PREDICTION OF CRUDE OIL PYROLYSIS PROCESS USING RADIAL BASIS FUNCTION NETWORKS

(Recibido el 15-06-2017. Aprobado el 04-09-2017)

Milad Norouzpour
Young Researchers and Elite Club, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran
milad. norouzpour@yahoo.com

Ahmad Reza Rasouli
Department of Oil and Gas Engineering, Shahid Bahonar University of Kerman, Iran

Abdolreza Dabiri
Department of Petroleum Engineering, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran

Amin Azdarpour
Department of Petroleum Engineering, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran

Mohammad Afkhami Karaei
Department of Petroleum Engineering, Marvdasht Branch, Islamic Azad University, Marvdasht, Iran

RESUMEN: Conocimiento del petróleo crudo La pirólisis y la combustión es una de las más importantes en la producción de petróleo utilizando el método de combustión in situ como una sección de los métodos de recuperación térmica de aceite mejorado (EOR). En este método, el petróleo crudo experimenta una serie de cambios físicos y químicos que pueden referirse a la pirólisis como la parte más importante de estos cambios. En este trabajo, hemos desarrollado modelos de Redes de Función Radial (RBFN) para predecir el peso restante de crudo durante el proceso de pirólisis del crudo. La densidad API, la viscosidad, la resina y el asfalteno y otros componentes del contenido de aceite crudo, la temperatura y la velocidad de calentamiento se seleccionan como parámetros de entrada RBFN, mientras que el peso restante de crudo en diferentes temperaturas se considera como salida de red. Los datos se obtuvieron mediante análisis termogravimétricos y experimentos de separación en seis muestras de diversos aceites crudos iraníes. Los resultados de este trabajo muestran que utilizando un RBFN, podemos predecir el peso restante de crudo durante su proceso de pirólisis con un error relativo absoluto medio (ARE) 5,88 por ciento y error cuadrático medio (MSE) 6,15 por la función newrbe y un promedio absoluto Error relativo (ARE) 7,25 por ciento y error cuadrático medio (MSE) 2,51 por función newrb para datos de prueba. Más aún, los resultados del análisis de regresión mostraron una muy buena coincidencia entre los resultados de laboratorio y los resultados previstos por la RBFN propuesta.

Palabras clave: Petróleo crudo, EOR, RBFN, Pirólisis

ABSTRACT: Knowledge of crude oil Pyrolysis and combustion is one of the most important in oil production using in situ combustion method as a section of thermal enhanced oil recovery (EOR) methods. In this method, crude oil undergoes a series of physical and chemical changes that can refer to pyrolysis as a most important part of these changes. In this work, we have developed Radial Basis Function networks (RBFN) models to predict remaining weight of crude oil during crude oil pyrolysis process. API density, viscosity, resin and asphaltene and other components of crude oil content, temperature and heating rate are selected as RBFN input parameters, whereas remaining weight of crude oil in different temperatures is considered as network output. The data were obtained by doing thermogravimetric analysis and separation experiments on six samples of various Iranian crude
oils. The results of this work show that using a RBFN, we can predict the remaining weight of crude oil during its pyrolysis process with an average absolute relative error (ARE) 5.88 percent and mean square error (MSE) 6.15 by newrb function and an average absolute relative error (ARE) 7.25 percent and mean square error (MSE) 2.51 by newrb function for test data. More over, the results of regression analysis showed a very good coincidence between the laboratory results and predicted results by the proposed RBFN.

**Keywords:** Crude oil, EOR, RBFN, Pyrolysis

1. **INTRODUCTION**

In-situ combustion is an important enhanced oil recovery process that has been studied extensively past years. This process has been considered particularly applicable for in-situ recovery of medium and heavy oil reservoirs. In in-situ combustion, heat is generated within the reservoir by igniting the formation oil and then propagating a combustion front through the oil reservoir. The fuel necessary to sustain the combustion front is supplied by the heavy residual material or coke that deposits on the sand grains during distillation, thermal cracking, pyrolysis etc. of the crude oil ahead of the combustion front [1-4]. The crude oil producing mechanisms in this method are: reducing oil viscosity due to the produced heating and solving produced gases in the oil, heating expansion of oil, oil distillation and cracking, solution gas drive and producing more pressure gradient because of injected and produced gases[5].

