Petroleum and Coal

Article **Open Access**

MODELING OF A COMMERCIAL SCALE THERMAL CRACKING PLANT USING GMDH NEURAL NETWORK

Reza Seif Mohaddecy[1], Sepehr Sadighi 'Ershad Amini[2]

1 Research Institute of Petroleum Industry (RIPI), Catalysis Research Division, P.O. Box14665137, Tehran, Iran
2 School of Chemical Engineering, College of Engineering, University of Tehran

Received December 25, 2016; Accepted April 5, 2017

Abstract

In this research, the group method of data handling (GMDH) networks is applied for estimating the significant process variables of a commercial scale thermal cracking (visbreaking) unit. This model can predict the fuel oil viscosity, severity and yield of products i.e. fuel oil, gasoline and light gasses by using a grand polynomial correlation which is a function of flow rate, viscosity and initial boiling point of feed, furnace outlet temperature, and steam flow rate. To this fifty data points research, are obtained from the target plant during a life cycle about 422 days. Then, the GMDH network uses 70% of these data points for self-training while using the remained ones for the validation step. The results show that the developed network can precisely estimate the fuel oil product properties during the life of the process. Moreover, it is confirmed that the proposed model is capable of predicting severity, gas, gasoline and fuel oil product yields, and fuel oil viscosity with the average absolute deviation (AAD%) of 0.462%, 0.448%, 0.458%, 0.051%, and 0.016%, respectively. Moreover, the root mean square error (RMSE %) of the mentioned parameters are 0.717%, 0.709%, 0.716%, 0.078% and 0.025%, respectively.

Keywords: Group method of data handling; Neural network; Visbreaking; Fuel oil; modeling.

1. Introduction

Visbreaking is a mild liquid phase pyrolysis of atmospheric or vacuum bottoms of crude oil distillation. The aim of this process is to reduce the viscosity of the residue and to significantly increase the production of lighter distillates, such as light gas, gasoline, and kerosene. The process with the name of "visbreaker" refers to the fact that the process reduces (i.e. breaks) the viscosity of the residual oil [1-2].

There are two types of visbreaking technology that are commercially available: the 'coil' or 'furnace' and the 'soaker' processes. In the coil process, conversion is achieved by high temperature cracking for a predetermined and relatively short period in the heater. In the soaker process, with low temperature and high residence time, the majority of conversion occurs in a reaction vessel or soaker drum where the two-phase heated effluent is held at a lower temperature for a longer period. Therefore, its heater duty, and in turn its fuel consum-ption, is only 70% of that for the coil-visbreaking process [3].

The kinetic modeling of visbreaking unit is limited by the complexity of the process. There is a large gap between fundamental studies and practical kinetic model reactions [4-7]. In this respect, artificial neural networks (ANNs) may constitute a powerful approach to develop estimators that can be used for on-line applications [8-10]. Moreover, as a data-based modeling approach, they are widely applied in process modeling and control which are even capable of modeling non-linear process [11-12]. However, ANN's structure contains a massive complicated of equations within its nodes and layers. Furthermore, the arrangement of the network is chosen manually or randomly which does not assure the best possible network.

On the contrary, the group method of data handling (GMDH) provides a self-organizing neural network to express the genome of the system. To do such a task, it uses the most suitable configuration by applying minimization process. In the other word, the GMDH utilizes feed-forward network whose coefficients are determined by using regression with the imitation of self-organizing activity. The algorithm chooses the most suitable polynomial expressions built by a combination of two independent variables at a time.

Up to now, ANNs have been utilized to monitor and model the performance of many refinery and petrochemical processes. However, there are few studies in which GMDH network can be applied for modeling such processes while no work is reported for modeling visbreaking unit by using this approach [13-15]. Therefore, the aim of the present study is to investigate the capability of GMDH network to predict the yield of visbreaking products, process severity, and fuel oil viscosity for a commercial scale visbreaking plant.

2. Process description

A commercial soaker-visbreaker unit was chosen as a case study. This unit was designed to process about 20,000 barrels per day of a mixture including vacuum residuum and slop vacuum gas oil; both are taken from a vacuum tower. The specifications of the feed are presented in Table 1 which can vary slightly with time from start of the run (SOR) to end of the run (EOR). As shown in Fig. 1, the fresh feed is charged to the coil furnace at the temperature about 340°C. This furnace is constructed from two sections, fired independently. After the coil furnace, two hot streams are drained into a transfer line, and the mixed product is entered into the soaker drum.

