VOL. 16, NO. 9, MAY 2021 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences
©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.
www.arpnjournals.com

ARTIFICIAL NEURAL NETWORK MODELS FOR THE PREDICTION OF ASPHALTENE ONSET PRESSURE (AOP) IN OIL RESERVOIRS

Cristian Loaiza, Jairo Sepúlveda, Germán Arce and Vladimir Mosquera
Universidad Surcolombiana, Avenida Pastrana, Neiva, Huila, Colombia
E-Mail: jsepulve@usco.edu.co

ABSTRACT

Formation damage caused by asphaltene precipitation is one of the most serious problems that can appear to a greater or lesser degree in practically all the stages of production, transportation, and processing of crude oil. Asphaltenes tend to destabilize under certain conditions of pressure, temperature, composition, and flow regime, causing them to aggregate, flocculate, precipitate, and deposit into the porous medium. One of the most popular techniques to evaluate this phenomenon is through the determination of asphaltene onset pressure (AOP). The term "onset" represents the point at which the first solid particle forms after asphaltene destabilization. Thus, in this work, highlights of different artificial neural network models (ANN) developed to predict AOP of oil reservoirs are presented. The research process was carried out in 3 different stages. The first approach was published in 2016. The ANN model was called Neuross-Asph, which was constructed with two self-organizing maps (SOM) and one feed-forward back-propagation network. Later, a modified version of the Neuross-Asph model was introduced in 2018. An extensive statistical analysis (ANOVA) was performed to determine the most appropriate ANN input parameters. Finally, a new ANN version was developed in 2019. In this stage, a fully connected neural network, called AsphPy, was implemented to predict the onset pressure of several crude oil samples. The main advantage of all proposed models was the use of few and accessible data as input parameters. Model predictions showed a good match with the experimental information. Consequently, the strategy designed was appropriate since the network topologies considered the relationships that have more influence on asphaltene precipitation.

Keywords: Asphaltene onset precipitation, artificial neural networks, self-organizing maps, feed-forward, fully connected.

1. INTRODUCTION

Asphaltene precipitation presents a severe challenge to the oil industry because its different effects can be found in all oil crude production line. Consequently, the understanding of asphaltene behavior, as well as the implementation of mechanisms to mitigate/inhibited its precipitation, have become one of the major research topics in petroleum engineering for decades.

Asphaltene is the heaviest and most polarized component of crude oil. This compound exists in petroleum partly dissolved and partly in colloidal form; the asphaltene solubility in oil is dependent on the polarity of the oil medium and the presence of other components (Akbarzadeh *et al*., 2007; Hosseini-Dastgerdi *et al*., 2015). Asphaltene is generally defined as n-alkene-insoluble and aromatic-soluble fraction (Abedini *et al*., 2010; Jewell *et al*., 1972; Kawanaka *et al*., 1991; Leontaritis, 1997; Mansoori, 1997; Vafaie-Sefti and Mousavi-Dehghani, 2006). Changes in oil thermodynamic properties such as fluid composition, reservoir pressure, and temperature can alter the stability of heavy organic compounds. Therefore, asphaltene can be precipitated from the crude oil as a solid, that is, the deposition process (Akbarzadeh *et al*., 2007; Garrouch and Al-Ruhaimani, 2005; Hirschberg *et al*., 1984). Accordingly, any change in oil thermodynamic equilibrium affects asphaltene solubility and leads to asphaltene precipitation (Akbarzadeh *et al*., 2007).

Asphaltene deposition, as a direct response of the asphaltene precipitation mechanism, can trigger multiple problems in the reservoir and production system. These solid precipitated particles cause formation damage (pore throat plugging and decrease in permeability), change the original rock wettability (from water-wet to oil-wet system), pipe plugging, and excessive pressure drop in pipelines (Ahmadi and Shadizadeh, 2012; Hosseini-Dastgerdi and Jafarzadeh-Ghoushchi, 2019; Rasuli *et al*., 2012; Soodbakhsh Taleghani *et al*., 2017). As a result, the efficiency of the oil recovery process is drastically affected, and the operating costs could be increased.

The major questions facing the asphaltene destabilization process are “How” and “How much” heavy organic compounds will precipitate under reservoir conditions. However, the accurate determination of the effect of each effective parameter on the asphaltene precipitation requires multiple experiments. Those procedures at real pressure, volume, and temperature (PVT) conditions tend to be time-consuming and expensive. Thus, the prediction of heavy organic precipitation using theoretical models has been an area of interest for many researchers. The modeling methods are usually classified into five groups including polymer solubility models, equation of state (EOS) models, colloidal techniques, thermodynamic micellization approaches, and molecular thermodynamic models (Zendehboudi *et al*., 2014). The main drawback of prediction models is the lack of experimental data to predict input parameters because their calculations required many ones. Despite that, these models have been used for many years to study asphaltene precipitation (Gharagheizi, F. Eslamimanesh *et al*., 2012; Pazuki and Nikookar, 2006).

Considering the limitations of the extended-traditional models as well as the fact that the asphaltene precipitation/deposition phenomenon is strongly non-

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

linear in terms of process and thermodynamic parameters, some researchers have investigated the potential application of artificial intelligence (AI) techniques to enhance the accuracy of the asphaltene prediction models (Abedini *et al*., 2010; Ahmadi, 2011; Ahmadi and Golshadi, 2012; Ahmadi and Shadizadeh, 2012; Alimohammadi *et al*., 2017; Ashoori *et al*., 2010; Hosseini-Dastgerdi and Jafarzadeh-Ghoushchi, 2019; Khamehchi *et al*., 2012; Mohammadi and Richon, 2008; Rasuli *et al*., 2012; Salahshoor *et al*., 2013; Soodbakhsh Taleghani *et al.*, 2017; Zahedi *et al.*, 2009; Zendehboudi *et al*., 2013, 2014).

Among the intelligent models, Artificial Neural Networks (ANN) have been the most employed to evaluate, analyze, and even predict asphaltene behavior. However, least–squares support vector machine (LSSVM), Bayesian Belief Network (BBN), Genetic Algorithm (GA), and their combinations with ANN, have been other proposed alternatives to model asphaltene precipitation. All efforts have been put into the determination of asphaltene onset pressure (AOP) at specific conditions, the whole asphaltene precipitation envelope (ADE), and the amount of asphaltene-solid precipitates (%wt.), in order to prevent/mitigate the formation damage caused by asphaltene deposition.

In 2008, Mohammadi and Richon proposed the first ANN approach to estimate the precipitation of dissolved asphaltene in a solution of solvent + precipitant. The model was developed based on a feed-forward ANN model, with one (1) output neuron (mass of precipitant/mass of asphaltene), one (1) input neurons (mass of solvent/mass of asphaltene), and one (1) neuron in the hidden layer (with a linear activation function). The acceptable agreement between the prediction of the ANN model and the experimental data reported in the literature, demonstrated the capability of the neural network (NN) technique to study the asphaltene precipitation phenomenon.

Afterward, the second ANN approach for estimation of asphaltene precipitation was proposed by Zahedi *et al*. (2009). This model was applied for a light crude oil at ambient temperature and atmospheric pressure with Multi-Layer Perceptron (MLP) NN architecture. A MLP is a NN with three layers: an input layer, a hidden layer, and an output layer. The results not only showed a good agreement with the experimental data but demonstrated better performance compared with two thermodynamical models.

