Original article

Estimation of minimum miscibility pressure of varied gas compositions and reservoir crude oil over a wide range of conditions using an artificial neural network model

Abouzar Choubineh[1], Abbas Helalizadeh[1], David A.Wood[2]*

[1]*Petroleum University of Technology, Ahwaz, Iran*
[2]*DWA Energy Limited, Lincoln, United Kingdom*

Abstract:
Minimum miscibility pressure (MMP) is a key variable for monitoring miscibility between reservoir fluid and injection gas. Experimental and non-experimental methods are used to estimate MMP. Available miscibility correlations attempt to predict the minimum miscibility pressure for a specific type of gas. Here an artificial neural network (ANN) model is applied to a dataset involving 251 data records from around the world in a novel way to estimate the gas-crude oil MMP for a wide range of injected gases and crude oil compositions. This approach is relevant to sequestration projects in which injected gas compositions might vary significantly. The model is correlated with the reservoir temperature, concentrations of volatile (C_1 and N_2) and intermediate (C_2, C_3, C_4, CO_2 and H_2S) fractions in the oil (Vol/Inter), C_{5+} molecular weight fractions in the oil and injected gas specific gravity. A key benefit of the ANN model is that MMP can be determined with reasonable accuracy for a wide range of oil and gas compositions. Statistical comparison of predictions shows that the developed ANN model yields better predictions than empirical-correlation methods. The ANN model predictions achieve a mean absolute percentage error of 13.46%, root mean square error of 3.6 and Pearson's correlation coefficient of 0.95. Sensitivity analysis reveals that injected gas specific gravity and temperature are the most important factors to consider when establishing appropriate miscible injection conditions. Among the available published correlations, the Yellig and Metcalfe correlation demonstrates good prediction performance, but it is not as accurate as the developed ANN model.

1. Introduction

1.1 Minimum miscibility pressure (MMP)

Oil production typically involves three stages of recovery: Primary, sceondary, and tertiary. During primary recovery, oil is produced from wells by either gas cap expansion, aquifer expansion, dissolved gas expansion, rock and fluid expansion, gravity drainage or a combination of these mechanisms. During secondary recovery external energy is introduced to maintain reservoir pressure or, at least, to slow down reservoir pressure depletion. Typically, the oil recovery factor reaches some 35% or higher after the combined effects of both primary and secondary recovery processes, depending upon reservoir and fluid properties. Tertiary or enhanced oil recovery methods (EOR) are used to mobilize at least some of the 65% or so of oil remaining in a reservoir and usually commences after many years of primary and secondary recovery from mature oil fields.

EOR techniques are increasing applied to mature oil fields, and in some cases EOR mechanisms are being instigated much earlier in a field's development cycles in order to gain maximum benefits from them. There are many different EOR techniques now applied and are commonly categorized into four groups: Chemical, thermal, microbial and gas flooding. CO_2 is the main fluid which is used in gas flooding. Other common fluids are hydrocarbon gas such as methane (often enriched with quantities of light natural gas liquids -NGLs, e.g., ethane (C_2), propane (C_3 and butane (C_4)), nitrogen and exhaust or flue gases.

The miscibility of injected gas, of whatever composition, and crude oil in a sub-surface reservoir, plays a considerable role in the success of gas injection projects. A key metric

to establish in such projects is minimum miscibility pressure (MMP). It is the lowest pressure at miscibility is established between gas and oil at a specified constant temperature. Crucially at MMP, the interfacial tension is nearly zero and no interface exists between oil trapped in the reservoir and the injected/displacing fluid. If the crude oil in a reservoir and the injection gas form a single phase as a consequence of their initial contact, they are referred to as first contact miscible (FCM) (Green and Willhite, 1998). FCM fluids are relatively unusual. More commonly injected gas and crude oil develop miscibility over time by progressive mass transfer between the phases. In such cases they are referred to as multi-contact miscible (MCM) (Green and Willhite, 1998). MCM is associated with various reservoir drive mechanisms (e.g., vaporizing gas, condensing gas, and the combination of condensing and vaporizing gas) (Zick, 1986).

1.2 Methods for determining MMP

There are different approaches to estimating/predicting MMP, which can be grouped into two categories: Experimental and non-experimental methods. MMP is often established through slim-tube experiments (Yellig and Metcalfe, 1980) and represent the standard procedure for its determination. However, other experimental techniques are used to determine MMP. These include: Rising bubble/falling drop method (Zhang et al., 2018); mixing-cell method (Stalkup, 1983), multiple-mixing cell method (Ahmadi and Johns, 2011), interface laser light scattering spectroscopy (ILLS) (Dorshow, 1995), surface laser light scattering spectroscopy (SLLS) (Dorshow, 1995); and, vanishing interfacial tension (VIFT) method (Rao, 1997; Ahmad et al., 2016). Experiments of this type are time-consuming and costly; therefore, it is beneficial where possible to employ accurate mathematical (non-experimental) determinations of MMP.

There are five distinct categories of non-experimental methods to estimate MMP (Ahmadi-Rahmatabadi, 2011). These are: Empirical correlations (see section 2); ternary representation/limiting tie-lines (Whitson and Michelsen, 1989); slim-tube simulations (El-Sharkawy, 1992), mixing-cell models (Yuan, 2003) and analytical methods (Orr, 2007).

Ahmed (2000) applied the Peng Robinson equation of state (EoS) to estimate MMP, introducing a miscibility function to aid in the prediction of the required injection pressure to achieve miscible gas injection. Fazali et al. (2013) combined random-mixing rules with the Peng-Robinson (PR) EoS to approximate the phase behavior of reservoir fluids (crude oil and gas) and to predict MMP for reservoir oil. Yuan and Johns (2005) proposed an analytical model from predicting MMP based upon characteristic theory. That approach decreased the number of equations and unknown variables involved in the calculation of MMP and helped to avoid false solutions. Yuan et al. (2005) developed a MMP prediction formula for pure and impure CO_2 injection gases applying equations of state. The input parameters for that formula were: Reservoir temperature, C_{7+} molecular weight of oil, intermediate hydrocarbon molecules (C_2 to C_6) in the crude oil expressed as a

percentage. The objective of that model was to be effective across a board range of the input variables.

In recent years, several artificial intelligence and optimization methods have been applied to MMP prediction, but these have been focused on relatively specific and narrow ranges of injection gas compositions. Huang et al. (2003) developed an ANN model to estimate CO_2 MMP of crude oils for pure and impure gas compositions. They demonstrated that their ANN model for CO_2 MMP prediction could distinguish the effects of different contaminants within the CO_2 gases, and the levels of contaminants that could be tolerated for miscible injection. Emera and Sarma (2005) demonstrated that a model applying a genetic algorithm could provide more accurate predictions of CO_2 MMP than empirical correlations. That model involved inputs of reservoir temperature, C_{5+} molecular weight of oil and the ratio of volatiles (C_1 and N_2) to intermediates (C_2-C_4, H_2S, and CO_2) in oil.

