Journal Pre-proofs

Original article

Development of new gas condensate viscosity model using artificial intelligence

F. Faraji, J.O. Ugwu, P.L. Chong

PII: S1018-3639(21)00165-3
DOI: https://doi.org/10.1016/j.jksues.2021.11.005
Reference: JKSUES 581

To appear in: *Journal of King Saud University - Engineering Sciences*

Received Date: 5 July 2021
Revised Date: 26 October 2021
Accepted Date: 10 November 2021

Please cite this article as: Faraji, F., Ugwu, J.O., Chong, P.L., Development of new gas condensate viscosity model using artificial intelligence, *Journal of King Saud University - Engineering Sciences* (2021), doi: https://doi.org/10.1016/j.jksues.2021.11.005

Development of new gas condensate viscosity model using artificial intelligence.

F.Faraji[1]*, J.O.Ugwu[1], P.L. Chong[1]
School of Computing, Engineering and Digital Technology, Teesside University, Middlesbrough, UK
*F.Faraji@tees.ac.uk (corresponding author); J.Ugwu@tees.ac.uk; P.Chong@tees.ac.uk

Development of new gas condensate viscosity model using artificial intelligence.

Abstract
Accurate estimation of gas condensate fluid properties is a challenging task due to evolving condensate liquid from the gas phase below the saturation pressure. Among the fluid properties viscosity of condensate liquid has the largest prediction uncertainty. The existing literature methods cannot cope with non-linearity and physics of gas condensate mixture (transition from single phase to two-phase) below the saturation pressure. Hence, in this study based on the experimental condensate viscosity data a simple linear equation as a function of pressure, temperature and solution gas to oil ratio was developed. For this purpose, comprehensive data source of 1368 experimental data points acquired from open literature has been used. For developing new condensate viscosity correlation an Artificial Intelligence method known as Takagi – Sugeno – Kang fuzzy algorithm was utilized. The accuracy of the developed correlation was compared with five previously published literature models. The superiority of new correlation over existing literature models is confirmed by statistical parameters of least root mean square error of 0.0194, mean average error of 0.0163 and average absolute relative deviation percentage of 7.123. The proposed condensate viscosity correlation is valid in a pressure rang of 0.25 – 75.84 MPa), temperature range of 303 – 443.15°K and Rs of 41.96 – 13496 scf/STB.
The proposed correlation can be used as an alternative approach to existing models for accurate estimation of gas condensate viscosity and produce reliable reservoir simulation studies.

Keywords
Artificial Intelligence (AI), Condensate viscosity, Fuzzy approach, K–mean clustering, Membership Functions (MFs), Takagi – Sugeno – Kang (TSK).

1. Introduction

Gas condensate reservoirs with depletion-mode of recovery are experiencing liquid drop out (liquid condensate) near the wellbore region when bottom-hole flowing pressure (BHFP) falls below the saturation pressure. This behaviour would directly affect the performance of the reservoir and in some cases lead to severe productivity decline (Strand and Bjørkvik, 2019). The viscosity of liquid condensate in such reservoir is an important parameter for evaluating performance, reservoir simulation, designing facilities and selecting the best production strategy (Dandekar, 2015). Hence, accurate prediction of this parameter is important for natural gas industry. Ideally, the viscosity should be measured experimentally by providing fluid samples using Constant Volume Depletion or Constant Composition Expansion (CCE) experiments. Nevertheless, experimental studies are limited in the industry because of the associated challenges, fluid behaviour of gas condensate mixture below the dew point, time and costs (Fattah and Lashin, 2018). When the fluid samples are not available to save cost and time, fluid properties (viscosity) are estimated by empirical semi-empirical correlations. The correlations are divided into two categories including those that developed based on corresponding state principle (CSP) and flow theory and gas-saturated oil or live oil viscosity models. The CSP correlations use fluid composition, pour point temperature, acentric factor, critical temperature, boiling point and molar mass for estimation of fluid (gas and liquid) viscosity (Baled et al., 2018). The two most prevalent CSP based viscosity models in any Pressure Volume Temperature (PVT) models (e.g., Eclipse "*Schlumberger*" and Landmark's Nexus "*Halliburton*") are Lohrenz et al., (1964) known as LBC and Pedersen and Fredenslund, (1987). These semi-empirical correlations are a strong function of reduced density and compositional variation and small changes in both aforementioned parameters may have a profound effect on their performance (Yang et al., 2007). Availability of the mixture composition data, which is a common problem as mentioned earlier is the key for using CSP viscosity models. In this case, the second category of correlations known as gas-saturated oil or live oil viscosity models can be used for the estimation of condensate viscosity (Yang et al., 2007). These correlations are simpler than CSP based models and mostly developed as a function of reservoir temperature, pressure, API gravity and solution gas to oil ratio "Rs". Some examples of these correlations are Beggs and Robinson, (1975); Bergman and Sutton, (2007); De Ghetto et al., (1994); Elsharkawy and Alikhan, (1999); Kartoatmodjo and Schmidt, (1991); Elsharkawy, (2006); Sutton, (2005); Ugwu et al., (2011). However, the accuracy of these models is questionable as some correlations are specific for a certain region and fail to predict viscosity in a wide range of operational conditions. It has been highlighted that the best performance of the prevalent existing viscosity models for prediction of liquid condensate is within 10 – 30% error and often increases to 50% (Whitson, 2006).

