Petroleum and Coal

Article Open Access

Determination of Compressibility Factor for Natural Gases Using Artificial Neural Network

Joseph C. Nwosu[1], *Stanley U. Ibeh*[*2], *Stanley I. Onwukwe*[1], *Boniface O. Obah*[1]

[1] *Department of Petroleum Engineering, Federal University of Technology Owerri, Nigeria*
[*2] *Department of Petroleum Engineering, Madonna University Akpugo, Enugu, Nigeria*

Received July 19, 2018; Accepted October 19, 2018

Abstract

This work proposes the use of data division and if statements in a programming language, as an effective classifier in Artificial Neural Network. The Standing and Katz chart was digitized to obtain input (pseudo reduced temperature and pressure) and output (gas compressibility factor) data points, which was used in developing the artificial neural network. A total of 114,120 input data points and 57,060 output data points were used. The dataset was divided into 4 groups, and each of the groups was assigned a neural network that corresponds to the value range of the grouped data using a Matlab nnet tool box.

***Keywords**: gas compressibility factor; artificial neural network; natural gas; gas correlation; Matlab.*

1. Introduction

Fluid properties are determined from the laboratory using applicable experimental procedures, with the aim of analyzing samples to meet conditions of interest. The unavailability of these samples brings rise to the use of empirical correlations. The determination of accurate compressibility factor is of great importance in the industry. This parameter has been a tool in determining the deviation of real gas from ideal behaviour. A compressibility factor is a useful tool in various engineering applications, which include; designing pipelines, gas flow rates, the design of oil and gas separators, gas reserves estimation, etc. According to Cengel and Boles [1], the principle of corresponding states indicates that all gases, when compared at the same reduced pressure (Ppr) and reduced temperature (Tpr), have approximately the same compressibility factor. The development of the general EOS and the elaboration of general experimental charts expressed in terms of the reduced properties were based on this principle. Standing and Katz [2] developed a chart for the compressibility factor and pseudo reduced properties (temperature and pressure) which is an industry standard. Other correlations exist and each of which shows better prediction at specific ranges of pseudo reduced temperatures.

Katz *et al.* [3] developed a chart of compressibility factor in terms of pseudo reduced properties for natural gases at pressures of 10,000 to 20,000psia. The chart displays a high level of linearity, which makes prediction less complex without the need for multiple data points selection across the pseudo reduced temperature curves.

Kamyab *et al.* [4] worked on digitizing the Standing-Katz and Katz compressibility chart to obtain input and output data, in which artificial neural network was used as a tool in predicting the output of compressibility factor. In their project, a two-hidden-layer feed forward network was designed and trained with back-propagation supervised learning. They developed a methodology to obtain z-factors for Natural Hydrocarbon Gases using Artificial Neural Networks (ANN). Data obtained directly from the Standing-Katz and Katz compressibility charts were used to train several topologies of ANN. The input parameters in the ANN are the pseudo-reduced pressure and temperature, and the output is the z-factor. Two of the successful networks have two hidden layers. The first ANN uses five neurons in each hidden layer and the second ANN uses ten neurons in each hidden layer (called 2-5-5-1 and 2-10-10-1 networks respectively). These topologies were trained with the data from the charts using a back-propagation training algorithm. This work had a limitation in understanding the level of accuracy of artificial neural networks as a predictive tool for complex datasets.

Beggs and Brill [5] developed a correlation for the determination of compressibility factor which is given below as:

$Z = A + (1 - A)e^{-B} + CP_{pr}^{D}$ (1), where:

$A = 1.39\,(T_{pr} - 0.92)^{0.5} - 0.36T_{pr} - 0.101$ (2)

$B = \left(0.62 - 0.23T_{pr}\right)P_{pr} + \left[\left(\frac{0.066}{T_{pr}-0.86}\right) - 0.037\right]P_{pr}^{2} + [\frac{0.32}{10^{9}(T_{pr}-1)}]P_{pr}^{6}$ (3)

$C = (0.132 - 0.32logT_{pr}$ (4)

$D = 10^{0.3106-0.49T_{pr}+0.1824T_{pr}^{2}}$ (5)

Dranchuk and Abou-Kassem [6] developed their equation for the determination of the compressibility factor as follows;