Table 1. Feed characterization

Property	Unit	Value
Specific gravity	-	1.006
Sulfur content	wt %	3.19
V + Ni content	wt ppm	188
Distillation (ASTM D1160)		
Vol %	Temperature (°C)	
IBP	303	
5	409	
10	457	
20	503	
30	543	
50	585	

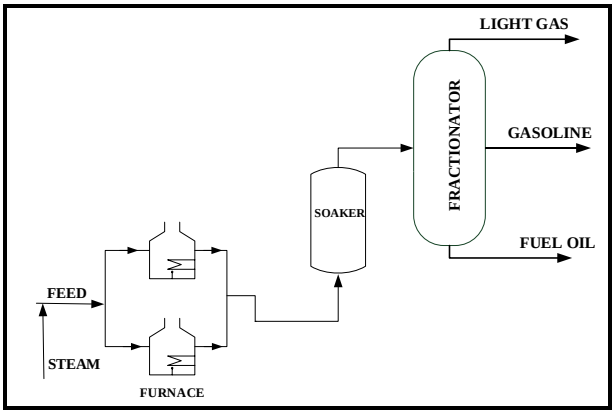

Figure 1. Scheme of the visbreaking plant

The specifications of coil and soaker drum are presented in Table 2. Then, the outlet stream from the soaker drum is quenched using the cold recycle stream to stop cracking reactions and to inhibit the coke formation. Finally, the combined stream is transferred to the fractionation tower and side strippers.

Table 2. Specifications of the coil and soaker of the visbreaking unit

		Coil specification
Number of tubes	-	128
Number of convection tubes	-	76
Number of radiation tubes	-	52
Tube length	m	18.745
Outside diameter	m	0.114
		Soaker specification
Outside diameter	m	2.405
Length	m	16.5

3. GMDH network modeling of experimental data

The basic structure of the brain has been widely employed in various fields such as modeling, control, and pattern recognition. The GMDH, introduced by Ivakhnenko [16-17], is a hierarchical and learning network structure that provides an effective approach to identify higher order non-linear systems. Its main purpose is the identification of relations in large complex non-linear multidimensional systems as well as their approximation and prediction. In the GMDH network, the part which corresponds to the neuron of a neural network is called the "N-Adaline", and is generally expressed by a polynomial. The N-Adaline is composed of two inputs and one output, and the output is generated by combinations of two inputs [18]. Two inputs, x_i and x_j, are then combined to produce a partial descriptor based on the simple quadratic transfer function as the following:

$$\hat{y}_n = a_0 + a_1 x_{i_n} + a_2 x_{j_n} + a_3 x_{i_n} x_{j_n} + a_4 x_{i_n}^2 + a_5 x_{j_n}^2 \quad (1)$$

where $\hat{y}_n$ is determined using the least squares method, and the coefficients a_0 to a_5 are determined statistically and are unique for each transfer function. The coefficients can be thought as analogous to weights found in other types of neural networks.

The GMDH topology is usually determined using a layer by layer pruning process based on a pre-selected criterion of what constitutes the best nodes at each level. The traditional GMDH method is based on an underlying assumption that data can be modeled by using an approximation of the Volterra Series or Kolmogorov-Gabor polynomial as shown as follows:

$$y = a_0 + \sum_{i=1}^{M} a_i x_i + \sum_{i=1}^{M}\sum_{j=1}^{M} a_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ijk} x_i x_j x_k \ldots \quad (2)$$

where X (x_1, x_2 ,…., x_M) is the vector of input variables, and A (a_1 , a_2 ,…., a_M) is the vector of summand coefficients [18].

