Journal of Natural Gas Science and Engineering 22 (2015) 661–669

Contents lists available at ScienceDirect

Journal of Natural Gas Science and Engineering

journal homepage: www.elsevier.com/locate/jngse

Smart correlation of compositional data to saturation pressure

Amin Gholami [a], Mohammad Afshar [b], Parisa Bagheripour [b], Mojtaba Asoodeh [b,*], Mohsen Vaezzadeh-Asadi [b]

[a] *Abadan Faculty of Petroleum Engineering, Petroleum University of Technology, Abadan, Iran*
[b] *Department of Petroleum Engineering, Kharg Branch, Islamic Azad University, Kharg, Iran*

A R T I C L E I N F O

Article history:
Received 16 September 2014
Received in revised form
8 January 2015
Accepted 10 January 2015
Available online 14 January 2015

Keywords:
Saturation pressure
Crude oil
Improved neural network

A B S T R A C T

Saturation pressure is one of the foremost parameters of crude oil which plays a key role in petroleum calculations. Experimentally, determination of this parameter in laboratory is costly and labor demanding. In this study, an improved intelligent model based on neural network optimized with genetic algorithm-pattern search technique is proposed for building quantitative formulation between saturation pressure and compositional data, including temperature, hydrocarbon and non-hydrocarbon compositions of crudes, and heptane-plus specifications. Genetic algorithm-pattern search technique is embedded in neural network formulation for finding optimal weights and biases of neural network. A comparison among the proposed model and published models in literature reveals the superiority of our model in terms of better accuracy and higher generalization. Improved neural networked showed R-square of 0.9892 and MSE of 17,617.99 which concludes that it is a promising alternative for determination of saturation pressure which is able to eliminating expenses of laboratory measurements and significantly saving time. This study showed GA considerably enhanced performance of conventional neural networks.

1. Introduction

Saturation pressure is one of the important parameters of crude oil. It's defined as the maximum pressure at which first few molecules of gas is departed from the liquid and make a small gas bubble (Ahmed, 2000). Attention to determination of this parameter has considerable effect on accuracy of reservoir engineering computation. Due to being highly cost, time, and labor intensive, cheap and rapid prediction of saturation pressure is of great interest (Danesh, 1998). Therefore, a growing attention has been paid to saturation pressure prediction and a variety of predictive models are presented in the last few years (Standing, 1947; Lasater, 1958; Glaso, 1980; Al-Marhoun, 1988; McCain, 1991; Dokla et al., 1991; Macary and El-Batanoney, 1992; Petrosky and Farshad, 1993; Gharbi and Elsharkawy, 1997a, 1997b; McCain et al., 1998; Al-Shammasi, 1999; Boukadi et al., 1999; Gharbi et al., 1999; Velarde et al., 1999; Al-Marhoun and Osman, 2002; Goda et al., 2003; Elsharkawy, 2003; Malallah et al., 2006; Al-Sebakhy et al., 2007; Al-Sebakhy, 2009; AlQuraishi, 2009; Dutta and Gupta, 2010; Asadisaghandi and Tahmasebi, 2011; Moghadam et al., 2011; Olatunji et al., 2011; Bandyopadhyay and Sharma, 2011; Khoukhi, 2012; Ghafoori et al., 2012; Asoodeh and Bagheripour, 2012a; Rafiee-Taghanaki et al., 2013; Moghadasi et al., 2013; Asoodeh and Kazemi, 2013; Kazemi et al., 2013; Farasat et al., 2013; Bagheripour et al., 2014; Afshar et al., 2014; Gholami et al., 2014a; Ghanji-Azad et al., 2014; Talebi et al., 2014; Arabloo et al., 2014; Bagheripour and Asoodeh, 2014; Ahmadi et al., 2014; Shojaei et al., 2014). Thermodynamic models are conventional approaches for modeling of saturation pressure. However, their use is restricted owing to their requirement to splitting and characterizing the heavy fractions of oil as well as high associated forecasting error of them (Elsharkawy, 2003). Thus this group of models is not suitable for estimation of saturation pressure. Some researchers focus on models in which the value of saturation pressure is formulated as a function of production data, including reservoir temperature, solution gas-oil ratio, reservoir oil gravity, and gas relative density (Standing, 1947; Lasater, 1958; Glaso, 1980; Al-Marhoun, 1988; McCain, 1991; Dokla et al., 1991; Macary and El-Batanoney, 1992; Petrosky and Farshad, 1993; Gharbi and Elsharkawy, 1997a, 1997b; McCain et al., 1998; Al-Shammasi, 1999; Boukadi et al., 1999; Gharbi et al., 1999; Velarde et al., 1999; Al-Marhoun and Osman, 2002; Goda et al., 2003; Malallah et al., 2006; Al-Sebakhy et al., 2007; Al-

* Corresponding author.
E-mail address: asoodeh.mojtaba@gmail.com (M. Asoodeh).

