



RECOVERY OF OIL VISCOSITY VALUES ACCORDING TO ITS ADDITIVE PARAMETERS

A. R. Valeev*, R. R. Tashbulatov, T. A. Barabanshchikova

Ufa State Petroleum Technological University, Ufa, Russia

ABSTRACT

Considering the complexity of mixtures formed within extensive and intricate pipeline networks, determining the viscosity of mixtures through a computational method that utilizes parameters conducive to additivity is advantageous. These parameters may include the density and component composition, as well as corresponding derivatives, such as the content of paraffins, resins, asphaltenes, and fractional composition. This study explored the relationship between the viscosity and these parameters. A linear regression model for estimating the logarithm of the viscosity was developed, where significant regression coefficients were identified, and the prediction error was assessed through cross-validation. A similar approach was adopted for the second-order regression analysis. This analysis incorporated linear regression including parameters pertaining to the fractional composition. Formulas for calculating the viscosity based on the identified properties were derived and the respective errors were evaluated.

Keywords: oil; viscosity; Arrhenius formula; cross-validation; paraffin; resins; asphaltenes; database.

© 2024 «OilGasScientificResearchProject» Institute. All rights reserved.

1. Introduction

In the design and operation of the main oil pipelines, the energy required by the oil pumping units is a key determinant of the economic pumping efficiency [1]. This energy requirement is influenced by the friction head loss within the pipeline, which is directly correlated with the oil viscosity [2].

Notably, even at consistent temperatures, the viscosities of different oils can vary significantly by one to two orders of magnitude. This variation can lead to substantial fluctuations in the energy consumed for pumping [3], highlighting the critical role of viscosity determination in oil transportation via pipelines.

Oil viscosity assessment at varying temperatures is a routine laboratory procedure because the technological evaluation of the main oil pipeline for a specific oil type is a standardized process.

This complexity increases when dealing with the extensive pipeline networks that are characteristically used by domestic oil and gas industries. These networks collect oil from geographically dispersed suppliers, resulting in various oil properties. Operational adjustments, changes in modes, and variations in the operating temperatures due to the extensive length of the pipeline system are common at individual oil pumping stations and pipeline sections. Moreover, repairs and alterations in production and consumption volumes can lead to changes in the oil flow in different sections. As oils from different sources are mixed within the pipeline network, the properties of the oil in any given section are not fixed.

Many oil properties such as the density are additive, whereas viscosity, in general, does not always follow this rule. Therefore, determining the viscosity of a mixed oil is a unique problem. Ideally, the logarithm of the viscosity of the mixture should be equivalent to a linear combination of the logarithms of the viscosities of the individual oils, as described by the Arrhenius equation:

$$\ln(v_{mix}) = x_1 \ln(v_1) + x_2 \ln(v_2) \quad (1)$$

In addition to the Arrhenius equation, other models such as those presented by Kendal and Monroe, Kern (Bingham, Zdanowski), and Walter are recognized as effective for describing the viscosity of mixtures, highlighting the complexity of this issue.

There are instances when the logarithm of the viscosity of a mixture is less than the linear combination it is based on, exhibiting characteristics known as a «viscosity minimum» (rational mixing, singular mixing) or «viscosity maximum» (irrational mixing).

Consequently, within the vast array of possible oil combinations entering a pipeline network, a multitude of mixtures is formed at any given pipeline section. The viscosities of these mixtures cannot be precisely calculated solely from the ratios of their components. Nonetheless, knowing the viscosities of these mixtures is essential for effective network operation planning and determining the energy consumption.

It is theoretically feasible to prepare all possible mixtures of extracted oils and measure their viscosities. However, even with a certain level of mixing discreteness, the number of potential combinations presents a formidable task in practice.

Therefore, predicting oil viscosity for additives based on

*E-mail: anv-v@yandex.ru

<http://dx.doi.org/10.5510/OGP20240100943>

other properties that are more reliable is of great significance. Such properties include the component composition and its derivatives (e.g., the content of paraffins, resins, asphaltenes, and fractional composition), as well as density.

Accurately solving this problem would enable the determination of the viscosity of oil mixtures at any point within the pipeline network via calculations through preliminary identification of the properties of oil from each supplier, thereby reducing the need for additional laboratory analyses.

