PREDICTION OF BUBBLEPOINT SOLUTION GAS/OIL RATIO IN THE ABSENCE OF A PVT ANALYSIS

a, b Elmabrouk, S.1; a Shirif, E.

a University of Regina, Faculty of Engineering and Applied Science, Petroleum System Engineering, Regina, Saskatchewan, Canada
b University of Tripoli, Faculty of Engineering, Petroleum Engineering Department, Tripoli, Libya

ABSTRACT

Several published correlations used to estimate the bubblepoint pressure and the bubblepoint oil formation volume factor of reservoir oils require that the value of the bubblepoint solution gas/oil ratio be one of the input variables. Consequently, engineers resort to an additional correlation in order to estimate this value. The majority of the published bubblepoint solution gas/oil ratio correlations are functions of bubblepoint pressure and gas gravity, which can be obtained either experimentally (pressure-volume-temperature, PVT analysis) or estimated from the existing correlations. Thus, it is difficult to apply the correlations in the absence of a PVT analysis. In this study, a multiple regression analysis technique was applied to develop two novel correlations to estimate the bubblepoint solution gas/oil ratio and stock-tank vent gas/oil ratio in the absence of a PVT analysis. The developed correlations can be directly applied by using readily available field data, thus, forgoing the requirement of additional correlations or a PVT analysis. The bubblepoint solution gas/oil ratio correlation is related to the separator gas-oil ratio, to the separator pressure, and to the stock-tank oil specific gravity. However, separator pressure and temperature with the stock-tank oil specific gravity were the only independent variables used in stock-tank vent gas/oil ratio correlation. Another additional and important application of the proposed stock-tank vent gas/oil ratio correlation was to estimate the stock-tank vent gas flow rate.

KEYWORDS

solution gas/oil ratio; stock-tank vent gas/oil ratio; bubblepoint pressure; PVT analysis; PVT correlations

1 To whom all correspondence should be addressed.
Address: University of Regina, Faculty of Engineering and Applied Science, Petroleum System Engineering, Regina, Saskatchewan, Canada
e-mail: saber_elmabrouk@yahoo.com; ezeddin.shirif@uregina.ca
doi:10.5419/bjjpg2011-0022
1. INTRODUCTION

Reservoir fluid pressure-volume-temperature (PVT) studies are designed to simulate the simultaneous fluid flow of oil and gas from the reservoir to the surface. This process is illustrated in Figure 1.

A bubblepoint solution gas/oil ratio ($R_{so}$) is a key parameter in reservoir and in production engineering calculations. The solution gas/oil ratio refers to the amount of gas dissolved in the oil at any given pressure. The solution gas/oil ratio increases with pressure until the attainment of the bubblepoint pressure, after which it remains constant and reaches the bubblepoint solution gas/oil ratio. Bubblepoint pressure ($p_b$) for reservoir oil is the pressure at which the gas begins to come out of solution at constant temperature. These important parameters are usually obtained experimentally (PVT analysis). Additionally, the bubblepoint solution gas/oil ratio is a basic parameter in many PVT correlations used to estimate other fluid properties such as the bubblepoint pressure ($p_b$), and the bubblepoint oil formation volume factor ($B_{ob}$). Numerous correlations available in the oil and gas industry that are used to estimate reservoir oil $p_b$ and $B_{ob}$ require a value of $R_{so}$ as an input variable.

