

# Estimation of Biodiesel Properties from Chemical Composition – An Artificial Neural Network (ANN) Approach

**M. I. Jahirul, W. Senadeera, R. J. Brown and L. Moghaddam**  
Biofuel Engine Research Facility, Science and Engineering Faculty  
Queensland University of Technology (QUT), Brisbane, Australia  
[md\\_jahirul@yahoo.com](mailto:md_jahirul@yahoo.com); [mj.islam@student.qut.edu.au](mailto:mj.islam@student.qut.edu.au);

**Abstract**— Biodiesel, produced from renewable feedstock represents a more sustainable source of energy and will therefore play a significant role in providing the energy requirements for transportation in the near future. Chemically, all biodiesels are fatty acid methyl esters (FAME), produced from raw vegetable oil and animal fat. However, clear differences in chemical structure are apparent from one feedstock to the next in terms of chain length, degree of unsaturation, number of double bonds and double bond configuration-which all determine the fuel properties of biodiesel. In this study, prediction models were developed to estimate kinematic viscosity of biodiesel using an Artificial Neural Network (ANN) modelling technique. While developing the model, 27 parameters based on chemical composition commonly found in biodiesel were used as the input variables and kinematic viscosity of biodiesel was used as output variable. Necessary data to develop and simulate the network were collected from more than 120 published peer reviewed papers. The Neural Networks Toolbox of MatLab R2012a software was used to train, validate and simulate the ANN model on a personal computer. The network architecture and learning algorithm were optimised following a trial and error method to obtain the best prediction of the kinematic viscosity. The predictive performance of the model was determined by calculating the coefficient of determination ( $R^2$ ), root mean squared (RMS) and maximum average error percentage (MAEP) between predicted and experimental results. This study found high predictive accuracy of the ANN in predicting fuel properties of biodiesel and has demonstrated the ability of the ANN model to find a meaningful relationship between biodiesel chemical composition and fuel properties. Therefore the model developed in this study can be a useful tool to accurately predict biodiesel fuel properties instead of undertaking costly and time consuming experimental tests.

**Keywords**— Biodiesel; Kinematic viscosity; Artificial Neural Network (ANN); Prediction model

## I. INTRODUCTION

Vegetable oil methyl or ethyl esters, commonly referred to as 'biodiesel', are a renewable liquid fuel alternative to petroleum diesel. In technical terms, biodiesel is diesel engine fuel comprised of mono-alkyle esters of long chain fatty acids derived from vegetable oil or animal fats, designated B100

and meeting the requirements of ASTM D 6751 or EN 14214 [1].

These mono-alkyl esters are the main chemical species that give biodiesel similar or better fuel properties compared with petroleum diesel [2]. It is also safer to handle, store and transport because it is biodegradable, non-toxic and has a higher flash point than diesel. One of the major advantages of biodiesel is that it has potential to reduce dependency on imported petroleum through the use of domestic feedstock for production [1-2]. However, as a fuel, there are currently several disadvantages to using biodiesel in diesel engine applications. The major disadvantage is its higher viscosity compared to conventional diesel fuel especially at lower temperatures. These differences mainly result from the difference in chemical compositions between petroleum diesel and biodiesel.

Biodiesels are usually made from vegetable oils and animal fat feedstock through chemical reaction called Transesterification. In this process the pure oil and fat is converted from natural oil (three long chain carbon molecules struck together by glycerin) into three monoalkyl esters (three separated long chain carbon molecules) shown in Fig. 1. Triglycerides are allowed to react with an alcohol (mostly methanol) under acidic or basic catalyst conditions producing fatty acid ester of the respective alcohol and free glycerol. After complete reaction, glycerol is removed as a byproduct and esters are known as biodiesel. However, biodiesels are not pure methyl esters whereas small amount of impurities are exist in commercial biodiesel depending on production process. The most common impurities are includes, free fatty acid, free glycerin, monoglyceride, diglyceride and triglyceride which eventually affects the physical properties of biodiesel [3].

