



# Journal of Applied Sciences

ISSN 1812-5654

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

## Estimation of Vegetable Oil-Based Ethyl Esters Biodiesel Densities Using Artificial Neural Networks

Saeid Baroutian, Mohamed Kheireddine Aroua,  
Abdul Aziz Abdul Raman and Nik Meriam Nik Sulaiman  
Department of Chemical Engineering, Faculty of Engineering, University Malaya,  
50603 Kuala Lumpur, Malaysia

**Abstract:** In this study a new approach based on Artificial Neural Networks (ANNs) has been designed to predict the density of various vegetable oil-based ethyl esters biodiesel. The experimental densities data measured at various temperatures from 15 to 90°C at 1°C interval were used to train the networks. The present work, applied a three layer back propagation neural network with nine neurons in the hidden layer. The results from the network are in good agreement with the measured data and the average absolute percent deviation are 0.35, 0.72, 0.54, 0.68 and 0.72% for the ethyl esters of palm, canola, corn and ricebran oil, respectively. The results of ANNs have also been compared with the results of theoretical estimations.

**Key words:** Biodiesel, ethyl ester, density, neural networks

### INTRODUCTION

Biodiesel is a renewable fuel produced from biological oils and fats, which has many characteristics of a promising alternative energy resource. It has properties similar to ordinary diesel fuel made from crude oil and can be used in conventional diesel engines. The most common process for making biodiesel is known as transesterification. This process involves combining any natural oil (vegetable or animal) with virtually any alcohol and a catalyst. Among the applicable oils and alcohols, vegetable oils and methyl and ethyl alcohol are most frequently used in the biodiesel production.

Density data are important in numerous chemical engineering unit operations. Biodiesel density data as a function of temperature is needed to model the combustion processes and other applications. The density of a ethyl ester biodiesel depends on its molecular weight, free fatty acid content, water content and temperature. As vegetable oil based biodiesel is getting popular as a fuel, a similar procedure for correcting measured density data will be needed. Results of measurements and predictions of specific biodiesel properties have been reported recently but ethyl ester biodiesel density measurements and predictions were rarely done.

Liew *et al.* (1992) determined the densities of the methyl esters of hexanoic, heptanoic, octanoic, decanoic and dodecanoic acids. Tate *et al.* (2006) obtained the

densities of canola, soy and fish oil methyl esters at temperatures up to 300°C. Nouredini *et al.* (1992) has measured the density as a function of temperature for a number of vegetable oils as well as eight fatty acids in the range C9 to C22. Baroutian *et al.* (2008a) measured the density of palm oil-based methyl ester at various temperatures and validated the data obtained was validated with the results of estimation methods.

Neural networks or simply neural nets are computing systems which can be trained to learn a complex relationship between two or more variables or data sets. Neural networks utilize a matrix programming environment making most nets mathematically challenging. The neuron model and the architecture of a neural network describe how a network transforms its input to output. This transformation can be viewed as a computation. Each model and architecture generate limitations on what a particular neural net can compute. The way a network computes its output, is in such a way that the products of the neuron's output and weight are summed with the neuron's bias and passed through the transfer function to get the neuron's output. Neurons may be simulated with or without biases.

The feed forward neural network is one of the most important historical developments in neurocomputing. One of the many benefits of this kind of network is the capability to approximate mathematical function or mapping. Subsequently, it was shown that a feed forward neural network with one hidden layer could approximate any continuous function to any degree of accuracy.

Back propagation was created by generalizing the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify vectors in an appropriate way as defined. Networks with biases, a sigmoid layer and a linear output layer are able to approximate any function with a finite number of discontinuities. Standard back propagation is a gradient descent algorithm, as is the Widrow-Hoff learning rule, in which the network weights are moved along the negative of the gradient of the performance function. The schematic for the back propagation model is shown in Fig. 1. The term back propagation refers to the manner in which the gradient computed for nonlinear multilayer networks.

There are two different ways in which the gradient descent algorithm can be implemented: incremental mode and batch mode. In the incremental mode, the gradient is computed and the weights are updated after each input is applied to the network. In the batch mode all of the inputs are applied to the network before the weights are updated.

Since, the first application to estimate chemical and physical properties of material, Artificial Neural Networks (ANNs) have been established as a dependable method for the achievement of this task. Neural network models have been used for the prediction of biodiesel characteristics, with very good results.