Shokir (2007) proposed an alternating conditional expectation optimization algorithm for improving MMP prediction of impure and pure CO_2 injected gases, validating the model with experimental analysis. Dehghani et al. (2008) developed a hybrid genetic algorithm-ANN model for predicting MMP applying it successfully to both CO_2 and natural gas streams. They demonstrated that it performed better than the conventional empirical equations particularly under conditions of limited field-specific data. In a similar approach Nezhad et al. (2011) successfully combined a feed-forward neural network with a particle swarm optimization (PSO) algorithm to predict CO_2 MMP.

Shokrollahi et al. (2013) used 147 data points from the literature to develop and evaluate a Least-Squares Support Vector Machine (LSSVM) algorithm to predict pure and impure CO_2 MMP. They considered outlier diagnosis for their dataset in order to identify unreliable data records. Ahmadi et al. (2015) also employed LSSVM combined with various evolutionary optimization algorithms (e.g., genetic algorithm, particle swarm optimization, and imperialist competitive algorithm) to develop and compare models to predict MMP. They demonstrated that a hybrid GA-PSO-LSSVM produced the most accurate results of the models they tested. The results showed that the new models had an appropriate precision. Zhong and Carr (2016) proposed a mixed-kernels function (MKF) based support vector regression model (MKF-SVR) model for predicting CO_2 MMP, demonstrating its superior performance compared to conventional empirical correlations. Tarybakhsh et al. (2018) applied a genetic algorithm to optimize their SVR model to achieve improved MMP predictions. Choubineh et al. (2016) combined ANN with a cuckoo optimization algorithm and a teaching learning-based optimization methodology to predict CO_2 MMP with a high level of accuracy. Ahmadi et al. (2017) employed gene expression programming and Alomair et al. (2015) adopted n alternating conditional expectation algorithm to predict MMP.

The large number of machine learning methodologies developed in recent years highlights the significance of achieving reliable MMP predictions. What the previously published artificial intelligence and optimization methodologies demonstrate convincingly is that when applied to a relatively narrow

range of injected gas compositions and/or reservoir conditions they outperform the conventional empirical correlations in predicting MMP. What has not been demonstrated is that such techniques can provide sufficiently accurate predictions of MMP when applied to a much wider range of injection gas compositions and crude oil reservoir conditions.

1.3 Study objectives

Here we develop and evaluate an artificial neural network (ANN) model to predict the gas-crude oil MMP for a dataset (251 data records from multiple published sources associated with crude oil reservoirs from around the world) including a wide range of injection gas compositions. This approach is relevant to situations where gases of varied compositions (e.g., flue gas collected from various industrial plants) are injected into oil reservoirs for the purposes of sequestration and enhanced oil recovery. As the injected gas composition varies over a wide range from time to time it is useful to be able to predict MMP for a wide range of potential injected gas compositions. We also compare this ANN model with experimental data and available literature correlations to demonstrate its superior performance. In addition, using a purpose-built Monte Carlo simulation sensitivity model, we evaluate the effects and significance of key independent metrics (i.e., reservoir temperature, concentration of volatile and intermediate fractions in oil, C_{5+} molecular weight of oil and injected gas specific gravity) on the accuracy of predicting the gas-crude oil MMP (i.e., the dependent variable) across such a wide range of conditions.

2. Empirical correlations

Most empirical correlations used historically for predicting MMP involve regressions applied to experimental data (Yellig and Metcalfe, 1980). Published miscibility correlations belong to two distinct groups: Correlations related to MMP for pure CO_2 or gas mixtures involving CO_2; correlations related to MMP for other gas compositions. These correlations usually require data inputs for just a few related metrics, e.g., reservoir temperature, reservoir crude oil properties, injection fluid properties.

Some of the more widely-applied correlations are considered here for comparison with our proposed analytical method.

2.1 Cronquist correlation (1978)

Cronquist proposed the following regression (Eqs. (1) and (2)) to predict MMP:

$$MMP = 0.11027 \times (1.8T + 32)^Y \qquad (1)$$

$$Y = 0.744206 + (0.0011038MW_{C5+}) + (0.0015279Vol) \quad (2)$$

where T is reservoir temperature (°C); MW_{C5+} is the molecular weight of the pentanes-plus fraction of crude oil; and, Vol is the mole percentage of volatile components (C_1 and N_2).

2.2 Lee correlation (1979)

Based on experimental data, Lee proposed an empirical equation for predicting CO_2 MMP. This correlation (Eqs. (3) and (4)) depends only on reservoir temperature as the input variable:

$$MMP = 7.3942 \times 10^b \qquad (3)$$

$$b = 2.772 - \frac{1519}{492 + 1.8T} \qquad (4)$$

where T is reservoir temperature (°C).

A key constraint in applying the Lee correlation is that if the predicted MMP is less than the bubble point pressure (Pb), then the CO_2 MMP = Pb.

2.3 Yellig & Metcalfe correlation (1980)

Yellig and Metcalfe derived a correlation (Eq. (5)) for estimating the CO_2 MMP based on experimental data, also depending only on reservoir temperature as the input variable:

$$MMP = 12.6472 + 0.015531 \times (1.8T + 3.2)$$
$$+ 0.000124192 \times (1.8T + 32)^2 - \frac{716.9427}{1.8T + 32} \qquad (5)$$

where T is reservoir temperature (°C).

Again, the key constraint in applying the Yellig & Metcalfe correlation is that if the predicted MMP is less than Pb, then the CO_2 MMP = Pb.

2.4 Alston et al. correlation (1985)

Alston et al. derived an empirical correlation (Eqs. (6) and (7)) to estimate MMP for pure CO_2 or gas mixtures involving CO_2; the pure or impure CO_2 MMP. The input metrics are: Reservoir temperature, molecular weight of the pentane-plus fraction of the oil, the mole fraction of the oil intermediate components (i.e., natural gas liquids C_2 to C_4, carbon dioxide and hydrogen sulfide) and the mole fraction of volatile oil components (i.e., methane and nitrogen).

$$MMP = 6.056 \times 10^{-6} \times (1.8T + 32)^{1.06}$$
$$\times MW_{C5+}^{1.78} \times \left(\frac{Vol}{Interm} \right)^{0.136} \qquad (6)$$

When bubble point pressure (Pb) < 0.345 MPa:

$$MMP = 6.056 \times 10^{-6} \times (1.8T + 32)^{1.06} \times MW_{C5+}^{1.78} \qquad (7)$$

where T is reservoir temperature (°C); MW_{C5+} is the molecular weight of the oil pentanes-plus fraction; Vol is the mole percentage of volatile (C_1 and N_2); and, $Interm$ is the mole fraction of the oil intermediate components (C_2-C_4, CO_2 and H_2S).

In order to take into account of the effects of impurities (i.e., the presence of C_1, C_2, C_3, C_4, N_2 or H_2S gases) in the CO_2 gas to be injected, the impure CO_2 MMP is correlated

Table 1. SF_i for different components used in the Dong Correlation.

