Scientific paper

Development of Artificial Neural Network Model for Diesel Fuel Properties Prediction using Vibrational Spectroscopy

Tomislav Bolanča,[1,*] Slavica Marinović,[2] Šime Ukić,[1] Ante Jukić[1] and Vinko Rukavina[2]

[1] *University of Zagreb, Faculty of Chemical Engineering and Technology, Marulićev trg 19, 10000 Zagreb, Croatia*

[2] *INA – Oil Industry Ltd., Refining and Marketing, Lovinčićeva bb, 10000 Zagreb, Croatia*

** Corresponding author: E-mail: tomislav.bolanca@fkit.hr*
Tel.: +385 1 4597209; fax: +385 1 4597250

Received: 31-05-2011

Abstract

This paper describes development of artificial neural network models which can be used to correlate and predict diesel fuel properties from several FTIR-ATR absorbances and Raman intensities as input variables. Multilayer feed forward and radial basis function neural networks have been used to rapid and simultaneous prediction of cetane number, cetane index, density, viscosity, distillation temperatures at 10% (T10), 50% (T50) and 90% (T90) recovery, contents of total aromatics and polycyclic aromatic hydrocarbons of commercial diesel fuels.

In this study two-phase training procedures for multilayer feed forward networks were applied. While first phase training algorithm was constantly the back propagation one, two second phase training algorithms were varied and compared, namely: conjugate gradient and quasi Newton. In case of radial basis function network, radial layer was trained using K-means radial assignment algorithm and three different radial spread algorithms: explicit, isotropic and K-nearest neighbour.

The number of hidden layer neurons and experimental data points used for the training set have been optimized for both neural networks in order to insure good predictive ability by reducing unnecessary experimental work.

This work shows that developed artificial neural network models can determine main properties of diesel fuels simultaneously based on a single and fast IR or Raman measurement.

Keywords: Artificial neural network, FTIR-ATR, Raman, diesel fuel

1. Introduction

The present determination of diesel fuel properties is based on standard methods such as ASTM, ISO and others. These methods can be time consuming (gas and liquid chromatography, distillation, etc.) and may require large sample size and the use of toxic and environmentally hazardous chemicals.[1–7] Vibrational spectroscopic methods (IR, NIR and Raman) may be an effective alternative to these standard procedures[8–14] since their utilization allows development of analytical methodologies that are fast and clean, use only a few millilitres of sample and do not require extensive use of reagents.[15] Multivariate regression analysis and vibrational spectroscopy has already been applied to predict the characteristics of gasoline. Usually combines infrared spectroscopy and the PLS (e.g. partial least squares, PLS) algorithm to determine properties of gasoline like research and motor octane number,[16] aromatics, olefins and saturated hydrocarbons, oxygenates, etc[17]. Cooper et al. have used FT-Raman spectroscopy and the PLS algorithm to determine the octane number and Reid vapor pressure in commercial gasolines,[18] and also a combination of near-IR, mid-IR and Raman spectroscopy to determine BTEX and weight percent oxygen.[19,20]

Some research dealing with the prediction of (bio) diesel fuel properties has already been presented. Most of published papers employ infrared spectroscopy and multi-

variate regression analysis to predict the characteristics of (bio) diesel fuel.[7,9,12]

As an alternative to established laboratory protocols, partial least squares (PLS) regression models based on Fourier transform infrared (FTIR) spectra were developed for the rapid and simultaneous determination of several middle distillate fuel properties.[21] On the basis of this work, the following middle distillate fuel properties may be confidently estimated by FTIR: gravity, API, density, viscosity, boiling point at 50% recovery, cetane index, carbon, hydrogen, carbon-hydrogen ratio, heat of combustion and aromatic compounds. Recently, infrared spectroscopy and multivariate calibration have been used to predict cetane index and distillation temperatures at 10, 50, 85 and 90% recovery of diesel fuel.[22] The infrared spectroscopy and PLS algorithm have also recently been used to predict properties of diesel/biodiesel blends.[23] Artificial neural networks (ANNs) are also powerful modelling tools, which can be used to correlate and identify highly complex relationships from input-output data only. Regression problems can be solved using following network types: multilayer feed forward (MLP), radial basis function (RBF), and general regression neural network. MLPs are most commonly used ones.[24] There are many algorithms for training MLP networks. The popular back propagation (BP) algorithm is simple but has a problem with slow convergence.[25] The second most commonly used neural network architecture is RBF network. Compared with MLP network, this network possesses certain advantages that make it very popular. The most important advantages are the simplicity of the network structure and the high speed of convergence during the training phase.[26]

Some researches have already been conducted on the analysis of gasoline[27,28] and diesel fuel using ANNs.[29,30] Neural networks were used to correlate and predict the cetane number and the density of diesel fuel from its chemical composition.[31] In this study, mean absolute errors with the test data set were 1.23 and 0.002 g/cm^3 for the cetane number and density, respectively. The diesel fuel lubricity has been determined using RBF network with other fuel properties as inputs.[32] The lubricity predicted by this neural network gave a coefficient of determination $R^2 = 0.94$, with more than 90% of the predicted lubricity values lying within the 95% confidence limit.

Furthermore, the cold filter plugging point of blended diesel fuel was determined using input parameters of kinematic viscosity, density, refractive intercept and the specific distillation range.[33] Finally, the distillation profile and cold properties of diesel fuel have been predicted employing mid-IR spectroscopy and MLP networks.[27] The developed models predict these properties based on the IR signal, with the level of accuracy characteristic for the repeatability of standard methods.

