were; epochs, time, goal, max_fail, max_reduc, min_grad, mu, mu_dec, mu_inc and mu_max, with values of 100, infinity, 0, 6, 1, 10^{-10}, 0.01, 0.1, 8 and 10 billion, respectively. **Figures 13 - 16** show the regression of data used for the training at various stages. The training sets had regression coefficients (R) of 1, the validation with R equaled to 0.99509, while the test sets had R equal 0.92283. The overall target and predicted output had R equal 0.98498, and the overall square of R that is R^2 equaled to 0.819. The deviation between the target and predicted output is shown in **Table 4**.

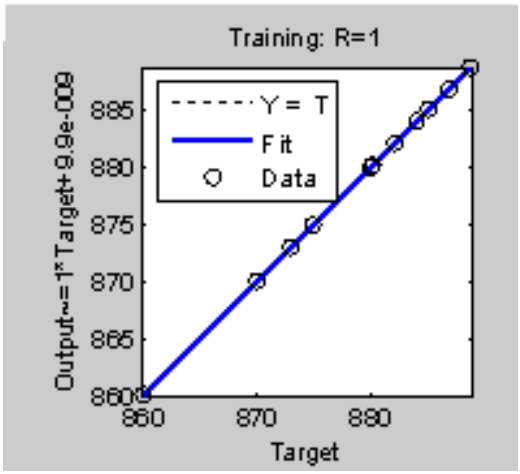

Figure 13 Regression of training data for density.

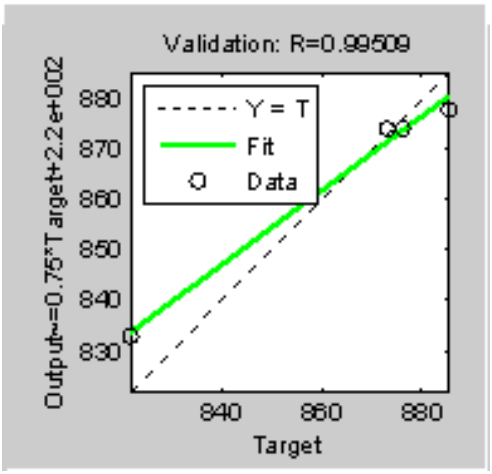

Figure 14 Regression of validation for density.

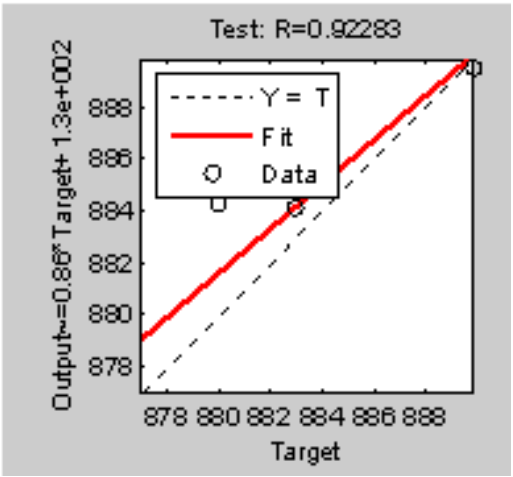

Figure 15 Regression of test data for density.

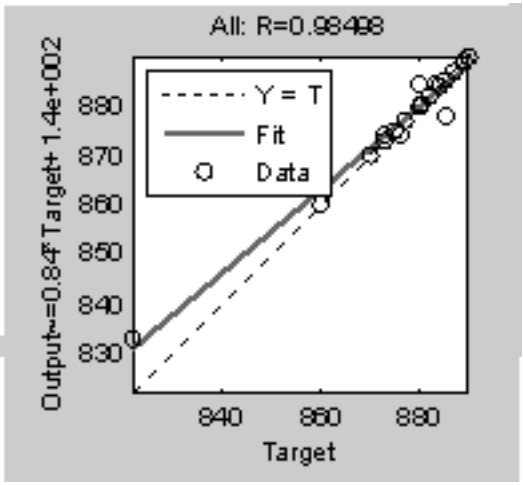

Figure 16 Regression of all data sets.

Table 4 Actual value against predicted value for flash point and density.

