American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-10, Issue-12, pp-80-88
www.ajer.org
Research Paper Open Access

Modeling Approach for Niger Delta Oil Formation Volume Factor Prediction using Support Vector Machine

Shaibu Samuel[1] and Ijeoma I. Mbachu[2]
[1] University of Port-Harcourt, Department Petroleum Engineering, Nigeria.
[2] University of Port-Harcourt, Department of Petroleum Engineering, Nigeria.

ABSTRACT : *Oil Formation Volume Factor(OFVF) is a very important fluid property in reservoir engineering computations. It is the volume of oil (and dissolved gas) at reservoir pressure and temperature required to produce one stock tank barrel of oil at the surface. It can be obtained either by conducting laboratory study on reservoir fluid samples or estimated, using empirically derived PVT correlations. Although laboratory results gives better prediction where controlled conditions are imposed but in situation where the experimental data are not available, artificial intelligence and published empirical correlations are used. Unfortunately, the development of published empirical correlations have many drawbacks and limitations as they were originally developed for certain ranges of reservoir fluid characteristics. This research work proposes the use of Support Vector Machine (SVM) as a new intelligence framework to predict oil formation volume factor as to address the limitations of empirical correlations and Neural Network models. 1402 data set was obtained from PVT report from Niger-Delta, out of which, 70% were used to train the model, 20% for testing and 10% for validation. A comparative study was carried out to compare support vector machine regression performance with the neural networks and other published empirical correlation techniques. The result revealed that the support Vector Machine Model performed better than the popularly used Feed Forward Back Propagation ANN and the empirical correlations in terms of quantitative and qualitative analysis employed. The new model performed best for the Niger-Delta Crude with highest correlation coefficient of 0.9812, Mean Absolute Error (Ea) of 1.2895, the best rank of 0.606267 with a better performance plot.*

Keywords - *Artificial Neural Network, Oil formation volume factor, statistical analysis, Support Vector Machine, Niger- Delta.*

Date of Submission: 08-12-2021 Date of acceptance: 23-12-2021

I. INTRODUCTION

An accurate knowledge of Pressure-Volume-Temperature (PVT) properties is essential in reservoir and production engineering calculations. Estimation of reserves, determination of oil reservoir performance, recovery efficiency, production optimization and design of production systems are some of the areas which require precise determination of fluid's physical properties at different conditions of pressure and temperature. Ideally, the physical properties of the reservoir fluids are determined experimentally in the laboratory. However, due to economical and/or technical reasons, quite often this information cannot be obtained from laboratory measured values. In this case, PVT properties must be estimated from empirically derived correlations. The correlations were developed using linear and non-linear regression or graphical techniques. The correlations are accurate within the range of data that were used to develop them [1]. Among those PVT properties is the Oil Formation Volume factor (OFVF), which is defined as the volume of reservoir oil that would be occupied by one stock tank barrel oil plus any dissolved gas at the bubble point pressure and reservoir temperature. Precise prediction of oil FVF is very important in reservoir and production computations.

Several correlations have been proposed for determining crude oil formation volume factor such as [2], [3], [4], [5], [6] and [7]. In order to find relationship between the input and output data driven from experiment, a powerful method than traditional modeling is necessary, hence, computational intelligence techniques, such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Fuzzy logic and Neural-Fuzzy among others, have been applied.

Investigators has recognized that the neural network can serve the petroleum industry to create a more accurate PVT model ([8], [9], [10], [11], [12], [13]).

[9] published neural network models for estimating bubble point pressure and oil formation volume factor for Middle East crude oils. They used two hidden layers neural networks to model each property separately. The bubble point pressure model had eight neurons in the first layer and four neurons in the second. The formation volume factor model had six neurons in both layers. Both models were trained using 498 data sets collected from the literature and unpublished sources. The models were tested by other 22 data points from the Middle East. The results showed improvement over the conventional correlation methods with reduction in the average error for the bubble point pressure oil formation volume factor.

[10] presented an Artificial Neural Network (ANN) for estimation of PVT properties of compounds. The data set was collected from Perry's Chemical Engineers' Handbook. Different training schemes for the back propagation learning algorithm; Scaled Conjugate Gradient (SCG), Levenberg-Marquardt (LM) and Resilient back Propagation (RP) methods were used. The accuracy and trend stability of the trained networks were tested. The LM algorithm with sixty neurons in the hidden layer proved to be the best suitable algorithm with the minimum Mean Square Error (MSE) of 0.000606. ANN is one of the best estimating method with high performance used in forecasting the PVT properties.

[12] developed a new artificial neural network model to predict oil formation volume factor using 802 data sets collected from the Niger Delta Region of Nigeria. One-half of the data was used to train the ANN models, one quarter to cross-validate the relationships established during the training process and the remaining one quarter to test the models to evaluate their accuracy and trend stability during the training process. Both quantitative and qualitative assessments were employed to evaluate the accuracy of the new model to the empirical correlation. The authors reported that the new correlation outperformed the best existing correlation by the statistical parameters used with a rank of 0.85 and better performance plot.

