J Am Oil Chem Soc (2011) 88:415–423
DOI 10.1007/s11746-010-1672-0

ORIGINAL PAPER

Cetane Number Prediction of Biodiesel from the Composition of the Fatty Acid Methyl Esters

Dongmei Tong · Changwei Hu · Kanghua Jiang · Yuesong Li

Received: 10 April 2010 / Revised: 25 July 2010 / Accepted: 23 August 2010 / Published online: 11 September 2010

Abstract In the present work, the measured cetane numbers (CN) of pure fatty acid methyl esters (FAME), as well as the FAME compositions and the reported CN of 59 kinds of biodiesels collected from literature were used to develop a simple model involving as more FAME component as possible for predicting CN of biodiesel from its FAME composition. Two different regression equations correlating the CN of pure FAME with the carbon number of fatty acid chain were obtained by regression analysis, which shows that the dependence of the CN on the carbon number varies with the unsaturated degree of fatty acid chain. The 59 biodiesels were divided into two categories and used, respectively to develop and test a multiple linear regression model (MLRM) correlating the CN of biodiesel with its FAME composition. A simple and convenient regression equation with a high accuracy and a good reproducibility (average absolute error of 0.49 CN for testing set and 1.52 CN for all data) were developed, showing excellent correlation (R^2: 0.9904 for testing set). The model developed in the present work can be used conveniently to give a satisfactory predicted CN of biodiesel from the FAME composition.

Keywords Biodiesel · Cetane number · Prediction · Carbon number · Composition of fatty acid methyl esters · Regression

D. Tong · C. Hu (✉) · K. Jiang · Y. Li
Key Laboratory of Green Chemistry and Technology, Ministry of Education, College of Chemistry, Sichuan University, Chengdu 610064, Sichuan, China
e-mail: chwehu@mail.sc.cninfo.net; gchem@scu.edu.cn

Introduction

Biodiesel, a mixture of long-chain fatty acids alkyl esters, is produced easily by either the transesterification of triacylglycerol from materials such as animal fats or vegetable oils with alcohols, or the esterification of free fatty acids (FFA) with alcohols. As a renewable and biodegradable fuel with low harmful emission, biodiesel has similar combustion properties to petroleum-based diesel and can be used in its pure form or in the form of blends with diesel in diesel engine [1, 2]. Utilizing biodiesel as an alternative fuel can reduce our dependency on petroleum-based energy resources and alleviate the air pollution owing to the wide use of fossil fuels.

Exhaust emission, one of the important focuses in the environmental protection, is determined by the combustion behavior of diesel in the combustion chamber of a diesel engine, for which the cetane number (CN) is one of the most significant indicators. The CN of diesel, specified by ASTM D613, is a measure of its ignition delay time. A higher CN, a desirable property in diesel engine, indicates shorter time between the ignition and the initiation of fuel injection into the combustion chamber. The higher CN is correlated with the reduction of nitrogen oxides (NO_x) and unburnt hydrocarbons (UHC) exhaust emissions [3], which is important for alleviating air pollution.

The CN of the test fuel is the percentage in volume of *n*-hexadecane (trivial name cetane) in a mixture of *n*-hexadecane (CN = 100) and heptamethylnonane (CN = 15), which matches the ignition quality of the test fuel in the same engine under standard conditions specified by ASTM D613. The CN test procedure is complex and involves running the fuel in a single-cylinder compression-ignition engine with a continuously variable compression ratio under a fixed set of conditions. Although the ASTM D613

method is the widely accepted test method for CN, it exhibits several inherent disadvantages, including a considerable amount of fuel sample requirement (~1 L), a time-consuming process, a relatively high reproducibility error, and a relatively high cost [4]. Therefore, there have been many attempts to develop theoretical models to predict the CN quickly and reliably from bulk properties of diesel such as density, viscosity, aniline point, distillation temperatures, chemical composition, saponification number, iodine value, and so on [5–8]. Ladommatos et al. [7] tested the accuracy of 22 equations for predicting the CN of diesel from the above mentioned properties, by comparing the predicted values of 563 fuels with the measured values. It is found that several equations are able to estimate the CN with good accuracy for diesel. But they also found that the predictive capability of these empirical equations was poor for untypical diesel fuels such as those containing vegetable oils and alcohols [7]. This could be attributed to the difference of the chemical components between petroleum-based diesel and biodiesel. Diesel is a mixture of a large variety of hydrocarbon compounds including paraffins, olefins, cycloparaffins and aromatics, while biodiesel is a mixture of fatty acids methyl esters (FAME) for methanol being widely used in biodiesel preparation.