$Z = 1 + C_1\left(T_{pr}\right)\rho_{pr} + C_2\left(T_{pr}\right)\rho_{pr}^{2} - C_3\left(T_{pr}\right)\rho_{pr}^{5} + C_4(\rho_{pr}T_{pr})$ (6), where:

$\rho_{pr} = 0.27\frac{P_{pr}}{zT_{pr}}$ (7)

$C_1\left(T_{pr}\right) = A_1 + A_2T_{pr}^{-1} + A_3T_{pr}^{-3} + A_4T_{pr}^{-4} + A_5T_{pr}^{-5}$ (8)

$C_2\left(T_{pr}\right) = A_6 + A_7T_{pr} + A_8T_{pr}^{-2}$ (9)

$C_3\left(T_{pr}\right) = A_9[A_7T_{pr}^{-1} + A_8T_{pr}^{-2}]$ (10)

$C_4\left(\rho_{pr}T_{pr}\right) = A_{10}\left(1 + A_{11}\rho_{pr}^{2}\right)(\rho_{pr}^{2}T_{pr}^{-2})$ (11)

The constants A1 through A11 are as follows: A1 = 0.3265; A2 = -1.07; A3 = -0.5339; A4 = 0.01569; A5 = -0.05165; A6 = 0.5475; A7 = -0.7361; A8 = 0.1844; A9 = 0.1056; A10 = 0.6134; A11 = 0.721 They used Newton's method to determine the solution of the above equations.

2. Artificial neural network

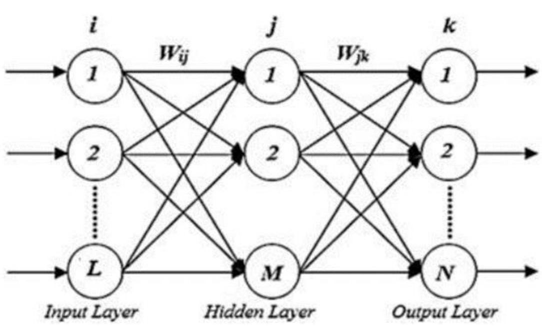

Fig. 1 Structure of an Artificial Neural Network

Artificial Neural Networks are simulators, which work on the basis of the human nervous system to carry out certain tasks like classification, pattern recognition, etc. The artificial neurons present in the network lies in constitutive layers of the network. Layers are linked to the next by specific weights (w). One of the most practical structures of the Artificial Neural Network is Multi-Layer Perceptron (MLP) (Fig. 1) in which the input and output layers are connected to each other by an additional layer called hidden layer.

This structure was adapted in the course of this work. Each input is multiplied by its corresponding weights. Weights carry the information needed by the neural network to solve a problem and also represent the strength of the interconnection between neurons inside the neural network.

Activation functions are set to serve as a transfer function used to get the desired output. There are linear as well as the non-linear activation function. Some of the commonly used activation function is - binary, sigmoidal (linear) and tan hyperbolic sigmoidal functions (non-linear).The structure of an Artificial Neural Network is given in Fig. 3.

3. Proposed approach

The Standing and Katz chart data has been read directly from a scanned figure of the original plot drawn by Standing and Katz in 1951.The software used in digitizing the chart and obtaining the dataset was GraphClick. According to Rakap *et al.* [7], this software has been tested and qualified to be able to digitize charts. Each one of the pseudo-reduced temperature curves from Standing and Katz chart was been digitized [4], some curves needed more points (Fig. 4) to describe the curvature better while other curves are almost linear and fewer points(Fig. 5) were necessary to define the curve. It should also be noted the need for data uniformity in the distribution of data points, so as to enable the neural network to to understand the data progression pattern.

