EasyChair Preprint

№ 7943

Modeling of Oxidative Desulfurization Process by Artificial Neural Network

Mehdi Gheisari

May 14, 2022

Modeling of Oxidative Desulfurization process by Artificial Neural Network

Abstract

In recent years Oxidative desulfurization process, having significant advantage against other well-known desulfurization process, have received considerable attention. In this study, modeling of Oxidative desulfurization of fuel oil was investigated using artificial neural network (ANN). It is found that ANN provides a useful method for developing nonlinear relations between variables. To determine effective parameters on ODS process; a principal component analysis was performed on data. The results showed that oxidant quantity, contact time and reactor temperature play important roles in determination of desulfurization performance. An artificial neural network, using back propagation (BP), was also utilized for modeling oxidative desulfuration process of fuel oil. Different structures were tried with several neurons in the hidden layer and the total error was calculated. Finally, eight hidden neurons were applied. The comparison between the outputs of ANN modeling being referred as BP-NN 5:8:1 and the experimental data showed satisfactory agreement.

Keywords: Artificial neural network, Modeling, desulfurization

1. Introduction

Petroleum derived fuels contain various organic sulfur compounds, which are usually divided into two main classes: the easy sulfur compounds and the refractory sulfur compounds. The first group is non-thiophenic sulfur compounds, thiophenes, benzothiophenes and non-beta dibezothiophenes. The refractory sulfur compounds include the benzothiophene and dibenzothiophenes and their mono-, di- and tri substituted homologues with alkyl group containing from one to 12 carbones. Sulfur compounds cause many problems such as catalyst deactivation, gum formation and corrosion. Sulfur oxides also contribute to environmental pollution problems like acid rain. In recent years to control air pollution due to fuel combustion, the US EPA has restricted sulfur levels in fuels requiring the use of low-sulfur fuels. Usually, hydrodesulfurization technique (HDS) is adapted as a method of removing sulfur compounds from petroleum fractions. According to this technology, sulfur is reacted with hydrogen gas in the presence of a catalyst under drastic conditions involving a high temperature and a high pressure to be turned into toxic hydrogen sulfide. Although HDS is inexpensive and effective in removing easy sulfur compounds, but it is rather inefficient in removing refractory sulfur compounds [1,2]. In order to produce ultra-low sulfur fuels, deep HDS techniques should be adopted. These techniques require HDS to be operated under more severe conditions, including the use of higher temperature, higher hydrogen pressure, more active catalysts, and longer residence time. So it is essential that a method, which can be operated under moderate conditions and has high efficiency in removing all kinds of sulfur compounds, be developed to produce ultra low sulfur products.

Various chemical processes for thoroughly removing sulfur compounds have been investigated in the past [3-5]. One idea that has drawn wide attention, as referred oxidative desulfurization (ODS), involves oxidizing the sulfur compounds then removing oxidized ones [6-8]. During the oxidation process the nucleophilic sulfur atom can be oxidized by the electrophilic addition reaction of oxygen atoms to sulfoxides (eq. 1) and sulfones (eq. 2) (containing common S=O units). The general form of the reaction may be written as:

R2S + [o] → R2SO Eq. 1

R2S + 2[o] → R2SO2 Eq. 2

Many papers and patents concerning oxidation of sulfur compounds have been published and claimed, including the use of various oxidants like Hydrogen peroxide [9-11], formic acid/hydrogen peroxide [12-15], acetic acid/ hydrogen peroxide [16-18], polyoxometalates/hydrogen peroxide [19,20] and nitrogen oxide [13]. However extraction was widely used to follow oxidation step [21,22], Funakoshi and et al. [23] have reported a process involves formic acid/H2O2 as oxidant, followed by distillation in high temperature like the boiling point range of examined gas oil. They have not implied the influence of using high temperature on properties of gas oil.