In this method, crude oil on heating and in inert atmosphere undergoes a series of physical and chemical changes that called pyrolysis[1]. Crude oil pyrolysis process and the factors affecting it play an important role in the performance prediction of an in situ combustion process [6]. Thermogravimetric analysis (TG) can be used to study the kinetics of a crude oil chemical reaction and determine weight loss and basic kinetic constants [7]. Bae investigated the thermo-oxidative behavior and fuel-forming properties of various crude oils using TG/DTG. He performed several examinations in air and nitrogen atmospheres and different pressure for studying of crude oil pyrolysis and combustion processes. The results indicated that oils can be classified according to their oxidation characteristics. No complete correlation could be established between viscosity, composition or density of the crude with the thermo-oxidative characteristics or pyrolysis and combustion processes of the oil [8]. Ciajolo and Barbella implemented an investigation on the pyrolysis and combustion of some heavy fuel oils and their separate paraffinic, aromatic, polar and asphaltene fractions using TG. Results showed that the thermal behavior of fuel oil can be interpreted in term of a low temperature (<400 °C) phase involving the volatilization of paraffinic and aromatic fractions, and a high temperature phase in which the polar and asphaltene fractions pyrolyse and leave a particulate carbon residue [9]. Kamal and Verkocy used TG/DTG and DSC on two Lloydminster regions, heavy-oil cores, and extracted oils and mineral matter. TG/DTG and DSC curves, obtained in helium and air atmospheres, for the two Lloydminster region cores and the extracted oils demonstrated at least three groups of chemical reactions occurring in three temperature regimes. Reactions in zone 1 were attributed to evaporation, distillation, thermalysis, and low-temperature oxidation (LTO), in zone 2 to distillation and thermal alteration of minerals, LTO, and combustion, and in zone 3 and/or 4 to pyrolysis, coking, polymerization, mineral matter decomposition and combustion [5]. Ranjar and Pusch studied the effect of the oil composition, based on their light hydrocarbon, resin and asphaltene contents, on the pyrolysis kinetics of the oil and the combustion kinetics of the fuel, using TG/DTG and DSC. The colloidal composition of oil, as well as the transferability and heat transfer characteristics of the pyrolysis medium, had a pronounced influence on the fuel formation and composition [10]. Ranjar and Pusch studied the effect of reservoir rock composition on the pyrolysis and combustion behaviour of crude oils in porous media. Pyrolysis and combustion tests were performed to examine how clays affect the amount of fuel and its reactivity. He concluded that clay minerals present in the matrix enhance fuel deposition during the pyrolysis process and also catalyze the oxidation of fuel [11]. Kok and Karacan presented the results of an experimental study on the determination of the pyrolysis behaviour and kinetics of six crude oils using DSC and TG/DTG. Two main temperature ranges accompanied by loss of mass were observed during pyrolysis. The first region, between ambient and 400 °C was distillation. The second region, between 400°C and 600 °C was vis-breaking and thermal breaking. As the crude oil becomes denser, the activation energy of cracking increases. The activation energy of cracking also shows a general trend with asphaltene content [12]. Studying this process using TG and other related examinations are very complicated, time consuming and expensive.

Beside, within the recent years, artificial neural networks (ANNs) have used successfully in many different areas of the oil industry such as hydrocarbon saturation, permeability, minimum miscibility pressure for Co2 flooding, residual water saturation and reservoir production performance with low permeability by numerous researchers [13-18]. This due to the ANNs features (learning and generalization, self organization, parallel process, high velocity and less expense) [19-21]. The aim of the present work is use of the ANNs features to develop Radial Basis Function network (RBFN) models that can be used to estimate remaining weight of
crude oil during crude oil pyrolysis process based on physical properties and chemical composition of crude oil, temperature and heating rate for the first time. This can be viewed as a multivariable interpolation problem in which it is required to estimate the function relating the input to the output using a set of input-output data. This kind of problem is referred to in the literature by different names such as nonparametric regression, function approximation and supervised learning in neural network terminology [22].

