Journal of Natural Gas Science and Engineering 22 (2015) 348–358

Contents lists available at ScienceDirect

Journal of Natural Gas Science and Engineering

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

On determination of natural gas density: Least square support vector machine modeling approach

Shayan Esfahani [a], Sina Baselizadeh [a], Abdolhossein Hemmati-Sarapardeh [b, *]

[a] Department of Petroleum Engineering, Amirkabir University of Technology, Tehran, Iran
[b] Young Researchers and Elite Club, Kerman Branch, Islamic Azad University, Kerman, Iran

ARTICLE INFO

Article history:
Received 19 October 2014
Received in revised form
1 December 2014
Accepted 4 December 2014
Available online 24 December 2014

Keywords:
Natural gas
Density
Least square support vector machine
Intelligent model

ABSTRACT

In this century, worldwide consumption of natural gas is expected to increase drastically because it is one of the cleanest and most available energy sources. Accurate knowledge of natural gas properties is of a vital significance in gas engineering. One of the most important properties of natural gas is density, which is traditionally measured through expensive, time consuming and cumbersome experiments. In this communication, a new reliable and accurate model for prediction of natural gas density is presented as a function of pseudo reduced pressure, pseudo reduced temperature and apparent molecular weight of gas. A supervised learning algorithm, namely least square support vector machine, has been employed for modeling the gas density, and the parameters of the model were optimized through coupled simulated annealing. The results of this study indicated that the developed model can satisfactorily predict gas density in a wide range of pressure (from 13.7 to 10,000 psia), temperature (from −25 to 460 °F) and gas composition (molecular weight from 16.04 to 129.66). Moreover, the accuracy and validity of the proposed model was compared to pre-existing models, and it was found that the proposed model is more accurate, reliable and superior to all the investigated models. In addition, the relevancy factor demonstrated that molecular weight has the greatest impact on gas density among the selected input parameters.

1. Introduction

Over the last few decades, natural gas has drawn much attention due to its abundance and clean burning (Kamari et al., 2013a). In the 21st century, worldwide consumption of natural gas is expected to increase drastically because it is one of the cleanest and most available energy sources. Accurate calculation in gas engineering like gas flow rate in porous media and reservoir simulation requires precise knowledge of natural gas properties. Applying accurate experiments is the most reliable method for measurement of gas properties; however, they are time-consuming, cumbersome, and costly. Moreover, natural gas is a mixture of light hydrocarbon and a variety of heavier hydrocarbon and non-hydrocarbon components which make the gas properties highly dependent on gas composition. Many researchers have attempted to develop reliable and accurate models for prediction of gas properties. In general, these models can be categorized as equations of state and empirically derived correlations (Sutton, 2007; Elsharkawy, 2004).

One of the most significant parameters in PVT calculations is gas density. Most of empirical correlations estimate gas density using PVT parameters such as pressure, temperature and Z-factor (Benedict et al., 1940; Dranchuk and Abou-Kassem, 1975; Nishiumi and Saito, 1975). The accuracy of correlations results depends on the accuracy of the inputs provided. Among the input parameters, precise calculation of Z-factor is challenging. There are many equations for prediction of Z-factor, however, they cannot predict it with high accuracy. Their error can be minimized by tuning the models for z-factor of the studied data set.

According to the "Law of corresponding state", pseudo reduced pressure and pseudo reduced temperature are calculated, and real gas Z-factor is obtained by an explicit form of these two parameters. It must be reflected on the regression formula that Z-factor is resolved as a root of equations of state. The density correlations based on Peng-Robinson equation of state (EOS) usually give results with an average error of 10% in a vast range of reservoir conditions (Chien and Monroy, 1986). The correlations based on chain-of-rotators (COR) equation of state can be applied on larger ranges

* Corresponding author.
E-mail addresses: aut.hemmati@aut.ac.ir, aut.hemmati@gmail.com (A. Hemmati-Sarapardeh).

http://dx.doi.org/10.1016/j.jngse.2014.12.003
1875-5100/

of reservoir conditions; however, this method is costly. There are also correlations based on the three-parameter Peng-Robinson (PR3) EOS and have a quite limited function range. Therefore, they are more accurate for the cases which require extensive computations such as compositional reservoir simulations (Chien and Monroy, 1986).

Some of the most common density correlations in petroleum industry are presented in the following part:

In the 1940s, Standing-Katz (Standing and Katz, 1942) performed an investigation on natural gas properties. They computed the gas density using ideal gas law corrected by compressibility factor. They also developed a graphical correlation for prediction of gas density. Reservoir temperature, pressure and gas specific gravity was the input data of their model.

