As. J. Energy Env. **2006**, 7(02), 299-306

Asian Journal on Energy and Environment

ISSN 1513-4121

Available online at www.asian-energy-journal.info

Research Paper

Predicting Properties of Biodiesels Using Statistical Models and Artificial Neural Networks

Kraipat Cheenkachorn*

Department of Chemical Engineering, King Mongkut's Institute of Technology North Bangkok, Bangkok, Thailand

*Author to whom correspondence should be addressed, email: kpc@kmitnb.ac.th

Abstract: Biodiesels (usually methyl esters) have continuously gained interests in many countries including Thailand. Vegetable oils are normally used as the main reactants for biodiesels via a transesterification process with short-chain alcohols. The fatty acid compositions in vegetable oils result in the properties of biodiesels such as viscosity, flash point, high-heating values, etc. In a present work, using the data mainly collected from previous literatures, statistical models and artificial neuron networks were used to predict the properties of biodiesels using the fatty acid compositions of various vegetable oils. The input variables of the models include different types of fatty acids such as palmitic acid, stearic acid acid, oleic acid, etc. The predicted properties include viscosity, high-heating value (HHV), and cetane number. The results show that both statistical models and the artificial neural networks precisely predict the properties of biodiesel derived from certain vegetable oils.

Keywords: Biodiesels, Artificial Neural Network, Linear Regression Model, Viscosity, Cetane Heat of Combustion.

Introduction

As energy demands extremely increase while the energy sources are limited, a number of current studies focus on a development of alternative fuels. Biodiesel is one of the potential alternative fuels and has been widely used in many countries. Depending on the climate and the availability, different vegetable oils are currently used as raw materials for a biodiesel production. In Thailand, a palm oil is mainly used in the transesterification process to produce the biodiesel because of its highest production rate. In addition to a direct use as an alternative fuel, biodiesels also offer other applications such as friction modifiers and wear additives.

Biodiesels offer a number of advantages over a regular diesel. For example, a previous study from U.S. DOE [1] showed that pure biodiesel and a blend of a conventional diesel with 20% (B20) can reduce a fuel economy and power approximately 10% and 2%, respectively. This study also showed that pure biodiesel can decrease the cancer risks by 94% while (B20) can reduce as much as 27%. The particulates are reduced as much as 55.4% using pure biodiesel while B20 offers 18% reduction. Although biodiesels offer several outstanding advantages over a regular diesel, they are not widely used because of their higher price compared to a regular diesel.With tax exemption and government subsidy, biodiesels can gain a wide acceptance in many countries such as Germany and other European countries.

Although the price of biodiesel blends can not complete with a conventional diesel, the strict regulations on the aromatics and sulfur content, which may require the capital intensive equipment and high cost operations, will result in higher cost of diesel fuel [2]. A comparative cost analysis and economic feasibility of biodiesel compared with compressed natural gas (CNG), and methanol was studied by Ahouissouss et al [3].They provided a complete cost analysis for operating a transit bus fleet including infrastructure, engine replacement, and maintenance cost. The Rust algorithm model showed biodiesel has a higher total cost than a regular diesel, but they have potential of completing with CNG and methanol. Raneses et al showed that the potential biodiesel demand in three specific markets including federal fleets, mining, marine/estuary areas has identified as likely candidates for commercialization [4].

It is generally known that the contents of different fatty acids in vegetable oils have strong effect on final properties of biodiesels. However, no correlation was made to predict the final properties of biodiesel based on their compositions. The present study focuses on the correlation between percent of fatty acids and final properties of biodiesel fuels.

Correlation of Biodiesel Properties

Biodiesel Kinetics

Biodiesels are usually produced from a transesterification process of vegetable oils with short chain alcohols (usually methanol). The transesterification reactions consist of three stepwise reactions. First, triglyceride reacts with methanol to form diglyceride and a methyl ester. Diglyceride then reacts with another methanol to form monoglyceride and another methyl ester. Monoglyceride finally reacts with methanol to form glycerol and methyl ester. An alkali or acid is commonly used as a catalyst. The overall reaction is assumed to proceed in the first-order reaction with respect to the concentration of triglyceride [5]. The reaction proceeds approximately 90-97% in an access of methanol at room temperature. Some of free fatty acid can be converted to soap and water.