Abedini *et al*. (2010) and Ashoori *et al*. (2010) compared asphaltene precipitation predictions of an ANN model and some scaling equations under the effect of temperature and solvent ratio. Core-flood tests were carried out with heavy oil samples from Iran. Firstly, they introduced a modified version of proposed scaling equations by Hu *et al*. (2000) and Rassamdana *et al.* (1996). The results obtained from the new scaling equation were more satisfied with the experimental data compared to the original ones. Additionally, an ANN model was designed and applied to predict the amount of asphaltene precipitation. The highlight of both published studies was the fact that ANN models could be more accurate than scaling equations to predict the asphaltene precipitation.

In 2011, Ahmadi developed a model based on a feed-forward ANN optimized by an imperialist competitive algorithm (ICA) to predict asphaltene precipitation. ICA was used to decide the initial weights of the NN. The model effectiveness was validated with a scaling model and a conventional ANN model (without ICA incorporation). One drawback when considering the combination of an ANN and an ICA is the determination of the optimal NN structure, which was determined manually for this scope.

To the best of our knowledge, Khamehchi *et al*. (2012) published the first AI work focused on the prediction of onset pressure, that is, the precise point at which the first solid particle forms after asphaltene destabilization. This model predicted the AOP and the bubble point pressure of a fluid reservoir during CO_2 injection using an adaptive neural fuzzy inference system (ANFIS) technique, which is a combination of fuzzy logic and NN. Nine (9) input parameters were fixed, which were a function of temperature, oil composition, and injected gas composition. The predictions were more accurate compared with other modeling methods, such as the thermodynamic solid model and the Standing correlation. Although Khamehchi *et al*. developed various AI models and the predictions showed good fit, the weight percentage of Saturated, Aromatic, Resin and Asphaltene (SARA) in the crude oil were not included. Many authors have proven the great impact of the content of SARA on the asphaltene precipitation phenomenon.

Later, Ahmadi and Shadizadeh (2012) presented a feed-forward ANN optimized by particle swarm optimization (PSO) as an AI modeling tool to predict asphaltene precipitation due to natural depletion. The main advantage of this strategy was the combination of the local searching ability of the backpropagation (BP) method with the global searching ability of PSO. The PSO-ANN model was applied to the experimental data from one of the northern Persian Gulf oil field. Although the average relative absolute deviation between the predictions from PSO–ANN and BP–ANN model and the experimental data was found to be less than 4%, the first one predicted asphaltene precipitation more accurately. Considering the NN structure was determined manually, the authors recommended the application of a substitute method for the ANN structure optimization.

Ahmadi and Golshadi (2012) presented an enhanced version of the previous investigation. This new ANN model was optimized by hybrid GA and PSO (HGAPSO). Due to the integrating global search and the local search abilities of GA and PSO, respectively, the HGAPSO strategy demonstrated higher robustness in optimization issues. The HGAPSO-ANN model soundness was ascertained by the comparison between the predictions and corresponding experimental data.

Taking into account that previous models were developed with dead oil and with not consider real pressure and temperature conditions, Rasuli *et al.* (2012) presented the first ANN approach based on live oil PVT

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

properties in the reservoir, production string, or flow-line conditions. Different neural models were designed and trained with several solution algorithms to find the best predictor for target samples. The best ANN structure could predict asphaltene deposition in a PVT cell with good accuracy. Hence, this study proved the ANN effectiveness against expensive and time-consuming experimental procedures, like traditional PVT tests.

In 2013, Salahshoor *et al.* proposed an adaptive neuro-fuzzy approach to predict the amount of asphaltene precipitation from relative permeability reduction and pressure drop. Different flow dynamic test scenarios, at reservoir conditions, were organized to perform on sandstone as well as carbonate rocks the practical exploration of the asphaltene deposition process. This model was the first one capable of describing the alteration of relative permeability due to asphaltene deposition during natural depletion and EOR techniques. Accordingly, two adaptive neuro-fuzzy models were sequentially developed in a nonlinear affine-type configuration to investigate the effect of multiple variables and parameters on asphaltene deposition based on the most recent input-output data. The outcomes demonstrated model efficacy to estimate asphaltene deposition in actual circumstance before facing the problems which could be serious at field scale.

Subsequently, a hybrid ICA-ANN model was developed to predict the amount of precipitated asphaltene and bubble point pressure for different light oil samples by Zendehboudi *et al.* (2013). This work was an enhanced version of the previous one developed by Ahmadi (2011) because it included experimental work, thermodynamic framework, and the ANN model combined with ICA. The predictive performance of the developed smart technique was significantly better than the numerical fluid model, one BP-ANN model, and one scaling equation.

One year later, Zendehboudi *et al.* (2014) developed one of the most complete studies about the understanding of asphaltene precipitation mechanism. Several static and dynamic experiments, with and without CO_2 injection, were conducted to investigate the effects of temperature, pressure, pressure drop, dilution ratio, and mixture compositions on asphaltene precipitation and deposition. Smart techniques (ANNs) linked with evolutionary algorithms (e.g., ICA and PSO) were also applied to the experimental results attained for both cases: (a) dynamic tests to simulate the production process and (b) statistic experiments for sensitivity analysis. The results showed that the ICA-ANN model attained more reliable outputs compared with PSO-ANN, the conventional ANN, and scaling models. An additional statistical investigation indicated that pressure and temperature are the most important contributing parameters in the static test runs. However, pressure drop, and temperature had a higher impact, among the important process variables in asphaltene deposition, during dynamic experimental work.

In 2016, the first major product of this research process was published (Loaiza and Quintero, 2016; Sepúlveda *et al.*, 2016). Neuross-Asph model, an ANN trained and tested with 58 PVT experiments, was constructed with two different algorithms: Self-organizing maps (SOM) and Feed-forward. The main advantage of this model was the use of a few accessible data as input parameters. The results showed good agreement with experimental data, which confirmed that the NN structure identified the relationships that have more influence on asphaltene precipitation.

To overcome the limitations of the previous models, which had not distinguished between different oil types, Alimohammadi *et al.* (2017) proposed a comprehensive ANN model to estimate the weight percent of precipitated asphaltene in three oil types: light, medium, and heavy. An MLP-NN with five (5) input variables and one (1) output variable was utilized in this study. Furthermore, Levenberg-Marquardt was employed as the training algorithm. The predictions of the proposed model were compared with others such as Flory-Huggins, Modified Flory–Huggins, and one scaling equation. The results proved that the ANN approach provides suitable predictions for different oil types over a wide range of pressure, which is a difficult task for most traditional techniques.

Through an adaptive neuro-fuzzy inference system (ANFIS), Soodbakhsh Taleghani *et al.* (2017) estimated the amount of precipitated asphaltene as a function of temperature, dilution ratio, and molecular weight of different n-alkanes. A PSO algorithm was used to optimize the membership function parameters of the network. Predictions of the intelligent model were compared with previous correlations and one new scaling model proposed by the authors. The obtained values of R^2 for the ANFIS and scaling models were 0.9912 and 0.9862, respectively.