Presently, several correlations exist to obtain the $R_{so}$. Examples of these correlations are presented by Standing (1947), Vasques and Beggs (1980), Petrosky and Farshad (1993), Frashad et al. (1996), Elsharawy and Alkhun (1997), Boukadi et al. (2002), and most recently, Hemmati and Kharrat (2007), Mazandarani and Asghari (2007) and Khamehchi et al. (2009). The basic assumption of $R_{so}$ correlations as a function of the following variables is:

$$R_{so} = f_1 (p_b, \gamma_g, \text{API gravity, } T_R) \quad (1)$$

These correlations are based on the assumption that $R_{so}$ has a strong relationship with bubblepoint pressure ($p_b$), gas specific gravity ($\gamma_g$), stock-tank oil API gravity and reservoir temperature ($T_R$). Accordingly, in order to apply the aforementioned correlations in the absence of a PVT analysis, one must know the $p_b$ and $\gamma_g$ values in advance. The difficulty of measuring parameters $p_b$ and $\gamma_g$ as field data leave engineers with the resort to perform additional correlations to estimate those parameters. Nevertheless, most $p_b$ correlations have its base on the relationship illustrated in Equation 2. Following this line of thought are Standing (1947), Lasater (1958), Vasquez and Beggs (1980), Vazquez and Beggs (1980), Glaso (1980), Al-Marhun (1988), McCain (1990), Dokla and Osman (1992), Macary

![Figure 1. Fluid flow of oil and gas from the reservoir to the separator and stock-tank (two-stage separation).](image-url)
and El-Batanoney (1992), Petrofsky and Farshad (1993), Omar and Todd (1993), De Ghetto and Villa (1994), Kartoatmodjo and Schmidt (1994), Frashad et al. (1996), Almeaid (1997), Hanafy et al. (1997), Boukadi et al. (1999), Velarde et al. (1999), Al-Shammasi (1999), Dindoruk and Christman (2001), Mehran et al. (2006), Hemmati and Kharrat (2007), Ikiensikimama and Ogboja (2009), and Moradi et al. (2010). However, the foresaid correlations are fundamental in order to estimate \( p_b \) and \( R_{sob} \) at temperatures other than \( T_R \) to design surface and subsurface facilities. This occurs due to reservoir fluid properties being found in the PVT analysis reports at \( T_R \) only. In such cases, even though the PVT analysis is available, correlations will provide the estimation of reservoir fluid properties needed.

\[
p_b = f_z (R_{sob}, \gamma_g, \text{API gravity, } T_n)
\]  

(2)

### 2. LITERATURE REVIEW

A review of the literature indicates that the prediction of reservoir fluid properties in the absence of a PVT analysis has received little attention. However, in 2010, Elmabrouk et al. proposed two correlations to estimate \( p_b \) and \( B_{sob} \) in the absence of a PVT analysis. The researchers used readily measured field parameters as input variables without the further requirement of additional correlations or a PVT experimental analysis. Presently, literature considers the \( p_b \) correlation as the only existing correlation to estimate \( p_b \) directly in the absence of a PVT analysis.

However, in most cases, \( R_{sob} \) correlations are derived by mathematically solving the \( p_b \) correlation to obtain \( R_{sob} \). An exception to this precedent are the correlations presented by Vazquez and Beggs (1980), Elsharkawy and Alikhán (1997), Boukadi et al. (2002), Hemmati and Kharrat (2007), Mazandarani and Asghari (2007), and Khamehchi et al. (2009), where the authors treated \( R_{sob} \) as an independent variable based on Equation 1. Nevertheless, the majority of the available \( R_{sob} \) correlations are impractical in the absence of a PVT analysis, simply because \( R_{sob} \) correlations require knowledge of \( p_b \) and \( \gamma_v \) values. As a result, it is difficult to apply the correlations in the absence of a PVT analysis.

Furthermore, an indirect way of obtaining the \( R_{sob} \) is to add the estimated stock-tank vent gas/oil ratio, \( R_{ST} \), to the field measured separator gas/oil ratio, \( R_{SP} \), as shown in Equation 3. This is only valid if the reservoir pressure is above the \( p_b \).