This study initiated a series of articles by the authors on predicting the mixture viscosity of different oils by leveraging their other additive properties.

The calculations were based on a database of oil properties referenced from the USSR Oil Handbooks [4], which contain data on the viscosity, density, and contents of paraffin, resins, asphaltenes, and the component composition of 773 oils produced in the 1970s in the USSR.

This database includes viscosity values at temperatures ranging from 0 to 70 °C, in 10 °C increments, noting that different oils may have varying properties.

2. Correlation between oil viscosity and its additive parameters

This study established the relationship between the kinematic viscosity of oil and its density and pour point, as well as the paraffin, resin, and asphaltene contents.

However, not all the listed properties are predictive of viscosity. The initial analysis explored the hypothesis that there is a correlation between the oil viscosity and one of these properties.

To test for correlations, statistical methods such as Pearson’s criterion, Spearman’s rank correlation coefficient, and Kendall’s rank correlation coefficient were employed. A correlation was inferred if at least two of the three tests yielded positive results.

The findings at the 99.73% confidence level (three-sigma) are collated in table 1. The general correlation trend persisted even when the confidence level was increased to 99.99994% (five sigma).

Based on the data obtained, there was a demonstrable correlation between the oil viscosity and the presence of paraffin, resins, asphaltenes, and the density. Conversely, there was almost no correlation between the viscosity and the pour point across most temperatures; thus, this parameter was excluded from subsequent analyses. It is also important

to acknowledge the limited reliability of the correlation data for temperatures of 0, 60, and 70 °C due to the small sample size; therefore, these temperatures will not be further considered.

3. Linear regression between oil viscosity and additive parameters

The subsequent phase involved developing a linear regression model for describing the relationship between the oil viscosity and corresponding additive parameters. A preliminary analysis revealed that the absolute value of the kinematic viscosity, measured in mm²/s, does not lend itself well to the approximation. In contrast, the logarithm of the kinematic viscosity presents a better fit, a finding that aligns with other viscosity-related formulas, such as the Arrhenius equation mentioned earlier. Consequently, all subsequent calculations were performed using the logarithmic value of the viscosity.

The analysis includes crafting a linear regression equation, with density (measured at 20 °C in kg/m³) and the content of paraffin, resins, and asphaltenes (expressed as a percentage) as variables [5]. The linear regression coefficients obtained are summarized in table 2. The significance of these coefficients, which was evaluated using Fisher’s criterion at a significance level of 0.05, is indicated in parentheses.

Subsequently, the error in determining the absolute viscosity and the associated confidence interval (also at a significance level of 0.05) were determined. The mean and median error rates were calculated. The median error was consistently lower than the mean error, suggesting the presence of outlier-specific oils for which the error was notably high.

Given that the mean and median errors represent the deviation of the known dataset or the training sample, there is a potential risk of overfitting the model, which could compromise its reliability in predicting other unseen data. To mitigate this risk, error evaluation using element-by-element cross-validation was conducted. This method of validation tends to yield a higher error rate than the mean but more accurately represents the prediction error for new data. This cross-validation error was considered while examining the results.

Although an error margin of 30–40 % in viscosity determination might appear substantial, this error diminishes when the viscosity is used to calculate the energy consumption and friction head losses. For instance, head losses in pipelines,

Information on correlation between oil viscosity and other parameters					
Parameter	Paraffin content	Resin content	Asphaltene content	Density at 20 °C	Solidification temperature
Viscosity at 0 °C	No (18)	No (12)	Yes (18)	Yes (18)	No (14)
Viscosity at 10 °C	Yes (136)	Yes (130)	Yes (137)	Yes (137)	No (88)
Viscosity at 20 °C	Yes (447)	Yes (401)	Yes (439)	Yes (472)	No (302)
Viscosity at 30 °C	Yes (416)	Yes (389)	Yes (416)	Yes (422)	No (285)
Viscosity at 40 °C	Yes (436)	Yes (393)	Yes (436)	Yes (440)	No (302)
Viscosity at 50 °C	Yes (514)	Yes (442)	Yes (504)	Yes (530)	No (333)
Viscosity at 60 °C	No (32)	Yes (12)	Yes (31)	Yes (31)	Yes (25)
Viscosity at 70 °C	No (34)	Yes (14)	Yes (34)	Yes (33)	Yes (25)