Ramadhas *et al.* (2006) successfully developed a multi-layer feed forward Artificial Neural Networks to predict the Cetane number of biodiesel. Durán *et al.* (2005) used neural networks for estimation of diesel particulate matter composition from transesterified waste oils blends. Baroutian *et al.* (2008b) employed artificial neural

networks to estimate the density of pure palm oil-based methyl ester biodiesel. The results from the network were in good agreement with the measured data and the average absolute percent deviation is 0.29%.

In this study, a new approach based on Artificial Neural Networks (ANNs) has been designed to estimate the density of vegetable oil based ethyl ester biodiesel obtained from various vegetable oils. Measured data of densities at various temperatures from 15 to 90°C were used to train the networks and test the results of it. The present study, applied a three layer back propagation neural network with nine neurons in the hidden layer. Predicted results were also compared with experimental densities and estimated results of theoretical method.

### MODEL SPECIFICATION

In this study, in order to train and validate the neural network, several measured density data of various vegetable oil ethyl esters were used (Durán *et al.*, 2005). The ethyl esters biodiesels obtained from palm, soybean, canola, corn and ricebran oil.

Biodiesels were prepared by transesterification of vegetable oils using ethanol as alcohol source and potassium hydroxide as catalyst in a batch system. The reactions were carried out using 100% excess ethanol, i.e., molar ratio of ethanol to oil is 6:1. Ethyl ester densities were measured at temperatures from 15 to 90°C, measurements were done three times to obtain mean values for each temperature.

The present study, applies a feed forward back propagation neural network in three layers. The input, hidden and output layers had 5, 9 and 1 neurons, respectively as showed in Fig. 2.

Deciding the number of neurons in hidden layer is a very important part of deciding our overall neural network architecture. Though the hidden layer does not directly

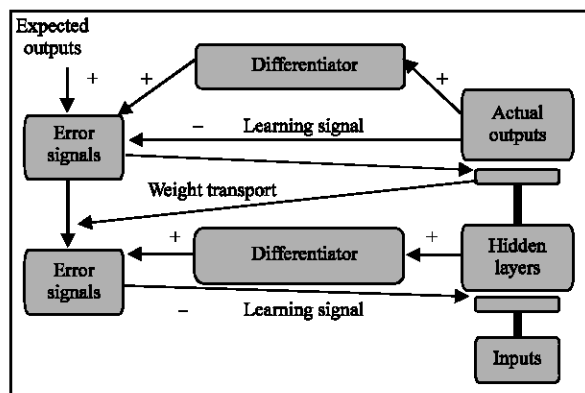


Fig. 1: Back propagation learning schematic (Grossberg, 1987)

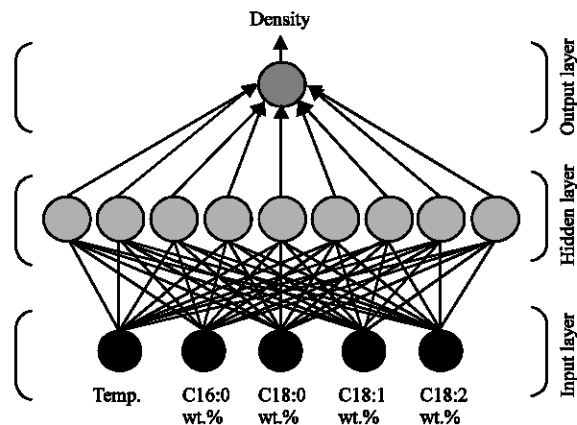


Fig. 2: Feed forward back propagation network with three layers

interact with the external environment, this layer has a tremendous influence on the final output and number of neurons of this hidden layer must be considered.

Using too few neurons in the hidden layers will result in something called underfitting. Underfitting occurs when there are too few neurons in the hidden layers to adequately detect the signals in a complicated data set. Using too many neurons in the hidden layers can result in several problems.

First too many neurons in the hidden layers may result in overfitting. Overfitting occurs when the neural network has so much information processing capacity that the limited amount of information contained in the training set is not enough to train all of the neurons in the hidden layers. The second problem is increasing the time of training.

Obviously, some compromise must be reached between too many and too few look neurons in the hidden layers. The selection of the architecture of the neural network has come down to trial and error. To organize the trial and error search for the optimum network architecture forward selection method was used. This method begins by selecting a small number of hidden neurons (only two hidden neurons).