When constructing GMDH, all combinations of the inputs are generated and sent into the first layer of the network. The outputs from this layer are then classified and selected as input for the next layer with all combinations of the selected outputs, sent into the layer 2. This process is continued as long as each subsequent layer (n+1) produces a better result than layer (n). When layer (n+1) is found not to be as good as layer (n), the process will be stopped. Now, each layer consists of nodes that a pair of inputs is its source. In GMDH topology, each node produces a set of coefficients (a_i & $i \in \{1,2,3,..,5\}$) that are estimated by using training data. Then, the fitness is tested by evaluating the mean square error of the predicted $\hat{y}$ and actual y values as follows:

$$\text{Error} = \sum_{n=1}^{N} (\hat{y}_n - y_n)^2 \quad (3)$$

To identify the coefficients with the best fit, the partial derivatives of equation (3) are calculated on each constant value a_i, and set it equal to zero as follows:

$$\frac{\partial \text{Error}}{\partial a_i} = 0 \quad (4)$$

Finally, to compare the predicted and actual values, average absolute deviations (AAD%) and root mean squared error (RMSE) were calculated as follows:

$$\text{ADD\%} = \frac{\sum_{i=1}^{N_t=50} \left| \left(\frac{Y_i^{\text{Actual}} - Y_i^{\text{Predict}}}{Y_i^{\text{Actual}}} \right) \right|}{N_t} \times 100 \quad (5)$$

$$RMSE\% = \sqrt{\frac{\sum_{i=1}^{N_t=50}\left(\frac{Y_i^{Actual} - Y_i^{Predict}}{Y_i^{Actual}}\right)^2}{N_t}} \times 100 \qquad (6)$$

where N_t , Y_i^{actual} , $Y_i^{Predict}$ are the number of test runs, actual variables, and the predicted values, respectively.

4. Results and discussion

To build up the GMDH model, fifty data points were gathered from the target visbreaking unit from the start of the run (SOR= 1st day) to the end of the run (EOR= 422th day). These points were included of gas, gasoline, and fuel oil product yields, severity, fuel oil viscosity, feed flow rate, outlet coil temperature, initial boiling point (IBP) of feed, steam flow rate, and viscosity of the feed.

Among these data points, 70% were selected for training the GMDH network, and the remained ones (i.e. 30%) were applied for validating that. To model the output variables of the plant (i.e. severity, the yield of products and viscosity of fuel oil), five different structures were selected. The input layer of each structure includes the significant process variables of the visbreaking unit i.e. feed flow rate, coil temp, steam flow rate, viscosity and IBP of feed.

According to the mentioned approach, it was found that a GMDH network with two neurons in the hidden layer (intermediate layer) was accurate enough to simulate outputs of the plant, and therefore, growing step of the network was stopped. The schematic of developed GMDH networks are shown in Fig.2. As seen, all models have one input layer, two intermediate layers, and one output layer. Additionally, the corresponding polynomial equations of the proposed models for the growth period are presented in Tables 3 to 7.