[13] researched on building an artificial neural network (ANN) model to predict oil formation volume factor for the different API gravity ranges. The new models were developed using combination of 448 published data from the Middle East, Malaysia, Africa, North Sea, Mediterranean basin, Gulf of Persian fields and 1389 data set collected from the Niger Delta Region of Nigeria. The data was divided into the following four different API gravity classes: heavy oils for API $\leq$ 21, medium oils for 21>API$\leq$26, blend oils for 26>API$\leq$35 and light oils for API> 35. The data set was randomly divided into three parts of which, 60% was used for training, 20% for validation, and 20% for testing for each particular API grade. The ANN models outperformed the existing empirical correlations by the statistical parameters used with the best rank and better performance plots.

In this research work, a new predictive tool which is more powerful than ANN was developed to estimate the oil formation volume factor. The powerful tool is Support Vector Machine (SVM).

II. SUPPORT VECTOR MACHINE (SVM)

Support Vector Machine is a set of related supervised learning methods used for classification and regression. They belong to a family of generalize linear classifiers. They can also be considered as a special case of Tikhonov Regularization. SVMs map input to a higher dimension space where a maximum separating hyperplane is constructed. The generalization of SVM is ensured by the special properties of the optimal hyperplane that maximizes the distance of training examples in a high dimension feature space. Recently, a new E-sensitive loss function technique that is based on statistic learning theory, and which adhere to the principle of risk maximization, seeking to maximize an upper bound of the generalization error was developed. This gave rise to the technique called support linear regression (SVR). It has been shown to exhibit excellent performance ([14], [15]).

Support vector machines regression is one of the most successful and effective algorithms in both machine learning and data mining communities. It has been widely used as a robust tool for classification and regression It has been found to be very robust in many applications, for example in the field of optical character recognition text categorization, and face detection in images. The high generalization ability of SVM regression is ensured by special properties of the optimal hyperplane that maximizes the distance to training examples in a high dimensional feature space ([15], [16]).

Due to the above merits, SVM has been successfully applied in many areas such as decision support, software reliability identification, pattern recognition, and in the prediction of oil and gas properties. [18] and [19] showed the special features of SVM in terms of its ability to handle small dataset and ease of training. In

terms of both execution time and correlation coefficient, they presented SVM to be a very close match to Functional Networks with very high mutually competitive edges. [15] presented SVM as a better model to improve the prediction of PVT properties of crude oil systems. The result showed the highest accuracy with the lowest error measures and highest correlation coefficient when compared with ANN and other published empirical correlations. In a similar study, [19] proposed ANN, SVM and Functional Networks to predict the Pressure Volume Temperature (PVT) properties of crude oil. The result showed that SVM regression and Functional Networks are competitive but SVM has the overall best result for gas and oil prediction. [20] proposed the application of SVM for prediction of toxic activity with different datasets When compared with ANN, SVM gave the highest correlation coefficient.[15] proposed ANN, SVM and Functional networks to predict the Pressure-Volume-Temperature (PVT) properties of crude oil. The result showed that SVR and FN are competitive but SVM has the overall best result for both gas and oil prediction. Other successful applications of SVM include [21].

The performance of SVR is based on the predefined parameters (so-called hyper-parameters). Therefore, to construct a perfect SVR forecasting model, SVR's parameters must be set carefully. Recently, SVR has emerged as an alternative and powerful technique to predict a complex nonlinear relationship problem. It has achieved great success in both academic and industrial platforms due to its many attractive features and promising generalization performance [15].

2.1 The Structure of Support Vector Machine

Recently, a regression version of SVM has emerged as an alternative and powerful technique to solve regression problems by introducing an alternative loss function. A brief description of SVM is given in this subsection and the detailed descriptions can be found in [23]. Generally, the SVR formulation follows the principle of structural risk minimization, seeking to minimize an upper bound of the generalization error rather than minimize the prediction error on the training set. This feature gives SVR a greater potential to generalize the input–output relationship. The SVR maps the input data x into a high-dimensional feature space F by nonlinear mapping, to yield and solve a linear regression problem in this feature space as it is shown in Fig.1. The regression approximation estimates a function according to a given data as shown in Equ. 1,

$$G = \{(x_i, y_i) : x_i \Re^p\}_i^n = 1 \quad (1)$$

Where; x_i denotes the input vector; y_i denotes the output (target) value and n denotes the total number of data patterns. The modeling aim is to build a decision function, where $\hat{y} = f(x)$ that accurately predicts the outputs $\{y_i\}$ corresponding to a new set of input–output examples, $\{(x_i, y_i)\}$.Using mathematical notation, the linear approximation function is approximated using the following function:

$$f(X) = (\omega^T \varphi(x) + b), \varphi: \Re^p \rightarrow F; \text{ and } \omega \in F, \quad (2)$$

Where, ω and b are coefficients; $\varphi(x)$ denotes the high-dimensional feature space, which is nonlinearly mapped from the input space b . Therefore, the linear relationship in the high-dimensional feature space responds to nonlinear relationship in the low-dimension input space, disregarding the inner product computation between ω and $\varphi(x)$ in the high-dimensional feature space.