Therefore, developing correlative models for predicting the CN of biodiesel fuel is necessary for studying its ignition quality. Researchers investigated the relationship between the CN of pure FAME and their properties, such as boiling point, viscosity, density and carbon number of the fatty acid moiety [9–11]. It is well known that the property is determined by the structure. The carbon number and the unsaturated property of fatty acid chain are two significant factors affecting the property of the CN. Normally, the CN of pure FAME increases with increasing chain length and the saturated degree of the fatty acid chain. Freedman et al. [9] obtained a quadratic equation correlating the CN of saturated FAME with the carbon number as follows:

$$Y = -57.26 + 14.892X - 0.4149X^2 \quad (1)$$

where Y is the CN, and X is the carbon number of fatty acid chain. Lapuerta et al. [11] also developed a correlation between the CN of pure FAME and the carbon number as well as the number of double bond as the following equation:

$$Y = -52.974 + \left(13.767 - 1.202db + 0.152db^2\right)X - 0.351X^2 \quad (2)$$

where Y is the CN, db is the number of double bonds and X is the carbon number. These equations reveal the effect of the carbon number and the unsaturated property on the CN of pure FAME. However, they are not suitable for estimating the CN of biodiesel, a mixture of FAME in essence.

The nature of biodiesel components inherently determines its combustion properties. The experimental results of many researchers indicated that the CN of biodiesel varied with the raw material in which the FAME composition varied [12–16]. The different distribution of FAME could be obtained even for the biodiesel derived from the same parent oil source, under different climatic conditions for the oil plants' growth and different oil processing methods. For example, a high CN was observed in biodiesels derived from palm oil and tallow, rich in saturated esters such as methyl palmitate (C16:0) and methyl stearate (C18:0) [17, 18]. While low CN was associated with more abundant unsaturated components such as methyl oleate (C18:1), linoleate (C18:2) and linolenate (C18:3) for some other biodiesels derived from soybean oil, sunflower oil and grape seed oil [18, 19]. The CN of biodiesel increased with an increasing weight percentage of saturated FAME and longer chain FAME [12–16]. It is obvious that the FAME composition plays an important role affecting the CN of biodiesel.

The level of oxidation of biodiesel through contact with the oxygen in air was also suggested as an important influencing factor resulting in higher CN [20, 21]. Hydroperoxides form during the oxidation of unsaturated FAME, which may split and form other secondary products such as aldehydes and carboxylic acids. These oxidative products contribute to the increase in the CN [20, 21]. Wadumesthrige et al. [21] measured comparatively the CN of undistilled biodiesel against distilled biodiesel. They found that the CN of undistilled biodiesel after 4 months storage under ambient condition was not significantly different from the value before storage, but the CN of distilled biodiesel increased significantly. This was ascribed to the fact that natural antioxidants such as tocopherol were removed from biodiesel by distillation, and so the distilled biodiesel was very susceptible to oxidation resulting in a higher CN. While the presence of natural antioxidants in undistilled biodiesel led to relatively stable CN. Based on these results, for undistilled biodiesel newly prepared or stored within a short time, the FAME composition will be the primary influence factor on the CN of biodiesel.

Consequently, developing a correlative model with a simple empirical equation for predicting the CN of biodiesel from the FAME composition, is important and necessary for studying the ignition quality of biodiesel instead of the complex ASTM D613 measurement. The detailed and quantitative correlation was investigated in the last few years [18, 22, 23]. An artificial neural network (ANN) model was developed by Ramadhas et al. [23] to predict the CN of some biodiesels, in which five kinds of FAME, methyl palmitate (C16:0), stearate (C18:0), oleate (C18:1), linoleate (C18:2) and linolenate (C18:3) were

involved. Bamgboye et al. [18] suggested a regression equation described as follows:

$$\mathrm{CN} = 61.1 + 0.088x_2 + 0.133x_3 + 0.152x_4 - 0.101x_5 - 0.039x_6 - 0.243x_7 - 0.395x_8 \quad (3)$$

where x_2 to x_8 are the weigh percentages of methyl myristate (C14:0), palmitate (C16:0), palmitoleate (C16:1), stearate (C18:0), oleate (C18:1), linoleate (C18:2) and linolenate (C18:3) in biodiesels, respectively. Gopinath et al. [22] also developed a multiple linear regression model as the following equation:

$$\mathrm{CN} = 62.2 + 0.017L + 0.074M + 0.115P + 0.177S - 0.103O - 0.279\mathrm{LI} - 0.366\mathrm{LL} \quad (4)$$

where L, M, P, S, O, LI and LL are the weight percentages of methyl laurate (C12:0), myristate (C14:0), palmitate (C16:0), stearate (C18:0), oleate (C18:1), linoleate (C18:2) and linolenate (C18:3) in biodiesels, respectively. In their work, they took into account the effect of 5–7 kinds of FAME with carbon number of fatty acid chain from 12 to 18. These FAME are the dominant chemical components in biodiesel, but not all of them. Some other FAME like methyl caprate (C10:0), arachidate (C20:0) and erucate (C22:1) also widely exist in biodiesel [19, 24, 25]. Since biodiesel is a mixture of FAME, each component may contribute to the CN. Disregarding the effect of other FAME results consequentially in an error of the predicted value from the actual value.

Therefore, there is a need to develop a model involving as many FAME components as possible for predicting the CN of biodiesel from the FAME composition. A simple empirical equation, the accuracy of which was verified by comparing with the models in literature and its application in predicting the CN of biodiesels derived from several conventional parent oil sources, was developed in the present work.