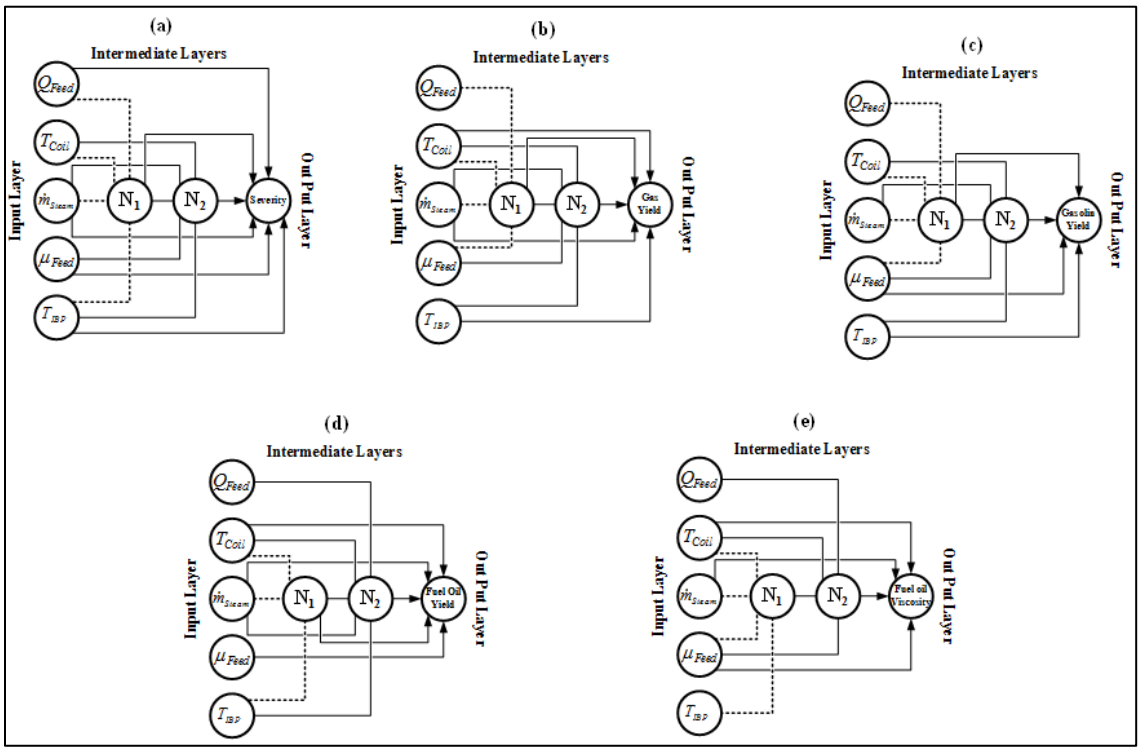

Fig. 2. A schematic of the proposed GMDH neural network, (a): severity, (b): gas yield (c): gasoline yield, (d): fuel oil yield and (e): fuel oil viscosity.

Table 3. Nodal expressions for GMDH neural network (severity)

$$Severity = -14.8478 + 0.00026 \times Q_{Feed}\ \mu_{Feed} - 0.54063 \dot{m}_{Steam} + 0.00488 \dot{m}_{Steam}\ N_1 - 0.00695 N_1^{\ 2} + 1.76735 N_2$$

$$N_2 = -48.3297 - 0.00680\ T_{Coil}\ \mu_{Feed} + 0.01351 N_1 T_{Coil} + 5.27376 - 0.01283\ T_{IBP}\ \dot{m}_{Steam} + 0.00728\ T_{IBP}\ \mu_{Feed} - 0.01099\ T_{IBP}\ N_1$$

$$N_1 = -570.764 - 0.00062\ \dot{m}_{Steam}\ Q_{Feed} - 0.00056 Q_{Feed}\ T_{IBP} + 0.00374 T_{Coil}^{\ 2}$$

Table 4. Nodal expressions for GMDH neural network (gas yield)

$$Gas_{\ Yield} = -58.8284 + 0.00091 \dot{m}_{Steam}\ T_{Coil} + 0.00331\ T_{Coil}\ N_2 - 0.00099 T_{IBP}\ \dot{m}_{Steam} - 0.03679 T_{IBP}\ N_1 + 0.00038 T_{IBP}^{\ 2} + 12.4756 N_1$$

$$N_2 = 124.944 + 1.58075 \times 10^{-5} T_{Coil}\ \dot{m}_{Steam} - 0.00075 T_{Coil}\ T_{IBP} + 0.00682 N_1 T_{Coil} - 0.311376\ \mu_{Feed} + 0.00076\ \mu_{Feed}\ T_{IBP}$$

$$N_1 = -24.7593 - 2.57761 \times 10^{-5} \dot{m}_{Steam}\ Q_{Feed} - 1.07009 \times 10^{-5} Q_{Feed}\ \mu_{Feed} + 0.00015 T_{Coil}^{\ 2}$$

Table 5. Nodal expressions for GMDH neural network (gasoline yield)

$$Gasoline_{\ Yield} = -1.42158 - 0.17208\ \mu_{Feed} + 0.00041\ \mu_{Feed}\ T_{IBP} - 0.03360\ N_2 T_{IBP} + 15.3966 N_1 - 3.1162 N_1 N_2 + 3.05182 N_2^{\ 2}$$

$$N_2 = 30.7535 + 0.00048 T_{Coil}\ \dot{m}_{Steam} - 0.00048 \mu_{Feed}\ T_{Coil} - 0.00051 \dot{m}_{Steam}\ T_{IBP} - 0.00838\ \mu_{Feed}\ N_1 + 0.00030\ \mu_{Feed}^{\ 2} + 0.01330 N_1 T_{IBP}$$

$$N_1 = 405.685 - 3.00727 \times 10^{-5} Q_{Feed}\ \dot{m}_{Steam} - 1.88247 \times 10^{-5} Q_{Feed}\ \mu_{Feed} - 1.98056 T_{Coil} + 0.00244 T_{Coil}^{\ 2}$$

