

# Viscosity prediction in selected Iranian light oil reservoirs: Artificial neural network versus empirical correlations

Mohammad Soleimani Lashkenari<sup>1</sup>, Majid Taghizadeh<sup>1\*</sup> and Bahman Mehdizadeh<sup>2</sup>

<sup>1</sup> Chemical Engineering Department, Babol University of Technology, P.O. Box 484, 4714871167 Babol, Iran

<sup>2</sup> National Iranian South Oil Company, Ahwaz, Iran

© China University of Petroleum (Beijing) and Springer-Verlag Berlin Heidelberg 2013

**Abstract:** Viscosity is a parameter that plays a pivotal role in reservoir fluid estimations. Several approaches have been presented in the literature (Beal, 1946; Khan et al, 1987; Beggs and Robinson, 1975; Kartoatmodjo and Schmidt, 1994; Vasquez and Beggs, 1980; Chew and Connally, 1959; Elsharkawy and Alikhan, 1999; Labedi, 1992) for predicting the viscosity of crude oil. However, the results obtained by these methods have significant errors when compared with the experimental data. In this study a robust artificial neural network (ANN) code was developed in the MATLAB software environment to predict the viscosity of Iranian crude oils. The results obtained by the ANN and the three well-established semi-empirical equations (Khan et al, 1987; Elsharkawy and Alikhan, 1999; Labedi, 1992) were compared with the experimental data. The prediction procedure was carried out at three different regimes: at, above and below the bubble-point pressure using the PVT data of 57 samples collected from central, southern and offshore oil fields of Iran. It is confirmed that in comparison with the models previously published in literature, the ANN model has a better accuracy and performance in predicting the viscosity of Iranian crudes.

**Key words:** Viscosity, crude oil, artificial neural network, empirical equations

## 1 Introduction

Viscosity is an important parameter needed for petroleum engineering analysis (Egbogah and Ng, 1990; Larter et al, 2008; Shi et al, 2010). The oil viscosity is a strong function of temperature, pressure, oil specific gravity, dissolved gases and composition of oil mixture (Riazi and Al-Sahhaf, 1996; Martín-Alfonso et al, 2007). Although oil viscosity can be measured isothermally at reservoir temperature and different pressures, viscosity data at different temperatures other than reservoir temperature are needed for design of processing equipments, tubing-string, gas-lift, pipelines and particularly for thermal recovery processes (Beal, 1946; Moharam et al, 1995; Das, 1998; Chang et al, 1999; Kilonzo and Margaritis, 2004; Obanijesu and Omidiora, 2009). In order to solve this problem, empirical correlations are used to predict the viscosity when experimentally measured data are not available. These correlations usually vary in complexity and accuracy depending upon the available crude oil data.

Several studies have been reported on the development of empirical correlations for prediction of crude oil viscosity

as a function of reservoir temperature, oil API gravity, and solution-gas oil ratio (Beal, 1946; Khan et al, 1987; Beggs and Robinson, 1975; Kartoatmodjo and Schmidt, 1994; Vasquez and Beggs, 1980; Chew and Connally, 1959; Elsharkawy and Alikhan, 1999; Labedi, 1992). Most of these correlations are developed for a given area or region using limited viscosity data. These correlations have limited accuracy in estimating crude oil viscosity when applied to new area or regions (Sutton and Farshad, 1990). Other empirical or semi-empirical models use reservoir fluid composition to predict oil viscosity (Lohrenz et al, 1964; Little and Kennedy, 1968).

The physical properties of crude oils determined by an experimental method have high accuracy. However, it is expensive and time-consuming. The oil industry requires a method which is fast, workable and more cost effective than the experimental method. Artificial neural network (ANN) is taken as the best alternative for predicting the physical properties of crude oils, as it takes a short time and is not costly (Van der Walt et al, 1993; Elsharkawy and Gharbi, 2000; Obanijesu and Omidiora, 2009; Omole et al, 2009; Dong et al, 2010). The ANN is currently used in prediction of properties in chemical and petroleum engineering (Roosta et al, 2012; Zendejboudi et al, 2012), prediction/estimation of the solubility of different solvents in supercritical carbon

\*Corresponding author. email: m\_taghizadehfr@yahoo.com

Received July 1, 2012

dioxide (Mehdizadeh and Movagharnejad, 2011), prediction of hydraulic fracturing (Mohaghegh and Ameri, 1994), predicting formation lithology (Wang and Zhang, 2008), and estimation of water saturation (Toomarian et al, 1997).

This paper develops an ANN model for predicting the viscosity of 57 crude oil samples collected from central, southern and offshore oil fields of Iran. This model is designed for estimation of saturated and under-saturated oil viscosities. Input parameters for this network are temperature, pressure, bubble point pressure, solution gas-oil ratio, and oil specific gravity. The results generated from the ANN model and selected empirical model (Khan et al, 1987; Elsharkawy and Alikhan, 1999; Labedi, 1992), were then plotted against

the existing data from the corresponding field on a cross plot to evaluate the model's degree of accuracy.