Note: The number of points in parentheses indicates the number of points in each dependency when performing the statistical test

Table 2

Results of viscosity approximation using linear regression equation

Parameter	Average error, %	Median error, %	Error considering element-by-element cross-validation, %	Number of points	Values of linear regression coefficients				
					Free term, %	Paraffin content, %	Resin content, %	Asphaltene content, %	Density at 20 °C, kg/m ³
Viscosity at 10 °C	34.9 ± 5.9	25.3 ± 8.0	36.5 ± 8.4	129	-25.252 (significant)	0.0535 (significant)	0.000675 (none)	-0.0235 (none)	0.0329 (significant)
Viscosity at 20 °C	42.5 ± 4.4	31.4 ± 1.9	43.2 ± 2.9	395	-27.168 (significant)	0.0796 (significant)	-0.00209 (none)	-0.00123 (none)	0.0346 (significant)
Viscosity at 30 °C	34.8 ± 3.7	25.8 ± 1.1	35.3 ± 2.1	385	-23.517 (significant)	0.0356 (significant)	0.000775 (none)	0.000882 (none)	0.0300 (significant)
Viscosity at 40 °C	30.6 ± 3.2	22.9 ± 1.1	31.1 ± 2.5	388	-21.576 (significant)	0.0245 (significant)	0.00253 (none)	0.000128 (none)	0.0274 (significant)
Viscosity at 50 °C	30.9 ± 3.4	22.6 ± 1.5	31.4 ± 3.0	433	-20.098 (significant)	0.0119 (significant)	0.00216 (none)	-0.00371 (none)	0.0255 (significant)

The result of assessing the significance of the regression coefficient using Fisher's criterion is given in parentheses.

especially those pumping oil in turbulent mode through hydraulically smooth pipes (which are most common in main oil pipelines), are affected by viscosity by a factor of 0.25, according to Leibenzon's formula [13]). Therefore, the error in determining the head losses is reduced to approximately 6.8–8.7%.

Based on the obtained results, the following equations were formulated for determining the viscosity:

$$v_{10} = \exp(-25.3 + 5.35P + 0.0675R - 2.3500A + 32.9p_{20}) \quad (2)$$

$$v_{20} = \exp(-27.2 + 7.96P - 0.2090R - 0.1230A + 34.6p_{20}) \quad (3)$$

$$v_{30} = \exp(-23.5 + 3.56P + 0.0775R + 0.0882A + 30.0p_{20}) \quad (4)$$

$$v_{40} = \exp(-21.6 + 2.45P + 0.2530R + 0.0128A + 27.4p_{20}) \quad (5)$$

$$v_{50} = \exp(-20.1 + 1.19P + 0.2160R - 0.3710A + 25.5p_{20}) \quad (6)$$

where viscosities v_{10} – v_{50} are represented in mm²/s; the contents of paraffins P , resins R , and asphaltenes A are in fractions; and the density p_{20} is in t/m³.

For the established formulas and those that follow, it is important to acknowledge that some of the regression coefficients are statistically insignificant. However, including all the determined coefficients in the calculations is recommended for preventing an increase in the formula error.

Notably, the influence of paraffin on the viscosity diminishes as temperature rises.

4. Construction of second-order regression equation for oil viscosity and additive parameters

Implementing a second-order regression is a practical approach for reducing the error in linear regression. The model not only considers the density and content of paraffin, resins, and asphaltenes, but also incorporates the squares of these values and their pairwise products, resulting in a total of 15 variables.

For the analysis, a method analogous to that previously employed was used. To communicate the results concisely, the table lists only the significant regression coefficients.

The following conclusions were drawn based on the obtained results. The significant regression coefficients for the different viscosities were different, and it was difficult to identify unambiguous patterns. These errors are smaller than those observed in the linear regression. The two-sample

Student's t-test at a significance level of 0.05 was used to objectively compare the errors of the two regressions. The results of these analyses are summarized in table 4.