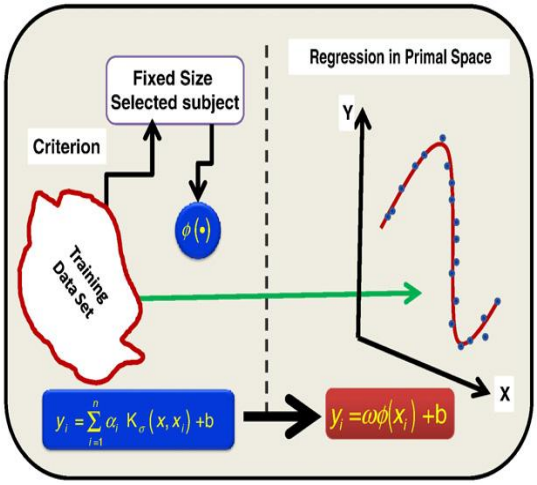

Fig. 2 Soft margin loss setting for a linear SVR [24].

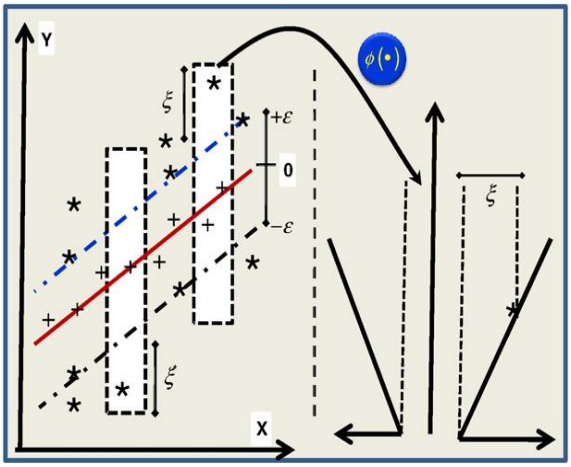

Fig.1. Mapping input space x into high-dimensional feature space [23].

Correspondingly, the original optimization problem involving nonlinear regression is transformed into finding the flattest function in the feature space F , and not in the input space, x . The unknown parameters ω and b in Equation (1) are estimated by the training set, G . SVR performs linear regression in the high-dimensional feature space by ε-insensitive loss. At the same time, to prevent over-fitting and, thereby, improving the generalization capability, following regularized functional involving summation of the empirical risk and a complexity term $\frac{\|\omega\|^2}{2}$, is minimized. The coefficients ω and b can thus be estimated by minimizing the regularized risk function.

$$R_{SVR}(C) = R_{emp} + \left(\frac{1}{2}\right) \| \omega^2 \| = \frac{c}{n} \sum_{i=1}^{n} l\varepsilon\ (yi, \hat{}yi) + (½) \| \omega^2 \|; \quad (3)$$

Where: R_{SVR} and R_{emp} represent the regression and empirical risks, respectively, $\frac{\|\omega\|^2}{2}$, denotes the Euclidean norm and C denotes a cost function measuring the empirical risk. In the regularized risk function given by equation (3), the regression risk (test set error), R_{SVR} , is the possible error committed by the function f in predicting the output corresponding to a test example input vector.

$$L_{\varepsilon}\left(y, \hat{y}\right) = \begin{cases} \left|y - \hat{y}\right| = \varepsilon, & if \left|y - \hat{y}\right| \geq \varepsilon \\ 0 \end{cases} \quad (4)$$

Other wise, the second item $\frac{\|\omega\|^2}{2}$,is the regularization term. The regularized constant C calculates the penalty when an error occurs, by determining the trade-off between the empirical risk and the regularization term, which represents the ability of prediction for regression. Raising the value of C increases the significance of the empirical risk relative to the regularization term. The penalty is acceptable only if the fitting error is larger than ε . The ε -insensitive loss function is employed to stabilize estimation. In other words, the ε -insensitive loss function can reduce the noise. Thus, ε can be viewed as a tube size equivalent to the approximation accuracy in training data as it is shown above in the empirical analysis, C and ε are the parameters selected by the users.

To estimate ω and b , we introduce the positive slack variables ξ_i and ξ_i^* , then according to Fig. 2 the sizes of the stated excess positive and negative deviations are represented by ξ_i and ξ_i^* respectively. The slack variables assume non-zero values outside the $[-\varepsilon, \varepsilon]$ region. The SVR fits $f(x)$ to the data such that;

III. METHODOLOGY

3.1 Data Description

The data used was obtained from conventional PVT reports that derive the various fluid properties through liberation process from the Niger-Delta Region of Nigeria. The data parameters include; Oil relative density (γ_o), Reservoir temperature (^oF), Solution Gas Oil Ratio (scf/stb), Gas relative density (γ_g) and Oil Formation Volume Factor (OFVF). The maximum, minimum and mean values of data used for validation, training and test data are shown in Tables 1 and 2.

