Journal of the Taiwan Institute of Chemical Engineers 55 (2015) 17–26

Contents lists available at ScienceDirect

Journal of the Taiwan Institute of Chemical Engineers

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

On accurate determination of PVT properties in crude oil systems: Committee machine intelligent system modeling approach

Amin Shokrollahi*, Afshin Tatar, Hossein Safari

Young Researchers and Elites Club, North Tehran Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Article history:
Received 18 October 2014
Revised 17 February 2015
Accepted 7 April 2015
Available online 11 May 2015

Keywords:
Oil formation volume factor (B_o)
Bubble point pressure (P_b)
Committee machine intelligent system (CMIS)
PVT properties
Empirical correlations

ABSTRACT

This study is focused on accurate determination of PVT properties of reservoir oil using an ensemble approach referred to as committee machine intelligent system. PVT properties of interest are oil formation volume factor (B_o) and bubble point pressure (P_b). Committee machine intelligent system model developed in this work combine multi-layer perceptron network, radial basis function network, and least squares support vector machine via a modified weighted averaging method, while optimizing these weights using genetic algorithm. Developed committee machine intelligent system model was found to accord excellently with experimental data yielding correlation coefficients (R^2) of 0.980 and 0.976 for bubble point pressure (P_b), and oil formation volume factor (B_o), respectively. Comprehensive comparisons were also carried out between a variety of PVT prediction models and committee machine intelligent system model developed in this study. At the end, leverage value statistics method was applied to detect and identify some probable outlying points from the gathered databases.

1. Introduction

Determination of the accurate and reliable phase behavior of petroleum reservoir fluids is an essential key in many aspects of oil industry and flow through porous media [1–3]. Most oil reservoirs are produced by depletion during which a decline in reservoir pressure is observed as fluids are recovered to the surface. However, reservoir temperature appears to be nearly constant during oil recovery. Thus, the fundamental variable which determines petroleum phase behavior under reservoir condition is reservoir pressure [2]. Volumetric data are usually obtained under reservoir and surface temperature, and are categorized as PVT data (pressure, volume, temperature). This information is then implemented in reserve evaluation, developing optimum recovery plans, obtaining quantitative and qualitative estimates of fluids being produced, and many other properties needed in computational reservoir engineering [1–3]. Fluid properties including bubble point pressure (P_b) and oil formation volume factor (B_o) play a significant role in many aspects of reservoir engineering especially in reserve evaluation and production capacity estimation [2,4]. These properties can either be determined experimentally through samples collected from bottom hole and from surface equipment or by using empirical correlations and predictive equations [2]. Nevertheless, experimental determination of these variables can be very costly and time consuming. Predictive correlations are widely applied in oil industry, particularly when PVT data are unavailable or uneconomical to obtain. In addition, some researchers have developed and applied equations of states (EOS) for predicting PVT properties of reservoir fluids, but it is generally believed that these equations do not yield accurate predictions at all reservoir conditions [1,2].

A wide numbers of studies have been conducted with the aim of devising a reliable and accurate empirical correlation capable of predicting PVT properties in petroleum reservoirs. Graphical correlation developed by Katz [5] to predict oil formation volume factor (B_o), can probably be considered as a benchmark in this field. This correlation employs reservoir pressure, temperature, API oil gravity, gas gravity, and solution gas–oil ratio. In later years, Standing [6] developed another graphical correlation to predict the bubble point pressure (P_b), oil formation volume factor (B_o), and total formation volume factor (B_t) based on laboratory analyses carried out on California crude oils. Standing's correlations [6], which are still most commonly used in petroleum industry, implement the reservoir temperature, API oil gravity, gas gravity, and solution gas–oil ratio as correlating parameters. A number of other empirical correlations have also been developed aiming to model the PVT properties of reservoir fluids, which are all based on these four parameters [4,7–19]. Each of these correlations underwent numerous modifications and improvements over time and new correlations have been evolved. The most recent empirical correlation was presented in 1997 by Almehaideb [20] for PVT properties of UAE crude oils employing 62 datasets obtained from different oil

* Corresponding author. Tel.: +98 936 382 6336.
E-mail address: shokrollahi.amin@gmail.com (A. Shokrollahi).

http://dx.doi.org/10.1016/j.jtice.2015.04.009

Nomenclature

AARD	average absolute relative deviations
ANN	artificial neural network
CMIS	committee machine intelligent system
GA	genetic algorithm
H	Hat matrix
H^*	warning leverage
LSSVM	least square supported vector machine
MLP	multilayer perceptron
R^2	correlation coefficient
RBF	radial basis function
ARD	average relative deviation
RMSE	root mean squared error
MSE	mean squared error

reservoirs. Almehaideb's correlations [20] predict oil formation volume factor (B_o), saturation pressure, oil compressibility, and oil viscosity. Reviews of empirical correlations presented in the literature to predict the bubble point pressure (P_b) and oil formation volume factor (B_o) have been summarized in Tables 1 and 2, respectively.

