Modelling viscosity of liquid dropout near wellbore region in gas condensate reservoirs using modern numerical approaches.

F.Faraji, J.O.Ugwu, P.L. Chong, F. Nabhani
School of Science, Engineering and Design, Teesside University, Middlesbrough, TS1 3BA, UK

Abstract

Liquid dropout occurs in gas condensate reservoirs below the dew point pressure around near wellbore region as a result of depletion from production of such reservoirs. Forecasting production as well as optimizing future recoveries of gas condensate reservoirs are highly desirable. This is not possible to achieve without accurate determination of liquid dropout viscosity (μ_c) below the dew point. The focus of research in past decades has been on the development of accurate viscosity prediction models below the dew point pressure to ensure accurate condensate production forecast. Gas condensate production forecast and optimisation around this region and condition are complicated due to unique gas condensate behaviour that violates thermodynamic laws.

Current methods are based on correlation estimation, however the accuracy of these correlations are less than satisfactory, and root cause is due to the miscapturing of complex behaviour of gas condensate reservoir near the wellbore region. These motivated the consideration of modern numerical approaches such as the Least Square Support Vector Machine (LSSVM) and Artificial Neural Network (ANN) used in this paper. These methods are considered as more data behaviour oriented, with the capability of capturing the fluid complexity of gas condensate in such conditions.

In this study viscosity of condensate phase near the wellbore region was modelled using machine learning techniques including ANN and LSSVM. For this purpose, over 300 viscosity data sets were collected from published literature and experimental studies worldwide. This databank includes API gravity, reservoir temperature, solution gas to oil ratio (Rs), specific gas gravity, fluid compositions and reservoir pressure.

Six well known previously published viscosity correlations refined using least-square approach to match the experimental data. Qualitative and quantitative error analysis of developed LSSVM and ANN showed their performance superiority over refined literature correlations. The new proposed models can be embedded as an extra feature of commercial reservoir simulation packages for optimization and future recoveries of gas condensate reservoirs.

Keywords:
Condensate viscosity, Gas condensate, Machine Learning (ML), Least Square Support Vector Machine (LSSVM), Artificial Neural network (ANN), Correlations

Nomenclature and units

API	Oil API gravity
ANN	Artificial Neural Network
LSSVM	Least Square Support Vector Machine
ML	Machine Learning
RMSE	Root Mean Square Error
GOR	Gas to oil Ratio (scf/STB)
Rs	Solution gas to oil ratio (scf/STB)
μ_d	Dead oil viscosity (cp)
μ_{ob}	Live oil viscosity (cp)
μ_c	Condensate viscosity (cp)
HPHT	High pressure high temperature
N	Number of data points
P	Reservoir pressure (psia)
T	Temperature (°F)
cp	Centipoise

1. Introduction

As reservoir pressure reduces in gas condensate to below the dew point due to production, the liquid evolved from gas phase and creates multi-phase flow near the wellbore region. Accumulation of the liquid in aforementioned region is increasing with time and is usually very high in rich gas condensate reservoirs. This phenomenon is called “liquid banking” and can cause severe productivity declines (Wheaton and Zhang, 2007). To understand this complex behaviour in depleting gas condensate reservoirs for forecasting production and optimizing future recoveries viscosity determination of the condensate liquid below the dew point is essential (Audonnet and Pádua, 2004; Kashefi et al., 2013).

In fact, inaccurate estimation of condensate liquid viscosity below the dew point has detrimental effect on cumulative production and can lead to large errors in reservoir performance. Previous studies show 1% error in reservoir fluid viscosity resulted in a 1% error in cumulative production (Al-Meshari et al., 2007; Whitson et al., 1999; Yang et al., 2007).

Measurement of condensate viscosity in gas condensate reservoirs is not made in a routine laboratory test and it may be very difficult to obtain due to unavailability of the samples, lack of high pressure high temperature (HPHT) facilities, small volume cell

viscometers and time and cost required for the measurements. Consequently this makes use of theoretical correlation more attractive (Al-Meshari et al., 2007; Hemmati-Sarapardeh et al., 2014; Whitson et al., 1999).

Depending on the input variables the correlations can be divided to two classes: 1). semi-empirical models that use reservoir fluid composition, critical temperature, acentric factor, pour point temperature, molar mass and boiling point. 2). the correlations which use field data such as reservoir temperature, pressure, API gravity and solution gas to oil ratio "Rs" (Chew and Connally, 1959; Khan et al., 1987). These correlations are deployed for three different conditions of under saturated, saturated and dead oil viscosity.

Condensate liquid viscosity is typically low for depleted gas condensate reservoirs, ranging from 0.1 to 1cp, in the near wellbore region (Al-Nasser and Al-Marhoun, 2012; Whitson et al., 1999). The API gravity of condensate reservoirs are between 40 to 60°API with gas to oil ratio (GOR) between 3000 – 150000scf/STB and temperature between critical temperature (127°C) and cricondentherm temperature (250°C) (Ahmed, 2010; Whitson et al., 2000). The above conditions were our constraint in selecting existing literature viscosity correlations for this study.

Variation of the condensate viscosity with reservoir composition is estimated using the correlation proposed by Lohrenz et al., (1964). This correlation is the most widely used viscosity model, especially in many commercial compositional simulators (ECLIPSE, 2014). Lohrenz et al., (1964), known commonly as LBC, is developed for predicting viscosity of dense gas mixture based on the original work of Jossi et al., (1962) for pure substances using corresponding state principle. Prediction performance of LBc model for viscosity prediction of gas phase in gas condensate reservoirs is reasonable, while prediction of condensate liquid viscosity by this method is very poor (Yang et al., 2007). Consequently, it is necessary to tune the LBC correlation by adjusting its coefficients to match the experimental data. This method is selected because it is taking into account compositional changes based on reduced density, which is characteristic of gas condensate reservoirs below the dew point (Fevang and Whitson, 1996; Mott, 2003).

Gas-saturated-oil (live oil) viscosity correlations are another alternative in literature that can be used to determine the condensate oil viscosity. Yang et al., (2007) suggested to use live oil (μ_{ob}) viscosity correlations to predict condensate liquid viscosity if the measured data is not available. These correlations are function of

solution gas to oil ratio Rs, reservoir pressure, reservoir temperature, fluid API gravity and gas specific gravity (γ_{Gas}). Subsequently these parameters are classed as input variables for developing our ANN and LSSVM models (Fig. 5). Rs is often the most significant component of the PVT correlations, which have big influence on the oil viscosity and should be precisely measured in any selected correlations (Hemmati-Sarapardeh et al., 2014). The solution gas to oil ratio is the amount of gas dissolved in the oil at any pressure. It increases linearly with pressure and it is a function of reservoir fluid composition (Fevang and Whitson, 1996; Jokhio et al., 2002).

The commonly used literature correlations for estimating gas-saturated-oil viscosities and comply with our defined constrained mentioned earlier are Beggs and Robinson, (1975), Kartoatmodjo and Schmidt, (1991), De Ghetto et al., (1994), Elsharkawy and Alikhan, (1999) and Bergman, (2000). The detailed formula of each correlation is given in Table 3. Further description of each correlation include their advantage and disadvantage is given in Appendix 1. These empirical correlations are used to estimate gas-saturated-oil viscosity as a direct function of dead oil viscosity. A brief discussion of each correlation is presented in following.

Beggs and Robinsons, (1975) developed a live oil viscosity correlation based on 2073 observations. The average error of -1.83% have been recorded during testing for proposed correlation. Their correlation is covering solution gas to oil ratio (Rs) within the range of 20 to 2070 scf/STB, oil gravity of 16 to 58°API, pressure range of 0 to 5250 and temperature of 70 to 295°F (Beggs and Robinson, 1975; El Aily et al., 2019).

Using 5321 gas-saturated-oil samples collected globally Kartoatmodjo and Schmidt, (1991) developed a gas-saturated-oil viscosity correlation as a function of dead oil viscosity and Rs. Their correlation can be applied for crude oils in the range of 14.4 to 59°API gravity, temperature range of 80 to 320°F, Rs range of 0 to 2890scf/STB and live oil viscosity range of 0.098 to 586cp (Kartoatmodjo and Schmidt, 1991).

De Ghetto et al., (1994) developed a correlation for light oil viscosity with gravity of API > 31.1 as a function of solution gas to oil ratio (Rs) and dead oil viscosity. His correlation is based on 195 data points collected globally. Their correlation is able to predicts live oil viscosity with less than 10% error within the temperature range of 80.6 to 334.6 °F, Rs of 8.61 to 3299scf/STB and 0.07< μ_{ob}< 295.9cp (De Ghetto et al., 1994).

Elsharkawy and Alikhan, (1999) developed their gas-saturated-oil viscosity correlation utilizing 254 datasets from Middle East oil samples. They concluded their research with 18.6% average absolute relative error obtained from proposed correlation. Their correlation covers the data range of 10 to 3600 for (Rs) and 0.05 to 20.89cp (μ_{ob}) (Elsharkawy and Alikhan, 1999).

Bergman, (2000) developed a gas-saturated crude oil viscosity using 2048 data points collected from worldwide. Bergman's correlation can be used in the range of 5 to 2890scf/STB solution gas to oil ratio (Rs) and live oil viscosity (μ_{ob}) range of 0.125 to 123cp with absolute average error of 9% (Whitson et al., 2000).

All aforementioned correlations developed from crude oil, which has compositional differences with gas condensate fluid composition. Moreover, they are direct function of dead oil viscosity, which is one of the most unreliable properties to be predicted by correlations due to the large effect that oil type (paraffinicity, aromaticity and asphaltene content) has on viscosity (Aily et al., 2019; Whitson et al., 2000). Condensate liquid viscosity in near wellbore region can change significantly during depletion in gas condensate reservoirs (Al-Meshari et al., 2007; Fevang, 1995; Whitson et al., 2000). Consequently, empirical and semi-empirical correlations do not fully reflect the viscosity changes with pressure in gas condensate reservoirs near wellbore region. Therefore, the utilized correlations in this study have tuned to match the experimental condensate liquid viscosity data.

The recent development and success of machine learning techniques in solving complex engineering problems has drawn attention to their various application in petroleum industry (Ahmadi et al., 2014; Ahmadi and Ebadi, 2014a; Ghiasi et al., 2014; Hemmati-Sarapardeh et al., 2014; Kamari et al., 2013; Naderi and Khamehchi, 2019; Shokir, 2008). For gas condensate reservoirs Ahmadi and Ebadi (2014), Elsharkawy and Foda (1998), Jalali et al. (2007) and Nowroozi et al. (2009) were using machine learning (ML) approach for predicting dew point pressure. Zendehboudi et al. (2012) used ML approach to model condensate-to-gas ratio (CGR) of gas condensate reservoirs. Recently Ghiasi et al. (2014) employed LSSVM to predict compressibility factor of gas condensate reservoirs.

