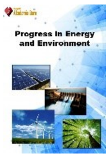
Progress in Energy and Environment

Journal homepage: https://www.akademiabaru.com/submit/index.php/progee/



Volume 15 (2021) 31-39

Original Article

Prediction of bubble point pressure for Sudan crude oil using Artificial Neural Network (ANN) technique

OPEN ACCESS

Sami Abdelrahman Musa Yagoub*, Gregorius Eldwin Pradipta, Ebrahim Mohammed Yahya

Department of Chemical and Petroleum Engineering, Faculty of Engineering, Technology and Built Environment, UCSI University, Jalan Puncak Menara Gading, 56000 Kuala Lumpur, Malaysia

Abstract

In this study, a new Artificial Neural Network predictive model was developed to determine the bubble point pressure P_b for Sudanese oil field using ANN tools in MATLAB software. Because of limitations of the experimental procedures and the time taken to obtain bubble point pressure value from the reservoir fluid samples analysis, an alternative is required where many researchers have been conducting research on the use of Artificial Neural Network (ANN) techniques. In the present study ANN model was developed and evaluated using 151 experimental data sets for Sudanese oil field, and more 61 data sets are used to compare the developed model with universal and regional published models. Comparing with universal and global empirical models, the developed model of ANN for P_b pressure has better precision index of correlation 94.63% with MSE and RMSE of 180 and 156, respectively. However, the results show that of ANN has a lower performance than regional PNN model, as the PNN shows index of correlation 97.57% with MSE and RMSE of 88 and 101, respectively. This difference may be due to the limitation in number of variables and number of data points used in each model developed. Thus, the ANN developed model in this study might be improved to predict the P_b especially for Sudan oil fields and similar oil field properties in regional by increasing the data point used in ANN model training.

Article Info

Received 4 August 2020

Received in revised form 8 January 2021

Accepted 11 January 2021

Available online 12 January 2021

Keywords

Artificial Neural Network

Pressure-Volume-Temperature (PVT)

Prediction of bubble point pressure

1 Introduction

Bubble point pressure is the highest pressure in the hydrocarbon system when the oil starts to form gas and leave the oil. It is one of the petroleum system's important properties in Pressure-Volume-Temperature (PVT). In a reservoir and production engineering, it has a role along with the other calculation indicators for creates PVT analysis. It consists laboratory procedure aimed at providing value of reservoir fluid properties that useful in material balance calculation, well test analysis, flow performance calculation, reservoir simulation, Enhance Oil Recovery (EOR) project, economic evaluation.

In several decades, the only ways to obtain bubble point pressure value is by experimental laboratory. Due to prejudices of experimental procedures and the possibility of human errors, many researchers have presented predictive correlations with intention to lighten the estimation of PVT properties of crude oil and in cases the experimental method is not available. The reasons for using these PVT Correlations are monetary matters, deficient non - representative sample quality, fluid, and human error

* **Corresponding author** SamiAbdelrahman@ucsiuniversity.edu.my

during sampling or fluid transfer, lack of sample volume in order to get complete analysis, validation in PVT experimental analysis to avoid significant error [1]. Linear regression model is used to predict the PVT correlations from the experimental data sets. These correlations can be obtained using software tools and statistical algorithms also.

In the recent past, researchers started to exploit artificial neural network as one of vigorous and dependable tools for collecting data and solving mathematical application for petroleum system prediction in terms of PVT properties. The accuracy gain from ANNs model shows that it has better precision in bubble point pressure prediction. Feed forward together with back propagation algorithms are the most common neural network and training algorithm [2].

As we mentioned above several models have been proposed calculating the bubble point pressure are available. They include Ahmed et al. [2], PNN [3], Nada et al. [4], Elmabrouk [5], Al-marhoun et al. [6], BP-ANN [7], Salem et al. [8], Heidarian et al. [9], and Abdelmajeed et al. [10]. These models vary between regional and universal or general models, which may not work or perform when applied in different regions. Thus, there is a massive need to develop regional models or correlations rather than developing universal general ones, and this depend on the availability of the data.

