

Fuzzy Sequential forward search for Oil Formation Volume Factor Predictive Tool Factor for Niger Delta Crude Oil

K. K. Salam^{a*}, D. Araromi^a, S. S. Ikiensikimama^b & A. O. Arinkoola^a

^a Petroleum Engineering Unit, Department of Chemical Engineering,
Ladoke Akintola University of Technology, Ogbomoso, Nigeria

^b Department of Petroleum & Gas Engineering, University of Port
Harcourt, Port Harcourt, Nigeria

Kazeem Kolapo SALAM (R. Eng)

Department of Chemical Engineering,

P. M. B. 4000,

Ladoke Akintola University of Technology(LAUTECH),

Ogbomlso. Oyo State.

Nigeria.

+234-805-706-6800

+234-703-438-4645.

emails: kaykaysalam@yahoo.co.uk

kaykaysalam@gmail.com

Abstract

Accurate prediction of fluid properties are essentials for all reservoir engineering calculations such as estimation of reserves, reserve forecasting, well testing analysis and in numerical reservoir simulation. Oil formation volume factor is one of the properties that can either be gotten from empirical or experimental method.

This work focuses on the use of fuzzy sequential forward techniques to develop a oil formation volume factor model using 1,316 data obtained from 45 different oil fields in the Niger Delta, Nigeria. The data set was randomly divided into two parts with 570 used for training and 566 for testing. The model developed has the lowest Root Mean Square Error (RMSE) of 0.0784 when compared with published correlation used for prediction. The accuracy of the developed model was tested with crossplot and statistical analysis. The model developed outperformed the existing correlations when subjected to statistical analysis.

Keyword: Formation volume factor, Fuzzy inference system, correlation, Artificial Intelligence, Root mean square error

1. Introduction

Accurate prediction of fluid properties are essentials for all reservoir engineering calculations such as estimation of reserves, reserve forecasting, well testing analysis and in numerical reservoir simulation (Al-Marhoun and Osman, 2002, Meisam *et al*, 2007). Any small error in measurement of these properties results in inappropriate reservoir estimation and poor performance of the reservoir model. One of these properties is Formation Volume Factor (FVF) which is defined as the volume of reservoir oil that would be occupied by one stock tank barrel plus dissolved gas at the bubble point pressure and reservoir temperature (Osman, *et al*, 2001).

Formation Volume Factor of oil can be obtained by either experimental method or estimated using empirically derived Pressure Volume Temperature (PVT) correlations. Experimental method gives accurate results under controlled conditions however, the results are heavily dependent on the validity of the reservoir fluid samples especially when reservoir has been depleted below bubble point pressure (Hemmati and Kharrat, 2007). Empirical correlations are mostly developed using either multiple regression techniques or graphical techniques with each correlation developed for a certain range of fluid characteristics and geographical areas with similar fluid compositions. However, the accurate prediction of the correlations are not known beforehand and the correlations often perform poorly outside the range of data used (Osman, *et al*, 2001, Hemmati and Kharrat, 2007). In the last half a century, a lot of correlations have been developed predict FVF and other reservoir physical properties. A comprehensive review can be found in the works of (Al-Shammasi, 1999, Goda *et al*, 2003).

Artificial neural network (ANN) is a computer-based algorithm consisting of a pool of simple processing units which communicate by sending signals to each other over a large number of weighted connections to learn the behavior of a data population by self-tuning its parameter in such a way that the trained ANN matches the employed data accurately (Obanijesu and Araromi, 2008). Some of ANN advantages can be found in the work of (Graupe, 2007). This principles has been widely used in solving petroleum engineering problems and example is the work of Garbi and Elsharkawy (1997) who used ANN to solve two problems. Other works on PVT properties prediction using ANN are Goda *et al*, 2003, Osman *et al*, 2001, Al-Marhoun and Osman, 2002. In other areas of petroleum engineering, Goda *et al*, 2007 used ANN for prediction of Irreducible water saturation in Australian Hydrocarbon basin, Dewpoint pressure

prediction for gas condensate reservoir (Meisam *et al*, 2007), Permeability prediction(Singh. 2005) and so on.

ANN suffers some limitations despite its wide application in solving petroleum engineering problem. Predictive tools developed based on ANN usually require a large number of adjustable parameters (network weights) with attendant problems such as relative high computational load for weight adaptation (network training). Also, insufficient data set might result to over fitting which may consequently lead to a loss of generalization capacity (Cubillos ,2001). Fuzzy logic which is also an AI tool is capable of overcoming the two limitation encountered in the use of ANN for solving problems (Aguilar and Cerrada, 2001). The fuzzy system uses linguistic variable classification such as LOW, HIGH and MEDIUM to incorporate the imprecision inherent in many real systems including human reasoning; thus providing the ability to capture knowledge from data that is inherently imprecise and maintaining a high level of performance in the presence of uncertainty (Brian and Ben, 2006). This is important because most measured data are inherently imprecise because of individual differences, measurement errors and noise. Other key advantageous feature of this approach lies in its ability to arrive at transparent model that is easily interpretable. Salam *et al* (2011) demonstrated superior performance of below bubble point viscosity model developed through Fuzzy logic techniques when compared with widely used empirical correlations in the industry. This work will develop a prediction tool using Fuzzy logic using data collected across fields in Niger delta region in Nigeria and the model will be tested and validated using existing correlations used in the region.