Table 6. Nodal expressions for GMDH neural network (fuel oil yield)

$$Fuel\ oil_{\ Yield} = 140.263 - 0.74931 T_{Coil} + 0.00185 T_{Coil}\ \mu_{Feed} + 0.00019\ \dot{m}_{Steam}\ N_1 - 6.55249 \times 10^{-5} \dot{m}_{Steam}^{\ 2} - 0.00094\ \mu_{Feed}^{2} + 1.13092 N_2$$

$$N_2 = -467.335 + 6.80173 \times 10^{-5} Q_{Feed}\ T_{IBP} - 0.00057\ \dot{m}_{Steam}\ T_{Coil} + 0.25608 \dot{m}_{Steam} + 11.341 N_1 - 0.0576588 N_1^{\ 2}$$

$$N_1 = 153.951 - 0.00192 T_{Coil}\ \dot{m}_{Steam} + 0.00209 T_{IBP} - 0.00037 T_{IBP}^{\ 2}$$

Table 7. Nodal expressions for GMDH neural network (fuel oil viscosity)

$$Fuel\ oil\ vis = 2.37973 - 0.00093\ \dot{m}_{Steam}\ T_{Coil} - 0.00010 T_{Coil}\ \mu_{Feed} + 0.00252\ N_2 T_{Coil} + 0.46694\ \dot{m}_{Steam} - 0.00076 N_2\ \dot{m}_{Steam} + 5.42687 \times 10^{-5}\ \mu_{Feed}^{2}$$

$$N_2 = -486.283 + 3.6551 Q_{Feed} - 0.00848 Q_{Feed}\ \mu_{Feed} - 0.00093\ \mu_{Feed}\ T_{Coil} + 0.00541 T_{Coil}\ N_1 + 1.53846\ \mu_{Feed} - 1.38076 N_1$$

$$N_1 = 3314.42 - 15.5018 T_{Coil} - 0.00275 T_{Coil}\ \mu_{Feed} + 0.01895 T_{Coil}^{\ 2} + 0.00103 \dot{m}_{Steam} + 1.30947\ \mu_{Feed} - 0.32258 T_{IBP}$$

Figures 3 to 7 show the comparison between the measured output variables obtained from the target visbreaking plant and the predicted ones using the GMDH network. As seen, a reasonable agreement can be observed.

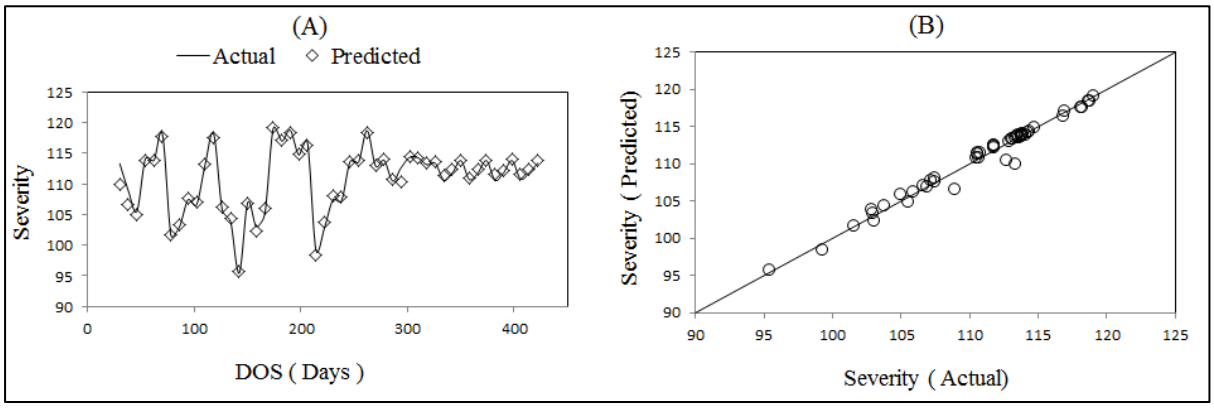

Fig.3. (A) Comparison of actual severity versus predicted, and (B) the regression plot of predicted severity versus actual