Subsequently, the neural network is trained and tested. The number of hidden neurons is then increased and the process is repeated so long as the overall results of the training and testing improved. The forward selection method is shown in Fig. 3.

Each layer of this network has its own weight matrix, its own bias vector, a net input vector and an output vector. This network can be used for general function approximation. It has been proven that three layer networks, with sigmoid transfer function in the hidden layer and output layer, can approximate virtually any function of interest to any degree of accuracy, provided that a sufficient amount of hidden units are available.

Therefore, the neuron model key component, the transfer function, is used to design the network and establish its behavior. Sigmoid functions are often used in neural networks to introduce nonlinearity in the model and/or to make sure that certain signals remain within a specified range.

A popular neural net element computes a linear combination of its input signals and applies a bounded sigmoid function to the result; this model can be seen as a smoothed variant of the classical threshold neuron.

A reason for its popularity in neural networks is that the sigmoid function satisfies this property. The derivative of the sigmoid function can be written (Mitchell, 1997):

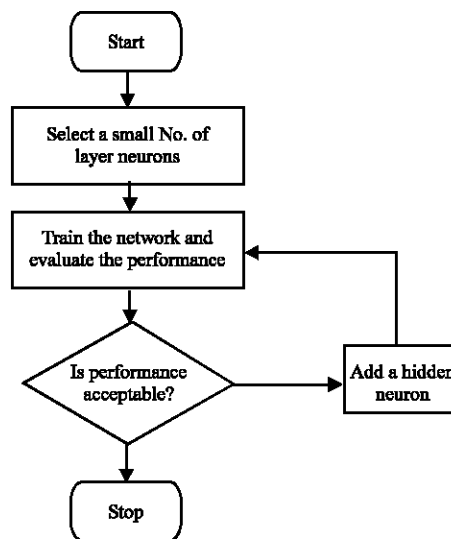


Fig. 3: Selecting the number of hidden neurons with forward selection

$$\frac{d}{dt} \text{sig}(t) = \text{sig}(t)(1 - \text{sig}(t)) \quad (1)$$

This simple polynomial relationship between the sigmoid function and its derivative is computationally easy to perform. It makes the back-propagation based learning easy because, it is differentiable, has simple relation between the function and its derivative, is flexible with easily changeable slope signs and the error convergence criteria of mean square error works well with it.

## MATERIALS AND METHODS

The methodology of the approach used in this study is accomplished by means of Matlab Toolbox. This software has extensive neural net capabilities. The study was conducted in 2008 at the Department of Chemical Engineering, University Malaya.

In this study, ethyl ester biodiesel prepared from transesterification of various vegetable oils and ethyl alcohol in presence of potassium hydroxide catalyst in a batch reaction system. The reactions were carried out using 100% excess ethanol, i.e., molar ratio of ethanol to oil is 6:1. After an hour reaction time in 50°C reaction temperature, the produced biodiesel were analyzed by gas chromatography method to determine the fatty acid ethyl ester composition.

Density measurements were carried out using a DMA 4500 density/specific gravity meter (Anton Paar, Austria). The adjustment of the density meter was checked using

degassed bi-distilled water, the measured value was compared with the corresponding value in the density tables and the accuracy was  $\pm 0.00003 \text{ g cm}^{-3}$ . Ethyl ester densities were measured at temperatures from 15 to 90°C, measurements were done three times to obtain mean values for each temperature, the uncertainty was  $\pm 0.00001 \text{ g cm}^{-3}$ .

Among the 380 density data points, 340 measured densities of various vegetable oil based ethyl esters at different temperature from 15 to 90°C were chosen to train the network and the rest for simulation and evaluate the accuracy of the newly trained network by providing the network a set of data it has never seen.

To improve the learning process fatty acid compositions of ethyl esters were used as input parameters. These fatty acid compositions include ethyl palmitate (C16:0), ethyl stearate (C18:0), ethyl oleate (C18:1) and ethyl linoleate (C18:2).