Although the aforementioned studies modelled some aspects of gas condensate reservoirs such as dew point pressure, CGR and compressibility factor, however there is a gap in literature for modelling viscosity of gas condensate reservoirs using ML approaches. In fact, to the best of the author's knowledge, there is not any published

work on modelling condensate liquid viscosity of gas condensate reservoirs using any ML approach. Therefore, the aim of this study is to develop novel models for prediction of condensate viscosity in gas condensate reservoirs based on machine learning techniques, namely, Least Squares Support Vector Machine (LSSVM) and Artificial Neural Network (ANN). For this purpose, more than 300 data sets from 13 PVT reports and experimental study were collected and a data bank was created. To establish accuracy of the proposed models an error analysis in terms of coefficient of determination (R^2), root-mean square error (RMSE) and mean square error (MSE) is carried out. In addition, in order to evaluate the performance of the newly proposed models against the existing empirical correlations, graphical and statistical error analysis are utilized (Hagan and Menhaj, 1994).

2. Methodology

2.1 Data acquisition

A database was developed in order to ascertain the accuracy of the proposed methods and examine the suitability of published viscosity correlations. Data from gas condensate PVT reports and also experimental investigation of gas condensate fluid is the base of our data bank. More than 300 data sets have been utilized for developing and testing the models. This data bank includes API gravity, gas specific gravity, reservoir fluid compositions, reservoir pressure, reservoir temperature and initial gas to oil ratio (GOR). Various techniques were used to measure viscosity of the condensate phase such as using electromagnetic pulse technology viscometer, rolling ball viscometer and capillary viscometer.

Ranges, sources and their corresponding statistical parameters of the data are presented in Table 1. The data base represents a comprehensive wide range of gas condensate systems obtained worldwide. Hence, the developed models in this study should be reliable to use in prediction of condensate viscosity below the dew points globally within the specified pressure and temperature.

Author	**Source of data**	**Pressure(psia)**	**Tem (°F)**	**Solution GOR(Rs)**	**μ(cp)**
Al-Meshari et al., (2007)	Saudi Arabia	0 – 5000	243	334 – 6759	0.264 – 0.561
Yang et al., (2007)	Norway	630 – 7014	338	1889 – 10279	0.178 – 0.271

Kashefi et al., (2013)	Binary Mixture	6011 – 20023	122 – 302	8125 – 25067	0.034 – 0.199
Thomas and Bennion, (2009)	Recombined fluid	2900 – 10600	246	2985 – 11812	0.076 – 0.62
Chen et al., (1995)	North Sea	4520 – 5733	259	7195 – 9264	0.1175 – 0.1572
Wheaton and Zhang, (2007)	Mixture C1-C7	304 – 2393	160	283 – 2661	0.04 – 0.141
Saeedi and Rowe, (1981)	US	253 – 2730	109 – 189	1889 - 10279	0.171 – 0.271
Gozalpour et al., (2005)	Binary Mixture	549 – 5019	100	6859 – 8592	0.0386 – 0.042
Guo et al., (1997)	Binary Mixture	2610 – 5366	110 – 262	5551 – 6000	0.45 – 0.67
O'Dell and Miller, (1967)	US, Texas	1500 – 3500	Unknown	2027 – 4731	0.075 – 0.27
Fetkovich et al., (1986)	North Sea	2827 – 6791	155	3686 – 9180	0.171 – 0.332
Ghahri et al., (2011)	Binary Fluid	800 – 5255	Unknown	1081 – 7103	0.0261 – 0.1411
Audonnet and Pádua, (2004)	Binary Fluid	14 – 10877	76 – 247	18 – 14703	0.086 – 1.672

Table 1. The origin and the ranges of data used for condensate liquid viscosity study.

2.2 Prediction of liquid dropout viscosity using literature correlations

The (Lohrenz et al., 1964) correlation shown in Eq. (1) is one of the most common methods in petroleum industry for estimating the viscosity of petroleum fluid and

commonly known as LBC method. The LBC is based on generalised relationship between viscosity and fourth degree polynomial of the reduced density.

$$[(\mu - \mu^*)\zeta + 10^{-4}]^{\frac{1}{4}} = A_0 + A_1\rho_r + A_2\rho_r^2 - A_3\rho_r^3 + A_4\rho_r^4 \quad (1)$$

Where ζ is the viscosity reducing parameter shown in Eq. (2), ρ_{pr} is reduced density calculated by Eq. (3), μ^* is low pressure gas mixture viscosity defined by Eq. (4), A_{0-4} are LBC coefficients of 0.1023, 0.023364, 0.058523, -0.040758 and 0.0093324 respectively.

$$\zeta = 5.35\left(\frac{T_{pc}}{Mi^3 P_{pc}^4}\right)^{1/6} \quad (2)$$

$$\rho_{pr} = \frac{\rho}{\rho_{pc}} = \frac{\rho}{M}\nu_{pc} \quad (3)$$

$$\mu^* = \frac{\sum_{i=1}^{N} z_i\mu_i}{\sum_{i=1}^{N} z_i\sqrt{M_i}} \quad (4)$$

Kay's mixing rule (Kay, 1936) is utilized to calculate the pseudocritical properties of temperature Tpc, pressure Ppc and volume ν_{pc}. In Eq. (4) z_i is the mole fraction of each pure components i and M_i is molecular weight of each component.

To establish special relation between C_{7+} fractions and critical volume Eq. (5) suggested by (Lohrenz et al., 1964) is used.

$$v_{cC_{7+}} = 21.573 + 0.015122\,M_{C_{7+}} - 27.65\gamma_{C_{7+}} + 0.070615\,M_{C_{7+}}\gamma_{C_{7+}} \quad (5)$$

Where $v_{cC_{7+}}$ is the critical molar volume, $M_{C_{7+}}$ is molecular weight and $\gamma_{C_{7+}}$ is specific gravity of C_{7+} fraction.

The component viscosities, μ_i in Eq. (4) is calculated using (Stiel and Thodos, 1962) expression as follows.

$$\begin{cases} \mu_i\zeta_i = (34\times10^{-5})Tr^{0.94} & for\ Tr \le 1.5 \\ \mu_i\zeta_i = (17.78\times10^{-5})(4.58Tr - 1.67)^{5/8} & for\ Tr > 1.5 \end{cases} \quad (6)$$

In LBC correlation viscosity unit 'μ' is in centipoise (cp), viscosity reducing parameter 'ζ' is in cp^{-1}, ρ is in lbm/ft^3, specific volume 'v_c' is in ft^3/lbm mol, temperature 'T' is in Rankine (°R), pressure 'P' is in psia, and molecular weight of each component 'Mi' is in lbm/lbm mol.

The prediction capability of the LBC for viscosity measurement of the hydrocarbon liquid especially in gas condensate reservoirs below the dew point is very poor and rapid increase in liquid viscosity cannot be represented by original LBC correlation (Ali, 1991; Hernandez; et al., 2002; Yang et al., 2007). The result of this study illustrated in Fig. (1a) also indicates LBC performance in predicting condensate liquid viscosity is very poor. Hence LBC correlation has been regressed using least-square approach to match the experimental viscosity data. The procedure for tuning of the LBC correlation recommended by Yang et al., (2007) followed in this study. The coefficients of A_{0-4} in LBC correlation Eq. (1) has tuned and new coefficient values are presented in Table. 2.

Coefficients	New values
A0	0.11364
A1	0.02173
A2	-0.20666
A3	0.06283
A4	0.17139

Table 2. The new coefficients for LBC correlations.

Fig (1a) depicts the prediction performance of LBC correlation with default and regressed values in predicting condensate viscosity. As it can be seen the performance of the LBC correlation improved significantly after tuning the coefficients. The second types of the empirical correlations, used in this study correlate gas-saturated-oil viscosity as a function of deal oil viscosity and solution gas to oil ratio. Six well known published literature correlations were selected for this purpose. The prediction performance of gas-saturated-oil correlations found to be poor in forecasting viscosity of condensate liquid and the results associate with large error. Therefore, in this study these correlations have been refined to match the experimental measurements. Table 3 depicts the original and tuned form of the utilized correlations for predicting condensate liquid viscosity.

Graphical error analysis of the refined literature correlations in predicting condensate viscosity is presented in Fig (1b-1f). The slope line of 45° in aforementioned figures representing zero error line in matching between measured and calculated values (Mansour et al., 2013). Qualitative error analysis in terms of coefficient of determination (R^2), absolute average relative deviation percentage (AARD%), mean square error (MSE) and root-mean square error (RMSE) has been applied. From Fig (1b-1f), and also quantitative error analysis in Table 6, Kartoatmodjo and Schmidt, (1991) outperforms other methods followed by Elsharkawy and Alikhan, (1999),

Bergman, (2000), De Ghetto et al., (1994), Beggs and Robinson, (1975) and LBC, (1964) correlation. The results of tuned correlations compared to the proposed LSSVM and ANN numerical methods, which will be discussed later.

Author	Correlation	Tuned correlation
Beggs & Robinson, (1975)	$\mu_{ob} = A(\mu_{od})^B$ $A = \frac{10.715}{(R_s + 100)^{0.515}}$ $B = \frac{5.44}{(R_s + 150)^{0.338}}$	$\mu_c = A(\mu_{od})^B$ $A = \frac{17.99}{(R_s + 100)^{0.515}}$ $B = \frac{4.056}{(R_s + 150)^{0.338}}$
Kartoatmodjo & Schmidt, (1991)	$\mu_{ob} = -0.06821 + 0.9824X_1 + 4.034 \times 10^{-4}{X_2}^2$ $X_1 = 0.43 + 0.5165 \times 10^{(-8.1\times10^{-4}R_s)}$ $X_2 = \left[0.2001 + 0.8428 \times 10^{(-8.1\times10^{-4}R_s)}\right]{\mu_{od}}^{X1}$	$\mu_c = -0.30612 + 1.174X_1 + 4.034 \times 10^{-4}{X_2}^2$ $X_1 = 0.43 + 0.5165 \times 10^{(-8.1\times10^{-4}R_s)}$ $X_2 = \left[0.2001 + 0.8428 \times 10^{(-8.1\times10^{-4}R_s)}\right]{\mu_{od}}^{X1}$
De Ghetto, (1994)	For (°API > 31.1) $\mu_{ob} = A(\mu_{od})^B$ $A = \frac{25.192}{(R_s + 100)^{0.6487}}$ $B = \frac{2.7516}{(R_s + 150)^{0.2135}}$	$\mu_c = A(\mu_{od})^B$ $A = \frac{62.96}{(R_s + 100)^{0.6487}}$ $B = \frac{2.1334}{(R_s + 150)^{0.2135}}$
Elsharkawy & Alikhan, (1999)	$\mu_{ob} = A(\mu_{od})^B$ $A = 1241.932(R_s + 641.026)^{-1.12410}$ $B = 1768.84(R_s + 1180.335)^{-1.06622}$	$\mu_c = A(\mu_{od})^B$ $A = 3978.167(R_s + 641.026)^{-1.12410}$ $B = 1361.93(R_s + 1180.335)^{-1.06622}$
Bergman, (2000)	$\mu_{ob} = A(\mu_{od})^B$ $A = e^{[4.768 - 0.8359\ln(R_s+300)]}$ $B = 0.555 + \frac{133.5}{R_s + 300}$	$\mu_c = A(\mu_{od})^B$ $A = e^{[4.6792 - 0.7772\ln(R_s+300)]}$ $B = 0.555 + \frac{133.5}{R_s + 300}$

Table 3. The original and tuned form of the employed literature correlations for predicting condensate liquid viscosity.