Accordingly, Rollins et al. (1990) developed a correlation (Equation 4) to estimate the \( R_{ST} \) as a function of \( \gamma_{ost}, \gamma_{gSP}, p_{sp} \) and \( T_{SP} \). However, this work has a practical limitation in the absence of a PVT analysis due to the difficulty in determining \( \gamma_{gSP} \). The \( \gamma_{gSP} \) value rarely is measured in the field and it is determined usually based on a gas composition analysis.

\[
R_{sob} = R_{SP} + R_{ST}
\]  

(3)

\[
\ln R_{ST} = 0.2126 - 4.916 \ln \gamma_{OST} + 3.469 \ln \gamma_{gSP} + 1.501 \ln p_{SP} - 0.9213 \ln T_{SP}
\]  

(4)

In 2003, Valko and McCain derived a correlation to estimate the \( R_{ST} \) as a function of \( p_{SP}, T_{SP} \) and \( \text{API} \). Applying a non-parametric regression analysis method, with 881 data points, they suggest the empirical Equation 5:

\[
\ln R_{ST} = 3.955 + 0.83z - 0.024z^2 + 0.075z^3
\]  

(5)

where, 

\[
z = \sum_{n=1}^{3} z_n
\]

\[
z_1 = -8.005 + 2.7 \ln p_{SP} - 0.161 (\ln p_{SP})^2
\]

\[
z_2 = 1.224 - 0.5 \ln T_{SP}
\]

\[
z_3 = -1.587 + 0.0441 \text{API} - 2.29 (\text{API})^2
\]

Valko and McCain (2003) estimated the \( R_{sob} \) indirectly using Equation 3 and reported an average RE of 0.0% and an average ARE of 5.2%. They proposed an additional correlation, Equation 6, to estimate the \( R_{sob} \) directly when the separator conditions are unknown. In estimating the \( R_{sob} \), they reported an average RE of 0.0% and an average ARE of 9.9%.

\[
R_{sob} = 1.1617 R_{SP}
\]  

(6)

### 3. NEWLY DEVELOPED CORRELATIONS

This study presents two multiple regression analysis models for predicting the \( R_{sob} \) in the...
absence of a PVT analysis. Both correlations are proposed as a function of readily available field data without the need for additional correlations or an experimental PVT analysis. An additional purpose of the proposed stock-tank vent gas/oil rate correlations is to estimate the stock-tank vent gas flow rate.

3.1 PVT data

A total of 480 separator tests, measured at the $P_{br}$, were obtained from 118 PVT reports and collected from various Libyan oil reservoirs located in the Sirte Basin. The majority of the data points were taken from a two-stage separation test. Normally, a separation flash test is carried out for the two stages of separation, see Figure 1. The volume of gas liberated at each stage and the volume of remaining liquid are measured. The first-stage separator pressure is generally varied to include at least four possible separator pressures at ambient temperature. The second stage is generally carried out under stock tank conditions of 14.7 psia (1 atmosphere) at ambient temperature. The stock tank is usually considered as one stage of separation. However, in single-stage separation, the test is carried out under separator conditions of 14.7 psia at ambient temperature. In this case, $R_{ST}$ is equal to zero.

Prior to the regression analysis, the collected datasets were divided randomly into two sets. The first set of 431 data points was used to construct the regression models, while the second set of 49 data points was used to test validity and compare the regression models against similar available correlations (Valko and McCain, 2003). The first dataset was used to generate the $R_{sob}$ correlation. The range of this data is presented in Table 1. Whereas, only 355 two-stage flash separation tests out of the 431 separator tests were used to generate the $R_{ST}$ empirical correlation, since the data from the single-stage flash separation tests were excluded. This dataset is described in Table 2.