## 2 Theory

### 2.1 Empirical viscosity correlations

Over the years, several empirical correlations have been developed for determining viscosity properties of crude oils using data from different geographical locations. The most popular empirical models applied for prediction of Middle East crude oil viscosity are those developed by Elsharkawy and Alikhan (1999), Khan et al (1987), and Labedi (1992). The correlations used for this study are given in Table 1.

**Table 1** Correlations for dead, saturated and under-saturated oil viscosity

Fluid property	Published empirical correlations
Dead oil viscosity	Elsharkawy and Alikhan (1999) $\mu_{od} = 10^x - 1$ $x = 10^y$ $y = 2.16924 - 0.02525 API - 0.68875 \log(T_f)$ (1)
	Labedi (1992) $\mu_{od} = \frac{10^{9.224}}{API^{4.7013} T_f^{0.6739}}$ (2)
Saturated oil viscosity	Elsharkawy and Alikhan (1999) $\mu_{ob} = A(\mu_{od})^B$ $A = 1241.932(R_s + 641.026)^{-1.1241}$ $B = 1768.841(R_s + 1180.335)^{-1.06622}$ (3)
	Khan et al (1987) At the bubble-point: $\mu_{ob1} = \frac{0.09(\gamma_g)^{0.5}}{R_s^{1/3}(\theta_r)^{4.5}(1-\gamma_o)^3}$ (4)
	Below the bubble-point: $\mu_{ob2} = \mu_{ob1}(P/P_b)^{-0.14} \exp[-2.5 \times 10^{-4}(P - P_b)]$ (5)
	Labedi (1992): At the bubble-point: $\mu_{ob1} = (10^{2.344-0.03542 API}) \times (\mu_{od}^{0.6447}) / (P_b^{0.426})$ (6)
	Below the bubble-point: $\mu_{ob2} = \frac{\mu_{ob1}}{1 - M_{ub}[1 - (P/P_b)]}$ $M_{ub} = 10^{-3.876} P_b^{0.5423} API^{1.1302}$ (7)
Under-saturated oil viscosity	Elsharkawy and Alikhan (1999) $\mu_o = \mu_{ob} + 10^{-2.0771}(P - P_b) \times (\mu_{od}^{1.19279} \mu_{ob}^{-0.40712} P_b^{-0.7941})$ (8)
	Khan et al (1987) $\mu_o = \mu_{ob} \exp[9.6 \times 10^{-5}(P - P_b)]$ (9)
	Labedi (1992) $\mu_o = \mu_{ob} - M_{ua}[1 - (P/P_b)]$ $M_{ua} = \frac{(10^{-2.488} \mu_{od}^{0.9036} P_b^{0.6151})}{10^{0.01976 API}}$ (10)

### 2.2 Artificial neural network

Neurons are the main building blocks of artificial neural networks. Each neuron is connected to some of its neighbors with a varying coefficient of connectivity. The connections receive a sum of inputs which generate the output by applying the transfer function. The type of neurons' transfer function used in this study is a sigmoid function in the hidden layer which is defined by the following equation:

$$\theta_j = \frac{1}{1 + e^{-\psi_j}} \tag{11}$$

In Eq. (11),  $\psi$  is the sum of weighted inputs to each neuron and  $\theta$  is the output of each neuron, and  $\psi$  can be calculated from Eq. (12):

$$\psi_j = \left( \sum_{i=1}^n w_{ij} \theta_i \right) + b_j \tag{12}$$

where  $w_{ij}$  denotes the connection between node  $j$  of interlayer  $l$  to node  $i$  of interlayer  $l-1$ ,  $b_j$  is a bias term and  $n$  is the number of neurons in each layer. In any interlayer  $l$  and neuron  $j$  input values are integrated to generate  $\psi_j$ . The bias term is considered as an extra unit to the model that is similar to the intercept term in regression modeling.

In order to minimize the difference between the experimental data and calculated values from the neural network, the above mentioned process was repeated for all the training data. After training and validation, the neural network can be tested with real data. The overall structure of the proposed network in our work is presented in Fig. 1.

There are numerous types of artificial neural networks such as multi-layer perceptron (MLP) network, radial basis function (RBF) network and recurrent neural network (RNN). This work uses a multi-layer perceptron network. The multi-layer perceptron network is one of the most popular and successful neural network architectures, which is suitable to a wide range of applications such as prediction and process modeling (Kim and Adali, 2002; Zhang, 2009).

