International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 3 ǁ 2022 ǁ PP. 51-61

Modeling Natural Gas Viscosity for Niger Delta Region Using Support Vector Machine (SVM) Algorithm

Atuokwu A. Fortune[*1] and Ijeoma I. Mbachu*[2]
[*1]*University of Port-Harcourt, Department of Petroleum and Gas Engineering, Nigeria.*
[*2]*University of Port-Harcourt, Department of Petroleum and Gas Engineering, Nigeria.*

Abstract
Viscosity of natural gas is an important parameter of theoretical and practical significance in the domain of natural gas recovery, transmission and processing. A small error in gas viscosity affect the inflow performance relationship (IPR) curves and eventually changes the reserves estimate of gas reservoirs. Majority of the exiting correlations give very significant error. This study centers on development of natural gas viscosity prediction model using support vector machine (SVM). MATLAB SVM module was employed in building the model using 332 data set from Niger Delta region of Nigeria. Quantitative and qualitative assessments were employed to evaluate the accuracy of the model both for artificial neural network (ANN) and some of the existing empirical correlations. The new SVM gas viscosity model gave an excellent prediction when compared to other natural gas viscosity models with a mean relative error of 1.0937 and Rank of 1.0934. It also gave the best forecast over artificial neural network model which has a mean relative error of 2.6126 and Rank of 1.0937.The statistical analysis and cross plots demonstrated the superiority of the proposed tool to other existing methods evaluated in this study. The proposed tool is an excellent choice for natural gas viscosity prediction.
Keywords: *Artificial Nural Network, Correlation, Natural Gas Viscosity, Niger- Delta, Support Vector Machine.*

Date of Submission: 09-03-2022 Date of acceptance: 25-03-2022

I. INTRODUCTION

The role of natural gas in meeting the world energy demand has been increasing because of its versatility, abundance, and clean burning [1]. An accurate knowledge of the thermophysical properties of natural gas is necessary for basic petroleum and chemical engineering calculations. One of these properties is the viscosity, which is an important parameter that is frequently used in the equations for single-phase and multiphase flow in gas and oil reservoirs, tubing, and transportation purposes [2]. Thus, the viscosity of a natural gas has to be evaluated for wide ranges of temperature, pressure, and composition [3]. The most reliable and accurate way to obtain thermophysical properties is from accurate experimental measurements. However, the wide range of possible natural gas mixtures and of conditions of interest impedes obtainment of the relevant data by experimental means alone [4]. Therefore, in the absence of experimentally measured properties, it is essential for the gas reservoir engineers to determine the properties from equations of state (EOSs), empirically derived correlations, and soft computing techniques. Many comparative studies have been carried out to ascertain the EOSs ability to estimate the PVT properties of reservoir fluids. The general conclusion is that EOSs have poor ability to predict the volumetric properties of hydrocarbon gas mixtures [5]. Several empirical correlations and corresponding state models have been developed for estimating the gas viscosity under various pressure and temperature conditions. However, correlations that are used to estimate the gas viscosity are sometimes too complex and also have significant error. Hence, introducing a more powerful, fast, and accurate method than the traditional ones is necessary.

Several well-known correlations are used in the petroleum and gas industry to determine the values of natural gas viscosity. Some of these available mathematical equations in the literature are ([6], [7], [8] [9], [5] and [10]. [6] graphical correlations has been the most popular charts in the petroleum industry, because their chart set is perhaps the most complete, including the atmospheric pressure chart, the viscosity ratio charts and correcting charts for non-hydrocarbons. Their correlation was developed, as a function of pseudo-reduced pressure, pseudo-reduced temperature and viscosity ratio. It was reported to have an average of 0.38 absolute error. Carr *et al*. correlation is recommended to be used for gases with specific gravity between 0.55 and 1.22 and a temperature range between 100 and 300^{o}F. [7] derived a correlation by modifying that of [8] to predict the gas viscosity at reservoir conditions. The reported accuracy of this semi-empirical correlation was a standard deviation of ±2.69% with a maximum deviation of 8.99%. [9] measured the viscosity of gas using Cambridge SPL440 viscometer applying metane sample for pressure at 5,000 to 30,000psia and temperatures from 100 to

400°F. From the measurements, [9] modified the [7] model and comparison was made using data from NIST. The results showed a good performance with the NIST data as to compare to the main [7] correlation.

