OPEN ACCESS

Saudi Journal of Engineering and Technology
Abbreviated Key Title: Saudi J Eng Technol
ISSN 2415-6272 (Print) |ISSN 2415-6264 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Development of Prediction Model for Oil Formation Volume Factor for Sudanese Crude Oil

Hassan Suliman[1*], Ibrahim Elamin[2], Mohamed Ali[1]

[1]Petroleum Labs Researches & Studies Ministry of Oil and Gas, Khartoum, Sudan
[2]Department of Chemical Eng, University of Science and Technology Omdurman, Sudan

DOI: 10.36348/sjet.2024.v09i07.005 | **Received:** 01.06.2024 | **Accepted:** 04.07.2024 | **Published:** 09.07.2024

***Corresponding author:** Hassan Suliman
Petroleum Labs Researches & Studies Ministry of Oil and Gas, Khartoum, Sudan

Abstract

Understanding Oil Formation Volume Factor βo is crucial for effective oil field development, impacting well performance analysis, reservoir simulation, and production engineering calculations. Traditionally, βo is determined through costly and time-consuming laboratory tests, prompting the need for accurate mathematical correlations. Existing correlations such as Vasquez-Beggs, Standing-Glaso, and others have been widely used but show varying degrees of accuracy across different operating conditions. In this study, these correlations were evaluated against 95 datasets of experimental βo data for Sudanese crude oils. Statistical analysis revealed that Vasquez-Beggs and Standing- Glaso models performed best, with average absolute errors of 3.4219 and 3.4477, and correlation coefficients of 0.7563 and 0.7213 respectively. Motivated by the limitations of existing correlations, a new ap- proach using Polynomial Neural Networks (PNN) was developed. This model utilized reservoir temperature, gas gravity, gas oil ratio, and API as input parameters, trained on 70% of the dataset and tested on the remaining 30%. The PNN model exhibited superior predictive performance with a relative average absolute error of 2.8607 and a correlation coefficient of 0.9080. This study contributes a robust predictive tool for estimating βo in Sudanese oil fields, offering enhanced accuracy over traditional correlations and facilitating more reliable reservoir management decisions.

Keywords: Oil Formation, Laboratory Tests, Polynomial Neural Networks (PNN), Temperature, Gas Gravity and API.

I. INTRODUCTION

The volume of oil that enters the stock tank at the surface is less than the volume of oil which flows into the wellbore from the reservoir. This change in oil volume which accompanies the change from reservoir condition to surface condition is due to three factors. The most important factor is the evolution of gas from the oil as pressure is decreased from reservoir pressure to surface pressure. This causes a rather large decrease in volume of the oil when there is a significant amount of dissolved gas. The reduction in pressure also causes the remaining oil to expand slightly, but this is somewhat offset by the contraction of the oil due to the reduction of temperature. The change in oil volume due to these three factors is expressed in terms of the formation volume factor of oil. Oil formation volume factor is defined as the volume of reservoir oil required to produce one barrel of oil in the stock tank. Since the reservoir oil include dissolved gas.

$$\beta_o = \frac{(V_o)_{P.T}}{(V_o)_{\triangle}}$$

The unit are barrels of oil at reservoir conditions per barrel of stock tank oil, res bbl/STB. The volume of stock tank oil is always reported at 60°F, regardless of the temperature of the stock. Thus, stock tank liquid volume, like surface gas volume, is reported at standard conditions. A typical oil formation factor curve, as a function of pressure for an undersaturated crude oil.

II. PROBLEM STATEMENT

Oil formation volume factor determination is important in the oil industry field, it is usually determined from laboratory measurements. Since experimental measurements are usually expensive and time consuming, engineers sometimes use published correlations for oil formation volume factor. These correlations have been proposed for various regions and crude oils with different nature and from various origins. They vary in their complexity and accuracy depending upon the available data of the crude oil. In Sudan, few studies have been done in this area of researches for Sudanese field.