Data Used

Data used were collected from previous literatures [6-13]. The collected data include some physical properties of biodiesels derived from each oil such as viscosity, high- heating value, and cetane number. It is assumed that a preparation of transesterification does not effect on these physical properties.

Artificial Neural Network

An artificial neural network (ANN) is information – processing paradigm that offers a diversified application and ideal solutions to a variety of classification problem as well as functional predictions. ANNs are composed of simple elements operating in parallel. The connections of complex system among these elements mostly determine the network functions. The ANNs are trained to perform a particular function by adjusting the values of the connections, or weights, between elements until a particular input leads to a specific output.

The ANN consists of three layers: input layer, hidden layer, and output layer. These three layer are connected each other. The input layer receives the input data outside the network and sends them to the hidden layer. The hidden layer contains interconnected neurons for the pattern recognition and the relevant information interpretation for adjusting the weights on the connections. Afterwards, the results from the hidden layer are sent to the output layer for the outputs. The object oriented programming is required to create an ANN because it contains different variables and methods. The neurons contain several functions and variable including weights, non-linear transfer functions, methods to add up all inputs and bias values. The sum of all products of all the inputs multiplying the weights and the bias values passes through a non-linear transfer function as the output of each neuron. A Basic concept of ANN is shown in Fig 1. The details of ANN architecture can be found elsewhere [14].

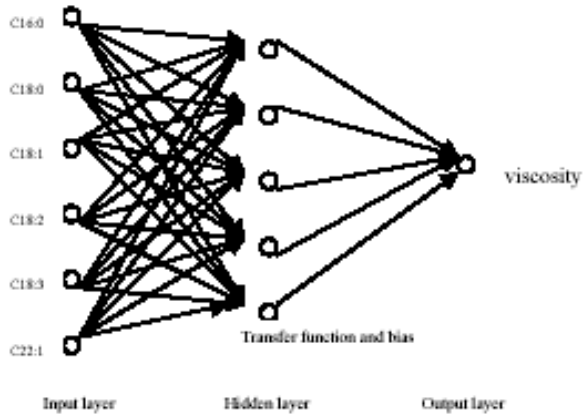

Fig. 1 Basic Concept of ANN.

Statiscal and Ann Models

Statistical Method

The physical properties of biodiesels were correlated with their different compositions using a statistical model. The best subset regression was used to determine which parameters affect the physical properties of biodiesels such as viscosity and cetane number. The parameters include percents of fatty acids which are the main components in vegetable oils. Table 1 shows the best subset regression results for viscosity of biodiesels. To identify the best subset regression model, the C_p criterions used are (1) the C_p value is small which correspond to a subset model

with a small total mean square error and (2) the C_p value is near a number of parameters in that subset. The model with low Cp value has small bias of the regression model [15]. The model should also have a high value of adjusted R^2.

Viscosity

Viscosity values for methyl esters of different vegetable oils at 40 °C are correlated using the best subset method. As seen from Table 1, the model with adjusted R^2 of 62.7 and Cp of 3.3 is chosen from the subsets to represent a correlation of viscosity with their fatty acid components. The factors that affect the viscosity are $(C18:1)^2$, $(C18:3)^2$, 2(C18:1), 4(C18:2), and (C22:1). The model suggests that only unsaturated fatty acids such as oleic acid, linoleic acid, linolenic acid, and erucic acid affect the viscosity of biodiesels. This is caused by the presence of double bonds and the chain length of those acids. Generally, the viscosity increases as the chain length of the acid in the triglyceride increases and decreases as the number of double bonds increase. The linear regression model of the best subset method is shown in equation (1).