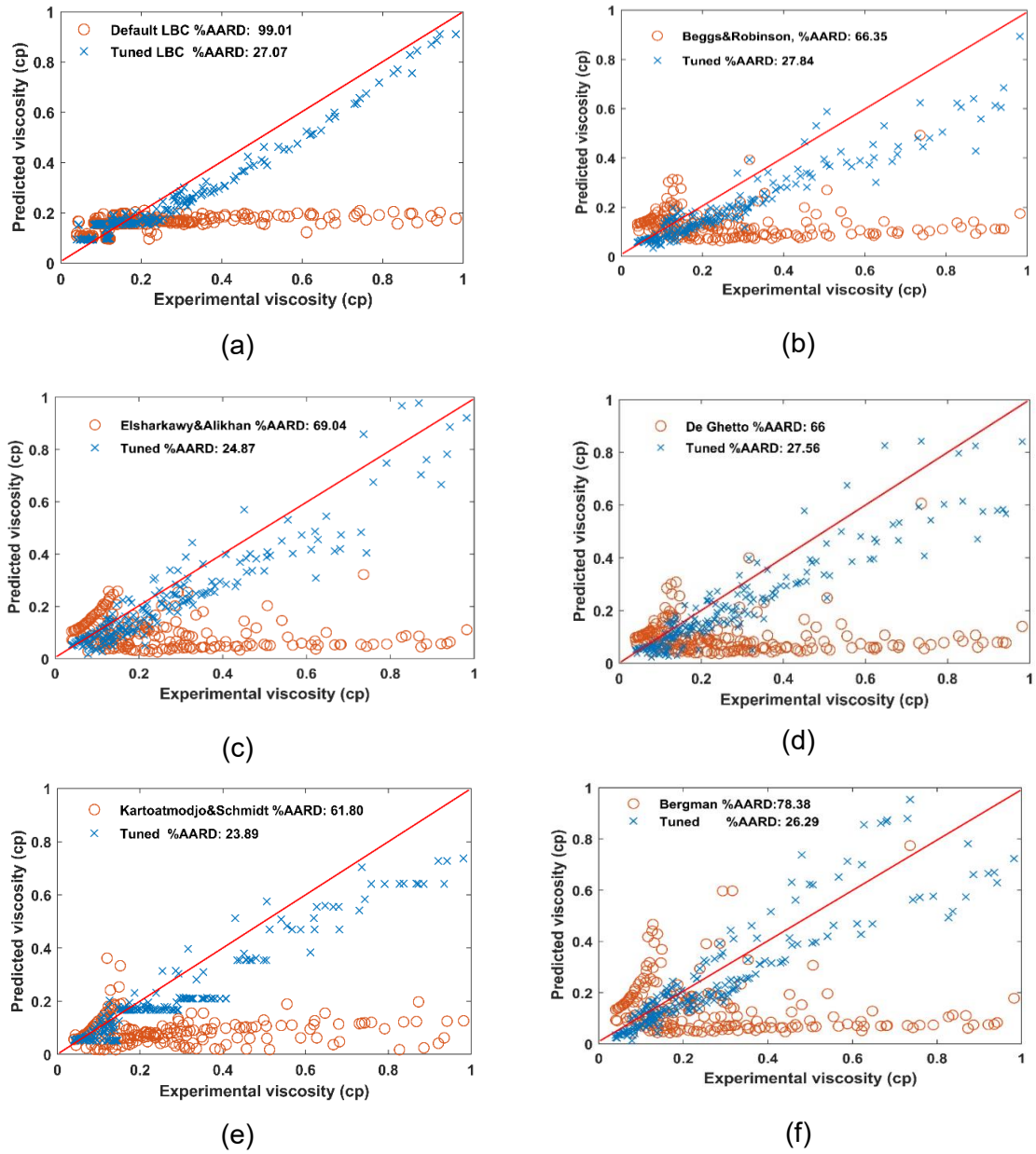

Fig. 1. Cross plot of the experimental viscosity versus predicted viscosity using employed correlations and their tuned results.

2.3 Least square support vector machine (LSSVM)

The support vector machine (SVM) has been identified as an efficient and powerful strategy developed from the machine-learning community (Cortes and Vapnik, 1995; Curilem et al., n.d.; Suykens et al., 2002). SVM is a tool for a set of related supervised learning methods that analyse data and recognize pattern using regression analysis and it is identified as a non-probabilistic binary linear classifier. The objective of this study is to develop a nonlinear relationship between the available experimental data considered as inputs (pressure, temperature, API gravity, gas to oil ratio and gas specific gravity) and the desired output (liquid dropout or condensate liquid viscosity)

(Ahmadi and Ebadi, 2014; Eslamimanesh et al., 2012; Hemmati-Sarapardeh et al., 2014; Kamari et al., 2013).

SVM method has many advantages over other machine learning techniques as follows: they are more likely to converge to the global optima, prior determination of the network is not required in this model and can be automatically determined as the training ends. Furthermore, the number of hidden layers and hidden nodes should not be determined and this algorithm has fewer adjustable parameters compared to ANN network (Eslamimanesh et al., 2012; Suykens et al., 2002).

Original SVM algorithm requires implementing set of nonlinear equations using quadratic programming, which is very hard to implement. Also the obtained outputs using SVM algorithm is much scattered for both linear and nonlinear regressions (Eslamimanesh et al., 2012; Suykens et al., 2002; Suykens and Vandewalle, 1999). To overcome abovementioned problems Suykens and Vandewalle, (1999) suggested a modification to the original SVM algorithm named Least-Squares Support Vector Machine (LSSVM). The LSSVM only requires solving set of linear equations, makes it easier to implement and faster alternative to the original SVM method (Eslamimanesh et al., 2011; Pelckmans et al., 2002; Suykens and Vandewalle, 1999). Suykens and Vandewalle, (1999) defined the cost function (J) for LSSVM by Eq. (7).

$$J = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^{N} {e^2}_k \tag{7}$$

Eq. (7) is subjected to the following constraint:

$$y_k = [w^T \varphi(x_k) + b + e_k], \qquad k = 1, \dots, N. \tag{8}$$

Where, x_k is input vector containing the input parameters (pressure, temperature, solution gas to oil ratio and gas specific gravity), y_k is output vector (condensate liquid viscosity), b stands for intercept of linear regression in LSSVM method, w stands for regression weight, e_k is the regression error for N training objects in least-squares error approach, γ is relative weight of the summation of the regression errors compared to the regression weight (right hand side of Eq. (7), φ is the feature map, mapping the feasible input region to the high dimensional feature space and transcript T stands for transposing the matrix.

Applying Lagrangian function, the regression weight w can be defined in Eq. (9).

$$w = \sum_{k=1}^{N} \alpha_k x_k \tag{9}$$

Where

$$\alpha_k = 2\gamma e_k \tag{10}$$

a_k denotes to the Lagrange multiplier, that may be either positive or negative, since LSSVM has equality restrictions. Assuming linear regression between the inputs and output parameters of LSSVM algorithm, Eq. (8) is re-written as follows (Pelckmans et al., 2002; Suykens et al., 2002; Suykens and Vandewalle, 1999).

$$\alpha_k = \frac{(y_k - b)}{x_k{}^T x + (2\gamma)^{-1}} \tag{11}$$

The linear regression in Eq. (11) can be converted to a nonlinear using the Kernel function in Eq. (12)

$$f(x) = \sum_{k,l=1}^{N} \alpha_k K(x, x_k) + b \tag{12}$$

Where $K(x, x_k)$ represents dependency of Kernel function to the inner values of two vectors x and x_k in the feasible region built by the inner product of the vectors $\phi(x)^T$ and $\phi(x_i)$ as follows: (Cortes and Vapnik, 1995; Eslamimanesh et al., 2012; Fazeli et al., 2013; Suykens et al., 2002; Suykens and Vandewalle, 1999).

$$K(x, xk) = \phi(x)^T \phi(x_k) \tag{13}$$

The radial basis function (RBF) Kernel defined in Eq. (14) has been executed. (Cortes and Vapnik, 1995; Eslamimanesh et al., 2012; Pelckmans et al., 2002; Suykens et al., 2002):

$$K(x, x_k) = \exp\left(-\frac{\|x_k - x\|^2}{\sigma^2}\right) \tag{14}$$

Where σ in Eq. (14) and γ in Eq. (7) are tuning parameters of LSSVM and can be determined by any external optimization algorithm. Robust Simulated Annealing (SA) algorithm in MATLAB optimization toolbox has been used to find the optimum values of these parameters. The root mean square error (RMSE) between the developed LSSVM model obtained results and experimental values, defined by Eq. (15), was considered as an objective function during the SA computation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(Vis_{esti} - Vis_{expi}\right)^2}{ns}} \tag{15}$$

Where Vis represents condensate viscosity, subscripts *est* and *exp* represent the predicted and actual value, *ns* is number of data points from the initial assigned population of 144 data sets. The optimized values of γ and σ^2 using SA optimization method for predicting the condensate liquid viscosity presented in Table 4.

LSSVM model	Input parameters	Model parameters	
		γ	σ^2
Condensate phase viscosity	Reservoir pressure, Temperature, API, gas SG, Rs	5625.256	23.65

Table 4. The optimum values of the LSSVM parameters.

In this study the data is divided into three subsets of "Training", "Optimization" and "Testing". Training set is used for generating the model structure, optimization is used for minimization of the error in trained model and test data is used to investigate the prediction capability of the developed model.