The procedure to create and train a network using this toolbox was as follows:

- Input (temperature, C16:0 wt.%, C18:0 wt.%, C18:1 wt.% and C18:2 wt.%) and target (density) vectors entered in its suitable format in the workspace of Matlab.
- The vectors normalized independently to assign a number between -1 and 1 to each element of vectors because inputs are sensitive in this range when sigmoid transfer function is used. On the other hand, inputs are normalized because the training domain may be biased toward one input variable or toward higher input. Furthermore, the sigmoid transfer function produces the output within the range of -1 to 1 and if the input is not normalized, bias may be generated.
- A three layer feed forward backpropagation network created in the Matlab neural networks toolbar. This was done because back propagation uses a gradient descent method and it allows the neural network to train with a higher degree of efficiency.
- Trainlm and Tansig chose as training and transfer function, respectively.
- Input and target vectors introduced to the created network and weight initialized.
- Training parameters such as the epochs and error goal adjusted.
- The specified network trained gradually. This process finished when the defined error was reached. During training the weights and biases of the networks were iteratively adjusted to minimize the network performance function.

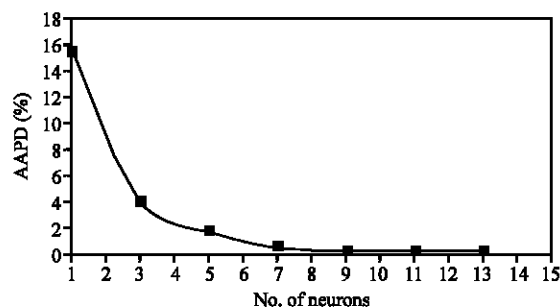


Fig. 4: Average Absolute Percent Deviation (AAPD) parameter of training results versus number of neurons in hidden layer

The optimized number of hidden layer neurons was determined during the learning and training processes by trial and error tests. On the other hand, different numbers of hidden neurons were tested, however since the performance did not change significantly with more neurons, the simplest network was chosen. To show optimization, 1, 3, 5, 7, 9, 11 and 13 neurons were chosen for training process.

As can be shown in Fig. 4, the best choice is the network with nine neurons in the hidden layer which is in good agreement with the experimental data. In Fig. 4, the Average Absolute Percent Deviation (AAPD) parameter of training process using different number of hidden layer neurons is shown according to the expression:

$$AAPD = \frac{1}{N} \sum_{i=1}^N \left| \left( \frac{\text{Density}_{\text{experimental}} - \text{Density}_{\text{estimated}}}{\text{Density}_{\text{experimental}}} \right) \right| \times 100 \quad (2)$$

After training the three-layer, feed forward, back propagation network, the vegetable oil ethyl esters biodiesel densities at other temperatures were predicted from the simulation of this network with the suitable inputs.

## RESULTS AND DISCUSSION

As can be shown in Fig. 5, there is a very good agreement between the measured data (normalized density) and the trained data.

This illustrates that the networks has been trained very well and can be used to simulate the biodiesel density at a wide range of temperatures.

There is a good agreement between the experimental data and the simulated data. The equations of the form  $y = f(x)$  in the Fig. 5 and 6 are the equations of the regression lines. When, all the points fall exactly on the line of 45°, the regression line is  $y = x$ . In this case, the

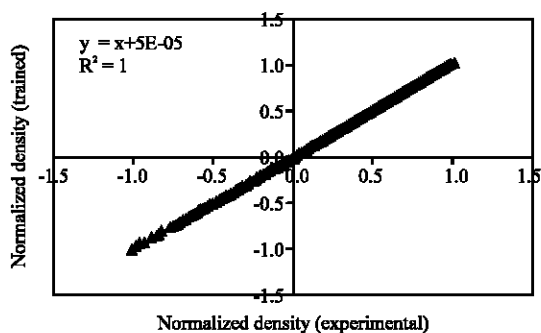


Fig. 5: Comparison between the training results and the measuring data (Normalized data)

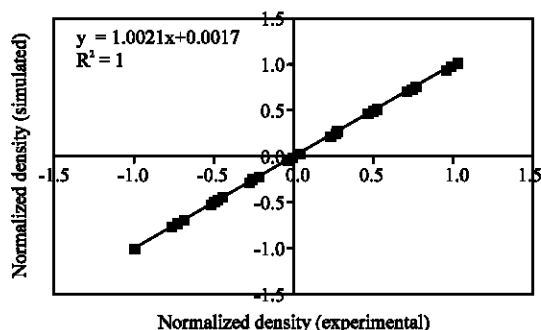


Fig. 6: Comparison between the simulating results and the measuring data (Normalized data)

network is trained or simulated very well. Otherwise, the regression line will have the form of  $y = ax + b$ . The regression constants ( $R^2$ -value) which are also shown in Fig. 5 and 6 show the agreement of trained and simulated data with experimental data.