Little work has been conducted on the modeling of oxidative desulfurization process [8,24]. Tam and et al. [24] have presented a mathematical kinetic model to describe the kinetics of sulfur removal in the oxidation of Arabian atmospheric gas oil (AGO) using a CSTR. As a matter of fact, Complexity, nonlinearity, and multivariable nature of chemical process modeling often make mathematical models very difficult or even impossible to develop. In recent years, the use of artificial neural networks (ANNs) for fitting complex kinetic data has been considered as a non-linear approach for the process design improvement or optimization, being performed the comparison with the mechanistic kinetic modeling. ANNs are computer-based systems that are designed to simulate the learning process of neurons in the human brain. Artificial neural networks are being extensively applied in many fields of science and engineering. Some applications of ANNs in chemical engineering are: process dynamics, modeling of processes [25], optimization of industrial chemical process [26], and modeling of chemical reactors [27]. To our knowledge, there is no report for application of ANNs for modeling of oxidative desulfurization process.

The present invention evaluate the efficiency of ANNs in modeling of ODS process using acetic acid/hydrogen peroxide system as the oxidant agent and distillation in the presence of water as separation mode under relatively low temperature.

2. Experimental methods

2.1. Materials

Fuel oil that was used for experiments was derived from Pyrolysis Gasoline of Tabriz Petrochemical Complex. The properties of fuel oil along with their related standard methods that were used to determine them are shown in table 1.

Table 1. Properties of fuel oil.

Property	Result	Method
Specific gravity	0.94	ASTM D1298-85

(g/cm2)		
H_2S	Negative	ASTM D853-91
Doctor test	Negative	ASTM D4952-94
Total sulfur (ppm)	280	BPC-216
SO_2	Negative	ASTM D853-91
Distillation range (°C)		ASTM D86-91
10%	174.2	
50%	186.1	
90%	209.1	

All of the reagents that were used in experiments were obtained from Merck (Germany), whereas in order to avoid the dangers connected with handling a higher purity of hydrogen peroxide like 50 or 70 wt%, the employed one had 30% of the purity.

2.2. Analysis

A total sulfur analyzer manufactured by Dohrman Co. was employed to determine any sample. A Chrompack gas chromatograph with flam ionization detector and CP-Sil-5CB column was used to analyze fuel oil. Separation conditions employed for the compounds reported in this paper were: column temperature program 60°C (3min), 5°C/min to 90°C, 8°C/min to 150°C, 15°C/min to 250°C and stay at 250°C for 36min; injector spilter, 280°C; detection FID, 300°C. Infrared spectroscopy analyziz of samples were carried out by using Shimatzo 405. The ANNs program was written in MATLAB running on a PC computer. PCA calculations used a singular-value decomposition algorithm and were also performed with the MATLAB software version 6.5. The operating system was Microsoft Windows 2000XP.

2.3. A typical process runs

In a typical run, the oxidation of sulfur compounds was performed by in situ generated acetic acid/H2O2 in the following way: definite moles of carboxylic acid and hydrogen peroxide, 0.2 ml of a mineral acid and 100 ml fuel oil were charged into the glass reactor and the mixture were stirred. Agitation of the reaction mixture at 125rpm provided uniform mixing of reactants. Reaction temperatures were maintained at definite temperature by immersing the bottom half of the reactor into a water bath. After the reaction time, the mixture was washed successively with water, 5% aqueous sodium bicarbonate and water, respectively. The resultant was distilled in the presence of an equal volume of water. During the distillation step the temperature stayed in the range about 95-96°C. The distillate of mixture formed two separated layers, aqueous layer (bottom) and fuel oil layer (top), which were separated by using a separatory funnel. Samples were finally subjected to total sulfur analysis. To investigate the process variables, the one at a time experimental design was used. The effects of amount of used H2O2, acetic acid, H2SO4, reactor temperature and reaction time process variables on desulfurization performance were investigated.