This paper aims to develop predictive ANN mode that is reliable in estimating bubble point pressure at certain conditions for Sudanese crude oil. ANN tools billed in MATLAB software was used to develop the P_b model and compared to published regionally and universally predictive correlations. Collected data points consist of 105 points. The data was divided into three groups. seventy percent of points was used to train the ANN models, fifteen percent was used for testing and fifteen percent was used for validating the developed ANN model.

2 Methodology

ANN tools built in MATLAB software could be used to develop ANN prediction model of Pb for Sudanese oil fields. Collection of 151 reservoir fluids datasets form Sudan crude oil field were used before to develop PNN predictive model [3], the same data pointes proposed to be utilized in this study for training, testing and validation. The dataset is divided into two sets of 70% for training, 15% for testing and 15% for validation. Then the process of ANN making was based on the datasets obtained, and the best performance achieved from many trials and errors after the process of training and testing. The best structure of ANN then gone through various statistical analysis to evaluate and validate its performance. The determination for ANN best structure will base on the statistical analysis from its error in training, testing and validation, then have the simplest structure. The statistical analysis used in this study are the root square (R^2), the root means square error (RMSE), the mean square error (MSE), and the coefficient factor (R). The statistical analysis equations are listed in Eqs. (1) – (4) as following.

$$R = \sqrt{1 - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(P_{\text{b,experimental}} - P_{\text{b,prediction}}\right)}{\frac{1}{n}\sum_{i=1}^{n}\left(P_{\text{b,experimental}} - P_{\text{b,avg,exp}}\right)}} \tag{1}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_{\text{b,experimental}} - P_{\text{b,prediction}}\right)^2} \tag{2}$$

$$R^2 = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(P_{\text{b,experimental}} - P_{\text{b,prediction}}\right)}{\frac{1}{n}\sum_{i=1}^{n}\left(P_{\text{b,experimental}} - P_{\text{b,avg,exp}}\right)} \tag{3}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(P_{\text{b,experimental}} - P_{\text{b,prediction}}\right)^2 \tag{4}$$

After the network processing techniques designed, the ANN model structure and Optimal model with highest correlation will be generated with optimally select model input parameters. The developed ANN model is gone through validation and evaluation processes. In this process, the model was compared with published model such as Polynomials Neural Network [3], with three commons empirical of Standing, Vasquez-Beggs, Abdelmajeed et al. [10]. Based on Table 1, the used input layer parameters for ANN different algorithms are three to four parameters. Thus, in this study we will use the following four parameters, which are reservoir temperature (T_R), oil gravity (API), gas specific gravity (γ_g) and solution gas-oil ratio (R_s), to predict the bubble point pressure (P_b).

Table 1 ANN Models developed for regional and universal application to determine the bubble point pressure (P_b).

No	Author/ANN Model	Year	Parameters	Target Area	Schematic
1	Nada et al. [4]	2012	$\gamma_0, \gamma_g, R_s, T_R$	Iraq Field	4 - 9 - 1
2	Elmabrouk [5]	2012	$P_{sp}, \gamma_{ost}, R_{sp}, T_R$	Libya	4 - 9 - 4 - 1
3	Al-Marhoun et al. [6]	2014	$\gamma_0, \gamma_g, R_s, T_R$	Canadian	4 - 20 - 1
4	BP-ANN [7]	2015	$\gamma_0, \gamma_g, R_s, T$	Middle East	4 - 25 - 10 - 5 -1
5	Salem et al. [8]	2015	$\gamma_0, \gamma_g, R_s, T_R$	Worldwide	4 - 11 - 22 - 1
6	PNN [3]	2016	γ_0, γ_g, R_s	Sudan	-
7	Heidarian et al. [9]	2017	$\gamma_0, \gamma_g, R_s, T_R$	Middle East	Genetic Algorithm (GA)
8	Ahmed et al. [1]	2018	$\gamma_0, \gamma_g, R_s, T_R$	Worldwide	4 - 6 - 1
9	SaDE-ANN [11]	2018	$\gamma_0, \gamma_g, R_s, T_f$	Worldwide	3 - 18 - 17 - 1