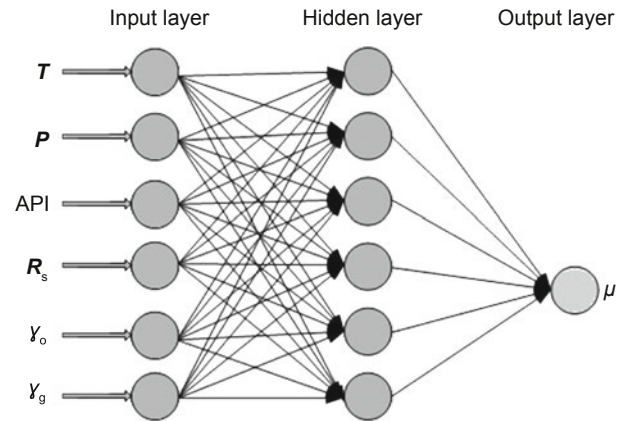


Fig. 1 Overall structure of the proposed network in our work

#### 2.2.1 Preparation of training dataset

In this work, an ANN model was developed for predicting the viscosity of Iranian crude oils using PVT data from samples collected from central, southern and offshore oil fields of Iran. Viscosity data were measured by rolling ball viscometry at different pressures by the Research Institute of Petroleum Industry of Iran (RIPI). Data included measurements of crude oil viscosity at reservoir temperature with 376 samples below the bubble-point pressure, 57 samples at the bubble-point pressure and 287 samples above the bubble-point pressure.

It is reported in the literature that the viscosity of crude oils strictly depends on temperature, pressure, dissolved gas-oil ratio, and oil specific gravity (Chew and Connally, 1959; Beggs and Robinson, 1975; Vazquez and Beggs, 1980; Khan et al, 1987; Labedi, 1992; Kartoatmodjo and Schmidt, 1994; Elsharkawy and Alikhan, 1999). Thus, these parameters were used in ANN in this work. The data range and source of the crude oils for some published models for saturated and under-saturated crude oil viscosity, as well as the proposed model in this work for the Iranian oil reservoirs, are shown in Tables 2 and 3.

Table 2 Data range of some existing models for saturated oil viscosity

Crude	Models			
	Labedi (1992)	Elsharkawy and Alikhan (1999)	Khan et al (1987)	This work
Source	Africa	Middle East	Saudi Arabia	Iran
Gas-oil ratio, scf/stb	13–3533	10–3600	–	315–1558
Bubble-point pressure, psi	60–6358	100–3700	–	800–5156
Saturated oil viscosity, cP	0.11–3.72	0.05–20.89	–	0.20–4.14

Table 3 Data range of some existing models for under-saturated oil viscosity

Crude	Models			
	Labedi (1992)	Elsharkawy and Alikhan (1999)	Khan et al (1987)	This work
Source	Africa	Middle East	Saudi Arabia	Iran
Pressure above bubble-point, psi	–	1287–10 000	Up to 5015	800–5156
Under-saturated oil viscosity, cP	–	0.2–5.7	0.13–71	0.3–6.47

### 2.2.2 ANN modeling

Programming, validation, training and testing of the ANN model was carried out with MATLAB 7.7.0, and all the programs were run on a Pentium IV (CPU 2.7 GHz and 2 GB RAM) personal computer with the Windows XP operating system.

For determination of optimal values of weights and biases of ANN, some steps need to be completed. They include:

1) Data need to be divided into three parts (training subset (60% of all data), validation subset (10% of all data) (see Appendix B) and testing subset (30% of all data)).

2) Data need to be normalized. Normalization can be done by several equations. In this work, data is scaled between [0.1–0.9] by means of Eq. (13):

$$(Scaled)_{value} = \frac{(Actual)_{value} - \min(Actual)_{value}}{\text{Max}(Actual)_{value} - \min(Actual)_{value}} \times 0.8 + 0.1 \quad (13)$$

The normalization step made the ANN models overcome the different magnitudes of the input and output variables.

3) Number of neurons in hidden layer should be optimized.

Some other parameters for neural network training need to be selected by user.

## 3 Results and discussion

In this work, we used a sigmoid function for the transfer function in the hidden layer and the lineal function (purelin) for transfer function in the output layer. The Levenberg-Marquardt back propagation learning algorithm was used for training. In most cases, one hidden layer is enough but the number of neurons in the hidden layer should be optimized for each problem. As mentioned above, optimization of the number of neurons in the hidden layer is one of the most important steps in ANN modeling. Fig. 2 shows the average relative deviations of ANN for testing the subset versus the number of neurons in the hidden layer for the samples at bubble-point pressure, and above and below the bubble-point pressure. Results strongly indicated that 6 neurons in hidden layer can predict with the best accuracy. So an ANN with 6 neurons in input layer, 6 neurons in the hidden layer and one neuron in the output layer was constructed in this study.