http://dx.doi.org/10.1016/j.jngse.2015.01.017
1875-5100/

Sebakhy, 2009; Dutta and Gupta, 2010; Asadisaghandi and Tahmasebi, 2011; Moghadam et al., 2011; Olatunji et al., 2011; Khoukhi, 2012; Asoodeh and Bagheripour, 2012a; Rafiee-Taghanaki et al., 2013; Moghadasi et al., 2013; Asoodeh and Kazemi, 2013; Bagheripour et al., 2014; Afshar et al., 2014; Ghanji-Azad et al., 2014; Talebi et al., 2014; Arabloo et al., 2014; Shojaei et al., 2014). Although this group of models is valuable, they can predict the saturation pressure of crude oil of specific region which their data is employed for developing models. Recently, many researchers attempted to introduce models by which saturation pressure is correlated to compositional data (Elsharkawy, 2003; Bandyopadhyay and Sharma, 2011). Elsharkawy (2003) presented a correlation for determining the saturation pressure from composition data, including temperature, hydrocarbon and non-hydrocarbon compositions of crudes, and heptane-plus specifications (Elsharkawy, 2003). Bandyopadhyay and Sharma (2011) introduced semi analytic model for estimation of saturation from compositional data (Bandyopadhyay and Sharma, 2011). In addition of aforementioned empirical correlations, several studies are devoted to intelligent models for formulating saturation pressure to composition data (AlQuraishi, 2009; Kazemi et al., 2013; Farasat et al., 2013; Gholami et al., 2014a; Bagheripour and Asoodeh, 2014; Ahmadi et al., 2014). Although previous approaches are utilizable and have acceptable performance; the quest for developing more reliable models for quantitative estimation of saturation pressure is exist. As a result, this study proposes an improved strategy based on the integrating neural network with genetic algorithm for formulating the saturation pressure to compositional data, including temperature, hydrocarbon and non-hydrocarbon compositions of crudes, and heptane-plus specifications. Results of this study showed improved neural network excels in terms of accuracy and generalization to previous published models. Transplanting genetic algorithm in conventional neural network can significantly enhance accuracy of final prediction. Stochastic capability of genetic algorithm wipes out sticking of neural network in local minima and converges neural network to global minimum.

2. Improved neural modeling

Artificial neural networks (ANNs) are computer models that attempt to simulate specific functions of human nervous system (Asoodeh and Bagheripour, 2013). This is carried out through nonlinear neurons, arranged in parallel structures called layers. Each neuron carries a bias value in its body and is connected to neurons of subsequent layer by means of connecting weights. Optimal values for these weights and biases determine optimal formulation between inputs and output. A three layered ANN with n input and m hidden neurons follows succeeding formulation (Nikravesh et al., 2003; Asoodeh and Bagheripour, 2012b; Gholami et al., 2014b)

$$OHL_{m\times 1} = \frac{1}{1+\exp(-(IW_{m\times n}\times I_{n\times 1}+Ib_{m\times 1}))} \tag{1}$$

$$O_{NN} = OW_{1\times m}\times OHL_{m\times 1}+b_{1\times 1} \tag{2}$$

Where, IW is connecting weights matrix between input layer and hidden layer; I is input vector; Ibis hidden layer bias vector; OHL is hidden layer net output vector; OW is connecting weights matrix between hidden layer and output layer; b is output layer bias; and O_{NN} is predicted output from neural network. Previous works (Nikravesh et al., 2003; Asoodeh and Bagheripour, 2012b; Gholami et al., 2014b) indicated that performance function of neural network decreases more rapidly in direction proportional to negative of gradient:

$$-\frac{\partial E}{\partial w_{ij}} = \delta_j o_i \tag{3}$$

here, E is performance function; δ is gradient of performance function; o_i is net output of ith layer; and w_{ij} is connecting weights between ith and jth layers. In other words, modification of weights and biases follows succeeding equation.

$$w(t+1) = w(t) - \alpha_t g_t \tag{4}$$

Where, w(t+1) and w(t) refer to weights and biases of (t+1)th and tth iterations, respectively. α_t and g_t are learning rate and gradient in tth iteration, correspondingly. This equation proves that in local minima, where g_t is equal to zero no modification over weights and biases is done. Therefore, neural network is trapped in local minima. To improve neural network, computational formulation of neural network is subjected to stochastic optimization, which is able to evade local minima and converges to global minimum. Hybrid genetic algorithm-pattern search (GA-PS) tool is utilized as global optimization tool for extracting optimal values of connection weights and biases of neural network. For this purpose, following fitness function is introduced to GA-PS.

$$MSE = \frac{1}{N}\sum_{\ell=1}^{N}(O_{NN\,\ell} - T_\ell)^2 \tag{5}$$

Here, MSE is mean square error of neural network's prediction. T is target value (i.e., measured saturation pressure), and N is number of training data. GA-PS tool finds optimal values of connecting weights and biases such that mean square error of neural network reaches its global minimum. For more details about GA-PS, refer to Asoodeh and Bagheripour (2012b).

2.1. Genetic algorithm

Genetic algorithm is a stochastic search technique which is capable of finding global minimum of a function among so many local minima. It emulates the biological process of natural evolution along with survival of the fittest theory of Darwin for finding the best solution for a desired function. Firstly in genetic algorithm a population of randomly generated chromosomes (solutions) is gathered. Each chromosome (solution) is evaluated by means of the fitness function (desired function that its global minimum is of interest) and a score is assigned to the chromosome (solution). Subsequently all chromosomes are ranked according to their score and genetic operators are applied to top ranked chromosomes (called parent) to produce new chromosome (called offspring). In this way, a population of randomly generated chromosome generates a new population of better quality. During these generations some of top ranked chromosomes (elite) are kept unchanged and they are handed down from one population to the next population. The mentioned process continues until the desired chromosome is achieved. Three types of genetic operators, including crossover, mutation, and inversion are applied during these generations. Crossover permits the exchange of information among chromosomes in the population and provides the innovative capability of GA. Mutation ensures desirable diversity (Asoodeh et al., 2014a, 2014b, 2014c).

3. Input/output data space

High-quality data guarantee the reliability of the evaluation of data-mining methodologies. In current study, the experimental observation for 131 crude oil samples are taken from literature for