The results show that the second-order regression decreases the error for temperatures of 20, 30, 40, and 50 °C. For the temperature of 10 °C, the change is insignificant.

The obtained dependencies for viscosity determination are given below:

$$v_{10} = \exp(-13.4 + 172P - 6.46R - 227A + 5.51p_{20} - 180P^2 + 375PA - 179Pp_{20} - 11.1RP - 0.835R^2 - 6.84RA + 9.24Rp_{20} - 25.9A^2 + 245p_{20}A + 15.2p_{20}^2) \quad (7)$$

$$v_{20} = \exp(40.0 - 2.30P - 6.78R + 51.6A - 117p_{20} + 48.6P^2 + 51.6PA + 2.72Pp_{20} + 0.650RP - 2.40R^2 + 14.6RA + 10.4Rp_{20} + 14.5A^2 - 76.1p_{20}A + 84.6p_{20}^2) \quad (8)$$

$$v_{30} = \exp(63.7 + 24.2P - 14.1R + 88.6A - 167p_{20} - 3.83P^2 + 23.9PA - 23.7Pp_{20} - 0.762RP - 3.57R^2 + 14.6RA + 20.2Rp_{20} + 11.3A^2 - 113p_{20}A + 110p_{20}^2) \quad (9)$$

$$v_{40} = \exp(44.8 + 34.6P - 23.2R + 97.4A - 119p_{20} - 3.14P^2 + 17.5PA - 37.9Pp_{20} - 0.744RP - 3.51R^2 + 9.74RA + 30.7Rp_{20} + 8.68A^2 - 119p_{20}A + 79.1p_{20}^2) \quad (10)$$

$$v_{50} = \exp(16.6 + 65.7P - 25.1R + 99.4A - 52.7p_{20} - 0.473P^2 + 0.462PA - 80.4Pp_{20} + 11.3RP - 3.22R^2 + 12.3RA + 31.8Rp_{20} + 10.1A^2 - 123p_{20}A + 40.9p_{20}^2) \quad (11)$$

where viscosities v_{10} – v_{50} are represented in mm²/s; the contents of paraffins P , resins R , and asphaltenes A are in fractions; and the density p_{20} is in t/m³.

5. Linear regression between oil viscosity and fractional composition

The presence of paraffins, resins, and asphaltenes indicates the amalgamated composition of oil. Consequently, considering the fractional composition of oil as a factor for predicting the viscosity is prospectively instructive.

Determining the most informative fractional composition data for viscosity prediction is a complex task that should be addressed in future studies.

Table 3

Results of viscosity approximation using the second-order regression equation

Parameter	Average error, %	Median error, %	Crossval. error, %	Number of points	Significant regression coefficients
Viscosity at 10 °C	32.0 ± 5.2	23.9 ± 6.0	36.7 ± 7.5	129	P^2 ,
Viscosity at 20 °C	39.0 ± 4.1	28.4 ± 3.1	42.5 ± 3.3	395	P^2, T^2, A^2 ,
Viscosity at 30 °C	30.6 ± 3.0	22.3 ± 2.3	32.5 ± 2.2	385	$a_0, p_{20}, T^2, T \cdot A, p_{20}^2$
Viscosity at 40 °C	26.9 ± 2.7	20.8 ± 2.2	28.7 ± 2.0	388	$a_0, T, p_{20}, T^2, T \cdot p_{20}, p_{20} \cdot A, p_{20}^2$
Viscosity at 50 °C	27.5 ± 2.9	18.3 ± 2.9	29.2 ± 2.8	433	$P, T, A, P \cdot p_{20}, T^2, T \cdot A, T \cdot p_{20}, A^2, p_{20} \cdot A$

The following notations are used in the table: a_0 – free regression term; P – paraffin content; T – tar content; A – asphaltene content; p_{20} – density at 20 °C

Table 4

Comparisons of linear regression and second-order regression accuracies for viscosity determination