1.1. Artificial neural networks (ANNs)

In the mid 1980s, the understanding of neural network models increased significantly after major developments in neuroscience [14]. They represent a wide category of computer algorithms that are designed to imitate the function of biological neurons in the human brain. They have a significant ability to find highly complex non-linear relationship between variables, even in situations where the input information is noisy or is less defined [23]. They have the ability to learn and generalize, whereby they can produce reasonable results for inputs not seen during training [24].

There are mainly two types of artificial neural networks, feed-forward ANNs and recurrent ANNs. Feedforward ANNs maintain a high level of research interest due to their ability to map any function to an arbitrary degree of accuracy. This has been demonstrated theoretically for both the RBFN [25] and the popular multilayer perceptron (MLP) network [26]. Feedforward ANNs have been applied to many diverse areas such as pattern recognition, time series prediction, signal processing, control and a variety of mathematical applications. The RBF [27, 28] was developed from an exact multivariate function interpolation [29] and has attracted a lot of interest since its conception. There are a number of significant differences between RBFs and MLPs:

- The RBF has one hidden layer while the MLP can have several.
- The hidden and output layer nodes of the RBF are different while the MLP nodes are usually the same throughout.
- RBFs are locally tuned while MLPs construct a global function approximation.

1.2. The Radial Basis Function network (RBFN)

The RBF architecture consists of a simple two layered network (a hidden layer and an output layer); each layer is fully connected to the one following [30]. A RBFN structure is depicted in Fig. 2. As we have mentioned, the hidden layer is composed of a number of nodes, RBF nodes, with radial activation functions, which shall be taken, in this analysis, as Gaussian functions. Two Parameters are associated with each RBF node, the centre and width. Both of these quantities refer to properties of the Gaussian function. Associated with the hidden to output connections, are conventional signal multipliers: the weights. The final output processing unit merely yields a weighted sum of its inputs [31].

![A typical view of a RBFN structure](image_url)

Their use in the literature is extensive, and its application varies from face recognition [32] to time series prediction [33]. The RBFNs are under continuously research, so we can find abundant literature about extensions and improvements of RBFNs learning and modeling [34-38]. Recently, we can find some work analyzing the behavior of RBFNs [39-41] and improving their efficiency [42-43]. As we can see from the recent and past literature, we can conclude that RBFNs are a widely employed and well-known model which is actually used.
Radial Basis Function Network (RBFN) is well suited for function approximation and pattern recognition due to its simple topological structure and its ability to reveal how learning proceeds in an explicit manner. A RBF is a function which has been built into a distance criterion with respect to a center [31]. Different basis functions have been proposed for the hidden-layer neurons, but normally the selected one is the Gaussian function. There are six basis functions, which are recognized as having useful properties for RBF networks [44–47]:

1. Multiquadratic:
\[ \varphi(x) = (\sqrt{x^2 + \sigma^2})^{\alpha/2}, \quad \text{For } \alpha > 0 \text{ and } x \in \mathbb{R} \]  
(1)

Which is a case of
\[ \varphi(x) = (\sqrt{x^2 + \sigma^2})^{\alpha}, \quad 0 < \alpha < 1. \]  
(2)

2. Gaussian:
\[ \varphi(x) = \exp(-\frac{x^2}{2\sigma^2}), \quad \text{For } \sigma > 0 \text{ and } x \in \mathbb{R}. \]  
(3)