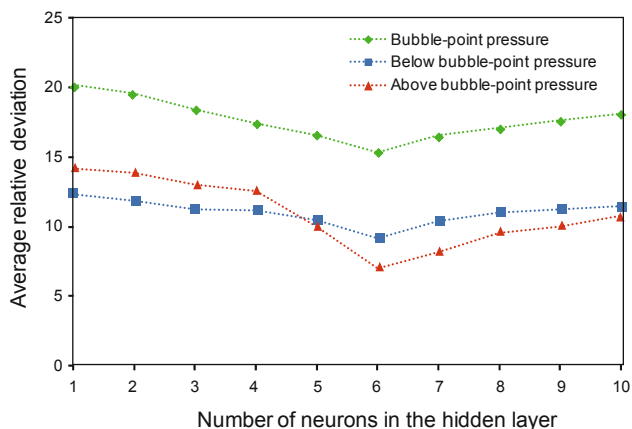


Fig. 2 Average relative deviations of ANN for testing subset versus number of neurons in the hidden layer

In order to compare the performance of the models, statistical error analysis was performed. The accuracy of ANN model presented in this study as well as the ones reported in literature for estimating the viscosity of Iranian crude oils are given in Tables 4-6. These tables show the average percent relative error ( $E_r$ ), average absolute percent relative error ( $E_a$ ), standard deviation ( $S$ ), and the correlation coefficient ( $r$ ) for the models (see Appendix A).

The performance plot (cross plot) is a graph of the predicted versus measured properties with a reference 45° line to ascertain fitness and accuracy of the correlations. A perfect correlation would show a straight line with a slope of 45° (see Figs. 3-8).

After training and evaluation of ANN, its ability to predict the viscosity of Iranian crude oils can be checked. The performance of the proposed architecture (for viscosity of crude oil at, above and below the bubble-point pressure) was studied by comparison with the experimental data and calculated values with the training and testing data sets and the results are shown in Figs. 3-5. The diagonal lines in these figures are the locations of exact predictions and the points in the figures show training and testing data. The error of each prediction is relative to the distance between each point and the diagonal line, so scattered points near the diagonal line can show the overall accuracy of the presented model.

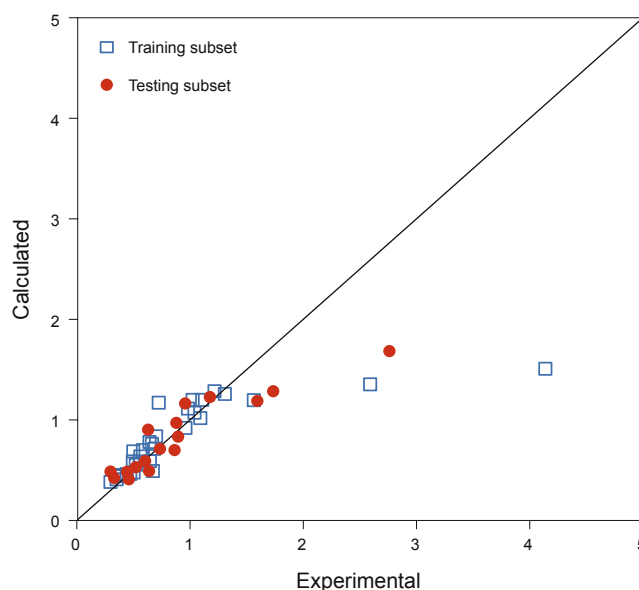


Fig. 3 Comparison of experimental data and calculated values with training and testing data for viscosity of oil at the bubble-point pressure