Viscosity is defined as the resistance to shear or flow which describes the behavior of a liquid in motion near a solid boundary such as the walls of a pipe. It is one of the most critical features of a fuel which plays a dominant role in fuel spray, fuel-air mixture formation and the combustion process [4-6]. In a diesel engine, fuel is sprayed into combustion

chamber and then atomized into small drops near the nozzle exit. The liquid fuel forms a cone-shaped spray at the nozzle exit and its viscosity has significant effects on the size of fuel drop and penetration. Therefore, fuel viscosity influences engine combustion, performance and emissions, especially carbon monoxide (CO) and unburnt hydrocarbon (UHC) [7]. However, very low fuel viscosity is not desirable because the fuel then doesn't provide sufficient lubrication for the precision fit of fuel injection pumps, resulting in leakage or increased wear. Therefore, all biodiesel standards specify the upper and lower limits of kinematic viscosity for biodiesel.

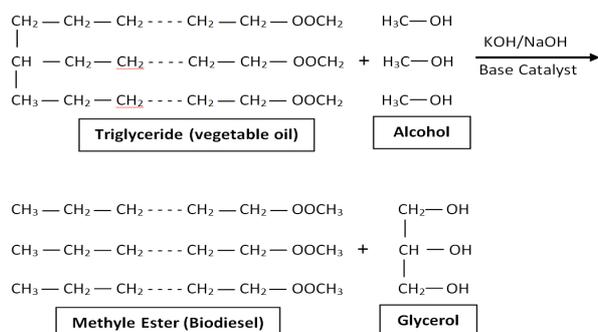


Fig 1. Transesterification reaction [3]

The viscosity of biodiesel is dependent on its complex chemical composition and impurities. A recent study showed that viscosity increases with increasing length of both the fatty acid chain and alcohol group [8]. As the lengths of the acid and alcohol segments in the ester molecules increased, so did the degree of random intermolecular interactions and consequently viscosity. Refaat [9] reported that shorter fatty acid chains with longer alcohol moieties display lower viscosity than ester with longer fatty acid chains and shorter alcohol moieties. Other factors that influence biodiesel viscosity include: number and position of double bonds, degree of saturation, molecular weight, branching hydroxyl groups and the amount of impurities, such as unreacted glycerides or glycerol etc [3].

In recent years, ANN modelling techniques have gained in popularity due to their ability to accurately predict from small data sets (examples) rather than from larger data sets requiring costly and time-consuming studies and experiments. The most important feature of artificial neural networks is their ability to solve problems through learning by example, rather than by becoming involved in the detailed characteristics of the systems. This feature of makes them very useful because they works like a 'black box' model and do not require detail or complete information about the problem, and when there are only sets of data inputs and outputs of the system. Therefore it has been successfully applied in various disciplines, including neuroscience [10], mathematical and computational analysis [11], learning systems [12] and engineering design and application [13-15]. In this paper, the potential of ANN modelling techniques in identifying sustainable future

generation biodiesel feedstock are identified based on the most recent literature. ANN also applied to predict important fuel properties of biodiesel. Baroutian *et al.* [16] predicted density from fatty acid profile for palm oil biodiesel. Ramadas *et al.* [17] used ANN to predict the cetane number of biodiesel based on fatty acid profile. ANN also has been used to predict viscosity, flash point, fire point based on diesel-biodiesel blend ratio [18]. However these prediction models are limited for a specific biodiesel and experiment condition. No investigation was to develop ANN model to predict kinematic viscosity of biodiesel for a wide ranges of biodiesel feedstock and consider the impurities generally contained in biodiesel. Such an ANN prediction model has been developed in this study. The developed ANN model has been simulated with new input data and prediction ability was presented graphically.

## II. ARTIFICIAL NEURAL NETWORKS (ANN)

The foundation of artificial neural networks (ANN) in a scientific sense begins with a biological neuron. In the brain, there is a flow of coded information (using electrochemical media, the so-called neurotransmitters) from the synapses towards the axon. The axon of each neuron transmits information to a number of other neurons. Groups of neurons are organized into sub-systems and the integration of these sub-systems forms the brain. On the other hand, an ANN is composed of a large number of simple processing units called neurons which are fully connected to each other through adoptable synaptic weight. This resembles a brain in two aspects. Knowledge can be acquired through training and knowledge can be stored.