Flash point			Density		
Biodiesel	**Actual FP**	**Predicted FP**	**Biodiesel**	**Actual Density**	**Predicted Density**
Peanut	176	177.5176	Kusum oil	870	869.9929
Palm	176	174.4666	Moringa oleifera	883	884.9760
Crude palm oil	174	178.1220	Croton megalocarpus	888.9	889.5178
Distilled palm oil	182	177.5846	Tobacco	886.8	886.3488
Sunflower oil	180	180.9252	Peanut	883	884.1717
Sunflower	177	178.0262	Soya bean	885	881.9125
High oleic sunflower	174	173.4548	Babassu	875	872.7696
Rape	170	171.2293	Palm	880	879.8725
Soybean oil	171	171.2427	Sunflower	860	859.8394
Jatropha curcas oil	166	161.5699	Tallow	877	877.0324
Grape	175	175.2883	Non erutic brasicca oil	888.8	886.2540
Almond	172	171.9011	Terminalia catappa oil	873	872.7521
Olive	178	170.3515	Soybean	885	885.1431
Corn	170	171.3549	Rapeseed	882	878.4606
Okra oil	156	157.2904	Yellow grease	873	873.0170
G abbysiica oil	157	157.2517	Soap stock	885	884.9819
Karanja oil	141	144.4006	Jatropha oil	880	884.3215
Cotton oil	150	150.0999	Cotton seed	875	876.6801
Soap stocks	169	166.6677	Rice bran oil	884	884.6287
Soya bean	178	175.5642	Roselle oil	880.1	883.6018
Conola oil	170	169.1864	Okra oil	876	874.9460
Non erutic brasica oil	163	162.9198			
Croton megalocarpus	189	187.2988			
Camelina oil	150	149.8009			
Waste cooking oil	141	141.0573			
ZBMSO	174	181.9058			
Used cooking oil	176	173.1657			
Polanga	140	140.6413			
Karanja	163	163.1270			
Polycarpa fruit oil	165	165.2751			
Pumpkin seed	174	177.0178			

Comparing present work to previous prediction

The prediction of biodiesel properties from the fatty acid composition using ANN has not been widely looked into. Many predictions of the properties from fatty acid compositions have been widely done with the aid of formulas. This has been a limitation on how to use the previous work done as a clue on how to improve a better neural network for biodiesel properties. Comparison was made with Damibras 2008 [5], which used formulae relationships between some properties of biodiesel. Ten feedstocks were used for this relationship.

Table 5 Actual value of kinematic viscosity against predicted values.

Feedstocks	Actual KV	D P KV	This work	D P error	This work error
Palm	4.5	4.41	4.34	0.09	0.16
Olive	4.5	4.47	4.31	0.03	0.19
Peanut	4.6	4.41	4.39	0.19	0.21
Rape	4.4	4.23	4.9	0.17	-0.5
Soybean	4.2	4.26	4.42	-0.06	-0.22
Sunflower	4.2	4.44	4.21	-0.24	-0.01
Grape	4.1	4.38	4.22	-0.28	-0.12
H. O SUN	4.4	4.35	4.44	0.05	-0.04
Almond	4.2	4.29	4.34	-0.09	-0.14
Corn	4.4	4.22	4.32	0.18	0.08

D P-Damibras prediction, KV-kinematic viscosity

Table 6 Actual value of flash point against predicted values.

Feedstocks	Actual FP	D P FP	This work	D P error	This work error
Palm	176	178.9	174.47	-2.9	1.53
Olive	178	178.9	170.55	-0.9	7.45
Peanut	176	182.17	177.43	-6.17	-1.43
Rape	170	175.64	171.23	-5.64	-1.23
Soybean	171	169.11	171.24	1.89	-0.24
Sunflower	177	169.11	178.82	7.89	-1.82
Grape	175	165.85	175.29	9.15	-0.29
H.O SUN	174	175.64	174.35	-1.64	-0.35
Almond	172	169.11	171.51	2.89	0.49
Corn	170	175.64	171.27	-5.64	-1.27

D P-Damibras prediction, FP-Flash point

The error difference in the actual kinematic viscosity and the predicted value is shown in **Table 5**. From this table, it can be seen that the 2 predictions, i.e., by Damibras prediction and this work, predicted 5 feedstocks, each very close to the actual values. It can be deduced from this table that the prediction by this present work is good enough, since the lowest error difference is 0.01, compared to the previous one, which was 0.03.

Table 6 shows the comparison between the actual flash point and the predicted values of the flash point of 10 biodiesel feedstocks. The prediction of the flash point of 10 feedstocks made by both works is shown in the table and, from the result, it can be seen that this work predicted 9 feedstocks very close to the actual value of the flash point of the selected feedstocks. The error difference from this work had a least error difference of 0.24, and others fall below this amount too, apart from olive, which was 7.45. The prediction from the work done by Dambras 2008 had a least error difference of 0.9, and it was with olive where this work predicted an error of 7.45; the high value in the prediction by this work might be as a result of modification done in the fatty acid value.