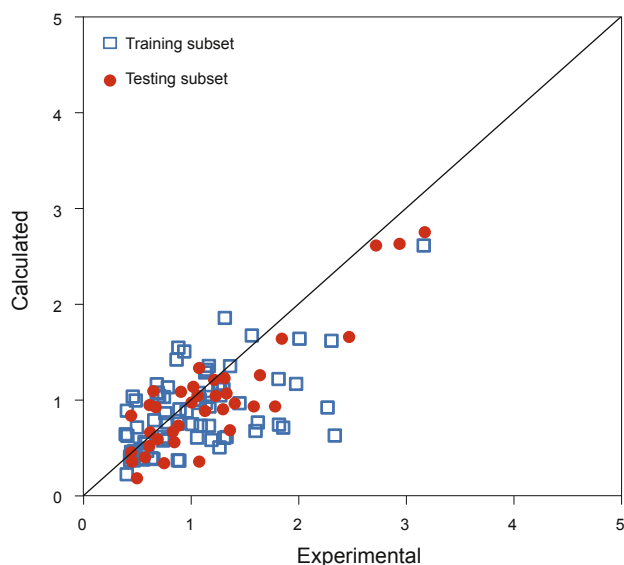
### 3.1 For saturated oil viscosity

#### 3.1.1 For oil at the bubble-point pressure

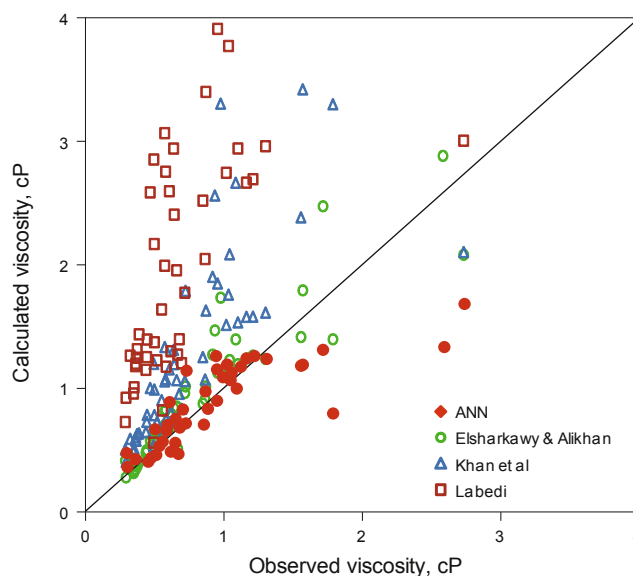
Statistical analysis parameters for all the models for estimating viscosity of Iranian crude oils at the bubble-point pressure show that the ANN model has the smallest average relative error, average absolute relative error, and standard deviation, followed by the model of Elsharkawy and Alikhan (1999), as shown in Table 4. The ANN model has  $E_a$  of 19.3, while Elsharkawy and Alikhan (1999), Khan et al (1987) and Labedi (1992) correlations have  $E_a$  of 20.4, 75.9 and 262.5

**Table 4** Statistical accuracy of models for estimating viscosity of Iranian crude oils at bubble-point pressure

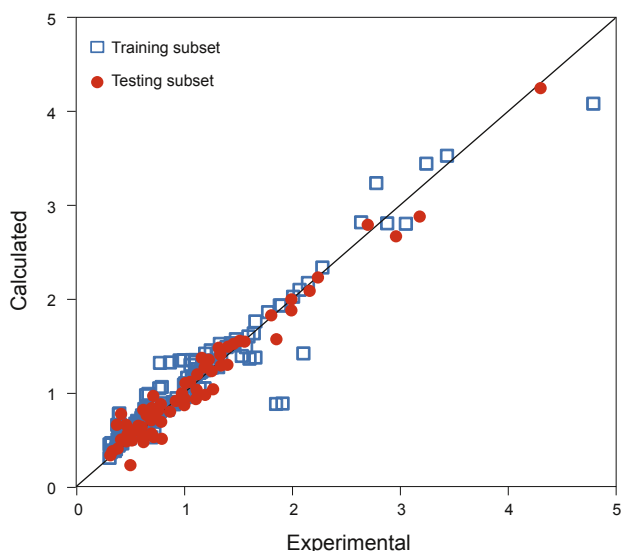
Model	Average percent relative error $E_r$	Average absolute percent relative error $E_a$	Standard deviation $S$	Correlation coefficient $r$
ANN (this work)	-9.3	19.3	17	0.95
Elsharkawy and Alikhan (1999)	-16.9	20.4	19.9	0.92
Khan et al (1987)	-75.1	75.9	49.1	0.69
Labedi (1992)	-262.5	262.5	164.7	0.40



**Fig. 4** Comparison of experimental data and calculated values with training and testing data for viscosity of oil below the bubble-point pressure



**Fig. 6** Performance of models for oil viscosity at the bubble-point oil pressure



**Fig. 5** Comparison of experimental data and calculated values with training and testing data for viscosity of oil above the bubble-point pressure

respectively, indicating that the results from the proposed ANN model are in better agreement with experimental data (see Fig. 6). Fig. 6 also shows that, the ANN model is better than the other models for saturated (bubble-point) oil viscosity below 1.5 cP.

**3.1.2 For oil below the bubble-point pressure**

Table 5 gives the results of the proposed ANN model and other ones for prediction of oil viscosity below the bubble-point pressure. This table shows that ANN model prediction of below the bubble-point viscosity of Iranian oil reservoirs has the largest value of correlation coefficient and the lowest average relative error, standard deviations, and average absolute relative error. As shown in Table 5, the new proposed model, with  $E_r$  of -5.8, has the smallest relative error, whereas Labedi’s correlation (1992), which was developed based on African crude oils, with  $E_r$  of -220.1, has the biggest relative error, among these correlations. Fig. 7, also shows that the new proposed model has the smallest error range. That is, the ANN model could be relied upon more for oil viscosity below the bubble-point pressure with viscosity below 3 cP.