III. RESEARCH OBJECTIVE

1. To calculate oil formation volume factor for Sudanese crude oil using the published correlation models and compares the results with the experimental data.
2. To develop a new predictive correlation model for oil formation volume factor of Sudanese crude oil.
3. Testing and validating the new developed correlation model.

IV. THESIS OUTLINES

This thesis is divided into five parts: part 1 general intro- duction of oil formation volume factor, problem statement and objectives of the study were provided. part 2 contains Importance of PVT Analysis and Oil Formation Volume Factor Measurements and Artificial Neural Network and Polynomial Neural Network then literature review of the Oil Formation Volume Factor correlations part 3 explains the methodology of the correlation models that has been applied in this study to predict the Oil Formation Volume Factor, the statically methods used to choose the suitable published correlation model and validate the new developed correlation and information about the Group Method of Data Handling (GMDH) program which was used to develop the new correlation. Part 4 shows the results and discussions. And finally, part 5 shows conclusion and recommendation.

V. OIL FORMATION VOLUME FACTOR MEASUREMENTS AND LITERATURE REVIEW

A. Importance of PVT Analysis

An accurate description of physical properties of crude oils is of a considerable importance in the fields of both applied and theoretical science and especially in the solution of petroleum reservoir engineering problems. Engineers typically require accurate estimates of crude oil properties in order to compute oil reserves, production capacity, and recovery efficiency of a reservoir. These properties are also used in the analysis of well test and production data, as well as for production engineering activities such as hydrocarbon system optimization and flow measurement. The best source of oil property data is a laboratory PVT (pressure- volume - temperature) analysis of a reservoir fluid sample. Knowledge of the PVT parameters is a requirement for all types of petroleum calculations such as determination of hydrocarbon flowing properties, predicting future performance, designing production facilities and planning methods of enhanced oil recovery. Over the last decade increased attention has been focused on models for predicting reservoir fluid properties from reservoir pressure, temperature, crude oil API gravity and gas gravity.

1) Oil Formation Volume Factor Measurements:

Oil for- mation volume factor is one of oil's most important physical properties. It is often one of the first parameters measured by oil analysis laboratories because of its importance to oil condition. Oil formation volume factor measured by GOR apparatus. The sample was subjected to flash separation from reservoir condition to laboratory temperature and separated into a gas and liquid phases.

2) Artificial Neural Network (ANN):

In biological definition, Natural neurons receive signals through synapses located on the dendrites or membrane of the neuron. When the signals received are strong enough (surpass a certain threshold), the neuron is activated and emits a signal though the axon. This signal might be sent to another synapse, and might activate other neurons. Artificial neural network is that technology initially grew from the full understanding of some ideas and aspects about how biological systems work, especially the human brain. Inspires neural network systems is drawn from many disciplines: primarily from neuroscience; engineering, and computer science; but also from psychology, mathematics, and physics stems are typically organized in layers.

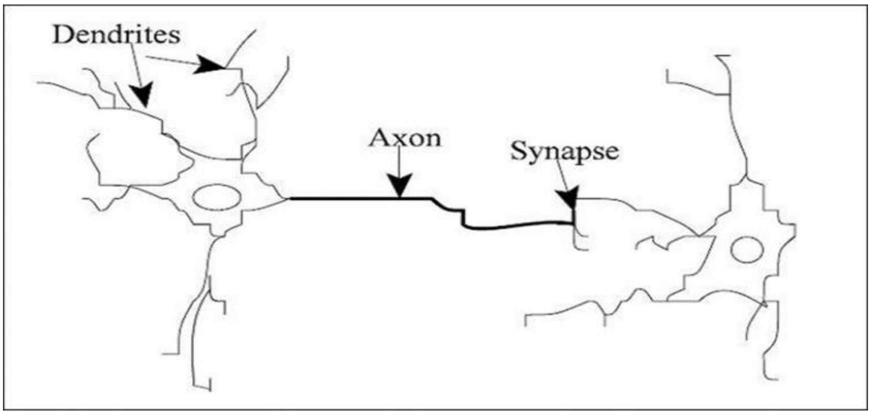

Fig 1: Biological neurons, (Rojas, March 1996)

B. Computing Definition

The simplest definition of a neural network, more properly referred to as an 'artificial' neural network (ANN), is provided by the inventor of one of the first neuro computers, Dr. Robert Hecht- Nielsen. He defines a neural network as: " a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dy- namic state response to external inputs", (Rojas, March 1996).