3. Results and discussion

3.1. Determination of effective parameters on ODS process

To determine the effective parameters on ODS process performance a principal component analysis was performed. The principal component analysis (PCA) method is widely used to select the interesting variables. PCA contains an orthogonalization procedure such as singular-value decomposition (SVD) that decomposes the primary data matrix by projecting the multi-dimensional data-set onto a new coordinates base formed by the orthogonal directions with data maximum variance. The data matrix consists of a number of experiments, each consisting of a number of variables. The eigenvectors of the data matrix are called principal components (PCs) and they are uncorrelated among them. The effect of each variable on process can be expressed as the linear combination of eigenvectors. Generally, the k-th Principal component, PCk, is a linear combination of the n variable (Xn,j) for the parameter under study and the coefficients an,k are called correlation coefficient (Eq. 3).

$$P\,C_k = \sum_{n=1}^{n} a_{n,k} X_{n,j} \qquad \text{Eq. 3}$$

The magnitude of each eigenvector is expressed by its own eigenvalue, which gives a measure of the variance related to that principal component. The main feature of PCA analysis is the coordinates of the data in the new base (score plot) and the contribution to each variable (correlation coefficient plot). The score plot is usually used for studying the classification of the data clusters; while the loads plot can be used for giving information on the relative importance of the variable to each principal component and their mutual correlation. In this study, the correlation coefficients were used to estimate the affect of considered parameters on ODS process, and subsequently the score's of effective PCs were used as input for modeling of ODS process. Table 2 shows eigenvalues and percentage variance for the principal component extracted. As it shown, all PCs keep almost the same amount of information in, so it is necessary to consider all PCs to determine the effective parameters. The correlation coefficient matrix of all factors with respect to the examined parameters including amount of used H2O2, acetic acid, H2SO4, reactor temperature and reaction time is shown in Table 3. As most parameters have a high correlation coefficient, which have correlation coefficients of more than 0.5 with the PCs were chosen.

Table 2. Percentage variance captured by the PCA analysis performed on the data associated to the effective parameters on ODS process.

PCs	Eigenvalue	variance (%)	cumulative variance (%)
1	6.1349	26.111	26.111
2	5.0092	21.321	47.432
3	4.9227	20.952	68.384
4	4.1888	17.828	86.212
5	3.2393	13.787	99.999

Table 3. The summary of the correlation coefficient of the PCs of all factors. 1) H2O2, 2) Acetic acid, 3) H2SO4, 4) Reactor temperature, 5) Reaction time.

Variables	PC1	PC2	PC3	PC4	PC5
1	-0.28303	0.73451	-0.03437	-0.58822	-0.18224
2	-0.16425	-0.51671	-0.67271	-0.485	-0.13515
3	-0.67775	-0.03355	0.002903	0.057171	0.73229
4	-0.63973	-0.05055	0.004677	0.43921	-0.6287
5	-0.15595	-0.4357	0.73908	-0.47181	-0.13039

3.2. Neural network model

Artificial neural networks (ANNs) represent a powerful tool for developing nonlinear relations

between variables. The advantage of the ANNs approach over conventional approaches is that the problem is directly modeled and has tolerance for even noisy data. ANNs possess the ability to 'learn' what happens in the process without actually modeling the physical and chemical laws that govern the system. An artificial neural network consists of interconnected layers of nonlinear processing elements, which are commonly referred to as neurons, as they resemble biological neurons and information flow channels between the neurons, usually denominated interconnects. Each processing neuron calculates the weighted sum of all interconnected signals from the previous layer plus a bias term, and then generates an output through its activation function. The most widely used networks are made up of three layers: the input, hidden, and output layers. The input layer acts as an input data holder, which distributes inputs into the network. The input data are propagated through the network via interconnections to processing elements in hidden layers where they are combined and modified by the activation function until the output signal can be obtained from the output layer. The feedforward neural networks, based on application of artificial neurons with a sigmoidal activation function, are usually employed for modeling and prediction. The back-propagation (BP) error algorithm, which is an iterative supervised-learning technique, is often used as error algorithm. In each epoch, the entire training set is presented to the network, case by case; errors are calculated and used to adjust the weights in the network using sigmoid transfer function. Because of the use of sigmoid functions in the ANN model, the values of the data variables must be normalized into range [0,1] before applying the ANN methodology. Of course, to calculate training and test error, all of outputs are performed an inverse range scaling to return the predicted response to their original value and computed them with experimental responses. Learning the neural network is an important step for developing a useful network. The network learns the trends contained in the data set and correlates the inputs and the outputs by finding the optimum set of weights that minimize the differences between the predicted and actual output values. The learning continues until the error between prediction and the actual data is minimized. The ability to approximate a complicated function and information processing of the neural network completely depends on the weight of the link between the neurons. The training algorithms may be prone to the problem of overlearning, in which the network learns to memorize and produce the desired outputs given in the training set but fails to generalize to new data. The performance of the network must be test with test set of data.