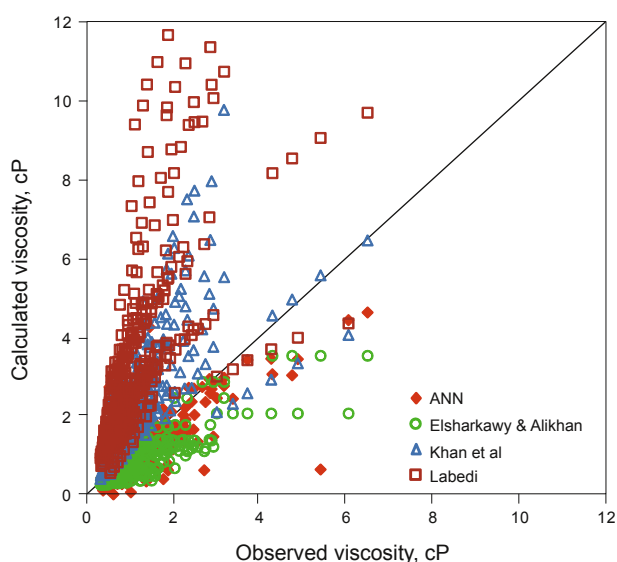
**3.2 For under-saturated oil viscosity**

Statistical analysis parameters for all the models for estimating viscosity of Iranian crude oils above the bubble-point pressure show that the ANN model has the smallest average relative error, average absolute relative error, and standard deviation, followed by the model of Elsharkawy and Alikhan (1999), as shown in Table 6. The ANN model has  $E_a$  of 19.4, while Elsharkawy and Alikhan (1999), Khan et al (1987) and Labedi (1992) correlations have  $E_a$  of

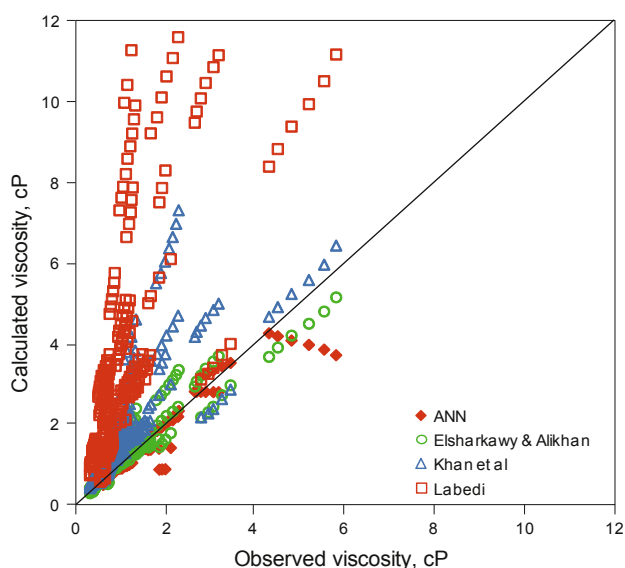
22.9, 77.9 and 265.7 respectively. Fig. 8 shows that Khan et al (1987) and Labedi (1992) correlations underestimate experimental data. This figure shows that the results of the ANN model developed in this work are in better agreement with experimental data compared with the other predictions.

**Table 5** Statistical accuracy of models for estimating viscosity of Iranian crude oils below the bubble-point pressure

Model	$E_r$	$E_a$	$S$	$r$
ANN (this work)	-5.8	13.7	15.2	0.63
Elsharkawy and Alikhan (1999)	19.8	25.6	24.1	0.43
Khan et al (1987)	-102.2	103.2	64.3	0.57
Labedi (1992)	-220.1	220.5	127.8	0.41



**Fig. 7** Performance of models for oil viscosity below the bubble-point pressure



**Fig. 8** Performance of models for oil viscosity above the bubble-point pressure

**Table 6** Statistical accuracy of models for estimating viscosity of Iranian crude oils above the bubble-point pressure

Model	$E_r$	$E_a$	$S$	$r$
ANN (this work)	-15.2	19.4	41.6	0.93
Elsharkawy and Alikhan (1999)	-20.5	22.9	47	0.92
Khan et al (1987)	-77.2	77.9	163.9	0.68
Labedi (1992)	-265.7	265.7	558.8	0.41

## 4 Conclusions

It seems that the most common method for calculating viscosity of crude oils is viscosity correlations. However, these correlations fail to predict oil viscosities in a wide range of operating conditions such as pressure and temperature. In this work, an artificial neural network (ANN) and three well known empirical equations were used to predict the viscosity of Iranian crude oils from three different regions. These models incorporate oil API gravity, reservoir temperature and pressure, and other parameters that can be obtained easily from PVT analysis. Our model was developed using 57 points for saturated oil viscosity, 376 points for viscosity below the bubble-point pressure, and 287 points for viscosity above the bubble-point pressure. It was shown how a multi-layer perceptron network can be trained to represent the viscosity of crude oils. Finally, it was found that in comparison with correlations previously published in the literature, the ability and accuracy of the ANN model for predicting oil viscosities of Iran are better. Our results indicated average relative error ( $E_r$ ) of ANN as 0.1%, -5.8% and -15.2% for the bubble-point pressure, below the bubble-point pressure, and above the bubble-point pressure, respectively.