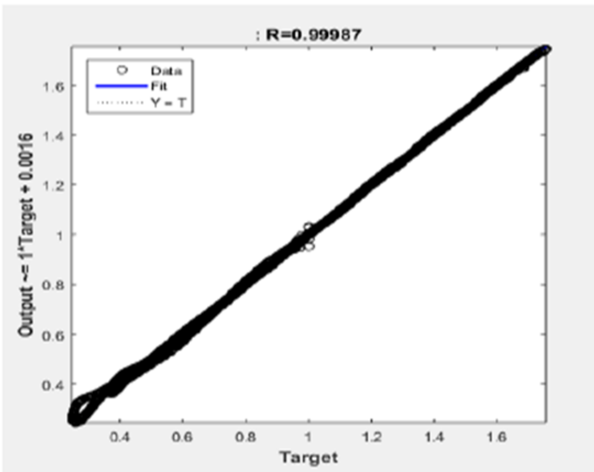

Fig 2. Regression plot for outputs derived from 1.05≤TPr≤1.2

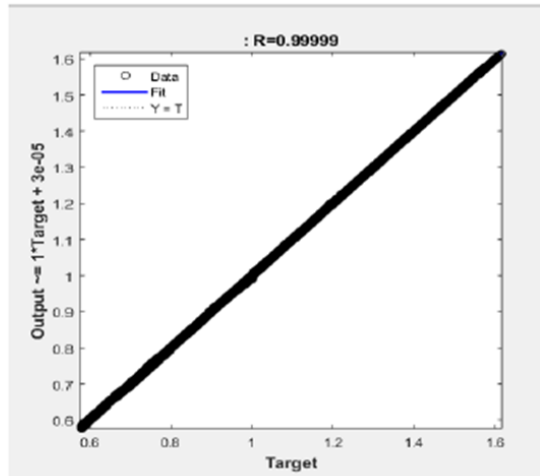

Fig. 3. Regression plot for outputs derived from 1.5≤Tpr≤2.0

The human nervous system works as a transmitter of information which perceives, interprets and transmits information to targets organs in the human body. A key fact that should be noted is that the information been perceived and interpreted are later classified (during the interpretation phase) before been transmitted to target organs, which optimizes the effectiveness of the nervous system in transmitting information. Classification in this context means allocating information to be transmitted to a particular target organ for execution depending on the information perceived. Artificial Neural Network works on this base principle, and the importance of data classification/grouping should always be noted especially when dealing with large datasets, as it increases the effectiveness in prediction.

This work made use of the Nnet toolbox in Matlab software to develop the needed Neural Networks using the back propagation algorithm. The Neural network assumed a total number of 30 hidden layers, using 99% data for training, 0.5% for Testing and 0.5% for validation. The dataset has 2 input data columns and 1 output data column, having a matrix of 57,060 × 2 and 57,060 × 1 respectively. The dataset was divided into 4 groups with the aim of developing a neural network for each group of dataset divided. The development of an individual neural network for each grouped data was to reduce the complexity in determining the output variables, in this case, the compressibility factor. Each of the Neural networks covered the ranges of pseudo reduced temperature values 1.05 ≤ Tpr ≤ 1.2, 1.2 < Tpr ≤ 1.5, 1.5 < Tpr ≤ 2.0 and 2.0 < Tpr ≤ 3.0 respectively, at Pseudo, reduced pressures of 0.5≤ Ppr ≤ 5.5.

This work proposes the use of "If Statements" in programming the desired neural network as a predictive tool for the compressibility factor when the Input data; Pseudo reduced pressure and temperature are provided. The results were compared with other existing correlations and showed an overall outstanding performance (Tables 1-5).

Table 1. Compressibility factors at different Tpr=1.2 using various correlations

Ppr	Standing and Katz	This approach	Kamyab *et al.* [4]	Beggs & Brill [5]	Dranchuk-Aboukassem [6]
0.5	0.8930	0.8960	0.8953923	0.9026461	0.8950631
1.5	0.657	0.6573	0.6607512	0.6758659	0.6532419
2.5	0.519	0.5196	0.5179963	0.4977865	0.5180675
3.5	0.565	0.5663	0.5676801	0.5605500	0.5631805
4.5	0.650	0.6499	0.6492856	0.6589953	0.6501377
5.5	0.741	0.7406	0.7424365	0.7567099	0.7453363

Table 2. Compressibility factors at different Tpr=1.3 using various correlations

Ppr	Standing and Katz	This approach	Kamyab *et al.* [4]	Beggs & Brill [5]	Dranchuk-Aboukassem [6]
0.5	0.916	0.9182	0.9196115	0.9266436	0.9203019
1.5	0.756	0.7584	0.7567070	0.7675523	0.7543694
2.5	0.638	0.6399	0.6394479	0.6526911	0.6377871
3.5	0.633	0.6329	0.6341957	0.6234648	0.6339351
4.5	0.684	0.6832	0.6857549	0.6921991	0.6898314
5.5	0.759	0.7604	0.7611212	0.7779095	0.7663247