In Fig. 7, the measured densities of ethyl ester biodiesels obtained from various vegetable oils, including palm, soybean, canola, corn and ricebran oil are compared with those predicted by the ANNs. It can be seen that there is a very good agreement between the results of the ANNs and the measured data. The AAPD parameters for neural network simulation are 0.35, 0.72, 0.54, 0.68 and 0.72% for the ethyl esters of palm, canola, corn and ricebran oil, respectively.

In Fig. 8-12, the predicted densities by ANNs are compared with those predicted by the Spencer and Danner method (Baroutian *et al.*, 2007c). The AAPD parameters for the result of Spencer and Danner method are 6.86, 13.5, 7.37, 8.09 and 6.34% for the ethyl esters of palm, canola, corn and ricebran oil, respectively. It can be seen that the ANNs give a better prediction with less deviation than those given by Spencer and Danner method. Although, the Spencer and Danner method has good prediction reliability at lower temperatures but at higher than 50°C ANNs are more reliable.

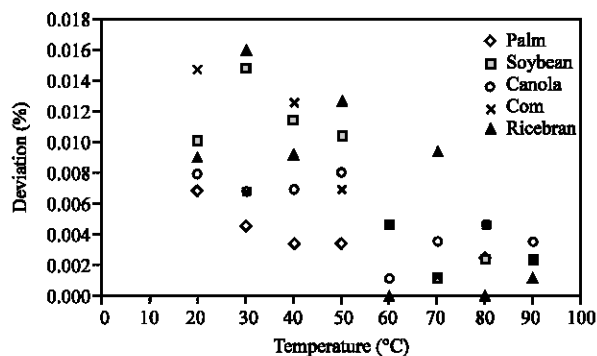


Fig. 7: Comparison of ethyl ester density estimation deviations from experimental data

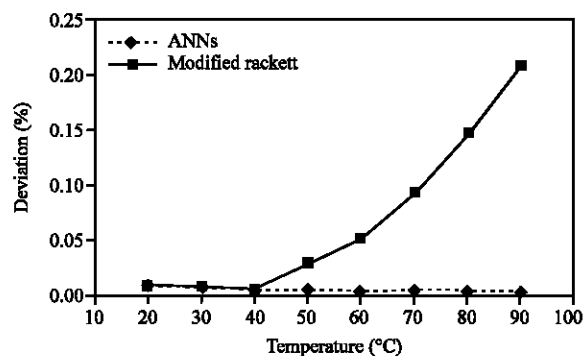


Fig. 8: Palm ethyl ester density deviation from experimental data, comparison of ANNs and modified rackett method

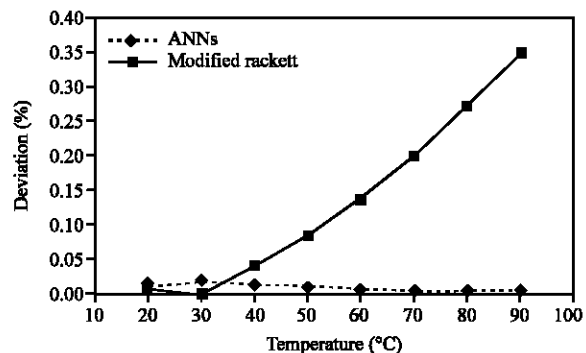


Fig. 9: Soybean ethyl ester density deviation from experimental data, comparison of ANNs and modified rackett method

The better predictability of ANNs can be attributed to its adaptability that changes its structure based on external or internal information that flows through the network during the learning phase. ANNs are rich and flexible systems that show robust performance in dealing

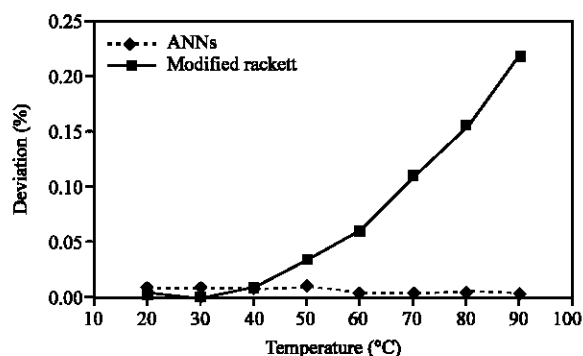


Fig. 10: Canola ethyl ester density deviation from experimental data, comparison of ANNs and modified rackett method