Application of Raman spectroscopy in determination of diesel fuel properties is generally poorly represented. Raman spectroscopy has not been extensively applied in the industry due to several constraints; high cost, low S/N ratio when compared to near-IR or mid-IR, fluorescence problems, etc. Although Raman spectroscopy is still quite expensive, most disadvantages have been solved thanks to its combination with the Fourier transform, more powerful laser and multivariate chemometric techniques.[34] One of the advantages of the Raman spectroscopy is the capability of glass vials application in order to obtain spectra, which compared to mid-IR spectroscopy accelerated and simplified measurement procedure, while near-IR spectroscopy is significantly less informative. In the literature we have found only one article that combine FT-Raman spectroscopy and MLP network for the prediction of cetane index, density, viscosity, T50, T90 and total sulphur content of diesel fuel.[35] The aim of the present work is to develop models based on infrared and Raman spectroscopy combined with MLP and RBF networks for rapid and accurate simultaneous determination of the most important physico-chemical properties of diesel fuel: cetane number, cetane index, density, viscosity, distillation temperatures at 10% (T10), 50% (T50) and 90% (T90) recovery, the contents of total aromatics and polycyclic aromatic hydrocarbons.

2. Theory

Neural networks are powerful modelling tools that have the ability to identify underlying highly complex relationships from input-output data only.[32] Using this approach it is possible to derive empirical models from a collection of experimental data, especially if the data that have to be correlated exhibit a complex nonlinear behaviour and cannot be described by linear mathematical models. Such models are obtained by training, i.e., the network is repeatedly presented with input/output pairs that have to be correlated. Although the training procedure can be quite time-consuming, once trained, the network produces an answer or prediction almost instantaneously.

Multilayer perceptrons (MLPs) and radial basis function (RBF) networks are the two most commonly-used types of feed forward network.[36,37]

2. 1. Multilayer Perceptron Neural Networks

MLP is one of the most popular network types, and in many problem domains seem to offer the best possible performance. It consists of several neuron layers: input layer, one or more hidden layers, and the output layer. Each MLP neuron performs a biased weighted sum of their inputs and passes this activation level through a transfer function to produce their output. The neurons are arranged in a layered feed forward topology. The network thus has a simple interpretation as a form of input-output model, with the weights and thresholds (biases) as free parameters of the model. Such networks can model functions of

almost arbitrary complexity, with the number of layers, and the number of neurons in each layer, determining the function complexity.[38]

The number of neurons in the first layer is equal to the number of input parameters and the values of input parameters are the layer's output. In case of other layers outputs are obtained in a further way. Let y_i^l be the output of i^{th} neuron of the l^{th} network layer, which can be computed according to further formulas as:

$$y_i^l = f\left(\sum_{j=1}^{N_{l-1}} w_{ij}^l y_j^{l-1} + \theta_i^l\right) \quad (1)$$

where N represents number of neurons in specific layer. Function f is the activation function, w_{ij}^l weight of the link between the j^{th} neuron of the $l-1^{\text{st}}$ layer and i^{th} neuron of the l^{th} layer, θ_i^l the bias parameter of i^{th} neuron of the l^{th} layer. Usually, as activation function for hidden layer(s), sigmoidal function is usually used, while linear function is applied in case of output layer.

Once the number of layers, and number of units in each layer, has been selected, the network's weights and must be set so as to minimize the prediction error made by the network.[37] This is the role of the training algorithms. The error of a particular configuration of the network can be determined by running all the training cases through the network, comparing the actual output generated with the desired or target outputs. The differences are combined together by an error function to give the network error. The most common error functions are the sum-squared error, where the individual errors of output units on each case are squared and summed together. Networks are trained using iterative algorithms, of which the best known is back propagation.[39] A considerable amount of research has been conducted into improved algorithms for training of multilayer perceptrons. The most influential of these are the second-order optimization algorithms[40,41] (conjugate gradient descent, quasi Newton). These algorithms are usually described as converging far more quickly than back propagation (one or two orders of magnitude faster). Conjugate gradient descent usually performs significantly better than , and it is the recommended technique for any network with a large number of weights (more than a few hundred) and/or multiple outputs and it is also a highly effective generic algorithm with low computer memory requirements and good stability. Quasi Newton is usually a little faster than conjugate gradient descent, but has substantially larger memory requirements and it is occasionally numerically unstable.[37]

2. 2. Radial Basis Function Neural Networks

A radial basis function (RBF) neural network has an input layer, a hidden and an output layer. The hidden layer neurons act as cluster centres, grouping similar training cases. The neurons in the hidden layer contain Gaussian transfer functions whose outputs are inversely proportional to the distance from the centre of the neuron.

The general form of the Gaussian function is:

$$Output = \exp\left(\frac{-x^2}{2\sigma^2}\right) \quad (2)$$

where *ó* (standard deviation) controls the spread of the function, and x is the Euclidean distance between the cluster centre and the input vector.

A radial basis function network (RBF), therefore, has a hidden layer of radial units, each actually modelling a Gaussian response surface. Since these functions are nonlinear, it is not actually necessary to have more than one hidden layer to model any shape of function: sufficient radial units will always be enough to model any function.[37]

RBF networks use a two stage training process – first, assignment of the radial centres and their deviations; second, optimization of the output layer. A classic RBF uses the identity activation function in the output layer, in which case linear optimization (pseudo-inverse, SVD) can be used, which is relatively quick compared with training.