3. Inverse multiquadratic:
\[ \varphi(x) = \frac{1}{(\sqrt{x^2 + \sigma^2})^{\alpha}}, \quad \beta > 0. \]  
(4)

4. Thin plate spline:
\[ \varphi(x) = \frac{1}{\sqrt{\sigma^2 + x^2}}. \]  
(5)

5. Cubic:
\[ \varphi(x) = x^3. \]  
(6)

6. Linear:
\[ \varphi(x) = x. \]  
(7)

The problem of estimating remaining weight of crude oil during crude oil pyrolysis process based on physical properties and chemical composition of crude oil, temperature and heating rate may be considered as a multivariable interpolation one. This entails finding an approximating function \( H(x) \) representing remaining weight of crude oil during crude oil pyrolysis process in which \( x \) is an \( N \)-dimensional vector. In this case, the \( x \) components are the independent variables, namely, API density, viscosity, resin and asphaltene and other components of crude oil content, temperature and heating rate (i.e. \( N=7 \)). In RBF networks, \( H(x) \) is expressed as a linear combination of multivariate Gaussian basis functions [48,49],

\[ H(x) = \sum_{j=1}^{M} W_j \varphi_j(x) + b_0 \]  
(8)

Where
\[ \varphi_j(x) = \exp \left[ -\sum_{i=1}^{N} \frac{(x_i-c_{ij})^2}{2\sigma_j^2} \right] \]  
(9)

\( \varphi_j(x) \)is the output of the jth node in the hidden layer, \( c_{ij} \) is the center of the jth RBF node for the kth input variable \( x_k \), \( \sigma_j \) is the width of the Gaussian function, \( W_j \) is the weight between jth RBF unit and output layer neuron, \( b_0 \) is the bias term i.e. data independent variable at output node, \( M \) is the number of hidden layer neurons.

The learning process of a radial basis function network involves using the input-output data to determine the parameters \( c_{ij}, \sigma_j \) and \( W_j \). One of the techniques used to obtain these parameters is based on assuming fixed radial basis functions. In this method, the centers are randomly selected from the training data set. On the other hand, the width of the Gaussian radial basis function is expressed in terms of the maximum distance between the chosen center \( d \) and the number of centers \( M \) as
\[ \sigma_j = \frac{d}{\sqrt{2M}} \]  
(10)

With \( c_{ij} \) and \( \sigma_j \) specified for all the hidden nodes, it remains to determine the weights. These can be computed using multiple linear regression techniques [49]. This involves processing the \( P \) training patterns through \( M \) hidden nodes to generate an \( M \times P \) matrix, \( Q \), say. The aim is to find the weights vector \( W \) (\( 1 \times M \) vector) such that the error between the computed output vector \( S \) (\( 1 \times P \) vector) and the target output vector, \( T \), is minimum. In matrix form, this translates to minimizing:
\[ ||T - S|| = ||T - WQ|| \]  
(11)

where \( || \cdot || \) represents the Euclidean norm.

The least-squares solution of (11) for the vector \( W \) can be found using the pseudoinverse of \( Q \) as follows:
\[ W = QTQ^T \]  
(12)

2. EXPERIMENTAL

2.1 Thermogravimetric analysis:

Experiments (TG/DTG and DSC) were performed using STA-409 PG thermal analysis system produced by Netzsch Company. Experiments were performed with a sample size of 30 mg at 1, 5 and 10 (˚C) ⁄min heating rates on six Iranian heavy crude oil samples. Nitrogen flow rate was kept constant at 50 ml ⁄min in the temperature range of 30 – 800 ˚C.

2.2. Separation experiments:

The asphaltene fractions of oils were separated by diluting the crude oil with forty volumes of n-heptane. The mixture was shaken about one hour and stored for two nights in dark. Filtration of this mixture through a 42 Whatman filter was the next step to obtain asphaltene. The asphaltene fractions were washed with n-heptane until no yellow color due to resins or oils were visible in the wash. Ethyl acetate acid is applied for precipitation of resin and asphaltene fractions. The crude oils were diluted with twenty volumes of ethyl acetate acid. The
mixture was shaken about one hour and stored for two nights in dark. And then it was filtered through a 42 Whatman filter to obtain resins and asphaltenes. Resins percent of crude oils were determined with subtraction of asphaltenes percent of previous step from resins and asphaltenes percent of this step.

Results of separation experiments and other crude oils specifications were listed in Table 1.