Hosseini-Dastgerdi and Jafarzadeh-Ghoushchi (2019) presented the most recent advancements in the implementation of ANN to evaluate the asphaltene precipitation. The simultaneous effects of many operative parameters were investigated using the combination of ANN and response surface methodology (RSM). The optimized model topology was set with nine (9) inputs parameters: temperature, pressure, oil saturation pressure at reservoir temperature, oil API gravity, the weight percentage of asphaltene and resin in oil, the molar percentage of injected methane, carbon dioxide, and nitrogen. Besides, seven (7) hidden neurons and one (1) output parameter (asphaltene precipitation percentage) were established. RSM was applied to investigate the simultaneous effects of variables on asphaltene destabilization behavior. Results showed that the combination of ANN and RSM can help to appropriately evaluate the simultaneous effects of different parameters on asphaltene precipitation as well as to predict the optimum condition to reach minimum asphaltene precipitation.

This work represents a novel way to approximate the AOP using different AI techniques. The primary use of this method is to reliably guarantee the dependence between the process factors and the outcome of the selected data of asphaltene precipitation experiments. This

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

methodology for predicting asphaltene precipitation has been a long-term research process. However, as input data, the most relevant parameters on the asphaltene destabilization phenomenon have always been considered. The last novelty of this work is using more than 300 experiments with an extensive range of changes that were introduced into ANN, to train the network and generate the model. Well training depends on the number of input experiments. More input experiments involved, the better the capability to approximate a real system (Hosseini-Dastgerdi and Jafarzadeh-Ghoushchi, 2019).

The structure of this paper is as follows. First, modeling strategy will be introduced followed by a brief description of the proposed neural network model. Then, the required steps to train the ANN model will be presented. Finally, the results of ANN will be evaluated and some of them will be compared to those of one scaling model.

2. MODEL DEVELOPMENT

2.1 Research Context

ANN is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system (Rasuli *et al.*, 2012). According to Hornik *et al.* (1990), ANNs are parallel information processing methods which can express complex and nonlinear relationship through number of input–output training patterns from the experimental data. ANN provides a non-linear mapping between inputs and outputs by its intrinsic ability (Ahmadi, 2011). In this sense, "Comportamiento de Fases (COFA)" research group of Universidad Surcolombiana has led the development of an accurate and robust intelligent model to predict AOP of oil reservoirs. Research process has been carried out in 3 different stages. Some details of AI models developed so far are shown in Figure-1.

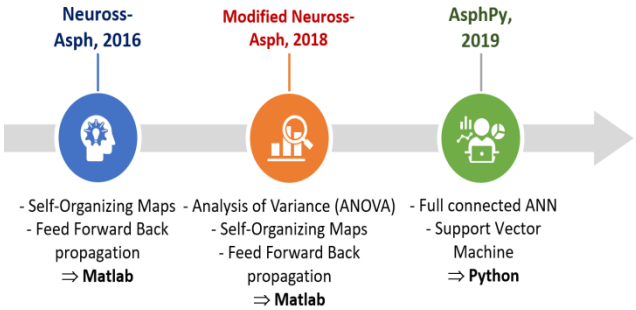


Figure-1. Some details of research process stages.

The first proposed ANN model, called Neuross-Asph, was published in 2016. ANN architecture was constructed in MATLAB (R2014b). Two self-organizing maps (SOM) and one feed-forward with a backpropagation algorithm were the ANN types selected. Specifics about the methodology, as well as the results of training/test procedures, can be found in Loaiza and Quintero (2016) and Sepúlveda *et al.* (2016). Later, a modified version of the Neuross-Asph model was introduced in 2018. An extensive statistical analysis (ANOVA) was performed in order to determine the most appropriate ANN input parameters. More information about the extra-process and its impact on the ANN predictions can be found in Olave and Quesada (2018). Finally, a new ANN version was developed in 2019. In this stage, a fully connected NN, called AsphPy, was implemented to predict the onset pressure of several crude oil samples. This intelligent model was implemented in Python (Python 3.7.0). Detailed information about the new ANN structure, as well as the updated model predictions, can be found in Perdomo and Arce (2019). It is worth mentioning that this paper contains the most remarkable and updated outcomes of the whole research process.

2.2 Data Collection

Generally, data mining is the process of analyzing data from different perspectives and summarizing it into useful information that can be used for several purposes (Rasuli *et al.*, 2012). Consequently, database is the most important part of data mining operations (Rasuli *et al.*, 2012). To construct the database for this research, all PVT properties and asphaltene deposition data of published works and field information were collected. These data were classified, categorized, and prepared to be fed into ANN after analysis.

Due to asphaltene precipitation involves several thermodynamic parameters, it was mandatory to take into consideration the ones related to pressure, temperature, and crude oil composition conditions. The experimental data collection covered all the effective parameters and formed the input data bank. Extrapolation procedures were carried out to the original database to get more points.

First model was developed with a data bank compound of only 58 PVT experiments. Meanwhile, modified Neuross-Asph model was constructed with 232 new points. Finally, the most recent ANN model was trained with 345 PVT experiments, which allowed us to construct a more reliable and robust intelligent model.

A common question in NN is, "what data sample size (or ratio of the total data) is required to train the network?" No simple procedure has been proposed yet to answer this question (Zendehboudi *et al.*, 2013). Therefore, considering a reasonable ratio for this selection can guarantee the validity of both the training and testing processes. For the predictions presented in this paper, 276 datapoints (80%) were randomly selected to train the ANN model and the remaining data, 69 datapoints, (20%) were used to test the prediction capability and accuracy of AsphPy model.

2.3 Input Effective Parameters

ANNs are relatively insensitive to data noise, as they can determine the underlying relationship between model inputs and outputs, resulting in good generalization ability. They can make extensive use of prior knowledge, and they are not limited by the assumptions of the underlying model (Ahmadi & Golshadi, 2012).

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

Considering that pressure, temperature, or composition variations can change stability of an oil sample, two main parameter groups affecting asphaltene stability were identified: fluid characteristics and reservoir properties (Lei *et al.*, 2010; Piroozan *et al.*, 2010). Consequently, temperature, pressure, saturation pressure at reservoir temperature, SARA analysis, and oil composition, were selected as input effective parameters. Table-1 shows required data to be fed into Asphy ANN model of 3 random samples. Table-2 shows the range of input parameters selected in AsphPy model.

Table-1. Properties of three different crude oils tested into AsphPy ANN model.

	Sample 1	**Sample 2**	**Sample 3**
T (°F)	188	224	116
P_b (psi)	2769	1963	748
P_{onset} (psi)	4146	2504	11094
Saturated (% wt.)	65.50	66.30	59.30
Aromatic (% wt.)	19.40	25.60	37.00
Resin (% wt.)	6.40	5.40	2.70
Asphaltene (% wt.)	8.70	2.80	0.90
N_2 (% mol)	0.06	0.16	1.41
CO_2 (% mol)	2.45	1.94	3.95
H_2S (% mol)	0.59	0.00	6.37
C_1 (% mol)	38.65	33.6	41.95
C_2 (% mol)	6.66	7.56	8.74
C_3 (% mol)	5.33	6.74	5.21
C_4 (% mol)	3.93	6.58	3.45
C_5 (% mol)	2.75	5.18	4.39
C_6 (% mol)	4.67	5.05	2.41
C_{7+} (% mol)	34.92	32.63	22.12

Input and output data were normalized before starting the modeling process to avoid any false influence of factors with higher orders of magnitude (Zendehboudi *et al.*, 2013). This procedure was carried out using "MinMaxScaler" which is a function from the Scikit-Learn library of Python.