Table 3. Compressibility factors at different Tpr=1.5 using various correlations

Ppr	Standing and Katz	This approach	Kamyab *et al.* [4]	Beggs & Brill [5]	Dranchuk-Aboukassem [6]
0.5	0.948	0.9493	0.9508509	0.9555248	0.9509373
1.5	0.859	0.8592	0.8607096	0.8618306	0.8593144
2.5	0.794	0.7974	0.7940885	0.7945385	0.7929993
3.5	0.770	0.7694	0.7685691	0.7691830	0.7710525
4.5	0.790	0.7910	0.7867923	0.7828753	0.7896224
5.5	0.836	0.8344	0.8323518	0.8248905	0.8331893

Table 4. Compressibility factors at different Tpr=2.0 using various correlations

Ppr	Standing and Katz	This approach	Kamyab *et al.* [4]	Beggs & Brill [5]	Dranchuk-Aboukassem [6]
0.5	0.982	0.9837	0.9839990	0.9853337	0.9824731
1.5	0.956	0.9575	0.9572277	0.9629020	0.9551087
2.5	0.941	0.9406	0.9414698	0.9471826	0.9400752
3.5	0.937	0.9365	0.9352303	0.9471826	0.9385273
4.5	0.945	0.9437	0.9453140	0.9404180	0.9497137
5.5	0.969	0.9686	0.9693022	0.9443010	0.9715388

Table 5. Compressibility factors at different Tpr=3.0 using various correlations

Ppr	Standing and Katz	This approach	Kamyab *et al.* [4]	Beggs & Brill [5]	Dranchuk-Aboukassem [6]
0.5	1.002	1.0036	1.0028553	1.0040392	0.9984498
1.5	1.009	1.0102	1.0095269	0.9514557	0.9995529
2.5	1.018	1.0187	1.0179196	0.7082371	1.0061111
3.5	1.029	1.0292	1.0286167	0.1399229	1.0176846
4.5	1.041	1.0412	1.0412701	-0.8897010	1.0336417
5.5	1.056	1.0570	1.0563968	-2.5178952	1.0532809

4. Results

The Matlab software produced 4 regression plots which contained the Mean Square error and Root Mean Square error of the individual Neural Network developed. The mean square error gives the squared mean deviation of the output data (actual value) from the target data (estimated value). The closer the mean square error value is to zero, the more accurate and the lesser the error of the estimated value. The root means square error measures how much error there is between two data sets, in other words, it compares an estimated value and the actual value. The R-squared value is an indicator of how well the model fits the data. An R-square close to 1 indicates that the model accounts for almost all the variability in the data. The first Neural Network for data group 1.05 ≤ Tpr ≤ 1.2 (Fig. 2), had a mean square error of 0.0016 and a root mean square error of 0.99987. The second Neural Network for data group 1.2 < Tpr ≤ 1.5 (Fig.4) had a mean square error of 0.000081, and a root mean square error of 1.0000. The third Neural Network for data group of 1.5 < Tpr ≤ 2.0 (Fig.3) had a mean square error of 0.00003, and a root mean square error of 0.99999 and the last Neural Network for data group 2.0 < Tpr ≤ 3.0 (Fig. 5) had a mean square error of 0.000011 and a root mean square error of 1.0000.