Another approach that recently become very popular for various modelling aspects of hydrocarbon reservoirs fluid properties are so-called Artificial Intelligence (AI) or Machine Learning (ML) techniques. Many scholars utilized AI for the accurate modelling of various reservoir fluid properties. Najafi-Marghmaleki et al., (2016) utilized an AI model known as a radial basis function to predict the dew point pressure. Zendehboudi et al., (2012) utilized

Artificial Neural Network (ANN) for the prediction of condensate to gas ratio (CGR). Al-Quraishi and Shokir, (2011) and Hemmati-Sarapardeh et al., (2014) used the ANN network and Least Square Support Vector Machine (LSSVM) to model hydrocarbon viscosity of gas and oil. Recently Faraji et al., (2020) implemented ANN and LSSVM for accurate estimation of liquid condensate viscosity in gas condensate reservoirs. Although the aforementioned approaches provide a satisfactory estimation of the studied properties, however, the major criticism is that they are black-box approaches and the visual relationship between input and output parameters cannot be interpreted (Bikmukhametov and Jäschke, 2020; Hemmati-Sarapardeh et al., 2020). In another word, a meaningful mathematical function from these methods cannot be derived. Therefore, this study aims to develop a novel condensate liquid viscosity correlation. The proposed correlation should provide scientific insight into fluid properties and cope with the non-linearity of gas condensate fluid mixture concerning liquid condensate viscosity without sacrificing accuracy.

For this purpose 1368 data points from open literature were collected and used for developing and testing the proposed model. Based on the experimental condensate viscosity data a simple linear equation as a function of pressure (P), temperature (T) and solution gas to oil ratio (Rs) $\mu_c = f(P, T, Rs)$ is proposed. Takagi – Sugeno – Kang (TSK) rule-based fuzzy approach is used for optimizing associate parameters of the proposed model.

2. Data collection

To assess the accuracy of the studied literature viscosity models and also develop a novel condensate viscosity correlation a data bank is gathered from experimental studies in the open literature. The sources of the data bank are (Al-Meshari et al., 2007; Audonnet and Pádua, 2004; Fevang, 1995; Gozalpour et al., 2005; Guo et al., 1997; Kariznovi et al., 2012; Kashefi et al., 2013; Khorami et al., 2017; Strand and Bjørkvik, 2019; Thomas et al., 2009; Yang et al., 2007). In the aforementioned studies, various methods include rolling ball viscometer, electromagnetic pulse technology viscometer and capillary viscometer and vibrating-wire sensor have been used for measurement of condensate phase viscosity. A binary mixture of methane and n-decane in various temperatures and pressure was considered as gas condensate fluid in most of those experimental studies. The data bank contains 1368 data points, which have been organized in matrix form. The databank is divided into two subsets of training (70%) and testing (30%). The statistical distribution of the data is summarized in Table 1.

Table 1. Statistical information of the databank.

Property	Min	Max	Average
Pressure, (MPa)	0.25	75.84	25.25

Reservoir Temperature, (°K)	303	443.15	353.15
Solution GOR, Rs, (scf/STB)	41.7	13496	3628
Condensate viscosity, μ_c, (cp)	0.0404	0.982	0.232

3. Development of new correlation using TSK fuzzy

3.1 TSK Fuzzy approach

The basic idea of fuzzy sets was initially proposed by Zadeh, (1965) to deal with vagueness, imprecision and uncertainty in a system. Unlike the classical binary logic that only admits true (1) or false (0) of an occurrence, fuzzy logic covers the degree of truth of a factor between 0 and 1. In contrast to the classical crisp set, where an object either belongs to a set or it does not, everything is a matter of degrees in a fuzzy set. The most well-known rule-based fuzzy inference systems (FIS) are linguistic Mamdani type and Takagi-Sugeno-Kang (TSK) (Mamdani and Assilian, 1975; Takagi and Sugeno, 1985). Both antecedents and consequence are fuzzy sets in the Mamdani approach, whereas in the TSK model antecedent contains fuzzy sets, and the consequence is a linear equation. TSK model is more compact and offers computational efficiency than Mamdani type fuzzy inference system (AL-Rousan et al., 2020). In TSK fuzzy model a linear relationship between input/output is defined by a set of fuzzy rules shown in the following equation.

$$R_1: if\ x_i\ is\ A_{i1} and \ldots and\ x_m\ is\ A_{im}\ \ then \\ y_i = a_{i1}x_1 + \cdots + a_{im}x_m + a_{i0} \tag{1}$$

Where $R_1 = (1, 2, \ldots, n)$ is the number of fuzzy rules, $x_i = (1,2, \ldots, m)$ are the input variables, y_i are the output variables whose values are inferred, $A_{i1}, \ldots, A_{im}$ are membership functions of the fuzzy sets in the premises and $a_{i0}, a_{i1}, \ldots, a_{im}$ are the model parameters in the consequence (Takagi and Sugeno, 1985). To determine aforesaid three items using input-output data of a respective system the design procedure of the TSK fuzzy model can be summarized in three steps as follow: 1) Fuzzy clustering; 2) Setting the membership functions and 3) Parameters estimation (Choudhury et al., 2020; Shokir, 2008).

Partitioning a set of input variables into some fuzzy subsets can be carried out in the first two steps. The relation of input/output of each fuzzy subset is defined in the last step (Takagi and Sugeno, 1985). TSK algorithm requires a combination of input/output variables called premise variables, where in this study the input variables are pressure (P), temperature (T) and solution gas to oil ratio (Rs), and the output variables are liquid condensate viscosities (μ_c). After preparing the dataset to two subsets of training (70%) and Testing (30%), the partitioning of the data to several clusters is required.

