Journal of Natural Gas Science and Engineering 20 (2014) 8–15

Contents lists available at ScienceDirect

Journal of Natural Gas Science and Engineering

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

Application of soft computing approaches for modeling saturation pressure of reservoir oils

Roya Talebi [a], Mohammad M. Ghiasi [b, *], Hossein Talebi [f, g], Mehrdad Mohammadyian [f], Sohrab Zendehboudi [d], Milad Arabloo [c, *], Alireza Bahadori [e]

[a] Department of Petroleum Engineering, Islamic Azad University, Omidieh Branch, Omidieh, Iran
[b] National Iranian Gas Company (NIGC), South Pars Gas Complex (SPGC), Asaluyeh, Iran
[c] Young Researchers and Elites Club, North Tehran Branch, Islamic Azad University, Tehran, Iran
[d] Department of Chemical Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA
[e] School of Environment, Science and Engineering, Southern Cross University, Lismore, NSW, Australia
[f] Department of Chemical Engineering, Islamic Azad University, Gachsaran Branch, Gachsaran, Iran
[g] South Pars Gas Complex, Asaluyeh, Bushehr, Iran

ARTICLE INFO

Article history:
Received 27 February 2014
Received in revised form
27 April 2014
Accepted 30 April 2014
Available online 21 June 2014

Keywords:
Saturation pressure
PVT
Intelligence model
Radial basis function
Statistical analysis

ABSTRACT

Accurate determination of bubble pressure of reservoir fluid at reservoir conditions is one of the important parameter which is necessary for various calculations in petroleum engineering. This study presents two improved algorithms based on machine learning approaches for efficient estimation of saturation pressure of reservoir oil. To achieve the research purpose, a large data set, comprising of more than 750 crude oil samples with different composition and geographical origins, was collected from the literature for development of the models. The efficiency of the proposed models was tested against sixteen well-known empirical correlations. The proposed models show good performance in terms of accuracy with the lowest error percentage and highest R^2 values.

1. Introduction

PVT properties of reservoir oils are of crucial importance in all stages of petroleum engineering computations including material balance calculations, reserve estimates, inflow performance calculations, well test analysis, surface facility design, and reservoir simulation (Danesh, 1998; Dindoruk and Christman, 2004). These data can be measured by conducting a laboratory analysis of the reservoir samples. However, this sampling and subsequent analysis comprises considerable costs and time, which is undesirable. In situations where direct measurements are not available, other predictive techniques are used to estimate the PVT properties (Arabloo et al., 2013; Asoodeh and Bagheripour, 2012; Dindoruk and Christman, 2004; El-Sebakhy, 2009; Fayazi et al., 2014, 2013).

Reservoir oil's PVT properties such as bubble point pressure (P_b) are of primary importance in reservoir and production calculations (Arabloo et al., 2014; Asoodeh and Bagheripour, 2012; Danesh, 1998). The development of predictive models for calculation of these PVT properties has been the subject of extensive studies (Al-Shammasi, 1999; Arabloo et al., 2014; Frashad et al., 1996; Kartoatmodjo and Schmidt, 1994; Lasater, 1958; Standing, 1947; Valkó and McCain, 2003; Vazquez and Beggs, 1980; Velarde et al., 1997).

A look at a number of extensive models that have been developed for predicting PVT properties of reservoir oil samples since the early 1940s confirms the importance of the subject from both academic and industrial perspectives. Unfortunately, large numbers of the developed models/correlations are often limited and cannot be utilized universally due to dissimilar characteristics of fluids in each region. Moreover, most of the developed models have been developed based on limited data and limited range of conditions. Several graphical and empirical correlations for calculating both saturation pressure and oil formation volume factor have been proposed during the past seven decades. These models are basically based on supposition that saturation pressure (P_b) is a function of solubility of gas (R_s), oil gravity (API), gas gravity (γ_g), and reservoir temperature (T). For further information, Al-Shammasi (1999), Arabloo et al. (2014) and Mahmood and Al-Marhoun (1996)

* Corresponding authors.
E-mail addresses: mm.ghiasi@gmail.com (M.M. Ghiasi), milad.arabloo@gmail.com (M. Arabloo).

http://dx.doi.org/10.1016/j.jngse.2014.04.023
1875-5100/

provide an excellent review and in-depth analysis of various P_b correlations.

The recent development and success of applying various machine learning modeling approaches to solve various difficult engineering problems has drawn attention to their potential applications in the petroleum industry (Ghiasi et al., 2014; Rafiee-Taghanaki et al., 2013; Zendehboudi et al., 2012; Zendehboudi et al., 2013).

The objective of this study is to develop reliable computer-based models for the prediction of saturation pressure of crude oil systems. To achieve the research purposes, a large data bank covering wide ranges of fluid and experimental data is collected to construct and validate the models. Moreover, comparative studies are conducted between the developed models in this study and the existing correlations. Highlighting the contribution of the paper, our developed ANN model covers wider ranges of input data compared to the previous models. The importance level of the input variables was also determined in this study as analysis of variance (ANOVA) technique was employed for this investigation and parametric sensitivity analysis purposes. On the other hand, different target function with different input variables is considered in this study as a sample of ANN application in oil and gas engineering. It is important to note that much higher accuracy in predicting saturation pressure is attained while employing the developed smart model compared to the predictive models (e.g., correlations) which would be an asset for engineering and research activities in this area.