Centres should be assigned to reflect the natural clustering of the data. The two most common methods are: sub-sampling and K-means algorithm. K-means algorithm tries to select an optimal set of points that are placed at the centroids of clusters of training data. Given K radial units, it adjusts the positions of the centres so that: (i) each training point belongs to a cluster centre; (ii) each training point is nearer to the belonging cluster centre than to any other centre. Once centres are assigned, deviations are set. The three most common methods to determine deviation (radial spread) are explicit deviation (chosen by the user), isotropic deviation (same for all units and selected heuristically to reflect the number of centres and the volume of space they occupy) and K-nearest neighbour (each unit's deviation is individually set to the mean distance to its K-nearest neighbours).[41,42,43] Hence, deviations are smaller in tightly packed areas of space, preserving detail, and higher in sparse areas of space.

RBF networks have a number of advantages over MLPs. First, as previously stated, they can model any nonlinear function using a single hidden layer, which removes some design-decisions about numbers of layers. Second, the simple linear transformation in the output layer can be optimized fully using traditional linear modelling techniques, which are fast and do not suffer from problems such as local minima which plague MLP training techniques. RBF networks can therefore be trained extremely quickly (i.e., orders of magnitude faster than MLPs). However, RBF's more eccentric response surface requires more units to adequately model most functions and, consequently, an RBF solution will tend to be slower to execute and more space consuming. The second disadvantage of RBF networks is its disability to extrapolate

beyond known data, since the response drops off rapidly towards zero if data points far from the training data are used. RBFs are also mores sensitive to the course of dimensionality and have greater difficulties if the number of input data is large.[37]

3. Experimental

3. 1. Samples and Experimental Procedures

93 diesel samples were collected from gas stations and the storage tanks (Croatia)[44] during a period of four months and stored in tightly sealed glass bottles at maximal temperature of 4 °C. Before instrumental analysis, samples were equilibrated at room temperature (22 ± 5 °C).

The standard methods[45–50] were mainly used for determination of the various diesel fuel properties. The samples were tested for: cetane number (internal method); cetane index (ASTM D 4737); density (ASTM D 1298); viscosity (ASTM D 445); distillation temperatures at 10% (T10), 50% (T50) and 90% (T90) recovery (ASTM D 86); and contents of total aromatics and polycyclic aromatic hydrocarbons (EN 12916).

FTIR spectra with attenuated total reflectance (ATR) and FT-Raman spectra were obtained on Nicolet 6700 Fourier transform instrument (Thermo Fisher Scientific Inc, USA). FTIR-ATR spectra were recorded using a Smart Performer sampling accessory and ZnSe cell, 50 scans with a resolution of 6 cm^{-1}, covering the 4000–650 cm^{-1} spectral range, DTGS detector and KBr beamsplitter. Before measuring each sample, a background spectrum is obtained using clean and dry cell following the same process as for the samples. The FT-Raman spectra were recorded in quartz cuvettes with a Teflon stopper, covering the 3700–350 cm^{-1} spectral range, 50 scans with a resolution of 8 cm^{-1}. A liquid nitrogen cooled Ge detector was used for signal detection. The laser excitation of 1064 nm was provided by NdYAG laser and laser power was set to 0.400 W. Typical FTIR-ATR and FT-Raman spectra of diesel fuels are presented in Figure 1.

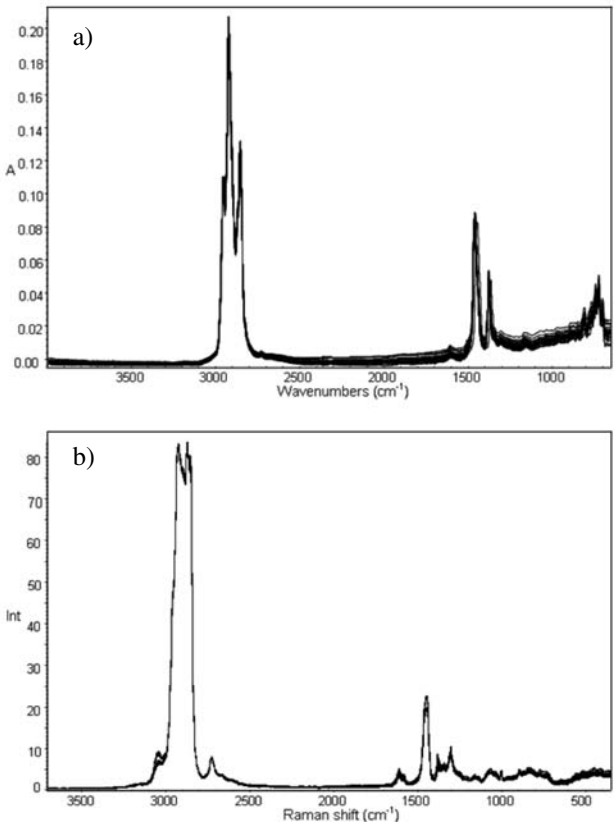

Figure 1. Typical spectra of diesel fuels: (a) FTIR-ATR, (b) FT-Raman.

3. 2. Neural Networks

The neural networks used in this work were MLP and RBF networks. As independent input variables for networks, several IR absorbances and Raman intensities were selected after visual examination of spectra (Figure 1). Physico-chemical properties of diesel fuels are the result of chemical composition. The main peaks observed in Figure 1 are associated with major functional groups present in this type of fuel. Since mid-IR and Raman spectroscopy are compatible techniques, spectral bands have similar wavenumbers, but show differences in intensity. Briefly, the bands between 3200–3000 cm^{-1} in the Raman spectrum are attributed to the C–H stretching of aromatic compounds and the region 3000–2800 cm^{-1} corresponds to C–H stretching of saturated n-alkyl groups. The bands in the region of 1500 to 1400 cm^{-1} are associated with the C–H deformation of CH_2 and CH_3 groups, and region 700–900 cm^{-1} is attributed to the C–H out of plane bending in different types of substituted benzene rings. The band around 1378 cm^{-1} in the Raman spectrum represents ring stretching of bicyclic aromatic fractions, and the maximum at 1302 cm^{-1} corresponds to twist and rock vibrations of n-alkanes. The most prominent band, occurring at 1002 cm^{-1}, arose from the symmetrical (trigonal) ring-breathing mode of monocyclic aromatic components in the fuel.