3.2 Bubblepoint solution gas/oil ratio, $R_{sob}$ correlation

A novel correlation to estimate the $R_{sob}$ was developed during this study. The data were correlated in terms of the $R_{SP}$, $P_{SP}$, and $\gamma_{oST}$ after the forward-backward search algorithm was carried out to select the best subset input variables. Numerous models were attempted as regression equations, however, the Equation 7 was found to be the most accurate option:

\[
R_{sob} = -152 + 0.99 R_{SP} + 0.256 P_{SP} + \\
+ 19.543 \ln P_{SP} - 557.2 \ln \gamma_{oST}
\]  

(7)

| Table 1. Range of data points for $R_{sob}$ Correlation (431 Data Points). |
|---------------------|----------|------|-------|------|--------|
| Variable | Units | Mean | SD | Minimum | Maximum |
| $R_{sob}$ | scf/STB | 409.3 | 256.6 | 17 | 1256 |
| $R_{SP}$ | scf/STB | 353 | 250.1 | 10 | 1256 |
| $P_{SP}$ | psia | 119.35 | 127.12 | 14.7 | 735 |
| $\gamma_{oST}$ | water=1 | 0.84035 | 0.02496 | 0.7999 | 0.921 |

| Table 2. Range of data points for $R_{ST}$ Correlation (355 Data Points). |
|---------------------|----------|------|-------|------|--------|
| Variable | Units | Mean | SD | Minimum | Maximum |
| $R_{ST}$ | scf/STB | 68.8 | 51.79 | 6.3 | 288.8 |
| $P_{SP}$ | psia | 141.17 | 129.25 | 29.7 | 735 |
| $\gamma_{oST}$ | water=1 | 0.84079 | 0.02561 | 0.7999 | 0.921 |
| $T_{SP}$ | °F | 111.19 | 30.79 | 73 | 194 |
scribed in Table 2, were Logarithms of residuals were calculated for Equation 7 and those obtained from Equation 8, using 355 data points described in Table 2, were compared with $R_{sob}$ values calculated from Valko and McCain’s correlations, Equations 5 and 6. This comparison is illustrated in Figure 4. The cross-plot shows that the majority of the plotted data points, estimated using the novel proposed correlations, fall on the $y=x$ line, indicating that the two proposed models have a higher accuracy than the Valko and McCain’s correlations providing far more reliable results. Moreover, the histogram of the relative error of the proposed correlations, along with the relative error from the Valko and McCain’s

### Table 3. Quantitative statistical analysis of the present correlations.

<table>
<thead>
<tr>
<th>Statistical Criterion</th>
<th>$R_{sob}$ Correlation</th>
<th>$R_{ST}$ Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$, %</td>
<td>99.3</td>
<td>99.2</td>
</tr>
<tr>
<td>Average RE, %</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Average ARE, %</td>
<td>6.8</td>
<td>7.6</td>
</tr>
<tr>
<td>Minimum ARE, %</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Maximum ARE, %</td>
<td>83.9</td>
<td>46.3</td>
</tr>
</tbody>
</table>

#### 3.3 Stock-tank vent gas/oil ratio, $R_{ST}$ correlation

Usually, the $R_{SP}$ is a field measurement, unlike the $R_{ST}$ which rarely is measured in the field. Thus, in order to obtain the $R_{sob}$ indirectly one must first estimate the $R_{ST}$, and then add it to the field measured $R_{SP}$. In this work, a novel correlation for estimating the $R_{ST}$ was developed. $P_{SP}$, $T_{SP}$, and $\gamma_{OST}$ were the only independent variables. A total of 355 two-stage separator tests were used to develop the $R_{ST}$ correlation. The range of the dataset is illustrated in Table 2. The natural logarithm of the $R_{ST}$ was regressed as a function of the natural logarithms of $P_{SP}$, $T_{SP}$, and $\gamma_{OST}$. Subsequently, the empirical Equation 8 was obtained:

$$R_{ST} = P_{SP}^{1.02} \gamma_{OST}^{-0.47} T_{SP}^{-0.524}$$

The model shows small values of average RE and average ARE of 1.1% and 7.6%, respectively, with a high $R^2$ of 99.2%, indicating that the model is sufficient for describing the data.