This paper is organized as follows. In the following section, the background of the MLP and RBF modeling approaches are discussed. Next, in Section 3 the methodology for development of predictive models is described. After that in Section 4, the accuracy of the newly developed models as well as previously published correlations is evaluated by means of different statistical and graphical quality measures. Finally, key findings of this study are presented in Section 5.

2. Artificial neural network (ANN) modeling

2.1. Multilayer perceptron networks

Multilayer perceptron (MLP) neural networks comprise of three different types of layers, namely, input layer, hidden layer(s), and output layer. A typical MLP network is shown in Fig. 1. Each layer is composed of some neurons as simple computing cells. Mathematically a neuron m could be represented by the following equations:

$$r_m = \sum_{i=1}^{n} (w_{mi} x_i + b_m) \quad (1)$$

$$y_m = f(r_m) \quad (2)$$

where $x_1, x_2, \ldots, x_n$ indicate the input signals; $w_{m1}, w_{m2}, \ldots, w_{mn}$ are synaptic weights of the neuron; r_m is the linear combiner output; b_m is the bias term; f is the activation function; and y_m is the neuron's output signal.

The number of neurons in the input and output layers are corresponded to the number of input and output data, respectively. The number of hidden layers and also neurons in them are optional and can be determined either intelligently or by trial and error to achieve the best performance (Ghiasi et al., 2013). Mean square error (MSE) indicates the performance of the developed network. In such networks, the error is back propagated through the network and the weights and biases are optimized through some iteration called epochs. The number of epochs should be such that the network neither undertrain nor overtrain. In the former, the network does not have enough time to complete the learning process. In the latter, the network does not learn but memorizes. This results in poor performance of network in prediction of test data set (Haykin, 1999).

To adjust the weights values a proper learning algorithm must be utilized. In this work, the back-propagation (BP) learning algorithm is employed. After feeding the data into the input layer of MLP network, the outputs will be computed at the last layer namely output layer. By employing the BP, the errors resulting from target and output values differences will propagate backward through the network. As a result, values of the weights will be adjusted so that the overall error minimizes. Detailed description of BP algorithm can be found elsewhere (Ghiasi et al., 2014; Yuan et al., 2014).

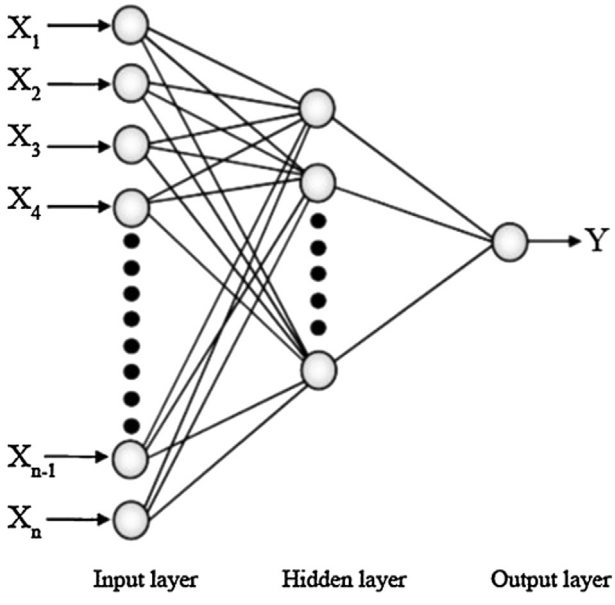

Fig. 1. Schematic representation of MLP neural network (Majidi et al., 2014).

2.2. Radial basis function networks

Both MLP and Radial Basis Function (RBF) networks have the same applications but different internal calculation structures. The main advantage of RBF networks is easy design that it has just three layers. They are capable of good generalization, high tolerance of input noises and ability of online learning (Santos et al., 2013). From the point view of generalization, RBFNs can respond very well to patterns that were not used for training (Hao et al., 2011).

RBF networks are neural networks based on localized basis functions and iterative function approximation (Dayhoff, 1990; Lowe and Broomhead, 1988; Zurada, 1992). Similar to MLPs, the RBF networks utilize supervise training technique and is a type of feed-forward neural networks. RBF has a simpler structure than MLP and training process is much faster. These features make RBF a popular alternative to the MLP. The origin of RBF is in performing exact interpolation of a set of data points in a multidimensional space (Powell, 1987). It is proved that RBF networks can be implemented by MLP networks with increased input dimensions (Wilamowski and Jaeger, 1996). The RBF architecture is similar to the classical regularization network (Poggio and Girosi, 1990). The regularization network has three desirable properties (Girosi and Poggio, 1990; Poggio and Girosi, 1990):

1. It is capable of approximating any multivariate continuous function on a compact domain to an arbitrary accuracy, given a sufficient number of units.