2. Data gathering

The data used for this study were obtained from conventional PVT reports that derived the various fluid properties through differential liberation process from different oil fields in the Niger Delta. A total of 245 PVT field reports were used to extract 1316 datasets for Fuzzy prediction of the undersaturated oil FVF. Table 1 shows the PVT input data range, average and standard deviation of the parameters used for this study.

Table 1: Range of parameter used to develop the model

	Min	Max	Average	Stad dev.
Formation volume factor above bubble point	1.022	3.2636	1.4887	0.3896
Formation volume factor at bubble point	0.134	3.2705	1.4758	0.412
Gas gravity	0.564	1.29	0.7461	0.136
Gas Oil Ratio	19	2948.8	861.76	645.15
Oil gravity	0.766	0.964	0.8565	0.0465
Bubble point Pressure	67	6560	2928.9	1226
Initial Pressure	115	8451	3794.1	1237.3
Temperature	104.4	264	172.79	29.76
API gravity	15.2	53.23	34.19	8.8018
Compressibility factor	3.4×10^{-7}	0.0017	1.656×10^{-5}	7×10^{-5}

3. Data Formulation for the model

Sequential Forward Search (SEQSRCH) in which each input is selected sequentially to optimize the total squared error was used to determine the influence of all the eight inputs on the prediction of Formation volume factor. This was achieved by employing SEQSRCH function of fuzzy logic toolbox in Matlab. Essentially, SEQSRCH builds an ANFIS model for each combination and trains it for one epoch and reports the performance achieved. The function SEQSRCH performs an exhaustive search within the available inputs to select the set of inputs that most influence the output. The variables are Gas gravity, API gravity, Oil gravity, Temperature, Above bubble point pressure, Gas oil ratio, bubble point Formation volume factor.

Table 2: Considering all parameters		
Parameter	Training	Checking
A	0.0896	0.1141
B	0.0975	0.3744
C	0.1527	0.6607
D	0.1244	0.596
E	0.1319	0.6604
F	0.1241	0.5954
G	0.1629	0.7222
H	0.1262	0.5985
AB	0.0884	0.8291
AC	0.0893	0.1108
AD	0.0891	0.1051
AF	0.1157	0.5733
FE	0.1183	0.5971
FG	0.1102	0.5839
FH	0.0998	0.1272
ABC	0.0864	0.6241
ABD	0.0874	2.203
FHE	0.0999	1.1259
FHG	0.1039	0.8919
ABF	0.0876	2.1528
ABCD	0.0843	0.7035
EFHG	0.0944	1.3304
ABCF	0.0857	1.4954

The results of the search were categorised into two set namely: Type 1 which comprises of A- Above bubble point pressure, B - Formation volume factor at bubble point, C - Oil gravity, and D - Gas oil ratio; and Type 2 comprises of E - API gravity, F - gas gravity, G - Temperature and H - compressibility factor. Table 2 show the effect of individual variable on the proposed model. Considering their RMS error the best three variable chosen are Oil gravity, Formation volume factor at bubble point, and Gas-oil ratio. The fourth variable chosen by the tool is Above bubble point pressure which have a higher RMS error value when compared with API gravity which should be fourth when considering the order of their RMS error. When the parameters in both groups are combined, the RMS value of the parameters in Type 1 are all less than 0.1 while all the parameters in Type 2 are all greater than 0.1. Based on the RMS error of single input and in

combination of inputs, inputs in Type 1 were selected for the prediction of above bubble point formation volume factor of oil.

3.1. Neuro-fuzzy modelling and the adaptive network fuzzy inference system

Neuro-fuzzy model adopted in this work is based on the Adaptive Network based Fuzzy Inference System (ANFIS) (Jang, 1993). The final output of the system is the weighted average of all rule outputs, computed as shown in Figure 1.

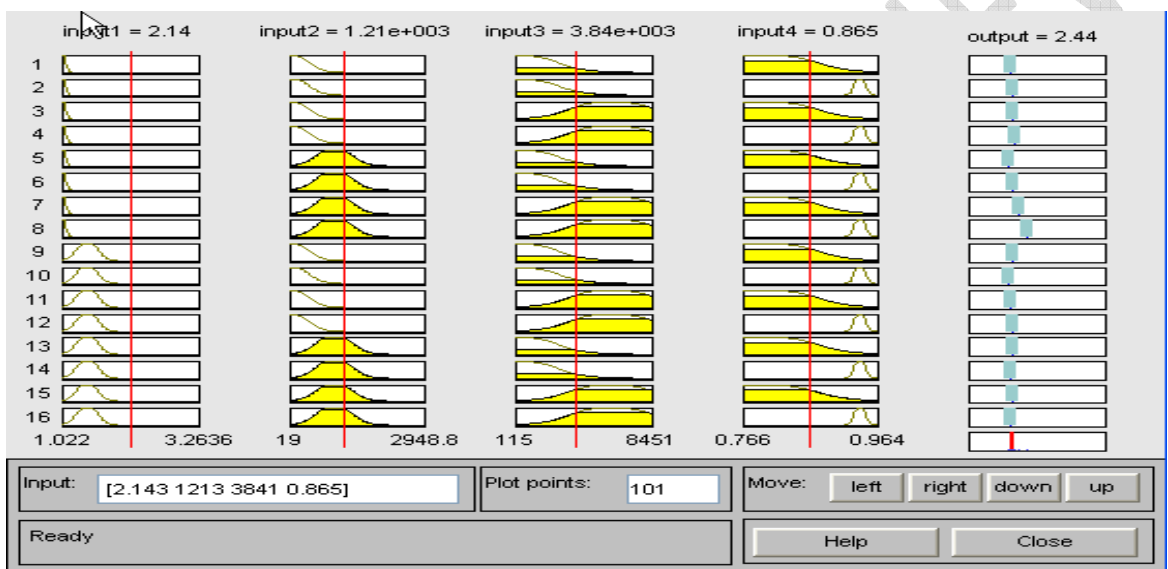


Figure 1: Rule view of fuzzy inference system.