[10] expressed the viscosity ratio μ_g/μ_1. It initially estimates the gas viscosity at the atmospheric pressure and prevailing temperature. The calculated viscosity at the atmospheric is then adjusted for pressure, using the gas pseudo reduced temperature and pressure. [11] used the[12] low- pressure gas viscosity to modify the [7] correlation by nonlinear regression. Their database includes viscosity data for methane, propane, methane/propane, methane/butane, methane/n-decane, and natural gases, plus gas condensate viscosities. They claimed that this method is suitable for both associated gas and gas condensate.

[2] modeled a viscosity relation for gas in surface reservoir condition and developed the correlation using experimental values from gas samples from Nigeria. The authors compared correlations formulated with experimental PVT viscosity and tested the performance by using it to solve two problems from which solutions by the complex Carr et al. [6] were available. [13] presented a model for prediction of the natural gas viscosity based on corresponding states theory. Their model is applicable for gases containing heptane plus and non-hydrocarbon components for a pressure range of 0.10−138.26 MPa, a temperature range of 241−456 K, and a gas specific gravity of 0.57−1.21.

Some authors have also done work on predicting gas viscosity using Artificial Neural Network [15] and [16]. [15] modelled a gas viscosity using Artificial neural network (ANN) based on back-propagation method. They built the model using 3841 experimental data both for testing and training. The designed neural network can predict the natural gas viscosity using pseudo-reduced temperature and pseudo-reduced pressure with AARD% of 0.221. The authors concluded that the comparison indicated that the proposed method can provide accurate results. [16] developed a model for predicting gas viscosity for high pressure high temperature. The authors built the intelligent predictive model using 154 laboratory measured data from Niger Delta using MATLAB ANN tool. The data used was randomly divided into three parts, of which 60% was used for training, 20% for validation, and 20% for testing. They used statistical assessments to assess the accuracy of the new model to the existing empirical correlations. The authors reported that the gas viscosity artificial neural network (ANN) model gave good prediction when compared to other gas viscosity models with a rank of 2.4639 with mean absolute error MAE of 4.3416 and correlation coefficient (R) of 0.995. [17] 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. [18] proposed the application of SVM for prediction of toxic activity with different datasets. When compared with ANN and RBF models, SVM gave the highest correlation coefficient (R). Other successful application of SVM model include [19].

From the literature, conclusion can be drowned that the correlations of gas viscosity available in open access are mainly for empirical and ANN model. Therefore, this paper focuses on modeling natural gas viscosity prediction using support vector machine which is more superior to the other forecasting tool.

1.1 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 hyper plane is constructed. The generalization of SVM is ensured by the special properties of the optimal hyper plane 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 ([20], [21], [22]).

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, ([23], [24], [21]). 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,[25] .The high generalization ability of SVM regression is ensured by special properties of the optimal hyper plane that maximizes the distance to training examples in a high dimensional feature space [24]. 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.

1.2 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 Figure 1. The regression approximation estimates a function according to a given data as shown in Equation 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:

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

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.

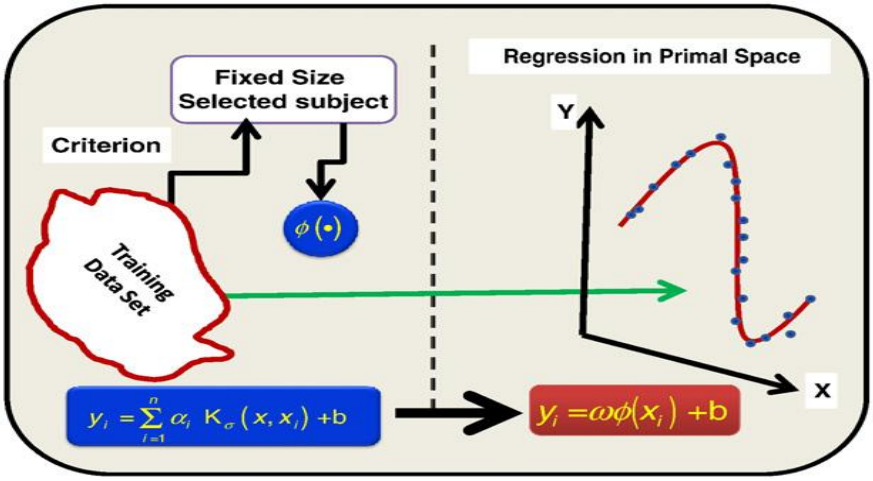

Figure 1: Mapping input space x into high-dimensional feature space [24]

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 (2) 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.