Table 1 Best Subset Method for Viscosity

1	[illegible]	[illegible]	[illegible]									■
1	[illegible]	[illegible]	[illegible]			■						
2	[illegible]	[illegible]	[illegible]		■							■
2	[illegible]	[illegible]	[illegible]	■								■
3	[illegible]	[illegible]	[illegible]					■	■			■
3	[illegible]	[illegible]	[illegible]		■				■			■
4	[illegible]	[illegible]	[illegible]				■	■	■			■
4	[illegible]	[illegible]	[illegible]					■	■		■	■
5	[illegible]	[illegible]	[illegible]				■	■	■	■		■
5	[illegible]	[illegible]	[illegible]				■	■	■		■	■
5	[illegible]	[illegible]	[illegible]		■		■	■	■	■		■
5	[illegible]	[illegible]	[illegible]			■	■	■	■		■	■
6	[illegible]	[illegible]	[illegible]		■		■	■	■	■	■	■
6	[illegible]	[illegible]	[illegible]		■	■	■	■	■		■	■
7	[illegible]	[illegible]	[illegible]		■	■	■	■	■	■	■	■
7	[illegible]	[illegible]	[illegible]	■	■		■	■	■	■	■	■
7	[illegible]	[illegible]	[illegible]	■	■	■	■	■	■	■	■	■

$$\text{Viscosity (cSt)} = 1.82 - 55.3\,(C18:3)^2 - 10.3\,(C18:1)^2 + 5.11\,(2\cdot C18:1) + 0.162\,(4\cdot C18:2) + 4.95\,(C22:1) \quad (1)$$

Statistical testing shows all coefficients except that of $(C18:1)^2$ and $(4\cdot C18:1)$ are significant according to their corresponding p-values as shown in Table 2. As seen from an analysis of variance of the viscosity, the model fairly predicts the viscosity of biodiesel. The non-linear behavior of biodiesel viscosity is observed and hard to obtain an accurate correlation to predict the viscosity. Therefore, an ANN is needed to precisely predict the viscosity of biodiesels from their compositions.

The ANN architecture is a multilayer feed forward network. A back propagation algorithm is used to train the data set. The input data include palmitic acid, stearic acid, oleic acid, linoleic

acid, linolenic acid, and erucic acid. The viscosity is simulated by a three-layer network with tan- sigmoid, logsigmoid, and purelin transfer functions in the first, second, and third layer, respectively. The sizes of three neurons are 10, 10, and 1, respectively. Figure 2 shows predicted

values of viscosity from the ANN compared with experimental data. The ANN model precisely predicts the viscosity of biodiesels. The mean square error of the network is less than 0.01.

Table 2 Viscosity Regression Results for Five-Parameter Model

Predictor	Coeff.	SE Coeff.	T	P
Constant	1.82	0.63	2.90	0.009
$(C18:3)^2$	-55.26	34.19	-1.62	0.122
$(C18:1)^2$	-10.26	2.71	-3.77	0.001
(2·C18:1)	5.11	1.29	3.97	0.001
(4·C18:2)	0.16	0.15	1.05	0.307
(C22:1)	4.95	0.88	5.63	0.000

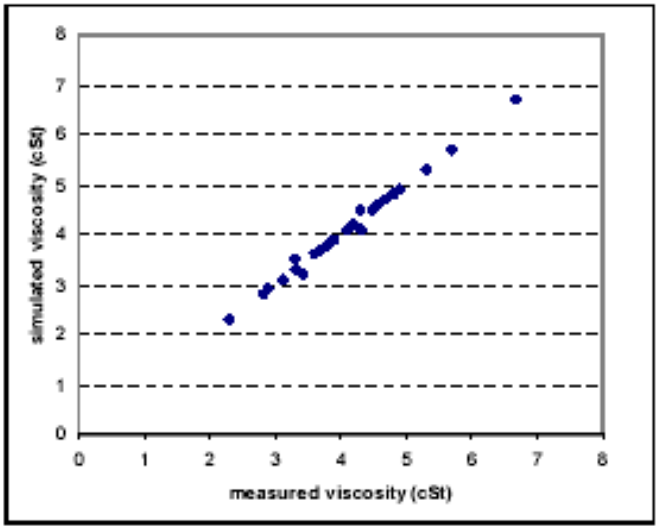

Fig. 2 ANN results for viscosity of biodiesels.