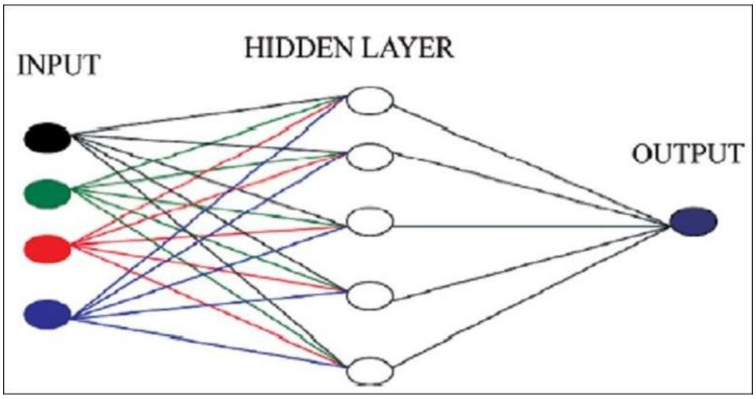

Fig 2: Structure of artificial neural network of present study, (Azubuike, 2013)

C. Polynomial Neural Network (PNN)

The Polynomial Neural Network (PNN) algorithm is also known as International Algorithm of Group Methods of Data Handling (GMDH). GMDH were originally proposed by Prof. A. G. Ivakhnenko. PNN correlates input and target variables using (non) linear regression. In this particular software the user can define the desired properties of the solution such as the number of terms and the maximum degree of polynomials using an approach proposed by Prof. Yu. P. Yurachkovsky, (Ladislav Zjavka, 2011).

D. Oil Formation Volume Factor correlations

Several empirical method have been proposed to estimate the oil formation volume factor, the correlations have been generated using laboratory data, and they were mostly developed originally in graphical forms and mathematical ex- pression of the graphical correlation. Those correlation models were developed by statistical, analytical or numerical methods. The published correlations for oil formation volume factor will be reviewed below. Standing (1947) presented a graphical correlation for estimating the oil formation volume factor with the gas solubility, gas gravity, oil gravity, and reservoir temperature as the correlating parameters. This graphical correlation originated from examining 105 experimental data points on 22 California hydrocarbon systems. An average error of 1.2% was reported for the correlation. Glaso (1980) these correlations were originated from studying PVT data on 45 oil samples. The average error of the correlation was reported at –0.43%.

VI. METHODOLOGY

A. Formation Volume Factor correlations Model

Several empirical methods have been proposed to estimate the Oil Formation Volume Factor; the correlations have been generated using laboratory data from different geographical areas over the world. Eight Oil Formation Volume Factor published correlation equations which were applied for oil formation volume factor of the experimental data of Sudanese crude oil in this study showed below.

B. Standing's Correlation

Standing (1947) presented a graphical correlation for esti- mating the oil formation volume factor with the gas solubility, gas gravity, oil gravity, and reservoir temperature as the correlating parameters. This graphical correlation originated from examining 105 experimental data point on 22 California hydrocarbon systems. An average error of 1.2% was reported for the correlation. Standing (1981) showed that the oil formation volume factor can be expressed more conveniently in a mathematical form by the following equation:

$$B_o = 0.9759 + 0.000120 \left(\frac{\quad}{\gamma_o} \right)^{0.5} + 1.25(T - 460) \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