A five- layer network in the frame of Sugeno type model (Ojala, 1994) can be used to implement ANFI network. The architecture of the fuzzy model is shown in Figure 2.

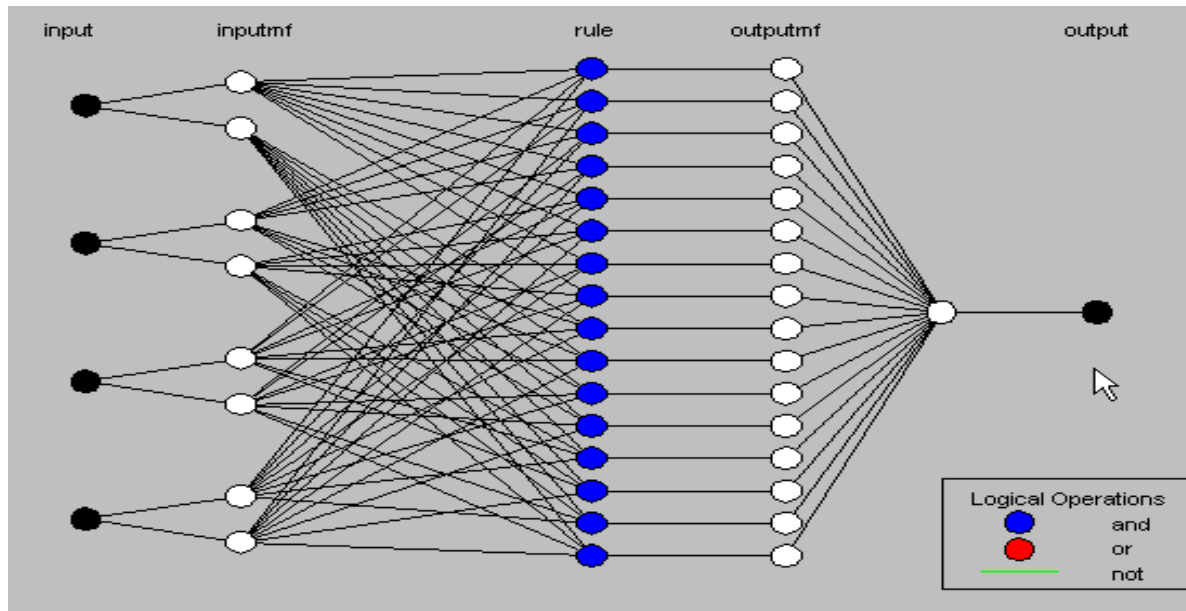


Figure 2: Schematic of fuzzy logic used for this work.

3.2. Data Partition

ANFIS used in this work is data-driven. This is because, the fuzzy rules were generated directly from the input-output data sets. In this case, the neuro-fuzzy model learnt the salient features in the data and automatically adjusted the system parameters so as to reduce the error between the calculated output and the measured output. In MATLAB scatter partition implemented as GENFIS2 was used to generate the inference system. The choice was based on its ability to model uncertainty of real measurements very well and also is well suited to ANFIS which can extract rules only where the premises are connected through differentiable t-norms to generates as many rules as the product of the number of linguistic terms per variable (Abraham, 2001). The model was run on a Duo core (1.6GHz, 2GB RAM) PC. The dataset was partitioned into 750 randomly selected data points for training and 566 randomly selected for checking the model.

4. Results and Discussion

After development of ANFIS model trained with 750 random data point from a total of 1316 data points. The other 566 data that were not used during training of the ANFIS model were used to test the stability and accuracy of the model. The accuracy of the model will be tested using the remaining 566 data points. There will be comparison between the existing empirical models and the developed model residual error, r -square values, AAPE and RMSE.

4.1. Residual Errors

Figure 3 and 4 show the trend of reduction in error generated by the model for training and testing of the ANFIS for first epoch numbers. The error generated by training data reduce from a RMS error of 0.0843 to 0.0447 and that of testing data reduces from a value of 0.7035 to a value of 0.045. The result recorded for training and testing error are in line with the value gotten when the parameters are combined and their maximum RMS error value are combined.