Based on assumed high correlation between spectral information (described above) and corresponding fuel properties it is possible to select 15 mid-IR absorbances from FTIR-ATR spectra and 17 intensities form FT-Raman spectra as input variables without additional need for principal component analysis and unnecessary prolongation of calculation procedures.

The output layer consists of nine neurons representing properties of diesel fuels (cetane number, cetane index, density, viscosity, T10, T50, T90, contents of total aromatics and polycyclic aromatic hydrocarbons) determined by standard methods.

The optimizations were performed in order to achieve precise and accurate model with respect to minimiza-

tion of unnecessary experimentation and time needed for the ANN training calculations.

For MLP network, the training algorithm, number of hidden layer neurons and number of experimental data points used for training were optimized. The advantage of the used MLP models was application of two-phase training approach. Two-phase training is a combination of two training algorithms, which enables to use advantages of both algorithms in same training procedure, resulting with better predictive ability obtained within shorter calculation time.[51] The first phase was 100 iteration steps of error back propagation training in order to achieve fast convergence to the region of global minimum on error surface. The second phase algorithm was varied between conjugate gradient (CG) and quasi Newton (QN) algorithm. The number of neurons in the hidden layer was varied from 2 to 14 (step 2) and number of experimental data in training set was varied from 15 to 45 (step 5). The second phase training procedure had been repeated until the global minimum on error surface was found. The logistic function was used as activation function connecting input and hidden layer and identity function was used as activation function connecting hidden and output layer.

For the RBF network, the radial layer was trained using K-means radial assignment algorithm and three different radial spread algorithms: explicit, isotropic, and K-nearest neighbour. The parameters for the radial spread training algorithms were optimized. The values 1, 1 and 10 were used for optimization as parameters for explicit; isotropic and K-nearest neighbour algorithms, respectively. The number of hidden layer neurons and experimental data points used for the training set were also optimized. The number of neurons in the hidden layer was varied from 5 to 17 (step 2) and number of experimental data in training set was varied from 20 to 45 (step 5).