### 4. Comparison with other correlations

In order to study the accuracy of the proposed correlations, the estimated $R_{sob}$ values from Equation 7 and those obtained from Equation 8, using 355 data points described in Table 2, were compared with $R_{sob}$ values calculated from Valko and McCain’s correlations, Equations 5 and 6. The model shows a small average RE of 1.2% and an average ARE of 6.8% with a high $R^2$ of 99.3%. Figure 2 shows the histogram plot of the residuals, indicating that the error terms follow a normal distribution with a mean almost equal to zero, and satisfy the normality assumption. The estimated $R_{sob}$ was plotted as a function of the actual values in Figure 3. The cross-plot shows the precision of the empirical correlation without bias. Table 3 summarizes the results of the quantitative analysis of the residuals.
correlations model, is plotted in Figure 5. The error terms from Equations 5, 7 and 8 satisfy the normality assumption with a mean equal to -1.958, 1.261 and 0.1568, respectively. However, the relative errors of the $R_{sob}$ values, obtained from Equation 6, deviate from the normal curve, failing to meet the normal distribution assumption. Table 4 summarizes the statistical results of this comparison. It is evident that the two proposed correlations give less average RE and average ARE than the Valko and McCain’s correlations. However, the two proposed correlations have approximately the same accuracy within this range of data.

Table 4. Statistical analysis of evaluation test (355 Data Points).

<table>
<thead>
<tr>
<th>Statistical Criterion</th>
<th>This Study</th>
<th>Valko and McCain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eq. (7)</td>
<td>Eq. (8)</td>
</tr>
<tr>
<td>Average RE, %</td>
<td>1.3</td>
<td>-0.2</td>
</tr>
<tr>
<td>Average ARE, %</td>
<td>7.3</td>
<td>6.7</td>
</tr>
<tr>
<td>Maximum ARE, %</td>
<td>56.1</td>
<td>37.0</td>
</tr>
<tr>
<td>Minimum ARE, %</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
However, Valko and McCain’s modes, expressed by high Rsob, estimated from Valko and McCain’s correlations. The data points remained the same in the process of correlation derivation. Figure 6 illustrates the behavior of Rsob, estimated from the two proposed correlations, against Rsob, estimated from Valko and McCain’s correlations. The cross-plot of the estimated values versus the experimental values shows that the majority of the data points plotted, estimated using the proposed models, fall on the 45 degree line. The proposed models, Equations 7 and 8, show a high $R^2$ of 99.61% and 99.97%, respectively. However, Valko and McCain’s modes, expressed by Equations 5 and 6, show an $R^2$ of 99.27% and 83.49%, respectively. Furthermore, the histogram seen in Figure 7 shows the distribution of the relative errors obtained from Equations 5, 6, 7, and 8. The error terms from Equation 5 deviate and skew to the right with a mean equal to −5.474, while the error terms from Equation 6 deviate from the normal distribution. On the other hand, the error terms from Equations 7 and 8 satisfy the normality assumption with arithmetic means equaling to 2.675 and 0.581, respectively. The statistical results of this validation test are illustrated in Table 6. As depicted, the calculated Rsob obtained from the two proposed empirical correlations, gives smaller average RE and average

<table>
<thead>
<tr>
<th>Wave Number</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rsob</td>
<td>473.2</td>
<td>329.9</td>
<td>56</td>
<td>1105.9</td>
</tr>
<tr>
<td>Rsob</td>
<td>421.6</td>
<td>328.3</td>
<td>21</td>
<td>1009.8</td>
</tr>
<tr>
<td>Rsob</td>
<td>51.58</td>
<td>31.39</td>
<td>9.8</td>
<td>155</td>
</tr>
<tr>
<td>Rsob</td>
<td>92.87</td>
<td>35.69</td>
<td>34.7</td>
<td>228</td>
</tr>
<tr>
<td>Rsob</td>
<td>103.83</td>
<td>27.35</td>
<td>73</td>
<td>158</td>
</tr>
<tr>
<td>Rsob</td>
<td>0.83324</td>
<td>0.02052</td>
<td>0.7999</td>
<td>0.8927</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statistical Criterion</th>
<th>This Study</th>
<th>Valko and McCain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eq. (7)</td>
<td>Eq. (8)</td>
</tr>
<tr>
<td>Average RE, %</td>
<td>2.675</td>
<td>0.548</td>
</tr>
<tr>
<td>Average ARE, %</td>
<td>4.177</td>
<td>3.174</td>
</tr>
<tr>
<td>Maximum ARE, %</td>
<td>19.96</td>
<td>25.432</td>
</tr>
<tr>
<td>Minimum ARE, %</td>
<td>0.036</td>
<td>0.031</td>
</tr>
</tbody>
</table>