## Acknowledgement

This work was financially supported by the Iran National Science Foundation (INSF).

## Nomenclatures

- $\gamma_g$  Gas specific gravity (air = 1.0)
- $\gamma_o$  Oil specific gravity (water = 1)
- API Oil API gravity
- $P$  Pressure, psi
- $P_b$  Bubble point pressure, psi
- $R_s$  Solution gas-oil ratio, scf/stb
- $\mu_b$  Oil viscosity below the bubble-point, cP
- $\mu_o$  Under-saturated oil viscosity, cP
- $\mu_{ob}$  Saturated oil viscosity, cP
- $\mu_{od}$  Dead oil viscosity, cP
- $T_f$  Reservoir temperature, °F
- $E_i$  Percent relative error
- $E_a$  Average absolute percent relative error
- $E_r$  Average percent relative error
- $S$  Standard deviation
- $n$  Number of data points
- $r$  Correlation coefficient
- $X_{exp}$  Experimental value for any fluid property
- $X_{est}$  Estimated value for any fluid property

$$\theta_r = (T_r + 459.67) / 459.67 = \text{Relative temperature}$$

$i$  Observation index

## References

- Beal C. The viscosity of air, water, natural gas, crude oil and its associated gases at oil field temperature and pressures. *Trans. AIME (Am. Inst. Min. Metall.)*. 1946. 165. 94-115
- Beggs H D and Robinson J R. Estimating the viscosity of crude oil systems. *J. Pet. Technol.* 1975. 27. 1140-1141
- Chang C, Nguyen Q D and Ronningsen H P. Isothermal start-up of pipeline transporting waxy crude oil. *J. Non-Newtonian Fluid Mech.* 1999. 87. 127-154
- Chew J and Connally Jr C A. A viscosity correlation for gas-saturated crude oils. *Pet. Trans. AIME.* 1959. 216. 23-25
- Das S K. Vapex: An efficient process for the recovery of heavy oil and bitumen. *SPE J.* 1998. 3. 232-237
- Dong X, Wang S, Sun R, et al. Design of artificial neural networks using a genetic algorithm to predict saturates of vacuum gas oil. *Pet. Sci.* 2010. 7. 118-122
- Egbogah E O and Ng J T. An improved temperature-viscosity correlation for crude oil systems. *J. Pet. Sci. Eng.* 1990. 5. 19-200
- Elsharkawy A M and Alikhan A A. Models for predicting the viscosity of Middle East crude oils. *Fuel.* 1999. 78. 891-903
- Elsharkawy A and Gharbi R. Comparing classical and neural regression techniques in modeling crude oil viscosity. *Adv. Eng. Software.* 2000. 32. 215-224
- Kartoatmodjo T and Schmidt Z. Large data bank improves crude physical property correlations. *Oil & Gas J.* 1994. 92. 51-55
- Khan S A, Al-Marhoun M A, Duffuaa S O, et al. Viscosity correlations for Saudi Arabian crude oils. *Society of Petroleum Engineers. SPE* 15720. Bahrain. 1987
- Kilonzo P M and Margaritis A. The effects of non-Newtonian fermentation broth viscosity and small bubble segregation on oxygen mass transfer in gas-lift bioreactors: A critical review. *Biochem. Eng. J.* 2004. 17. 27-40
- Kim T and Adali T. Fully complex multi-layer perceptron network for nonlinear signal processing. *The Journal of VLSI Signal Processing.* 2002. 32. 29-43
- Labedi R. Improved correlations for predicting the viscosity of light crudes. *J. Pet. Sci. Eng.* 1992. 8. 221-234
- Larter S, Adams J, Gates I, et al. The origin, prediction and impact of oil viscosity heterogeneity on the production characteristics of tar sand and heavy oil reservoirs. *J. Can. Pet. Technol.* 2008. 47. 52-61
- Little J E and Kennedy H. A correlation of the viscosity of hydrocarbon systems with pressure, temperature and composition. *Old SPE J.* 1968. 8. 157-162
- Lohrenz J, Bray B and Clark C. Calculating viscosities of reservoir fluids from their compositions. *J. Pet. Technol.* 1964. 16. 1171-1176
- Martín-Alfonso M J, Martínez-Boza F J, Navarro F J, et al. Pressure-temperature-viscosity relationship for heavy petroleum fractions. *Fuel.* 2007. 86. 227-233
- Mehdizadeh B and Movagharnjad K. A comparative study between LS-SVM method and semi empirical equations for modeling the solubility of different solutes in supercritical carbon dioxide. *Chem. Eng. Res. Design.* 2011. 89. 2420-2427
- Mohaghegh S and Ameri S. Artificial neural network as a valuable tool for petroleum engineers. *Unsolicited Paper for Society of Petroleum Engineers.* 1994.
- Moharam H M, Al-Mehaideb R A and Fahim M A. New correlation for predicting the viscosity of heavy petroleum fractions. *Fuel.* 1995. 74. 1776-1779
- Obanijesu E and Omidiora E. The artificial neural network's prediction of crude oil viscosity for pipeline safety. *Pet. Sci. Technol.* 2009. 27. 412-426
- Omole O, Falode O and Deng D A. Prediction of Nigerian crude oil viscosity using artificial neural network. *Pet. Coal.* 2009. 51. 181-188
- Riazi M R and Al-Sahhaf T A. Physical properties of heavy petroleum fractions and crude oils. *Fluid Phase Equilib.* 1996. 117. 217-224
- Roosta A, Setoodeh P and Jahanmiri A. Artificial neural network modeling of surface tension for pure organic compounds. *Ind. Eng. Chem. Res.* 2012. 51. 561-566
- Shi L, Ye Z, Zhang Z, et al. Necessity and feasibility of improving the residual resistance factor of polymer flooding in heavy oil reservoirs. *Pet. Sci.* 2010. 7. 251-256
- Sutton R and Farshad F. Evaluation of empirically derived PVT properties for Gulf of Mexico crude oils. *SPE Reservoir Eng.* 1990. 5. 79-86
- Toomarian N, Barhen J, Glover C, et al. Oil reservoir properties estimation using neural networks. *Proceedings of NEURAP '97.* Marseilles. France. 1997. 133-136.
- Van der Walt T, Van Deventer J and Barnard E. The estimation of kinematic viscosity of petroleum crude oils and fractions with a neural net. *The Chem. Eng. J.* 1993. 51. 151-158
- Vazquez M and Beggs H D. Correlations for fluid physical property prediction. *J. Pet. Technol.* 1980. 32. 968-970
- Wang K and Zhang L. Predicting formation lithology from log data by using a neural network. *Pet. Sci.* 2008. 5. 242-246
- Zendehboudi S, Ahmadi M A, James L, et al. Prediction of condensate-to-gas ratio for retrograde gas condensate reservoirs using artificial neural network with particle swarm optimization. *Energy Fuels.* 2012. 26. 3432-3447
- Zhang Y. An improved QSPR study of standard formation enthalpies of acyclic alkanes based on artificial neural networks and genetic algorithm. *Chemom. Intell. Lab. Syst.* 2009. 98. 162-172