Materials and Methods

Data Collection for Pure FAME and Biodiesels

The measured CN of pure FAME with carbon numbers of fatty acid chains from 6 to 22 shown in Table 1 were collected from the literature [9, 12, 13, 26–29]. The FAME compositions and the reported CN of 59 kinds of biodiesel from different parent oil sources were also obtained from the literature, as shown in Table 2 [15].

Procedures

The correlative model was developed by two steps. Firstly, the relationship between the CN of pure FAME and the carbon number of fatty acid chain for saturated esters, as well as the relationship for unsaturated esters containing one carbon–carbon double bond, were established using the

Table 1 CN of pure FAME

Fatty acid methyl ester	Reported CN	Reference	The present work		Literature work[b]	
			Calc. CN	Relative err. (%)	Calc. CN	Relative err. (%)
Caproate (C6:0)	18.0	[9]	16.3	9.4	17.0	5.6
Caprylate (C8:0)	33.6	[26]	36.2	7.6	34.7	3.3
Caprate (C10:0)	47.2	[26]	49.3	4.3	49.6	5.1
Laurate (C12:0)	60.8	[9]	58.0	4.6	61.7	1.5
Myristate (C14:0)	66.2	[26]	64.7	2.2	71.0	7.3
Palmitate (C16:0)	74.3	[9]	71.9	3.2	77.4	4.2
Palmitoleate (C16:1)	51.0	[29]	51.0	–	60.6	18.8
Stearate (C18:0)	75.6	[9]	82.0	8.4	81.1	7.3
Oleate (C18:1)	56.5	[13]	56.6	–	62.2	10.1
Linoleate (C18:2)	38.2	[13]	–	–	–	–
Linolenate (C18:3)	22.7	[12, 27]	–	–	–	–
Arachidate (20:0)	100	[11]	97.2	2.8	82.0	18.0
Eicosenoate (C20:1)	(64.8)[a]		–	–	–	–
Erucate (C22:1)	76.0	[28]	76.0	–	56.9	25.1
Average relative error for saturated FAME (%[c])				5.3		6.5

[a] Calculated CN according to Eq. 7 in the present work

[b] Calculated CN according to Eq. 2 obtained from Ref. [11]

[c] Relative error = |reported CN-predicted CN|/(reported CN) × 100%

Table 2 FAME composition and reported CN of 59 biodiesel fuels [15]