Moreover, a validation test was carried out using a comparison of the 49 two-stage flash separation tests described in Table 5 against the Valko and McCain’s correlations. The data points remained the same in the process of correlation derivation. Figure 6 illustrates the behavior of Rsob, estimated from the two proposed correlations, against Rsob, estimated from Valko and McCain’s correlations. The cross-plot of the estimated values versus the experimental values shows that the majority of the data points plotted, estimated using the proposed models, fall on the 45 degree line. The proposed models, Equations 7 and 8, show a high $R^2$ of 99.61% and 99.97%, respectively. However, Valko and McCain’s modes, expressed by Equations 5 and 6, show an $R^2$ of 99.27% and 83.49%, respectively. Furthermore, the histogram seen in Figure 7 shows the distribution of the relative errors obtained from Equations 5, 6, 7, and 8. The error terms from Equation 5 deviate and skew to the right with a mean equal to −5.474, while the error terms from Equation 6 deviate from the normal distribution. On the other hand, the error terms from Equations 7 and 8 satisfy the normality assumption with arithmetic means equaling to 2.675 and 0.581, respectively. The statistical results of this validation test are illustrated in Table 6. As depicted, the calculated Rsob obtained from the two proposed empirical correlations, gives smaller average RE and average

![Figure 6. Validation test contrasting actual vs. estimated Rsob (49 data points).](image-url)
ARE, indicating that the proposed correlations have a higher accuracy than Valko and McCain’s correlations, offering more reliable results.

5. CONCLUSIONS

(1) This study proposed two novel and simple empirical correlations, based on a multiple regression analysis, to estimate $R_{sob}$ and $R_{ST}$ in the absence of a PVT analysis. These correlations were developed as a function of field data readily available. Both correlations can be applied straightforwardly in the absence of a PVT analysis without the need for any additional correlations or experimental measurements.

(2) The proposed bubblepoint solution gas/oil ratio correlation was developed based on 480 two-stage and single-stage flash separation tests. The majority of the samples were taken from the two-stage flash separation test. In the single-stage separation test, the separator pressure was the atmospheric pressure, and the stock-tank vent gas/oil ratio was equal to zero. The proposed stock-tank vent gas/oil ratio correlation was developed based on 355 two-stage separation tests.

(3) The two correlations in exam were developed specifically for Libyan oil reservoirs located in the Sirte basin. However, those correlations may be considered for use in other reservoirs throughout the world.

(4) Quantitative and graphical analyses of the residuals were investigated with the purpose of verifying the accuracy of the proposed correlations. The present $R_{sob}$ correlation provides calculated values that have 1.2% average RE of and 6.8% average ARE with a 99.3% $R^2$. The $R_{ST}$ correlation provides calculated values with 1.1% average RE and 7.6% average ARE with a 99.2% $R^2$. The residuals from both correlations satisfy the normal distribution assumption and describe the data well.

(5) A comparison between the proposed correlations and similar correlations available in the literature indicates that the correlations developed in this research offer a better approximation to the normal distribution. The two proposed correlations achieve approximately the same accuracy when predicting $R_{sob}$ values. The $R_{sob}$ values estimated from the proposed correlations can be used as a basic input variable in many PVT imperial correlations in order to estimate other fluid properties such as the $p_b$ and $B_{sob}$.