A widely used ANN model called the multi-layer perception (MLP) neural network, which consist of one input layer, one or more hidden layer (s) (middle) in between input and output layers and one output layer. Each layer employs several neurons (nodes), and each neuron in a layer is connected to the neurons in the adjacent layer with different weights. Signals flow into the input layer, pass through the hidden layer(s), and arrive at the output layer. With the exception of the input layer, each neuron receives signals from the neurons of the previous layer. The incoming signals or input ( $x_{ij}$ ) are multiplied by the weights ( $v_{ij}$ ) and summed up with the bias ( $b_j$ ) contribution. Mathematically it can be expressed as:

$$\text{net}_j = \sum_{i=1}^n X_i V_{ij} + b_j \quad (1)$$

The output of a neuron is determined by applying an activation function to the total input and calculated using equation 1. If the computed outputs do not match the known (*i.e.* target) values, NN model is in error. Then, a portion of this error is propagated backward through the network. This error is used to adjust the weight and bias of each neuron throughout the network so the next iteration error will be less for the same units. The procedure is applied continuously and

repetitively for each set of inputs until there are no measurable errors, or the total error is smaller than a specified value.

The most important feature of artificial neural networks is their ability to solve problems through learning by example, rather than by fully understanding the detailed characteristics of the systems. This feature makes it very useful because it works like a 'black box' model, and does not require detail or complete information about the problem, and can be utilized when all that is available are sets of data inputs and outputs of the system. It has a natural propensity to store experiential knowledge and to make it available for use (Fig. 2). Therefore, this nonlinear computer algorithm can model large and complex systems with many interrelated parameters.

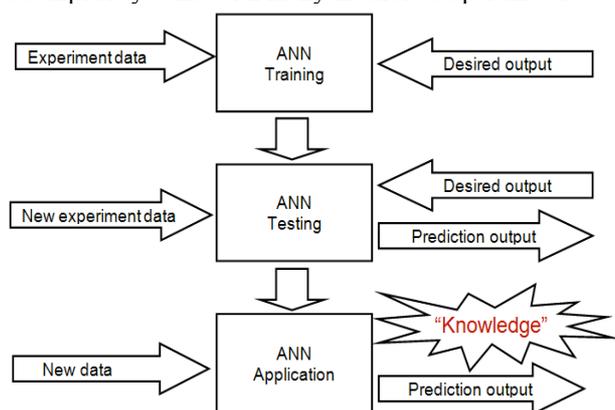


Fig. 2. Working principle of an ANN [3]

Fig. 3 shows several stages are involved in ANN prediction model developing process. An important stage of a neural network is the training step, in which an input is introduced to the network together with the desired output: the weights and bias values are initially chosen randomly and the weights adjusted so that the network produces the desired output. After training, the weights contain meaningful information, contrary to the initial stage where they are random and meaningless. When a satisfactory level of performance is reached, the training stops, and the network uses the weights to make decisions.

### III. DATA COLLECTION

Necessary data to train an ANN for kinematic viscosity prediction of biodiesel were collected from papers published in recognized international Journals, conferences and report of renowned research centers around the world. The popular scientific and electronic databases including Elsevier, Taylor and Francis, DieselNet, Scopus, Springer, Wiley international, American chemical society, IEEE, SAGA Publication, MDPI etc. were searched for relevant papers for this study. More than 120 papers were collected mostly published in last decade which contain experimental results of chemical composition and kinematic viscosity of biodiesels at 40 °C. Fig. 1 shows the absolute number and the percentage of the various biodiesel datasets included in the current investigation. The collected data comprises biodiesels (fatty acid methyl esters) from 55 different feedstock, 5 biodiesel-

biodiesel blends and 59 pure methyl esters. From the Fig. 1 it can be seen a large number of feedstock were investigated worldwide for biodiesel production, those includes edible