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

Otherwise, 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 Figure 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;

The SVR fits f(x) to the data such that,

-the training error is minimized by minimizing ξ_i and ξ_i^*

-$\frac{\|\omega\|^2}{2}$,is minimized to raise the flatness of $f(x)$, or to penalize excessively complex fitting functions. Thus, SVR is formulated as minimization of the following functional (Equations 5 and 6):

$$Minimimize \quad R_{SVR}(\omega,C)=\frac{1}{2}\left\|\omega^2\right\|+C\sum_{i=1}^{n}L_{\varepsilon}\left(\xi_i+\xi_i^*\right), \tag{5}$$

$$\text{Subjected to } \begin{Bmatrix} y_i-\omega\phi(x_i)-b_i\varepsilon+\xi_i \\ \omega\phi(x_i)-b_i+y_i\varepsilon+\xi_i^* \end{Bmatrix}\xi^{(*)}\geq 0 \tag{6}$$

Where; ξ_i and ξ_i^* denote slack variables that measure the error of the up and down sides, respectively. The above formulae indicate that increasing ε decreases the corresponding ξ_i and ξ_i^* in the same constructed function $f(x)$ thereby reducing the error resulting from the corresponding data points. Finally, by introducing Lagrange multipliers and exploiting the optimality constraints, the decision function given by equation (1) has the following explicit form [22]:

$$f\left(x,\alpha i,\alpha i^*\right)=\sum_{i=1}^{n}\left(\alpha i-\alpha i^*\right)K\left(x_i-x_j\right)+b \tag{7}$$

Where the parameters α_i and α_i^* are called the lagrangian multipliers as represented in equation (6), which satisfy the equalities $\alpha_i\alpha_i^*=0$, $\alpha_i>0$ and $\alpha_i^*\geq 0$ where $i=1,2,\ldots,n$ ([22], [23]),). The term $K(x_i x_j)$ in Equation (6) is defined as kernel function, where the value of kernel function equals the inner product of two vectors x_i and x_j in the feature space $\varphi(x_i)$ and $\varphi(x_j)$ meaning that $k(x_i,x_j)=\varphi(x_i)\varphi(x_j)$. The kernel function is intended to handle any dimension feature space without the need to calculate $\varphi(x)$ accurately. If any function can satisfy Mercer's condition, it can be employed as a kernel function. The typical examples of kernel

function are the polynomial kernel $\left(K(x,y) = [x * y + 1]^d\right)$ and the Gaussian kernel $K(x,y) = \exp\left[-(x-y)^2 \,/\, 2\sigma^2\right]$

In these equations, d represents the degree of the polynomial kernel, and σ^2 represents bandwidth of the Gaussian kernel. These parameters must be selected accurately, since they determine the structure of high-dimensional feature space and govern the complexity of the final solution.

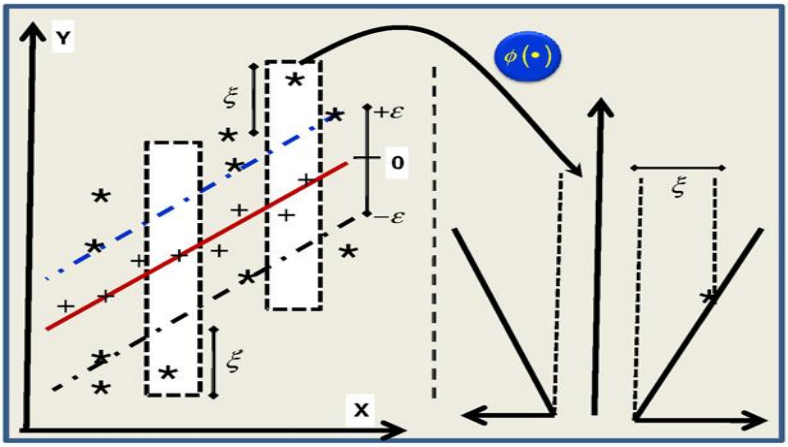

Figure 2: Soft margin loss setting for a linear SVR [25].

II. METHODOLOGY

2.1 Support Vector Machine Regression Model

The model's task is to estimate the functional dependence of the dependent variable y (Viscosity) on a set of independent variables x (reduced temperature and pressure and gas specific gravity). Like other regression model, SVM model assumes that a relationship between the independent and dependent variables exist and it is given by a deterministic function f(x) plus the addition of some additive noise defined in Equation (8):

$$y = f(x) + noise \tag{8}$$

The work is then to find a functional form for f(x) that can correctly predict new cases that the SVM has not been presented with before. This can be achieved by training the SVM model on a sample set (training data).