Based on inputs/outputs source a data sets N=1, 2, …, n have been organized in a matrix form as shown in the following equation.

$$N = \begin{bmatrix} P_1 & T_1 & Rs_1 & \mu_{c1} \\ P_2 & T_2 & Rs_2 & \mu_{c2} \\ \vdots & \vdots & \vdots & \vdots \\ P_N & T_N & Rs_N & \mu_{cN} \end{bmatrix} \tag{2}$$

Before applying any clustering techniques (e.g., K-mean clustering), to the created matrix data structure the optimum number of clusters needs to be defined in a data set. This would also determine the number of fuzzy IF-THEN rules. In this study Calinski and Harabasz, (1974) cluster evaluation method was used to find the optimum number of clusters.

3.2 Fuzzy clustering

There are several fuzzy clustering methods in literature such as fuzzy c-means (FCM), Gustafson-Kessel (GK), K-means clustering and subtractive clustering. In this study, the K-means clustering method was used as one of the most popular classification algorithms for the data without any defined categories (unlabelled data). This algorithm, is an iterative, hill climbing data-partitioning algorithm, where N observations can be partitioned to "c" clusters where an objective function "J" can be estimated as follow (Klawonn et al., 2015).

$$J = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij} d_{ij} \tag{3}$$

where J should be minimized under the following constraints:

$$\sum_{i=1}^{c} u_{ij} = 1 \quad for\ all\ \ j \in \{1, \ldots, n\} \tag{4}$$

Where $u_{ij} \in \{0,1\}$ indicates whether data vector x_j is assigned to cluster i $(u_{ij} = 1)$ or not $(u_{ij} = 0)$; $d_{ij} = \|x_j - v_i\|^2$ is squared Euclidean distance between data vector x_j and cluster prototype v_i. The number of clusters "c" in this method must be known in advance. Our criteria for assigning an initial number of clusters is based on the assumption that nonlinearity in the data can be approximated by 12 clusters (Shokir, 2008).

In general, there is no specific rule for defining the optimum number of clusters "c", however, several techniques such as the elbow method, the silhouette method G-means algorithm and Calinski-Harabasz exist in literature. In this study, the Calinski-Harabasz cluster evaluation method was used as an efficient technique (Calinski and Harabasz, 1974). The criteria of Calinski-Harabasz also called variance ratio criterion (VRC), which is defined in the following.

$$VRC_c = \frac{SS_B}{SS_W} \times \frac{(N-c)}{(c-1)} \tag{5}$$

Where SS_B and SS_w are between and within overall cluster variance respectively and defined in Eq. (6) and c is the number of clusters, and N represents the data points.

$$
\begin{cases} SS_B = \sum_{i=1}^{c} n_i \|m_i - m\|^2 \\ SS_W = \sum_{i=1}^{c} \sum_{x \in c_j} \|x - m_i\|^2 \end{cases} \tag{6}
$$

Where n_i is the number of observations in cluster i, m_i is the centroid of cluster i, m stands for the mean of the data, x is the number of data samples, c_j is the ith cluster and $\|m_i - m\|^2$, $\|x - m_i\|^2$ is Euclidean distances between two vectors. Large SS_B and a smaller SS_W are representing well-grouped clusters, which means the larger the VRC_c ratio, the better partitioning of the data (Calinski and Harabasz, 1974). Therefore, to achieve the optimum number of clusters, the validity measure of VRC_c is maximized for the number of clusters c. Hence, the highest Calinski-Harabasz index is the optimum number of clusters. In this study using our training data the best number of clusters returned by validity faction, VRC_c is 9 as shown in Fig.1.

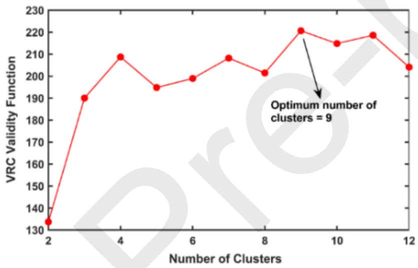


Fig. 1. The results of validity function VRC_c, for liquid condensate viscosity input data.

Having defined the optimum number of clusters for our training data, the k-means algorithm, presented in Eq. (11) can proceed in the following three steps for a training data point, $A = \{x_1, x_2 \ldots, x_n\}$ in n-dimensional space $\mathbb{R}^n$: Step 1: Choose an initial cluster centres $z_1, z_2, z_3, \ldots, z_K$ randomly from n points $\{x_1, x_2 \ldots, x_n\}$; Step 2: assign data points $a = A$ to its closest centre and obtain k-partition of A; Step 3: Recalculate centres for the new partition and go to step 2 until no more data, change their clusters, or the algorithm is converged.

3.3 Setting the membership function

To determine the membership degree of an object (data point) to a certain set (cluster), $A_{i1}, \ldots, A_{im}$ the membership functions (MFs) have to be set between 0 and 1 (Nazari et al., 2015). This is a binary issue, which states an object is either belong to a set or not and can be represented by the following expression.

$$
\mu_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases} \tag{7}
$$

where $\mu_A(x)$ represents an ambiguous membership of component *x* in set *A*, and $\in$, $\notin$ represent contained or not contained in set *A*, respectively. Zadeh, (1965) extended classical binary membership of only 0 and 1 to a real continuous interval where the numbers between 0 and 1 can represent various degree of a membership of an object (data point) to a set (cluster) as follow.

$$\mu_A: U \to [0,1] \tag{8}$$

Where U represents a universal set defined for a specific problem in fuzzy set A. For instance, if $U = \{x_1, x_2, \ldots, x_n\}$, then the degree of membership of $x_1, x_2, \ldots, x_n$ in U can be defined by the following equation.

$$A = \{(\mu_A(x_1)|x_1), (\mu_A(x_2)|x_2), \ldots, (\mu_A(x_n)|x_n)\} \tag{9}$$

The relation between the input/output is defined by fuzzy IF-THEN rules, where a conclusion can be achieved based on the hypothesis. This explains the principle of an inference mechanism which, , if a hypothesis is known then another fact or conclusion can be reached (Shokir, 2008).