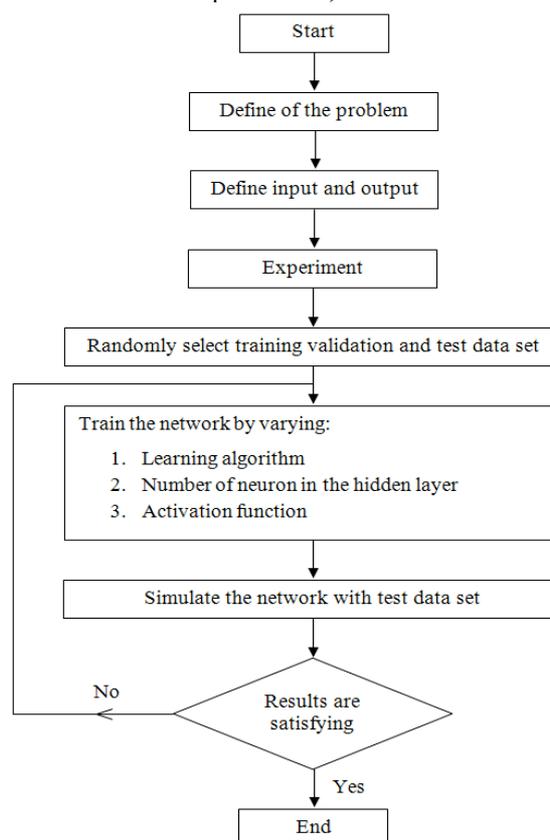


Fig. 3. Proposed flow chart of an ANN prediction model development

non-edible vegetable oils, west cooking oils, beef tallow, chicken fats, fish oils, algae etc. It is also interesting to see that, as a representative of actual biodiesel, many investigations have used pure methyl esters which are mostly artificially through chemical process. The most popular feedstock of biodiesel investigated worldwide is soybean (11.36%) followed by palm (6.25%), sunflower (6.25%), canola (5.11%), rapeseed (4.83%) as edible oil feedstocks. Among the non-edible oil, most investigated feedstock is Jathropa found in 15 papers. During data collection stage, care have been taken to only include data from those researchers who actually measured themselves the values reported, following the internationally accepted experimental standards.

Altogether, 352 data sets of biodiesel chemical compositions and corresponding kinematic viscosity were selected to use in this study. The main chemical compositions in biodiesels are mono-alkyle esters of fatty acids which are feedstock dependent and vary significantly from to the next. Therefore a wide range of fatty acid profiles found in the collected data that are tabulated in Table I with their structure and name. These fatty acids are straight-chain compounds ranging in size from 8–24 carbons and are mainly the three



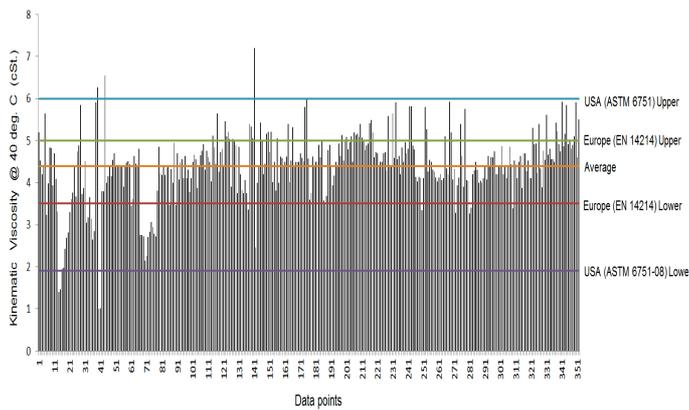


Fig. 6. Kinematic viscosities of the samples with biodiesel standards

Due to the variation in chemical composition, the kinematic viscosities of biodiesels differ significantly which is evident in Fig. 6. The kinematic viscosity of biodiesels collected in this study ranges from 1.4 to 7.2 cSt with the mean value of 4.39 cSt. Most of the viscosity data are within the acceptable range of USA (ASTM 67) biodiesel standard which is from 1.9 to 6 cSt. However, the European biodiesel standard (EN 14214) impose more strict kinematic viscosity regulation for biodiesel which a ranges from 3.5 to 5 cSt and many biodiesels shown in Fig. 6 fail to meet this standard