2.2 Data Collection

The 332 data set used in this research work were obtained from convectional PVT reports that derive the various fluid properties through the liberation process from the Niger-Delta Region of Nigeria. The parameters used are temperature, pressure and gas viscosity. Tables 1, 2 and 3 show the maximum, and minimum data set used in training, testing and validation of the newly developed model.

TABLE 1. Summary of maximum and minimum values of training and test data for SVM gas viscosity model

Range	P (psia)	T (°R)	γ_g
Max	7115.0	734.6	1.45702
Min	130.00	611.0	0.60560

TABLE 2. Summary of maximum and minimum values of validation data for SVM gas viscosity model

Range	P (psia)	T (°R)	γ_g
Max	5160	691	1.00402
Min	510	616	0.61606

TABLE 3. Summary of average values used for training, test and validation of the SVM gas viscosity model

Data	P (psia)	T (°R)	γ_g
Training & Test	2602	656	0.7329
Validation	2554	663	0.7182

2.3 Modeling Process

MATLAB support vector machine module was used to build the natural gas viscosity model. Nine steps were adopted in building this powerful predictive model.

STEP 1: Import the Data: The input data was imported to the MATLAB environment using the import command.

STEP 2: Select the Variables: The variable comprise of Input variables which are Pressure (P), Temperature (T), Gas gravity while the output variables is natural gas viscosity.

$$Input = \left[P \quad data \quad ,T \quad data \ , \ \gamma_g \quad data \right] \tag{9}$$

$$T \text{ arg } et = \left[Gas \quad Vis \text{ cos } ity \quad data \right] \tag{10}$$

Step 3: Division of Data point: The data points were divided into training, testing and validation. 75% of the data was used for training the model, 15% was used for testing the trained model and 10% was used for validation.

Step 4: Choosing Kernel Function: Imbedded in the support vector machine is a function that is design to estimate the model parameters. This function is called the kernel function. Different kernel function exist, (Linear, Polynomial, Radial Basis Function and Sigmoid Function) but Radial Basis Function (RBF) has been far the most popular choice of kernel types used in SVM. This is mainly because of their localized and finite responses across the entire range of the real independent variables. The RBF function is defined in Equation (11)

$$K(x_i, x_i) = e^{-\gamma |x_i - x_j|} \tag{11}$$

Step 5: Estimating Model Parameters: The function contains parameters which must be estimated from modeling and simulating the SVM. These constant are: Capacity (C) and epsilon (ε). The SVM was modeled such that a search for the model parameters was 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 are is fed into the model, it adjusts its weight until the model has being fitted appropriately, which occurs as part of the cross validation process. Unlike supervised learning, unsupervised learning uses unlabeled data. From the data, it discovers patterns that help solve for clustering or association problem. This work used supervised learning approach for the SVM modelling.

STEP 7: Choosing Stopping Criteria: The SVM simulates until a stopping criteria is met, this stopping criteria is defined by the tolerance (t) or number of iteration (N).

STEP 8: Determining Model Accuracy: Several statistical parameters were used to measure the performance of the model such as coefficient of correlation.

STEP 9: Simulation of the SVM Model: This allows performing additional tests on the model or putting it to work on new inputs.

2.3 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 R) into a single comparable parameter called Rank was used [26]. 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;

$$\mathbf{Min\ Z_i} = \sum_{j=1}^{m} \mathbf{S_{i,j} q_{i,j}} \tag{12}$$

Subject to

$$\sum_{i=1}^{n} \mathbf{S_{i,j}} \tag{13}$$

with

$$0 \leq \text{Sij} \leq 1 \tag{14}$$

Where Si,j is the strength of the statistical parameter j of correlation i and qij, the statistical parameter j corresponding to correlation i. j = MRE, MAE, …. R1, where R1 = (1-R) and Zi is the rank, (or weight) of the desired correlation. The optimization model outlined in equations 1 to 3 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 3 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. 12 to 14) 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.

III. RESULTS AND DISCUSSION

3.1 Quantitative Screening Result

The trained SVM model was tested with 49 data points that were not previously used during training and validation. These data points were randomly selected by the MATLAB tool to test the accuracy and stability of the support vector machine model. The performance of the SVM was compared with field data and the prediction from other empirical correlations such as, [2], [7], [10], [13] and ANN model.