Component	SF_i
H_2S	0.7
C_1	2.5
N_2	7.5
SO_2	0.5
O_2	5

(Eqs. (8) and (9)) with the weighted-average pseudo-critical temperature T_{CM} of the injected gas and the pure CO_2 MMP:

$$\frac{MMP_{impure}}{MMP_{pure}} = \left(\frac{87.8}{1.8T_{CM} + 32} \right)^{\left(\frac{1.935 \times 87.8}{1.8T + 32} \right)} \tag{8}$$

$$T_{CM} = \sum_{i=1}^{n} W_i \times T_{ci} \tag{9}$$

where T_{CM} is the weighted-average pseudo-critical temperature (°C).

The critical temperatures used for each component gas in Eqs. (8) and (9) are the true critical temperatures except for H_2S and CO_2, which are both assigned a uniform value of 51.85 °C.

2.5 Dong correlation (1999)

Dong proposed a correlation (Eqs. (10) and (11)) for predicting the impure CO_2 MMP involving the mole average critical temperature T_{ac} as the input variable:

$$\frac{MMP_{impure}}{MMP_{pure}} = \left(\frac{T_{ac}}{304.2} \right)^4 \tag{10}$$

$$T_{ac} = \sum_{i=1}^{n} X_i \times SF_i \times T_{ci} \tag{11}$$

where SF_i is the strength of species "i" in changing the apparent critical temperature of the flue gas relative to the critical temperature of CO_2. The SF_i values for various gases are listed in Table 1. For CO_2 and other components not listed in Table 1, SF_i is equal to 1. T_{ci} is the critical temperature of gas component "i" (°K); and, X_i is the mole fraction of gas component "i".

2.6 Hudgins et al. correlation (1990)

Hudgins et al. derived a correlation for pure N_2 MMP (Eqs. (12) to (14)) based on an experimental study of nitrogen miscible flooding associated with an EOR application for light crude oil:

$$MMP = 5568 \times e^{-R_1} + 3641 \times e^{-R_2} \tag{12}$$

$$R_1 = 792.06 \times \left(\frac{C_2 - R_5}{W} \right) \tag{13}$$

$$R_2 = 2.158 \times 10^6 \times \left(\frac{C_1^{5.632}}{W} \right) \tag{14}$$

$$W = MW_{C7+} \times T^{0.25} \tag{15}$$

where T is the temperature (°F); C_1 is the mole fraction of methane; and, C_2-C_5 is the mole fraction of C_2-C_5.

3. Methodology

Both experimental methods and empirical correlations for estimating MMP have their limitations. Experimental methods are expensive and time-consuming. Empirical correlations generally tend to work well for the specific oil reservoirs and conditions from which their underlying dataset is derived; they are typically unable to be applied globally across all types of crude oil reservoir (Yellig and Metcalfe, 1980). Therefore, it is beneficial to develop analytical models to calculate MMP for various injection gases under a wide range of conditions. Various machine learning and artificial intelligence techniques lend themselves to the type of analysis required. Artificial neural networks (ANN) are one such method.

3.1 Overview of artificial neural network (ANN)

ANN represents a machine learning tool well suited to the prediction of variables dependent on complex, often poorly defined, nonlinear functions that define their relationships with a number of independent or variably correlated input variables for which measurements are available (Krenker et al., 1995). Back propagation (BP) is well-established ANN training algorithm. There are two pathways that constitute the typical ANN; the feed-forward pathway and the backward pathway. The feed-forward pathway is configured when setting up a feed-forward ANN. This involves defining the initial weights, which are varied through various iterations of a simulation involving BP training of a neural network (Krenker et al., 1995). The network weights/biases are progressively adjusted during back-propagation training, leading to the gradual improvement of the networks prediction accuracy. Training requires a number of data records from a dataset for which both input metric values and associated dependent variable values (i.e., gas-crude oil MMP) are known (i.e., measured experimentally). As training progresses, the weights applied to nodes within the ANN are sequentially adjusted to minimize an objective function (e.g., the mean square error (MSE) of the estimated versus measured value of the dependent variable, in this case gas-crude oil MMP).

A five-layer feed-forward ANN is illustrated in Fig. 1. This is feed-forward neural network (FNN) consists of three distinct types of processing layers; the input layer, output layer and three hidden layers between the input and output layers. Each layer consists of a number of nodes or neurons, each with an associated activation function. Fig. 1 consists of an input layer involving four neurons, three hidden layers with nine, ten and nine neurons, respectively, and an output layer with a single neuron. It is possible to vary the number of neurons in the input and hidden layers and the number of hidden layers to
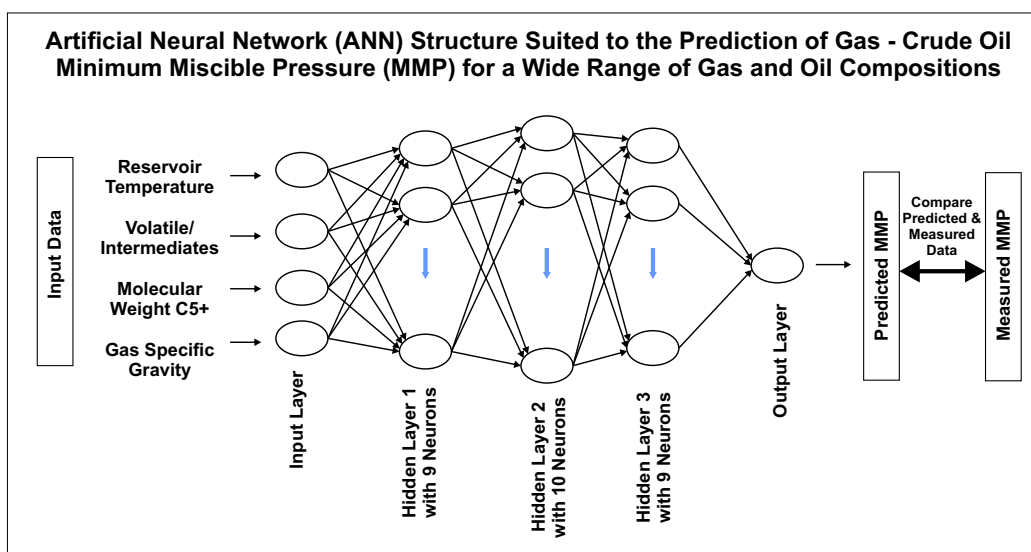
Fig. 1. Schematic of multilayer feed-forward neural network (FNN) applied in this study.

suit the complexity of the model and number of input variables involved. Input variables are introduced to the network via specific nodes in the input layer. The input nodes are then each connected to the nodes of the first hidden layer. The nodes of the first hidden layer are connected to the nodes of the second hidden layer, and so on. The nodes of the third hidden layers are linked to the single node of the output layer which delivers the prediction of the dependent variable. The mathematics of the algorithms involved in FNN applying a BP training algorithm are now well documented and widely applied (Goh, 1995; Al-Alawi et al., 2005; Zarenezhad and Aminian, 2011) and are therefore not repeated here. Clearly, the results obtained using FNN can only be as accurate as the measurements of the input variables in the underlying training data records. Careful scrutiny of the input data quality is therefore essential if accurate predictions are to be produced via the FNN method.

