Energy Conversion and Management 65 (2013) 255–261

Contents lists available at SciVerse ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

Prediction of the cetane number of biodiesel using artificial neural networks and multiple linear regression

Ramón Piloto-Rodríguez [a,*], Yisel Sánchez-Borroto [a], Magin Lapuerta [b], Leonardo Goyos-Pérez [a], Sebastian Verhelst [c]

[a] Faculty of Mechanical Engineering, Technical University of Havana, Calle 114, No. 11901 e/119 y 127, Cujae, Marianao 15, 19390 Ciudad de la Habana, Cuba
[b] Escuela Técnica Superior de Ingenieros Industriales, University of Castilla-La Mancha, Avda. Camilo José Cela s/n, 13071 Ciudad Real, Spain
[c] Department of Flow, Heat and Combustion Mechanics, Faculty of Engineering, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 23 May 2012
Received in revised form 18 July 2012
Accepted 26 July 2012
Available online 17 October 2012

Keywords:
Cetane number
Biodiesel
Neural network
Fatty acid
Ester composition

ABSTRACT

Models for estimation of cetane number of biodiesel from their fatty acid methyl ester composition using multiple linear regression and artificial neural networks were obtained in this work. For the obtaining of models to predict the cetane number, an experimental data from literature reports that covers 48 and 15 biodiesels in the modeling-training step and validation step respectively were taken. Twenty-four neural networks using two topologies and different algorithms for the second training step were evaluated. The model obtained using multiple regression was compared with two other models from literature and it was able to predict cetane number with 89% of accuracy, observing one outlier. A model to predict cetane number using artificial neural network was obtained with better accuracy than 92% except one outlier. The best neural network to predict the cetane number was a backpropagation network (11:5:1) using the Levenberg–Marquardt algorithm for the second step of the networks training and showing $R = 0.9544$ for the validation data.

1. Introduction

Several physical properties of biodiesel fuels depend on their fatty acid ester composition [1–3]. Also related to the ester composition is the cetane number which is one of the most cited indicators of diesel fuel quality [3–6]. The cetane number measures the readiness of the fuel to autoignite when it is injected into the combustion chamber. It is generally dependent on the composition of the fuel and can influence the engine stability, noise level, and exhaust emissions.

The cetane number (CN), determined by a standard (diesel engine) test ASTM D613, is a measure of the ignition quality of a diesel fuel in a compression ignition engine. A fuel with higher cetane number has a shorter ignition delay period and starts the combustion shortly after it is injected into the chamber [4]. While the ignition delay can be influenced by engine type and operation conditions, the cetane number mainly depends on the chemical composition of the fuel.

The cetane number of biodiesel is generally higher than the standard diesel fuel. Experimental data shows values varying between 45 and 67 for biodiesel and ranged between 40 and 49 for diesel fuel [7,8]. A single fatty acid alkyl ester molecule can have a cetane number between 42 and 89, depending on its molecular structure [7].

Van Gerpen [7] studied the effect of adding pure esters to diesel fuel. A linear regression fit on the CN data for each ester as a function of the percent of ester in the blend was used. The obtained values of the coefficient of correlation were ranged between 0.4889 and 0.9965 depending on the fatty acid added to the blend.

Equations for predicting the cetane number of diesel or biodiesel fuels have been published [4,9–15], correlating this parameter with different input experimental factors or using different mathematical methods. Yang et al. [9] developed multiple linear correlation equations for predicting the CN for 12 hydrocarbons in order to compare with a model developed using artificial neural networks (ANNs).

A model for the estimation of the cetane number of biodiesel fuels based on a literature review was proposed by Lapuerta et al. [13,14]. The model was built up from experimental data obtained using different methods, initially divided in those from a diesel engine called Cooperative Fuel Research engine (CFR) and those from an Ignition Quality Tester (IQT) device, and finally brought together. A quadratic correlation with the number of carbon atoms in the original fatty acid and the number of double bonds was statistically selected as the most suitable. The R^2 obtained were ranged between 0.918 and 0.947.

Bamgboye et al. [15] applied multiple linear regression for obtaining a model for predicting cetane number were the R^2

* Corresponding author. Tel.: +53 7 2663624.
E-mail address: rpiloto@ceter.cujae.edu.cu (R. Piloto-Rodríguez).

0196-8904/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.enconman.2012.07.023

Nomenclature

CN	cetane number	*M*	percent of myristic
ASTM	American standard	*P*	percent of palmitic
ANNs	artificial neural networks	*Pt*	percent of palmitoleic
CFR	Cooperative Fuel Research engine	*S*	percent of stearic
IQT	Ignition Quality Tester	*O*	percent of oleic
R^2	coefficient of determination	*Li*	percent of linoleic
FAMEs	fatty acid methyl esters	*Ln*	percent of linolenic
MLR	multiple linear regression	*Ei*	percent of eicosanoic
BP	backpropagation	*Er*	percent of erucic
CGD	conjugate gradient descend	*Ot*	sum of residual FAMEs to reach 100%
QP	quick propagation	*wt*	weight percent
BD	biodiesel	*R*	coefficient of correlation
La	percent of lauric		

obtained was 0.883 and he validated the model using data from literature. Ramos et al. [16] reports the use of a previously published equation to predict cetane number from the cetane number of the individual fatty acid methyl esters (FAMEs). Its use cannot avoid the engine tests or collecting cetane numbers of pure FAMEs from literature reports.