The database was randomly split into three sub data sets of 80% training,10% testing and 10% validation. The allocation percentage of the data is selected according to the recommendations by Ahmadi and Ebadi, (2014) and Eslamimanesh et al., (2012). During the training of the model cross validation has been performed where, the training data sets into several folds and accuracy of each fold checked. Table 5 is presenting the statistical error analysis of the LSSVM in each stage of training, optimizing and testing.

Input variables for this model are as pressure, temperature, API gravity, gas specific gravity and solution gas to oil ratio "Rs". The acceptable distribution of the data is one with homogeneous accumulations of the data on the domain of the three data sets (Eslamimanesh et al., 2011; Gharagheizi et al., 2014).

The MATLAB code for trained LSSVM model generated and prediction capability of the trained model was tested for new data sets. The graphs in Fig. 2 and Fig. 3 are indicating the performance of LSSVM model in training stage and in predicting new experimental set of data (testing stage), respectively. The majority (73%) of the data points in this study are within lower viscosity range of 0 – 0.4cp. Therefore, the testing of the data is toward lower viscosity region, which is more realistic characterisation of gas condensate viscosity below the dew point near wellbore region (Whitson et al., 1999; Yang et al., 2007). The viscosity of condensate liquid in near wellbore region, where condensate liquid in mobile is very low. This is due to existence of more lighter

C_{7+} fractions in mobile condensate liquid composition in aforementioned region (Fevang, 1995, p. 44). Even though the higher viscosity prediction in Fig. (3) has higher error than the lower viscosity prediction, the AARD% is still reasonably small because the majority (73%) of the values are in lower viscosity region.

Fig. 4 is representing residual plot of LSSVM trained data. Ability of the trained LSSVM in predicting new data sets are also analysed by presenting graph of standard deviation error in Fig. 5 and standard error from the mean in Fig. 6.

Stage of the process	R^{2a}	RMSE[b]	MSE[c]	AARD%[d]
Training set	0.9139	0.10845	0.01176	13.96
Optimization set	0.87256	0.111121	0.012348	14.12
Testing set	0.7723	0.121037	0.01465	14.25

Table 5. Statistical error performance of the LSSVM.

a $R^2 = 1 - \frac{\sum_i^N (cal.((i))/Est.(i) - exp.(i))^2}{\sum_i^N \left(cal.((i))/Est.(i) - average\ (exp.(i))\right)^2}$

b $RMSE = \left(\frac{\sum_i^N (cal.((i))/Est.(i) - exp.(i))^2}{N}\right)^{0.5}$

c $MSE = \left(\frac{\sum_i^N (cal.((i))/Est.(i) - exp.(i))^2}{N}\right)$

d $AARD\% = \frac{100}{N} \sum_i^N \frac{|(cal.((i))/Est.(i) - exp.(i)|}{exp.(i)}$

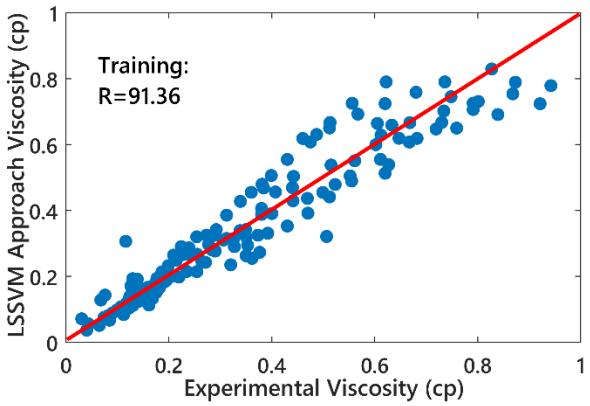

Fig. 2. Performance of the LSSVM trained model (R^2=0.9136).

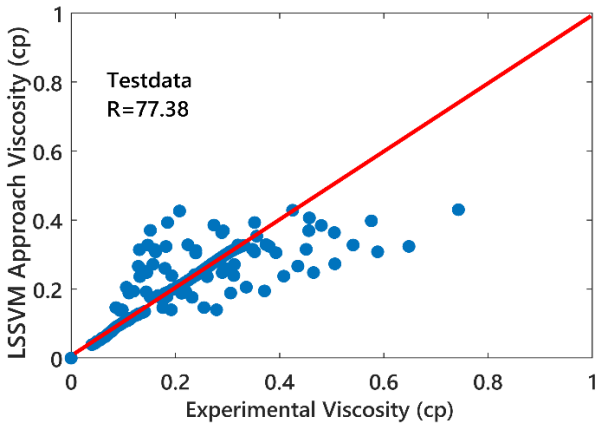

Fig. 3. Performance of LSSVM in predicting new data (R^2=0.7738).

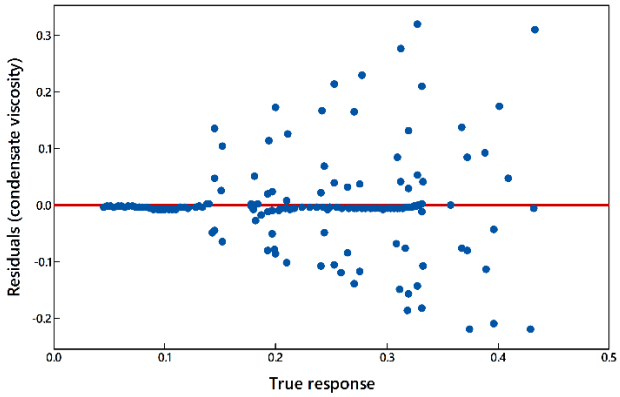

Fig. 4. Residual plot of LSSVM trained data.

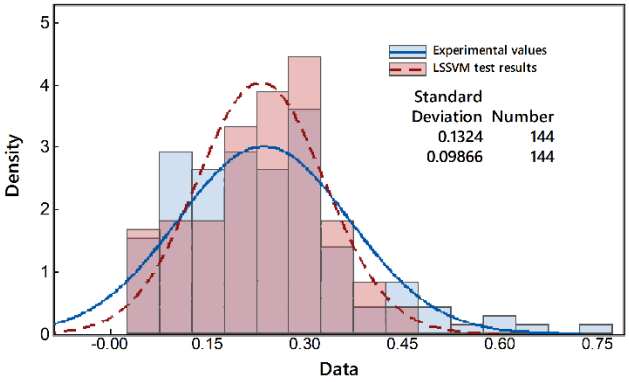

Fig. 5. Graph of standard deviation of LSSVM method against experimental data.

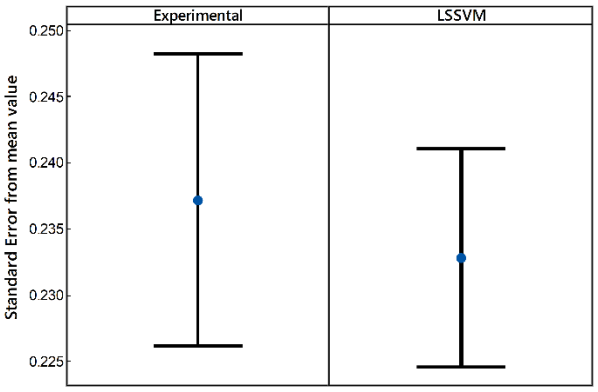

Fig. 6. Interval plot of experimental data against LSSVM approach.

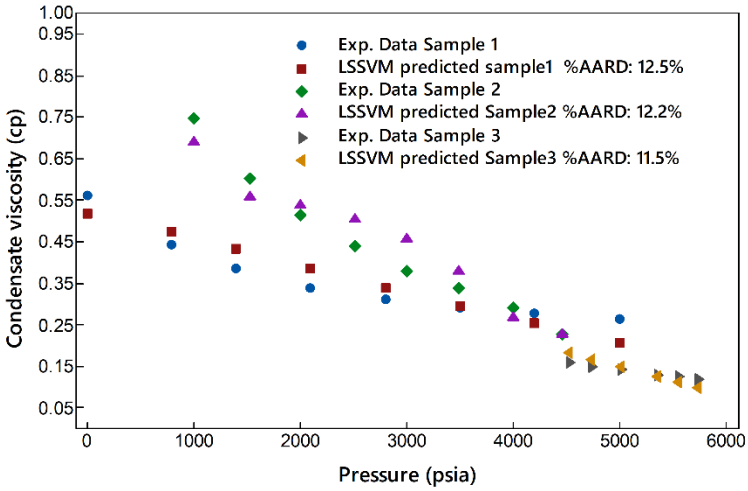

Fig. 7. Comparison between experimental and predicted values provided by LSSVM for three samples of condensate fluid viscosity.

The primary aim of this study was to develop a model that predicts viscosity of the condensate liquid in depleted gas condensate reservoirs with high accuracy using machine learning techniques. The results indicate that LSSVM is performing better than tuned literature correlations. However, the error is still high, approximately about 23% in testing stage, where the capability of the model assessed using new data sets. Therefore, to certify the effectiveness and accuracy of the suggested LSSVM model for estimation of condensate viscosity among smart approaches in another attempt an Artificial Neural Network (ANN) was developed, which is presented in following section.

2.4 Artificial Neural Networks (ANN)

A detailed description of neural networks can be found in Cios and Shields (1997), Dreyfus (2005) and Haykin (1994). ANN is a computational technique in artificial intelligence that uses complex computation system for predicting the output responses. ANNs are inspired by biological networks, performing in a massive parallel

connection between nonlinear, parametrized, and bounded functions called neurons (Cios and Shields, 1997; Mesbah et al., 2017).

Such a network is a massively parallel-distributed processor that has a natural tendency for storing experimental knowledge and making it available for future use. In ANN system knowledge is acquired by the network through a learning process and synaptic weights will store this knowledge (Haykin, 1994). Hence, mathematical interpretation of the problem does not required. Neurons in such a system coordinate their work, and they transfer information by using synapses “electromagnetic communications” (Ghaffari et al., 2006). Through a set of known input (5 in this study) and output data (1 in this study), the network will be trained. Through a learning process the network monitors the error between the predicted and desired outputs and continue to adjust the weights until the optimization criteria are reached. This process is usually carried out in two stages: first the input variables are linearly combined, then the result is used as argument of non-linear *activation function (a)*. The activation function must be non-decreasing and differentiable function; the most common choices are either the identity function $(y = x)$, or bounded sigmoid (s-shaped) function, as the logistic $[y = 1/(1 + e^{-x})]$ (Eslamimanesh et al., 2011; Ghaffari et al., 2006; Haykin, 1994; Hippert et al., 2001).