Despite their satisfactory performances in predicting PVT properties, developed empirical correlations usually carry some deficiencies and uncertainties as they are either limited to a certain range or specific operational condition [2]. Empirical correlations are merely based on observations rather than theory in which a dependent parameter is related to a number of independent parameters and model parameters are subsequently tuned employing the same databank used in model development. In recent years, inductive machine learning algorithms have widely been applied in petroleum industry to model macroscopic properties through learning the physical patterns in a set of experimentally measured data [21–23]. In contrast to empirical correlations, machine learning methodologies are branches of artificial intelligence primarily concerned with the development of rigorous and comprehensive models capable of learning the physical patterns between dependent and independent parameters [24]. These models initially observe the fed databank to discover a meaningful relationship between input and output parameters, and then generalize to perform satisfactorily for the unseen data [24]. A number of researchers have employed these machine learning methods to predict the PVT properties of petroleum reservoir fluids along with reducing the deficiencies associated with empirical correlations and equations of state. Gharbi and Elsharkawy [25] probably did the first work by employing the neural networks for quantifying the saturation pressure and oil formation volume factor (B_o) of Middle East crude oils. In their model, 498 datasets all pertinent to Middle East crudes were used to construct the predictive model and the model was subsequently tested by employing another 22 dataset obtained from the same source. Several other investigators have also used a similar approach based on machine learning algorithms to obtain a reliable and accurate prediction of PVT properties of petroleum fluids employing huge datasets [25–31]. However, employed artificial neural networks may in some cases fail in accurate prediction of these properties, partly due to random initializations of networks weights and biases and variations in stopping criteria of the optimization process. Support vector machine algorithms have recently been applied in solving many complex problems in science and engineering due to their ease of handling large scale problems [21–23,32].

Owing to the fact that each of these models produces different individual errors while approximating PVT values, an ensemble of intelligent models based on different learning algorithms could be an outstanding effort in order to obtain more accurate and reliable models. Thus, the primary aim of this study is to develop a committee machine intelligent system (CMIS) based on three distinctive intelligent algorithms, namely, radial basis function (RBF) neural network, multilayer perceptron (MLP) neural network, and least square support vector machine (LSSVM). Details of each mentioned experts are available in literature [32–34]. These algorithms are subsequently combined through a modified weighted averaging method in order to improve the individual results. Genetic algorithm (GA) is then used to optimize the weights associated with each of these models, mainly due to its flexibility and well performance in solving large scale optimization problems. Proposed methodology is then compared with experimental PVT data and those predicted by other correlations. Results show that this approach outperforms previous correlations and modeling methodologies.

2. Details of intelligent method

Artificial intelligence (or AI) is "the study and design of intelligent agents", where an intelligent agent is a system that perceives its environment and takes actions which maximizes its chances of success. An artificial neural network (ANN) is an information-processing paradigm inspired by the way the densely interconnected, parallel

Table 1
Common correlations and their restrictions used to predict bubble point pressure (P_b) .

Author	Correlation	Origin of sample	Restrictions
Standing [6]	$P_b = 18.2[(R_s/\gamma_g)^{83} 10^a - 1.4]$ $a = 0.0009(T - 460) - 0.0125(\text{API})$	California oil fields	$130 < P_b < 7000$ psia
Lasater [7]	$P_b = [P_f(T + 459.67)]/\gamma_g$ $Y_g = \dfrac{R_s/379.3}{R_s/379.3 + 350\gamma_o/M_o}$ $M_o = 725.32143 - 16.03333\gamma_{oAPI} + 0.09524\gamma_{oAPI}^2$ $P_f = 0.38418 - 1.20081Y_g + 9.64868{Y_g}^2$	Canada west and mid-cont. U.S. and S.A.	$P_b < 5780$ psia
Vazquez and Beggs [18]	$P_b = [(C_1 R_s/\gamma_{gs})10^{-C_3 \text{API/T}}]^{C_2}$ $\text{API} <= 30 : C_1 = 27.624, C_2 = 0.91433, C_3 = 11.172$ $\text{API} > 30 : C_1 = 56.18, C_2 = 0.84246, C_3 = 10.393$	Worldwide	Separation condition must be known
Glaso [8]	$\log(P_b) = 1.7669 + 1.7447[log(G)] - 0.30218[log(G)]$ $G = (R_s/\gamma_g)^{0.816}(T - 460)^{0.172}(\text{API})^{-0.989}$	North Sea	
Al-Marhoun [10]	$P_b = 0.00538 * {R_s}^{0.715082}{\gamma_g}^{-1.87784}{\gamma_o}^{3.1437}T^{1.32657}$	Middle East	
Dokla and Osman [54]	$P_b = 8363.86 * {R_s}^{0.724047}{\gamma_g}^{-1.01049}{\gamma_o}^{0.107991}T^{-0.952584}$	UAE	$P_b < 3573$ psia
Petrosky and Farshad [17]	$P_b = 112.727\left[\dfrac{{R_s}^{0.5774}}{\gamma_g^{0.8439}}10^x - 12.340\right]$ $x = 4.561 \times 10^{-5}T^{1.3911} - 7.916 \times 10^{-4}\text{API}^{1.5410}$	Gulf of Mexico	$590 < P_b < 4640$ psia
Al-Shammasi [27]	$P_b = \gamma_o^{5.527215}[R_s T\gamma_g]^{0.783716}\exp(-1.841408\gamma_o\gamma_g)$	Worldwide	$1574 < P_b < 6523$ psia, $31.7 < P_b < 7127$ psia