3 Results and Discussion

3.1 ANN Architecture

This study applied the multilayer perceptron (MLP) Artificial Neural Network (ANN) technique to developed ANN bubble point pressure predictive model, developed model will named as Levenberg Marquardt- Artificial Neural Network (LM-AN). The network was trained using 70% of obtained data points. The network processing techniques designed has four (T_R, API°, R_s, γ_g) and three (API°, R_s, γ_g) inputs parameters, one hidden layer, with different nodes numbers to select the best model structure based on Coefficient Factor (R) and RMSE values. The network was trained with the Levenberg Marquardt training algorithm and activated by the Hyperbolic Tangent transfer function. Refer to Table 2, the MLP-ANN generated model with structure of four inputs, sixteen neurons in a hidden layer and single output (4 – 16 – 1) has a higher coefficient factor 0.9950, precision and low RMSE 89.8, that indicate it has lower error than the other models. In Fig. 1, it could be seen the distribution of cross plot of 3 input parameters is closer to the 45-degeee line which means it almost has similar result with the target set, also as in Table 3 the 3 input parameters model shows higher coefficient 0.9926 value and lower RMSE 81.5 values compare to the 3 input parameters model. In this study the selected LM-ANN model for the 4 input parameters was selected as best model base on the results of the validation and testing stage of both models as shown in Fig. 2, Fig. 3, Table 7, and Table 8.

Table 2 Comparison of each LM-ANN Model for training data.

LM-ANN Model	Coefficient Factor (R)	RMSE
4 – 2 – 1	0.9753	262.5358
4 – 15 – 1	0.9722	184.8795
4 – 16 – 1	0.9950	89.7832
4 – 20 – 1	0.9857	183.6388
4 – 100 – 1	0.6525	276.49
4 – 15 – 5 – 1	0.9885	129.5475
4 – 10 – 10 – 1	0.8827	597.5163
4 – 15 – 10 - 1	0.9627	246.7843

Table 3 Comparison of 3 and 4 input of LM-ANN Model for training data.

No	LM-ANN Model	R	R^2	RMSE	MSE
(1)	4 – 16 – 1	0.9950	0.9900	89.8	4.86
(2)	3 – 16 – 1	0.9963	0.9926	81.55	76.91

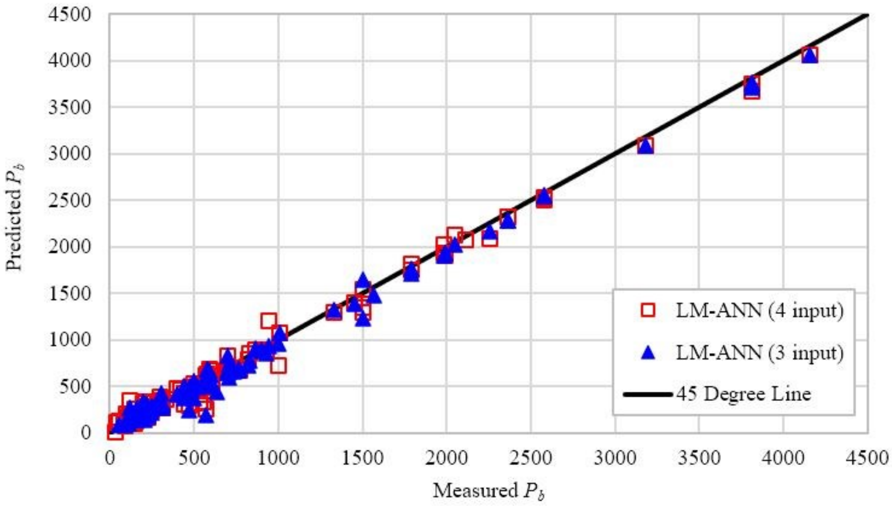


Fig. 1 Cross Plot of developed predicted bubble point pressure models by using three and four input parameters vs. measured values (training dataset).

