

PREDICTION OF DENSITY AND CETANE NUMBER OF DIESEL FUEL FROM GC-FIMS AND PIONA HYDROCARBON COMPOSITION BY NEURAL NETWORK

Hong Yang, Yevgenia Briker, Renata Szykarczuk and Zbigniew Ring

The National Centre for Upgrading Technology, Devon, AB,
Canada, T9G 1A8

Introduction

Modeling and optimization of bitumen upgrading processes require reliable correlations to predict the key properties of the product stream based on its computer-generated composition. In the approach taken by the National Centre for Upgrading Technology (NCUT), these models are designed so that this computer-generated composition is equivalent to the various available chromatographic tests. This effectively decouples the development of the main parts of any process model: the reactor and product quality models. As part of its process modeling program, NCUT has assembled a large database of diesel fuels. The chemical compositions of the diesel fuels were characterized by several chromatographic methods. A neural network approach was then taken to correlate fuel properties such as density, aniline point, cloud point, viscosity, refractive index, molecular weight and cetane number with chemical compositions. This work presents the preliminary results on the correlation of density and cetane number with the hydrocarbon compositions of diesel fuels given by the sum of GC-FIMS (Gas Chromatograph-Field Ionization Mass Spectrometry) and PIONA results (normal paraffins, isoparaffins, olefins, naphthenes and aromatics).

Method

One hundred fourteen diesel samples were prepared by blending 15 original diesel components obtained from Canadian refineries and derived from both conventional crude oil and oil-sands bitumen. Cetane number and density (g/cm^3 , 15.6°C) of the diesel fuels were measured using the ASTM D613 and ASTM D4052 methods, respectively. Hydrocarbon type compositions of the fuels were determined by GC-FIMS and PIONA. In each sample, the material boiling in the $200 - 343^\circ\text{C}$ range was analyzed by GC-FIMS while PIONA was used to analyze the material boiling at temperatures lower than 177°C . The sum of GC-FIMS and PIONA results, weighted with the corresponding mass fractions, gave the total hydrocarbon composition of the blend. For the GC-FIMS determinations, a $30 \text{ m} \times 250\mu\text{m} \times 0.25\mu\text{m}$ HP1-MS non-bonded column was used. The injection ($0.2\mu\text{L}$; 19:1 split) was made with the oven at 45°C . The AC PIONA analyzer based on HP GC 5890 instrument was used to perform the analysis. It was operated under the 'mode 20' conditions (normal paraffins, isoparaffins, naphthenes, and aromatics). Further details are reported elsewhere¹.

A three-layer Ward backpropagation network with three hidden slabs (WSGN – NeuroShell® software, Ward System Group Inc. MD, USA) was used in neural network correlations². Including the original diesel fuels and the diesel blends, a total of 129 samples were used to construct the neural network correlations. The 129 samples were divided into three data sets: training set, test set and production set. Two steps were required to create the three-layer backpropagation neural network model: a training step and a validation step. In the training step, the neural network was supplied with the training data set, including the input and corresponding output values. The network learned the trends contained in the data

set and correlated the inputs and outputs by finding the optimum set of weights that minimized the differences between the predicted outputs and the actual outputs. The test set was used with calibration during the training process to prevent over-training of networks, such that they would generalize well on new data. During the validation step, the neural network was provided with the production data set, not seen during the training step, to compute an average error for the test set of this model. The training terminated when 20,000 epochs had passed since reaching of the minimum average error for the test data set.

Twelve hydrocarbon types were used as neural network inputs, and density or cetane number was the network output. Table 1 lists the maximum and minimum values of the inputs and outputs for all the diesel samples used in this study. As shown in the table, the diesel samples used in this work cover a wide range of chemical composition and physical properties.