Item	Sources	FAME composition/wt%												Reported CN
		10:0	12:0	14:0	16:0	16:1	18:0	18:1	18:2	18:3	20:0	20:1	22:1	
1	*Annona reticulate* Linn	–	–	1.0	17.2	42.0	7.5	48.4	21.7	–	–	–	–	53.47
2	*Thevetia peruviana* Merrill	–	–	–	15.6	–	10.5	60.9	5.2	7.4	0.3	–	–	57.48
3	*Canarium commune* Linn	–	–	–	29.0	–	9.7	38.3	21.8	1.2	–	–	–	55.58
4	*Celastrus paniculatus* Linn	–	–	–	25.1	–	6.7	46.1	15.4	3.0	–	–	–	51.9
5	*Vernonia cinerea* Less	–	–	8.0	23.0	–	8.0	32.0	22.0	–	3.0	–	–	57.51
6	*Putranjiva roxburghii*	–	–	–	8.0	–	15.0	56.0	18.0	–	3.0	–	–	54.99
7	*Calophyllum apetalum* Wild	–	–	–	8.0	–	14.0	48.0	30.0	–	–	–	–	51.57
8	*Calophyllum inophyllum* Linn	–	–	–	17.9	2.5	18.5	42.7	13.7	2.1	–	–	–	57.3
9	*Mesua ferrea* Linn	–	–	0.9	10.8	–	12.4	60.0	15.0	–	0.9	–	–	55.1
10	*Azadirachta indica*	–	–	–	14.9	–	14.4	61.9	7.5	–	1.3	–	–	57.83
11	*Moringa oleifera* Lam	–	–	–	9.1	2.1	2.7	79.4	0.7	0.2	–	–	–	56.66
12	*Pongamia pinnata* Pierre	–	–	–	10.6	–	6.8	49.4	19.0	–	4.1	2.4	–	55.84
13	*Sapindus trifoliatus* Linn	–	–	–	5.4	–	8.5	55.1	8.2	–	20.7	–	–	59.77
14	*Mimusops hexendra* Robx	–	–	–	19.0	–	14.0	63.0	3.0	–	1.0	–	–	59.32
15	*Pterygota alata* Rbr	–	–	–	14.5	–	8.5	44.0	32.4	–	–	–	–	51.09
16	*Holoptelia integrifolia*	–	–	3.5	35.1	1.9	4.5	53.3	–	–	1.1	–	–	61.22
17	*Vallaris solanacea* Kuntze	–	–	–	7.2		14.4	35.3	40.4	–	1.8	–	–	50.26
18	*Balanites roxburghii* Planch	–	–	–	17.0	4.3	7.8	32.4	31.3	7.2	–	–	–	50.46
19	*Mappia foetida* Milers	–	–	–	7.1	–	17.7	38.4	36.8	–	–	–	–	50.7
20	*Aphanamixis polystachya* Park	–	–	–	23.1	–	12.8	21.5	29.0	13.6	–	–	–	48.52
21	*Meyna laxiflora Robyns*	–	–	–	18.8	–	9.0	32.5	39.5	–	–	–	–	50.42
22	*Terminalia chebula* Retz	–	–	–	19.7	–	2.4	37.3	39.8	–	0.6	–	–	49.6
23	*Aleurites moluccana* Wild	–	–	–	5.5	–	6.7	10.5	48.5	28.5	–	–	–	34.18
24	*Euphorbia helioscopia* Linn	–	2.8	5.5	9.9	–	1.1	15.8	22.1	42.7	–	–	–	34.25
25	*Garcinia echinocarpa* Thw	–	–	–	3.7	–	43.7	52.6	–	–	–	–	–	63.1
26	*Garcinia morella* Desr	–	–	–	0.7	–	46.4	49.5	0.9	–	2.5	–	–	63.52
27	*Saturega hortensis* Linn	–	–	–	0.4	–	0.4	12.0	18.0	62.0	–	–	–	25.46
28	*Actinodaphne angustifolia*	4.3	87.9	1.9	0.5	–	5.4	–	–	–	–	–	–	63.2
29	*Litsea glutinosa* Robins	–	96.3	–		–	–	2.3	–	–	–	–	–	64.79
30	*Neolitsea cassia* Linn	0.3	85.9	3.8		–	–	4.0	3.3	–	–	–	–	64.05
31	*Neolitsea umbrosa* Gamble	1.7	59.1	11.5		–	–	21.0	6.7	–	–	–	–	60.77
32	*Swietenia mahagoni* Jacq	–	–	–	9.5	–	18.4	56.0	–	16.1	–	–	–	52.26
33	*Anamirta cocculus* Wight & Hrn	–	–	–	6.1	–	47.5	46.4	–	–	–	–	–	64.26
34	*Broussonetia papyrifera Vent*	–	–	–	4.0	–	6.1	14.8	71.0	1.0	3.0	–	–	41.25
35	*Argemone mexicana*	–	–	0.8	14.5	–	3.8	18.5	61.4	–	1.0	–	–	44.45
36	*Salvadora oleoides Decne*	0.8	35.6	50.7	4.5	–	–	8.3	0.1	–	–	–	–	66.13
37	*Salvadora persica Linn*	0.1	19.6	54.5	19.6	–	–	5.4	–	–	–	–	–	67.47
38	*Nephelium lappaceum Linn*	–	–	–	0.2	–	13.8	45.3	–	–	34.7	4.2	–	64.86
39	*Madhuca butyracea Mac*	–	–	–	66.0	–	3.5	27.5	3.0	–	–	–	–	65.27
40	*Madhuca indica JF Gmel*	–	–	0.1	17.8	–	14.0	46.3	17.9	–	3.0	–	–	56.61
41	*Rhus succedanea* Linn	–	–	–	25.4	–	–	46.8	27.8	–	–	–	–	52.22
42	*Ervatamia coronaria* Stapf	–	–	–	24.4	0.2	7.2	50.5	15.8	0.6	0.7	0.2	–	56.33
43	*Basella rubra* Linn	–	–	0.4	19.7	0.4	6.5	50.3	21.6	1.1	–	–	–	54
44	*Corylus avellana*	–	–	3.2	3.1	–	2.6	88.0	2.9	–		–	–	54.5
45	*Jatropa curcas* Linn	–	–	1.4	15.6	–	9.7	40.8	32.1	–	0.4	–	–	52.31
46	*Moringa concanensis* Nimmo	–	–	–	9.7	–	2.4	83.8	0.8	–	3.3	–	–	56.32
47	*Ziziphus mauritiana* Lam	–	–	–	10.4	–	5.5	64.4	12.4	–	1.8	2.6	1.7	55.37

Table 2 continued

Item	Sources	FAME composition/wt%												Reported CN
		10:0	12:0	14:0	16:0	16:1	18:0	18:1	18:2	18:3	20:0	20:1	22:1	
48	*Croton tiglium* Linn	–	–	11.0	1.2	–	0.5	56.0	29.0	–	2.3	–	–	49.9
49	*Princeptia utilis* Royle	–	–	1.8	15.2	–	4.5	32.6	43.6	–	–	–	–	48.94
50	*Aegle marmelos correa* Roxb	–	–	–	16.6	–	8.8	30.5	36.0	8.1	–	–	–	48.3
51	*Vernonia cinerea* Less	–	–	8.0	23.0	–	8.0	32.0	22.0	–	3.0	–	–	57.51
52	*Joannesia princeps* Vell	–	–	2.4	5.4	–	–	45.8	46.4	–	–	–	–	45.2
53	*Garcinia combogia* Desr	–	–	–	2.3	–	38.3	57.9	0.8	0.4	–	0.3	–	61.5
54	*Garcinia indica* Choisy	–	–	–	2.5	–	56.4	39.4	1.7	–	–	–	–	65.16
55	*Illicium verum* Hook	–	–	4.43		–	7.93	63.24	24.4	–	–	–	–	50.71
56	Melia *azadirach* Linn	–	–	0.1	8.1	1.5	1.2	20.8	67.7	–	–	–	–	41.37
57	*Myristica malabarica Lam*	–	–	39.2	13.3	–	2.4	44.1	1.0	–	–	–	–	61.85
58	*Urtica dioica* Linn	–	–	–	9.0	–	–	14.6	73.7	2.7	–	–	–	38.73
59	*Tectona grandis* Linn	–	–	0.2	11.0	–	10.2	29.5	46.4	0.4	–	2.3	–	48.31