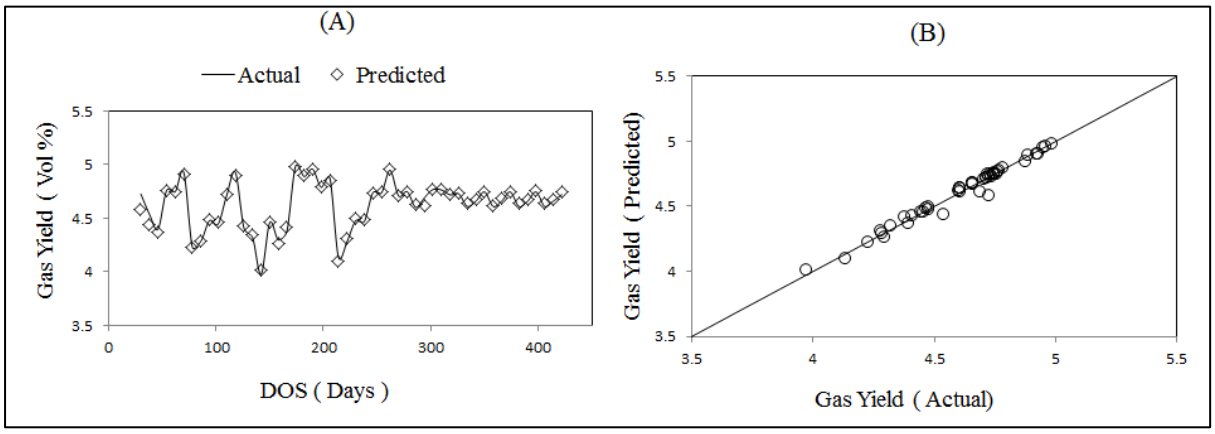

Fig.4. (A) Comparison of actual gas yield versus predicted, and (B) the regression plot of predicted gas yield versus actual.

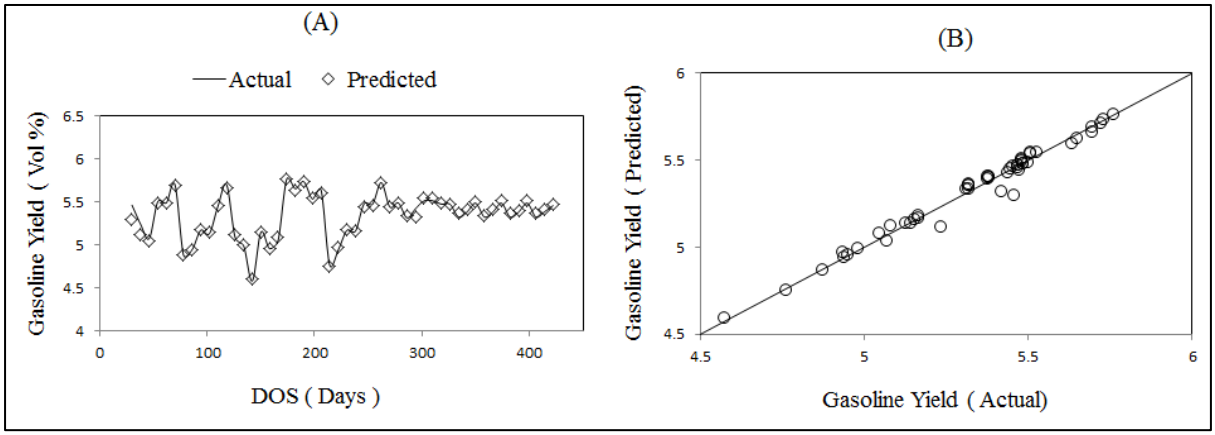

Fig.5. (A) Comparison of actual gasoline yield versus predicted, and (B) the regression plot of predicted gasoline yield versus actual.