2.4 Structure of the Artificial Neural Network

In general, developing a reliable and robust network strongly depends on the appropriate data preprocessing, proper architecture selection, and right network training choice (Zendehboudi *et al.*, 2013). As shown before, data preprocessing was carefully done based on the selection of the most appropriate input parameters, which influence asphaltene precipitation, as well as the normalization process which allowed to handle all parameters in a specific range. Architecture selection process will be explained now while network training choice will be presented in the Results and Discussion section.

Table-2. Ranges of input parameters selected in AsphPy ANN model.

	Max	**Min**
T (°F)	260	116
P_b (psi)	5692	214
Saturated (% wt.)	75.60	24.80
Aromatic (% wt.)	68.00	11.60
Resin (% wt.)	18.80	2.50
Asphaltene (% wt.)	16.30	0.20
N_2 (% mol)	7.86	0.01
CO_2 (% mol)	50.87	0.01
H_2S (% mol)	8.26	0.00
C_1 (% mol)	57.41	8.78
C_2 (% mol)	11.02	3.66
C_3 (% mol)	9.59	2.67
C_4 (% mol)	6.58	1.10
C_5 (% mol)	8.85	1.92
C_6 (% mol)	8.04	1.76
C_{7+} (% mol)	57.77	16.52

Final ANN structure depends on the number and type (input, output or hidden) of layers to be used, the interconnection scheme between layers and between layers-neurons, and the number of neurons. In general, the number of neurons and hidden layers depends on the complexity and nonlinearity of the problem. As in nature, the network function is determined largely by the connections between elements because each neuron is connected to input and output by a corresponding weight.

Different topologies are available to be used according to the purpose; however, the selection of the most appropriate one rests on the specific application. Indeed, ANN robustness is thanks to its knowledge-learning basis. ANN topologies offer the possibility to learn, classify and generalize data sets whose relationship corresponds to the description or association with properties of dynamic systems. In this case, asphaltene precipitation due to thermodynamic properties alterations represents the system.

The first two developed models, Neuross-Asph (2016) and modified Neuross-Asph (2018), were constructed with two topologies: self-organizing maps (SOM) and feed-forward neural network (FNN). SOM network, introduced by Teuvo Kohonen (Kohonen, 1990), is an evolutionary algorithm of unsupervised competitive learning that identifies common features, regularities, correlations, or categories into data. Mapping a high-dimensional data manifold onto a regular, low-dimensional (typically 2D) grid is the main function of

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

SOM architecture. SOM model is made up of one input layer (with N neurons, one for each input variable) and one output layer (with M neurons). Figure-2 shows a schema of a typical SOM network.

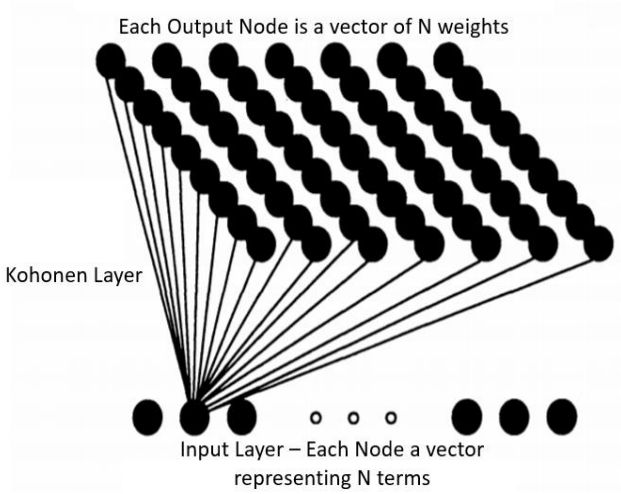


Figure-2. Typical architecture of SOM networks. Taken from Lin *et al.*, (1999).

The FNN, which is designed with one input layer, one output layer and hidden layers, is the most common architecture used in engineering applications (Zendehboudi *et al.*, 2013). In a FNN model, the information or signals propagate only in one direction, from input to output. A three-layered FNN with a back-propagation algorithm can approximate any nonlinear continuous function to an arbitrary accuracy (Brown and Harris, 1994; Hornik *et al.*, 1989). Feed-forward topology incorporates a supervised learning process and a sigmoidal-type activation function. Typical architecture of the FNN is presented in Figure-3.

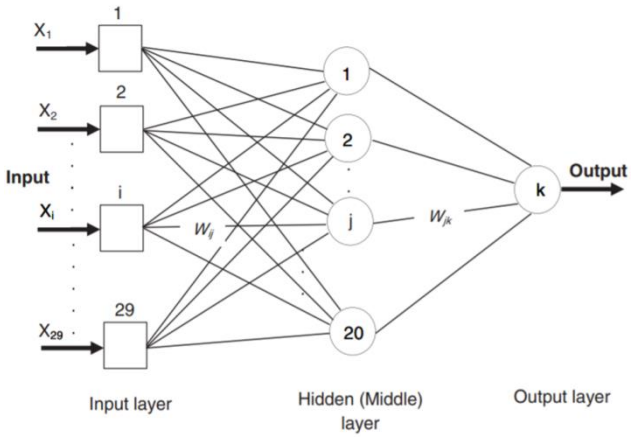


Figure-3. Diagram of the typical structure of FNNs. Taken from El-Gohary *et al.*, (2017).

The final model was constructed with a fully connected architecture, where all nodes (or neurons) in one layer, are connected to the neurons in the next layer. Thus, all ANN connections are connected to the node of the output layer. Analog to human brain behavior, all neurons make synapses between them in a fully connected ANN. Synapse is a space between a neuron and another node (neuron or not), where different things are continually happening. Physically it is a separation, but functionally is a connection that transfers information from one node to another (Perdomo and Arce, 2019).

The structure of AsphPy model is shown in Figure-4. ANN topology consists of three stages. The first one involves the data input layer and contains N number of nodes, each node introducing non-linearity through an activation function. Generally, activation function is considered the core of any ANN model due to it decides whether the neuron should fire depending on its inputs (Alimohammadi *et al.*, 2017). The second stage has N number of intermediate or hidden layers; each one has N number of nodes with its respective activation function. Finally, AsphPy model contains one output layer with only one node (asphaltene onset pressure). This architecture was constructed using the Python Karas library.

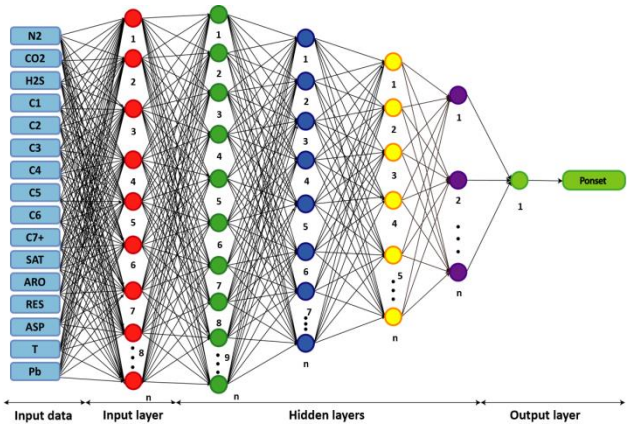


Figure-4. Architecture of the artificial neural network used in AsphPy model.