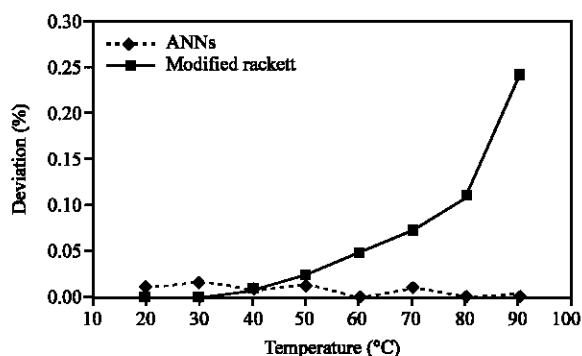


Fig. 12: Ricebran ethyl ester density deviation from experimental data, comparison of ANNs and modified rackett method

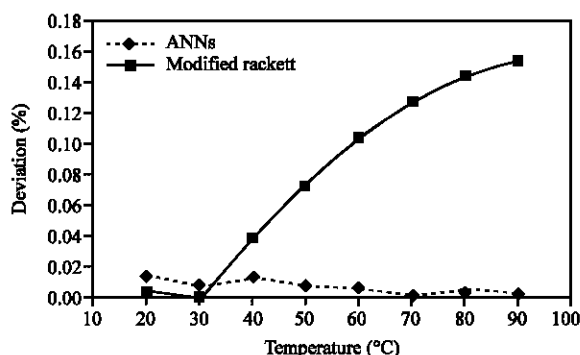


Fig. 11: Corn ethyl ester density deviation from experimental data, comparison of ANNs and modified rackett method

with noisy or incomplete data and have the ability to generalize from the input data. They may be better suited than other modeling systems to predict outcomes. In this study due to the sufficient amount of training data, the network trained properly and the outputs were accurate. Spencer and Danner method use the modified Rackett equation with critical properties of mixtures to estimate liquid density (Spencer and Danner, 1972).

The Rackett equation modified by Spencer and Danner to estimate liquid density  $\rho$  ( $\text{g cm}^{-3}$ ) is (Spencer and Danner, 1972):

$$\rho = \frac{P_R}{Z_{RA}^*} \quad (3)$$

Where:

$$\phi = (1 - T/T_c)^{2/7} - (1 - T_R/T_c)^{2/7} \quad (4)$$

In Eq. 3 and 4,  $Z_{RA}$  is Rackett compressibility factor and can be determined for biodiesel using measured

densities and applying the Eq. 3.  $T(\text{K})$ ,  $T_R(\text{K})$ ,  $T_c(\text{K})$  and  $\phi$  are temperature, reduced temperature, critical temperature and fugacity coefficient, respectively.

To determine the mixture critical properties, the Lee-Kesler mixing rules can be used (Plocker *et al.*, 1978). The equations to calculate critical temperature of mixture  $T_{cm}$  (K),  $P_{cm}$  (bar) and  $V_{cm}$  ( $\text{cm}^3 \text{mol}^{-1}$ ) are:

$$T_{cm} = \frac{1}{V_{cm}^{1/4}} \sum_i \sum_j x_i x_j V_{cij}^{1/4} T_{cij} \quad (5)$$

$$V_{cm} = \sum_i \sum_j x_i x_j V_{cij} \quad (6)$$

$$\omega_m = \sum_i x_i \omega_i \quad (7)$$

$$T_{cij} = (T_a T_g)^{1/2} k'_{ij} \quad (8)$$

$$V_{cij} = \frac{1}{8} (V_a^{1/3} + V_g^{1/3})^3 \quad (9)$$

$$P_{cm} = (0.29050.085 \omega_m) R T_{cm} / V_{cm} \quad (10)$$

where,  $x$  is mole fraction,  $\omega$  is acentric factor and  $R$  is gas constant ( $\text{mL} \cdot \text{bar} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ ). The fatty acid composition of the ethyl esters are shown in Table 1.