TABLE 1. Summary of maximum and minimum values of validation data for oil formation volume factor Support Vector Machine

PVT Properties	Maximum Values	Minimum Values	Mean Values
γ_g	0.8210	0.5630	0.6490
γ_o	0.9670	0.7950	0.8720
T (°F)	260.00	127.00	165.80
Rs (scf/stb)	1448.0	7.0000	357.98

TABLE 2. Summary of maximum and minimum values of training and test data for oil formation volume factor Support Vector Machine

PVT Properties	Maximum Values	Minimum Values	Mean Values
γ_g	1.3710	0.5310	0.6790
γ_o	0.9670	0.7560	0.8570
T(°F)	264.00	104.40	173.47
Rs (scf/stb)	1616.0	7.0000	447.64

3.2 Data Validation

Before any experimental PVT data are used for design or study purposes, it is necessary to ensure that there are no error or major inconsistencies that would render any subsequent work useless. Two such means of data validation are the Campbell diagram (Buckley plot) and the Mass Balance Diagram which are otherwise known as cross plot. These techniques were used to validate the data set used in this work.

3.3 Modeling Technique

Support vector machine regression was used to build the Oil Formation Volume Factor model using kernel functions with the Radial Basis function (RBF) procedure using MATLAB (2019b) version. The steps involved are as follows;

Step 1: Import the data; the input data was imported to the MATLAB environment using the import command. The following variables; Solution GOR (scf/stb), Reservoir temperature (^{o}F), Gas relative density, Oil relative density, were imported into the MATLAB environment.

Step 2: Select the variables; this is to arrange a set "P input" vector and "T output" vectors as columns into first and second matrix in the MATLAB workspace as follows;

(P) Input data = [GOR ; T ; γ_g ; γ_o] (7)

(T) Target data = [OFVF] (8)

Step 3: Division of Data Point: The size of the data points was 1402 and were divided into three parts which are training, testing and validation. The model was trained with 70% of the data points, 20% was used for testing the trained model and 10% was used for validating the model. Step 4: Choose the kernel function: The kernel function used in this study is Radial Basis Function (RBF). It is the most popular choice of kernel types used in SVM modeling because of its high level of accuracy. This is because of its localized and finite responses across the entire range of the real independent variables (equations 2 and 9).

$$K(x_i, x_j) = e^{-\gamma|x_i - x_j|^2} \quad (9)$$

Step 5: Estimating Model Parameters: The function contains parameters which must be estimated from modelling and simulating the SVM. These constants are: Capacity (C) and epsilon (ε). The SVM was modelled such that a search for the model parameters were initiated between the interval of 1 to 100 for capacity (C) and 0.1 to 1.5 for epsilon(ε).

Step 6: Method of simulation: The two methods applicable are supervised and unsupervised learning. Supervised, also known as supervised machine learning is defined by the use of labelled datasets to train algorithms that classify data or predict outcomes accurately. As input data is fed into the model, it adjusts its weight until the model has been fitted appropriately, which occurs as part of the cross-validation process.

Unlike supervised learning, unsupervised learning uses unlabeled data. From the data, it discovers pattern that help solve for cluttering or association problem. This work used supervised learning approach for the SVM modelling. Supervised method which means that the output is already known i.e., oil formation volume factor. Supervised model keeps iterating the provided value to obtain a near criteria.

Step 7: Choosing stopping criteria/Determination of model accuracy: the SVM simulates until a stopping criterion is met, this stopping criterion is chosen based on two factors; Number of iterations (N) and Tolerance (t). Several parameters were used to measure the performance of the model.

Step 8: Simulation of support vector machine model; This allows performing additional tests on the model or putting it to work on new inputs.

3.3 Correlation Comparison

To compare the performance and accuracy of the new model to other empirical correlations, two forms of analyses were performed which are quantitative and qualitative screening. For quantitative screening method, statistical error analysis was used. The statistical parameters used for the assessment were percent mean relative error (MRE), percent mean absolute error (MAE), percent standard deviation relative (SDR), percent standard deviation absolute (SDA) and correlation coefficient (R).