[13] correlation was developed based on corresponding data theory and is used to prediction gas viscosity for gases containing heptane plus and non-hydrocarbon components. Form Table 4 and Figure 1, this correlation did not perform well which may be due to the range of temperature and pressure data used in this study. The SVM forecast is very good when compared to other empirical correlations evaluated. The support vector machine model ranked first with a numerical value of 1.0937 and a mean relative error of 1.2481 while the ANN ranked second with a mean relative error of 2.6126 and rank (Z) of 1.9346. The support machine vector model gave better prediction than ANN, followed by Dempsey [10], Ohirhian and Abu [2], Lee et al. [7] and Heidaryan et al. [13].

TABLE 4: Statistical Accuracy of Natural Gas Viscosity Models

Authors	MAE	MRE	SRE	SAE	R	RANK
This Study	0.0185	1.2481	1.6073	1.0129	0.9980	1.0937
ANN	0.2928	2.6126	2.9700	1.4425	0.9917	1.9345
Dempsey (1965)	1.7790	3.7486	4.2364	2.6571	0.9823	2.9078
Ohirhian and Abu (2005)	-2.1967	6.8180	7.3716	4.0832	0.9791	4.3443
Lee et al. (1966)	-9.0053	9.1802	7.7331	7.5173	0.9705	5.2556
Heidaryan et al. (2013)	-35.704	41.378	70.760	67.590	0.5019	33.830

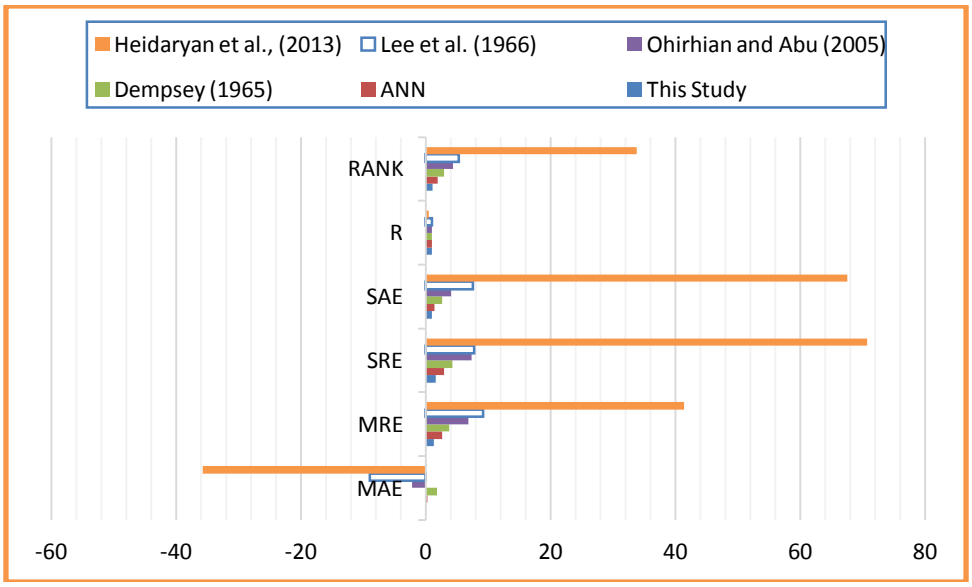

Figure 3: Graphical Representation of Statistical Evaluation for Different Gas Viscosity Correlations

3.2 Qualitative Screening

Figure 9 shows a tight cloud of points around the 45° line indicating a good agreement between the experimental and the calculated data as to compare to Figures 4 to 8. Support Vector Machine model plot (Figure 9) has the tightest cloud of points around the 45° indicating its excellent performance in predicting natural gas viscosity over the ANN model and some of the empirical correlations evaluated.