For the purpose of the present application, a feed forward back-propagation (BP) neural network structure has been trained with a Levenberg-Marquardt backpropagation algorithm. Table 4 shows the properties of designed network. In this study, the linear function (yi= Ii) is used in the output layer and the sigmoid nonlinear function is used in the hidden layers:

$$y_i = f(I_i) = \frac{1}{1+\exp(-I_i)} \qquad \text{Eq. 3}$$

So all inputs were scaled in the range [0,1] by using the following transformation before feeding in to the network:

$$P = \frac{x_i - \bar{x}_i}{s_{x_i}} \qquad \text{Eq. 4}$$

where, P is scaled data, is are mean and standard deviation of x . In this x and $_{x_i}$ I way, the output data were returned to their original value through the following scaling:

$$x_i = s_{x_i}.P + \bar{x}_i \qquad \text{Eq. 5}$$

The number of the input layer was equal to the number of effective parameters, and the number of the output neurons was equal to the performance of ODS described as total sulfur of fuel oil after ODS process. But in the case of the number of hidden neurons, the estimation of the optimum number of hidden layer was necessary. The optimum structure of the neural network

can be determined by a trial and error method. Different structures having 1 to 10 neurons in the hidden layer were tried. Mean square error (MSE) was considered to calculate errors according to following equation:

$$MSE = {}^{1}/_{2}\sum_{i}(T_i - O_i)^2$$

Where, Ti and Oi stand for the output and target value, respectively. Figure 1 shows the obtained errors by applying various neurons in hidden layer. As it is shown, with increasing number of the neurons in the hidden layer beyond 7, the error of the estimation didn't decreased considerably. On the other hand it was found that the ANN with 8 neurons in the hidden layer need the least epoch to show its best performance, so the ANN with 8 neurons was chosen as the optimum structure and used in the rest of modeling process. Subsequently an artificial neural network with 5 input neurons in the input layer, 1 output neuron in the output layer and 8 hidden neurons in the hidden layer was developed to model the oxidative desulfurization process being referred as back-propagation neural network (BP-NN) 5:8:1. A schematic diagram of the feed forward neural network developed for modeling oxidative desulfurization, composed with neurons arranged in layers, is shown in Figure 2. The elements of the input layer were the variables: amount of used H2O2, acetic acid, H2SO4, reactor temperature and reaction time. The output layer corresponded to the sulfur content of fuel oil after oxidative desulfurization.

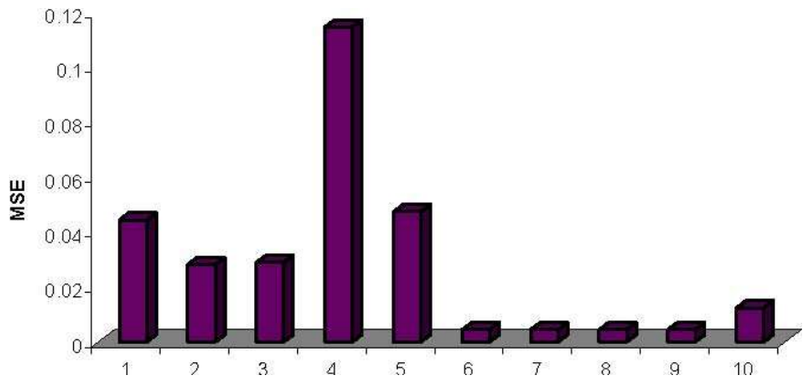

Figure 1. Determination of optimum number of neurons in hidden layer.

The data preparation program assigns data in sequence to the learning set or to the test set according to each variable range and the learning: test set data ratio of 3:1.

The learning would stop when the mean square error of the system is less than 0.00001. All adjusted parameters for designed ANN-BP are listed in table 4.