3.2 LM-ANN Developed Model

From the model, an equation could be generated to estimate the P_b using API, γ_g, R_s, T_f as the parameters. The information of each parameter was displayed in Table 4. The equations are exposed as follows:

$$P_b = \frac{P_{bn}+1}{0.0005} + 31 \tag{5}$$

$$P_{bn} = \sum_{i=1}^{n}\left(\omega_{2,i}x_i\right) + b_2 \tag{6}$$

$$x_i = \frac{2}{1+\exp\left[-2\left(\omega_{i,1}T_{fn} + \omega_{i,2}\mathrm{API}_n + \omega_{i,3}\gamma_{gn} + \omega_{i,4}R_{sn} + b_{1,i}\right)\right]} - 1 \tag{7}$$

$$\mathrm{API}_n = 0.0407\left(\mathrm{API} - 15.9\right) - 1 \tag{8}$$

$$\gamma_{gn} = 2.0582\left(\gamma_g - 0.54\right) - 1 \tag{9}$$

$$R_{sn} = 0.0023\left(R_s - 3.2\right) - 1 \tag{10}$$

$$T_{fn} = 0.0147\left(T_f - 107.6\right) - 1 \tag{11}$$

where:

n = Number of the hidden layer (n=1)
i = Neuron index, as shown in Table 5 and Table 6
$\omega_{1,i}, b_{1,i}$ = Bias and weight of the input and the first hidden layer, as shown in Table 5
$\omega_{2,i}, b_2$ = Bias and weight of the first hidden layer and the output layer, as shown in Table 6
API_n = Normalized oil API gravity, as calculated by Eq. (8)
γ_{gn} = Normalized gas gravity, as calculated by Eq. (9)
R_{sn} = Normalized oil/gas ratio, as calculated by Eq. (10)

T_{fn} = Normalized temperature in Fahrenheit, as calculated by Eq. (11)

P_{bn} = Normalized bubble point pressure, as calculated by Eq. (6)

Table 4 Statistical descriptions of all data samples.

Parameters	Units	Minimum	Average	Maximum
Measured P_b	Psi	31	724.431	4155
Temperature T	ºF	107.6	178.4	244
Oil/Gas Ratio R_s	SCF/STB	1.2	135.2	877.7
API Gravity	^{0}API	15.9	31.4	65
Gas Specific Gravity	Dimensionless	0.54	0.9	1.53

Table 5 LM-ANN weight bias from input layer to hidden layer.

i	**$\omega1$ (Input to Hidden Layer)**				***$b_{1,i}$***
	$\omega1_{i,1}$ (T_f)	$\omega1_{i,2}$ (API)	$\omega1_{i,3}$ (Gas gravity)	$\omega1_{i,4}$ (Oil-gas ratio)	
1	0.9895	0.9527	0.9568	0.9891	1.0088
2	1.0179	1.1605	0.5483	1.2157	1.0440
3	1.0549	0.7927	0.8353	0.9604	0.8733
4	0.2034	0.5269	0.1076	0.7346	1.5157
5	0.9290	0.9689	0.9229	1.0299	1.1505
6	0.8321	0.6839	0.3871	1.9296	0.4843
7	0.9286	1.4191	0.7031	1.7412	0.9700
8	0.7925	1.2139	0.5088	0.4990	0.5056
9	0.9868	0.9536	0.9567	0.9906	1.0128
10	-1.7754	-0.2449	-1.4173	2.0527	-0.3466
11	1.9312	0.0061	1.5675	-1.7250	0.7535
12	0.9180	0.9888	0.7730	1.0347	1.3132
13	1.1252	0.6174	0.4884	1.8687	-0.9983
14	0.7165	0.7929	0.4576	1.4502	0.8820
15	0.4018	0.6291	1.6061	3.5431	-0.8629
16	1.0497	1.1834	0.5149	1.2566	0.9973

3.3 Evaluation of LM-ANN Developed Model Performance

3.3.1 Validation of LM-ANN Model

LM-ANN developed model has been validated by using 15% of collected data to verify the equation accuracy and performance. Fig. 2 shows the cross plots with 45-degree line and the validation data for the 4 input and 3 input parameters models, the 4 input has a perfect matching with the fitted line compare to the 3 input parameters model data. If refer to Table 7, LM-ANN (4 input) model scores really high accuracy and it has 110.5 of RMSE which is the lowest compare to the LM-ANN (3 input) model which has 177.2 of RMSE.