The critical properties of each pure constituent (fatty acids of ethyl esters) are estimated by the Joback modification of Lydersen's method (Poling *et al.*, 2000) using the following relations:

$$T_c = T_b \left[ 0.584 + 0.965 \left\{ \sum_k N_k (tck) \right\} - \left\{ \sum_k N_k (tck) \right\}^2 \right]^{-1} \quad (11)$$

**Table 1: Fatty acid mole fraction of vegetable oil ethyl esters**

Fatty acid ethyl ester	Palm x	Soybean x	Canola x	Corn x	Ricebran x
C 16:0	0.47	0.14	0.06	0.11	0.19
C 18:0	0.04	0.05	0.02	0.03	0.05
C 18:1	0.38	0.25	0.71	0.29	0.32
C 18:2	0.11	0.56	0.21	0.57	0.44

$$P_c = \left[ 0.113 + 0.0032N_{atoms} - \sum_k N_k (pck) \right]^{-2} \quad (12)$$

$$V_c = 17.5 + \sum_k N_k (vck) \quad (13)$$

## CONCLUSION

The approach presented in this study speeds up the process of prediction of the densities of various vegetable oil-based ethyl ester biodiesels. This new approach is based on artificial neural networks to estimate the density. This method is able to predict the density at various temperatures. The comparison of the results obtained by Artificial Neural Networks (ANNs) with those predicted by Spencer and Danner method (Spencer and Danner, 1972) shows the reliability of ANNs over the theoretical methods. Finally, good agreement between measured data and the result of artificial neural network shows that the ANNs can be a powerful model for predicting density for the ethyl esters biodiesel of different vegetable oils.

Furthermore, this approach provides a new way to estimate the density of ethyl ester biodiesel with respect to available methods accurately.

## REFERENCES

- Baroutian, S., M.K. Aroua, A.A. Raman and N.M. Sulaiman, 2008a. Density of palm oil-based methyl ester. *J. Chem. Eng. Data*, 53: 877-880.
- Baroutian, S., M.K. Aroua, A.A. Raman and N.M. Sulaiman, 2008b. Prediction of palm oil-based methyl ester biodiesel density using artificial neural networks. *J. Applied Sci. Data*, 8: 1938-1943.
- Durán, A., M. Lapuerta and J. Rodríguez-Fernández, 2005. Neural networks estimation of diesel particulate matter composition from transesterified waste oils blends. *J. Fuel.*, 84: 2080-2085.
- Grossberg, S., 1987. Competitive learning: From interactive activation to adaptive resonance. *J. Cognitive Sci.*, 11: 23-63.
- Liew, K.Y., C.E. Seng and L.L. Oh, 1992. Viscosities and densities of the Methyl Esters of some n-alkanoic acids. *J. Am. Oil Chem. Soc.*, 69: 155-158.
- Mitchell, T.M., 1997. *Machine Learning*. 1st Edn., WCB-McGraw-Hill, Boston, ISBN: 0070428077, pp: 96-97.
- Noureddini, H., B.C. Teoh and L.D. Clements, 1992. Densities of vegetable oils and fatty acids. *J. Am. Oil Chem. Soc.*, 69: 1184-1188.
- Plocker, U., H. Knapp and J. Prausnitz, 1978. Calculation of high-pressure vapor-liquid equilibria from a corresponding states correlation with emphasis on asymmetric mixtures. *J. Ind. Eng. Chem. Process Design Dev.*, 17: 324-332.
- Poling, B.E., J.M. Prausnitz and J.P. O'Connell, 2000. *The Properties of Gases and Liquids*. 5th Edn., McGraw-Hill, New York, ISBN: 0070116822/9780070116825, pp: 14-15.
- Ramadhas, A.S., S. Jayaraj, C. Muraleedharan and K. Padmakumari, 2006. Artificial neural networks used for the prediction of the octane number of biodiesel. *J. Renewable Energy*, 31: 2524-2533.
- Spencer, C.F. and R.P. Danner, 1972. Improved equation for prediction of saturated liquid density. *J. Chem. Eng. Data*, 17: 236-241.
- Tate, R.A., K.C. Watts, C.A.W. Allen and K.I. Wilkie, 2006. The densities of three biodiesel fuels at temperatures up to 300°C. *J. Fuel.*, 85: 1004-1009.