least-squares fitting method. In the regression analysis, the measured CN of pure FAME and the carbon number were used as inputs to give two unary nonlinear regression equations. Secondly, a multiple linear regression model was developed to predict the CN of biodiesel from its FAME composition. The 59 kinds of biodiesels were divided into two categories. The first set of 40 biodiesels from item 1 to 40 in Table 2 was used to develop the correlation. In the process, the reported CN and the corresponding FAME compositions of the 40 biodiesels as well as the CN of pure FAME, were used as inputs to obtain a regression equation. The least-squares fitting method was applied to obtain a multiple linear regression equation described as follows.

$$\mathrm{CN} = \alpha \sum (\mathrm{CN}_i w_i) + \alpha_0 \quad (5)$$

where CN is the predicted CN of biodiesel, CN_i is the CN of each pure FAME as shown in Table 1, w_i is the weight percentage of each FAME in biodiesel as independent variable of regression equation, α is the regression coefficient, and α_0 is the constant term. Then the second set, containing the remaining 19 biodiesels from item 41 to 59 in Table 2, was used to test the accuracy of the regression equation. The predicted CN was compared with the reported CN. The results from the present work were also compared with those reported in the literature.

Results and Discussion

CN of Pure FAME

From the data shown in Table 1, it appears that the CN increases with increasing chain length and the saturated degree of the fatty acid molecule. In order to investigate in depth the dependence of the CN of FAME with various saturated degrees on the structure of the fatty acid chain, the data in Table 1 were divided into two categories, saturated FAME and unsaturated FAME.

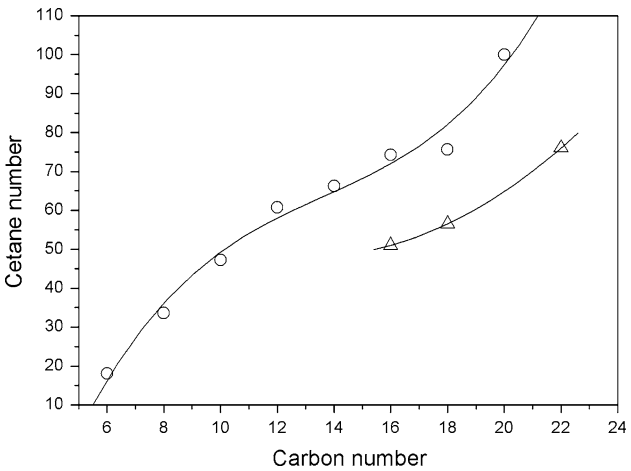

Fig. 1 Plot of CN of pure FAME versus carbon number of fatty acid chain. *Circle*, saturated FAME; *triangle*, unsaturated FAME containing one carbon–carbon double bond

For saturated FAME, there is a nonlinear increase in CN with increasing chain length of fatty acid moiety, as shown in Fig. 1. The CN was given as the dependent variable Y, while the carbon number of fatty acid chain was given as variable X. The fitted curve is a cubic polynomial described as the following regression equation:

$$Y = -107.71 + 31.126X - 2.042X^2 + 0.0499X^3 \quad (6)$$

A high correlation coefficient (R^2: 0.9835) indicates the excellent correlation between the CN and the carbon number for saturated FAME.

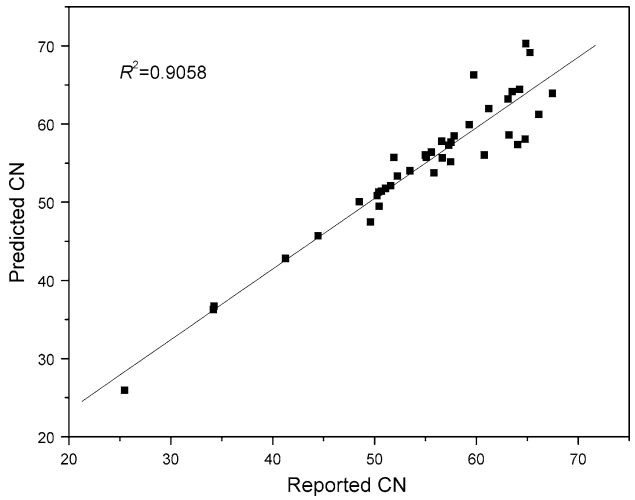

Fig. 2 Comparison of reported CN and predicted CN

For unsaturated FAME, the dependence of the CN on the carbon number is different from that for saturated FAME. The FAME containing one double bond like methyl palmitoleate (C16:1), methyl oleate (C18:1) and methyl erucate (C22:1) were selected to develop the correlation, because their CN were easily obtained from the literature comparing with other unsaturated FAME. As shown in Fig. 1, the CN increases nonlinearly with increasing carbon number. The nonlinearity can be described as the following quadratic equation:

$$Y = 109.000 - 9.292X + 0.354X^2 \quad (7)$$

where Y is the fitted CN and X is the carbon number of unsaturated fatty acid chain containing one carbon–carbon double bond. It is clear that the quantitative relationship between the CN of unsaturated FAME and the carbon number differs with that of saturated FAME. This could be attributed to the varying unsaturated degree. It may be presupposed that the regression equations for unsaturated FAME containing two or three double bond will also differ from Eq. 6 and Eq. 7.