C. Glaso's Correlation

Glaso (1980) proposed the following expressions for cal- culating the oil formation volume factor: These correlations were originated from studying PVT data on 45 oil samples. The average error of the correlation was reported at –0.43% with a standard deviation of 2.18%.

$$B_o = 1.0 + 10^4 \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$A = -6.58511 + 2.91329 \log B^*_{ob} - 0.27683 \left(\log B^*_{ob}\right)^2$$ ……………. (3)

$$B^*_{\infty} = R_s \left(\frac{\gamma_R}{\gamma_o}\right)^{0.526} + 0.968(T - 460)$$ ………………….……….. (4)

Where: T= temperature °R, *γo* = specific gravity of the stock- tank oil, 60°F/60°F.

D. Marhoun's Correlation

Marhoun (1988) developed a correlation for determining the oil formation volume factor as a function of the gas solubility, stock-tank oil gravity, gas gravity, and temperature. The empirical equation was developed by use of the nonlinear multiple regression analysis on 160 experimental data points. The experimental data were obtained from 69 Middle Eastern oil reserves. The author proposed the following expression:

$B = 0.497069 + 0.000862963T + 0.00182594F + 0.00000318099F^2$ ………………..……. (5)

$$F = R_s^a \gamma_g^b \gamma_o^c$$ ……………….. (6)

Where T is the system temperature in °R and the coefficients a, b, and c have the following values: a= 0.742390 b= 0.323294 c= –1.202040

E. Petrosky and Farshad's Correlation

Petrosky and Farshad (1993) proposed a new expression for estimating oil formation volume factor. The proposed relation- ship is similar to the equation developed by Standing; however, the equation introduces three additional fitting parameters to increase the accuracy of the correlation. The authors used a nonlinear regression model to match experimental crude oil from the Gulf of Mexico hydrocarbon system. Their correlation has the following form:

$$B_o = 1.0113 + 7.2046 \times 10^{-5} A$$ …………………………..…. (7)

$$A = \left[R_s^{0.3738} \left(\frac{\gamma_s^{0.2914}}{\gamma_s^{0.6265}}\right) + 0.24626(T - 460)^{0.5371}\right]$$ …………… (8)

F. almehaideb

In (1997) Almehaideb published a new set of correlations for UAE crudes. He used 62 data sets from UAE reservoirs to develop the new correlations.

$$B_0 = a_1 + a_2 R_s T/\gamma_0^2$$ …………………………… (9)

Where, T= temperature °F, and *γo*= specific gravity of the stock-tank oil, 60°F/60°F. a1= 1.122018 a2= 1.41E-06

G. Macary and El-Batanoney

In (1992) Macary and El-Batanoney presented new correlations for oil formation volume. 90 data sets from 30 independent reservoirs in the Gulf of Suez were used to develop the correlations.

$B_o = (a_1 + a_2 T)\, N$ ……………… (10)

$$N = \exp\left[a_3 R_s + a_4 (\gamma_o/\gamma_g)\right]$$ ………………. (11)

a1= 1.0031 a2= 0.0008 a3= 0.0004 a4= 0.0006

H. Vazquez and Beggs

In (1976) Vazquez and Beggs published correlations for gas oil ratio and oil formation volume factor they started categorizing oil mixtures into two categories, above 30 API gravity and below 30 API gravity. More than 6000 data points from 600 laboratory measurements were used in developing the correlations.

$$B_o = 1 + a_1 R_s + a_2 \left(\gamma_{o\,API}/\gamma_g\right)(T - 60) + a_3 R_s \left(\gamma_{oAPI}/\gamma_g\right)(T - 60)$$ ………………. (12)

Where: a1= 0.0004677 a1= 0.000467 a2= 0.00001751 a2= 0.000011 a3= -1.8106E-08 a3= 1.337E-09 If API ¡= 30 If API ¿= 30