#### IV. PREDICTION MODEL DEVELOPMENT

The kinematic viscosity prediction model was developed for biodiesel using an ANN which represents a mathematical relationship between input and output parameters of a system as such a black box model. The selection of input parameters which contribute to the output is therefore a crucial task. It is also desirable to minimize the number of input parameters for an ANN system in order to reduce the computational time. In general, the best input parameters are being selected based on an understanding of the physics of a problem. Published literature suggested that the kinematic viscosity of biodiesel is a function of chemical composition- the fatty acid profile of methyl ester and impurities. In the present study, 27 variables are used as input parameters in developing the ANN for viscosity prediction. These parameters include: mass percentage of 18 fatty acid methyl ester that is commonly found in the biodiesel and 9 parameters listed in Table 2. Among 352 data sets that have been collected from literature, 327 sets are used in the ANN model training process whereas other 25 data sets are randomly selected for simulation.

The three data sets needed to develop ANNs: for training, validation and testing the network. The usual approach is to prepare a single data-set, and differentiate it by a random selection. The learning algorithm called the feed-forward back-propagation was applied for the single hidden layer shown in the Fig. 7. The Neural Network was optimized using the MATLAB Version R2012a Neural Network

Toolbox. In the training stage, to define the output accurately, the number of neurons in hidden layer and learning algorithm has been optimized using trial and error methods. In the optimisation process, networks were trained with 12 algorithms available in the software. These are, Bayesian Regularization (BR), Quasi-Newton (QN), Resilient Backpropagation (RB), Scaled Conjugate Gradient (SCG), Levenberg-Marquardt (LM), Conjugate Gradient with Powell Restarts (CGPR), Fletcher-Powell Conjugate Gradient (FPCG), Polak-Ribière Conjugate Gradient (PRCG), One Step Secant (OSS), Variable Learning Rate Gradient Descent (VLRGD), Gradient Descent with Momentum (GDM) and Gradient Descent (GD). For each algorithm, the prediction accuracy of the model was observed for 10 to 60 neurons in the single hidden layer with 5 intervals. It was found that prediction accuracy improved with the number neurons in the hidden layer. Fig. 8 shows the results of the network optimisation process. The best results was found using LM learning algorithm with 45 neuron in the hidden layer. Moreover, among the 12 algorithm use in this study, LM learning algorithms showed better results followed by GD, OSS and CGPR. With a very high number of neurons in hidden layer, the computational time increased significantly along with the over fitting problems which eventually reduced the models effectiveness. Neurons in the input layer have no transfer function. Logistic sigmoid (logsig) transfer function has been used in the hidden layer while a purelinear (purelin) transfer function has been used in output layer. The optimized ANN structure is shown in Fig. 9. The training was stopped after 53 iterations as shown in Fig. 10 to avoid over fitting of the network. After stopping the training the software has identified the best ANN network at 47 iterations.