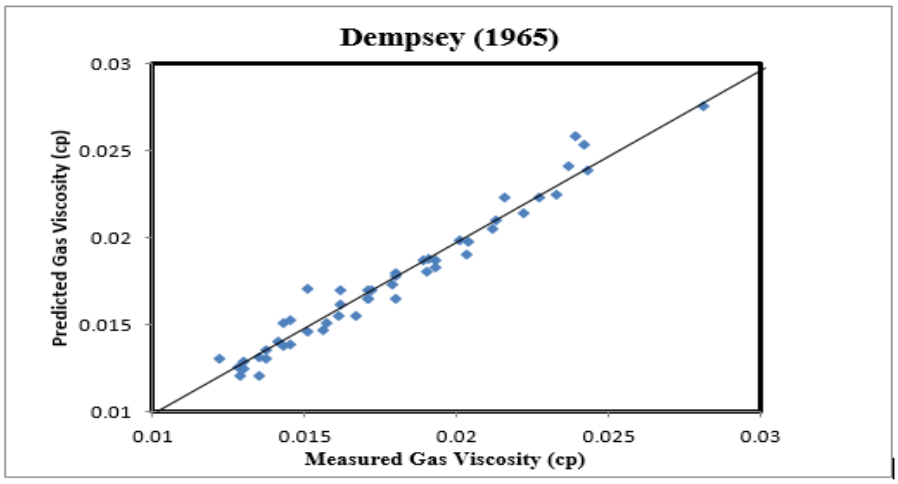

Figure 4: Plot of predicted against measured gas viscosity for [10]

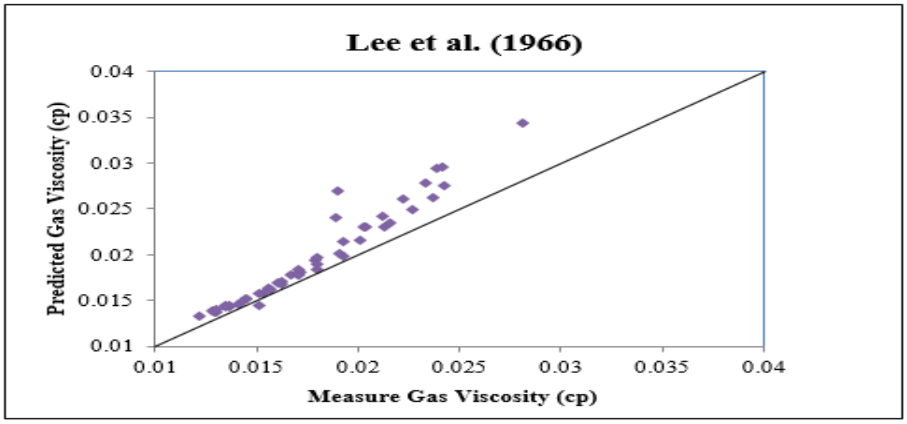

Figure 5: Plot of predicted against measured gas viscosity for [7]

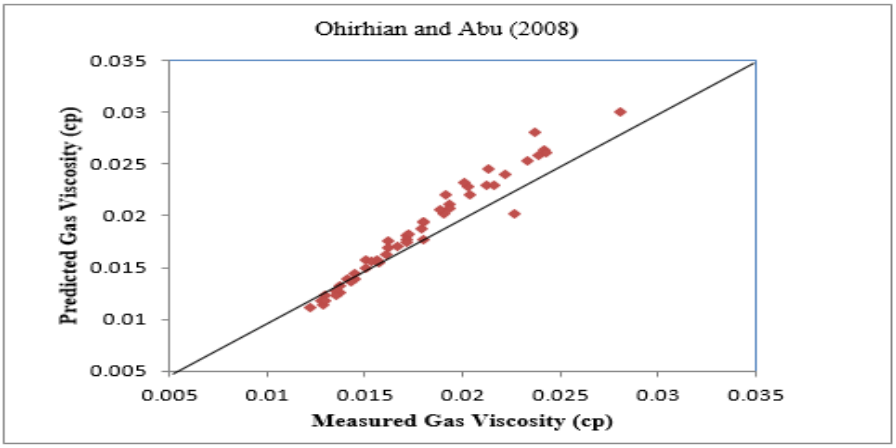

Figure 6: Plot of predicted against measured gas viscosity [2]