I. Kartoatmodjo and Schmidt

In (1994) Kartoatmodjo and Schmidt presented what should be considered the most comprehensive study of black oil PVT properties. Kartoatmodjo and Schmidt developed a new set of empirical correlations based on a large data collection developed from reservoirs all over the world.

$$B_o = 0.98496 + 0.0001 \times F^{1.50} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (13)$$

With the term F as given by:

$$F = R_{sb}^{0.755}\gamma_g^{0.25}\gamma_o^{-1.5} + 0.45T \quad \ldots\ldots\ldots\ldots (14)$$

J. Error Analysis

Absolute Average Deviation (%AAD), Root Mean Squared Error (RMSE), Standard Deviation (STD) and Regression Co- efficient (R^2) are used to compare and evaluate the prediction ability of correlations, which are defined as below:

$$\%AAD = \frac{\sum_i^n \left| y_i^{exp} - y_i^{pred} \right| / y_i^{exp}}{n} \times 100 \quad \ldots\ldots\ldots\ldots (15)$$

$$RMSE = \sqrt{\frac{\sum_i^n \left(y_i^{exp} - y_i^{pred} \right)^2}{n}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (16)$$

Where n is the number of data points, Yi exp is the oil formation volume factor obtained experimentally and Yi exp is oil formation volume factor. Where Y- is the mean of the observed values of Y.

$$R^2 = 1 - \frac{\frac{1}{n}\sum_i^n \left(y_i^{exp} - y_i^{pred} \right)^2}{\frac{1}{n}\sum_{i=1}^n \left(y_i^{exp} - \bar{y} \right)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (17)$$

The lower value of an error measure (exceptR^2), he closer the model is to the particular data points. While RMSE represents model's deviation from the data, the STD captures how irregular the problem is.

1) Group Method of Data Handling (GMDH) Method:

VariReg is a software tool for general-purpose multidimen- sional regression modeling with the main emphasis on meth- ods used in surrogate modeling. VariReg is primarily intended for use on small and moderately sized numerical data sets. Gints Jekabsons at the Riga Technical University was devel- oped it. It is. The tool provides means for creating "full" polynomial regression models, sparse polynomial models (also called partial polynomial models) employing subset selection algorithms and Polynomial Neural Networks (PNN) induced by Group Method of Data Handling (GMDH). A regression model describes a relation between a vector of d real valued input variables (features) (x1 , x2 ,... x d) d=2,3 and a single real valued output variable y. Using a finite number n of training observations (data cases or data points) (Xi, Yi), i=1,2,...,n. One wants to build a model F that allows predicting the output value for yet unseen input values as closely as possible (Cherkassky and Mulier 2007).

2) Polynomial Neural Networks (PNN):

The GMDH method is mainly referred as a method for self - organizing polynomial neural networks. The most widely known of its variations work exclusively with polynomials and therefore also the result of the method can be written in polynomial form. The building blocks of GMDH usually are second or third degree polynomials of two or three input variables. Such building blocks, also called partial descriptions (PDs), like neurons in neural networks, are arranged in layers. The coefficients of each PD are calculated using the OLS by trying to approximate the original dependent variable y of training data. The exact number of layers and connections of the network as well as the structure of PDs is not set a priori but is the object of search layer by layer. The maximal number of PDs selected in each layer is usually set equal to the number of original input variables (Gints Jekabsons, 2009).

VII. RESULTS AND DISCUSSIONS

From PVT reports for Sudan crude oil , 95 datasets were used in this research for analyzing oil formation volume factor, and to know what is the best empirical among most popular empirical correlations by using the statistical analysis, then developed new correlation using 77 datasets (70% As train data) by using polynomial neural network PNN, and testing the model with 18 datasets (30% As test data), and finally the comparison was done between the best common empirical correlation s and the new PNN model.