The neurons are organized in a way that define the network architecture. We used multilayer perception (MLP) type, in which the neurons are organized in layers Fig. (8). The neurons in each layer may share same inputs, but they are not connected to each other. The neural networks consist of hidden layers, output layer, inputs and bias units. Number of hidden layers and number of neuron of each layers can be arbitrary (Khosrojerdi et al., 2016). However, increasing number of neurons may cause overfitting while decreasing their numbers may result on poor performance of the network. The main advantage of ANN is ability to process large amount of data sets (Ghaffari et al., 2006; Khosrojerdi et al., 2016; Mesbah et al., 2017; Hippert et al., 2001).

Fig. (8) depicts the schematic diagram of ANN structure for predicting viscosity of condensate liquid fluid. This design has one layer for inputs consists of five input parameters, one hidden layer, two bias units and one output unit. This architecture recommended by Hagan et al. (2014), Hagan and Menhaj (1994) and Hippert et al. (2001) as an efficient and the most popular multilayer feed-forward architecture. Nevertheless, there is large number of other designs, which might be considered

suitable for other applications. Further information about ANN network architecture used in this study is presented in Appendix B.

The network is designed in MATLAB and calculations carried out by implementing different number of neurons in hidden layer (layer 2). To select the best architecture in terms of number of neurons in a hidden layer a trial and error procedure was implemented. The performance of each structure was assessed by comparing coefficient of determination (R^2) and root man square error (RMSE). We came up with the proposed structure in Fig. (8) (5 neurons in layer 2) as the best topology.

The aforementioned architecture performance evaluation is required to determine the complexity of a neural network as one of the important factors. Hagan et al, (2014) and Soroush et al, (2015) highlighted importance of level of complexity in neural network structure to ovoid overfitting with higher number of neurons and poor performance with not enough number of neurons.

Our input parameters are API gravity, solution gas to oil ratio (Rs), pressure, temperature and gas specific gravity. The output layer is viscosity of condensate fluid calculated by the ANN network. There are many algorithms available to train the network and minimize the error and find the optimum values of the weights and biases; including Levenberg–Marquardt (LM), scaled conjugate gradient (SCG), and resilient back propagation (Hippert et al., 2001; Soroush et al., 2015).

The LM backpropagation algorithm introduced by Kenneth, (1944) and recommended by Behera and Chattopadhyay, (2012) as one of the fastest and most popular backpropagation algorithm was used for adjusting the weights in this study. The tangent sigmoid transfer functions set for the neurons in hidden layer.

For training of the model 70% of whole data bank (210 data points) randomly selected and split to three data sets of 80% (168 data points) for training, 10% (21 data points) for validation and 10% (21 data points) for testing.

The ANN network is trained to map input data by iterative adjustment of the weight function. Information from inputs feed forwarded through the network to optimize the weight between the neurons. Optimization of the weight function is carried out by back propagation of the error during training or learning stages. The ANN reads the inputs and output values in training stage and changes the value of weight functions to minimise the difference in predicted and the target (observed) values. The error in prediction is minimized across training iterations (epochs) and training continues to the point that the network reaches a specified level of accuracy (Ghaffari et al., 2006).

Once the model has reached satisfactory accuracy or the model is converged, the training will stop. The performance of the ANN trained model for the training stage is presented in Fig. (9) and Fig. (10).

Fig. (7) and Fig. (11) depict the performance of the developed LSSVM and ANN models respectively in predicting the condensate viscosity data. As it can be seen from the aforementioned figures both LSSVM and ANN network predict the independent sample data with satisfactory accuracy. This will be discussed in details in results section.

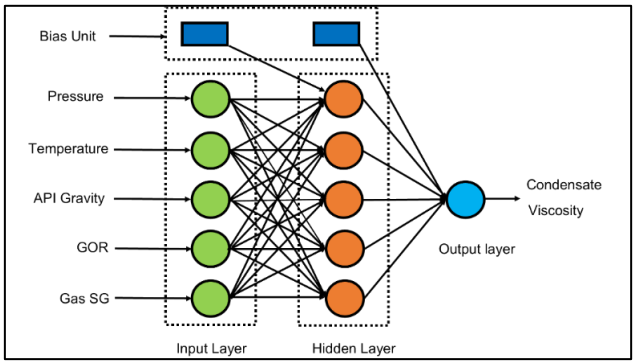

Fig. 8. Developed ANN model architecture for prediction of condensate liquid viscosity.

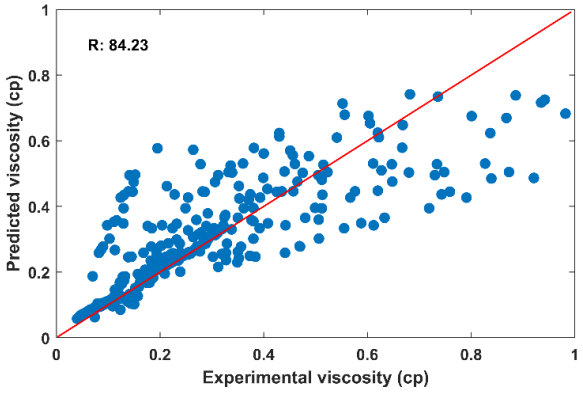

Fig. 9. Prediction performance of developed ANN network for condensate liquid viscosity in training stage.

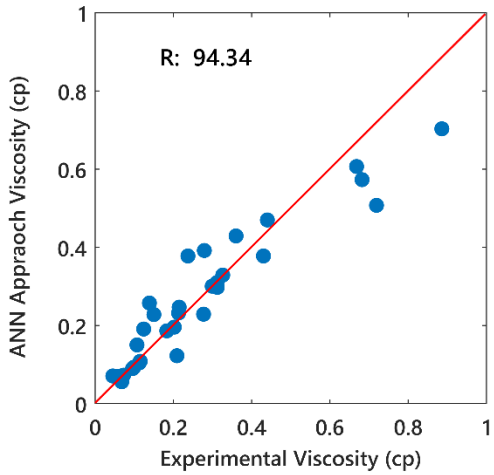

Fig. 10. Prediction performance of ANN network for condensate liquid viscosity in testing stage.

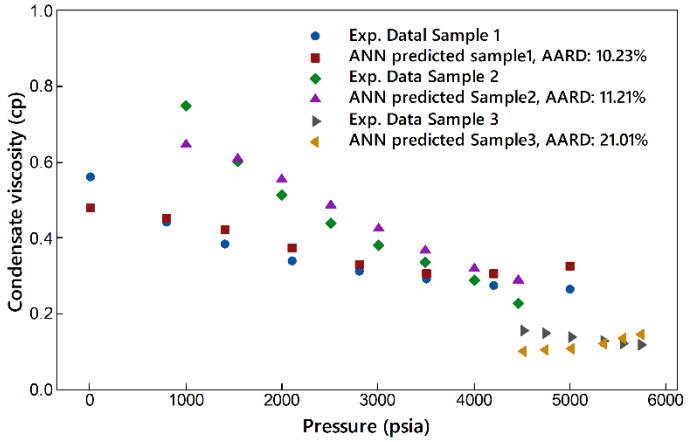

Fig. 11. Prediction performance of ANN model for 3 condensate liquid viscosity samples as a function of pressure.

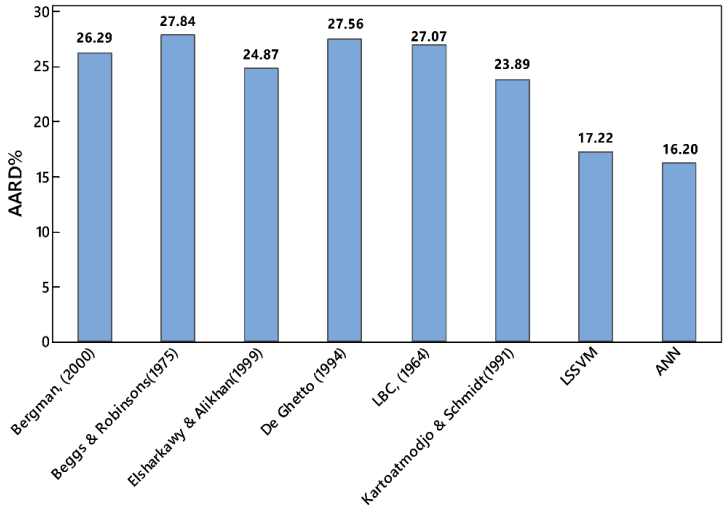

Fig. 12. Performance comparison of employed methods in this study in predicting experimental condensate liquid viscosity.

Method	R^2	RMSE	MSE	AARD%
LBC (1964)	0.7241	0.1240	0.0154	27.07
Bergman (2000)	0.7297	0.1236	0.0153	26.29
Beggs and Robinson (1975)	0.7207	0.1244	0.0155	27.84
Elsharkawy and Alikhan (1999)	0.7344	0.1228	0.0151	24.87
De Ghetto (1994)	0.7243	0.1240	0.0154	27.56
Kartoatmodjo and Schmidt (1994)	0.7412	0.1220	0.0149	23.89
LSSVM	0.7738	0.1208	0.0146	17.22
ANN	0.8423	0.1144	0.0131	16.20

Table 6. Statistical parameters of developed models and utilized correlation for prediction of condensate liquid viscosity.

3 Results and discussion

In this study two intelligent based models of LSSVM and ANN were developed to predict condensate liquid viscosity in depleted gas condensate reservoirs near wellbore region.

In first phase of this study prediction performance of LBC compositional model and 5 gas-saturated-oil empirical literature correlations were investigated for prediction of condensate viscosity.

The prediction performance of the compositional method of LBC, (1964) in predicting condensate liquid viscosity is very poor (Yang et al., 2007) and adjustment of LBC coefficients are usually necessary to match the experimental condensate viscosity (Fevang and Whitson, 1996; Whitson et al., 1999; Yang et al., 2007). The statistical analysis of the results shown in Fig. (1a) confirm the poor performance of compositional based LBC model. The reason for this is might due to the sensitivity of LBC method to mixture density and critical volumes of the heavy components. Hence, in this study the coefficients of the LBC correlation have tuned using least-square approach to match the experimental condensate viscosity data. Fig. (1a) representing the prediction performance of LBC, (1964) with default and adjusted coefficients.

The coefficients of five well-known gas-saturated-oil viscosity literature correlations regressed to match the condensate experimental data. The results of these regressions presented in Fig. (1b – f). These empirical correlations are function of dead oil viscosity and solution gas to oil ratio. It should be noted that dead oil viscosity

is one of the most “difficult” properties to be estimated by correlations due to its dependency to paraffin, aromatic, naphthalene and asphaltene content (Hemmati-Sarapardeh et al., 2014; Whitson et al., 2000). This might be one of the reasons for poor performance of the default empirical gas-saturated-oil viscosity correlations. Moreover, these correlations were originally developed using crude oil samples, which its properties are fundamentally different from condensate liquid.