Parameter	Linear regression		Second-order regression		Criterion statistics	Table value	Comparison result
	Average value, %	Standard deviation, %	Average value, %	Standard deviation, %			
Viscosity at 10 °C	36.5	3.2	36.7	3.1	-0.461	1.969	Not different
Viscosity at 20 °C	43.2	2.3	42.5	2.8	3.808	1.963	Different
Viscosity at 30 °C	35.3	1.9	32.5	1.8	21.588	1.963	Different
Viscosity at 40 °C	31.1	1.7	28.7	1.7	19.226	1.963	Different
Viscosity at 50 °C	31.4	1.8	29.2	1.7	18.887	1.963	Different

The general feasibility of using the fractional composition data was assessed through an experiment similar to that previously conducted. Prior research has indicated that the viscosity can be estimated through linear regression by utilizing parameters such as the density and paraffin content. By incorporating the fraction yield data at a specific temperature, e.g., 300 °C, we can further explore this predictive approach. The analysis results are presented in table 5.

The two-sample Student’s t-test with a significance level of 0.05 was applied. This statistical test assists in determining whether the calculation error for viscosity is indeed reduced by using this additional data. The results of the statistical analyses are presented in table 6.

According to the data presented in the table, incorporating the fraction yield data up to 300 °C markedly decreased the error in the viscosity calculation. Therefore, the fraction yield is a valuable parameter for predicting oil viscosity. The

authors will focus on this aspect in subsequent studies.

The following formulas for viscosity were derived from the results:

$$v_{10} = \exp(-20.3 + 5.76P + 28.1p_{20} - 2.25F_{300}) \quad (12)$$

$$v_{20} = \exp(-10.2 + 7.23P + 17.2p_{20} - 5.02F_{300}) \quad (13)$$

$$v_{30} = \exp(-8.83 + 1.83P + 15.2p_{20} - 4.46F_{300}) \quad (14)$$

$$v_{40} = \exp(-7.65 + 0.825P + 13.5p_{20} - 4.43F_{300}) \quad (15)$$

$$v_{50} = \exp(-6.15 + 0.0341P + 11.4p_{20} - 4.30F_{300}) \quad (16)$$

where the viscosities $v_{10}-v_{50}$ are represented in mm²/s; the content of paraffin P , resins R , asphaltenes A , and the yield of the fraction up to 300 °C F is in fractions, and the density p_{20} is in t/m³.

As the temperature increases, the effect of the paraffins on the viscosity decreases.

Table 5

Results of viscosity approximation using linear regression equation and information on density, paraffin content, and yield fraction up to 300 °C

Parameter	Average error, %	Median error, %	Crossval. error, %	Number of points	Values of linear regression coefficients			
					Free term, %	Paraffin content, %	Density at 20 °C, kg/m ³	Fraction yield fraction up to 300 °C, %
Viscosity at 10 °C	32.4 ± 6.8	25.1 ± 6.3	34.2 ± 7.9	80	-20.253 (significant)	0.0576 (significant)	0.0281 (significant)	-0.0225 (significant)
Viscosity at 20 °C	34.7 ± 4.7	26.6 ± 3.4	35.8 ± 3.5	205	-10.173 (significant)	0.0723 (significant)	0.0172 (significant)	-0.0502 (significant)
Viscosity at 30 °C	27.3 ± 3.1	21.8 ± 2.8	28.0 ± 2.9	212	-8.826 (significant)	0.0183 (significant)	0.0152 (significant)	-0.0446 (significant)
Viscosity at 40 °C	24.9 ± 3.2	19.4 ± 2.9	25.4 ± 3.0	219	-7.652 (significant)	0.00825 (none)	0.0135 (significant)	-0.0443 (significant)
Viscosity at 50 °C	24.6 ± 4.3	18.2 ± 3.3	25.2 ± 3.4	232	-6.150 (significant)	0.000341 (none)	0.0114 (significant)	-0.0430 (significant)

The result of assessing the significance of the regression coefficient using Fisher’s criterion is given in parentheses.