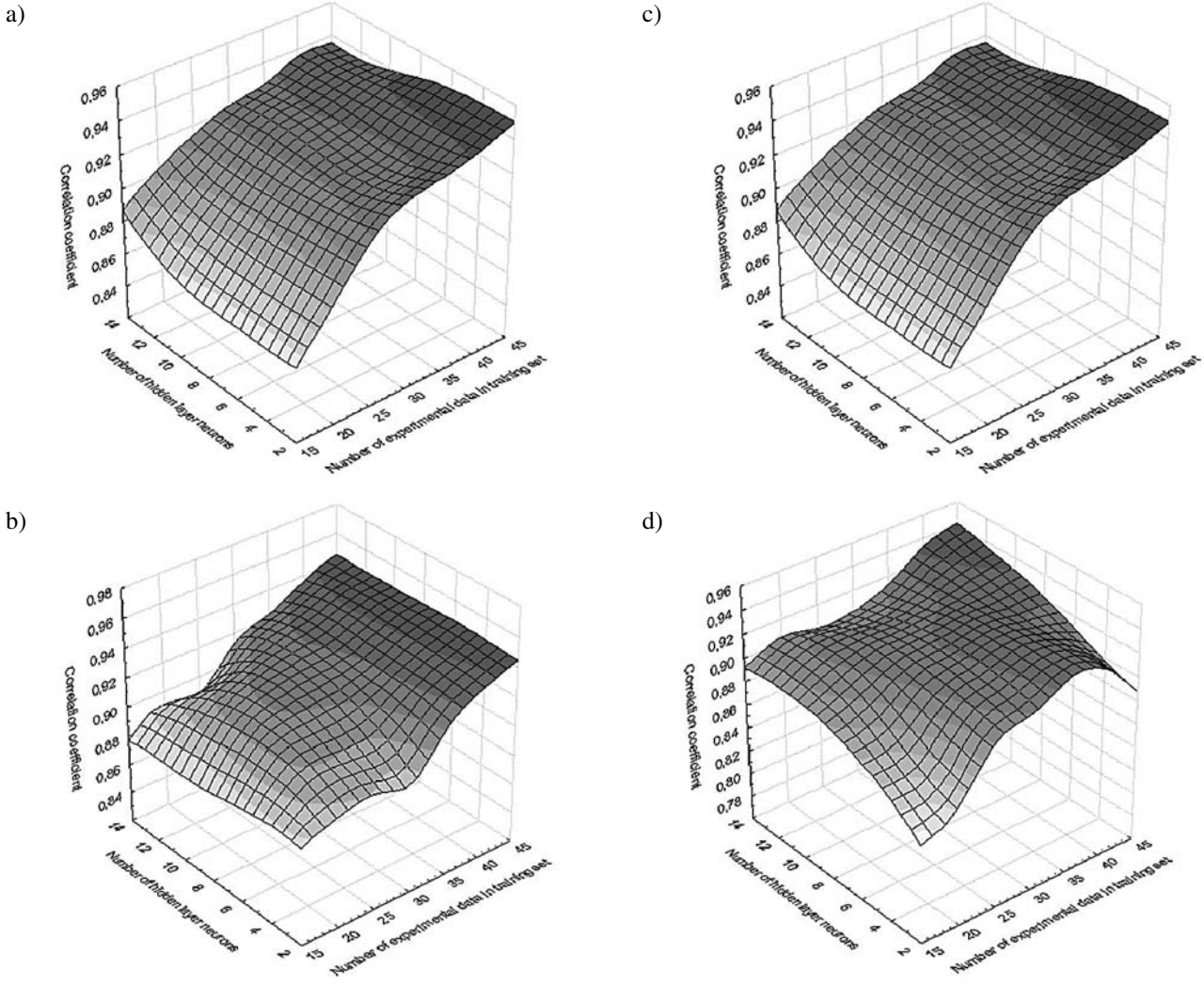

Figure 2. Influence of number of hidden layer neurons and number of experimental data in training set on the correlation coefficient of the multilayer feedforward artificial neural network: (a) conjugate gradient training algorithm using FTIR-ATR spectral input data; (b) quasi Newton training algorithm using FTIR-ATR spectral input data; (c) conjugate gradient training algorithm using FT-Raman spectral input data; (d) quasi Newton training algorithm using FT-Raman spectral input data.

The predictive performance of developed neural network models was tested using Pearson's correlation coefficient *R* (between the predicted and observed output values), average error and average absolute error of the output variable.

All ANN calculations in this work were performed using Statistica 7.1 software (StatSoft Inc., USA).

4. Results and Discussion

The experimental data set was split into three subsets: training set, selection set and validation set. The first set was used to train networks, second one to prevent over-training process, and the third one to validate prediction ability of the developed ANN model. The physicochemical properties of diesel fuel were modelled simultaneously and the presented results describe average values for correlation coefficients (for all modelled properties) based on external validation data set only. In Figure 2, the optimization for MLP models using FTIR-ATR and FT-Raman spectral input data is illustrated, including the effects of the training algorithm, number of hidden layer neurons and number of experimental data points used for training.

As it can be seen, all correlation coefficients between actual and predicted values were acceptable, for both vibrational spectroscopies. When comparing these two spectroscopy techniques, it is clearly visible that FTIR-ATR/MLP model is slightly more accurate than the FT-Raman/MLP model. Also, one can observe that neural network models obtained by using CG training algorithm give more stabile results and slightly higher correlation coefficients than those obtained by using QN algorithm. This indicates CG as highly effective algorithm with very good stability.

However, MLP network that uses FTIR-ATR input spectral data, CG training algorithm, 45 experimental data points in the training set and 8 hidden layer neurons, produces model with maximal correlation coefficient (R = 0.9577; Figure 2a). It is generally preferable to diminish the number of experimental data points used for training in order to reduce the overall experimental effort. Figure 2 shows that the number of experimental data points used for training procedure can be reduced to 30 without significant impact on the model accuracy.

Figure 3 illustrates the results of optimization for RBF models using FTIR-ATR and FT-Raman spectral input data. However, the FTIR-ATR spectral input data gives better and more stabile results compared with FT-Raman input data. It is probably a consequence of a very small number of spectral vibrations of samples in Raman spectra (Figure 1). It can be seen (Figure 3a) that maximal correlation coefficient was obtained for FTIR-ATR input data using K-means radial assignment algorithm in combination with explicit radial spread algorithm. The Figure 3 also presents optimization of number of hidden layer neurons and number of experimental data points needed for the training set. The amount of experimental data used for the training was varied from 20 to 45 and number of hidden layer neurons from 5 to 17. It is also shown that optimal configuration was achieved using maximal number of experimental data in the training set and maximal number of hidden layer neurons.

From Figures 2 and 3 is clearly visible that MLP models are more accurate than RBF ones. Furthermore, use of CG training algorithm gives slightly better results than QN one. In accordance with previous discussion, the performance characteristic of optimal developed ANN models in prediction diesel fuel properties are shown in Table 1.

The optimal model (maximal average correlation coefficient) using FTIR-ATR spectral input data were achieved using MLP network, CG training algorithm, 45 experimental data points in the training set, and 8 hidden layer neurons. Furthermore, optimal models using FT-Raman spectral input data was achieved using MLP neural network, CG training algorithm, 45 experimental data points in the training set, and 12 hidden layer neurons.

As it can be seen (Table 1), the correlation coefficient between actual and predicted values were acceptable for all diesel fuel properties, but when comparing these

Table 1. The performance characteristic of optimal developed artificial neural network models to prediction diesel fuel properties.

	FTIR-ATR				FT-Raman		
Diesel property	**Property range**	**Error mean**	**Abs. error mean**	**Correlation coefficient**	**Error mean**	**Abs. error mean**	**Correlation coefficient**
Cetane number	50.1–55.9	0.0301	0.3228	0.9597	0.0860	0.3669	0.9406
Cetane index	48.0–56.3	0.0708	0.3828	0.9820	–0.0033	0.4304	0.9816
Density (kg/m^3)	827.2–841.3	0.1441	0.7467	0.9315	0.0183	0.9399	0.9092
Viscosity (mm^2/s)	2.24–3.79	0.0159	0.0658	0.9815	0.0019	0.0969	0.9635
Total aromatics (wt %)	23.2–34.3	0.0042	0.5709	0.9662	0.1274	0.6544	0.9644
PAH (wt %)	1.8–6.1	–0.0142	0.2887	0.9343	–0.0411	0.2836	0.9509
T10 (C)	194.6–246.5	0.2400	3.3062	0.9792	0.3183	3.6331	0.9547
T50 (C)	253.9–288.9	0.4412	1.6811	0.9838	0.1047	2.0155	0.9772
T90 (C)	329.1–348.5	–0.1982	1.7768	0.9007	–0.1133	1.8389	0.8909