3.2 Data collection

For the ANN model proposed the following input variables are involved in gas-crude oil MMP prediction: Reservoir temperature, concentration of volatile components (C_1 and N_2) in the injected gas; concentrations of intermediate (C_2, C_3, C_4, CO_2 and H_2S) fractions in oil; C_{5+} molecular weight of oil; and, injected gas specific gravity.

The dataset used here involves 251 data records compiled from 32 published studies (Rathmell et al., 1971; Jacobson, 1972; Dicharry et al., 1973; Holm and Josendal, 1974; Yellig and Metcalfe, 1980; Cobb and Goodrich, 1981; Gardner et al., 1981; Graue and Zana, 1981; Frimodig et al., 1983; Cardenas et al., 1984; Alston et al., 1985; Glaso, 1985; Sebastian et al., 1985; Firoozabadi and Aziz, 1986; Hanssen, 1988; Hand and Pinczewski, 1990; Hudgins et al., 1990; Sebastian and Lawrence, 1992; Thomas et al., 1994; Yurkiw and Flock, 1994; Harmoon and Grigg, 1998; Srivastava and Huang, 1998; Rao

Table 2. Minimum and maximum values of the data variables for the 251 data records in the dataset.

Data	Unit	Min	Max
Reservoir temperature (Input)	degrees fahrenheit	90	340
Vol/Inter (Input)	unitless	0	37.84
MW_{C5+} (Input)	lb/lbmol	113.5605	302.5
Injected gas specific gravity (Input)	unitless	0.556067	1.574349
MMP (Output)	psia	948	9115.622

and Lee, 2000; Srivastava, 2000; Dong et al., 2001; Jaubert et al., 2002; Ayirala, 2005; Bon et al., 2005; Bon et al., 2006; Al-Netaifi, 2008; Ren et al., 2011; Adekunle, 2014) with each record containing data for the input variables selected. The dataset includes: Sixty-seven records involving pure carbon dioxide, thirty-six records involving pure nitrogen, eight records involving pure methane, and one record involving pure ethane. The remaining data records in the data set are impure gas (i.e., mixtures of two or more gases). Due to the diversity of gases represented in the dataset the, specific gravity of the gases relative to air ranges from 0.55 to 1.57. The diversity and number of data records involved in the compiled dataset are significant. They set this study apart from previously published machine learning attempts to predict gas-crude oil MMP. The details of the data set are provided in Appendix A and the individual data records, their sources and gas compositions are provided in a supplementary data file.

The data records of the dataset are divided, 75%, 10% and 15% into training, validation and testing subsets, respectively. The ranges of values represented by the input and output variables in the full dataset are listed in Table 2. These values

Table 3. The FNN structure and adjustment/transfer functions applied to the gas-crude oil MMP dataset. "logsig" refers to log-sigmoidal transfer function and "purelin" refers to a transfer function.

Parameter	Value
Number of neuron in hidden layer 1	9
Number of neuron in hidden layer 2	10
Number of neuron in hidden layer 3	9
Activation function of input-hidden layer 1	logsig
Activation function of hidden layer 1-hidden layer 2	purelin
Activation function of hidden layer 2-hidden layer 3	purelin
Activation function of hidden layer 3-output layer	purelin

highlight the wide ranges of input-variable values and gas-crude oil MMP (output) values covered by this comprehensive dataset.

3.3 ANN model developed to predict MMP

The data for each variable in each data record (Table 2) is normalized to values between 0 and 1 using the minimum and maximum values. Using normalized data in ANN avoids the introduction of biases associated with the different scales and units of the variables involved. By testing different numbers of hidden layers and neurons in the network, a range of adjustment/transfer functions are assigned to the input-hidden, hidden-hidden-hidden and hidden-output layers, resulting in different prediction accuracies for the gas-crude oil MMP. Table 3 shows the network structure established for this dataset based upon those tests.

4. Results and discussion

4.1 Gas-crude oil MMP analysis based on ANN model

Results of the ANN model developed for the prediction of gas-crude oil MMP applied to the training, validation and testing subsets of the dataset (as described in section 3) are illustrated in Figs. 2 to 4. These figures compare the calculated MMP (output/prediction) values with the actual measured MMP values for the dataset.

The predicted versus measured MMP values displayed in Figs. 2 to 4 are significantly correlated, and the linear correlation line (red) positioned close to the $y = x$ line. Moreover, the Pearson's correlation coefficients for the training, validation and testing subsets are 0.96, 0.93 and 0.93, respectively. This indicates that the prediction performance of the ANN model is acceptable, even though it is applied to such a diverse dataset in terms of injected gas compositions. Nevertheless, there are some significant outlier points which the proposed model is incapable of predicting sufficient preciseness, most of them are related to high minimum miscibility pressures. The main reason for such outlying predictions is that there are a low number of data records available for reservoir pressures greater

than 6000 psia. Another contributing factor to these outliers is

Table 4. Statistical analysis of the accuracy of MMP prediction by six empirical models and the ANN model proposed in this study applied to those data records relevant to the specific correlation range. Full data set of 251 data records applies to the ANN model only.

Prediction methods	MAPE	RMSE	R
Cronquist	21.31	5.51	0.6669
Lee	20.41	5.86	0.6881
Yellig and Metcalfe	14.92	4.46	0.7302
Alston	27.55	5.99	0.6254
Dong	35.40	7.15	0.4766
Hudgins	49.04	18.93	0.5754
ANN Model	13.46	3.60	0.9500

the wide range of values associated with each input variable across the entire dataset (Table 2).

Figs. 5 to 7 illustrate predicted and measured MMP versus data index. These trends demonstrate that overall the ANN model achieves a high level of accuracy in its MMP prediction. The outlier data records (i.e., greatest mis-match between predicted and measured MMP values) that do exist are spread across the data range in each of the data subsets.