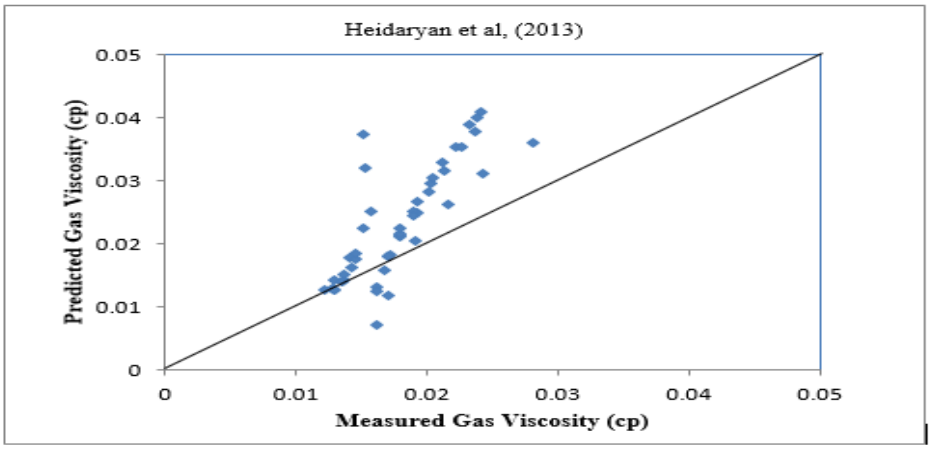

Figure 7: Plot of predicted against measured gas viscosity for [13]

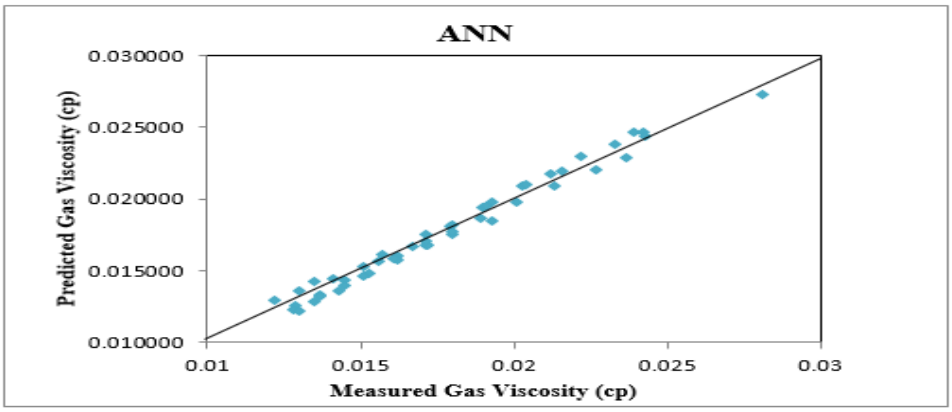

Figure 8: Plot of predicted against measured gas viscosity for ANN

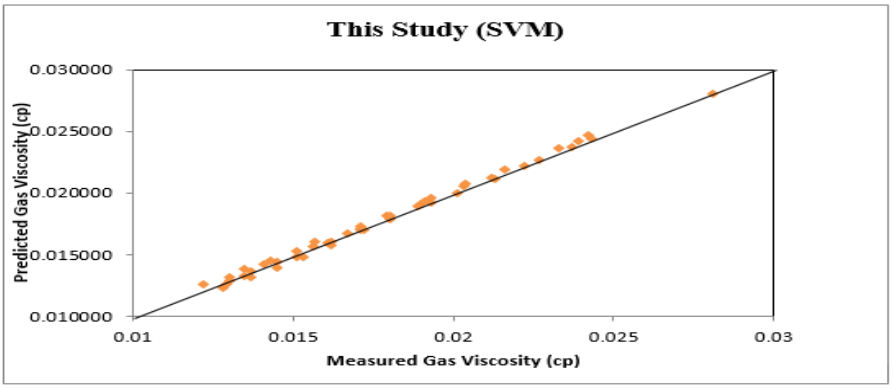

Figure 9: Plot of predicted against measured gas viscosity for this study

IV. CONCLUSION

The SVM regression model can predict gas viscosity from reservoir properties (pressure, temperature and specific gravity). The results show that the SVM regression model developed gives a much better prediction with very high accuracy than the published empirical correlations. The statistical analysis shows the SVM Regression model performs better than the Artificial Neural Network Model and can be employed in predicting gas viscosity. It is recommended that other PVT properties should be modeled using support vector machine algorithm.