Table 4. Adjusted parameters for designed ANN-BP

Maximum number of epochs to train	100
Performance goal	10^{-5}
Maximum validation failures	5
Factor to use for memory/speed	1
Minimum performance gradient	10^{-10}
Initial adaptive value (N)	0.001
Decrease factor of N	0.1
Increase factor of N	10
Maximum N	10^{+10}
Maximum time to train in seconds	Infinitive

The results obtained after learning the ANN model were compared with the experimental values of sulfur content of fuel oil. The reliability of the predicted sulfur content obtained by this prediction was evaluated using a randomly generated test set.

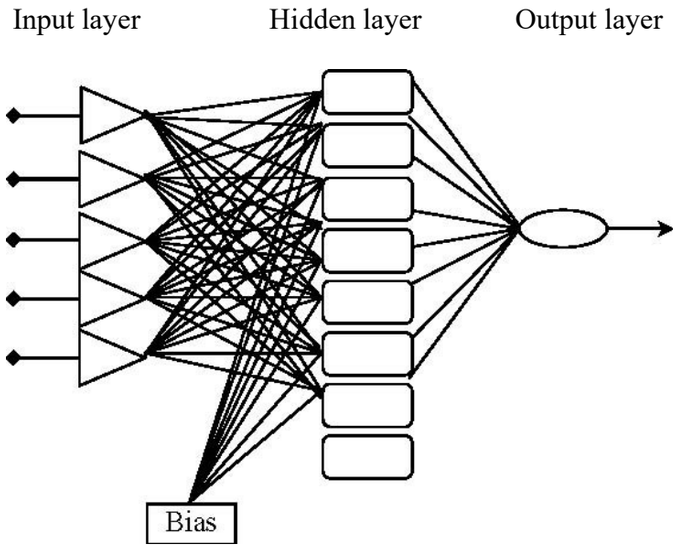

Figure 2. Scheme of the neural network architecture BP-NN 5:8:1.

The correlation between the experimental and the predicted values of sulfur content of fuel oil are shown in the graph, which can give an overview of the actual output and the predicted output using the optimized BP-ANN (Figure 3). As it can be seen two lines were used to show the success of the prediction. The one is the perfect fit (calculated data equal to experimental data), on which all the data of an ideal model should lay. The other line is the line that best fits on the data of the scatter plot with equation Y=ax+b. The results obtained are excellent and very similar to the actual figures. The correlation coefficients of the fitting are 0.995, 0.938 for training and test sets, respectively that are within a quite acceptable limit. The results confirmed that the neural network simulator developed in this work could illustrate the ability of modeling the oxidative desulfurization process.

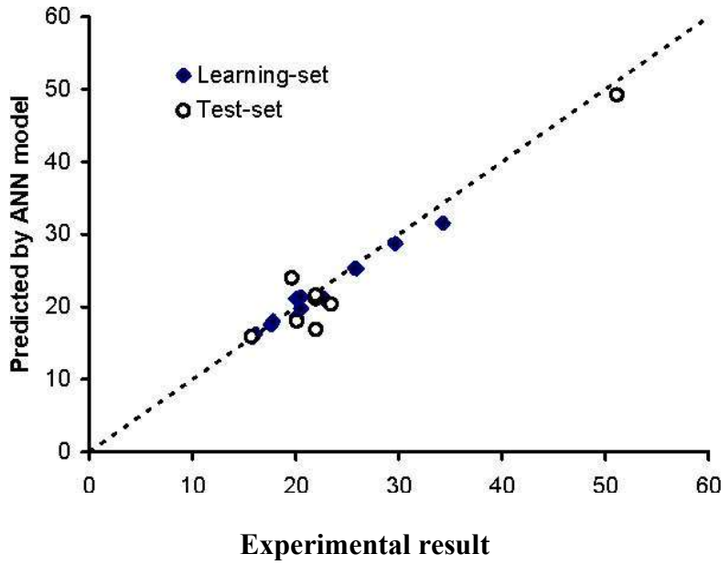

Figure 3. Comparison of ANN model prediction with experimental data.