Table 1. Maximum and minimum values of the neural network inputs and outputs

	Max	Min
Inputs (mass%)		
Isoparaffins	33.09	0.70
n-Paraffins	20.95	0.37
Monocycloparaffins	29.42	1.60
Dicycloparaffins	28.38	0.27
Polycycloparaffins	31.80	1.29
Alkylbenzenes	18.31	5.81
Benzocycloalkanes	31.36	3.08
Benzodicycloalkanes	10.37	0.09
Diaromatics	62.51	0.74
Triaromatics	3.92	0.00
Tetraaromatics	0.63	0.00
Aromatic sulfurs	5.07	0.01
Outputs		
Density (g/ml @ 15.6°C)	0.9569	0.7985
Cetane (CN)	58.4	18.7

Results and Discussion

The consistency of the blending process was checked by comparing the density of the diesel blends measured by ASTM D4052 with the calculated weight averaged density derived from original diesel fuels by the following equation:

$$d = \sum_{i=1}^n d_i W_i$$

where n is the number of original diesel blending components that contributed to the blend, d_i is the density of the i^{th} blending component, W_i is the weight fraction of the i^{th} component in the blend. Figure 1 plots the calculated densities of the blends versus those measured by ASTM D4052. Good agreement between the two sets of data is demonstrated there.

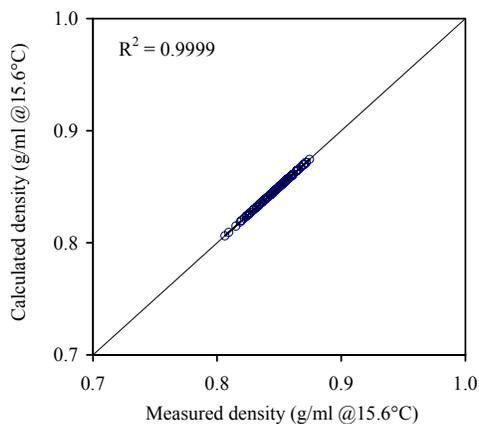


Figure 1.
Comparison of weighted average density with measured density

Density correlation

The twelve hydrocarbon types listed in Table 1 were used as inputs, and density as the output for the three-layer Ward backpropagation network. Seventy-five diesel blends were chosen as the neural network training data set. To cover the whole range of the diesel fuels in this study, the training data set included the 15 original diesel blending components that contained the maximum and minimum values of inputs and outputs. Twenty diesel blends were included in the test set and 34 diesel blends were in the production set.

The densities of all the diesel blends were calculated with the obtained neural network models. The results are plotted in Figure 2 (open circle) showing parity between the calculated and measured densities. Unfortunately, the correlation coefficient obtained was only 0.8560.

In order to improve the predictions, we carefully examined the boiling range (Simdist ASTM D2887) of the diesel blends, and found that 30 diesel blends contained substantial amount of material boiling higher than 343°C. Since the GC-FIMS method we employed could only give reliable hydrocarbon compositions in the range between 200 to 343°C, the hydrocarbon compositions of these 30 heavier diesel blends were found erroneous. A new neural network model was completed after we removed the 30 blends from the data set. The new model was used to calculate the densities of the remaining 99 diesel blends. The results were compared with the previous ones in Figure 2 (close circle). A significant improvement was achieved. A correlation coefficient of 0.9875 was obtained in this case.

Cetane number correlation

Neural network correlations for predicting cetane numbers from the diesel fuels' chemical compositions were established using both full diesel blends (129 sample data) and diesel blends without the 30 heavy samples (99 sample data). Figure 3 shows the neural network predicted versus measured cetane numbers in both cases. The results indicate that the removal of the 30 heavy diesel blends benefits the cetane number correlation.

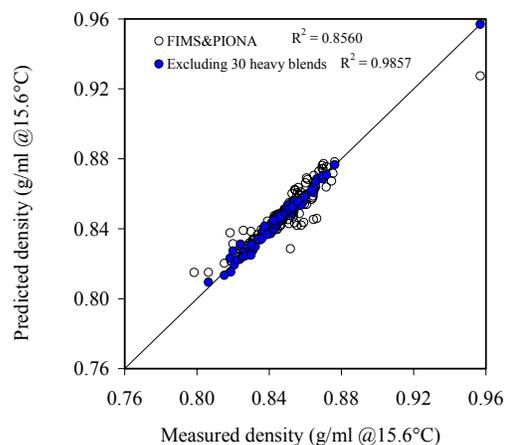


Figure 2.
Comparison of calculated density with measured density with (○) and without (●) 30 heavy diesel blends