3.3.2 Testing of LM-ANN Model

LM-ANN developed model performance was tested by using 15% of collected data. From Fig. 3, the LM-ANN (4 input) developed model shows closer distribution of P_b prediction to the measured P_b, also by referring to Table 8, LM-ANN (4 input) model has higher accuracy and correlation coefficient factor than LM-ANN (3 input) model with 0.9525 and 0.9404, respectively. LM-ANN (4 input) model also score really high accuracy and it has 93 RMSE which is the lowest compare to the LM-ANN (3 input) model which has 175 of RMSE.

Table 6 LM-ANN weight bias from hidden layer to output layer.

i	*ω*2 (Hidden Layer to Output)	b_2
1	-0.0134	-1.1593
2	0.4281	
3	-0.2508	
4	1.0729	
5	0.1552	
6	-0.7685	
7	0.6767	
8	-1.4776	
9	-0.0082	
10	1.7725	
11	1.6916	
12	0.3564	
13	2.2548	
14	-0.4273	
15	-1.3749	
16	0.4449	

Table 7 Statistical analysis of LM-ANN developed model for (validation dataset).

Model	R^2	R	RMSE
LM-ANN (4 input)	0.9865	0.9932	110.5
LM-ANN (3 input)	0.9103	0.9541	177.2

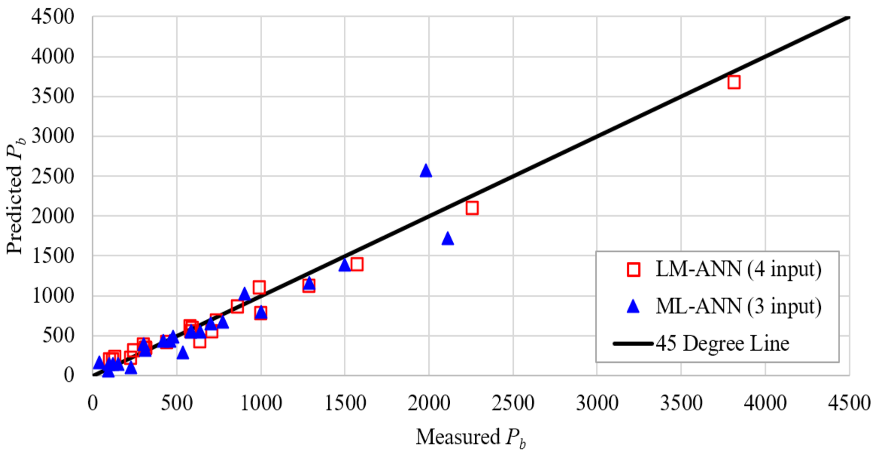


Fig. 2 Cross plot of developed predicted bubble point pressure models by using three and four input parameters vs. measured values (validation dataset).

3.3.3 Extra Validation of LM-ANN Model

For further validation, an additional collection of data not used in constructing the LM-ANN developed model compare to the global and regional published models was used in this sub-section. This will give us an idea on how suitable our model predictive ability and its performance is compared to the currently available models.

The common empirical models used in this study were Standing, Vasquez and Beggs, and the regional predictive models were PNN and Abdelmjeed et al. [10] for Sudanes oil field. Based on Fig. 3 and Table 9, the validation was focused on Standing due to this empirical were the most suitable for the data compare to Vaquez Beggs. So that the LM-ANN compare to Standing empirical models the ML-ANN (4 input) is scores really high accuracy and it has 0.9463 of correlation coefficient factor and lowest 158 of RMSE compare to Standing empirical model. Also refer to Fig. 3 and Table 9, the LM-ANN models performance was evaluated with statistical analysis compare to PNN and Abdelmjeed et al. [10] regional models of Sudanese oil fields. Comparison shows that the ML-ANN (4 input) has higher accuracy and correlation coefficient factor than Abdelmjeed et al. [10] regional model with 0.9463 and 0.7210 respectively. But the PNN regional model has higher accuracy and correlation coefficient factor compare to ML-ANN (4 input) developed model, with correlation coefficient factor with 0.9757 and 0.9463 respectively, this due the amount of data point used in construct both models, since PNN models developed based on 212 data points and 151 data point was used to develop the LM-ANN models. Thus, we recommend to improve LM-ANN models by increasing the data amount to generate more an accurate ANN predictive models.