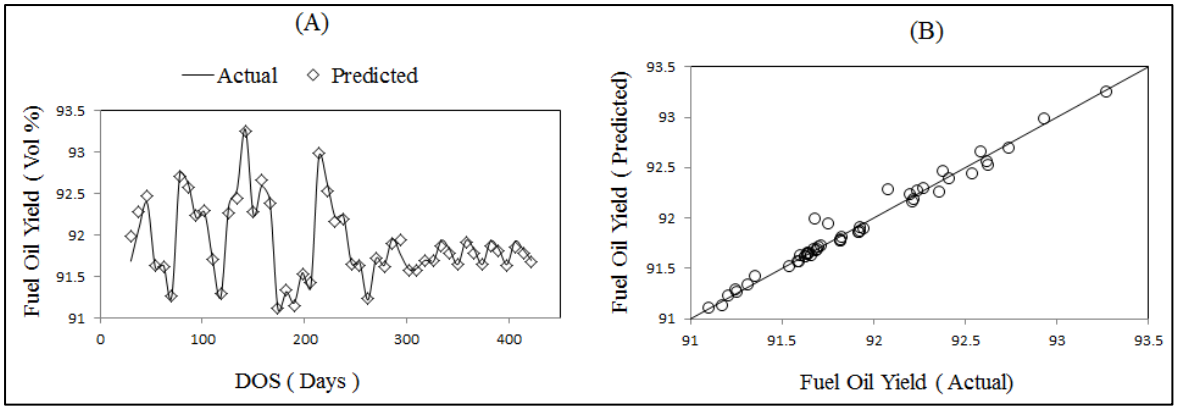

Fig.6. (A) Comparison of actual fuel oil yield versus predicted, and (B) the regression plot of predicted fuel oil yield versus actual.

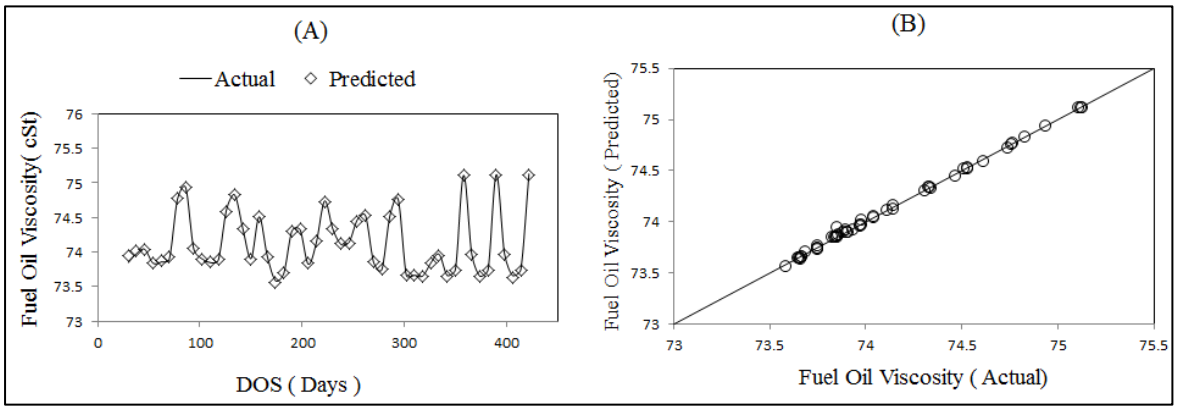

Fig.7. (A) Comparison of actual fuel oil viscosity versus predicted, and (B) the regression plot of predicted fuel oil viscosity versus actual.

To have a better justification, Table 8 reveals the AAD% and RSME% of the results for the mentioned process variables. As observed, the AAD% and RSME% of predicting are less than 1%, and therefore, the GMDH network is reliable to simulate the outputs of the target commercial scale visbreaking plant.

Table 8. RMSE% and AAD% of proposed GMDH model

No.	Process output variable	ADD %	RMSE %
1	Severity	0.462	0.717
2	Gas yield	0.448	0.709
3	Gasoline yield	0.458	0.716
4	Fuel oil yield	0.051	0.078
5	Fuel oil viscosity	0.016	0.025

It is supposed that the main deviation of the model is mainly related to some outliers, such as power fluctuation of instruments, calibration of analysis devices, human errors, and signal transmission that cannot be excluded from the measured data points. However, from the presented results, it can be concluded that the proposed approach i.e. GMDH network is reliable enough to be applied for engineering applications such as process optimization.