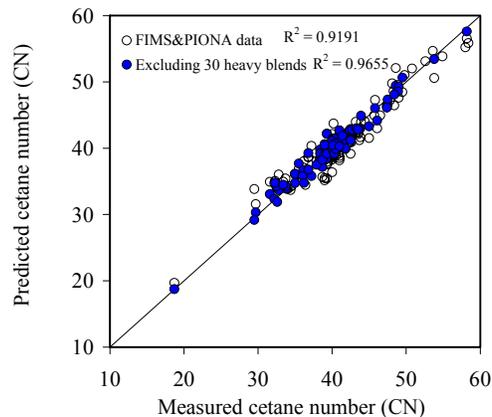


Figure 3.
Comparison of calculated cetane number with measured cetane number with (○) and without (●) 30 heavy diesel blends

NCUT had previously developed neural network correlations for densities and cetane numbers of diesel fuels from chemical compositions using the same three-layer Ward backpropagation network³. A smaller database with 69 diesel blends was used in that case. LC/GC-MS (liquid chromatography (LC)/gas chromatography-mass spectrometry) was used to determine the chemical compositions of diesel fuels. Good correlation results were obtained in our previous study. In the present work, the chemical composition was determined by the sum of PIONA and GC-FIMS. GC-FIMS has several advantages over GC-MS. In GC-FIMS the sample can be analyzed entirely without separating into saturate and aromatic fractions, which reduces the error caused by lost of lighter and heavy ends during the separation procedure. GC-FIMS can also distinguish n-paraffins from iso-paraffins, which increases the prediction accuracy of the correlations since n-paraffins and iso-paraffins have different effects on cetane number and density. The statistical results

of density and cetane number obtained in the present work and previous work are compared in Table 2.

Table 2. Comparison of the statistical results for density and cetane number predictions using neural network models derived from GC-MS and GC-FIMS +PIONA results

	Density		Cetane number	
	GC-MS	GC-FIMS	GC-MS	GC-FIMS
	PIONA		PIONA	
R-square	0.96	0.99	0.91	0.97
Mean absolute error	0.004	0.002	1.32	0.72
Max absolute error	0.009	0.007	6.9	2.85
Percentage with 5%	100	100	81.2	94.9

Two conclusions can be drawn from Table 2. First, the mean absolute error for cetane number prediction from both neural network models were below the reproducibility limits of the ASTM D613 engine test method, which was between 2.8 to 4.8 depending on the cetane number of diesel fuel. However, the mean error for the density prediction was not as good as the reproducibility of the ASTM D4052 test (0.0005). Second, the neural network correlations developed using more detailed hydrocarbon composition (sum of GC-FIMS and PIONA) as well as a larger database (the current work) resulted in substantial improvements of density and cetane number predictions over the neural network correlations developed using GC-MS and the smaller database. The mean absolute error for density decreased from 0.004 to 0.002 when the current model replaced the previous neural network model. The mean absolute error for cetane number reduced from 1.32 to 0.72 when the new model was used. These correlations could probably be further improved by introducing more inputs or slightly different inputs to the correlations. The former requires a greater experimental database but the latter could be done using the same database.

Conclusions

A neural network method was used to establish correlations for density and cetane number for a diesel fuel from its chemical composition determined by GC-FIMS and PIONA. The results show that reliability of the hydrocarbon compositions is very important to create accurate neural network correlations. Significant improvement was obtained for both density and cetane number correlation after the removal from the database of 30 heavy diesel blends that could not be characterized correctly by the GC-FIMS method. The neural network correlation could predict the cetane number with mean absolute error well below the reproducibility limit of the ASTM engine test method. However, further effort is needed to develop a better correlation for density prediction. Our results also showed that significantly better neural network correlations were obtained using hydrocarbon compositions derived from GC-FIMS + PIONA and a large database.

Acknowledgement. Partial funding for this work has been provided by the Canadian Program for Energy Research and Development (PERD), the Alberta Research Council and The Alberta Energy Research Institute. The authors wish to thank NCUT's analytical staff for all the analyses.

References

- (1) Briker, Y.; Ring, Z.; Iacchelli, A.; McLean, N.; Fairbridge, C., Malhotra, R.; Coggiola, M.; Young, S.; *Energy & Fuels*, **2001**, *15*, 996.
- (2) NeuroShell® 2 Manual. Ward Systems Groups, Inc. MD. 1996.
- (3) Yang, H.; Ring, Z.; Briker, Y.; Mclean, N.; Friesen, W.; Fairbridge, C.; *Fuel*, **2002**, *81*, 65.