A. Developed (GMDH) Correlation Model Equation

Using VariReg software the equation below generated by tow method of regression models, "full" polynomials and Polynomial Neural Networks (PNN) or (GMDH) methods. They have been given same results (see Appendix A).

$$\beta o = -0.0265288561205113 * \gamma g + 1.01947834371573 * H \ldots\ldots\ldots (18)$$

With the term H as given by:

$$H = a_1 + a_2 GOR - a_3 GOR^2 + a_4 GOR^3 + a_5 GOR^2 \gamma g \ldots\ldots\ldots (19)$$

Where

$a_1 = 1.03769628622416$
$a_2 = 0.000653592632236022$
$a_3 = 7.16588244252448 \times 10^{-7}$
$a_4 = 6.12323207174588 \times 10^{-11}$
$a_5 = 4.19595696934857 \times 10^{-7}$

1) Comparison Results of Published Correlation:

The comparison was made in terms of Absolute Average Deviation, Root Mean Squared Error and relative root mean square error and Regression Coefficient in order to find the most suitable one to be applied for Sudanese light crude oil as shown in Table. It has been found from the analysis of Table that the glasos Correlation oil formation volume factor (2) is a suitable model to calculate the oil formation volume factor for data of Sudanese crude oil since %AAD is the lowest values. It can also been shown from Figure that glasos Correlation results is the closest to the 45 line.

Table I: Statistical Analysis for Published Correlations Models

CORRELATION	R2	RMSE	AAD
Standing	0.7213	0.0565	3.4477
glasos	0.7178	0.0568	3.1161
marhouns	0.6120	0.0667	3.7945
petrosky and farshads	-3.3678	0.2238	16.3918
almehaideb	0.4558	0.0790	6.3646
macary and el-batanoney	-0.2220	0.1184	9.8419
vazquez & beggs	0.7563	0.0527	3.4219
kartoatmodjo & schmidt	0.6689	0.0616	4.1269