Conclusions

In this study, a remodeled ANN was used to predict the properties of biodiesel from the fatty acid composition in other to reduce the cost and time spent on experimental analysis of these properties. Five fatty acids were considered to be the major factors that affect the fuel properties of biodiesel. These properties have been remodeled with consideration to other fatty acid composition that might be present in the composition of biodiesel and are included in the neural network depending on its chain length and level of saturation. Four neural networks were designed and trained to predict the cetane number, flash point, kinematic viscosity, and density of biodiesel, with the aid of an artificial neural network with a logsig and a purelin transfer function in the hidden layer of all the properties being predicted. All networks were able to give a predicted result which was close to the actual experimental result that has been reported. Five, six and seven neurons were used as the number of neurons to achieve the best result for the model of density, flash point, cetane number, and kinematic viscosity, respectively, with a logsig transfer function in the first hidden layer where 2 layers were used. The 5 fatty acid composition used as the input parameters were palmitic acid (C16:0), stearic acid (C18:0), oleic acid (C18:1), linoleic acid (C18:2), and linolenic acid (C18:3), except for density, where temperature was considered as the sixth parameter.

It is seen that the result from this study has a closer value to the actual values of the properties to be predicted, compared to other experimental values seen in the literature. Therefore, this approach is a more accurate, time-efficient and cost-efficient method for predicting biodiesel fuel properties.

Acknowledgements

The authors gratefully acknowledge Microbiologist Olayinka John Ekundayo for providing the necessary research facilities.

References

[1] MJ Ramos, CM Fernández, A Casas, L Rodríguez and Á Pérez. Influence of fatty acid composition of raw materials on biodiesel properties. *Bioresource Technol.* 2009; **100**, 261-8.
[2] K Pramanik. Properties and use of jatropha curcas oil and diesel fuel blends in compression ignition engine. *Renew. Energ.* 2003; **28**, 239-48.
[3] WMJ Achten, L Verchot, YJ Franken, E Mathijs, VP Singh, R Aerts and B Muys. Jatropha biodiesel production and use- Review. *Biomass Bioenerg.* 2008; **32**, 1063-84.
[4] GEl Diwani, NK Attia and SI Hawash. Development and evaluation of biodiesel fuel and by-products from jatropha oil. *Int. J. Environ. Sci. Tech*. 2009; **6**, 219-24.
[5] A Demirbas. Mathematical relationships derived from biodiesel fuels. *Energ. Sourc.* 2008; **30**, 56-69.
[6] LC Meher, DV Sagar and SN Naik. Technical aspect of biodiesel production by transesterification-A review. *Renew. Sustain. Energ. Rev.* 2006; **109**, 248-68.
[7] AZ Abdullah, N Razali, H Mootabadi and B Salamatinia. Critical technical areas for future improvement in biodiesel technologies. *Environ. Res. Lett.* 2007; **2**, Article ID 034001.
[8] K Cheenkachorn. Predicting properties of biodiesel using statistical models and artificial neural networks. *In*: Proceedings of the Joint International Conference on Sustainable Energy and Environment, Hua Hin, Thailand, 2004.
[9] A Mariod, S Klupsch, IH Hussein and B Ondruschka. Synthesis of alkyl esters from three unconventional Sudanese oils for their use as biodiesel. *Energ. Fuel.* 2006; **20**, 2249-52.
[10] G Knothe, JV Gerpen and J Krahl. *The Biodiesel Handbook*. AOCS Press, Urbana, 2005.
[11] E Alptekin, M Canakci. Determination of the density and the viscosities of biodiesel-diesel fuel blends. *Renew. Energ.* 2008; **33**, 2623-30.
[12] G Knothe. Dependence of biodiesel fuel properties on the structure of fatty acid ethyl esters. *Fuel Process. Tech.* 2005; **86**, 1059-70.

[13] M Canakci and H Sanli. Biodiesel production from various feedstocks and their effects on the fuel properties. *J. Ind. Microbiol. Biotechnol.* 2008; **35**, 431-41.
[14] S Fernando, P Karra, R Hernandez and SK Jha. Effect of incompletely converted soybean oil on biodiesel quality. *Energy* 2007; **32**, 844-51.
[15] AS Ramadhas, S Jayaraj, C Muraleedharan and K Padmakumari. Artificial neural networks used for the prediction of the cetane number of biodiesel. *Renew. Energ.* 2006; **31**, 2524-33.
[16] S Baroutian, MK Aroua, AA Raman and Nik Sulaimon. Estimation of vegetable oil based methyl esters biodiesel density using artificial neural network. *J. Appl. Sci.* 2008; **8**, 3005-11.
[17] J Kumar and A Bansal. Application of artificial neural network to predict properties of diesel biodiesel blends. *Kathmandu Univ. J. Sci. Eng. Tech.* 2010; **6**, 98-103.
[18] X Wei, A Kusiak and H Rahil. Prediction of influent flow rate: a data-mining approach. *J. Energ. Eng.* 2013; **139**, 118-23.
[19] A Verma, X Wei and A Kusiak. Predicting the total suspended solids in wastewater: a data-mining approach. *Eng. Appl. Artif. Intell.* 2013; **26**, 1366-72.
[20] A Kusiak, X Wei, A Verma and E Roz. Modeling and prediction of rainfall using radar reflectivity data: A data-mining approach. *IEEE Trans. Geosci. Rem. Sens.* 2013; **51**, 2337-42.