The information that how the data points are distributed in the input space provides the guideline for creating several fuzzy clusters and their detection. Cluster centres and eigenvalues of fuzzy covariant matrices can be used for capturing this information (Shokir, 2008, 2006; Takagi and Sugeno, 1985). In this study, the gaussian membership function is employed to define the antecedent fuzzy sets as follow (Nazari et al., 2015).

$$\mu_{Ai}(v_i) = exp\left(-\frac{(v_i - c_i)^2}{2\sigma_i^2}\right) \tag{10}$$

Where v_{ij} is the scalar values of inputs, σ_i is the standard deviation and c_i is the mean of the ith fuzzy set Ai. Fig. 2 illustrates the output response of gaussian membership functions to one of the inputs (pressure). Other parameters of the training data set including temperature, solution gas to oil ratio and liquid condensate viscosity were similarly given the degree of membership function. MF1 to MF9 in Fig. 2 represent 9 membership functions that were constructed according to 9 identified clusters using Eq. (5).

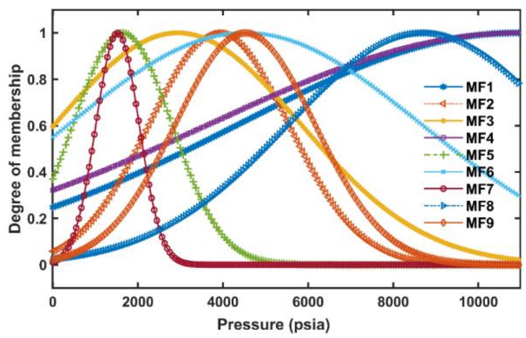


Fig. 2. Gaussian membership function for input 1 “pressure”.

To optimize the consequent parameters of the TSK model $a_{i0}, a_{i1}, \dots, a_{im}$ in Eq. (1) are using least-square approximation if X denotes a matrix having ith row in the input vector known as xi (inputs) and if Y represents a vector column with y_i (output) as its ith component. Also if w_i denotes to $N \times N$ real matrix the degree of firing β_{ij} can be defined as follow (Takagi and Sugeno, 1985).

$$\beta_{ij} = \frac{\beta_i(x_i)}{\sum_{z=1}^{c} \beta_i(x_j)} \tag{11}$$

And if $\theta_i = [a_{i1}, \dots, a_{im}, a_{i0}]$ represents consequent parameters of ith rule in each vector, to determine a_{i0} in θ_i, a unitary column I is added to the matrix X, $X_e = [X, I]$. This is an extended matrix for the input values, then θ_i is calculated by the following expression.

$$\theta_i = \left[{X_e}^T . W_i . X_e\right]^{-1} {X_e}^T . W_i . Y \tag{12}$$

Where ${X_e}^T$ is the transpose of matrix X_e. The obtained parameters θ_i for each matrix, substituted in the following equation to approximate the output value of Y.

$$Y \approx X . \theta_i \tag{13}$$

This output value Y is representing the constants that limit the degree of membership functions within the various clusters. These optimized constant parameters are substituted in the linear equation proposed in this study Eq. (14) for the estimation of condensate viscosity. The range of each cluster for input variables (P, T, Rs) and the values of constants parameters (A, B, C, D) in the proposed equation are presented in Table 2. For instance, using the constants in Table 2, the function introduced for rule 1 can be used as follow:

$$\begin{cases} \mu_c = AP + BT + CR_s + D \\ R1: if\ 44.99 < P < 75.15\ and\ 303.15 < T < 405.37\ and\ 5245 < Rs < 6101 \\ Then\ \mu_c = -0.0063P + 0.0025T + 0.0452Rs + 0.0032 \end{cases} \tag{14}$$

Table 2. The range of inputs and constant values of new condensate viscosity correlation.

Rule No	Pressure (MPa)	Temperature (°K)	Rs (scf/STB)	A	B	C	D
1	44.99 – 75.15	303.15 – 405.37	5245 – 6101	-0.0063	0.0025	0.0452	0.0032
2	11.29 – 26.93	348 – 404.6	714 – 9732	0.0003	0.0025	-0.0123	-0.0063
3	19.91 – 20.24	390.72 – 393.15	1167 – 1465	0.0024	0.0022	0.000124	-7.007
4	50.07 – 75.44	303.15 – 315.92	971 – 3646	0.0011	-0.0056	-0.0017	23.84
5	7.99 – 11.29	348.15 – 353.92	1160 – 9869	0.0008	-0.0019	0.000012	-0.5807
6	28.52 – 31.16	323.15 – 353.15	4955 – 6267	0.0006	-0.0034	0.0001	-0.7915

7	3.55 – 10.49	255.37 – 303.15	4695 – 5425	0.0007	0.0014	0.0001	0.2219
8	21.90 – 60.08	315.28 – 405.37	1235 – 9186	0.0001	0.00004	- 0.001	0.6126
9	11.29 – 31.16	323.15 – 443.15	2222 – 2618	0.00001	0.0036	-0.0053	13.97

The developed TSK fuzzy architecture is shown in **Fig. 3** where 9 fuzzy rules interrelate the inputs parameters (P, T and Rs) to the output (μ_c). Output is represented as a dynamic linear relation for the estimation of liquid condensate viscosity.