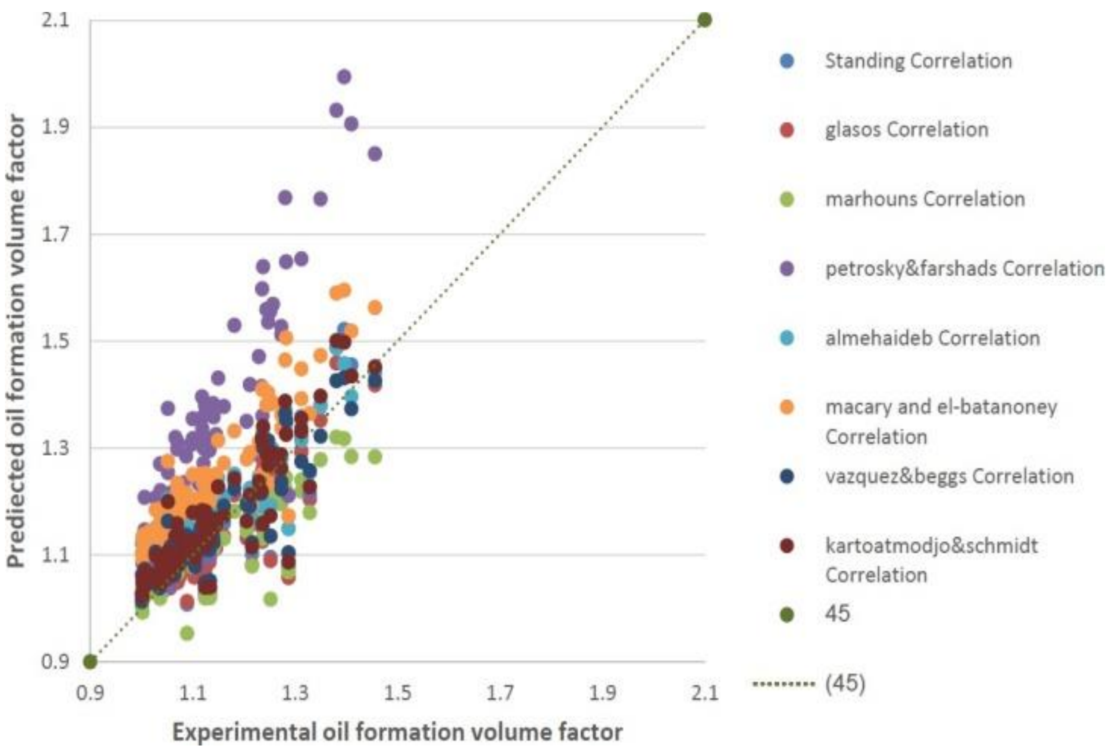

Fig 3: The Results Predicted from Published Correlation vs. the Experimental Data

B. Validation

Validation process is an essential part of developing any new model or correlation. It has been found that the new developed correlation model showed better results in terms of STD, R2, RMSE and AAD% compared with glasos as shown in Table below. And that confirm the performance of the developed correlation model accurst. For a much certainty to the analyzed results accuracy of developed model as shown in figure, which compares the scatter diagrams of this work with glasos correlation and the experimental data. The figure covers 95 data points of the whole data-set.

Table II: Statistical Analysis for Results of Oil Formation Volume Factor Predicted From (GMDH) Model and Glasos Model Compare with Experimental Data

CORRELATION	R2	RMSE	AAD
developed (GMDH) model	0.9080	0.0442	2.8607
glasos	0.7178	0.0568	3.1161

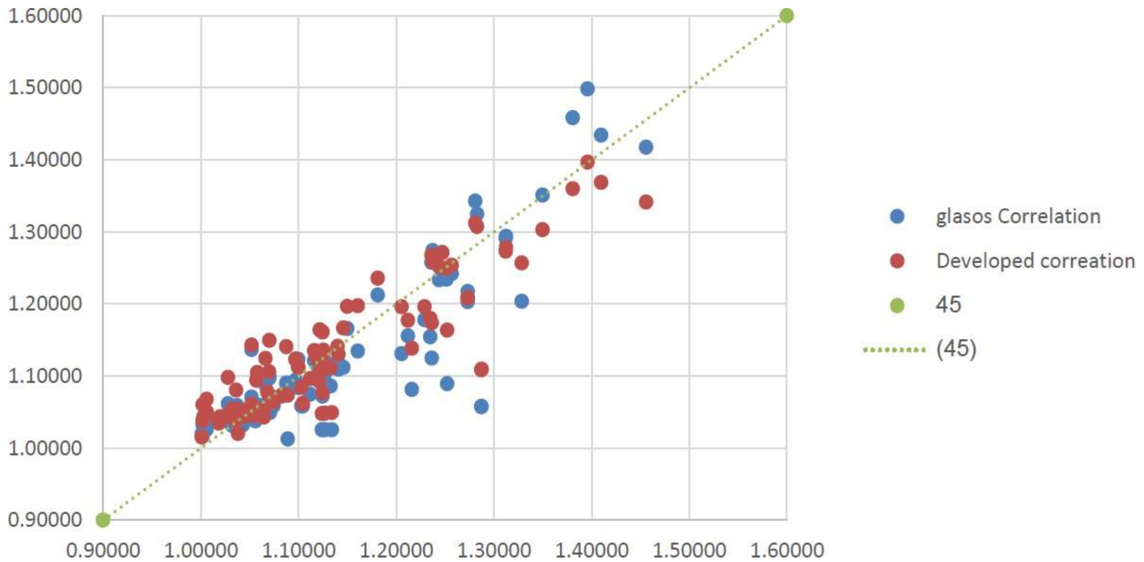

Fig 4: The Results Predicted from (GMDH) Developed Correlation Model glasos Correlation vs. the Experimental Data

