Fluid Phase Equilibria 437 (2017) 14–22

Contents lists available at ScienceDirect

Fluid Phase Equilibria

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

Application of radial basis function neural networks in bubble point oil formation volume factor prediction for petroleum systems

Aref Hashemi Fath

Young Researchers and Elite Club, Gachsaran Branch, Islamic Azad University, Gachsaran, Iran

ARTICLE INFO

Article history:
Received 2 July 2016
Received in revised form
20 December 2016
Accepted 10 January 2017
Available online 11 January 2017

Keywords:
PVT
Bubble point oil FVF
Radial basis function neural networks
Empirical correlation
Outlier detection

ABSTRACT

This paper presents a powerful and comprehensive predictive model based on radial basis function (RBF) neural networks to predict the bubble point oil formation volume factor (FVF), which is one of the most important pressure–volume–temperature properties of crude oils. For this purpose, a large reliable data bank covering a wide range of various crude oil samples was used, with the data collected from the open literature. The performance of the proposed model for the prediction of the bubble point oil FVF was evaluated, using statistical and graphical error analyses, against a number of well-known predictive empirical correlations. The results indicated that, the developed RBF model is able to provide a strong agreement between the predicted values and corresponding experimental data, with an average absolute percent relative error and a coefficient of determination of 1.4562% and 0.9887, respectively, making it more accurate and reliable than the published empirical correlations. In addition, the leverage approach showed that the developed model was statistically acceptable and valid, and only six data points may be considered as probable outliers.

1. Introduction

Accurate determination of the pressure–volume–temperature (PVT) properties of reservoir fluids, such as oil formation volume factor (FVF) at the bubble point pressure is important for several petroleum engineering calculations including well tests investigation, production facility design, mass balance, reservoir simulation, fluid flow in porous media, recovery efficiency, and reservoir future performance forecast [1–6]. The bubble point oil FVF is defined as the volume of reservoir oil that would be occupied at the bubble point pressure and reservoir temperature by one stock tank barrel oil plus any dissolved gas in the oil at that pressure and temperature.

Physical properties of crude oils can be determined through laboratory experiments on reservoir fluid samples [7]. Usually, the bubble point oil FVF is determined in laboratory following a procedure called differential liberation test [7]. However, for some cases, the required laboratory equipment is unavailable; moreover, the procedure not only is costly and time-intensive, but also needs extremely careful sampling for wells with asphaltene and sand production [8]. As a solution, one may refer to equations of state (EOSs) and empirical correlations proposed to calculate PVT properties. EOSs often involve complicated computations and require an extended set of data from the initial composition of reservoir fluid along with some experimental measurements. In addition, the results of an equation of state deem accurate and acceptable only when a proper tuning process (i.e. to match the calculated results with the experimental data) is engaged. Besides, as long as oil and gas viscosities are concerned, EOSs are known to provide poor accuracies, necessitating the application of viscosity correlations such as the correlation proposed by Lohrenz et al. [9].

On the contrary, empirical correlations involve simple calculations and do not require neither to be matched with experimental data nor detailed fluid data. These correlations are mostly developed for a specific geographical region with given chemical composition of reservoir fluid and data range.

During the past decades, researchers have made much effort to introduce a general correlation for the calculation of oil FVF based on linear regression, nonlinear multiple regression, and graphical techniques. The first correlation for the calculation of crude oil's physical properties was proposed in 1942 when Katz [10] presented a correlation for oil FVF.

In 1947, Standing [11] used 105 experimental data sets collected from 22 hydrocarbon mixtures taken from different locations across California to propose graphical correlations for the

E-mail address: aref.hashemifath@yahoo.com.

http://dx.doi.org/10.1016/j.fluid.2017.01.010
0378-3812/© 2017 Elsevier B.V. All rights reserved.

calculation of bubble point pressure, oil FVF, and total oil FVF. Standing [11] reported the average errors of 4.8%, 1.17%, and 5% for bubble point pressure, oil FVF, and total oil FVF, respectively.

In 1980, Glasø [4] prepared correlations to predict bubble point pressure, oil FVF, total oil FVF, and dead oil viscosity. The correlations were developed on the basis of 45 crude oil samples most of which were collected from North Sea. Glasø [4] further presented a correction method for bubble point pressure in the presence of H_2S, CO_2, and N_2 components and reported average relative errors of 1.28%, −0.43%, and −4.56% for the calculated bubble point pressure, oil FVF, and total oil FVF values, respectively, as compared to corresponding experimental data.

The efficiency and capability of the mentioned correlations and other similar empirical correlations proposed by various researchers (e.g. Vazquez and Beggs [12] for a global data bank, Obomanu and Okpobiri [13] for Nigerian oil fields, Al-Marhoun [14] for Middle Eastern oil fields, Labedi [15] for African oil fields, Macary and El-Batanoney [16] for Gulf of Suez oil fields, Dokla and Osman [17] for United Arab Emirates oil fields, Frashad et al. [18] for Colombian oil fields, Omar and Todd [19] for Malaysian oil fields, Petrosky and Farshad [20] for Gulf of Mexico oil fields, Kartoatmodjo and Schmidt [21] for Middle Eastern, Indonesian, North and Latin American oil fields, Khairy et al. [22] for Egyptian oil fields and Dindoruk and Christman [23] for Gulf of Mexico oil fields) depends on the domain of data on which basis they are developed, so that, predicted values by these correlations are sometimes associated with large errors, i.e. they may not serve as comprehensive and perfect approaches to oil FVF prediction for crude oils of various geographical locations with different properties.