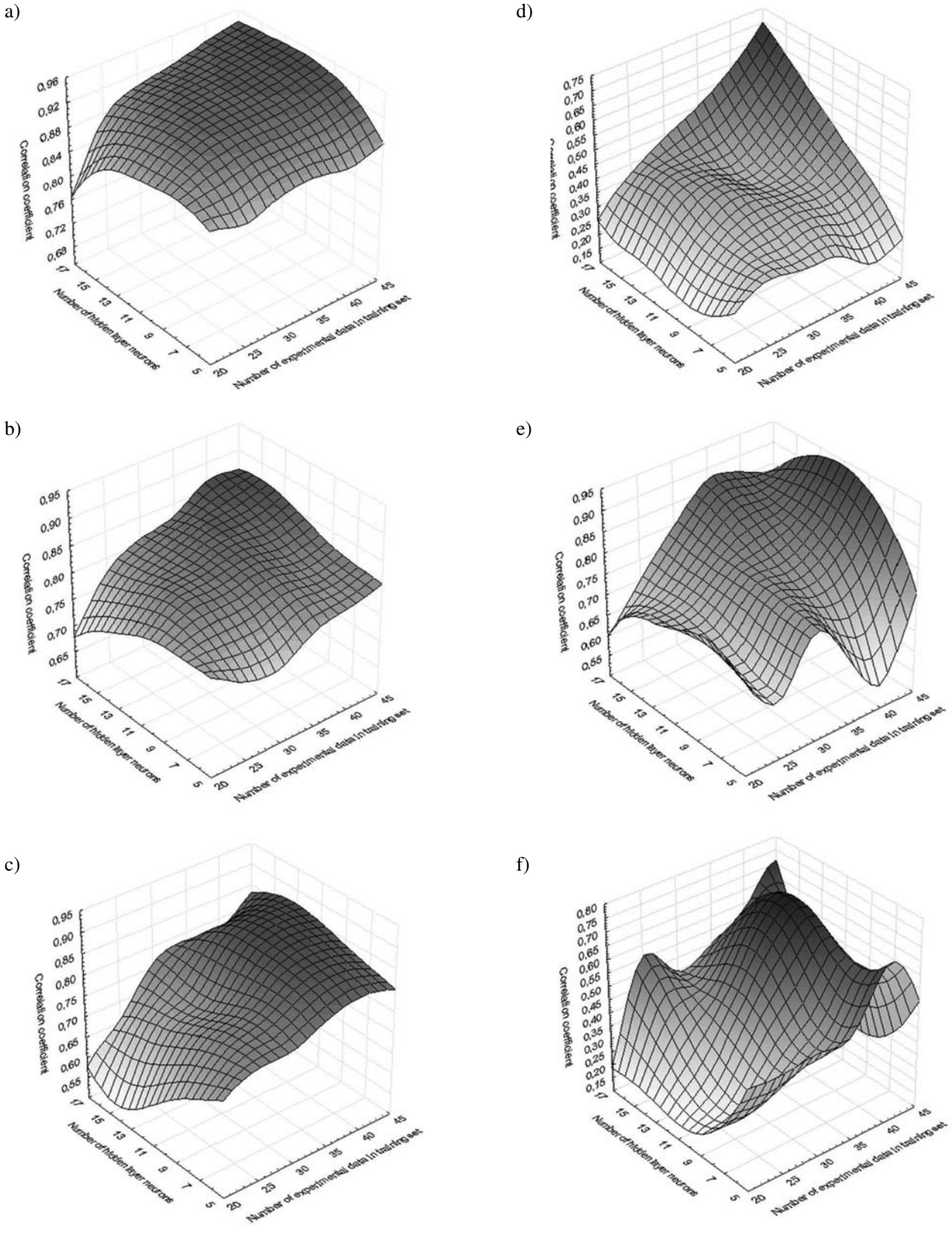

Figure 3. Influence of number of hidden layer neurons and number of experimental data in training set on the correlation coefficient of the radial basis function artificial neural network using K-means radial assignment algorithm and: (a) explicit radial spread algorithm using FTIR-ATR spectral input data; (b) isotropic radial spread algorithm using FTIR-ATR spectral input data; (c) K-nearest neighbour radial spread algorithm using FTIR-ATR spectral input data; (d) explicit radial spread algorithm using FT-Raman spectral input data; (e) isotropic radial spread algorithm using FT-Raman spectral input data; (f) K-nearest neighbour radial spread algorithm using FT-Raman spectral input data.

two spectroscopy techniques, it is clearly visible that model using FTIR-ATR spectral input data is more accurate than model using FT-Raman input data. An exception was the correlation coefficient for polycyclic aromatic hydrocarbons that was slightly better using FT-Raman input data. The cause of this exception should be sought in a separate well-defined maximum at 1378 cm^{-1} in the Raman spectrum that represents ring stretching of bicyclic aromatic fractions in diesel fuel (Figure 1).

The obtained results are basically in agreement with other recent publications (Santos et al. Santos et al.[35]). Although, the performance characteristic seems to be comparable, the approaches based on singe neural network model (presented in this work), predicting simultaneously all required properties include interactions modelling. This could additionally raise the prediction ability.

The accuracy of the obtained models is comparable to the reproducibility values of the standard methods,[33–38] which were used for experimental determination of diesel fuels properties. Despite the relatively high value of correlation coefficient for density (Table 1), the model values are out-of-range due to the very small reproducibility of the standard method.