Table 1. specification of used crude oils

<table>
<thead>
<tr>
<th>oil name</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>20.83</td>
<td>21.87</td>
<td>18.67</td>
<td>20.26</td>
<td>13</td>
<td>30.39</td>
</tr>
<tr>
<td>Viscosity (cp)</td>
<td>157.30</td>
<td>117.90</td>
<td>306.51</td>
<td>185.87</td>
<td>3345.57</td>
<td>18.85</td>
</tr>
<tr>
<td>Resins (wt %)</td>
<td>8.09</td>
<td>6.01</td>
<td>13.73</td>
<td>14.04</td>
<td>16.42</td>
<td>4.01</td>
</tr>
<tr>
<td>Asphaltenes (wt %)</td>
<td>4.73</td>
<td>5.42</td>
<td>10.71</td>
<td>9.66</td>
<td>18.95</td>
<td>3.25</td>
</tr>
<tr>
<td>Dispersant (wt %)</td>
<td>87.18</td>
<td>88.57</td>
<td>75.56</td>
<td>76.30</td>
<td>64.63</td>
<td>92.74</td>
</tr>
</tbody>
</table>

3. THE PROPOSED RBF MODEL

In this work, using MATLAB software, we have applied a RBFN for prediction of remaining weight of crude oil during crude oil pyrolysis process. API density, viscosity, resin and asphaltene and other components of crude oil content, temperature and heating rate are selected as RBFN input parameters, whereas remaining weight of crude oil in different temperatures is considered as network output. Estimation of parameters (spread and number of neuron in hidden layer) in neural network model is not systematic process; rather it is possible by try and error and repeat of tests. In this work, the optimum network is determined by constructing several models and repeat them. Each ANN has been trained with 3/4 of data set and 1/4 of data set have been used to evaluate their accuracy and trend stability. For a proper working of the neural network a preprocessing of the input and output data is performed, that after exerting them, the input and output data are normalized between [-1, 1]. It is clear that after finishing simulation, converse of the above functions were exerted. During learning process, regularly the learning rate of network was measured by target functions and at last, the network was accepted that had the least error and the best regression analysis coefficients. Network error was calculated through the following equation:

\[
ARE = \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{|t(i) - a(i)|}{t(i)} \right] \\
MSE = \frac{1}{n} \sum_{i=1}^{n} (t(i) - a(i))^2
\]

Where t and a are target and predicted output, respectively, and n is number of used data for network training or testing.

Since in each run, training and testing data are selected randomly, noting this point is necessary that the resulted answers in each run are different from the other runs. Because of this, each of the used networks was operated 10 times and their average of results was considered as the networks finally result.

4. RESULTS AND DISCUSSION

At first, the structure of network was determined and then prediction of remaining weight of crude oil during crude oil pyrolysis process has been performed.

4.1. Effect of the number of neurons in hidden layer on network performance

MATLAB software use newrb and newrbe functions to make of a RBFN. The function newrb iteratively creates a RBFN one neuron at a time. Neurons are added to the network until the sum-squared error falls beneath an error goal or a maximum number of neurons (number of training set) have been reached whereas the newrbe function produces only a network with as many hidden neurons as there are input vectors. In this work, the newrbe function is used for investigation of Effect of the neurons number of hidden layer on network performance. The value, 0.84, is used as the Gaussian spread of networks. Fig. 2 shows the mean of network ARE versus the number of hidden layer after ten runs.
4.2. Determining the optimum Gaussian spread:

Determining the optimum Gaussian spread is the most important topic in design of RBFN. The Gaussian spread determines the width of an area in the input space to which each neuron responds. It is important that the spread parameter be large enough that the RBFN neurons respond to overlapping regions of the input space, but not so large that all the neurons respond in essentially the same manner. In this work, the Gaussian spread is determined by both newrb and newrbe functions.

For determining the optimum Gaussian spread, different RBFNs are constructed with the variant spread in range (0,1) and spacing 0.02. Fig. 3 and Fig. 4 show the mean of network AREs versus the Gaussian spread after ten runs for newrb and newrbe functions, respectively.

As it is observed in Fig. 3 and Fig. 4, the least error was obtained in the networks with the Gaussian spread 0.12 and 0.70 for newrb and newrbe functions, respectively.