Table 6

Comparisons of linear regression accuracies with and without information on the fraction of fraction yield up to 300 °C

Parameter	Linear regression		Second-order regression		Criterion statistics	Table value	Comparison result
	Average value, %	Standard deviation, %	Average value, %	Standard deviation, %			
Viscosity at 10 °C	36.5	3.2	34.2	3.6	4.582	1.971	Different
Viscosity at 20 °C	43.2	2.3	35.8	2.6	34.003	1.964	Different
Viscosity at 30 °C	35.3	1.9	28.0	1.6	49.528	1.964	Different
Viscosity at 40 °C	31.1	1.7	25.4	1.7	39.468	1.964	Different
Viscosity at 50 °C	31.4	1.8	25.2	2.3	36.329	1.964	Different

Similar to the discussion presented above, the derived errors of 25–35 % (table 6) when calculating the energy consumption and friction head losses will result in an error of 5.7–7.7 % (because according to Leibenzon's formula, the head losses are affected by the viscosity by a factor of 0.25 when pumping oil in turbulent mode through an area of hydraulically smooth pipes), which is relatively insignificant.

6. Discussion

Notably, the conclusions drawn are contingent on the specific database used; with an expanded dataset, the results could vary. Some values may have been determined with less precision, and the significance of the applied statistical tests may vary. However, the overall trends of these findings remain consistent. It is also crucial to acknowledge that the initial database contains only a limited number of oils with viscosity data at temperatures of 0, 60, and 70 °C. Consequently, the conclusions drawn for these temperatures are less reliable and should not be indiscriminately extrapolated. This limitation can potentially be addressed by acquiring additional experimental data.

The research conducted led to the following conclusions:

1. A correlation between oil viscosity and the contents of paraffin, resins, asphaltenes, as well as density was established using an available database of oil properties. No correlation between the viscosity and pour point was observed at most temperatures.

2. A linear regression model was developed for predicting the logarithm of the oil viscosity based on these parameters. In the resultant linear regression model, only the density and paraffin content proved to be significant predictors for

each temperature. The models for viscosity prediction were developed with an approximate error of 30–40 %, which translated to an error of 6.8–8.7 % for calculating the head losses and energy consumption. The influence of paraffins on the viscosity decreased with increasing temperature.

3. Second-order regression models were developed, revealing that different parameters became significant at different temperatures. These models achieved a modest reduction in error compared with linear regression.

4. A linear regression model incorporating the paraffin content and fractional yield up to 300 °C was formulated. This model reduced the error in the viscosity calculation to 25–35 %, leading to an error of 5.7–7.7 % in calculating the head losses and energy consumption.

The findings of this study are applicable to the calculation and prediction of operating conditions within a pipeline network that transports oil from different fields, each with distinct properties. The lack of comprehensive viscosity data in the field makes predicting the viscosity of mixed oil flows from various sources challenging, given that viscosity is not additive during mixing. This gap makes it difficult to accurately estimate the head losses and energy consumption within the pipeline. However, because the component composition is additive, it is possible to predict the resulting composition after mixing within a pipeline system. Therefore, the models developed in this study can estimate the viscosity at any point in the pipeline using only the oil component composition. Ultimately, this could minimize the need for extensive laboratory viscosity testing of various oil mixtures and facilitate the determination of the most energy-efficient operating modes for an oil pipeline system as a whole.

The research was carried out at the expense of a grant from the Eurasian Scientific and Educational Center (project No. ENOC-06-22).

References

- Chen, Y., Jin, L., Zhang, M., Minglan, H. (2023). Experimental study on the rheological characteristics and viscosity-enhanced factors of super-viscous heavy oil. *Liquid and Gaseous Energy Resources*, 3(2), 67-75.
- Chen, Y., He, M., Zhang, M., Luo, J. (2023). Prediction and evaluation of polymer turbulent drag reduction efficiency based on BP neural network. *Liquid and Gaseous Energy Resources*, 3(2), 76-88.
- Langbauer, C., Langbauer, T., Fruhwirth, R., Mastobaev, B. (2021). Sucker rod pump frequency-elastic drive mode development – from the numerical model to the field test. *Liquid and Gaseous Energy Resources*, 1(1), 64-85.
- (1971-1974). Oils of the USSR. Handbook. Volume 1-4. Moscow: Khimiya.
- Khuramshina, R. (2023). Analysis of hydrocarbon solvents for removing various types of asphalt, resin, and paraffin deposits. *Liquid and Gaseous Energy Resources*, 3(1), 7-13.