REFERENCES

[1]. Wang, X. and Economides, M. (2009)"Advanced Natural Gas Engineering," Gulf Publishing Company: Houston, TX, pp. 400.
[2]. Ohirhian, P.U and Abu I.N. (2008) "A New Correlation for the Viscosity of Natural Gas Viscosity" SPE, 106391.
[3]. Davani, E., Ling, K. Teodoriu, C., McCain, J and G. Falcone, G. (2009) "Inaccurate Gas Viscosity at HP/HT Conditions and its Effects on Unconventional Gas Reserves Estimation", SPE 122827.
[4]. Vesovic, V. (2001)."Predicting the Viscosity of Natural Gas". International Journal of Thermo physical. 2001, 22 (2), pp. 415−426.
[5]. Olatunji, S. O, Selamat, A and. Raheem, A. A (2011) "A Predicting correlations properties of crude oil systems using type-2 fuzzy logic systems". Expert Syst., 38 (9), pp. 10911−10922.
[6]. Carr, N. L., Kobayashi, R. and Burrows, D. B. (1954) "Viscosity of Hydrocarbon Gases under Pressure", Trans-AIME, Vol.201, pp., 270-275.
[7]. Lee, A. J. Gonzalez, M. H. and Eakin B. E. (1996) "The Viscosity of Natural Gases, JPT", Trans., AMIE, 37 pp. 997-1000.
[8]. Starling, K. E., and Ellington, R. T., (1964): Viscosity correlations for nonpolar dense fluids. AIChE J. 1964, 10 (1), pp. 11−15.
[9]. Viswanthan, A. (2007) "Viscosities of Natural Gases at high pressure and high Temperature" Texas A&M University.
[10]. Dempsey, J. R. (1965) "Computer Routine Treats Gas Viscosity as a variable", Oil and Gas Journal, pp. 141-143.
[11]. Sutton, R. P. (2007) "Fundamental PVT Calculations for Associated and Gas/Condensate Natural-Gas Systems". SPE Reservoir. Eng., 10 (3), pp. 270−284.
[12]. Londono, F. E., Archer, R. A. and Blasingame, T. A. (2002): "Simplified Correlations for Hydrocarbon Gas Viscosity and Gas Density- Validation and Correlation of Behavior Using a Large-Scale Database", SPE Paper 75721, presented at SPE Gas Technology Symposium held in Calgary, Alberta, Canada.
[13]. Heidaryan, E. Esmaeilzadeh, F. and Moghadasi, J. (2013) "Natural gas viscosity estimation through corresponding states based models". Fluid Phase Equilib., 354 (0), pp.80−88.
[14]. Mohamad, A. Osfouri S.,and. Mohamadi, B. and Zarei, R., (2015) "Prediction of gas compressibility factor using intelligent models". Natural Gas Industry, pp 283-294.
[15]. Azubuike, I. I. and Ikensimama, S. S. (2020) "Modeling Gas Viscosity of High Pressure High Temperature Condition Using Artificial Neural Network" American Journal of Engineering Research (AJAR), Volume 9, Issue 3.

[16]. Oloso, M.A, Khoukhi, A., Abdulraheem A and Elshafei, M. (2010), "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, pp. 19-21.
[17]. Zhoa, C. Y., Zhang, H.X., Liu, M.C, Hua, Z.D.,and Fan, B.T (2006) "Application of Support Vector Machine (SVN) for Prediction Toxic Activity of Different Data Sets", Toxicology 217, pp. 105-119.
[18]. Xinjun, P. and Yifei W. (2009) "A Normal Least Square Support Vector Machine (NLS-SVM) and its Learning Algorithm", Neuron computing 72: pp. 3734-3741.
[19]. Anifowose, X. (2011) "Prediction of Oil and Gas properties using Support Vector Machine" IPTC 14514, 2011.
[20]. El-Sabahky, E. A. "PVT Properties of Crude Oil Systems Base on Support Vector Machine Modeling Scheme". Medical Artificial Intelligence (MED), Inc., an Elsevier Company, 2009
[21]. Osuna, E., Freund, R. and Giwsi, F."Training Support Vector Machine, an Application to Face Detection, *IEEE Conf. Computer, Pattern Recognition*, 1997, pp. 130-136.
[22]. Joachims, T. (1998) "Text categorization with Support Vector Machines: Learning with many relevant features". Machine Learning: ECML-98. Lecture Notes in Computer Science. Springer. 1398: pp. 137–142.
[23]. Vapnik, V. 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.
[24]. Peng, K.L., Wu. C.H. and Goo, Y.J (2004) "The development of a new statistical technique for relating financial information to stock market returns", *Int.J. Manag. 21 (4)*, 492-505.
[25]. Scholkopf, P. and Smola, A.J. (2002) "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond". *MIT Press, Cambridge, MA.*
[26] Ikensikimama, S. S. (2009)"Reservoir Fluid Property Correlations" Advances in Petroleum Engineering, Chi Ikoku Petroleum Engineering Series, IPS Publications, Port Harcourt. p.53-57.