For correlation comparison, a new approach of combining all the statistical parameters mentioned above (MRE, MAE, SDR, SDA and Rank) into a single comparable parameter called Rank was used [17]. The use of multiple combinations of statistical parameters in selecting the best correlation can be modeled as a constraint optimization problem with the function formulated as;

$$Min \quad Rank = \sum_{j=1}^{m} S_{i,j} q_{1,j} \quad (10)$$

Subject to $$\sum_{i=1}^{n} S_{i,j} \quad (11)$$

With $$0 \leq S_{ij} \leq 1 \quad (12)$$

Where $S_{i,j}$ is the strength of the statistical parameter j of correlation i and q_{ij}, the statistical parameter j corresponding to correlation i. j = MRE, MAE, …. R^1, where $R^1 = (1-R)$ and Z_i is the rank, (or weight) of the desired correlation. The optimization model outlined in equations 18 to 20 was adopted in a sensitivity analysis to obtain acceptable parameter strengths. The final acceptable parameter strengths so obtained for the quantitative screening are 0.4 for MAE, 0.2 for R, 0.15 for SDA, 0.15 for SDR, and 0.1 for MRE. Finally, equation 20 was used for the ranking. The correlation with the lowest rank was selected as the best correlation for that fluid property. It is necessary to mention that minimum values were expected to be best for all other statistical parameters adopted in this study except R, where a maximum value of 1 was expected. Since the optimization model (equations 10 to 12) is of the minimizing sense a minimum value corresponding to R must be used. This minimum value was obtained in the form (1-R). This means the correlation that has the highest correlation coefficient (R) would have the minimum value in the form (1-R). In this form the parameter strength was also implemented to 1-R as a multiplier. Ranking of correlations was therefore made after the correlations had been evaluated against the available database.

For qualitative screening, performance plots were used. The performance plot is a graph of the predicted versus measured properties with a 45^o reference line to readily ascertain the correlation's fitness and accuracy. A perfect correlation would plot as a straight line with a slope of 45^o.

IV. RESULTS AND DISCUSSION

The support vector machine models were tested with 20% of training data (280 data) points that were not previously used during training and validation. These data were randomly selected by the Support vector machine tool to test the accuracy and stability of the model. The performance of the support vector machine model was compared with predictions from Artificial Neural Network model and some selected empirical correlations such as Obomanu and Okpobiri (1987) Al-Marhoun (1988), Ikiensikimama (2009), Petrosky and Farshad (1993).

These predictive correlations were carefully selected, having been developed specifically for the prediction of oil FVF and some of which were recommended for the estimation of oil FVF for API gravity range.

The results of the statistical assessment as presented in Figure 3 gives the statistical accuracies for all the oil formation volume factor correlations and ANN model examined. The results show that the support vector machines regression algorithm has both reliable and efficient performance as to compare to other existing correlations and ANN model. Table 3 shows the numerical values of the model accessed with SVM having the best rank of 0.606267 with Mean absolute Error (E_a) of 1.2895 and correlation coefficient (R) of 0.9812.

In addition, it outperforms [9] which is one of the popular indigenous empirical correlation and the feedforward neural networks in terms of root mean squared error, absolute average percent error, standard deviation, correlation coefficient and Rank.

TABLE 3. Statistical Accuracy of Oil Formation Volume Factor Using Niger-Delta Data

Correlations	%MRE	%MAE	%SRE	%SAE	R	RANK
SVM (This Study)	0.28950	1.2895	0.192524	0.1925	0.9812	0.6062
ANN	0.36540	1.1654	0.19856	2.0324	0.9723	0.8428
Ikiensikimama (2009)	0.98654	2.6696	2.887085	1.3215	0.9505	1.6104
Al-Marhoun (1988)	1.49157	3.6627	1.514832	1.3192	0.9651	2.0463
Petrosky and Farshad (1993)	10.1086	10.108	8.580097	8.5800	0.9494	7.6384
Obomanu and Okpobiri (1987)	14.9844	14.984	10.882	10.8820	0.9134	10.774

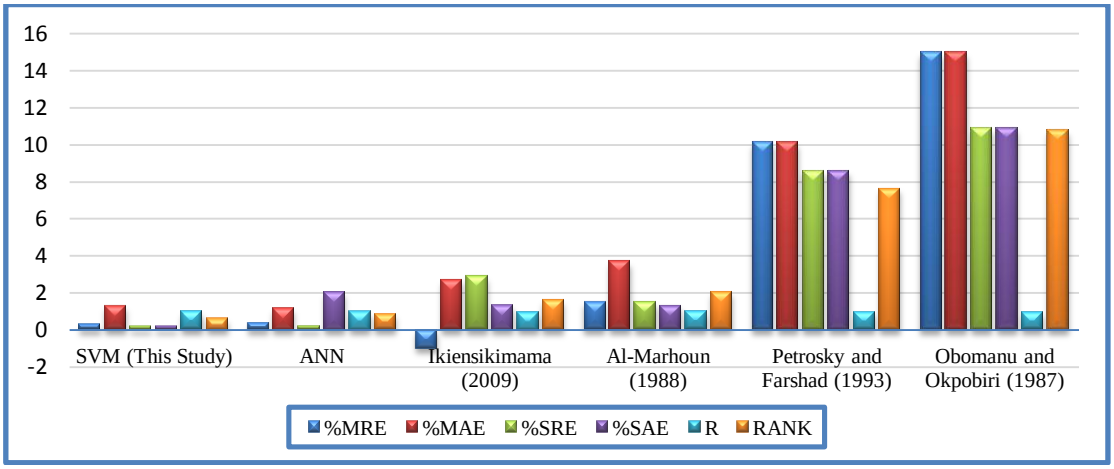

Fig. 3. Comparison of the Statistical Accuracy for Different correlations using Niger -Delta Data

Figs. 4 to 7 illustrate cross plots of the predicted versus experimental oil formation volume factor (OFVF) values. A cross plot is graph of predicted versus measured properties with a 45° reference line to readily ascertain the correlation's fitness and accuracy.