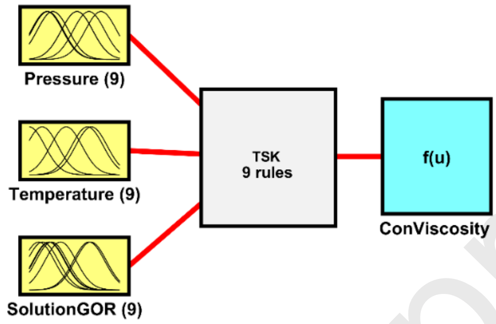


Fig. 3. The architecture of TSK fuzzy AI model for predicting liquid condensate viscosity.

4. Results and discussion

In this study, a new gas condensate viscosity correlation based on the TSK rule-based fuzzy approach has been developed as a function of reservoir pressure (P), temperature (T) and solution gas to oil ratio (Rs). For this aim, a comprehensive data source was collected and divided into two subsets of training (70%) and testing (30%) in a randomised approach. Training data points (958) are partitioned into optimum nine clusters using the k-mean clustering technique. Consequently, nine rules were defined and fired into the partitioned data points to interrelate input parameters of P, T and Rs to the output parameter of condensate viscosity using the gaussian membership function. The proposed correlation can be used to estimate condensate viscosity within P, T and Rs range of [37.7 – 11000 psi (0.25 – 75.84 MPa)], [86 – 338°F (303 – 443.15°K)] and (41.96 – 13496 scf/STB) respectively.

The newly developed model compared with well-known literature correlations of LBC, (1964); Beggs and Robinsons, (1975); Kartoatmodjo and Schmidt, (1991); De Ghetto et al. (1999); Elsharkawy and Alikhan, (1999); Elsharkawy, (2006); Sutton, (2005); Ugwu et al., (2011) and optimized LBC correlation by Yang et al., (2007). These empirical and semi-empirical correlations were developed for a certain region and different operational conditions (pressure and temperature). Hence to implement a fair comparison with the developed AI model in this study, constants of each model are optimized (tuned) using the least square approach to match the utilized experimental data. This approach is similar to the viscosity prediction in reservoir fluid simulation studies (ECLIPSE, 2014; Yang et al., 2007). The testing data set has been used to evaluate the prediction performance of the developed and utilized literature models. For this purpose, statistical error parameters of Root Mean Square Error (RMSE), Mean Average Error (MAE) and Average Absolute Relative Deviation percentage (AARD %) shown in Eq. (15) were used. To visualize the performance of the newly developed model against other methods graphical error analysis include error distribution and graph of the cross plot have been utilized.

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

$$MAE = \left(\frac{\sum_i^N \left|\mu_{exp}.(i) - \mu_{cal}.(i)\right|}{N}\right) \quad (15)$$

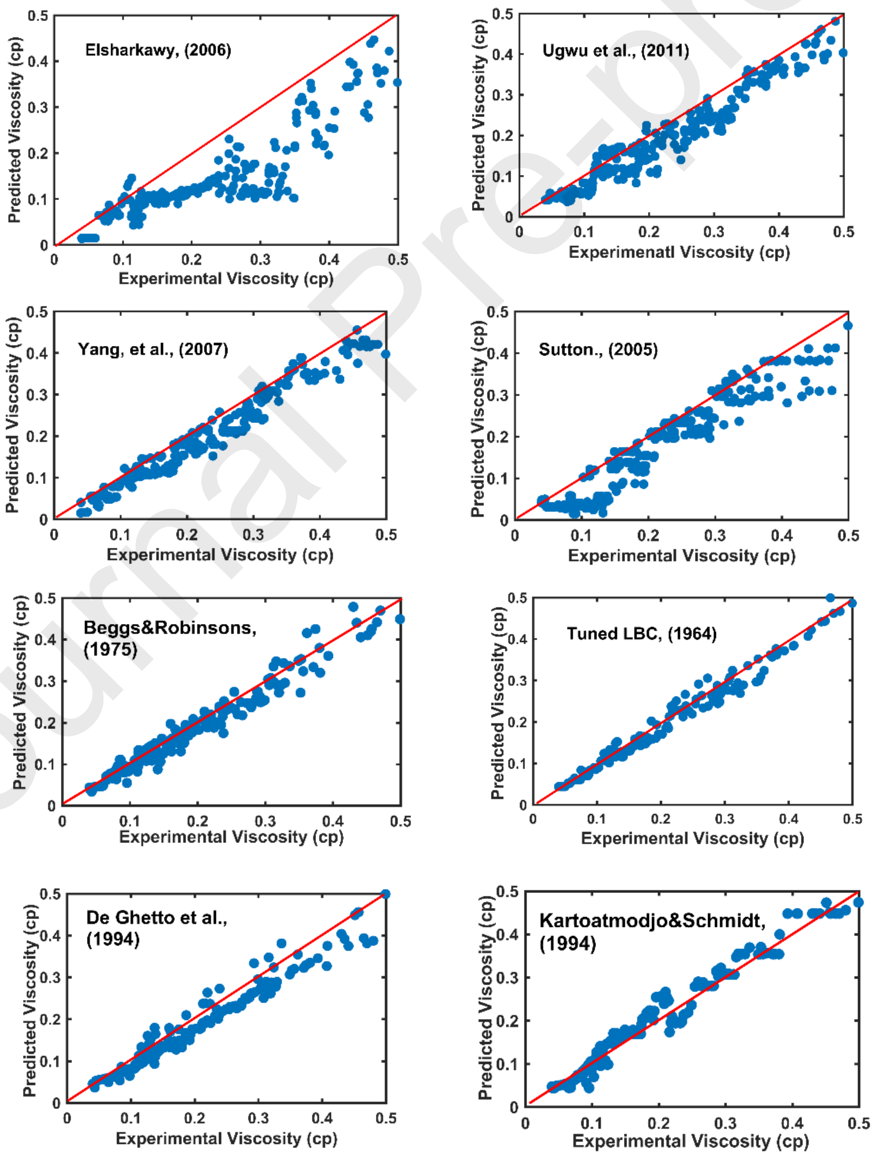
$$AARD\% = \frac{100}{N}\sum_i^N \frac{\left|(\mu_{exp}.(i) - \mu_{cal}.(i))\right|}{\mu_{exp}.(i)} \times 100$$