Equation 6 and Eq. 7 developed in the present work as well as Eq. 2 obtained from the literature [11] were used to calculate the CN of pure FAME, as shown in Table 1. For the CN estimation of saturated FAME, the average relative error of calculated value from reported value is 6.5% according to Eq. 2. The more accurate result is observed from the present model, because the lower average relative error of 5.3% is given by Eq. 6 in our work. For the CN estimation of unsaturated FAME, the higher accuracy of Eq. 7 than that of Eq. 2 is observed for the lower relative error as shown in Table 1. These results indicate that the method used in the present work to divide pure FAME into saturated esters and unsaturated esters can reveal more clearly the dependence of the CN on the structure of the fatty acid chain, and give more accurate calculated CN of pure FAME.

Development of a Correlative Model for Biodiesel CN Prediction

Through the regression analysis based on the developing set of 40 biodiesels in Table 2 which was applied to develop the correlative model, the regression coefficient α and the constant term α_0 in Eq. 5 were obtained. Then Eq. 5 was changed into the following multiple linear regression equation:

$$\mathrm{CN} = 1.068 \sum (\mathrm{CN}_i w_i) - 6.747 \quad (8)$$

The predicted results of this model, the plots of the predicted CN versus the reported CN of the developing set, are depicted in Fig. 2. The linear correlation coefficient

Table 3 Statistic results of correlative models for CN prediction of 19 biodiesels in the testing set

Statistic constants[a]	The present work		Literature work 1[b]	Literature work 2[c]
	Testing set	All data		
Minimum absolute error	0.01	0.01	1.98	0.04
Maximum absolute error	1.65	6.71	4.02	2.78
Average absolute error	0.49	1.52	3.18	0.78
Minimum relative error (%)	0.02	0.02	3.43	0.07
Maximum relative error (%)	2.98	10.87	10.38	4.26
Average relative error (%)	0.96	2.74	6.23	1.48
Percentage with 2% (%)	78.9	59.3	0	63.2
R^2	0.9904	0.9234	0.9969	0.9782
Standard deviation	0.65	2.26	0.37	1.04

[a] Absolute error = |reported CN-predicted CN|. Relative error = |reported CN-predicted CN|/(reported CN) × 100%. Percentage with 2% is the percentage of the predicted value that has less than 2% relative error from the reported value

[b] Statistic results of predicted CN of the testing set according to Eq. 3 obtained from Ref. [18]

[c] Statistic results of predicted CN of the testing set according to Eq. 4 obtained from Ref. [22]

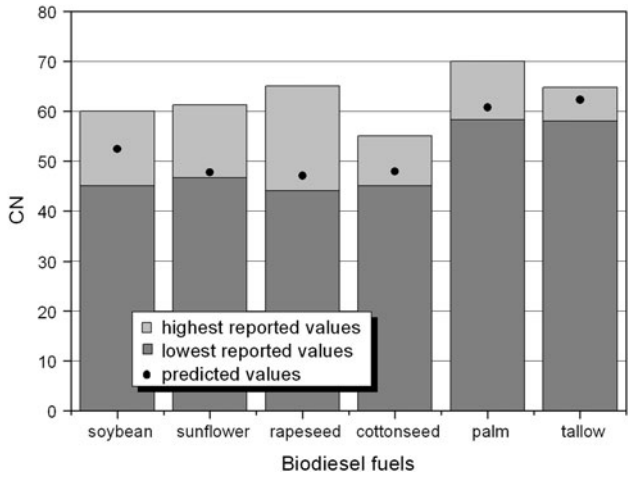

Fig. 3 CN of biodiesels derived from several conventional feed stocks and predicted CN. *Open squares*, highest reported value; *filled squares*, lowest reported value; *filled circles*, predicted value

($R^2 = 0.9058$) with a standard deviation of 2.73 was observed. The statistical result of the multiple linear regression model in Table 3 shows that the average absolute error of 1.52 CN even for all data of 59 biodiesels is well below the reproducibility limits of the ASTM D613 engine test method (2.5 ~ 3.3 CN).