Poor performance of the published literature correlations in predicting liquid dropout viscosity, motivated to develop two machine leaning models of LSSVM and ANN network in this study. The performance of the newly proposed models LSSVM and ANN were compared against refined previously published correlations through graphical and statistical error analysis. The statistical error analysis results carried out in terms of coefficient of determination (R^2), Root Mean Square Error (RMSE), Average Absolute Relative Deviation (AARD%) and Mean Square Error (MSE). The result of this error analysis is tabulated in the Table 6. Graphical representation of AARD% is also provided in Fig. (12). The results in Table 6 and Fig. (12) indicate ANN model outperforms other methods with AARD of 16.20%, R^2 of 0.8423, RMSE of 0.1144 and MSE of 0.0131. ANN followed by LSSVM, Kartoatmodjo and Schmidt (1994), Elsharkawy and Alikhan (1999), Bergman (2000), LBC (1964), De Ghetto et al. (1994) and Beggs and Robinson (1975).

The results show using either compositional model of LBC or gas-saturated-oil viscosity literature correlations require significant tuning of coefficients for viscosity prediction of condensate liquid. Whereas developed two intelligent approaches were able to monitor condensate liquid viscosity with appropriate precision and integrity.

Non-linear relationship between the available experimental data and the desired outputs created using developed LSSVM model. The optimum values of two important tuning parameters of LSSVM include $\sigma^2$and γ are presented in Table 2. Simulated Annealing optimization (SA) algorithm was applied to achieve these two optimum values.

The ability of proposed LSSVM and ANN models for calculating condensate liquid viscosity as a function of changing pressure has been investigated for three gas condensate samples from literature. Fig. (7) and Fig. (11) are demonstrating experimental and predicted condensate liquid viscosities using LSSVM and ANN models respectively. The results show that both models are able to forecast physical trend of experimental condensate viscosity. The accuracy of the models for predicting

condensate viscosity of independent samples determined by AARD%. The error analysis show that both models perform well with acceptable level of accuracy.

From Fig. (7) and Fig. (11) it is evident that increasing pressure decreases the condensate viscosity. The pressure changes due to depletion in gas condensate reservoirs can have significant effect on condensate viscosity variation near wellbore region (Fevang and Whitson, 1996). This changes can be due to the complex behaviour of gas condensate reservoir below the dew points, which violate thermodynamic laws. The developed LSSVM and ANN models successfully captured the trend of condensate viscosity while utilized correlations were not accurate enough in tracking these changes.

Although the prediction performance of the LSSVM was better than published literature correlations, however the error was still high with R^2 of 0.7738 and AARD of 17.22%. Therefore, Artificial Neural Network (ANN) method was used aiming for more accurate ML modelling approach. Performance prediction of ANN network is a function of number of neurons that is used in hidden layer (layer 2 in Fig. 8). A trial and error approach were implemented to find the optimum number of neurons. For this study the ANN architecture with five neurons provide the most satisfying results with least RMSE and the highest R^2.

4 Conclusion

Better modelling of condensate viscosity is very important for optimizing future recoveries, simulation studies, PVT calculations and accurate production performance forecast of gas condensate reservoirs. Current techniques in literature are providing poor prediction performance of condensate viscosity in near wellbore region. Hence in this study efforts have been made to model this liquid dropout viscosity using numerical artificial intelligence based methods including Least Square Support Vector Machine (LSSVM) and Artificial Neural Network (ANN). Both LSSVM and ANN models are capable of simulating the actual physical trend of the condensate viscosity in gas condensate reservoirs with variation of condensate API gravity, reservoir pressure, reservoir temperature, solution gas to oil ratio (Rs) and gas specific gravity. The advantage of LSSVM is that overfitting is not possible with this method. The robust simulated annealing optimizer implemented to find two important tuning parameters $\sigma^2$and γ and tune LSSVM method.

The results of this study indicated that proposed ANN and LSSVM are more robust, efficient and reliable than literature correlations. In ANN approach care should be taken to not over fit the data. This can be done by designing a network with appropriate level of complexity such as number of neurons and hidden layers.

Tuning the evolved LSSVM and ANN approach with other optimization method such as Genetic Algorithm (GA) or Coupled Simulated Annealing (CSA) to reduce the error can be considered for future studies.

Simplicity and flexibility of the developed model make them a good candidate to determine the viscosity of the condensate liquid in depleted gas condensate reservoirs. The developed models can be implemented in PVT calculation of gas condensate reservoirs for more accurate and reliable modelling of such reservoirs.

References

Ahmadi, M.A., Ebadi, M., 2014a. Fuzzy Modeling and Experimental Investigation of Minimum Miscible Pressure in Gas Injection Process. Fluid Phase Equilib. 378, 1–12. https://doi.org/10.1016/j.fluid.2014.06.022

Ahmadi, M.A., Ebadi, M., 2014b. Evolving smart approach for determination dew point pressure through condensate gas reservoirs. Fuel 117, 1074–1084. https://doi.org/10.1016/J.FUEL.2013.10.010

Ahmadi, M.A., Ebadi, M., Hosseini, S.M., 2014. Prediction breakthrough time of water coning in the fractured reservoirs by implementing low parameter support vector machine approach. Fuel 117, 579–589. https://doi.org/10.1016/j.fuel.2013.09.071

Ahmed, T.H., 2010. Reservoir engineering handbook, 4th ed. Gulf Professional Pub, Oxford.

Al-Meshari, A., Kokal, S., Al-Muhainy, A., Ali, M., 2007. Measurement of Gas Condensate, Near-Critical and Volatile Oil Densities and Viscosities at Reservoir Conditions, in: Proceedings of SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, California. https://doi.org/10.2523/108434-ms

Al-Nasser, K.S., Al-Marhoun, M.A., 2012. Development of New Gas Viscosity Correlations, in: SPE International Production and Operations Conference & Exhibition. Society of Petroleum Engineers, Doha. https://doi.org/10.2118/153239-ms

Ali, J.K., 1991. Evaluation of correlations for estimating the viscosities of hydrocarbon fluids. J. Pet. Sci. Eng. 5, 351–369. https://doi.org/10.1016/0920-4105(91)90053-P

Audonnet, F., Pádua, A.A.., 2004. Viscosity and density of mixtures of methane and n-decane from 298 to 393 K and up to 75 MPa. Fluid Phase Equilib. 216, 235–244. https://doi.org/10.1016/J.FLUID.2003.10.017

Beggs, H.D., Robinson, J.R., 1975. Estimating the Viscosity of Crude Oil Systems. J. Pet. Technol. 27, 1140–1141. https://doi.org/10.2118/5434-PA

Behera, S.S., Chattopadhyay, S., 2012. A Comparative Study of Back Propagation and Simulated Annealing Algorithms for Neural Net Classifier Optimization. Procedia Eng. 38, 448–455. https://doi.org/10.1016/J.PROENG.2012.06.055

622 Bergman, D.F., Sutton, R.P., 2007. An Update to Viscosity Correlations for Gas-
623 Saturated Crude Oils, in: SPE Annual Technical Conference and Exhibition.
624 Society of Petroleum Engineers, Anaheim. https://doi.org/10.2118/110195-MS
625 Chen, H.L., Wilson, S.D., Monger-McClure, T.G., 1995. Determination of Relative
626 Permeability and Recovery for North Sea Gas Condensate Reservoirs, in: SPE
627 Annual Technical Conference and Exhibition. Society of Petroleum Engineers,
628 Dallas. https://doi.org/10.2118/30769-MS
629 Cios, K.J., Shields, M.E., 1997. The handbook of brain theory and neural networks:
630 By Micheal A. Arbib (Ed.), MIT Press, Cambridge, MA, 1995, ISBN 0-262-
631 01148-4, 1118 pp. Neurocomputing 16, 259–261. https://doi.org/10.1016/S0925-
632 2312(97)00036-2
633 Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297.
634 https://doi.org/10.1007/BF00994018
635 Curilem, M., Acuña, G., … F.C.-C.E., 2011, undefined, n.d. Neural networks and
636 support vector machine models applied to energy consumption optimization in
637 semiautogeneous grinding. folk.ntnu.no.
638 De Ghetto, G., Paone, F., Villa, M., Spa, A., 1994. Reliability Analysis on PVT
639 Correlations, in: European Petroleum Conference. SPE, London.
640 Dreyfus, G., 2005. Neural Networks : Methodology and Applications, 1st ed.
641 Springer-Verlag Berlin Heidelberg, Paris.
642 ECLIPSE, 2014. Eclipse Reservoir Simulation Reference Manual.
643 El Aily, M., Mansour, E.M., Desouky, S.M., Helmi, M.E., 2019. Modeling viscosity of
644 moderate and light dead oils in the presence of complex aromatic structure. J.
645 Pet. Sci. Eng. 173, 426–433. https://doi.org/10.1016/J.PETROL.2018.10.024
646 Elsharkawy, A.M., Alikhan, A.A., 1999. Models for predicting the viscosity of Middle
647 East crude oils. Fuel 78, 891–903. https://doi.org/10.1016/S0016-
648 2361(99)00019-8
649 Elsharkawy, A.M., Foda, S.G., 1998. EOS simulation and GRNN modeling of the
650 constant volume depletion behavior of gas condensate reservoirs, in: SPE Asia
651 Pacific Conference on Integrated Modelling for Asset Management. Society of
652 Petroleum Engineers, Kuala Lumpur. https://doi.org/10.1021/ef970135z
653 Eslamimanesh, A., Gharagheizi, F., Illbeigi, M., Mohammadi, A.H., Fazlali, A.,
654 Richon, D., 2012. Phase equilibrium modeling of clathrate hydrates of methane,
655 carbon dioxide, nitrogen, and hydrogen+water soluble organic promoters using
656 Support Vector Machine algorithm. Fluid Phase Equilib. 316, 34–45.
657 https://doi.org/10.1016/j.fluid.2011.11.029
658 Eslamimanesh, A., Gharagheizi, F., Mohammadi, A.H., Richon, D., Illbeigi, M.,
659 Fazlali, A., Amir, ||, Forghani, A., Yazdizadeh, M., 2011. Phase Equilibrium
660 Modeling of Structure H Clathrate Hydrates of Methane + Water
661 "Insoluble" Hydrocarbon Promoter Using Group Contribution-Support
662 Vector Machine Technique. Ind. Eng. Chem. Res 50, 12807–12814.
663 https://doi.org/10.1021/ie2011164
664 Fazeli, H., Soleimani, R., Ahmadi, M.-A., Badrnezhad, R., Mohammadi, A.H., 2013.
665 Experimental Study and Modeling of Ultrafiltration of Refinery Effluents Using a
666 Hybrid Intelligent Approach. Energy & Fuels 27, 3523–3537.
667 https://doi.org/10.1021/ef400179b
668 Fetkovich, M.D., Guerrero, E.T., Of Tulsa, U., Fetkovich, M.J., Thomas, L.K., 1986.
669 SPE Oil and Gas Relative Permeabilities Determined From Rate-Time
670 Performance Data. Society of Petroleum Engineers, New Oreleans.
671 Fevang, Ø., 1995. Gas Condensate Flow Behavior and Sampling. October.