Comparing the developed ANN models with the standard methods they yielded good predictions and satisfied requirements of reproducibility, except in the determination of density. Density could be determined using those models if less accuracy would be acceptable.

5. Conclusion

This work presents the development of models for prediction of diesel fuels properties using MLP and RBF neural networks. The developed models predict cetane number, cetane index, density, viscosity, T10, T50, T90, the contents of total aromatics and polycyclic aromatic hydrocarbons based on FTIR-ATR and FT-Raman input data. The training algorithms, number of hidden layer neurons and experimental data points used for the training set were optimized for both neural networks in order to insure good predictive ability by reducing unnecessary experimental work.

MLP network using FTIR-ATR spectral data and CG training algorithm, 45 experimental data points in the training set and 8 hidden layer neurons produces model with maximal correlation coefficient. It was found that MLP models are more accurate than the RBF ones and the usage of FTIR-ATR input spectral data gives slightly better results in comparison to FT-Raman ones. Correlation coefficients are ranged between 0.9007–0.9838 and 0.8909–0.9816 for the FTIR-ATR spectral input data, i.e. the FT-Raman spectral input data. Obtained absolute error mean of the neural network models are within range for the reproducibility of standard methods. From these results it can be concluded that developed neural network models can be used for rapid and simultaneous prediction of main diesel fuels properties.

6. References

1. R. M. Balabin, R. Z. Safieva, E. I. Lomakina, *Chemometr. Intell. Lab. Syst*. **2007**, *88*, 183–188.
2. R. M. Balabin, R. Z. Safieva, E. I. Lomakina, *Chemometr. Intell. Lab. Syst.* **2008**, *93*, 58–62.
3. R. M. Balabin, R. Z. Safieva, E. I. Lomakina, *Anal. Chim. Acta*. **2010**, *671*, 27–35.
4. R. M. Balabin, R. Z. Safieva, *Fuel* **2008**, *87*, 1096–1101.
5. R. M. Balabin, R. Z. Safieva, *Fuel* **2008**, *87*, 2745–2752.
6. M. Tariq, S. Ali, F. Ahmad, M. Ahmad, M. Zafar, N. Khalid, M. A. Khan, *Fuel Process. Technol.* **2011**, *92*, 336–341.
7. M. F. Ferrão, M. de Souza Viera, R. E. P. Pazos, D. Fachini, A. E. Gerbase, L. Marder, *Fuel* **2011**, *90*, 701–706.
8. R. M. Balabin, R. Z. Safieva, E. I. Lomakina, *Microchem. J.* **2011**, *98*, 121–128.
9. R. M. Balabin, R. Z. Safieva, E. I. Lomakina, *Fuel* **2011**, *90*, 2007–2015.
10. M. R. Monteiro, A. R. P. Ambrozin, M. da Silva Santos, E. F. Bofo, E. R Pereira-Filho, L. M. Lião, A. G. Ferreira, *Talanta* **2009**, *78*, 660–664.
11. S. L. C. Ferreira, P. A. de Paula Pereira, J. A. Nóbrega, O. Fatibello-Filho, M. A. Feres, B. F. Reis, R. E. Bruns, F. Radler de Aquino Neto, *Anal. Lett.* **2008**, *41*, 1494–1545.
12. R. M. Balabin, R. Z. Safieva, *Anal. Chim. Acta* **2011**, *689*, 190–197.
13. L. A. F. de Godoy, E. C. Ferreira, M. P. Pedroso, C. Henrique de V. Fidélis, F. Augusto, R. J. Poppi, *Anal. Lett.* **2008**, *41*, 1603–1614.
14. M. R. Monteiro, A. R. P. Ambrozin, L. M. Lião, A. G. Ferreira, *Talanta* **2008**, *77*, 593–605.
15. Borin, R. J. Poppi, *Vib. Spectrosc*. **2005**, *37*, 27–32.
16. J. J. Kelly, C.H. Barlow, T.M. Jinguji, J.B. Callis, *Anal. Chem*. **1989**, *61*, 313–320.
17. G. E. Fodor, K.B. Kohl, R.L. Mason, *Anal. Chem*. **1996**, *68*, 23–30.
18. J. B. Cooper, K. L.Wise, J. Groves, W. T. Welch, *Anal. Chem*. **1995**, *67*, 4096–4100.
19. J. B. Cooper, K. L.Wise, W. T. Welch, R. R. Bledsoe, M. B. Sumner, *Appl. Spectrosc*. **1996**, *50*, 917–921.
20. J. B. Cooper, K. L.Wise, W. T. Welch, M. B. Sumner, B. K. Wilt, R. R. Bledsoe, *Appl. Spectrosc*. **1997**, 51, 1613–1620.
21. G. E. Fodor, R. A. Mason, S. A. Hutzler, *Appl. Spectrosc*. **1999**, *53*, 1292–1298.
22. F. B. Gonzaga, C. Pasquini, *Anal. Chim. Acta* **2010**, *670*, 92–97.
23. L. F. B. de Lira, F. V. C. Vasconcelos, C. F. Pereira, A. P. S. Paim, L. Stragevitch, M. F. Pimentel, *Fuel* **2010**, *89*, 405–409.
24. T. Bolanča, Š. Cerjan-Stefanović, M. Regelja, H. Regelja, S. Lončarić, *J. Chromatogr. A*. **2005**, *1085*, 74–85.