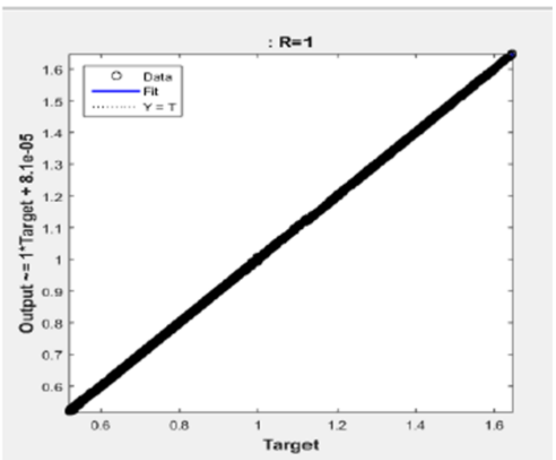

Fig 4. Regression plot for outputs derived from 1.2<Tpr≤1.5

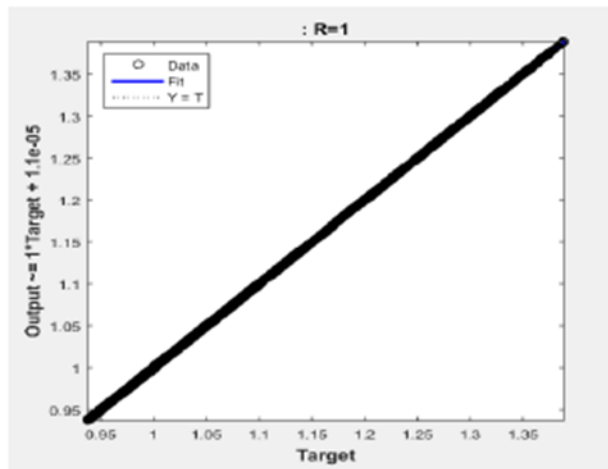

Fig. 5. Regression plot outputs derived from 2.0<Tpr≤3.0

The average absolute error (mean square error), was determined using the following formula

$$MAE = \frac{1}{n}\sum_i^n |V_e - V_a| \qquad (12)$$

where: V_e = estimated value; V_a = actual value; MAE = mean absolute error.

The mean absolute error determination gives the deviation of the estimated output data from the actual output data. The lower the value of this error, the higher the precision. This was determined for all the correlations used in the comparison and proved this approach to have the least mean absolute error of about 0.00111. The value of the mean square error proves the precision of this approach to other methods and can be seen in Table 6.

Table 6. Average absolute errors

This approach	Kamayab *et al.* [4]	Beggs & Brill [5]	Dranchuk-Aboukassem [6]
0.00111	0.00281249	0.233375273	0.0033042

The data obtained from this method was compared with existing correlations like; Begs and Brill, Dranchuk-Aboukassem [6] and Kamyab *et al.* [4]. These correlations were proven by previous research to have a high degree of accuracy in the determination of the compressibility.

The test data used for this comparison was derived from previous literature that proved these correlations to have a high level of accuracy. The comparisons were made and given below in the table of values (Tables 1–5).

5. Conclusion

The values derived from this approach proved to be better than existing methods. This approach has justified the uniqueness and accuracy of Artificial Neural Network in data analysis and pattern recognition, especially in data division. The theory from this approach can be deployed in computer based applications to give precision in the determination of the compressibility factor.

Nomenclature

P - pressure, psia
P_c -critical pressure, psia
P_{pr} -pseudo-reduced pressure
ρ_g- density of gas, lbm/ft3
ρ_{pr}-pseudo-reduced density
T -temperature, R
T_c -critical temperature, R
T_{pr} -pseudo-reduced temperature
z -gas compressibility factor

Reference

[1] Cengel YA, and Boles MA. Thermodynamics an Engineering Approach. 6th ed., Mc Graw-Hill Book Co. 2007.
[2] Standing MB, Katz DL. 1942. Density of natural gases. Transactions of the AIME. 1942; 146 (1): 140-149. doi:10.2118/942140-G.
[3] Katz DL, Cornell D, Kobayashi,R, Poettmann FH, Vary JA, Elenbaas JR, Weinaug CF. Handbook of Natural Gas Engineering. McGraw-Hill 1959, New York, p. 107.
[4] Kamyab M, Sampaio JH, Qanbari F, Eustes AW 2010. Using artificial neural networks to estimate the z-factor for natural hydrocarbon gases. Journal of Petroleum Science and Engineering, 2010; 73 (3-4): 248-257.
[5] Beggs H, and Brill J. 1978. Two-phase flow in pipes, Tulsa 1978, University of Tulsa.
[6] Dranchuk PM, and Abou-Kassem H. 1975. Calculation of Z factors for natural gases using equations of stat". *Journal of Canadian Petroleum Technology, 1975;* 14 (3): 34-36.
[7] Rakap S, Evran D, Cig O. 2016. Comparative evaluation of the reliability and validity of three data extraction programs: UnGraph, GraphClick, and DigitizeIt. Computers in Human Behavior, 2016; 55 (A): 159-166.

To whom correspondence should be addressed: Stanley U. Ibeh, Department of Petroleum Engineering, Madonna University Akpugo, Enugu, Nigeria