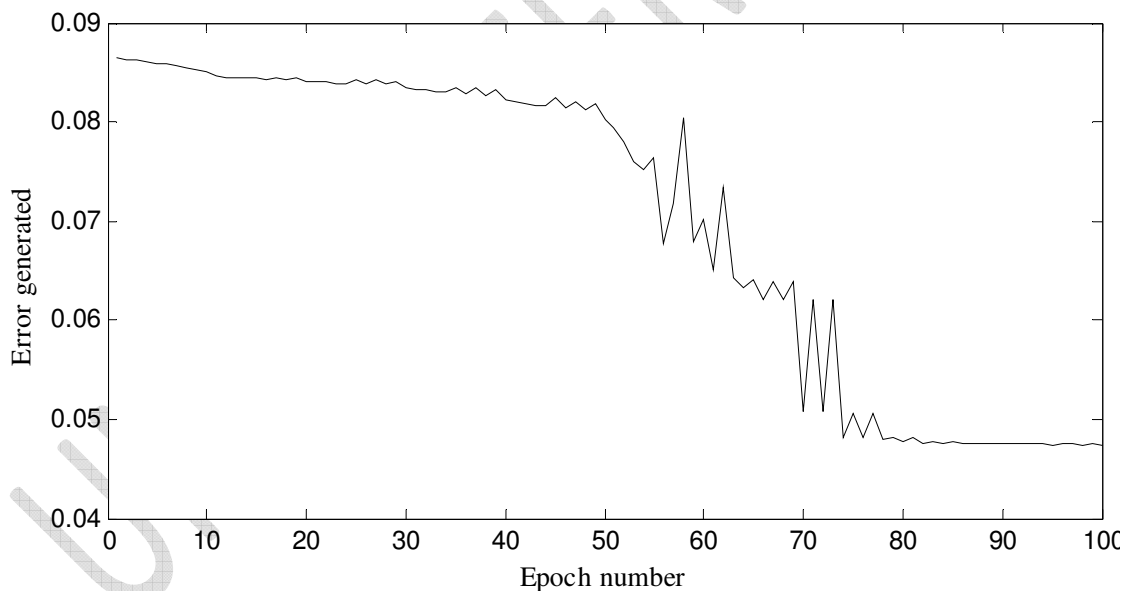


Figure 3: Error trend of the training data for 100 epoch number

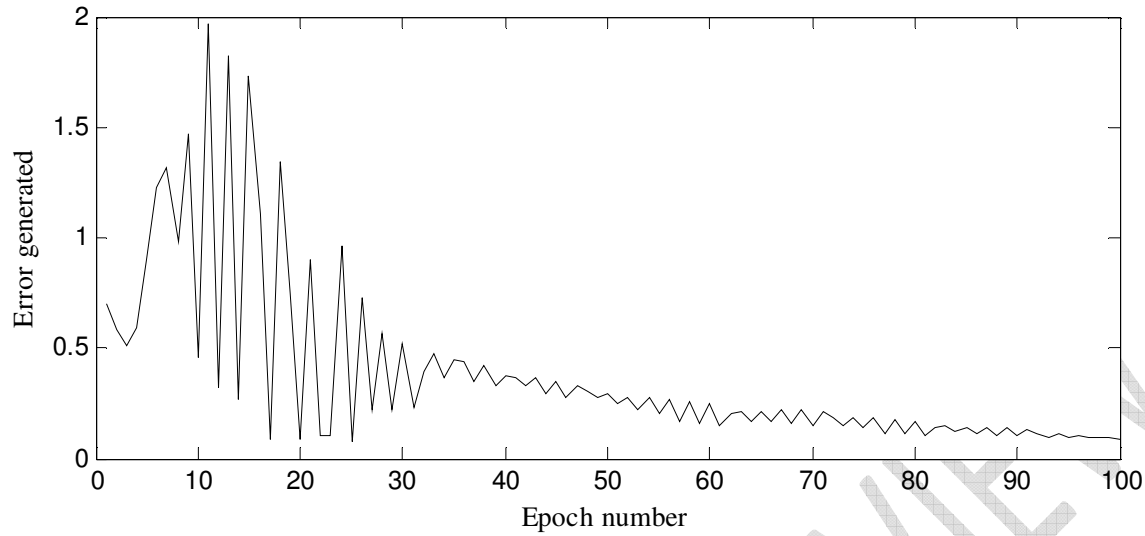


Figure 4: Error trend of the testing data for 100 epoch number

Figures 5 – 10 show the residual error distribution of the the training, testing and whole data and also of the empirical correlations about the zero line to check the error trend of the models. From the residual error of the training model (Figure 9), the values reduces from -0.7302 to 0.2218. It was also observed that highest residual was more pronounced between data 200 and 230.