Cetane Number

Cetane number is a measure of an ignition delay of the fuels. Fuels with higher cetane number process a short ignition delay which is desirable for a diesel engine. Table 2 shows the best subsets for cetane number of biodiesels. As seen from the best subset table, the model with four variables is chosen to predict the cetane number of biodiesels. The variables that affect cetane number are stearic acid, oleic acid, linoleic acid, and erucic acid. The linear regression model for cetane number is shown in equation (2) with adjusted R^2 of 81.9% which is fairly reasonable. Statistical testing also shows all coefficients except that of C18:2 are significant according to their corresponding p-values presented in the Table 3 The results agree with previous study that cetane number increases as the chain length increases and decreases as a number of double bonds decreases [10].

$$CN = 33.6 + 0.539\ (C18:0) + 0.303(C18:1) + 0.0878\ (C18:2) + 0.233(C22:1) \quad (2)$$

Table 3 Best Subset Method for Cetane Number

Vars	R-Sq	R-Sq(adj)	C-p	C16:0	C18:0	C18:1	C18:2	C18:3	C22:1
1	51	45.5	18.6				X		
1	35.6	28.4	26.7			X			
2	63	53.7	14.3			X			X
2	61.3	51.6	15.2	X			X		
3	85	78.6	4.8		X	X			X
3	85	61.3	11.2	X			X	X	
4	82.9	81.9	4.7		X	X	X		X
4	87	78.3	5.8	X	X	X			X
5	92.3	84.6	5	X	X	X	X		X
5	90	80	6.2		X	X	X	X	X
6	92.3	80.9	7	X	X	X	X	X	X

Heat of Combustion

Biodiesels generally have lower heat of combustion than regular diesel because of the presence of oxygen in those oils. However, the volumetric fuel economy is lower on biodiesels and biodiesel blends. Since the correlation between high- heating values (HHV) and the fatty acid compositions of biodiesels is non-linear, it is impossible to obtain the simple statistical model for HHV. The ANN with Levenberg- Marquardt training method is used to predict the heating value based on the contents of palmitic acid, stearic acid, oleic acid, linoleic acid, linolenic acid, and erucic acid. The HHV is simulated by a two-layer network with tan-sigmoid and purelin transfer functions in the first and second layers, respectively. The sizes of three neurons are 10 and 1, respectively. The ANN model accurately predicts heating values with the mean square error of 0.0047. The predicted results are very close to the experimental values as shown in Figure 3. The results show that HHV increases as the chain length increases and decrease as the double bonds decrease. This observation agrees with the previous study [10].

Table 4 Viscosity Regression Results for Five-Parameter Model

Predictor	Coeff.	SE Coeff.	T	P
Constant	33.56	5.16	6.51	0.001
(C18:0)	0.54	0.15	3.58	0.012
(C18:1)	0.32	0.07	4.43	0.005
(C18:2)	0.09	0.06	1.51	0.182
(C22:1)	0.23	0.06	3.78	0.009

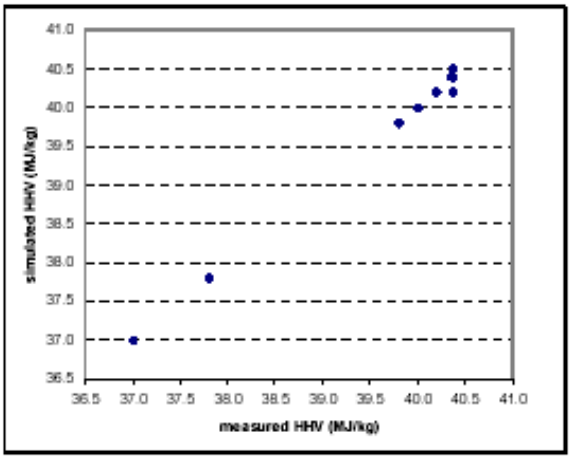

Fig. 3 ANN results for HHV of biodiesels.