5. Conclusions

In this research, a group method of data handling (GMDH) network was applied to predict momentous output variables of a commercial scale visbreaking unit located in an oil refinery.

These process output variables were severity, gas, gasoline, and fuel oil product yields, and fuel oil viscosity. GMDH considered the influence of feed flow rates, outlet coil temperature, steam flow rate, feed viscosity and IBP of feed.

To establish the model, 50 data points from the SOR to the EOR were used to build and train the GMDH network. It was found that a GMDH network with two neurons in the intermediate layer was accurate enough to simulate outputs of the plant i.e. severity, gas, gasoline, and fuel oil product yields, and fuel oil viscosity with the AAD% and RSME% of less than 1%. Therefore, the GMDH network is reliable to simulate the outputs of the target comercial scale visbreaking plant, and it is also applicable to model such complex systems.

Nomenclature

$\hat{y}_n$	*Predicted value*	Q_{Feed}	*Flow rate of feed (m^3/h)*
x_{i_n}, x_{j_n}	*Two input neurons*	μ_{Feed}	*Kinematic viscosity of feed (cSt)*
a_0 *to* a_5	*Coefficients*	*ANN*	*Artificial neural network*
y	*Actual value*	*GMDH*	*Group method of data handling*
N_t	*Number of test run*	*SOR*	*Start of run*
$\overset{0}{m}_{Steam}$	*Mass flow rate of steam (kg /h)*	*EOR*	*End of run*
T_{Coil}	*Outlet coil temperature (°C)*	*RMSE*	*Root mean square error*
T_{IBP}	*Initial boiling point of feed (°C)*	*AAD*	*Absolute average deviation*

References

[1] Gary J H, Handwerk GE. Petroleum Refining Technology and Economics, Fourth ed., Marcel Dekker Inc., New York, 2001, 88, p.125.
[2] Speight J. The Chemistry and Technology of Petroleum, fifth ed., CRC Press, Boca Raton, 2014, p.442.
[3] Sawarkar AN, Pandit AB, Samant SD, Joshi J. B. The Canadian Journal of Chemical Engineering, 2007; 85: 1.
[4] Sadighi S, Mohaddecy SR. International Journal of Technology, 2013; 4: 102.
[5] Singh J, Kumar S, Garg M. Fuel Processing Technology, 2012; 94: 131.
[6] Chacón R, Canale A, Bouza A, Sánchez Y. Braz. J. Chem. Eng., 2012; 29: 135.
[7] Mohaddecy SR, Sadighi S. Chemical Technology: An Indian Journal, 2012; 6: 7.
[8] Assidjo, E.: Yao, B.: Kisselmina, K.: Amané, D., Braz. J. Chem. Eng., 2008, 25, 515.
[9] Sadighi S, Mohaddecy SR. Journal of Petroleum and Gas Engineering, 2013; 4: 8.
[10] Fonseca AP, Oliveira JV, Lima EL. Braz. J. Chem. Eng., 2000; 17: 517.
[11] Sadighi S, Mohaddecy SR. Petroleum Technology Quaterely, 2014; Q2: 117.
[12] Sedighi M, Keyvanloo K, Towfighi J. Industrial & Engineering Chemistry Research Journal, 2011; 50:1536.
[13] Atashrouz S, Amini E, Pazuki G. Ionics, 2015; 21: 1595.
[14] Bennett FR, Crew P, Muller JK. Int. J. Mol. Sci., 2004; 5: 101.
[15] Chen L, Wang T. Indian Journal of Engineering & Material Science, 2010; 17: 179.
[16] Ivakhnenko AG. Polynomial Theory of Complex Systems, IEEE Transactions on Systems, Man and Cybernetics., 1971; 4: 364.
[17] Ivakhnenko AG. Soviet Automatic Control, 1986; 1: 43.
[18] Fujii K, Yamamoto T. Electr. Eng. Jpn., 2014; 188:31.

To whom correspondence should be addressed. E-mail: Seifsr@ripi.ir