Although the sigmoidal is the classical nonlinearity function in fully connected networks, in this work a linear function was selected. Rectified Linear Unit (ReLU) (Khan *et al.*, 2019) activation function worked better to AOP predictions. Additionally, Adam (Kingma and Ba, 2014) optimizer was used to optimize the weights during the training process.

3. RESULTS AND DISCUSSIONS

As mentioned in the previous section, different ANN have been used to estimate asphaltene onset pressure from the experimental data that were obtained from previous studies reported in the literature and real-field information collected from different sources. Based on a trial-and-error procedure and the training set of data, optimum values of the required network parameters were found including one input layer (50 neurons) with an input dimension equal to sixteen (16) corresponding to the input effective parameters: N_2 (molar percentage of nitrogen), CO_2 (molar percentage of carbon dioxide), H_2S (molar percentage of hydrogen sulfide), C_1 (molar percentage of

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

methane), C_2 (molar percentage of ethane), C_3 (molar percentage of propane), C_4 (molar percentage of butane), C_5 (molar percentage of pentane), C_6 (molar percentage of hexane), C_{7+} (molar percentage of heptane and heavier components), SAT (weight percentage of Saturated), ARO (weight percentage of Aromatic), RES (weight percentage of Resin), ASP (weight percentage of Asphaltene), T (temperature) and Pb (saturation pressure at reservoir temperature). Seven (7) hidden layers (with 45, 40, 35, 30, 25, 20, 15 neurons, respectively) and one output layer with only one neuron (AOP).

AsphPy ANN was trained by performing optimization of weights for each node interconnection and bias terms until the output values at the output layer neurons were as close as possible to the actual outputs (Ahmadi, 2011). The trained network was then used to apply what it has learned to approximate or predict the corresponding output (Ahmadi and Golshadi, 2012). The performance of AsphPy model was determined through calculating some important statistical parameters such as root Mean Square Error (MSE) and the coefficient of determination (R^2). The mathematical definitions of the parameters are defined in Equation (1) and Equation (2).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(k_{i,\exp} - k_{i,pred}\right)^2 \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(k_{i,\exp} - k_{i,pred}\right)^2}{\sum_{i=1}^{N}\left(k_{i,\exp} - k_{mean}\right)^2} \quad (2)$$

where N is the number of data, $k_{i,\exp}$ is the *i-th* experimental AOP, $k_{i,pred}$ is the *i-th* predicted AOP by the neural network model, k_{mean} is the average value of the experimental AOP data. A neural network model is acceptable if the MSE parameter is as small as possible and if R^2 approaches unity (Hosseini-Dastgerdi and Jafarzadeh-Ghoushchi, 2019).

One of the challenges during the determination of optimal structure or topology of the NN was the number of epochs. One epoch is when the entire dataset is passed forward and backward through the NN only once. Considering that a limited dataset was used, updating the weights with a single pass or one epoch was not enough. Unfortunately, there is no definitive answer about what the right numbers of epochs are. It depends on database diversity. Different numbers of epochs (between 500 and 200000) were tested for both training and testing stages. Table-3 contains the results of the network performance based on the statistical parameters described above. Only the performance of the testing process in terms of MSE is shown in Figure-5.

According to the data presented in Table-3 and Figure-5, an increase in the number of epochs from 500 to 10000 led to an increased in the accuracy simulation results obtained by AsphPy model. However, the performance of ANN had to be carefully analyzed for 50000, 100000, and 2000000 epochs. For the training process, with 50000 and 200000 epochs the network had the least MSE (8.12E-05 and 4.91E-05, respectively), as well as the highest R^2 (very close to unity). The results of the testing process showed that the best ANN performance was obtained with 100000 epochs. The MSE was the least (0.00171) and the R^2 was the highest (0.9665). Taking into account that the testing stage assessed the ability prediction of ANN, and for this study, ANN performance with 100000 epochs was the most successful for the training stage (MSE parameter was small and R^2 parameter approached unity), the optimum ANN topology required 100000 epochs. In this sense, ANN with 50000 epochs or less could be considered underfitting structures, and the ANN with 200000 epochs was an overfitting topology.

Table-3. Performance of AsphPy ANN Model for training and testing stages based on the number of epochs.

Epochs	Training		Testing	
	MSE	R^2	MSE	R^2
500	0.01029	0.8907	0.06674	0.4361
1000	0.00549	0.9470	0.04712	0.4956
5000	0.00107	0.9904	0.02172	0.7727
10000	0.00057	0.9957	0.00229	0.9438
50000	8.12E-05	0.9996	0.00287	0.9155
100000	0.00034	0.9970	0.00171	0.9665
200000	4.91E-05	0.9999	0.00286	0.9270

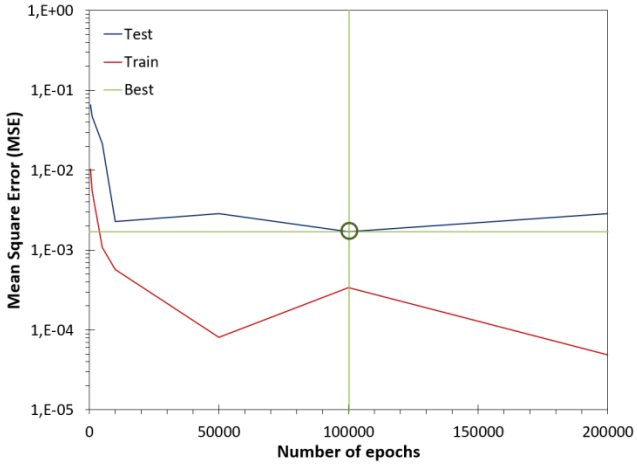
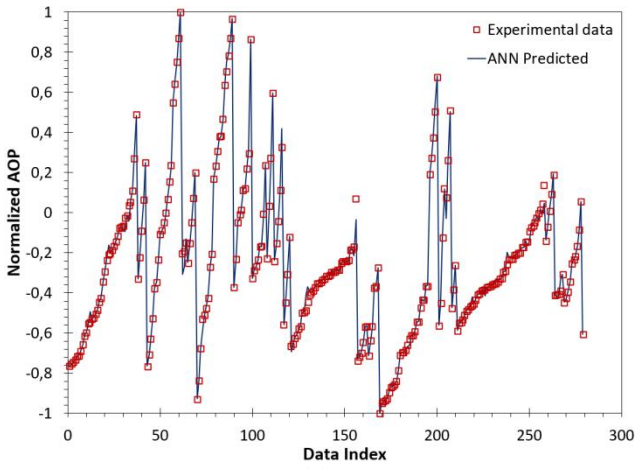


Figure-5. Performance plot of AsphPy ANN model in terms of the number of epochs.