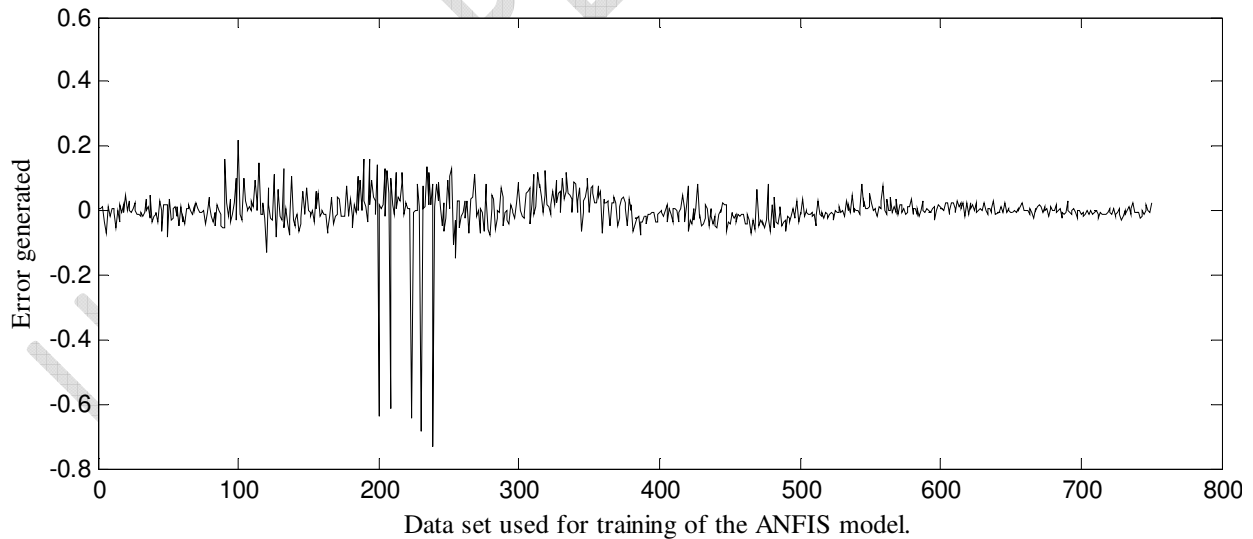


Figure 5: Error Analysis for training data set.

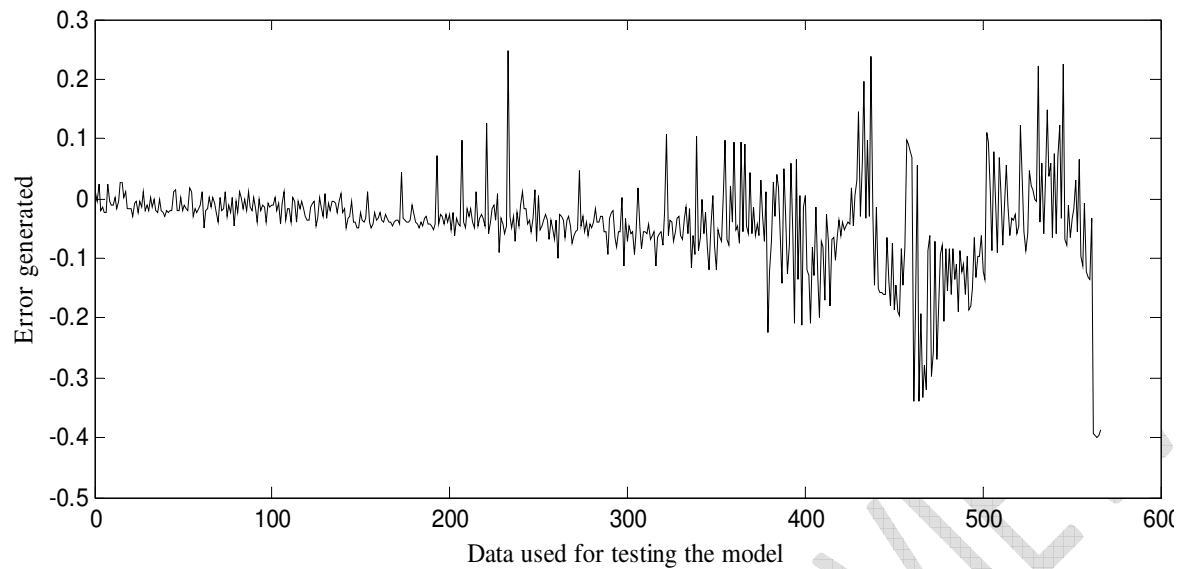


Figure 6: Error Analysis of data used in testing the model

This could be as a result of imprecise measurement of the parameter and are cathered for during fuzzifurised stage. The residual of testing model are between -0.40 and 0.2459. For the whole data used for the model development, the residual error are from -0.7302 to 0.2459.