Compare to other cross plots, Fig. 4 shows the tightest cloud of points around the 45° line with very good clusters at low band, indicating the excellent agreement between the experimental and the calculated data values when compared to Figs. 5 to 7. In addition, this indicates the superior performance of the Support vector Machine model over empirical correlations evaluated. The accuracy of the model indicates that the Support Vector Machine intelligent model does not over fit the data, which implies that it was successfully trained.

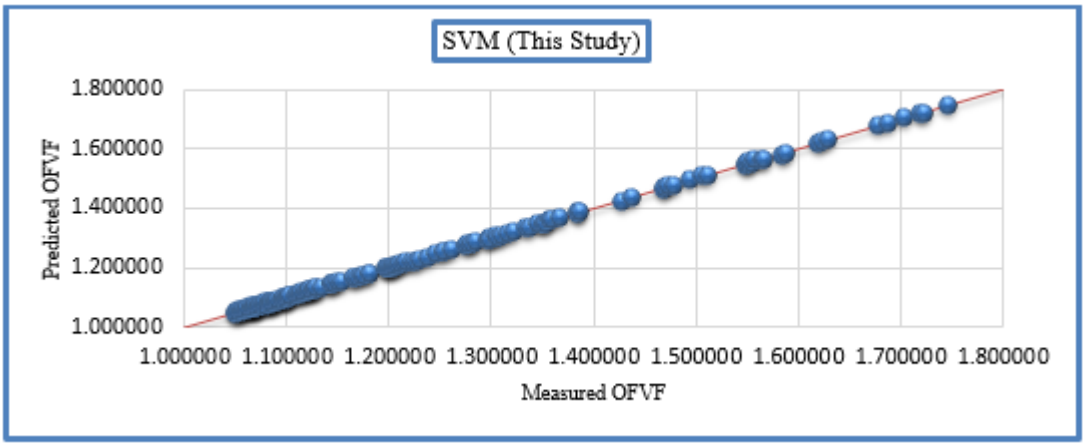

Fig. 4. Cross plot of SVM (This Study) Correlation using Niger-Delta Data

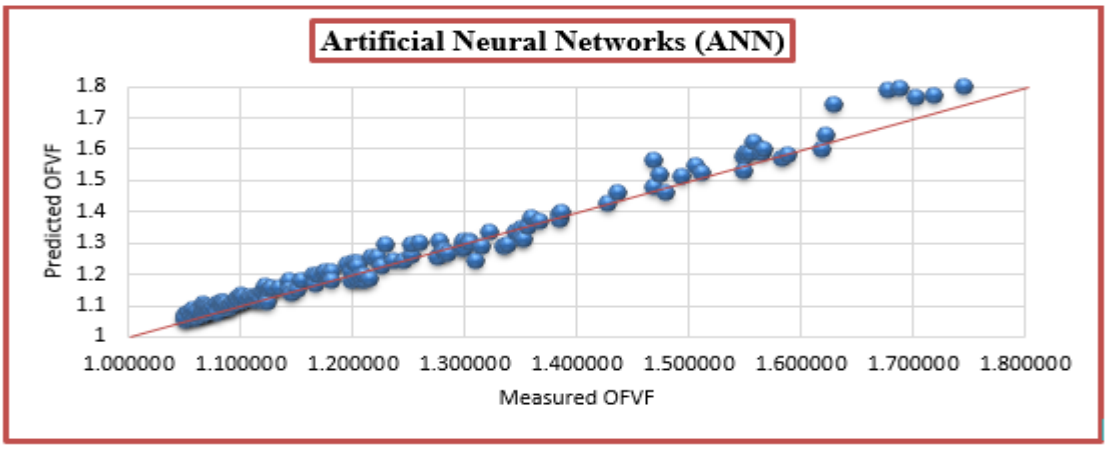

Fig. 5. Cross plot of Artificial Neural Networks using Niger-Delta Data

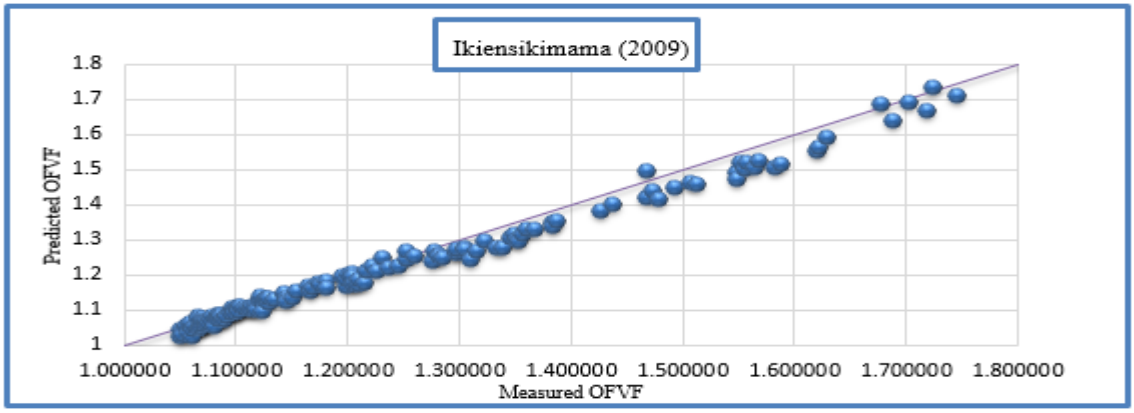

Fig. 6. Cross plot of Ikiensikimama (2009) correlation using Niger-Delta Data

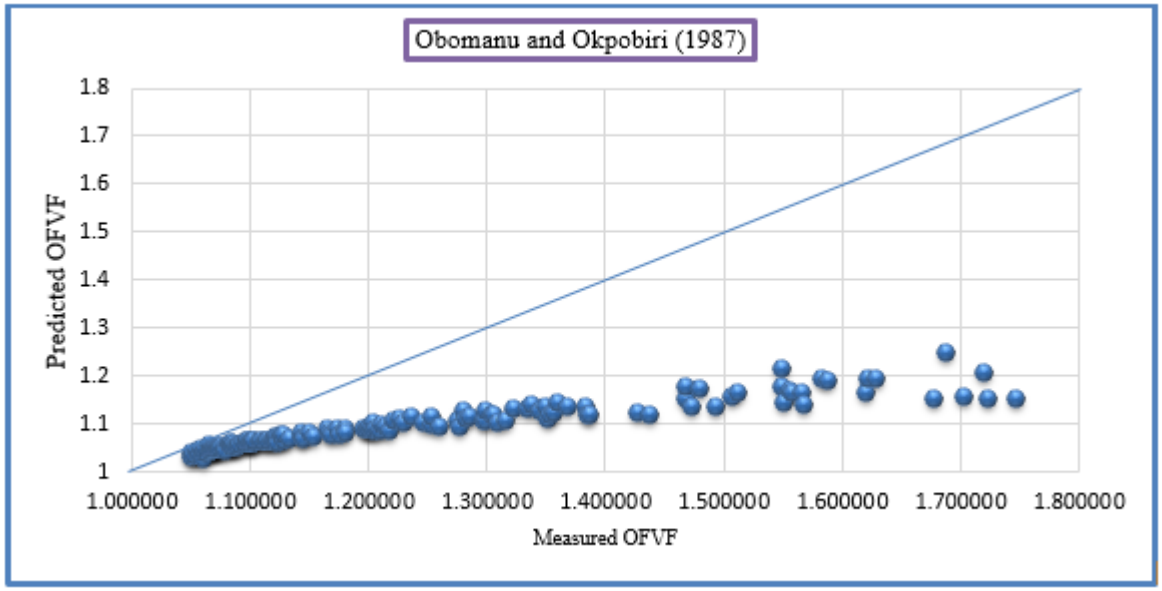

Fig. 7. Cross plot of Obomanu and Okpobiri (1987) Correlation using Niger-Delta Data

V. CONCLUSION

The newly developed support vector machine model for predicting crude oil formation volume factor for Niger-Delta region was developed in this study using MATLAB 2019 Version. The kernel function imbedded in the support vector machine was used to estimate the model parameters. The new tool outperformed the existing correlations and Artificial Neural Network model by the statistical parameters used. It shows a best rank with a numerical value of 0.6063 and better performance plot as compared to the existing empirical correlations for those regions where the data was used. This leads to a bright light of support vector machines modeling and will assist petroleum exploration engineers to estimate various reservoir properties with better accuracy, leading to reduced exploration time and increased productions.

REFERENCES

[1] E. A. Osman, O. A. Abdel-Wahhab, and M. A. Al-Marhoun, Prediction of Oil Properties using Neural Networks, *SPE Paper 68233, Presented at the SPE Middle East Oil Show Conference, Bahrain*, 2001.

[2] M. B. Standing, A Pressure-Volume-Temperature Correlation for Mixtures of California Oils and Gases, *Drill & Prod. Pract., API* ,1947, pp 275-87.

[3] D. A Obomanu, and G. A Okpobiri, Correlation the PVT Properties of the Nigeria Crudes, *Trans. ASME Vol. 109*, 1987, 14-24.

[4] Al-Marhoun, M. A. (1988): "PVT Correlations for Middle East Crude Oils," *J. Pet. Tech*, 1988, pp. 650-666.