Conclusion

The statistical models using best subset method fairly predicts viscosity and cetane number of biodiesels from their fatty acid compositions. A deviation from linear regression model is caused by the presence of double bonds in the fuels. The ANN models accurately predict those properties. The predicted values are very close to the experimental results.

Symbols

C16:0 percent of palmitic acid content

C18:0 percent of stearic acid content

C18:1 percent of oleic acid content

C18:2 percent of linoleic acid content

C18:3 percent of linolenic acid content

C 22:1 percent of erucic acid content

CN cetane number

HHV high heating value, MJ/kg

ANN artificial neural network

References

[1] www.biodiesel.org

[2] Srivastara, A. and Prasad, R.(2000), riglycerides-based diesel fuels, *Renewable and Sustainable energy Reviews*, pp. 4 111-133.

[3] Ahouissouss, N.B.C. and Wetstein, M.E. (1997), A comparative cost analysis of biodiesel, compressed natural gas, methanol, and diesel for transit bus systems, *Resource and Energy Economics*, **20** pp. 1-15.

[4] Raneses, A.R., Glaser, L.K., Price, J.M., and Duffield, J.A. (1999), Potential biodiesel markets and their economic effects on the agricultural sector of the United States, *Industrial Crops and Products*, **9** pp. 151-162.

[5] Kusdiana, D. and Saka, S. (2001), Kinetics of transesterification in rapeseed oil to biodiesel fuel as treated in super critical methanol, *Fuel* **80**, pp. 693-698.

[6] Zhang, X., Peterson, C.L., Reece, D., Moller, G. and Haw, R. (1995), Biodegradability of biodiesel in the aquatic environment, *ASAE paper 95-6742*, American Society of Agricultural Engineers, St. Joseph, MI.

[7] Demirbas, A. (2003), Biodiesel from vegetable oils via transesterification in supercritical methanol, *Energy Conversion & Management*, **43**, pp. 2349-2356.

[8] Da Silva, F.N., Prata, A.S. and Teixeira, J.R. (2003), Technical feasibility assessment of oleic sunflower methyl ester utilization in diesel bus engine, *Energy Conversion & Management*, **44**, pp. 2347-2365.

[9] Saha, B. and Woodward, J (1997), ACS Symposium series 666, American Chemical Society, Washington, D.C.

[10] Graboski, M.S. and McCormick, R.L. (1998), Combustion of fat and vegetable oil derived fuels in diesel engines, Progess in Energy and Combustion, *Science*, **24**, pp. 125-164.

[11] Allen, C.A.W., Watts, K.C., Ackman, R.G. and Pegg, M.G. (1999), Predicting the viscosity of biodiesel fuels from their fatty acid ester composition, *Fuel*, pp. 781319-1326.

[12] Foidl, N., Sanchez, M., Mittelbach, M., Heckel, S. And Jatropha Curcas, L. (1996), As a source for the production of biodiesel in Nicaragua, *Bioresource Technology*, **58**, pp. 78-82.

[13] Demirbas, A. (2003), Biodiesel fuels from vegetable oils via catalytic and non-catalytic supercritical alcohol transesterifications and other methods:a survey, *Energy Conversion & Management*, **44**, pp. 2093-2109.

[14] www.mathwork.com/access/helpdesk/help/toolbox/nnet.s html.

[15] Neter, J., Kutner, M. Nachsheim, C.J. and Wasserman, W. 1996, *Applied Linear Statistical Models*, 4th Edition, McGrawHill