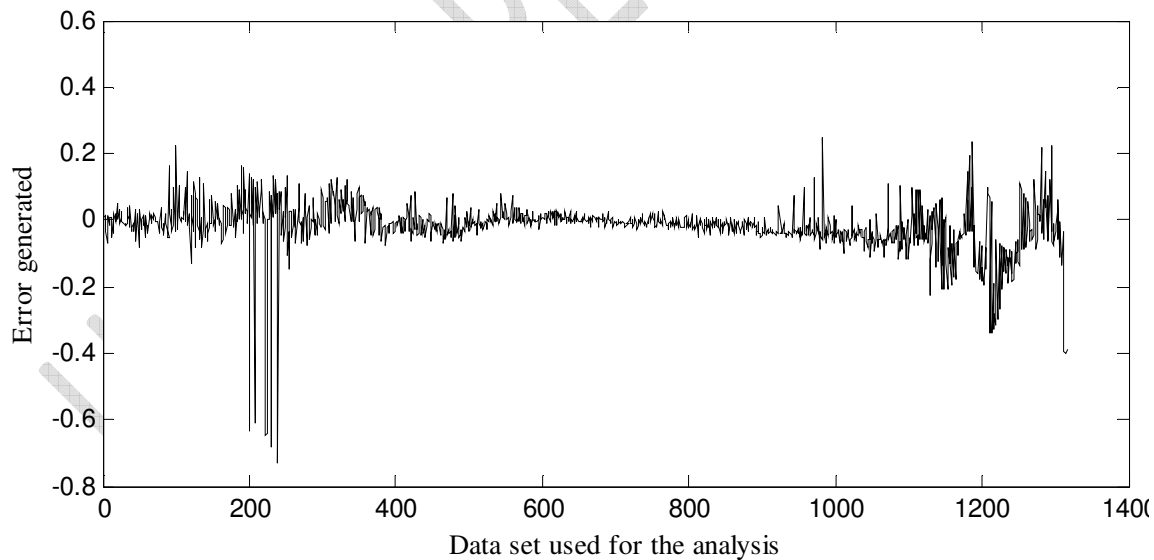


Figure 7: Error Analysis of the ANFIS model using all the data

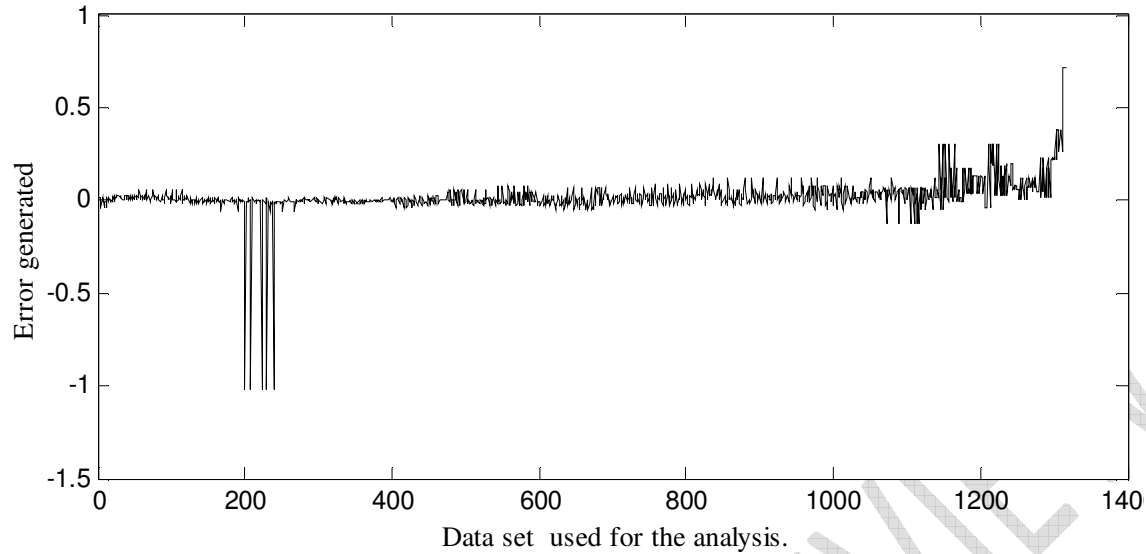


Figure 8: Error analysis for Al marhoun (1988) model

The residual analysis were also conducted on the predictions of three empirical correlations used for determination FVF which are: Al marhoun (1988), Almaarhoun (1992) and Vasquez and Beggs (1980). The residual of Almarhoun are from -1.0154 to 0.7180. The effect of imprecise measurement of data is pronounced in the outcome of the residual error because the highest residual without 200 and 230 data points are -0.0156 which was further reduced to -1.0154 when the imprecise data points are considered.