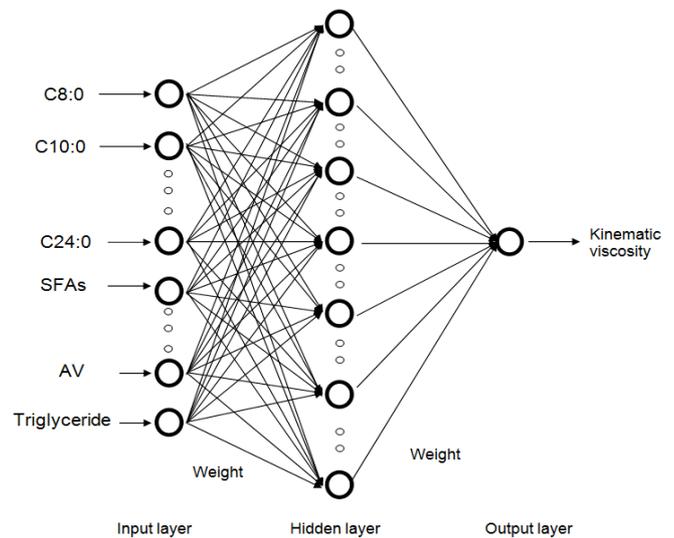


Fig 7. Structure of ANN

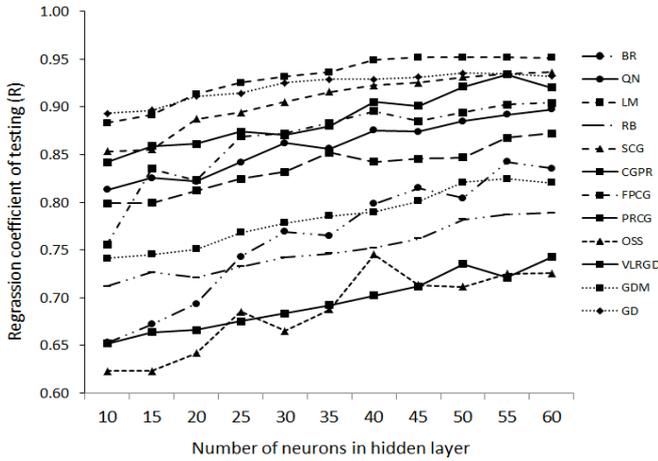


Fig. 8. Optimisation of ANN model

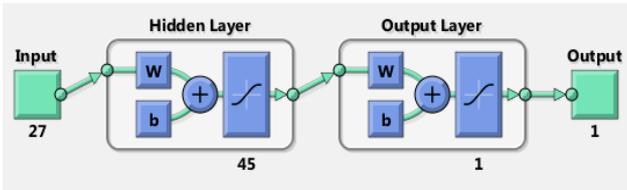


Fig. 9. Optimised ANN for kinematic viscosity prediction

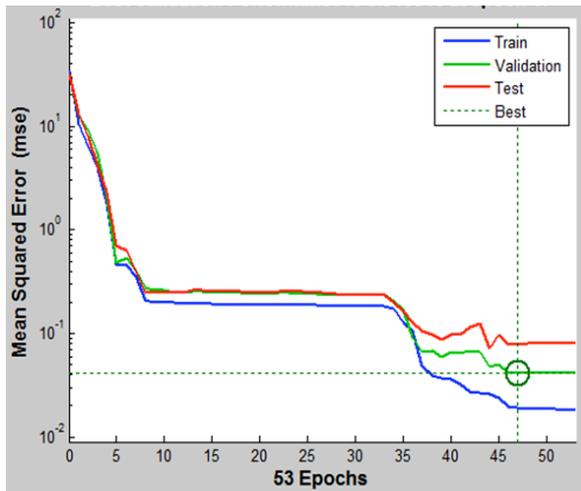


Fig. 10. Training performance of ANN model

## V. SIMULATION

The ability of the ANN model to predict kinematic viscosity has been verified by simulating it with unknown input variables. For this purpose, 25 sets of data selected randomly from collected data sets which was not used during the training of ANN model. After simulation, the predicted kinematic viscosity of biodiesel was compared with the corresponding experimental values and shown in Fig. 6. Using the results produced by the network, statistical methods have been used to investigate the prediction performance of NN results. To judge the prediction performance of a network, several performance measures are used. Those include

statistical analysis in terms of absolute fraction of variance ( $R^2$ ), root mean squared (RMS) and maximum average error percentage (MAEP). Formulas to calculate the error parameters are shown in Equation 2 to 4.

$$R^2 = 1 - \left( \frac{\sum_{i=1}^{i=N} (E_a - E_p)^2}{\sum_{i=1}^{i=N} (E_a - E_m)^2} \right) \quad (2)$$

$$RMS = \sqrt{\frac{\sum_{i=1}^{i=N} (E_a - E_p)^2}{N}} \quad (3)$$

$$MEP = \frac{1}{N} \sum_{i=1}^{i=N} \left( \frac{|E_a - E_p|}{E_a} \times 100 \right) \quad (4)$$

where

- $E_a$ -Actual result
- $E_p$ -Predicted result
- $E_m$ -Mean value
- $N$ -Number of pattern