672 University of Trondheim.
673 Fevang, Ø., Whitson, C.H., 1996. Modeling Gas-Condensate Well Deliverability.
674 SPE Reserv. Eng. 11, 221–230. https://doi.org/10.2118/30714-PA
675 Ghaffari, A., Abdollahi, H., Khoshayand, M.R., Bozchalooi, I.S., Dadgar, A., Rafiee-
676 Tehrani, M., 2006. Performance comparison of neural network training
677 algorithms in modeling of bimodal drug delivery. Int. J. Pharm. 327, 126–138.
678 https://doi.org/10.1016/J.IJPHARM.2006.07.056
679 Ghahri, P., Jamiolahmady, M., Sohrabi, M., 2011. Gas Condensate Flow Around
680 Deviated And Horizontal Wells, in: SPE EUROPEC/EAGE Annual Conference
681 and Exhibition. Society of Petroleum Engineers, Vienna, pp. 2–23.
682 Gharagheizi, F., Ilani-Kashkouli, P., Sattari, M., Mohammadi, A.H., Ramjugernath,
683 D., Richon, D., 2014. Development of a LSSVM-GC model for estimating the
684 electrical conductivity of ionic liquids. Chem. Eng. Res. Des. 92, 66–79.
685 https://doi.org/10.1016/J.CHERD.2013.06.015
686 Ghiasi, M.M., Shahdi, A., Barati, P., Arabloo, M., 2014. Robust modeling approach
687 for estimation of compressibility factor in retrograde gas condensate systems.
688 Ind. Eng. Chem. Res. 53, 12872–12887. https://doi.org/10.1021/ie404269b
689 Gozalpour, F., Danesh, A., Todd, A.C., Tohidi, B., 2005. Viscosity, density, interfacial
690 tension and compositional data for near critical mixtures of methane + butane
691 and methane + decane systems at 310.95 K. Fluid Phase Equilib. 233, 144–
692 150. https://doi.org/10.1016/J.FLUID.2005.03.032
693 Guo, X., Wang, S., Rong, T., Guo, T., 1997. Viscosity model based on equations of
694 state for hydrocarbon liquids and gases. Fluid Phase Equilib. 139, 405–421.
695 https://doi.org/10.1016/S0378-3812(97)00156-8
696 Hagan, M.T., Demuth, H.B., Beale, M.H., De Jes s, O., 2014. Neural network
697 design, 2nd ed. Frisco.
698 Hagan, M.T., Menhaj, M.B., 1994. Training feedforward networks with the Marquardt
699 algorithm. IEEE Trans. Neural Networks 5, 989–993.
700 https://doi.org/10.1109/72.329697
701 Haykin, S.S., 1994. Neural networks : a comprehensive foundation, 1st ed.
702 Macmillan, New York.
703 Hemmati-Sarapardeh, A., Shokrollahi, A., Tatar, A., Gharagheizi, F., Mohammadi,
704 A.H., Naseri, A., 2014. Reservoir oil viscosity determination using a rigorous
705 approach. Fuel 116, 39–48. https://doi.org/10.1016/J.FUEL.2013.07.072
706 Hernandez;, J.C., Vesovic, V., Carter, J.N., Lopez, E., 2002. Sensitivity of Reservoir
707 Simulations to Uncertainties in Viscosity, in: SPE/DOE Improved Oil Recovery
708 Symposium. Society of Petroleum Engineers, Oklahoma, pp. 2–10.
709 https://doi.org/10.2118/75227-ms
710 Hippert, H.S., Pedreira, C.E., Souza, R.C., 2001. Neural networks for short-term load
711 forecasting : a review and evaluation. IEEE Trans. Power Syst. 16, 4333.
712 Jalali, farhang, Abdy, Y., Akbari, M., 2007. Dewpoint Pressure Estimation of Gas
713 Condensate Reservoirs, Using Artificial Neural Network (ANN), in: Proceedings
714 of EUROPEC/EAGE Conference and Exhibition. Society of Petroleum
715 Engineers, London. https://doi.org/10.2523/107032-MS
716 Jokhio, S.A., Tiab, D., Escobar, F., 2002. Forecasting Liquid Condensate and Water
717 Production In Two-Phase And Three-Phase Gas Condensate Systems. Society
718 of Petroleum Engineers, San Antonio, pp. 1–13. https://doi.org/10.2118/77549-
719 ms
720 Jossi, J.A., Stiel, L.I., Thodos, G., 1962. The viscosity of pure substances in the
721 dense gaseous and liquid phases. AIChE J. 8, 59–63.

https://doi.org/10.1002/aic.690080116
Kamari, A., Hemmati-Sarapardeh, A., Mirabbasi, S.-M., Nikookar, M., Mohammadi, A.H., 2013. Prediction of sour gas compressibility factor using an intelligent approach. Fuel Process. Technol. 116, 209–216. https://doi.org/10.1016/J.FUPROC.2013.06.004
Kartoatmodjo, T.R.S., Schmidt, Z., 1991. New Correlations For Crude Oil Physical Properties.
Kashefi, K., Chapoy, A., Bell, K., Tohidi, B., 2013. Viscosity of binary and multicomponent hydrocarbon fluids at high pressure and high temperature conditions: Measurements and predictions. J. Pet. Sci. Eng. 112, 153–160. https://doi.org/10.1016/J.PETROL.2013.10.021
Kay, W.B., 1936. Density of Hydrocarbon. Ind. Eng. Chem. 28, 1014–1019. https://doi.org/10.1021/ie50321a008
Kenneth, L., 1944. A Method for the Solution of Certain Non-Linear Problem in Least Squares. https://doi.org/10.1090/qam/10666
Khan, S.A., Al-Marhoun, M.A., Duffuaa, S.O., Abu-Khamsin, S.A., 1987. Viscosity Correlations for Saudi Arabian Crude Oils, in: Middle East Oil Show. Society of Petroleum Engineers, Bahrain. https://doi.org/10.2118/15720-ms
Khosrojerdi, S., Vakili, M., Yahyaei, M., Kalhor, K., 2016. Thermal conductivity modeling of graphene nanoplatelets/deionized water nanofluid by MLP neural network and theoretical modeling using experimental results. Int. Commun. Heat Mass Transf. 74, 11–17. https://doi.org/10.1016/J.ICHEATMASSTRANSFER.2016.03.010
Lohrenz, J., Bray, B.G., Clark, C.R., 1964. Calculating Viscosities of Reservoir Fluids From Their Compositions. J. Pet. Technol. 16, 1171–1176. https://doi.org/10.2118/915-PA
Mansour, E.M., Farag, A.B., El-Dars, F.S., Desouky, S.M., Batanoni, M.H., Mahmoud, M.R.M., 2013. Predicting PVT properties of Egyptian crude oils by a modified Soave–Redlich–Kowng equation of state. Egypt. J. Pet. 22, 137–148. https://doi.org/10.1016/J.EJPE.2012.09.005
Mesbah, M., Soroush, E., Rostampour Kakroudi, M., 2017. Predicting physical properties (viscosity, density, and refractive index) of ternary systems containing 1-octyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)imide, esters and alcohols at 298.15 K and atmospheric pressure, using rigorous classification techniques. J. Mol. Liq. 225, 778–787. https://doi.org/10.1016/J.MOLLIQ.2016.11.004
Mott, R., 2003. Engineering Calculations of Gas-Condensate-Well Productivity. SPE Reserv. Eval. Eng. 6, 298–306. https://doi.org/10.2118/86298-PA
Naderi, M., Khamehchi, E., 2019. Fuzzy logic coupled with exhaustive search algorithm for forecasting of petroleum economic parameters. J. Pet. Sci. Eng. 176, 291–298. https://doi.org/10.1016/J.PETROL.2019.01.049
Nowroozi, S., Ranjbar, M., Hashemipour, H., Schaffie, M., 2009. Development of a neural fuzzy system for advanced prediction of dew point pressure in gas condensate reservoirs. Fuel Process. Technol. 90, 452–457. https://doi.org/10.1016/j.fuproc.2008.11.009
O'Dell, H., Miller, R., 1967. Successfully Cycling a Low-Permeability, High-Yield Gas Condensate Reservoir. J. Pet. Technol. 19, 41–47. https://doi.org/10.2118/1495-PA
Pelckmans, K., Suykens, J.A.K., Van Gestel, T., De Brabanter, J., Lukas, L., Hamers, B., De Moor, B., Vandewalle, J., 2002. LS-SVMlab: a MATLAB/C