Lately, EOS models have received more attention. Poettmann-Carpenter (Poettman and Carpenter, 1952) graphically predicted the real gas Z-factor as an explicit function of pseudo reduced pressure and pseudo reduced temperature. Equations of state are implicit in terms of Z-factor. In 1973, Beggs and Brill (Beggs and Brill, 1973) performed a study on gas density and Z-factor. This attempt provided a correlation with average absolute error of 0.19%. This model is only valid in the range of $1.2 < Tr < 2.4$ and $0.0 < Pr < 10$. In 1975, Dranchuk-Abou-kassem (DAK) (Dranchuk and Abou-Kassem, 1975) presented a new correlation for estimating Z-factor. Their correlation was derived based on a Han-Starling form of Benedict-Webb-Robin EOS. They used 1500 data points and reported an average absolute error of 0.486%. As this correlation agrees with the data trend, it can be used for the most engineering calculations, but it leads to dramatic errors in critical conditions. In 1975, Nishiumi-Saito (NS) (Nishiumi and Saito, 1975) presented a new correlation by adding a few more terms to Dranchuk-Abou-kassem (Dranchuk and Abou-Kassem, 1975) expression in order to provide better performance in the vicinity of critical condition. Afterward, Londono et al. (Londono et al., 2002) compared the performance of DAK (Dranchuk and Abou-Kassem, 1975) and NS-EOS (Nishiumi and Saito, 1975) by regression of these correlations using a new gas density (Z-factor) data set. First, they performed regression on a pure component data set. Then, regression was performed on a mixture data and the EOS was calibrated and two new correlations were presented for each relation. In 2008, Shokir (Shokir, 2008) introduced a new model for prediction of gas density using fuzzy logic approach. He used 5350 data points to develop his model. Input data including pseudo reduced pressure, pseudo reduced temperature and molecular weight were partitioned by fuzzy clustering. He optimized cluster numbers and proposed separate correlations for each cluster. The average absolute error of this model was reported 2.37%. One year later in 2009, Al-Quraishi – Shokir (AlQuraishi and Shokir, 2011) developed a new model for gas density using Alternating Conditional Expectations (ACE) algorithm. 800 random data points were used in order to build ACE density model and they reported 4.93% of average absolute error. Furthermore, they compared the performance of ACE model with other density models and their model came in the third place after DAK (Dranchuk and Abou-Kassem, 1975) and Shokir (Shokir, 2008) model. In 2009, AlQuraishi and Jummah (AlQuraishi and Jumma, 2009) developed their new models for gas density based on genetic programming (GP) technique, by means of the same data base as Al-Quraishi – Shokir (AlQuraishi and Shokir, 2011) used for their model development. They reported the average absolute error of 11.6%.

Equations of state, in general, have poor capability to estimate volumetric properties of gas mixtures (Kamari et al., 2013a). Furthermore, correlations which are used to predict gas density are sometimes too complex, requiring initial values and longer computations. These correlations have significant error in some conditions of temperature, pressure and composition. Even a small error in gas density calculation leads to a large error in predicting properties and parameters. As a result, developing a simple, robust and accurate model for prediction of gas density is of a vital importance. Artificial intelligent models are robust techniques in developing a predictive model for handling complex systems (Lashkarbolooki et al., 2012). Artificial intelligence techniques are fast and accurate methods for prediction of gas and oil properties and have been successfully applied in various gas and petroleum engineering problems (Chamkalani et al., 2013; Kamari et al., 2014a; Tehrani and Fariba, 2011; Zendehboudi et al., 2013). Very recently, least square support vector machine (LSSVM) modeling approach has been used and showed excellent performance in solving different problems in oil and gas engineering (Kamari et al., 2013b, 2014b; Kamari et al., 2014; Hemmati-Sarapardeh et al., 2014a, 2013a; Shokrollahi et al., 2013). However, to the best of the authors' knowledge, no research work has yet been published on the prediction of gas density using the LSSVM algorithm.

The main objective of this study is to develop a fast, reliable and accurate model for the prediction of gas density. To this end, a large data bank, covering a wide range of thermodynamic conditions, is collected from variety of literature sources. Afterward, the LSSVM algorithm is employed to construct the gas density model. The parameters of the model are optimized through an external optimizer, namely coupled simulated annealing (CSA) technique. The performance and accuracy of the developed CSA-LSSVM model is evaluated and compared to previously published models by means of statistical and graphical error analyses. Besides, the reliability of the proposed model as well as existing models across the spectrum of the independent variables is investigated. And finally, the relevancy factor is employed to find the relative effect of input parameters on the gas density.

2. Model development

2.1. Data acquisition and preprocessing

To develop a comprehensive and robust model for prediction of gas density, a large data bank which covers a wide range of pressure, temperature and gas composition is required. Such a large data bank (3819 data points) was collected from literature sources (Assael et al., 2001; Langelandsvik et al., 2007; Schley et al., 2004; Sanaei et al., 2011), which covers the temperature range of −25 to 460 °F, pressure range of 13.7–10,000 psia, apparent molecular weight of 16.04–129.66, and density range of 0.0003–0.7525 g/cc. This data set comprises pure hydrocarbon and natural gases containing a small concentration of heavy hydrocarbon components and impurities such as nitrogen, carbon dioxide, helium, and hydrogen sulfide. The ranges of data used in this study and their corresponding statistical parameters are summarized in Table 1.

Pseudo reduced temperature, pseudo reduced pressure and apparent molecular weight were selected as the independent input variables and gas density has been considered as the desirable output. Therefore, the following general relation has been considered for the gas density:

Table 1
Summarized statistical description of the datasets used for model development.

Parameter	Min	Average	Max	Stan. Dev.
Molecular weight	16.04	36.49	129.66	27.22
Pressure (psia)	13.7	3456	10000.00	2877.56
Reservoir temperature (°F)	−25.70	217.85	460.00	131.94
T_{pr}	0.54	1.43	2.68	0.40
P_{pr}	0.02	5.93	29.29	5.25
Density (g/cc)	0.00038	0.26	0.75	0.21