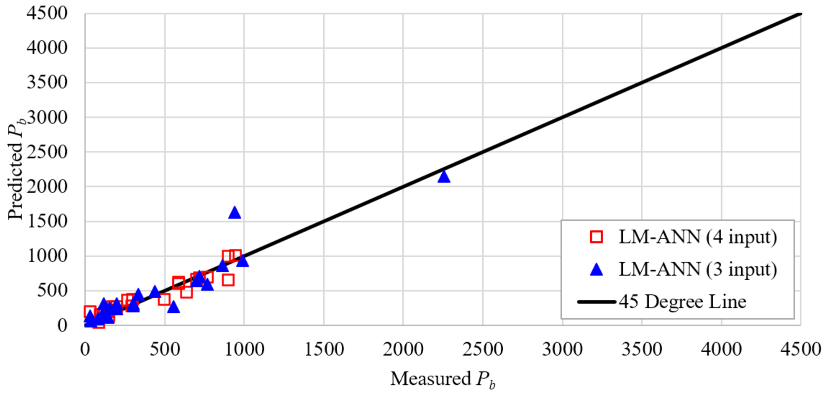


Fig. 3 Cross plot of developed predicted bubble point pressure models by using three and four input parameters vs. measured values (testing dataset).

Table 8 Statistical analysis of LM-ANN developed model for (testing dataset).

Model	R^2	R	RMSE
LM-ANN (4 input)	0.9073	0.9525	93
LM-ANN (3 input)	0.8843	0.9404	175

Table 9 Statistical analysis of LM-ANN and published models for extra dataset.

Model/Empirical	Target Area	R	R^2	RMSE
LM-ANN (4 input)	Sudan	0.9463	0.8954	156
LM-ANN (3 input)	Sudan	0.9149	0.8370	189
PNN	Sudan	0.9757	0.9520	101
Standing	Common	0.9186	0.8439	225
Vasquez Beggs	Common	0.8888	0.7901	1797
Abdelmajeed et al. [10]	Sudan	0.7210	0.5199	587

4 Conclusion

A conclusion can be made based on the explanation and the results above. The LM-ANN models were proposed depending on 151 data points to predict the P_b models for Sudanese crude oil. This model has four inputs, one hidden layer (16 neurons) and one output. After, through several modifications, LM-ANN predictive model achieved 99.50% of accuracy with mean square error and root mean square error, 4.86 and 89.8, respectively. LM-ANN developed model performances is outperforming the others common and regional empirical predictive models stated in this research. However, the regional PNN predictive model for Sudanese oil field shows better performance than LM-ANN models, this due to the amount of data points used in construct PNN and LM-ANN models which are 212 and 151, respectively. Thus, the LM-ANN model can be enhanced in the future with applying a greater number for training the model.