4. Conclusion

The attempt to model oxidative desulfurization process (ODS) showed that:

1) It is very useful to apply nonlinearity approach for modeling chemical process.
2) Principal component analysis (PCA) can be used to select effective parameters on desired process.
3) Comparing the variances showed that all PCs remained the same variance and all of them must be considered.
4) Using the value of correlation coefficient of five PCs, It was concluded that the amount of used H2O2, acetic acid, H2SO4, reactor temperature and reaction time are effective parameters on performance of ODS process.
5) Back-propagation neural network (BP-NN) was applied to model ODs process.
6) Errors obtained by applying various neurons (1-10) in hidden layer showed that the structure with 8 neurons in hidden layer had the best performance.
7) BP-NN 5:8:1 was used for modeling of ODS process.
8) The correlation coefficients of the fitting of experimental and predicted value by designed BP-NN were 0.995, 0.938 for training and test sets, respectively.
9) The results showed that the BP-NN 5:8:1 has very good ability to model ODS process.

References:

[1] C. Song, X. Ma, Applied Catalysis B: Environmental 41 (2003) 207.
[2] W.C. Baird, G.B. Mcvicker, J.J. Schorfheide, D.P. Klein, S.S. Hantzer, M. Daage, M.S. Touvelle, E.S. Ellis, D.E.W. Vaughan, J. Chen, US patent 6,245,221, 2001.
[3] K.A. Gray, G.T. Mrachko, C.H. Squires, Current Opinion in Microbiology 6 (2003) 229.
[4] C. Huang, B. Chen, J. Zhang, Z. Liu, Y. Li, Energy & Fuels 18 (2004) 1862.
[5] A.B. Martin, A. Alcon, V.E. Santos, F. Garcia-Ochoa, Energy & Fuels 18 (2004) 851.
[6] A. Attar, W.H. Corcoran, Ind. Eng. Chem. Prod. Res. Dev. 17 (1978) 103.
[7] K. Yazu, Y. Yamamoto, T. Furuya, K. Miki, K. Ukegawa, Energy & Fuels 15 (2001) 1535.
[8] D. Wang, E.W. Qian, H. Amano, K. Okata, A. Ishihara, T. Kabe, Applied Catalysis A: General 253 (2003) 91.
[9] J.A. Kocal, US patent 6,277,271, 2001.
[10] A.V. Anisimov, E.V. Fedorova, A.Z. Lesnugin, V.M. Senyavin, L.A. Aslanov, V.B. Rybakov, A.V. Tarakanova, Catalysis Today 78 (2003) 319.
[11] D.L. Yu, J.B. Gomm, Neural Comput & Applic 10 (2002) 289.
[12] P.D. Filippis, M. Scarsella, Energy & Fuels 17 (2003) 1452.
[13] P.S. Tam, J.R. Kittrell, J.W. Eldridge, Ind. Eng. Chem. Res. 29 (1990) 321.
[14] A.S. Rappas, V.P. Nero, S.J. Decanio, US patent 6,406,616, 2002.
[15] C.A. Audeh, US patent 5,310,479, 1994.
[16] Y. Shiraishi, T. Hirai, Energy & Fuels 18 (2004) 37.
[17] O. Sonmez , E.S. Giray, Fuel Process. Technol. 70 (2001) 159.
[18] B. STOVE, G. WALTER, S. EBBER, D. GEOFFREYE, US patent 6,596,914.
[19] M. Te, C. Fairbridge, Z. Ring, Applied Catalysis A: General 219 (2001) 267.
[20] S. Murata, K. Murata, K. Kidena, M. Nomura, Energy & Fuels 18 (2004) 116.
[21] H. Mei, B.W. Mei, T.F. Yen, Fuel 82 (2003) 405.
[22] F. Zannikos, E. Lois, S. Stournas, Fuel Process. Technol. 42 (1995) 35.
[23] I. Funakoshi, T. Aida, EP 565,324, 1993.
[24] P.S. Tam, J.R. Kittrell, J.W. Eldridge, Ind. Eng. Chem. Res. 29 (1990) 324.
[25] D.L. Yu, J.B. Gomm, Neural Comput & Applic 10 (2002) 289.
[26] C.A.O. Nascimento, R. Giudici, R. Guardani, Computers and Chemical Engineering 24 (2000) 2303.
[27] E.J. Molga, Chemical Engineering and Processing 42 (2003) 675.