The prediction performance of the developed correlation using the TSK fuzzy AI method as well as other models in the range of (0 – 0.5cp) have been tested. This lower range is the true representative of mobile condensate liquid that flows in the gas phase toward the wellbore (Fevang, 1995; Whitson and Brulé, 2000). In higher viscosity range the interfacial tension of the liquid condensate is high, which cause the liquid droplets to be absorbed by the formation. **Table 3** summarizes statistical error analysis results of the newly developed model and tuned literature correlations for estimating experimental condensate viscosity data. The results indicate that the developed condensate viscosity correlation yields good agreement between the predicted condensate viscosity and measured condensate viscosity values with the lowest RMSE of 0.0194, MAE of 0.0163, and AARD % of 7.123. The developed model is a function of P, T and Rs. From **Fig. 4** it can be seen that among the literature correlations Kartoatmodjo and Schmidt, (1994), has the highest scattering around zero error line (diagonal line) while tuned LBC, (1964) provides the least spreading for prediction of liquid condensate viscosity. The reason for the high error by utilized literature models is because of the individual limitation of each model for a specific range of viscosity and also different hydrocarbon mixture that was used in their development.

From **Fig. (4)**, the consisting scattering of the data along the zero-error line can be observed using new developed correlation. The new liquid condensate viscosity model responds very well to the pressure and temperature change with the least amount of error. This new correlation can be used in PVT packages of reservoir simulators as an alternative approach for the determination of liquid condensate viscosity.

Table 3. The statistical accuracy of optimized existing literature correlations and the developed TSK model for estimating condensate viscosity using test data.

Method	RMSE	MAE	AARD%
Tuned LBC (1964)	0.0196	0.0153	8.32
Beggs and Robinson (1975)	0.0264	0.0192	9.95
Elsharkawy and Alikhan (1999)	0.0248	0.0176	8.52
De Ghetto et al., (1994)	0.0293	0.0241	12.79
Kartoatmodjo and Schmidt, (1994)	0.0232	0.0194	11.66
Elsharkawy, (2006)	0.1101	0.0908	33.47
Ugwu et al., (2011)	0.0616	0.0714	17.66
Sutton, (2005)	0.3673	0.0869	15.84
Yang et al., (2007)	0.0544	0.0396	14.80
New Correlation	0.0194	0.0163	7.123



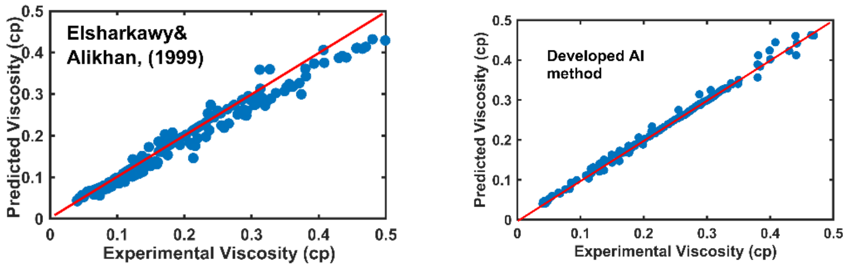


Fig. 4. Cross plot of estimated against condensate liquid viscosity measurements of literature correlations and new developed method.

The validity of the developed correlation using AI for the prediction of two different hydrocarbon mixtures viscosities were examined. Both examples were taken from Pisarev and Kondratyuk, (2019), where they measured viscosities of n-butane (C_4H_{10}) and n-pentane (C_5H_{12}) in different pressure ranges. The viscosity measurements were taken at 360K and 310.95K respectively. The results of AARD% show that the developed model predict the viscosity of n-butane mixture with 10.04% error and viscosity of n-pentane mixture with 16.32%. The graph in Fig. 5 shows the residual plot of relative error calculation. Although the error is still in the acceptable margin, however, the accuracy of the developed model for prediction two mixtures viscosities deteriorates. This is reiterating the fact that a wider spectrum of hydrocarbon mixture viscosity data is required to train the TSK fuzzy model and subsequently extends the range of each input in developed model, which could be subject to further study.

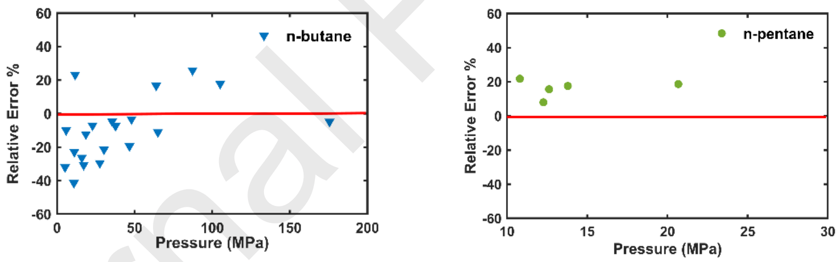


Figure 5. Residual plot of relative error percentage of AI model in predicting n-butane (C_4H_{10}) and n-pentane (C_5H_{12}) viscosities as a function of pressure.