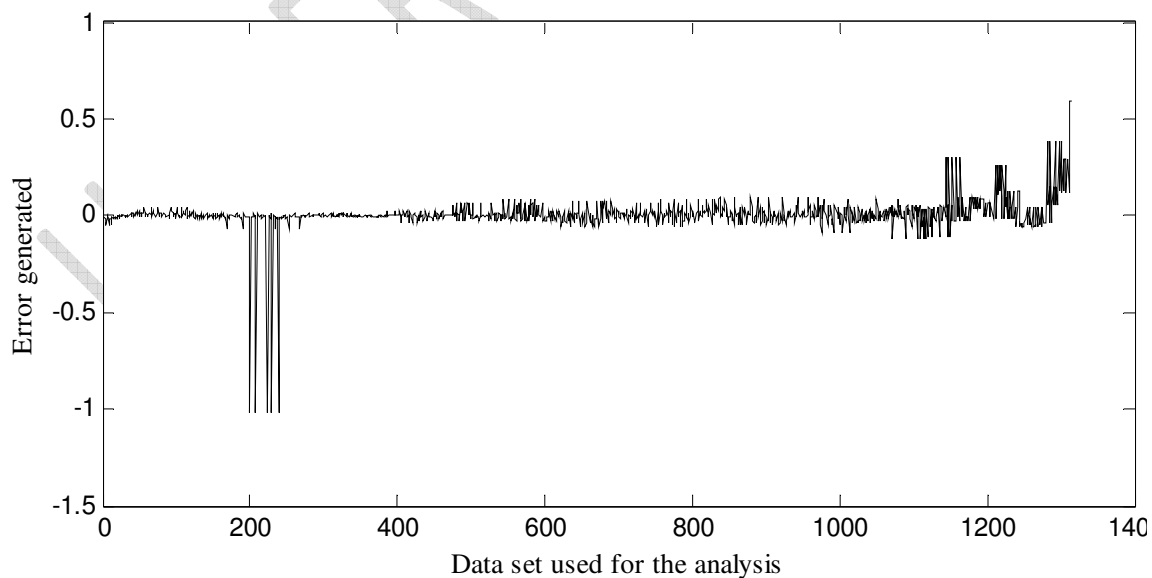


Figure 9: Error analysis of Al marhoun(1992) model

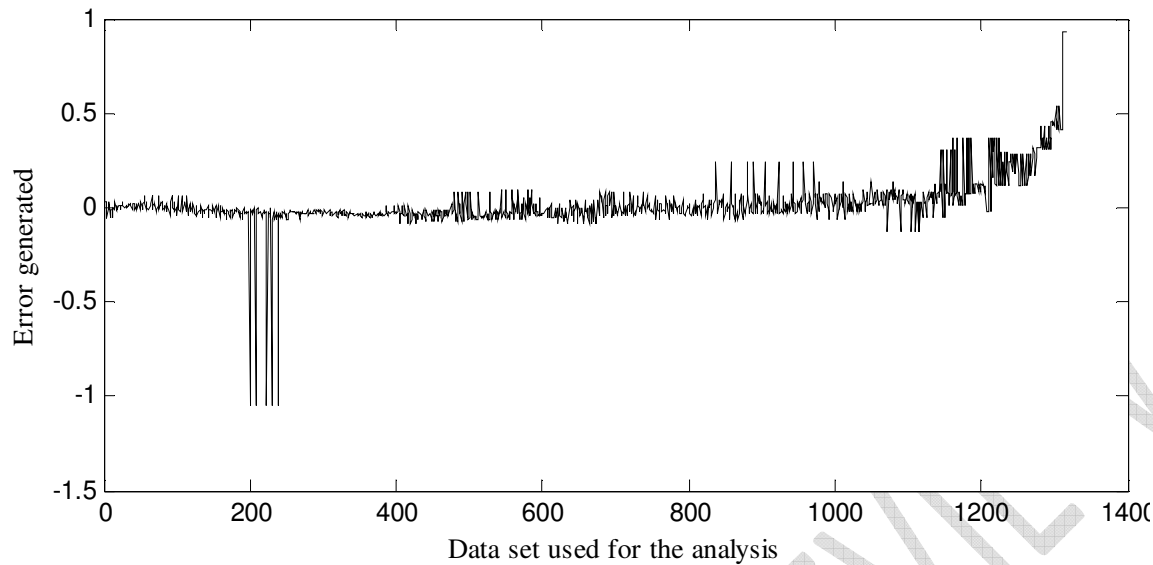


Figure 10: Error Analysis of Van and Beggs (1980) model

The above imprecision in data is pronounced in the model of Almarhoun (1992) and Vasquez and Beggs (1980). The error of Almarhoun (1992) are from -1.0154 to 0.5930 and that of Vasquez and Beggs (1980) are from -1.0489 to 0.9355. From the outcome it was discovered that the ANFIS model at all stage outperform the empirical correlations of Al marhoun (1988) , Almarhoun (1992) and Vasquez and Beggs (1980).

4.2. Statistical Analysis

The developed model was compared with existing empirical correlation using statistical analysis in order to evaluate the performance, suitability, and accuracy of the model. The analysis made use of the following statistical tools: average absolute percentage relative error, root mean square error (RMSE), correlation coefficient and minimum and maximum absolute percentage error.

Table 3: Comparison of Proposed model and existing empirical correlation

Model	Emin (%)	Emax (%)	AAPE	RMSE	R
Training	4.90E-06	0.7302	6.2176	0.0714	0.964
testing	3.04E-05	0.4001	3.0107	0.0868	0.983
whole	4.90E-06	0.7302	4.8384	0.0784	0.985
Almarhoun(1992)	2.01E-06	1.0154	5.5576	0.103	0.949
Almarhoun(1998)	1.29E-05	1.0169	5.2989	0.0951	0.950
Vasquez and Beggs (1980)	1.66E-05	1.0489	7.0672	0.1367	0.927

The values of testing and whole data sets show in table 3 a high accuracy in predicting OFVF with their correlation coefficient values of 0.983 and 0.985 while the values from the three empirical correlations of Al marhoun (1988) , Almarhoun (1992) and Vasquez and Beggs (1980) were 0.975, 0.974 and 0.963, respectively. A correlation coefficient value closer to 1 indicates that there is a good agreement between the experimental and predicted OFVF values. Also show in table 3 are AAPE, RMSE, Emin, and Emax of the model developed as well as the three empirical correlations. The lower value of AAPE signifies a better accuracy of the correlation.

5. Conclusion

A model for the prediction of below point oil formation volume factor (OFVF) is developed through Fuzzy network and the following are deduced from the developed model that:

1. The trend of the error tend to reduce from 0.8 to 0.0023 for 100 epoch number which uphold better prediction of the model.
- 2.The model recorded the lowest value of Minimum percentage error(Emin%), Maximum percentage error(Emax%), AAPE and RMSE values when compared with three widely used correlation with values 4.90E-06, 0.7302, 4.8384 and 0.0784.
- 3.The prediction from the model gives a better prediction and accuracy than the published correlations with correlation coefficient of 0.985.
4. Error analysis conducted on all the model show that the developed model consistently predicted below bubble point viscosity with the lowest error.