(6) Another relevant application of the proposed stock-tank vent gas/oil ratio correlation is to estimate the stock-tank vent gas flow rate.
ACKNOWLEDGEMENTS

The authors wholeheartedly appreciate and thank the Faculty of Graduate Studies and Research (FGSR) at the University of Regina for its generous financial support. Moreover, the authors gratefully acknowledge the management of AGOCO, Sirte Oil, Waha Oil, Millita Oil and Gas, Repsol Oil, and Harouge Oil for providing the data and giving permission to publish this work.

NOMENCLATURE

ARE absolute percent relative error, %
GOR gas-oil ratio, Scf/STB
ln natural logarithm
RE percent relative error, %
SD standard deviation
PVT pressure-volume-temperature
API API stock tank oil gravity, degree
Bρbubblepoint oil formation volume factor, bbl/STB
ρb bubblepoint pressure, psia
ρr reservoir pressure, psia
PR separator pressure, psia
R2 correlation coefficient, 100%
Rρsolution gas oil ratio, Scf/STB
Rρ solution gas oil ratio, Scf/STB
RSP separator gas oil ratio, scf/STB
RST stock-tank vent gas oil ratio, scf/STB
TR reservoir temperature, °F
TSP separator temperature, °F
TST Stock-tank temperature, °F
γb gas specific gravity, air=1
γGSP separator gas specific gravity, air=1
γGST stock-tank gas specific gravity, air=1
γGST stock-tank oil specific gravity, water=1

6. REFERENCES


Al-Shammasi, A.A. Bubblepoint pressure and oil formation volume factor correlations. Paper SPE 53185 Presented at the SPE Middle East Oil Show, Bahrein, Feb. 20-23, 1999.

http://dx.doi.org/10.1081/LFT-120002089


http://dx.doi.org/10.1016/S0920-4105(96)00075-7


Mazandarani, M.T.; Asghari, S.M. Correlations for predicting solution gas-oil ratio, bubblepoint pressure and oil formation volume factor at bubblepoint of Iran crude oils. European Congress of Chemical Engineering (ECCE-6), Sep. 2007, p.16-20, Copenhagen.


Appendix A: statistical error analysis

Statistical error analysis is used to analyze the errors (εj) to check the performance and the accuracy of the developed models. This analysis can be determined by various statistical means. The criteria used in this study are; error standard deviation (SD), multiple regression coefficients (R²) and minimum, maximum and average values of both relative and absolute relative error. The εj can
be defined as the deviation of the calculated value from the true value.

\[ e_i = y_i - \hat{y}_i \]

where, \( y_i \) is an observation and \( \hat{y}_i \) is a corresponding fitted value obtained from the fitted regression model.

**Relative error (RE)**

This is an indication of the relative deviation from the true values and is given by:

\[ RE = \frac{e_i}{y_i} \]

**Absolute relative error (ARE)**

It indicates that the relative absolute deviation from the true values as a percentage, defined by:

\[ ARE = |RE| \times 100 \]

**Maximum, minimum and average of RE and ARE**

After the RE and ARE for each data points are calculated, the minimum, maximum, and average values are scanned to know the error range of each model. Lower values imply a better model.

**Standard deviation (SD)**

Standard deviation is a measure of dispersion, defined by:

\[ SD = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2} \]

where \( \bar{x} \) is the average value of the \( n \) observations of \( x_i \). A low SD indicates that the data points tend to be very close to the average, whereas high SD indicates that the data are spread out over a large range of values.

**Correlation coefficient (R²)**

\( R^2 \) is a statistic that will give some information about the goodness of fit of a model. In other words, it represents the degree of success in reducing the SD by regression analysis. An \( R^2 \) of 100% indicates that the regression model perfectly fits the data.

\[ R^2 = \left( 1 - \frac{\Sigma (y_i - \hat{y}_i)^2}{\Sigma (y_i - \bar{y})^2} \right) \times 100 \]

where \( \bar{y} \) is the average of \( y_i \) values.