[5] J. Petrosky, and F. Farshad, Pressure Volume Temperature Correlation for the Gulf of Mexico, *Paper SPE 26644 presented at the 1993 SPE Annual Technical Conference and Exhibition, Houston, TX*, 1993.

[6] M. I. Omar, and Todd, A. C., Development of Modified Black oil Correlation for Malaysian Crudes, *Paper SPE 25338, presented at the 1993 SPE Asia Pacific Oil and Gas Conference, Singapore*,1993.

[7] S. S. Ikiensikimama, (2009), *Reservoir Fluid Property Correlations*, (Chi Ikoku Petroleum Engineering Series, IPS Publications, Port Harcourt, 2009).

[8] M. A. Al-Marhoun, and E. A. Osman, Using Artificial Neural Networks to Develop New PVT Correlations for Saudi Crude Oils, *Paper SPE 78592, presented at the 10th Abu Dhabi International Petroleum Exhibition and Conference (ADIPEC), Abu Dhabi, UAE,* 2002.

[9] R. B. Gharbi, and E. A El-sebakhy, "Neural-Network Model for Estimating the PVT Properties of Middle East Crude Oils," *SPE 37695 presented at the SPE Middle East Oil Show and Conference, Bahrain,* 1997.

[10] Mohammad and Mohammad, A new heuristic model for estimating the oil formation volume factor. Faculty of Petroleum Engineering, Amirkabir University of Technology (Tehran Polytechnic), Hafez Avenue, Tehran, Iran.

[11] O. Omole, O. A. Falode, and A. D. Deng, Prediction of Nigerian Crude Oil Viscosity Using Artificial Neural Network, *Petroleum and Coal ISSN 151(3), 2009* 181-188.

[12] I. I. Azubuike and S. S. Ikiensikimama, Modelling Approach for Niger Delta Oil Formation Volume Factor Prediction using Artificial Neural Network" *Refereed Proceedings, Paper SPE 162987. 36th Annual International Conference of the SPE Nigerian Council, 6 – 8th August 2012, Expo Centre, Eko Hotel and Suites, Victoria Island, Lagos, Nigeria.* 2012.

[13] I. I. Azubuike and S. S. Ikiensikimama, Forcasting Oil Formation Volume Factor for API Gravity Ranges Using Artificial Neural Network, *Advances in Petroleum Exploration and Development Journal Vol. 5, No. 1,* 2013, pp 1-9.

[14] F. Anifowose, "Prediction of Oil and Gas properties using Support Vector Machine" IPTC 14514, 2011.

[15] E. A. El-sebakhy , Forecasting PVT Properties of Crude Oil System Based on Support Vector Machines Modeling Scheme, *Journal of Petroleum Science and Engineering,* 2009 pp. 25-34

[16] E. Osuna, R. Freund, and F. Giwsi, Training Support Vector Machine, an Application to Face Detection, *IEEE Conf. Computer, Pattern Recognition*, 1997, pp. 130-136.

[17] F. Anifowose, "Prediction of Oil and Gas properties using Support Vector Machine" *IPTC 14514,* 2011.

[18] H. H. Hanafy, S.A. Macary, Y.M., Elnady, A. A.Bayomi,and M.H. El-Batanoney Empirical PVT Correlation applied to Egyptain crude oils exemplify significance of using regional correlations, *Paper SPE 37295 presented at the SPE Oilfield Chemistry International Symposium, Houston*, 1997 pp. 18-21.

[19] M.A Oloso, A. Khoukhi, A Abdulraheem and M. Elshafei, Prediction of Crude Oil Viscosity and Gas/Oil Ratio Curves Using Recent Advances to Neural Networks, *Paper SPE 125360-MS presented at the SPE/EAGE Reservoir Characterization and Simulation Conference, Abu Dhabi, UAE,* 2010, pp. 19-21.

[20] C. Y. Zhoa, H.X Zhang,, M.C Liu,., Z.D Hua,., and B.T Fan, "Application of Support Vector Machine (SVN) for Prediction Toxic Activity of Different Data Sets", Toxicology 217, pp. 105-119, 2006.

[21] P. Xinjun, and W. Yifei, "A Normal Least Square Support Vector Machine (NLS-SVM) and its Learning Algorithm", Neuron computing 72: pp. 3734-3741, 2009.

[22] V. Vapnik, Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing, *Advances in Neural Information Processing Systems 9, Cambridge, MIT Press.* 1997 pp. 281– 287.

[23] K.L., Peng, C.H. Wu. and Y.J Goo, The development of a new statistical technique for relating financial information to stock market returns, *Int.J. Manag. 21 (4),* 2004. 492-505.

[24] P. Scholkopf, A.J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. *MIT Press, Cambridge, MA,* 2002.

Shaibu Samuel, et. al. "Modeling Approach for Niger Delta Oil Formation Volume Factor Prediction using Support Vector Machine."*American Journal of Engineering Research (AJER),* vol. 10(12), 2021, pp. 80-88.