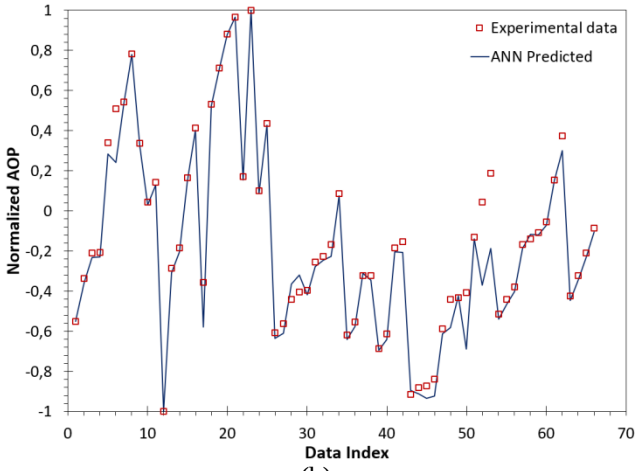
A comparison between predicted and measured normalized amounts of AOP at training and testing phases for AsphPy model is presented in Figure-6. Clearly, the outputs during the training stage are in good agreement with the database collected; see Figure-6(a). There is only one predicted AOP point with a high difference to the experimental point, specifically index 156. Experimental

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

reported AOP was 8177 psi, whereas the outcome of ANN was approximately 10% less, both measured at 116 °F. This could be explained because, usually at low temperatures, asphaltene onset pressure increases significantly (following an exponential tendency), which could make difficult the ANN predictions. Besides, the outputs of the model in the testing phase confirm an acceptable agreement with the experimental asphaltene onset data. As can be seen, Figure-6(b) presents only two cases (indexes 52 and 53) with a sharp difference between real and predicted AOP. Considering that more than 65 samples were tested by AsphPy model, an accurate prediction capability of the neural network could be validated.



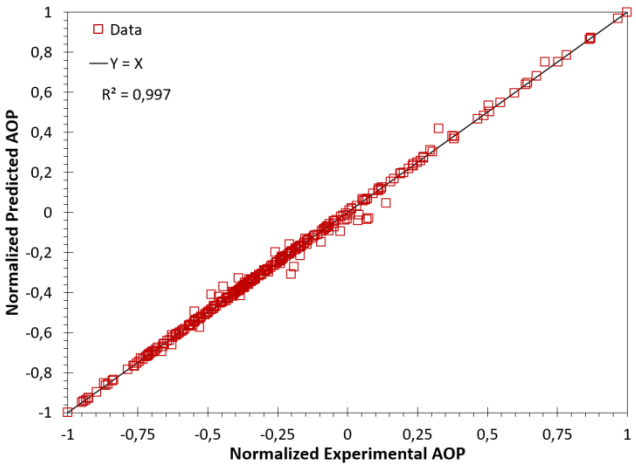
(a)



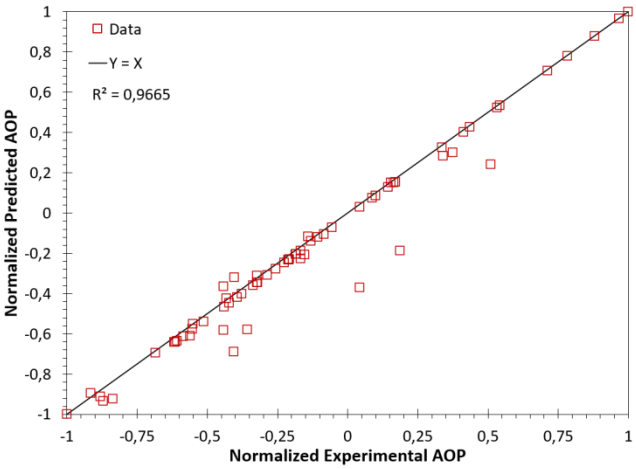
(b)

Figure-6. Comparison between measured and predicted asphaltene onset pressure (AsphPy model): (a) training stage and (b) testing stage.

In Figure-7, the measured amounts of AOP along with the predicted values using AsphPy-ANN network in the form of scatter diagrams are presented. Generally, based on correlation coefficient (R^2) value, for R^2 value greater than 0.9 indicates a very satisfactory model performance, while for R^2 value in the range of 0.8-0.9 signifies a good efficiency, and value less than 0.8 indicates an unsatisfactory model precision (Coulibaly and Baldwin, 2005). Consequently, the parameter R^2=0.9970 for training process and R^2=0.9665 for testing stage validated very good accuracy of the developed model.



(a)

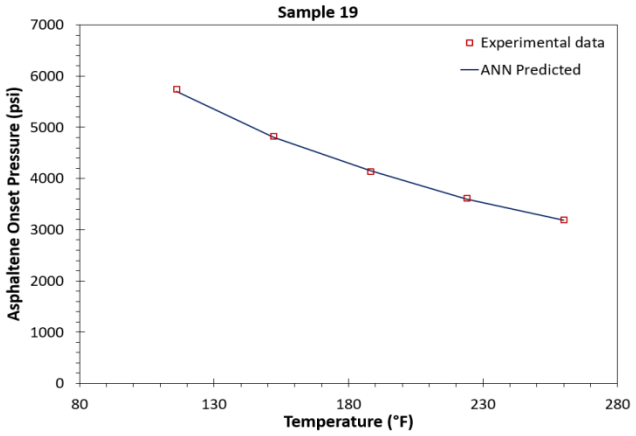


(b)

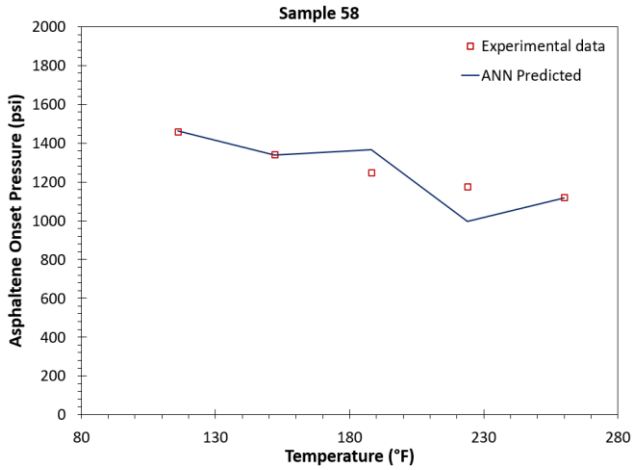
Figure-7. Scatter plot of predicted values of AOP versus real data based on AsphPy model: (a) training stage and (b) testing stage.

Figure-8 shows the experimental and predicted values of asphaltene onset for two samples at different temperatures. As can be seen, in Sample 19 all values are, practically, the same. On the other hand, in Sample 58 can be identified some differences between experimental data and the predictions obtained by AsphPy ANN model. In fact some of the cases with the highest differences between experimental data and predictions of ANN showed in Figure-6 belongs to Sample 58 (Figure-8(b)). For example, at 360 °F, experimental AOP is 1249 psi, but the network overestimated this value by 9.5% approximately. At 380 °F, the corresponding experimental value was 1177 psi, but the predicted AOP was 995 psi. That is, the ANN underestimated the value by 15%, approximately. This could be explained by the fact that the weight percentage of Aromatics of Sample 58 was the highest value among all data points used during the testing process, which could affect the prediction ability of the AsphPy model.

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.



(a)



(b)

Figure-8. Comparison between measured and predicted AOP, at different temperatures, based on AsphPy model: (a) Sample 19 and (b) Sample 58.

In this sense, it is worth remembering that asphaltene precipitation phenomenon is strongly non-linear in terms of process and thermodynamic parameters. There are several factors which could trigger it, and all of them could act simultaneously. However, considering that statistical parameters for training and testing stages showed excellent results, it could be considered that AsphPy network can accurately determine the AOP in oil reservoirs.