5. Conclusion

A new correlation has been developed for prediction of condensate viscosity of gas condensate reservoirs undergoing depletion. An AI tool called TSK fuzzy approach has been utilized to develop a new liquid condensate viscosity correlation as a function of reservoir pressure, temperature and solution gas to oil ratio. The prediction capability of the proposed correlation has been compared to five tuned existing literature models for condensate viscosity of gas condensate reservoirs. Statistical and graphical error analyses have been used for the comparison of the obtained results. The results in this study show that the new model follows the actual physical trend of the experimental data with very good accuracy. The following conclusion can be drawn based on the obtained results.

- A newly developed correlation based on TSK fuzzy approach provides an accurate prediction of experimental measurements and outperformed existing literature models with the lowest RMSE of 0.0194, MAE of 0.0163 and AARD% of 7.123.
- The new correlation in this study is independent of gas reduced density which is required in CSP based literature correlations and it only needs three parameters of P, T and Rs.
- The obtained results indicate that the new correlation is capable of modelling nonlinear and multivariate condensate (oil) viscosity with only three input parameters of reservoir pressure, temperature and solution gas to oil ratio.
- The validity of the developed model for the prediction of other hydrocarbon mixture viscosities was confirmed using two independent samples.
- The proposed model is valid for a pressure range of 0.25 – 75.84 MPa, temperature range of 303 – 443.15 Kelvin and solution gas to oil ratio of 41.7 – 13496 scf/STB.
- The validity of the model for a wider range of operating conditions can be further investigated in future studies using a larger data bank.
- The developed correlation can be used in PVT calculations of gas condensate reservoirs for reliable modelling and to perform simulation studies.

Statement: The authors declare that they have no conflict of interest.

References

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-Quraishi, A.A., Shokir, E.M., 2011. Artificial neural networks modeling for hydrocarbon gas viscosity and density estimation. J. King Saud Univ. - Eng. Sci. 23, 123–129. https://doi.org/10.1016/j.jksues.2011.03.004

AL-Rousan, N., Mat Isa, N.A., Mat Desa, M.K., 2020. Efficient single and dual axis solar tracking system controllers based on adaptive neural fuzzy inference system. J. King Saud Univ. - Eng. Sci. 32, 459–469. https://doi.org/10.1016/j.jksues.2020.04.004

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

Baled, H.O., Gamwo, I.K., Enick, R.M., McHugh, M.A., 2018. Viscosity models for pure hydrocarbons at extreme conditions: A review and comparative study. Fuel 218, 89–111. https://doi.org/10.1016/j.fuel.2018.01.002

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

Bergman, D.F., Sutton, R.P., 2007. An Update to Viscosity Correlations for Gas-Saturated Crude Oils, in: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, Anaheim. https://doi.org/10.2118/110195-MS

Bikmukhametov, T., Jäschke, J., 2020. First Principles and Machine Learning Virtual Flow Metering: A Literature Review. J. Pet. Sci. Eng. https://doi.org/10.1016/j.petrol.2019.106487