4.2 Comparison of ANN model and empirical correlations

The prediction performance comparison of the ANN model developed here compared with the six empirical correlations described in section 2 are assessed using the standard statistical accuracy metrics: Pearson's correlation coefficient (R); root mean square error (RMSE); and, mean absolute percentage error (MAPE). As these empirical correlations were each developed for quite specific conditions (i.e., gas types) it is not surprising that they do not perform well when applied to the diverse dataset evaluated in this study. Indeed, as shown in Table 4 and Figs. 8 to 10, the proposed ANN model performs much better than the empirical correlations
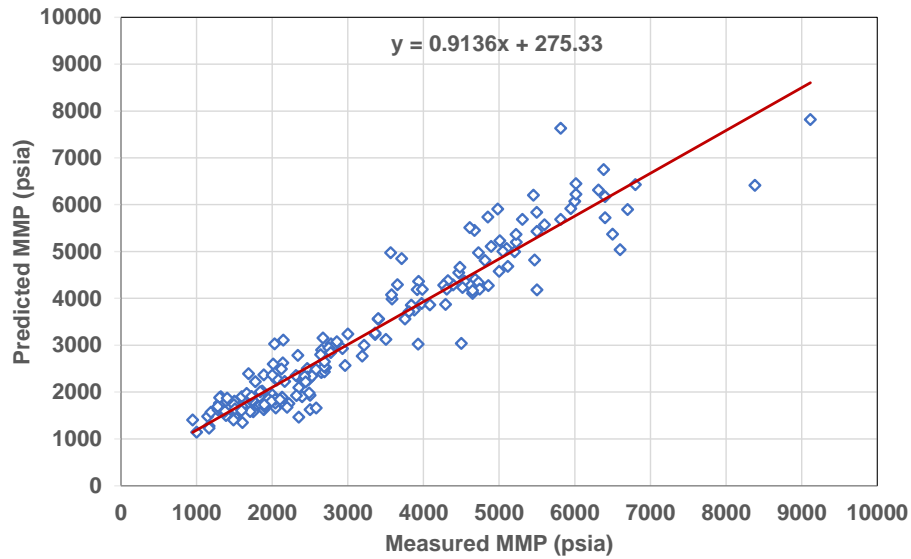
Fig. 2. ANN model MMP prediction versus measured MMP for the 188 data records of the training subset of the database.
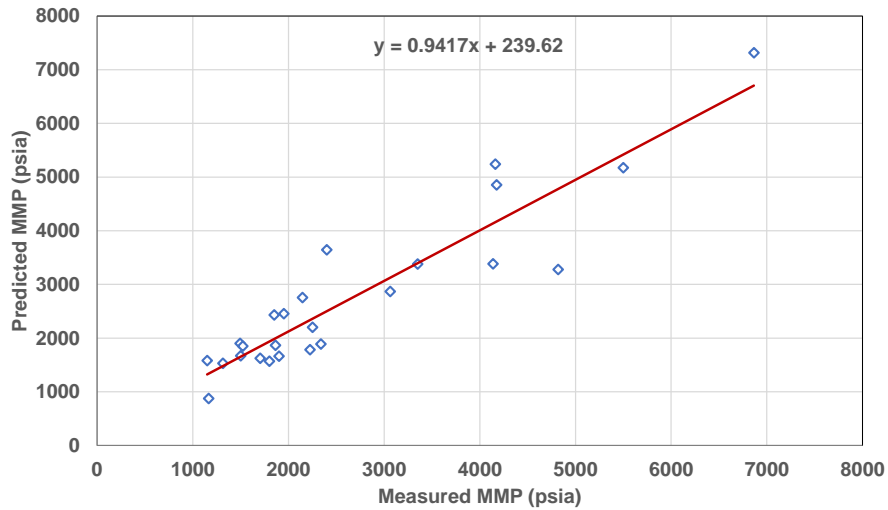

Fig. 3. ANN model MMP prediction versus measured MMP for the 25 data records of the validation subset of the database.
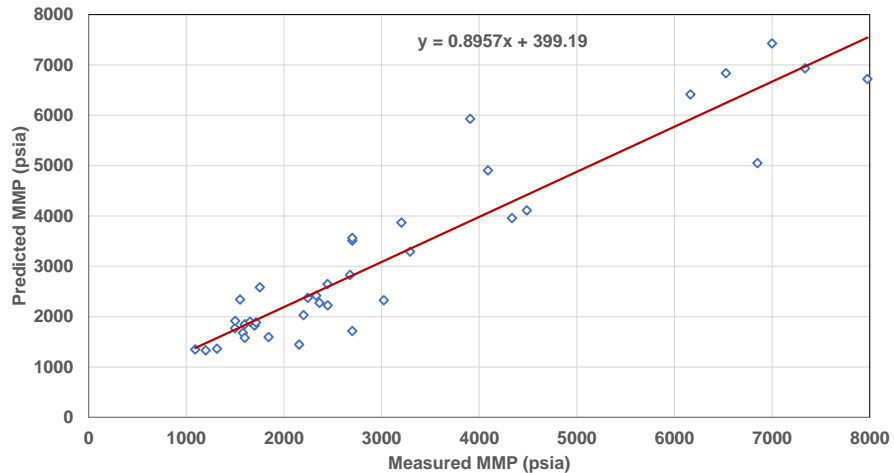

Fig. 4. ANN model MMP prediction versus measured MMP for the 38 data records of the testing subset of the database.
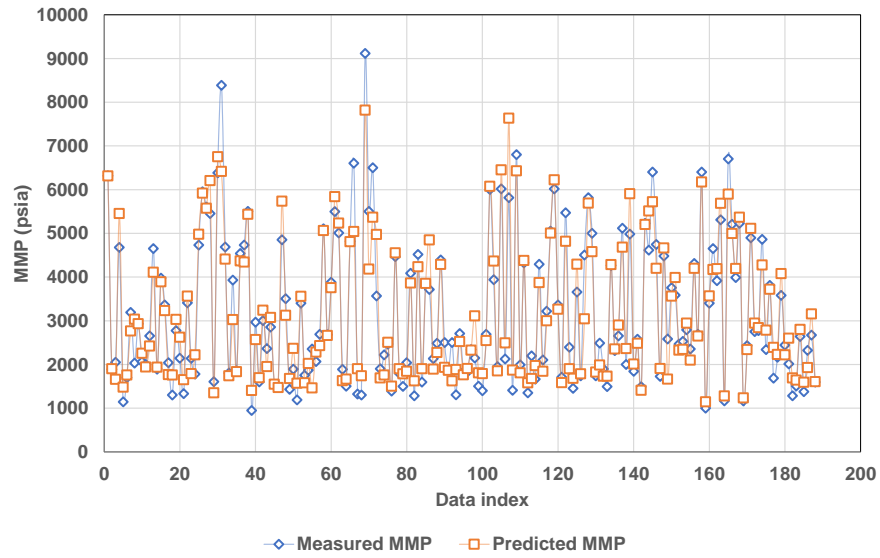
Fig. 5. Comparison of predicted versus measured MMP for the 188 data records of the ANN training subset displayed according to their data index number.
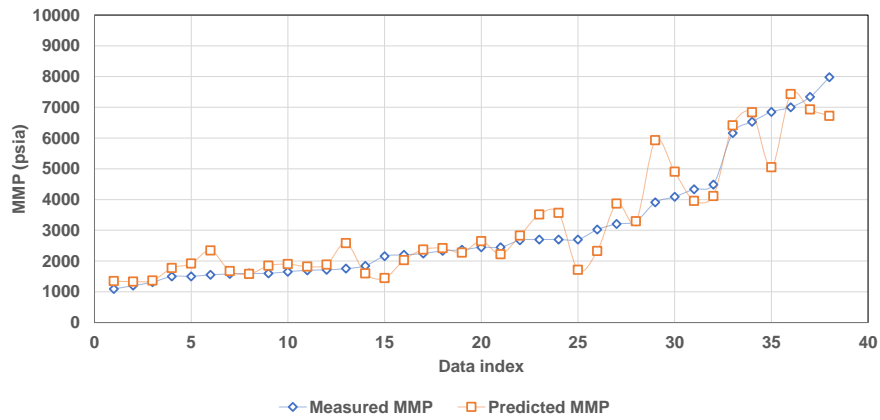