In addition to the comparison of predictions made by the above ANN approach, an empirical equation based on the model developed by Fahim (2007) is also presented here to assess the effectiveness of the AsphPy model through a systematic statistical analysis. This scaling equation was developed to avoid the tuning process of some experimental Pressure Volume Temperature (PVT) data of the reservoir fluid required for conventional equations of state. This model was developed through 33 crude oil samples, primarily from the Middle East, at different temperatures.

Fahim's equation for AOP was applied to the database selected to test our ANN model. The predicted values versus the experimental data plotted in Figure-9 shows poor agreement. Moreover, the MSE obtained to be 0.0448, confirmed that Fahim's model was not capable of predicting the experimental data used in this work. This could be explained because some variables used in the empirical method had a limited range compared with AsphPy model. For example, molar percentage of nitrogen, molar percentage of carbon dioxide, molar percentage of hydrogen sulfide, molar percentage of pentane, molar percentage of hexane, and molar percentage of heptane and heavier components.

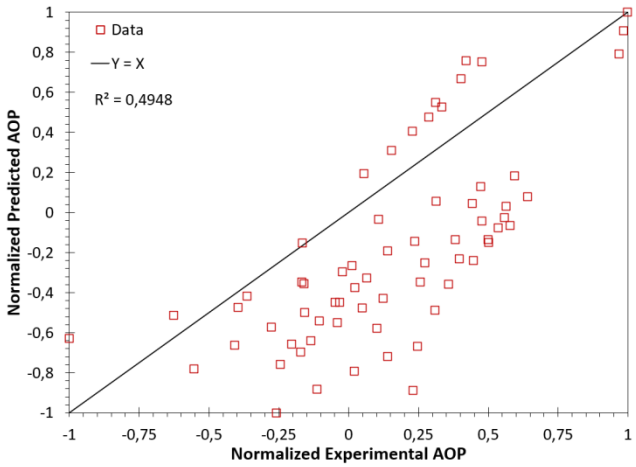


Figure-9. Scatter plot presenting predicted AOP versus the experimental data based on empirical equation model proposed by Fahim (2007).

CONCLUSIONS

A long-term research process was carried out to develop neural network models to predict the asphaltene onset pressure in oil reservoirs. A comprehensive bank of experimental data borrowed from the open literature, along with a set of allowed field data, was used to construct ANN topologies. Based on results obtained from this research, following main conclusions can be drawn:

a) Three different models have been developed to estimate AOP: Neuross-Asph (2016), Modified Neuross-Asph (2018), and AsphPy (2019). This paper presented the outcomes of the most recent model. The proposed neural network structure described in this work was determined manually by trial-and-error procedure based on two statistical criteria (MSE and R^2). It was demonstrated the high-performance of AsphPy ANN model to predict AOP during training and testing stages.

b) The predictive performance of the developed smart technique is significantly better than one of the available scaling equation approach (the empirical equation proposed by Fahim). The Mean Square Error between AsphPy model predictions and the experimental data was 0.00034 and 0.00171 during training and testing stages, whereas the scaling approach was 0.0448.

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

c) Asphaltene precipitation phenomenon is strongly non-linear in terms of process and thermodynamic parameters. Consequently, to develop an accurate model, temperature, pressure, saturation pressure at reservoir temperature, SARA analysis, and oil composition, were selected as input effective ANN parameters.

d) The high prediction capability of AsphPy model can be explained by the wide range of input parameters selected to develop AsphPy ANN model. Clearly, the smart technique presented in this work had the ability to estimate AOP in oil reservoirs with a wide range of properties in terms of fluid composition, reservoir condition, and content of asphaltene.

ACKNOWLEDGEMENTS

The authors thank "Comportamiento de Fases (COFA)" research group of Universidad Surcolombiana and the inter-administrative agreement among CMG-CPIP-Universidad Surcolombiana for all support during the development of this research.

REFERENCES

Abedini A., Ashoori S. & Saki Y. 2010. Application of neural network model for prediction of asphaltene precipitation. Trinidad and Tobago Energy Resources Conference.

Ahmadi M. A. 2011. Prediction of asphaltene precipitation using artificial neural network optimized by imperialist competitive algorithm. Journal of Petroleum Exploration and Production Technology, 1(2-4): 99-106. doi:10.1007/s13202-011-0013-7

Ahmadi M. A. & Golshadi M. 2012. Neural network based swarm concept for prediction asphaltene precipitation due to natural depletion. Journal of Petroleum Science and Engineering, 98-99, 40-49. doi: 10.1016/j.petrol.2012.08.011

Ahmadi M. A. & Shadizadeh S. R. 2012. New approach for prediction of asphaltene precipitation due to natural depletion by using evolutionary algorithm concept. Fuel, 102, 716-723. doi:10.1016/j.fuel.2012.05.050

Akbarzadeh K., Hammami A., Kharrat A., Zhang D., Allenson S., Creek J., Kabir S., Jamaluddin A., Marshall A., Rodgers R. & Mullins O. 2007. Asphaltenes-problematic but rich in potential. Oilfield Review, 19(2): 22-43.

Alimohammadi S., Sayyad Amin J. & Nikooee E. 2017. Estimation of asphaltene precipitation in light, medium and heavy oils: experimental study and neural network modeling. Neural Computing and Applications, 28(4): 679-694. doi:10.1007/s00521-015-2097-3

Ashoori S., Abedini A., Abedini R. & Nasheghi K. Q. 2010. Comparison of scaling equation with neural network model for prediction of asphaltene precipitation. Journal of Petroleum Science and Engineering, 72(1-2): 186-194. doi:10.1016/j.petrol.2010.03.016

Brown M. & Harris C. J. 1994. Neurofuzzy adaptive modelling and control. Prentice-Hall, Englewood Cliffs.
Coulibaly P. & Baldwin C. K. 2005. Nonstationary hydrological time series forecasting using nonlinear dynamic methods. Journal of Hydrology, 307(1-4): 164-174. doi:10.1016/j.jhydrol.2004.10.008

El-Gohary K. M., Aziz R. F. & Abdel-Khalek H. A. 2017. Engineering approach using ANN to improve and predict construction labor productivity under different influences. Journal of Construction Engineering and Management, 143(8).