During the past two decades, artificial neural networks (ANNs) have been extensively applied in various fields of petroleum engineering including the prediction of PVT properties of crude oil, porosity and permeability assessments, the prediction of minimum miscibility pressure, pressure drop determination in production wells and pipelines, and the calculation of reservoir characteristics. As an example, in 2008, Rasouli et al. [24] presented two multilayer perceptron neural network models to predict the bubble point pressure and oil FVF based on the solution gas-oil ratio, oil gravity, reservoir temperature, and gas specific gravity for Iranian crude oil samples. Both models were trained using 106 experimental data sets and tested by 9 data sets. They reported that their models had lower average absolute error compared with the proposed correlations by Standing [11], Glasø [4], and Al-Marhoun [14] to calculate the bubble point pressure and oil FVF.

The present research is aimed at developing an accurate and reliable system based on radial basis function neural network (RBF-NN), as an acceptable alternative to experimental methods, EOSs, and empirical correlations, to calculate the bubble point oil FVF for different oil samples with various conditions. To achieve the research purpose, the model was constructed and evaluated using a total of 756 experimental data sets representing crude oil samples from around the world; the data sets were collected from the literature.

Accordingly, the proposed RBF model was compared against a number of well-known empirical correlations commonly used to predict the bubble point oil FVF. Furthermore, trend tests were employed to investigate the influence of independent variables on the bubble point oil FVF predictions. Finally, leverage approach was performed to detect probable doubtful data points (outliers) and to find the applicability domain of the proposed model.

2. Radial basis function neural networks (RBF-NNs)

ANNs are computational tools used to model complicated nonlinear relationships and find proper behavioral patterns among data points. The networks present a combination of neural units (which are also known as nodes or neurons which are distributed between layers and act as processing units), biases, activation functions, interconnections (links), and weights. Inspired by human brain, ANNs have found applications in solving a wide range of problems including optimization, classification, function approximation, prediction, pattern recognition, etc. These networks are to be trained by examples; i.e. they are fed with a set of input data along with target data based on which they produce a well-complicated mathematical model representing other new inputs.

Based on the structure of neural networks and the way neurons are connected to one another, different neural networks can be formed; Multilayer perceptron (MLP) and RBF neural networks are two of the most well-known neural networks with wide ranges of application in solving problems [25].

RBF neural networks are neural networks based on localized basis functions and iterative function approximation. In terms of structure, a RBF-NN is composed of three layers, namely an input layer, an output layer, and a hidden layer (see Fig. 1). These types of networks are of fixed architecture with a single hidden layer; this is while MLP-NNs may be of more than one hidden layers. Indeed, a RBF-NN represents a special case of a MLP-NN [26]. Owing to their simple design, extremely strong tolerance to input noises, and fast yet pervasive training capabilities, these networks have attracted a large deal of attention. In RBF-NNs, there is a single input layer wherein no processing is undertaken. The hidden layer, however, contains radial basis functions, with the output layer solely containing collectors. In fact, the output layer linearly combines all outputs from neurons in the hidden layer to generate the network output. Compared to MLP networks, this type of network requires larger number of neurons, even though they enjoy shorter designs, with the principal distinction being the application of activation functions to be used by neurons [27].

MLP networks provide global approximations to nonlinear input(s)-output(s) mapping whereas RBF networks act as local approximations [28].

Network output for an input pattern, x, can be expressed as follows:

$$y_j(x) = \sum_{i=1}^{I} w_{ij}\phi(\|x - c_i\|) \quad j = 1, 2, \ldots, n \tag{1}$$

Where $y_j(x)$ represents the network's jth output, I is the number of units in the hidden layer, w_{ij} represents the weight of the link between the ith hidden unit and jth output node. Furthermore, ϕ is the RBF employed by the ith hidden unit. x denotes the input vector

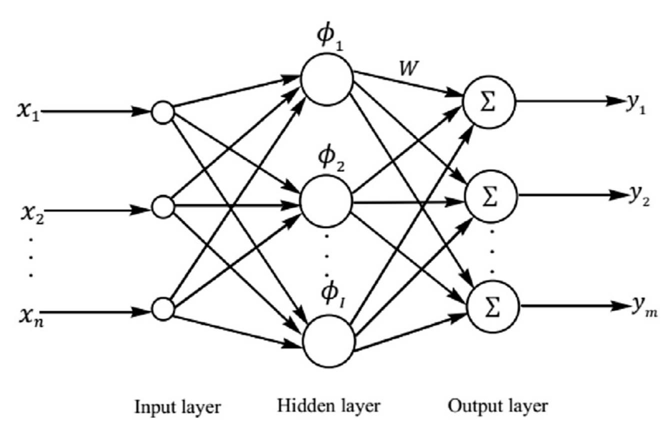

Fig. 1. Architecture of a radial basis function neural network.