Calinski, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Commun. Stat. - Theory Methods 3, 1–27. https://doi.org/10.1080/03610927408827101
Choudhury, S., Dash, T.P., Bhowmik, P., Rout, P.K., 2020. A novel control approach based on hybrid Fuzzy Logic and Seeker Optimization for optimal energy management between micro-sources and supercapacitor in an islanded Microgrid. J. King Saud Univ. - Eng. Sci. 32, 27–41. https://doi.org/10.1016/j.jksues.2018.03.006
Dandekar, A., 2015. Critical evaluation of empirical gas condensate correlations. J. Nat. Gas Sci. Eng. 27, 298–305. https://doi.org/10.1016/j.jngse.2015.08.072
De Ghetto, G., Paone, F., Villa, M., Spa, A., 1994. Reliability Analysis on PVT Correlations, in: European Petroleum Conference. SPE, London.
ECLIPSE, 2014. Eclipse Reservoir Simulation Reference Manual.
Elsharkawy, A.M., 2006. Efficient methods for calculations of compressibility, density, and viscosity of natural gases, in: Journal of Canadian Petroleum Technology. Petroleum Society of Canada, pp. 55–61. https://doi.org/10.2118/06-06-04
Elsharkawy, A.M., Alikhan, A.A., 1999. Models for predicting the viscosity of Middle East crude oils. Fuel 78, 891–903. https://doi.org/10.1016/S0016-2361(99)00019-8
Faraji, F., Ugwu, J.O., Chong, P.L., Nabhani, F., 2020. Modelling viscosity of liquid dropout near wellbore region in gas condensate reservoirs using modern numerical approaches. J. Pet. Sci. Eng. 185, 1–14. https://doi.org/10.1016/j.petrol.2019.106604
Fattah, K.A., Lashin, A., 2018. Improved oil formation volume factor (Bo) correlation for volatile oil reservoirs: An integrated non-linear regression and genetic programming approach. J. King Saud Univ. - Eng. Sci. 30, 398–404. https://doi.org/10.1016/j.jksues.2016.05.002
Fevang, Ø., 1995. Gas Condensate Flow Behavior and Sampling (PhD Thesis). October. Petroleum Engineering and Applied Geophysics, University of Trondheim.
Gozalpour, F., Danesh, A., Todd, A.C., Tohidi, B., 2005. Viscosity, density, interfacial tension and compositional data for near critical mixtures of methane + butane and methane + decane systems at 310.95 K. Fluid Phase Equilib. 233, 144–150. https://doi.org/10.1016/J.FLUID.2005.03.032
Guo, X., Wang, S., Rong, T., Guo, T., 1997. Viscosity model based on equations of state for hydrocarbon liquids and gases. Fluid Phase Equilib. 139, 405–421. https://doi.org/10.1016/S0378-3812(97)00156-8
Hemmati-Sarapardeh, A., Hajirezaie, S., Soltanian, M.R., Mosavi, A., Nabipour, N., Shamshirband, S., Chau, K.-W., 2020. Modeling natural gas compressibility factor using a hybrid group method of data handling. Eng. Appl. Comput. Fluid Mech. 14, 27–37. https://doi.org/10.1080/19942060.2019.1679668
Hemmati-Sarapardeh, A., Shokrollahi, A., Tatar, A., Gharagheizi, F., Mohammadi, A.H., Naseri, A., 2014. Reservoir oil viscosity determination using a rigorous approach. Fuel 116, 39–48. https://doi.org/10.1016/J.FUEL.2013.07.072
Kariznovi, M., Nourozieh, H., Abedi, J., 2012. Experimental and thermodynamic modeling study on (vapor+liquid) equilibria and physical properties of ternary systems (methane+n-decane+n-tetradecane). Fluid Phase Equilib. 334, 30–36. https://doi.org/10.1016/j.fluid.2012.06.028
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
Khorami, A., Jafari, S.A., Mohamadi-Baghmolaei, M., Azin, R., Osfouri, S., 2017. Density, viscosity, surface tension, and excess properties of DSO and gas condensate mixtures. Appl. Petrochemical Res. 7, 119–129. https://doi.org/10.1007/s13203-017-0183-4
Klawonn, F., Kruse, R., Winkler, R., 2015. Fuzzy clustering: More than just fuzzification. Fuzzy Sets Syst. 281, 272–279. https://doi.org/10.1016/J.FSS.2015.06.024
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
Mamdani, E.H., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man. Mach. Stud. 7, 1–13. https://doi.org/10.1016/S0020-7373(75)80002-2
Najafi-Marghmaleki, A., Tatar, A., Barati-Harooni, A., Choobineh, M.J., Mohammadi, A.H., 2016. GA-RBF model for prediction of dew point pressure in gas condensate reservoirs. J. Mol. Liq. 223, 979–986. https://doi.org/10.1016/j.molliq.2016.08.087
Nazari, D., Abadi, M., Khooban, M.H., 2015. Design of optimal Mamdani-type fuzzy controller for nonholonomic wheeled mobile robots. J. King Saud Univ. - Eng. Sci. 27, 92–100. https://doi.org/10.1016/j.jksues.2013.05.003
Pedersen, K.S., Fredenslund, A., 1987. An improved corresponding states model for the prediction of oil and gas viscosities and thermal conductivities. Chem. Eng. Sci. 42, 182–186. https://doi.org/10.1016/0009-2509(87)80225-7
Pisarev, V., Kondratyuk, N., 2019. Prediction of viscosity-density dependence of liquid methane+n-butane+n-pentane mixtures using the molecular dynamics method and empirical correlations. Fluid Phase Equilib. 501, 112273. https://doi.org/10.1016/J.FLUID.2019.112273
Shokir, E., 2008. Novel Density and Viscosity Correlations for Gases and Gas Mixtures Containing Hydrocarbon and Non-Hydrocarbon Components. J. Can. Pet. Technol. 47. https://doi.org/10.2118/08-10-45
Shokir, E.M., 2006. A Novel Model for Permeability Prediction in Uncored Wells. SPE Reserv. Eval. Eng. 9, 266–273. https://doi.org/10.2118/87038-pa
Strand, K.A., Bjørkvik, B.J.A., 2019. Interface light-scattering on a methane–decane system in the near-critical region at 37.8 °C (100 °F). Fluid Phase Equilib. 485, 168–182. https://doi.org/10.1016/j.fluid.2018.12.016
Sutton, R.P., 2005a. Fundamental PVT Calculations for Associated and Gas-Condensate Natural Gas Systems, in: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, Dallas. https://doi.org/10.2118/97099-MS
Sutton, R.P., 2005b. Fundamental PVT Calculations for Associated and Gas/Condensate Natural-Gas Systems. SPE Reserv. Eval. Eng. 10, 270–284. https://doi.org/10.2118/97099-pa
Takagi, T., Sugeno, M., 1985. Fuzzy Identification of Systems and Its Applications to Modeling and Control. IEEE Trans. Syst. Man Cybern. SMC-15, 116–132. https://doi.org/10.1109/TSMC.1985.6313399
Thomas, F.B., Bennion, D.B., Andersen, G., 2009. Gas Condensate Reservoir Performance. J. Can. Pet. Technol. 48, 18–24. https://doi.org/10.2118/09-07-18
Ugwu, J., Mason, E., Gobina, E., 2011. Modified gas condensate down-hole PVT property correlations.
Whitson, C., 2006. Pressure/Volume/Temperature (PVT) treatment, in: Petroleum Engineering Handbook Volume V, Reservoir Engineering and Petrophysics. Society of Petroleum Engineers, Texas, pp. 1458–1465.
Whitson, C., Brulé, M., 2000. Phase Behavior, 1st ed, Society of Petroleum Engineers. Society of Petroleum Engineers, Richardson. https://doi.org/10.1021/ma00080a014
Yang, T., Fevang, O., Christoffersen, K., Ivarrud, E., 2007. LBC Viscosity Modeling of Gas Condensate to Heavy Oil, in: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, Anaheim. https://doi.org/10.2523/109892-ms
Zadeh, L.., 1965. Fuzzy Sets. Inf. Control 8, 338–353.
Zendehboudi, S., Ali Ahmadi, M., James, L., Chatzis, I., 2012. Prediction of Condensate-to-Gas Ratio for Retrograde Gas Condensate Reservoirs Using Artificial Neural Network with Particle Swarm Optimization. Energy & Fuels 26, 3432–3447. https://doi.org/10.1021/ef300443j

Declaration of interests

The authors declare that they have no conflict of interest.