References

- A. Abraham. 2001. Neuro-Fuzzy Systems: State-of-the-Art Modeling Techniques. In Connectionist Models of Neurons, Learning Processes, and Artificial Intelligence, Springer, Berlin, J. Mira, A. Prieto (Eds.)
- J. Aguilar, and M. Cerrada. (2001). Fuzzy classifier system and generic programming on system identification. Proceedings of the Generic and Evolutionary Computation Conference, San Francisco, July 7–11, 1245–1251.
- M.A. Al-Marhoun. 1988. PVT Correlations for Middle East Crude Oils, JPT, 650.
- M. A. Al-Marhoun. 1992. New Correlation for formation Volume Factor of oil and gas Mixtures, JCPT 22.
- M.A. Al-Marhoun and E.A. Osman. 2002. Using Artificial Neural Network to develop new PVT correlation for Saudi Crude Oil, Paper No. 78592, 10th Abu Dhabi International Petroleum Exhibition and conference, Abu Dhabi, United Arab Emirate, October 13-16, 1-10.
- A.A. Al-Shammasi. 1999. Bubble Point Pressure and Oil Formation Volume Factor Correlations. Paper No. 53185, SPE Middle East Oil Show, Bahrain, Kingdom of Bahrain, February 20-23, 1-17.
- F. Cubillos, H. Callejas, E.L. Lima, and M.P. Vega. (2001). Adaptive Control Using a Hybrid Neural Model: Application to a Polymerisation Reactor. Brazilian Journal of Chemical Engineering, 18, 1.
- R. B. Garbi and A.M. Elsharkawy. (1997). Universal Neural Network Model for estimating the PVT properties for Crude Oil System, Paper No. 38099, 1997 SPE Asia Pacific Oil & Gas Conference, Kuala Lumpur, Malaysia. April 14-16, 1-10.
- H.M Goda, E. M. El-M Sohir, K. A. Fattah, M. H. Sayyouch. 2003. Prediction of PVT Data using Neural Nstwork Computing Theory. Paper No. 85650, 27th Annual SPE International Technical Conference and Exhibition, Abuja, Nigeria, August 4-6, 1-14.

265 M. H Goda, H.R. Meier, and P. Behrenbruch. 2007. Use of Artificial Intelligence Technique for
 266 Predicting Irreducible Water Saturation. Paper No. 109886, Asia Pacific Oil and Gas
 267 Conference, Jakarta, Indonesia, October 30- November 1, 1-10.
 268 D. Graupe. 2007. Principles of Artificial Neural Networks Advanced Series on Circuits and
 269 Systems, 2nd Ed., Vol. 6, World Scientific Publishing, Singapore.
 270 P. Hegeman, D. Cheng, and N. Varotsis. 2007. Application of Artificial Neural Network to
 271 Downhole Fluid Analysis. IPTC Paper No. 78592, Abu Dhabi International Petroleum
 272 Technology Conference (IPTC), Abu Dhabi, United Arab Emirate, December 4-6, 1-6.
 273 M.N. Hemmati and R. Kharrat. 2007. A Correlation Approach for Prediction of Crude-Oil
 274 PVT Properties. Paper No. 104543, 2007 15th SPE Middle East Oil & Gas Show and
 275 Conference, Bahrain, Kingdom of Bahrain, March 11-14, 1-15.
 276 S. S. Ikiensikimama, J. Madu and L. Dipeolu. 2006. Black Oil Empirical PVT Correlations
 277 Screening for the Niger Delta Crude Oil. Paper No. 105984. 30th SPE Annual SPE International
 278 Technical Conference and Exhibition, Abuja, Nigeria, July 30-August 2, 1-14.
 279 Jang, J.S.R. 1993. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665-685.
 280 K. A. Meisam, J.F. Farhang, and A. Yaser. (2007). Dewpoint Pressure Estimation of Gas
 281 Condensate Reservoir, Using Artificial Neural Network. Paper No. 107032, 2007 SPE
 282 Europe/EAGE Annual Conference and Exhibition, London, United Kingdom, June 11-14, 1-10.
 283 E. O. Obanijesu. and D. O. Araromi. (2008). Predicting Bubble-point Pressure and Formation-
 284 volume Factor of Nigerian Crude Oil System for Environmental Sustainability, Journal
 285 Petroleum Science and Technology, 26 (7), 1993-2008
 286 T. Ojala. 1994. Neuro fuzzy systems in control, Tampere University of Technology, Finland.
 287 Ph.D. thesis.
 288 E.A. Osman, O.A. Abdel-Wahhab, and M.A. Al-Marhoun. 2001. Prediction of Oil PVT
 289 Properties Using Neural Networks, Paper No. 68233 SPE Middle East Oil Show, Bahrain,
 290 Kingdom of Bahrain, March 17-20, 1-15.
 291
 292

293 B. Roffel and B. Betlem. 2006. Process Dynamics and Control: Modelling for Control and
294 Prediction, John Wiley and Sons Ltd.

295 K. K. Salam, D. Araromi & S. S. Ikiensikimama (2011): Neuro-Fuzzy Modeling for the
296 Prediction of Below-Bubble-Point Viscosity, Petroleum Science and Technology, 29:17, 1741-
297 1752.

298 S. Singh. 2005. Permeability Prediction Using Artificial Neural Network: A Case Study of Uinta
299 Basin. Paper No. 99286, SPE International Student Contest at the SPE Annual Technical
300 Conference and Exhibition, Dallas, Texas, October 9-12, 1-8.

301 M. Vazquez and H.D. Beggs. 1980. Correlation for Fluid Physical Property Prediction, JPT
302 968.
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318