The reported CN and the FAME composition of the testing set of 19 biodiesels were applied to validate the accuracy of the present model against the models in literature [18, 22]. The statistic results of predicted CN according to Eq. 3 [18] and Eq. 4 [22] obtained from the literature, and the present model Eq. 8, respectively, are given in Table 3. The maximum absolute error and the maximum relative error in the present model are 1.65 CN and 2.98%, all lower than the values in the other two models from the literature (4.02 CN and 10.38%, 2.78 CN and 4.26%, respectively). The average absolute error of 0.49 CN and the average relative error of 0.96% in the present model are also lower than that in the two models (3.18 CN and 6.23%, 0.78 CN and 1.48%, respectively). The percentage of the predicted value that has less than 2% relative error from the reported value is 78.9%, also higher than the highest value of 63.2% in the models from the literature. Comparing our work with the literature, it is observed that the present fitting results are more accurate. The higher accuracy may be attributed to the fact that the effect of more pure FAME like methyl caprate (C10:0), arachidate (C20:0), eicosenoate (C20:1) and erucate (C22:1) on the CN was calculated in the present model, comparing with the literature. Disregarding the effect of these FAME, widely existed in biodiesel, will result consequentially in the error of the predicted value from the actual value. The above results indicate that the multiple linear regression model (MLRM) developed in this study is suitable for estimating the CN of a biodiesel with high accuracy from its FAME composition.

CN Prediction of Biodiesels Derived from Several Conventional Parent Oil Sources

As mentioned above, the presence of the natural antioxidant in biodiesel results in a relatively stable CN. The FAME composition is considered as the primary factor leading to CN change against the oxidation of FAME. Even for the biodiesel derived from the same kind of parent oil source, the differences in the climatic conditions during the oil plants' growth and the oil processing methods can result in a variable FAME composition in the biodiesel, which brings on a wide variation of the CN [12–16, 18]. As shown in Fig. 3, biodiesels derived from several conventional parent oil sources display different CN values [30]. For instance, the CN of biodiesel derived from soybean oil ranges from 45 to 60, and that derived from sunflower oil ranges from 46.6 to 61.2, while the CN of biodiesel derived from palm oil rich in saturated fatty acid esters has a high value over the range from 58.3 to 70.

Several kinds of biodiesel derived from soybean oil, sunflower oil, rapeseed oil, cottonseed oil, palm oil and tallow, respectively, were selected to verify the practicability of the present model. Their typical FAME compositions are shown in Table 4 and were taken from the literature [24, 30–33]. The predicted CN of these biodiesels are also plotted in Fig. 3. A desired result is observed.

Table 4 Typical composition of FAME in biodiesels derived from several conventional feed stocks

Biodiesel	FAME composition/wt%									Reference
	12:0	14:0	16:0	16:1	18:0	18:1	18:2	18:3	20:0	
Soybean	–	0.56	14.17	1.27	5.19	48.20	22.19	1.45	0.28	[30]
Sunflower	–	0.10	5.40	0.10	5.20	48.20	41.00	–	–	[33]
Rapeseed	–	–	3.49	–	8.50	64.40	22.30	8.23	–	[24], [31]
Cottonseed	–	–	28.33	–	0.89	13.27	57.51	–	–	[24]
Palm	0.10	1.00	42.80	–	4.50	40.00	10.10	0.20	–	[24]
Tallow	0.60	2.91	24.34	3.44	19.1	40.23	2.58	0.33	0.29	[32]

namely, that all of the predicted values are in the range of the experimental values. This indicates that the present correlative model is a simple and practical method to give a satisfactory predicted CN of a biodiesel from its FAME composition without arduous and expensive experimental determination.

Conclusions

The relationship between the CN of pure FAME and the carbon number of fatty acid chain was established in the present work. Two different regression equations were obtained which show that the dependence of the CN on the carbon number varies with the unsaturated degree of the fatty acid chain. A convenient and practical regression equation with a high accuracy was developed, which shows excellent correlation ($R^2 = 0.9904$ for testing set) between the CN and the FAME composition and a good reproducibility (average absolute error: 0.49 CN for the testing set and 1.52 CN for all data). There are 78.9% of the predicted values having less than a 2% error from the reported values and a standard deviation of 0.65 for the testing set and 2.26 for all data. The practicability of the model was verified by its application in predicting CN of biodiesels derived from several conventional parent oil sources. All of these results indicate that the MLRM model developed in the present work can be used conveniently to give a satisfactory predicted CN of biodiesel without arduous and expensive experimental determination.

Acknowledgments This work was financially supported by the National Basic Research Program of China (973 program, No. 2007CB210203) and the Scientific Research Foundation for Young Teachers of Sichuan University (No. 2009SCU11104). The supervision of academician Qingshi Zhu is highly appreciated.