772 toolbox for Least Squares Support Vector Machines, Tutorial. KULeuven- ESAT.
773 Leuven-Heverlee.
774 Saeedi, J., Rowe, A.M., 1981. Viscosity Correlations for Compositional Reservoir
775 Simulators (SPE9643), in: Middle East Oil Technical Conference of the Society
776 of Petroleum Engineers. Society of Petroleum Engineers, Bahrain, pp. 645–649.
777 https://doi.org/10.2118/9643-MS
778 Shokir, E.M., 2008. Novel Density and Viscosity Correlations for Gases and Gas
779 Mixtures Containing Hydrocarbon and Non-Hydrocarbon Components. J. Can.
780 Pet. Technol. 47. https://doi.org/10.2118/08-10-45
781 Soroush, E., Mesbah, M., Shokrollahi, A., Rozyn, J., Lee, M., Kashiwao, T.,
782 Bahadori, A., 2015. Evolving a robust modeling tool for prediction of natural gas
783 hydrate formation conditions. J. Unconv. Oil Gas Resour. 12, 45–55.
784 https://doi.org/10.1016/J.JUOGR.2015.09.002
785 Stiel, L.I., Thodos, G., 1962. The viscosity of polar gases at normal pressures.
786 AIChE J. 8, 229–232. https://doi.org/10.1002/aic.690080220
787 Suykens, J.A.K., De Brabanter, J., Lukas, L., Vandewalle, J., 2002. Weighted least
788 squares support vector machines: robustness and sparse approximation.
789 Neurocomputing 48, 85–105. https://doi.org/10.1016/S0925-2312(01)00644-0
790 Suykens, J.A.K., Vandewalle, J., 1999. Least Squares Support Vector Machine
791 Classifiers. Neural Process. Lett. 9, 293–300. https://doi.org/10.1023/A
792 Thomas, F.B., Bennion, D.B., 2009. Gas Condensate Reservoir Performance. J.
793 Can. Pet. Technol. 10.
794 Wheaton, R.J., Zhang, H.R., 2007. Condensate Banking Dynamics in Gas
795 Condensate Fields: Compositional Changes and Condensate Accumulation
796 Around Production Wells. https://doi.org/10.2118/62930-ms
797 Whitson, C., W, J., Brulé, M., 2000. Phase Behavior, 1st ed, Society. Society of
798 Petroleum Engineers. https://doi.org/10.1021/ma00080a014
799 Whitson, C.H., Fevang, Ø., Yang, T., 1999. Gas Condensate PVT – What's Really
800 Important and Why?, in: Optimisation of Gas Condensate Fields. Norwegian U.
801 of Science and Technology, London. https://doi.org/10.2118/117930-PA
802 Yang, T., Fevang, O., Christoffersen, K., Ivarrud, E., 2007. LBC Viscosity Modeling
803 of Gas Condensate to Heavy Oil, in: SPE Annual Technical Conference and
804 Exhibition. Society of Petroleum Engineers, Anaheim.
805 https://doi.org/10.2523/109892-ms
806 Zendehboudi, S., Ali Ahmadi, M., James, L., Chatzis, I., 2012. Prediction of
807 Condensate-to-Gas Ratio for Retrograde Gas Condensate Reservoirs Using
808 Artificial Neural Network with Particle Swarm Optimization. Energy & Fuels 26,
809 3432–3447. https://doi.org/10.1021/ef300443j
810

811 # Appendix A
812

Researcher	Number of data points	Fluid sample	Reported error	Advantages and applicability	Disadvantages
Lohrenz-Bary-Clark (1964)	520 data points used to develop oil viscosity and 300 data samples used to develop dense gas viscosity.	Black to highly volatile oil samples. High pressure gas mixture.	16% of average error for oil and 4% of average error for gases.	• Can be used to determine both gas and hydrocarbon liquid viscosity. • The LBC correlation uses reservoir fluid composition to determine the fluid viscosity. • Most widely used correlation due to its simplicity and flexibility. • Take account of compositional changes in reservoirs fluids.	• Very sensitive to mixture density and critical volume of heavy components. • Prediction performance of the LBC is poor for oil viscosity. • The tuning of coefficients is usually required to match the experimental data. • The tuning procedure is not straight forward especially for gas condensate fluids. • Heavy tuning of LBC coefficients can cause non-monotonic relations between viscosity and reduced density.
Bergman (2000)	2048 data points from worldwide used to develop gas-saturated-oil viscosity.	Crude oil	9% absolute average error and 11.58% standard deviation.	• Ability to predict the wide range of crude oil viscosity 0.125 – 123cp. • Simple and flexible to use. • One of the most accurate method over wide range of conditions (Bergman and Sutton, 2007).	• Limited range of solution gas to oil ratio 5 – 2890scf/STB. • Applicable to crude oil and need tuning for other type of hydrocarbon liquids such as condensate liquid. • Inaccurate dead oil calculation can reduce the accuracy.

Elsharkawy and Alikhan (1999)	254 crude oil samples from Middles East	Crude oil	Average relative error of 2.8% an average absolute error of 18.6%.	• Ability to predict the gas-saturated-oil viscosity in lower range 0.05 – 20.89cp. • Less input parameters in computation process (API, reservoir pressure and reservoir temperature).	• Limited applicability to specific geographical region. • Function of dead oil viscosity, which reduce the accuracy. • Require accurate solution gas to oil ratio.
Beggs and Robinson (1975)	2073 data points used in development of correlation	Crude oil	Average error of -1.83% and standard deviation of 27.25.	• Covers good range of solution gas to oil ratio (Rs) of 20 – 2070scf/STB. • Widely used in industry. • Simple calculation procedure.	• Unknown applicability to the specific region. • Unknown ability of predicting different viscosity ranges.
Kartoatmodjo & Schmidt (1991)	5321 crude oil data from Indonesia, America, Middles East & Latin America	Crude oil	Absolute error of 0.08% and 16.08% absolute average deviation	• Comprehensive data bank has been used in developing the correlation. • Covers a wide range of viscosity between 0.096 – 586cp.	• Cannot accurately predict the viscosity at low gas-oil ratio when reservoir pressure becomes atmospheric.

De Ghetto et al. (1994)	195 oil samples from Mediterranean Basin, Africa, Persian Gulf, North America (3700 data points)	Light crude oil API >31.1	Absolute error of 15.2% and standard deviation of 14.8% for oil with API > 31.1	• Developed for light crude oil with API >31.1, which its properties are close to condensate fluid. • Able to predict the gas-saturated-oil viscosity within the range of 0.1 – 120cp. • Simple calculation procedure.	• Function of dead oil viscosity, which is hard to predict accurately. • Poor performance for predicting of gas condensate reservoirs.

Table A1. Description of utilized empirical gas-saturated-oil correlations.

Appendix B

This section covers mathematical hypothesis of simple neural network architecture shown in Fig. B1, where superscripts are values associated with each layer.

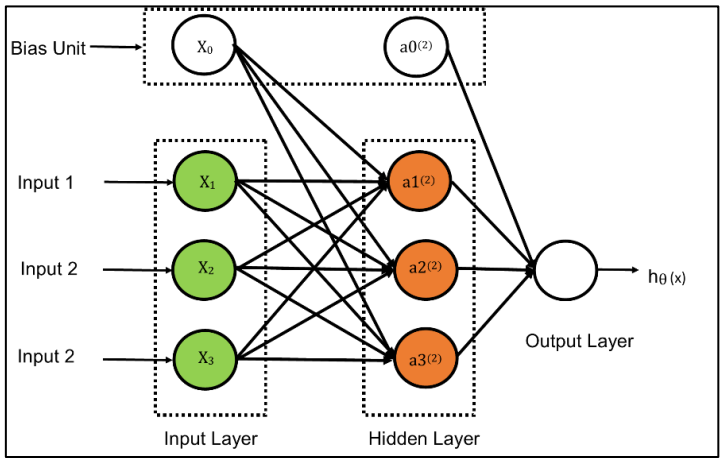

Fig. B1. Schematic illustration of the ANN structure and computational steps to measure any output.

In graph shown in Fig. B1:

$ai^{(j)} = \text{activation" } of\ unit\ i\ in\ layer\ j$

$\Theta^{(j)} = \text{matrix of weights controlling function mapping from layer j to layer j+1}$

In order to calculate each activation function (a) a sigmoid function (g) is multiplied by sum of linear combination of inputs for each neuron; these inputs include ($x_1, x_2, x_3\ and\ bias\ unit\ x_0$) in hidden layer. Eq. (B1) to Eq. (B3) are representing the calculation of the activation functions.

Then the output function $h_\theta(x)$ shown in Eq. (B4) is a sigmoid function of sum of each neuron's weight multiplied by activation function of same neuron in layer 2. The neurons of the output layer have linear transfer functions.

$$a_1^{(2)} = g\left(\Theta_{10}^{(1)} x_0 + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + \Theta_{13}^{(1)} x_3\right) \quad \text{(B1)}$$

$$a_2^{(2)} = g\left(\Theta_{20}^{(1)} x_0 + \Theta_{21}^{(1)} x_1 + \Theta_{22}^{(1)} x_2 + \Theta_{23}^{(1)} x_3\right) \quad \text{(B2)}$$

$$a_3^{(2)} = g\left(\Theta_{30}^{(1)} x_0 + \Theta_{31}^{(1)} x_1 + \Theta_{32}^{(1)} x_2 + \Theta_{33}^{(1)} x_3\right) \quad \text{(B3)}$$

$$h_\Theta(x) = a_1^{(3)} = g\left(\Theta_{10}^{(2)} a_0^{(2)} + \Theta_{11}^{(2)} a_1^{(2)} + \Theta_{12}^{(2)} a_2^{(2)} + \Theta_{13}^{(2)} a_3^{(2)}\right) \quad \text{(B4)}$$

In above equation g is a sigmoid type function and can be evaluated from Eq. (B5).

$$g(z) = \frac{1}{(1+e^{-z})} \quad \text{(B5)}$$

To vectorise the above mathematical definition of neural network presented in Equation (B1) to (B3), the following relations can be defined:

If:

$$\begin{Bmatrix} \Theta_{10}^{(1)} x_0 + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + \Theta_{13}^{(1)} x_3 = Z_1^{(2)} \\ \Theta_{20}^{(1)} x_0 + \Theta_{21}^{(1)} x_1 + \Theta_{22}^{(1)} x_2 + \Theta_{23}^{(1)} x_3 = Z_2^{(2)} \\ \Theta_{30}^{(1)} x_0 + \Theta_{31}^{(1)} x_1 + \Theta_{32}^{(1)} x_2 + \Theta_{33}^{(1)} x_3 = Z_3^{(2)} \end{Bmatrix} \quad \text{(B6)}$$

Substituting Equation (B6) into Eq. (B1) to Eq. (B3) defines the activation functions in Equation (B7).

$$\begin{Bmatrix} a_1^{(2)} = g\left(Z_1^{(2)}\right) \\ a_2^{(2)} = g\left(Z_2^{(2)}\right) \\ a_3^{(2)} = g\left(Z_3^{(2)}\right) \end{Bmatrix} \quad \text{(B7)}$$

And If:

$$\begin{Bmatrix} x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ Z^{(2)} = \begin{bmatrix} Z_1^{(2)} \\ Z_2^{(2)} \\ Z_3^{(2)} \end{bmatrix} \end{Bmatrix} \quad \text{(B8)}$$

And then input functions substitute with $a^{(1)}$ in layer one:

$$\begin{Bmatrix} Z^{(2)} = \Theta^{(1)} x = \Theta^{(1)} a^{(1)} \\ a^{(2)} = g\left(Z^{(2)}\right) \end{Bmatrix} \quad \text{(B9)}$$

In equation B9, $a^{(2)}$ is [3x3] matrix without bias function, and if $a_0^2 = 1$ for bias unit in layer 2, $Z^{(3)}$ defined as follow:

$$Z^{(3)} = \Theta^{(2)} a^{(2)} \quad \text{(B10)}$$

The value of the final function or output layer is sigmoid function of $Z^{(3)}$, as shown in Eq. (B11).

$$h_\Theta(x) = g(Z^{(3)}) \quad \text{(B11)}$$

The values of *x* are considered as input of activation function. The above calculation was carried out and completed in MATLAB, to determine the output values.