4.3. The optimum network

After performing the above steps, different networks are constructed using newrb and newrbe functions with the Gaussian spread 0.84 and 0.70, respectively (the number of neurons in hidden layer of newrb function fix to 14 neurons).
Since in each run, both the network weights and training and testing data are selected randomly, hence the network ARE and MSE are different in each run. Minimum, maximum and mean values of network ARE and MSE after 10 runs are given in Table 2.

### Table 2. Network errors for training and testing sets

<table>
<thead>
<tr>
<th>Data set</th>
<th>Training Data <em>(newrb function)</em></th>
<th>Testing Data <em>(newrb function)</em></th>
<th>Training Data <em>(newrbe function)</em></th>
<th>Testing Data <em>(newrbe function)</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum of network ARE (%)</td>
<td>0.0267</td>
<td>1.418</td>
<td>2.21</td>
<td>4.05</td>
</tr>
<tr>
<td>Maximum of network ARE (%)</td>
<td>0.0619</td>
<td>5.024</td>
<td>2.74</td>
<td>7.76</td>
</tr>
<tr>
<td>Mean of network ARE (%)</td>
<td>0.484</td>
<td>3.793</td>
<td>2.52</td>
<td>5.88</td>
</tr>
<tr>
<td>Minimum of network MSE</td>
<td>0.0000949</td>
<td>0.232</td>
<td>0.29</td>
<td>0.85</td>
</tr>
<tr>
<td>Maximum of network MSE</td>
<td>0.000386</td>
<td>4.244</td>
<td>0.448</td>
<td>8.893</td>
</tr>
<tr>
<td>Mean of network MSE</td>
<td>0.000193</td>
<td>2.173</td>
<td>0.96</td>
<td>3.16</td>
</tr>
</tbody>
</table>

Fig. 5 shows experimental data that obtained by thermogravimetric analysis and predicted data that provided by the proposed RBFN *(newrb function)* for testing set. As it is observed in Fig. 5, there is a good agreement between experimental data and predicted data that express performance of proposed RBFN in prediction of remaining weight of crude oil during crude oil pyrolysis process.

Performance of trained network can be measured using error of training and testing sets, but it is better to survey network reaction details with more accuracy. Regression analysis is designed for this project. For performing this analysis, the network output and target vector introduced to respective function *(postreg)* and its output parameters were calculated. This function gives three parameters which the first two parameters are m and b means slope and y-intercept. If there was perfect proportion within output and target (means output equal with target completely), slope will be one and y-intercept will be zero. The third variable returned by postreg is coefficient correlation of outputs and targets. If this number is one, in this case it is distinguished that there is perfect relationship between output and target [50]. Fig. 6 and Fig. 7 shows regression analysis of the modeling of remaining weight of crude oil during crude oil pyrolysis process for train and test data.
Table 3 shows results of regression analysis for train and test data.

**Table 3.** results of regression analysis for train and test data

<table>
<thead>
<tr>
<th>Slope</th>
<th>Train data</th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y-intercept</td>
<td>1.3×10^{-5}</td>
<td>0.38</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>1</td>
<td>0.9996</td>
</tr>
</tbody>
</table>

Regarding the values resulted from regression analysis, it can be concluded that a very good coincidence exists between the laboratory results which achieved from thermogravimetric analysis and predicted results by proposed RBFN.

5. CONCLUSION

In this study, the ability of ANN with RBF neural network for prediction of remaining weight of crude oil during crude oil pyrolysis process has been investigated and the optimum network structure is determined by constructing several models. API density, viscosity, resin and asphaltene and other components of crude oil percentage, temperature and heating rate were selected as network input parameters, whereas remaining weight of crude oil in different temperatures considered as network output. The results show a good agreement between experimental data and predicted data by proposed network.

6. REFERENCES

[31] Julián Luengo, Salvador García, Francisco Herrera, A study on the use of imputation methods for experimentation with Radial Basis Function Network classifiers handling missing attribute values: The good synergy between RBFNs and EventCovering method, Neural Networks, 2009 Elsevier.(3)