Actual and predicted data shown in Fig. 11 indicates the good prediction ability of the developed ANN for predicting kinematic viscosity of biodiesel for wide ranges of biodiesel feedstocks. In statistical analysis it was found that the absolute fraction of variance ( $R^2$ ) was close to unity 0.93, Root-Mean-Squared (RMS) error was 0.18 cSt and maximum average error percentage (MAEP) was 4.65%. This is a good correlation between the measured and predicted kinematic viscosity. The network has been trained well and can be used to simulate biodiesel kinematic viscosity over a wide range of feedstock. However, the prediction accuracy of the model should be further improved by increasing the number and ranges of training data set.

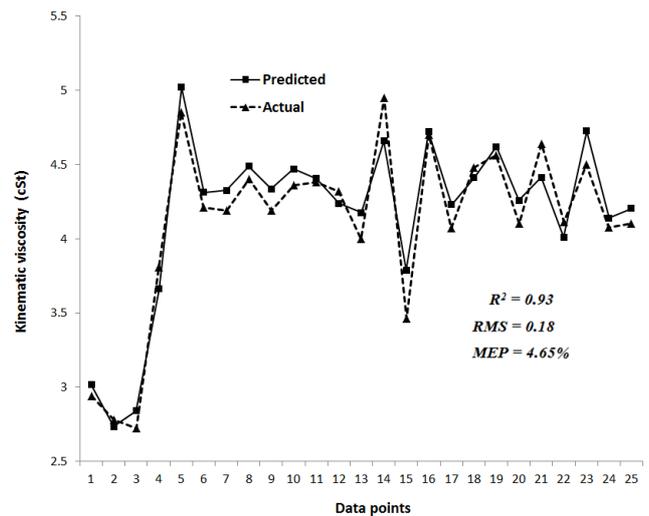


Fig. 11. Experimental and predicted kinematic

## VI. CONCLUSION

The aim of this paper was to investigate the suitability of using artificial neural networks (ANN) for the prediction of biodiesel kinetic viscosity from its chemical composition. To train the network, experimental data was collected from more than 120 peer reviewed paper published in recognized journal, conferences and reports. Altogether, 352 data sets which included biodiesel chemical compositions and corresponding kinematic viscosity at 40 °C of biodiesel were collected. The collected data comprises biodiesels (fatty acid methyl esters) from 55 different feedstocks, 5 biodiesel-biodiesel blends and 59 pure methyl esters. In the training stage, to define the output accurately, the number of neurons in hidden layer and learning algorithm was optimized using trial and error methods. The best network for this study was a standard back propagation (BP) neural network model with LM algorithms and 45 neurons in hidden layer. The performance of the developed ANN prediction model was evaluated by prediction with the 25 data sets which were not used in the training process. The network produced the predicted results in good argument to the experimental ones. The overall results show that the networks can be used as an alternative way for predicting kinematic viscosity of biodiesel at different temperature conditions. The absolute fraction of variance (R<sup>2</sup>), Root-Mean-Squared (RMS) and maximum average error percentage (MAEP) was values were found 0.93, 0.18 cSt and 4.65% respectively which is within an acceptable range of accuracy. The results of this study also show that an ANN has the ability to learn and generalize a wide range of experimental conditions. Therefore, the usage of ANNs may be recommended to optimize the chemical composition of biodiesels to optimize fuel quality for internal combustion engine application. However the network should be further improved by including additional robust data set during the training process.

## VII. ACKNOWLEDGMENTS

The authors wish to express their thanks to QUTPRA scholarship, Faculty of Science and Engineering (SEF), Queensland University of Technology (QUT) for proving fund for this study.