25. T. Bolanča, Š. Cerjan-Stefanović, Š. Ukić, M. Rogošić, M. Luša, *J. Chemometrics* **2008**, *22*, 106–113.
26. H. Sarimveis, A. Alexandridis, G. Bafas, *Neurocomputing* **2003**, *51*, 501–505.
27. N. Pasadakis, V. Gaganis, Ch. Foteinopoulos, *Fuel Process. Technol.* **2006,** *87*, 505–509.
28. L. C. Côcco, C. I. Yamamoto, O. F. von Meien, *Chemom. Intell. Lab. Syst*. **2005,** *76*, 55–63.
29. H. Yang, C. Fairbridge, Z. Ring, *Petrol. Sci. Technol.* **2001,** *19,* **573–586.**
30. F. C. C. Oliveira, C. R. R. Brandão, H. F. Ramalho, L. A. F. da Costa, P. A. Z. Suarez, J. C. Rubim, *Anal. Chim. Acta* **2007**, *587*, 194–199.
31. H. Yang, Z. Ring, Y. Briker, N. McLean, W. Friesen, C. Fairbridge, *Fuel* **2002**, *81*, 65–74.
32. D. M. Korres, G. Anastopoulos, E. Lois, A. Alexandridis, H. Sarimveis, G. Bafas, *Fuel* **2002**, *81*, 1243–1250.
33. Ch. Wu, J. Zhang, W. Li, Y. Wang, H. Cao, *Fuel Process. Technol.* **2006**, *87*, 585–590.
34. J. M. Andrade, S. Garrigues, M. de la Guardia, M. Gómez-Carracedo, D. Prada, *Anal. Chim. Acta* **2003**, *482*, 115–128.
35. V. O. Santos Jr., F. C. C. Oliveira, D. L. Lima, A. C. Petry, E. Garcia, P. A. Z. Suarez, J. C. Rubim, *Anal. Chim. Acta,* **2005**, *547*, 188–196.
36. P. Melin, O. Castillo, *Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing*, Springer, Verlag Berlin Heidelberg, **2005**.
37. T. Hill, P. Lewicki, *Statistics: Methods and Applications*, StatSoft Inc., Tulsa, **2005**.
38. C. Satchwell, *Pattern Recognition and Traiding Decisions*, McGraw Hill, New York, **2005**.
39. D. Graupe, *Principles of Artificial Neural Networks*, 2nd ed., World Scientific, New Jersey, **2007**.
40. C. Bishop, *Neural Networks for Pattern Recognition*, University Press, Oxford, **1995**.
41. A. J. Shepard, Second-Order Methods for Neural Networks, Springer, New York, **1997**.
42. S. Haykin: *Neural Networks: A Comprehensive Foundation*, Macmillan Publishing, New York, **1994**.
43. T. Bolanča, Š. Cerjan-Stefanović, M. Luša, H. Regelja, S. Lončarić, *Chemom. Intell. Lab. Syst.* **2007**, *86*, 95–101.
44. EN 14275-03, *Automotive fuels – Assessment of petrol and diesel fuel quality – Sampling from retail site pumps and commercial site fuel dispensers*. European Committee for Standardization (CEN), **2003**.
45. *PetroSpec Calibration Update Software*, TX 77064 USA Houston. Petroleum Analyzer Company L. P., **2001**.
46. ASTM D 4737-10, *Standard Test Method for Calculate Cetane Index by Four Variable Equation*. American Society for Testing Materials, 2010.
47. ASTM D 1298-99, *Standard Practice for Density, Relative Density (Specific Gravity), or API Gravity of Crude Petroleum and Liquid Petroleum Product by Hydrometer Method.* American Society for Testing Materials, 1999.
48. ASTM D 445-06, *Standard Test Method for Kinematic Viscosity of Transparent and Opaque Liquids (the Calculation of Dynamic Viscosity).* American Society for Testing Materials, 2006.
49. ASTM D 86-05, *Standard Test Method for Distillation of Petroleum Products at Atmospheric Pressure*. American Society for Testing Materials, 2005.
50. EN 12916-06, *Petroleum products – Determination of aromatic hydrocarbon types in middle distillates – High performance liquid chromatography method with refractive index detection.* European Committee for Standardization (CEN), 2006.
51. G. Srečnik, Ž. Debeljak, Š. Cerjan-Stefanović, M. Novič T. Bolanča, *J. Chromatogr. A* **2002**, *973*, 47–59.

Povzetek

Opisan je razvoj modelov za usklajevanje in napoved lastnosti dizelskega goriva iz absorbanc IR in Ramanskih intenzitet kot vhodnih spremenljivk za umetne nevronske mreže. Večnivojska nevronska mreža s progresivnim tokom podatkov ter radialno funkcijska nevronska mreža sta bili uporabljeni za napoved cetanskega števila, cetanskega indeksa, gostote, viskoznosti, destilacijske temperature pri 10 % (T10), 50 % (T50) in 90 % (T90) obnovljivosti, vsebnost aromatov in policikličnih aromatičnih ogljikovodikov v komercialnih dizelskih gorivih.

V tej študiji smo uporabili dvostopenjski učni algoritem. Prva stopnja je temelji na običajnem vzvratnem širjenju napake, druga stopnja pa bodisi na konjugirani gradientni ali na kvazi Newtonovi metodi. Trije različni algoritmi, uporabljeni za določitev radialno bazne funkcije, so: eksplicitni, izotropni in K-najbližnje-sosedni.

Število nevronov v skritem nivoju in število učnih vzorcev smo oprimizirali pri vsaki uporabljeni metodi, tako da smo zagotovili dobro napovedno zmogljivost ob zmanjšani potrebi po eksperimentalnem delu.

Pokazali smo, da modeli na osnovi nevronskih mrež lahko določajo glavne lastnosti dizelskega goriva simultano, na osnovi ene same hitre meritve srednjega IR ali Ramenskega spektra.