VIII. CONCLUSION AND RECOMMENDATION

A. CONCLUSION

- Eight different correlation of oil formation volume factor were applied for crude oil for the 95 data points of Sudanese crude oil. The results showed significant differences compared with the experimental data. This is because these correlations have been proposed for various areas and crude oils and might not give reasonable results for other crude oils and areas.
- Statistical analysis results showed that glasos Correlation is the closest for applicability for Sudanese crude oil prediction amongst the other seven correlations.
- Distinct correlation using (GMDH) method has been proposed for Sudanese crude oils. The equation model de- velopment was a function of the gas oil ratio and gas gravity of the crude oils. The new developed model equation:
  $\beta o = -0.0265288561205113 * \gamma g + 1.01947834371573 * H$ ………………… (20)

With the term H as given by:

$H = a_1 + a_2GOR - a_3GOR^2 + a_4GOR^3 + a_5GOR^2\gamma g + 1$ …………………………………… (21)

- The new model has been validated against experimental data and against glasos model. It has been found that the new model showed better results in terms of statistical factors compared with glasos model.
- The new developed correlation model will be useful and effective tool for estimate Oil Formation volume factor of crude oil in the applications of reservoir simulation, formation evaluation and in designing surface facilities.

B. RECOMENDATIONS

- There is a lack of studies in Sudan on this area of researches for Sudanese crude oils.
- This study, recommends to repeated the study using more experimental data, also to develop this correlation with increase R^2 and decrease AAD.
- Furthers studies are required to include API gravity effect in oil formation volume factor correlation.

REFERENCES

- Ahmed, T. (2013). Equations of state and PVT analysis. Elsevier.
- Velarde, J. J. (1996). Correlation of black oil properties at pressures below the bubble-point (Doctoral dissertation, Texas A M University).
- Al-Shammasi, A. A. (1999, February). Bubble point pressure and oil formation volume factor correlations. In SPE Middle East Oil and Gas Show and Conference (pp. SPE-53185). SPE.
- Jekabsons, G. (2010). VariReg: a software tool for regression modeling using various modeling methods, Riga Technical University.
- Jekabsons, G. (2010). VariReg: a software tool for regression modeling using various modeling methods, Riga Technical University.
- Elkatatny, S., & Mahmoud, M. (2018). Development of new correlations for the oil formation volume factor in oil reservoirs using artificial intelligent white box technique. *Petroleum, 4*(2), 178-186.
- BinMerdhah, A. B., & Baarimah, S. O. Improvement of Oil Formation Volume Factor and Solution Gas Oil Ratio Correlations for Yemeni Crude Oils.
- Saghafi, H. R., Rostami, A., & Arabloo, M. (2019). Evolving new strategies to estimate reservoir oil formation volume factor: Smart modeling and correlation development. Journal of Petroleum Science and Engineering, 181, 106180.
- Baarimah, S., Mustafa, M., Khattab, H., Tantawy, M., & Gawish, A. (2016). Assessment of Oil Reservoir Properties Empirical Correlations. *IJMTER*, 129-147.
- Emara, R. (2015). A New Oil Formation Volume Factor Correlation of Egyptian Crude Oils. Int. J. Adv.
- Singh, T., & Hosein, R. (2012, June). PVT Correlations for Trinidad oil offshore the Southwest Coast. In SPE Trinidad and Tobago Section Energy Resources Conference? (pp. SPE-158805). SPE.
- Vazquez, M., & Beggs, H. (1980). Correlations for fluid physical property prediction. *JPT*, *32*(6), 968–970. SPE-6719-PA. DOI: 10.2118/6719- PA.
- Ayoub, M. A., Elhadi, A., Fatherlhman, D., Saleh, M. O., Alakbari, F. S., & Mohyaldinn, M. E. (2022). A new correlation for accurate prediction of oil formation volume factor at the bubble point pressure using Group Method of Data Handling approach. *Journal of Petroleum Science and Engineering*, 208, 109410.
- Al-Fattah, S. M., & Al-Marhoun, M. A. (1994). Evaluation of empirical correlations for bubblepoint oil formation volume factor. *Journal of Petroleum Science and Engineering*, *11*(4), 341-350.