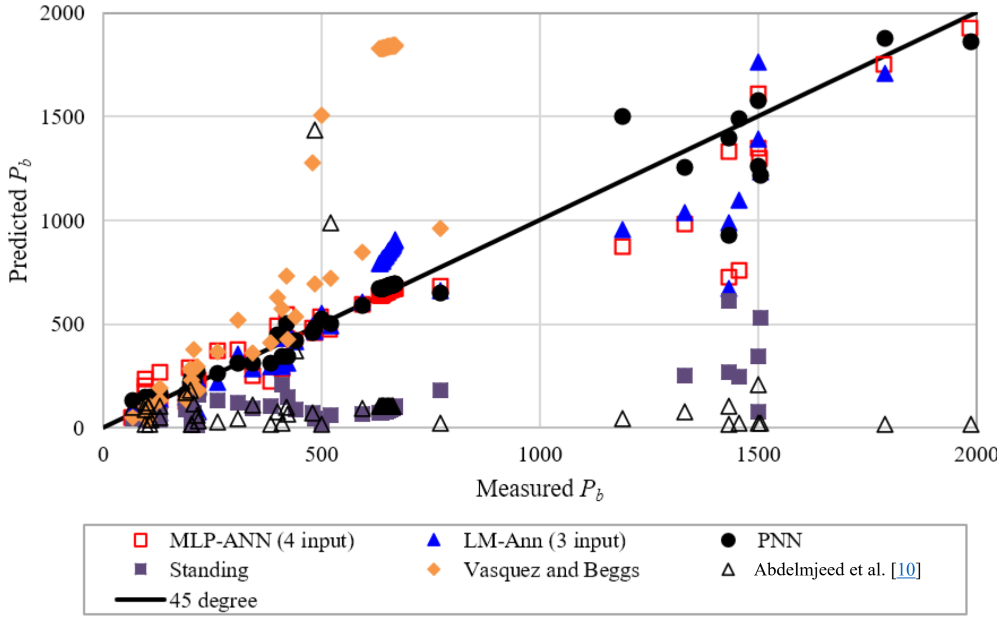


Fig. 4 Cross Plot of LM-ANN developed model, Global and Regional published models vs. measured values (extra dataset).

Declaration of Conflict of Interest

The authors declared that there is no conflict of interest with any other party on the publication of the current work.

ORCID

Sami Abdelrahman Musa Yagoub https://orcid.org/0000-0001-7642-5812

References

[1] A. Ahmed, E.-B. Ahmed, H.S. Mohamed, Development of an expert system for reservoir fluid PVT properties correlations, in: 60th Annual Southwestern Petroleum Short Course, 2013.
[2] A. Hashemi Fath, A. Pouranfard, P. Foroughizadeh, Development of an artificial neural network model for prediction of bubble point pressure of crude oils, Petroleum. 4 (2018) 281–291. https://doi.org/10.1016/j.petlm.2018.03.009.
[3] H. aldeen M.M. Mustafa, M.A.M. Elgasim, M.A.K. Idrees, New correlation for prediction of bubble point pressure for Sudan crude oil, 2016. http://repository.sustech.edu/handle/123456789/15198 (accessed January 12, 2021).
[4] S.A. Nada., M.R. Iqdam, M.A. Yazan, T.A. Zainab, Development of PVT correlation for Iraqi crude oils using artificial neural network, Iraqi Journal of Chemical and Petroleum Engineering. 13 (2012) 9–16. https://ijcpe.uobaghdad.edu.iq/index.php/ijcpe/article/view/337 (accessed January 12, 2021).
[5] S.K. Elmabrouk, Application of function approximations to reservoir engineering, 2012.
[6] M.A. Al-Marhoun, S.S. Ali, A. Abdulraheem, S. Nizamuddin, A. Muhammadain, Prediction of bubble point pressure from composition of black oils using artificial neural network, Petroleum Science and Technology. 32 (2014) 1720–1728. https://doi.org/10.1080/10916466.2012.707267.
[7] K.S. Arjun, K. Aneesh, Modelling studies by application of artificial neural network using matlab, Journal of Engineering Science and Technology. 10 (2015) 1477–1486.

[8] O.B. Salem, A.G. Ahmed, B.B. Amer, Artificial intelligence techniques for predicting the reservoir fluid properties of crude oil systems, International Research Journal of Engineering and Technology. 2 (2015) 373–382.
[9] M. Heidarian, M. Karimnezhad, M. Schaffie, M. Ranjbar, A new empirical correlation for estimating bubble point pressure using the genetic algorithm, Geology, Geophysics & Environment. 43 (2017) 33. https://doi.org/10.7494/geol.2017.43.1.33.
[10] A. Mohamed, M. Ahmed, I. Mohamed, M. Abdein, The performance of the popular PVT correlations on Sudanese crude oils: A statistical study, in: SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, Society of Petroleum Engineers, 2018. https://doi.org/10.2118/192438-MS.
[11] S. Elkatatny, T. Moussa, A. Abdulraheem, M. Mahmoud, A self-adaptive artificial intelligence technique to predict oil pressure volume temperature properties, Energies. 11 (2018) 3490. https://doi.org/10.3390/en11123490.