Fig. 6. Comparison of predicted versus measured MMP for the 38 data records of the ANN validation subset displayed according to their data index number.
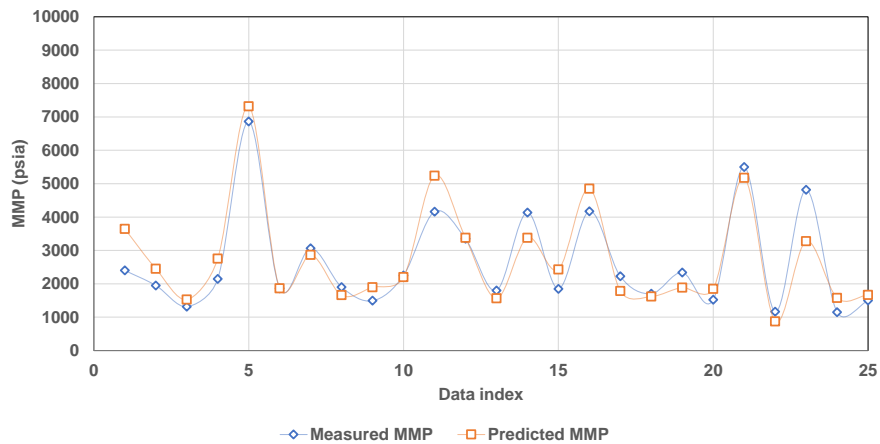

Fig. 7. Comparison of predicted versus measured MMP for the 25 data records of the ANN testing subset displayed according to their data index number.
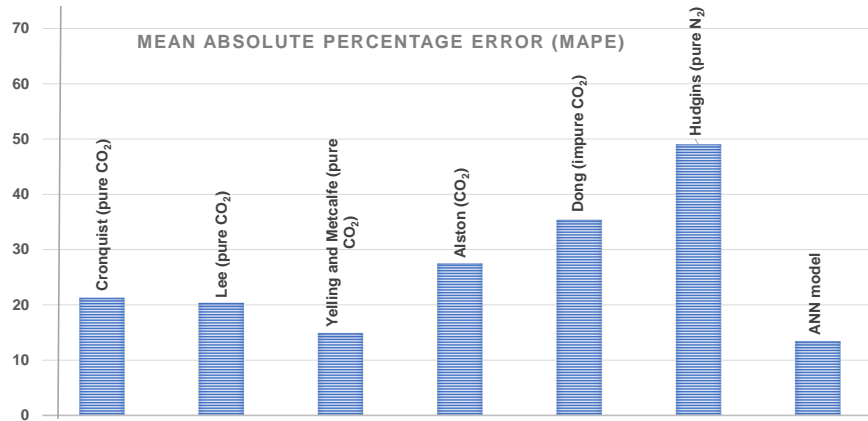
Fig. 8. MMP prediction performance in terms of MAPE of the ANN model compared to the six empirical correlations (section 2) applied to those data records of the dataset (with multiple gas compositions from published gas-crude oil MMP measurements) relevant to the specific correlation range.

Fig. 9. MMP prediction performance in terms of RMSE of the ANN model compared to the six empirical correlations (section 2) applied to those data records of the dataset (with multiple gas compositions from published gas-crude oil MMP measurements) relevant to the specific correlation range.
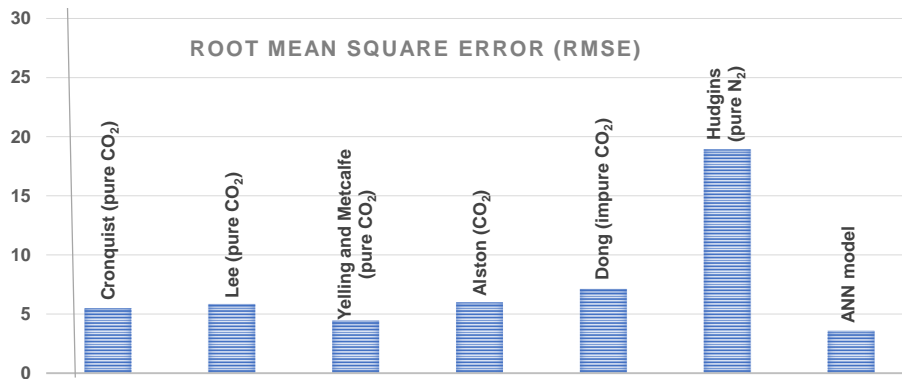

Fig. 10. MMP prediction performance in terms of Pearson's correlation coefficient of the ANN model compared to the six empirical correlations (section 2) applied to those data records of the dataset (with multiple gas compositions from published gas-crude oil MMP measurements) relevant to the specific correlation range.
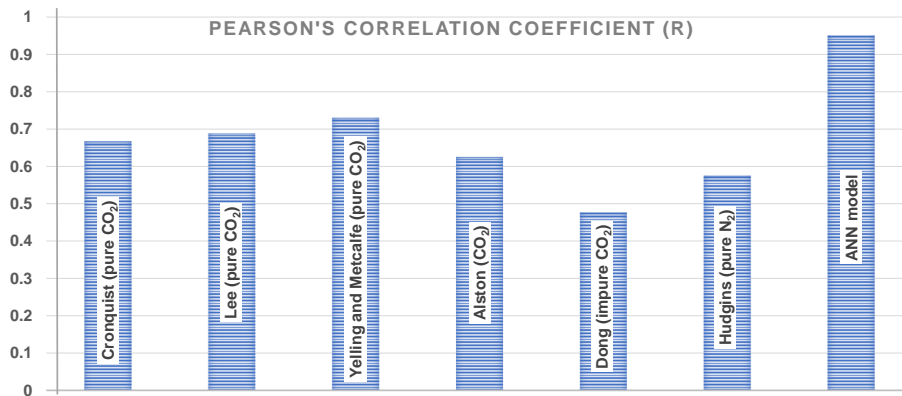
when applied to this diverse dataset. The Yellig & Metcalfe empirical correlation performs better than the other empirical correlations when applied to this diverse dataset, but not as well as the ANN model developed here.