Most of the models published for cetane number prediction were developed with Multiple Linear Regression (MLR) techniques. That procedure requires the user to specify a priori a mathematical model to fit the data in order to obtain the empirical correlation. An alternative to avoid that problem is the use of artificial neural networks.

Unlike the correlation techniques, the neural network can identify and learn the correlative patterns between the input and output data once a training set is provided. The use of ANNs for predicting and modeling of energetic and mechanical systems is reported [17–24]. Their use in the modeling of engines combustion processes is also reported [25–28]. Very few reported the use of ANNs for obtaining models to predict the cetane number of diesel fuels [9,29,30], and only one for its prediction in biodiesel fuels [12].

Yang et al. [9] used a backpropagation neural network model with a training step and a validation step. The results shown a higher coefficient of determination ($R^2 = 0.97$) than using MLR. Basu et al. [29] obtained relationships between the CN of diesel fuels using nuclear magnetic resonance. The cetane number was determined using an IQT. Ramadhas et al. [12] used an ANN to predict cetane number selecting four types of networks. Santana et al. [30] estimated the CN of individual components of diesel fuel using ANNs. The neural networks have also been applied to the prediction of other fuel properties [31].

Determination of the CN by an experimental procedure at present is an expensive and time consuming process. Therefore, the obtaining of accurate models to predict the CN of a biodiesel from its FAME composition in a wide range of feedstocks characteristics would be useful for the scientific community.

The purpose of this work is to obtain models for the estimation of the cetane number of biodiesel from their FAME composition using MLR and ANNs searching for the best suitable model to predict cetane number in the range of biofuels studied, covering biodiesels from 63 feedstocks.

2. Experimental set-up and procedures

In the present work 48 different biodiesel fuels (including 10 pure fatty acids) were taken from references as input and output data for the obtaining of a MLR and for the implementation of ANNs for predicting the cetane number. The FAME main composition presented in biodiesel obtained from different feedstocks is covered by ten FAMEs selected [10,12,15,16,32–36]. The input data covers FAME composition and the output covers the cetane number. The validation of the models obtained was done using a separate data set selected from literature reports, which was not used for developing the models. The data selected for validation covers 15 samples.

The degree of relationship between measured and fitted cetane number data was expressed as the *R* and R^2. The best fit was expressed as the higher *R* and the lower mean absolute error. The obtained model using MLR was compared with two models available in literature [10,15]. Due to the type of data inputs collected for this work, the comparison with correlations as those proposed by Lapuerta et al. [14] and Tong [33] was discarded.

With the aim of comparing the models obtained using MLR and ANNs, different networks were developed using two basic topologies (11:5:1) and (11:7:1). As an example, one of the topologies used in this work is shown in Fig. 1. The ANNs used were the multilayer Perceptrons, with one hidden layer and five or seven units. The inputs of the network were ten, representing the chemical composition of 10 FAME and one input representing the total amount of the other FAMEs found in the biodiesel sample. The CN was the unique variable output of the network.

The chemical formula and the structure of the FAMEs on which this research is focused are shown in Table 1. The ten FAMEs listed represent the inputs for the CN modeling. The basic structural

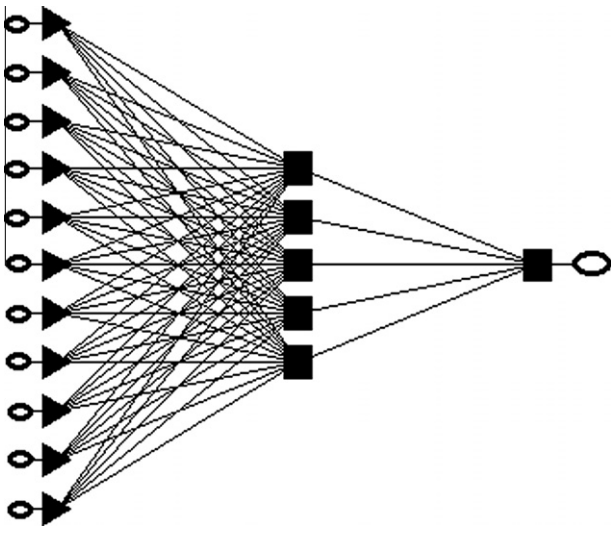

Fig. 1. Network (11:5:1) for the prediction of the cetane number.