## Appendix A. Statistical analysis

### 1. Percent relative error

$$E_i = \left( \frac{X_{\text{exp}} - X_{\text{est}}}{X_{\text{exp}}} \right)_i \times 100 \quad (i = 1, 2, \dots, n)$$

### 2. Average percent relative error

$$E_r = \frac{1}{n} \sum_{i=1}^n E_i$$

### 3. Average absolute percent relative error

$$E_a = \frac{1}{n} \sum_{i=1}^n |E_i|$$

### 4. Standard deviation

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (E_i - E_r)^2}$$

### 5. Correlation coefficient

$$r = \sqrt{1 - \frac{\sum_{i=1}^n [X_{\text{exp}} - X_{\text{est}}]_i^2}{\sum_{i=1}^n [X_{\text{exp}} - \bar{X}]_i^2}}$$

with

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n [X_{\text{exp}}]_i$$

**Appendix B. Experimental data used in validation subset**

Pressure psi	Viscosity above the bubble-point pressure, cP	Pressure psi	Viscosity below the bubble-point, cP	Pressure psi	Viscosity at the bubble-point pressure, cP
940	2.6	255	3.2	800	2.6
1250	2.7	500	2.9	1558	0.9
1615	2.8	745	2.6	2514.8	0.9
2045	2.9	313	1.3	2975	0.5
2485	3.0	613	1.1	4400	0.4
2815	3.2	913	1.0	3066	0.9
2018	0.9	1218	0.9		
2518	0.9	1508	0.9		
3023	0.9	185	1.3		
3523	1.0	695	1.1		
4023	1.0	1215	1.1		
5023	1.1	1665	1.0		
2615	0.9	2215	0.9		
3000	0.9	225	1.2		
3440	1.0	450	1.1		
3965	1.0	750	1.0		
4440	1.1	1220	0.7		
4940	1.1	1725	0.7		
3115	0.6	2215	0.6		
3540	0.6	2725	0.6		
4005	0.6	210	0.7		
4490	0.6	520	0.7		
4985	0.6	815	0.6		
4815	0.4	1305	0.5		
5140	0.4	1815	0.5		
5440	0.4	2315	0.5		
5715	0.4	2815	0.5		
5975	0.4	3290	0.4		
3210	0.9	3815	0.4		
		4330	0.4		
		225	2.3		
		575	1.8		
		860	1.6		
		1342	1.4		
		1840	1.2		
		2222	1.0		
		2807	0.9		