Fahim M. A. 2007. Empirical equations for estimating ADE of crude oils. Petroleum Science and Technology, 25(7): 949-965. doi:10.1080/10916460500526981

Garrouch A. A. & Al-Ruhaimani F. A. 2005. Modeling Permeability Reduction Caused by Asphaltene Deposition. SPE Annual Technical Conference and Exhibition. doi:10.2118/96697-MS

Gharagheizi F., Eslamimanesh A., Sattari M., Mohammadi A. H. & Richon D. 2012. Corresponding states method for evaluation of the solubility parameters of chemical compounds. Industrial & Engineering Chemistry Research, 51(9): 3826-3831. doi: 10.1021/ie202633m

Hirschberg A., DeJong L. N. J., Schipper B. A. & Meijer J. G. 1984. Influence of temperature and pressure on asphaltene flocculation. Society of Petroleum Engineers Journal, 24, 283-293. doi:10.2118/11202-PA

Hornik K., Stinchcombe M. & White H. 1989. Multilayer feedforward networks are universal approximators. Neural Networks, 2(5): 359-366. doi: 10.1016/0893-6080(89)90020-8

Hornik K., Stinchcombe M. & White H. 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Networks, 3(5): 551-560. doi: 10.1016/0893-6080(90)90005-6

Hosseini-Dastgerdi Z. & Jafarzadeh-Ghoushchi S. 2019. Investigation of Asphaltene Precipitation using Response Surface Methodology Combined with Artificial Neural Network. Journal of Chemical and Petroleum Engineering, 53(2): 153-167. doi:10.22059/jchpe.2019.261438.1238

Hosseini‐Dastgerdi Z., Tabatabaei‐Nejad S. A. R., Khodapanah E. & Sahraei E. 2015. A comprehensive study on mechanism of formation and techniques to diagnose asphaltene structure; molecular and aggregates: a

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

review. Asia‐Pacific Journal of Chemical Engineering, 10(1): 1-14. doi:10.1002/apj.1836

Hu Y. F., Chen G. J., Yang J. T. & Guo T. M. 2000. A study on the application of scaling equation for asphaltene precipitation. Fluid Phase Equilibria, 171(1-2): 181-195. doi:10.1016/S0378-3812(00)00348-4

Jewell D. M., Weber J. H., Bunger J. W., Plancher H. & Latham D. R. 1972. Ion-exchange, coordination, and adsorption chromatographic separation of heavy-end petroleum distillates. Analytical Chemistry, 44(8): 1391-1395. doi:10.1021/ac60316a003

Kawanaka S., Park S. J. & Mansoori G. A. 1991. Organic Deposition From Reservoir Fluids: A Thermodynamic Predictive Technique. SPE Reservoir Engineering, 6(2): 185-192. doi: 10.2118/17376-PA

Khamehchi E., Behvandi R. & Rashidi F. 2012. Prediction of Bubble Point Pressure & Asphaltene Onset Pressure During CO2 Injection Using ANN & ANFIS Models. Journal of Petroleum Science and Engineering, 1(2): 35-45. doi:10.22078/JPST.2011.44

Khan F. N., Fan Q., Lu C. & Lau A. P. T. 2019. An optical communication's perspective on machine learning and its applications. Journal of Lightwave Technology, 37(2): 493-516. doi:10.1109/JLT.2019.2897313

Kingma D. P. & Ba J. 2014. Adam: A method for stochastic optimization. 3rd International Conference for Learning Representations.

Kohonen T. 1990. The self-organizing map. Proceedings of the IEEE, 78(9): 1464-1480. doi:10.1109/5.58325

Lei H., Pingping S., Ying J., Jigen Y., Shi L. & Aifang B. 2010. Prediction of asphaltene precipitation during CO2 injection. Petroleum Exploration and Development, 37(3): 349-353. doi:10.1016/S1876-3804(10)60038-9

Leontaritis K. J. 1997. PARA-based (paraffin-aromatic-resin-asphaltene) reservoir oil characterizations. International Symposium on Oilfield Chemistry.

Lin C., Chen H. & Nunamaker J. F. 1999. Verifying the proximity and size hypothesis for self-organizing maps. Journal of Management Information Systems, 16(3): 57-70. https://www.jstor.org/stable/40398444

Loaiza C. S. & Quintero S. 2016. Desarrollo de un modelo para la determinación de la presión de inicio de precipitación de asfaltenos (onset) en yacimientos de petróleo. Universidad Surcolombiana.

Mansoori G. A. 1997. Modeling of asphaltene and other heavy organic depositions. Journal of Petroleum Science and Engineering, 17(1-2): 101-111. doi:10.1016/S0920-4105(96)00059-9

Mohammadi A. H. & Richon D. 2008. Estimating Onset of Precipitation of Dissolved Asphaltene in the Solution of Solvent + Precipitant Using Artificial Neural Network Technique. The Open Thermodynamics Journal, 2(1): 82-88. doi:10.2174/1874396x00802010082

Olave N. F. & Quesada N. H. 2018. Determinación de la presión onset usando análisis estadístico anova en modelo de red neuronal. Universidad Suroclombiana.

Pazuki G. R. & Nikookar M. 2006. A modified Flory-Huggins model for prediction of asphaltenes precipitation in crude oil. Fuel, 85(7-8): 1083-1086. doi:10.1016/j.fuel.2005.10.005

Perdomo J. D. & Arce G. C. 2019. Determinación de la presión onset en yacimientos petrolíferos con redes neuronales implementadas en python basado en modelo estadístico. Universidad Surcolombiana.

Piroozan Z., Kharrat R. & Emamzadeh A. 2010. Study of asphaltene precipitation-deposition due to pressure depletion. Brazilian Journal of Petroleum and Gas. 4, 71-81.

Rassamdana H., Dabir B., Nematy M., Farhani M. & Sahimi M. 1996. Asphalt flocculation and deposition: I. the onset of precipitation. AIChE Journal, 42(1): 10-22. doi:10.1002/aic.690420104

Rasuli N., Khishvand M. & Naseri A. 2012. An artificial neural network approach to predict asphaltene deposition test result. Fluid Phase Equilibria, 329, 32-41. doi:10.1016/j.fluid.2012.06.001

Salahshoor K., Zakeri S., Mahdavi S., Kharrat R. & Khalifeh M. 2013. Asphaltene deposition prediction using adaptive neuro-fuzzy models based on laboratory measurements. Fluid Phase Equilibria, 337, 89-99. doi:10.1016/j.fluid.2012.09.031

Sepúlveda J. A., Mosquera V., Loaiza C. S. & Quintero S. 2016. Neural model development for the determination of asphaltene onset pressure in oil reservoirs. ARPN Journal of Engineering and Applied Sciences. 11(21).

Soodbakhsh Taleghani M., Saeedi Dehaghani A. H. & Shafiee M. E. 2017. Modeling of precipitated asphaltene using the ANFIS approach. Petroleum Science and Technology, 35(3): 235-241. doi:10.1080/10916466.2016.1252773

Vafaie-Sefti M. & Mousavi-Dehghani S. A. 2006. Application of association theory to the prediction of asphaltene deposition: Deposition due to natural depletion and miscible gas injection processes in petroleum

©2006-2021 Asian Research Publishing Network (ARPN). All rights reserved.

reservoirs. Fluid Phase Equilibria, 247(1-2): 182-189. doi:10.1016/j.fluid.2006.06.019

Zahedi G., Fazlali A. R., Hosseini S. M., Pazuki G. R. & Sheikhattar L. 2009. Prediction of asphaltene precipitation in crude oil. Journal of Petroleum Science and Engineering, 68(3-4): 218-222. doi:10.1016/j.petrol.2009.06.023

Zendehboudi S., Ahmadi M. A., Mohammadzadeh O., Bahadori A. & Chatzis, I. 2013. Thermodynamic investigation of asphaltene precipitation during primary oil production: Laboratory and smart technique. Industrial and Engineering Chemistry Research, 52(17): 6009-6031. doi:10.1021/ie301949c

Zendehboudi S., Shafiei A., Bahadori A., James L. A., Elkamel A. & Lohi A. 2014. Asphaltene precipitation and deposition in oil reservoirs - Technical aspects, experimental and hybrid neural network predictive tools. Chemical Engineering Research and Design, 92(5): 857-875. doi:10.1016/j.cherd.2013.08.001