References

1. Demirbas A (2007) Importance of biodiesel as transportation fuel. Energ Policy 35:4661–4670
2. Johnston M, Holloway T (2007) A global comparison of national biodiesel production potentials. Environ Sci Technol 41: 7967–7973
3. Ladommatos N, Parsi M, Knowles A (1996) The effect of fuel cetane improver on diesel pollutant emissions. Fuel 75:8–14
4. Ghosh P, Jaffe SB (2006) Detailed composition-based model for predicting the cetane number of diesel fuels. Ind Eng Chem Res 45:346–351
5. DeFries TH, Indritz D, Kastrup RV (1987) Prediction of cetane number by group additivity and carbon-13 nuclear magnetic resonance. Ind Eng Chem Res 26:188–193
6. Gülder ÖL, Glavincevski B (1986) Prediction of cetane number of diesel fuels from carbon type structural composition determined by proton NMR spectroscopy. Ind Eng Chem Prod Res Dev 25:153–156
7. Ladommatos N, Goacher J (1995) Equations for predicting the cetane number of diesel fuels from their physical properties. Fuel 74:1083–1093
8. Yang H, Ring Z, Briker Y, McLean N, Friesen W, Fairbridge C (2002) Neural network prediction of cetane number and density of diesel fuel from its chemical composition determined by LC and GC–MS. Fuel 81:65–74
9. Freedman B, Bagby MO (1990) Predicting cetane numbers of n-alcohols and methyl esters from their physical properties. J Am Oil Chem Soc 67:565–571
10. Bose PK (2009) Empirical approach for predicting the cetane number of biodiesel. International J Autom Technol 10(4): 421–429
11. Lapuerta M, Rodríguez-Fernández J, de Mora EF (2009) Correlation for the estimation of the cetane number of biodiesel fuels and implications on the iodine number. Energ Policy 37:4337–4344
12. Graboski MS, McCormick RL (1998) Combustion of fat and vegetable oil derived fuels in diesel engines. Prog Energ Comb Sci 24:125–164
13. Knothe G (2008) "Designer" biodiesel: optimizing fatty ester composition to improve fuel properties. Energ Fuels 22:1358–1364
14. Marchetti JM, Miguel VU, Errazu AF (2007) Possible methods for biodiesel production. Renew Sust Energ Rev 11:1300–1311
15. Mohibbe Azam M, Waris A, Nahar NM (2005) Prospects and potential of fatty acid methyl esters of some non-traditional seed oils for use as biodiesel in India. Biomass Bioenerg 29:293–302
16. Sharma YC, Singh B, Upadhyay SN (2008) Advancements in development and characterization of biodiesel: a review. Fuel 87:2355–2373
17. Knothe G (2005) Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. Fuel Process Technol 86:1059–1070
18. Bamgboye AI, Hansen AC (2008) Prediction of cetane number of biodiesel fuel from the fatty acid methyl ester (FAME) composition. Int Agrophys 22:21–29
19. Ramos MJ, Fernández CM, Casas A, Rodríguez L, Pérez A (2009) Influence of fatty acid composition of raw materials on biodiesel properties. Bioresour Technol 100:261–268
20. Van Gerpen J (1996) Cetane number testing of biodiesel. In: Proceedings of the 3rd liquid fuel conference, American society of agricultural engineers, Nashville, TN, pp 197–206
21. Wadumesthrige K, Smith JC, Wilson JR, Salley SO, Ng KYS (2008) Investigation of the parameters affecting the cetane number of biodiesel. J Am Oil Chem Soc 85:1073–1081
22. Gopinath A, Puhan S, Nagarajan G (2009) Relating the cetane number of biodiesel fuels to their fatty acid composition: a critical study. Proc I MechE Part D J Autom Eng 223:565–583
23. Ramadhas AS, Jayaraj S, Muraleedharan C, Padmakumari K (2006) Artificial neural networks used for the prediction of the cetane number of biodiesel. Renew Energ 31:2524–2533
24. Ma F, Hanna MA (1999) Biodiesel production: a review. Bioresour Technol 70:1–15
25. Rashid U, Anwar F, Moser BR, Knothe G (2008) *Moringa oleifera* oil: a possible source of biodiesel. Bioresour Technol 99:8175–8179
26. Klopfenstein WE (1985) Effect of molecular weights of fatty acid esters on cetane numbers as diesel fuels. J Am Oil Chemi Soc 62:1029–1031
27. Knothe G, Bagby MO, Ryan TW (1997) Cetane numbers of fatty compounds: influence of compound structure and of various potential cetane improvers. SAE Technical Paper Series 971681
28. Knothe G, Dunn RO, Bagby MO (1997) Biodiesel: the use of vegetable oils and their derivatives as alternative diesel fuels. ACS Symposium Series 666, pp 172–208. doi:10.1021/bk-1997-0666.ch010

29. Knothe G, Matheaus AC, Ryan TW (2003) Cetane numbers of branched and straight-chain fatty esters determined in an ignition quality tester. Fuel 82:971–975
30. Geller DP, Goodrum JW (2004) Effects of specific fatty acid methyl esters on diesel fuel lubricity. Fuel 83:2351–2356
31. Goering CE, Schwab A, Daugherty MJ, Pryde EH, Keakin AJ (1981) Fuel properties of eleven vegetable oils. ASAE Tech Paper, OSTI ID: 6493624
32. Kinast JA (2003) Production of biodiesels from multiple feedstocks and properties of biodiesels and biodiesel/diesel blends. NREL Repots, NREL/SR-510-31460
33. Soumanou MM, Bornscheuer UT (2003) Improvement in lipase-catalyzed synthesis of fatty acid methyl esters from sunflower oil. Enzyme Microb Technol 33:97–103