## VIII. REFERENCE

- [1]. A. Demirbas, "Relationships derived from physical properties of vegetable oil and biodiesel fuels" *Fuel*, 2008, Vol. 87, pp. 1743-1748.
- [2]. S. Fernando, P. Karra, R. Hernande, and J. S. Kumar, "Effect of incompletely converted soybean oil on biodiesel quality" *Energy*, 2007, Vol. 32, pp. 844-851.
- [3]. M. I. Jahirul, R. J. Brown, W. Senadeera, Z D Ristovski, and I. O'Hara, "Artificial neural network approach in identifying sustainable future generation biofuel feedstock" *Energies*. Special issue: Alternative Fuels for the Internal Combustion Engines (ICE), 2013, Vol. 6, pp. 3764-3806.
- [4]. C. Y. Choi and R. D. Reitz, "A numerical analysis of the emissions characteristics of biodiesel blended fuels" *J Eng Gas Turbines Power*, 1999, Vol. 121, pp. 31-37.
- [5]. S. Lee, D. Tanaka, J. Kusaka, and Y. Daisho, "Effects of diesel fuel characteristics on spray and combustion in a diesel engine" *JSAE*, 2002, Vol. 23, pp. 407-414.
- [6]. M. E. Tat, and J.V. Gerpen, Biodiesel blend detection with a fuel composition sensor. *Trans ASAE* 2002, Vol. 19, pp. 30-36.
- [7]. G. Knothe, and K. R. Steidle, "Kinematic viscosity of biodiesel fuel components and related compounds Influence of compound structure and comparison to petrodiesel fuel components" *Fuel*, 2005, Vol. 84(9), pp. 1059-1065.
- [8]. D. P. Geller, and J. W. Goodrum, "Effects of specific fatty acid methyl esters on diesel fuel lubricity" *Fuel* 2004, Vol. 83, pp. 2351-56.
- [9]. A. A. Refaat, Correlation between the chemical structure of biodiesel and its physical properties. *Int J Environ Sci Technol* 2009, Vol. 6, pp. 677-694.
- [10]. E. Alkim, E. Gürbüz, and E. Kılıç, "A fast and adaptive automated disease diagnosis method with an innovative neural network model" *Neural Networks*, 2012 Vol. 330, pp. 88-96.
- [11]. M. A. Costa, A. P. Braga, and B. R. Menezes, "Convergence analysis of sliding mode trajectories in multi-objective neural networks learning" *Neural Networks*, 2012, Vol. 330 pp.21-31.
- [12]. S. Carrillo, J. Harkin, L. McDaid, S. Pande, S. Cawley, B. McGinley, and F. Morgan. "Advancing interconnects density for spiking neural network hardware implementations using traffic-aware adaptive network-on-chip routers" *Neural Networks* 2012, Vol. 330, pp. 42-57.
- [13]. T. Samura, and H. Hayashi, "Directional spike propagation in a recurrent network: Dynamical firewall as anisotropic recurrent inhibition" *Neural Networks*, 2012, Vol. 330, 236-246.
- [14]. D. Gao, Z. Yang, C. Cai, and F. Liu, "Performance evaluation of multilayer perceptrons for discriminating and quantifying multiple kinds of odors with an electronic nose" *Neural Networks* 2012, Vol. 330 pp. 204-215.
- [15]. R. C. J. Minnett, A. T. Smith, W. C. Lennon, R. Hecht-Nielsen, "Neural network tomography: Network replication from output surface geometry" *Neural Networks* 2011, Vol. 245, pp. 484-492
- [16]. S. Baroutian, and M. K. Aroua, "Estimation of Vegetable Oil-Based Ethyl Esters Biodiesel Densities using Artificial Neural Networks" *J Appl Sci*, 2008, Vol. 8 pp. 3005-3011
- [17]. A. S. Ramadhas, S. Jayaraj, C. Muraleedharan, and K. Padmakumari, Artificial neural networks used for the prediction of the cetane number of biodiesel. *Renewable Energy* 2006, Vol. 3115 pp. 2524-2533.
- [18]. J. Kumar, Bansal A, and M. K. Jha, "Comparison of Statistical and Neural Network Techniques in Predicting Physical Properties of Various Mixtures of Diesel and Biodiesel" *Proceedings of the World Congress on Engineering and Computer Science*. San Francisco, USA. October 24-26, 2000.