At first inspection of the results it may seem surprising that the Yellig and Metcalfe (1980) correlation which uses only one input variable (reservoir temperature) has an average absolute error of 14.9% compare to 13.5% for the ANN model developed here, which uses five input variables. However, the Yellig and Metcalfe (1980) correlation is only used for pure CO_2 data records in our dataset, whereas the ANN model is applied to all 251 data records. Moreover, the ANN model outperforms the Yellig and Metcalfe (1980) correlation for the three-error metrics considered: For the Yellig and Metcalfe (1980) correlation, MAPE, RMSE, and R are 14.92, 4.46 and 0.73, respectively. On the other hand, for the ANN model MAPE, RMSE, and R are 13.46, 3.6, and 0.95, respectively.

Table 4 highlights that the proposed ANN model has the smallest MAPE (13.46%), RMSE (3.6%) and highest R (0.95) of the models evaluated when applied to the diverse MMP dataset.

4.3 Sensitivity analysis applied to the ANN prediction model

A Monte Carlo simulation was performed to provide sensitivity analysis of the proposed ANN model for predicting MMP for a wide range of injected gas compositions. The focus of the sensitivity analysis is to establish the individual and collective impacts of the four input variables on the MMP predictions. By running thousands of iterations of the simulation selecting values for each input variable (normalized on a scale of 0 to 1) using random numbers between 0 and 1 and monitoring the impact on the predicted MMP values it is possible to identify the dependence of the dependent variable (gas-crude oil MMP) on each of the input variables (i.e., reservoir temperature, concentration of volatile and intermediate fractions in oil, C_{5+} molecular weight of oil and injected gas specific gravity).

Fig. 11 illustrates the impact of each input variable individually on the MMP prediction, revealing that these impacts are highly non-linear and distinct for each input variable.

There is a direct positive relationship between MMP and reservoir temperature (Fig. 11a), highlighting that as reservoir temperature rises MMP increases. The relationships between the other three input variables and MMP are more complex. In the lower half of their value ranges the concentrations of volatile and intermediate fractions in oil (Fig. 11b) and C_{5+} molecular weight of oil (Fig. 11c) both demonstrate positive relationships with the predicted MMP. However, in the upper half of their value ranges these two input variables both show a reversal of that relationship, albeit determined by relatively few data points. On the other hand, the specific gravity of the injected gas (Fig. 11d) shows a negative relationship with predicted MMP over the upper 70% of its value range versus a moderate positive relationship over the lower 30% of its value range. Of course, the lowest values in the specific gravity

range evaluated are related to methane, the intermediate values belong mainly to nitrogen and the highest values are for impure and pure carbon dioxide.

Fig. 12 displays the results of the simulation focused on the relative impacts of the four input variables on predicted gas-crude oil MMP. The lower graph reveals that injected gas specific gravity has the most significant impact on the MMP prediction, followed by reservoir temperature. The other two variables clearly contribute to a lesser degree to the MMP prediction. The upper graph (Fig. 12) expresses these relative contributions in percentage terms. Clearly, if the ANN methodology is applied to a much narrower range of gas compositions, as has been the case in previous machine learning methods applied to MMP prediction, reservoir temperature would be the most influential input variable on the MMP predictions over that limited range.

5. Conclusions

A key parameter in miscible gas flooding of crude oil reservoirs is the minimum miscibility pressure (MMP). The ANN model developed in this study successfully estimates the gas-crude oil MMP for a diverse dataset (251 data records) encompassing a wide range of injected gas compositions (specific gravity relative to air ranging from 0.55 to 1.57) and reservoir crude oil compositions. This novel model is based on four distinct input variables: Reservoir temperature; concentrations of volatile (C_1 and N_2) and intermediate (C_2, C_3, C_4, CO_2 and H_2S) fractions in oil; C_{5+} molecular weight fractions in the oil; and, injected gas specific gravity.

From the application of the ANN model developed to the diverse MMP dataset the following conclusions can be drawn:

1) A statistical comparison of the prediction accuracy of the proposed model and six other commonly used empirical correlations for MMP prediction reveals that the new ANN model achieved significantly more accurate predictions of MMP. ANN model has the smallest MAPE (13.46%), RMSE (3.6%) and highest R (0.95) of the models evaluated when applied to the diverse MMP dataset. These results confirm that meaningful predictions of MMP can be derived for a wide range of injected gas compositions by applying this approach, making it potential useful for sequestration-EOR projects using injected gas compositions that might vary over time.

2) Among the input parameters, simulation conducted for sensitivity analysis revealed that injected gas specific gravity and reservoir temperature are the input variables that have the most significant impact on the predicted MMP values by the ANN model. The volatile/intermediate fractions and C_{5+} molecular weight of oil input metrics have much less significant impacts on the predicted MMP values.

3) Comparing the MMP prediction performance of the published empirical correlations, based for the most part on experimental studies of a narrow range of injected gas compositions, when applied to the diverse MMP dataset evaluated here, it is the Yellig and Metcalfe correlation that performs better than the other empirical
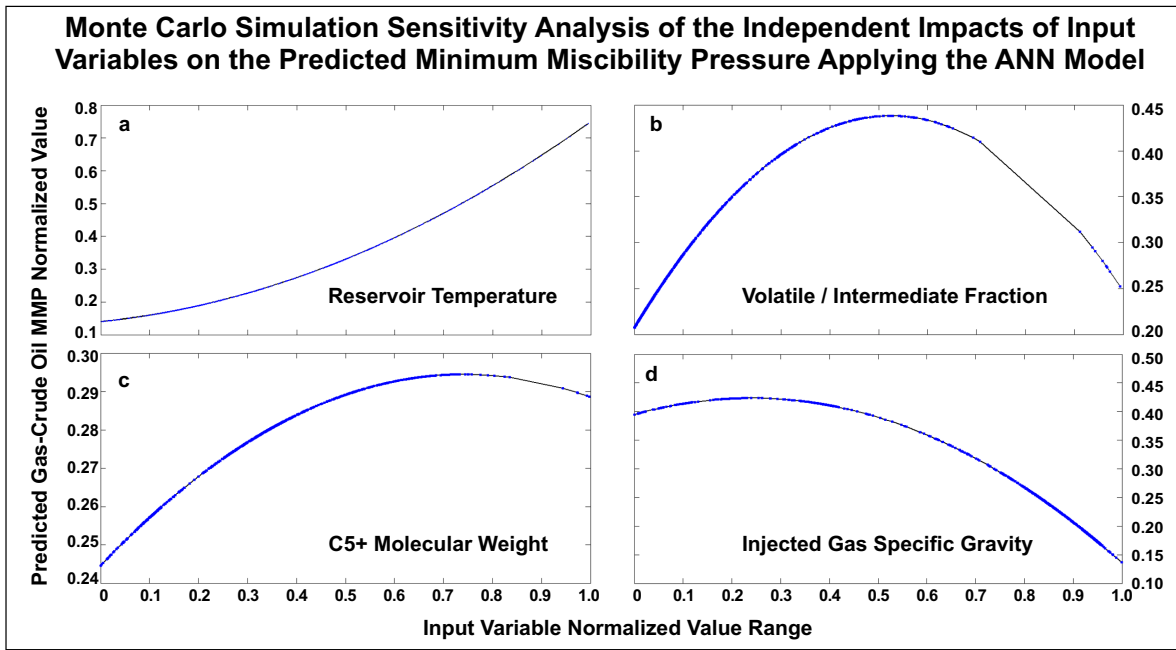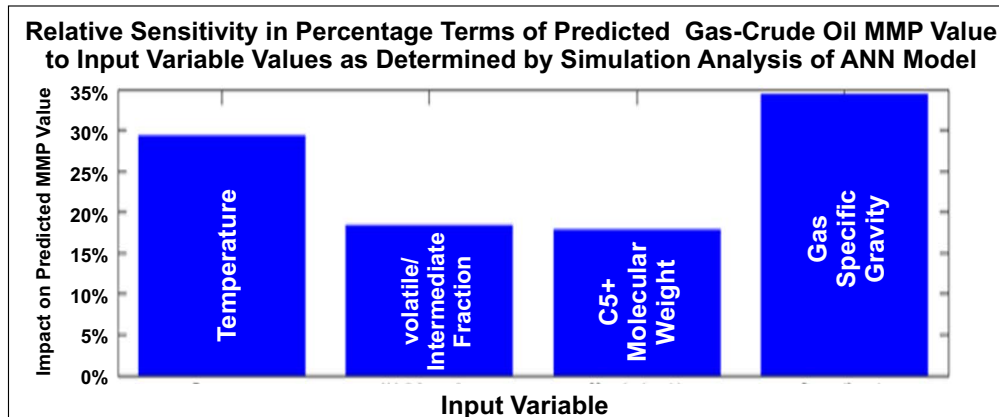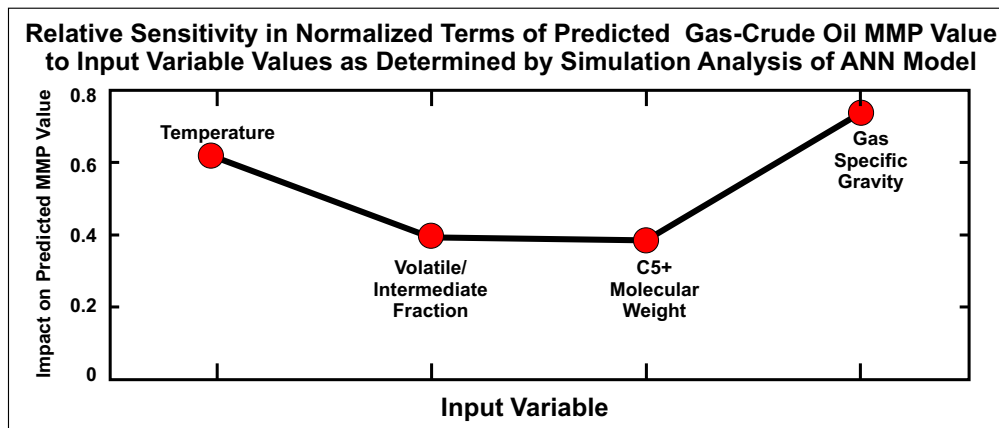
Fig. 11. Sensitivity of predicted gas-crude oil MMP individually to each input variable established by analyzing the results of multiple iterations of a Monte Carlo simulation applied to the ANN model.

(a)

(b)

Fig. 12. Relative sensitivity of the prediction of MMP by the ANN model to each input metric when applied to the diverse MMP dataset established for Monte Carlo simulation analysis.

correlations evaluated when applied to this wide range of gas compositions. The Yellig and Metcalfe correlation achieved a mean absolute percentage error of 14.92%, root mean square error of 4.46 and Pearson's correlation coefficient of 0.7302.

Nomenclature

ANN = Artificial Neural Network
BP = Backpropagation
EOR = Enhanced Oil Recovery
EoS = Equation of State
FNN = Feed Forward Neural Network
FCM = First Contact Miscibility
ILLS = Interface Laser Light Scattering Spectroscopy
LSSVM = Least-Squares Support Vector Machine
MAPE = Mean Absolute Percentage Error
MCM = Multi Contact Miscibility
MKF = Mixed-Kernels Function
MMP = Minimum Miscibility Pressure
MSE = Mean Square Error
PSO = Particle Swarm Optimization
PR = Peng-Robinson
R = Pearson's Correlation Coefficient
RMSE = Root Mean Square Error
SLLS = Surface Laser Light Scattering Spectroscopy
VIFT = Vanishing Interfacial Tension

Acknowledgments

References

Adekunle, O. Experimental approach to investigate minimum miscibility pressures in the Bakken. USA, Colorado School of Mines, 2014.

Ahmadi, K., Johns, R.T. Multiple-mixing-cell method for MMP calculations. SPE J. 2011, 16(4): 733-742.

Ahmadi, M.A., Zahedzadeh, M., Shadizadeh, S.R., et al. Connectionist model for predicting minimum gas miscibility pressure: Application to gas injection process. Fuel 2015, 148: 202-211.

Ahmadi, M.A., Zendehboudi, S., James, L.A. A reliable strategy to calculate minimum miscibility pressure of CO_2-oil system in miscible gas flooding processes. Fuel 2017, 208: 117-126.

Ahmadi-Rahmatabadi, K. Advances in calculation of minimum miscibility pressure. USA, University of Texas at Austin, 2011.

Ahmad, W., Vakili-Nezhaad, G., Al-Bemani, A.S., et al. Experimental determination of minimum miscibility pressure. Procedia Eng. 2016, 148: 1191-1198.

Ahmed, T. Minimum miscibility pressure from EOS. Paper

PETSOC-2000-001 Presented at the Canadian International Petroleum Conference, Calgary, Alberta, 4-8 June, 2000.

Al-Alawi, S., Al-Badi, A., Ellithy, K. An artificial neural network model for predicting gas pipeline induced voltage caused by power lines under fault conditions. COMPEL Int. J. Comput. Math. Electr. Electron. Eng. 2005, 24(1): 69-80.

Al-Netaifi, A.S. Experimental investigation of CO_2-miscible oil recovery at different conditions. Saudi Arabia, King Saud University, 2008.

Alomair, O., Malallah, A., Elsharkawy, A., et al. Predicting CO_2 minimum miscibility pressure (MMP) using alternating conditional expectation (ACE) algorithm. Oil Gas Sci. Technol. 2015, 70(6): 967-982.

Alston, R.B., Kokolis, G.P., James, C.F. CO_2 minimum miscibility pressure: A correlation for impure CO_2 streams and live oil systems. SPE J. 1985, 25(2): 268-274.

Ayirala, S.C. Measurement and modeling of fluid-fluid miscibility in multicomponent hydrocarbon systems. USA, Louisiana State University, 2005.

Bon, J., Emera, M.K., Sarma, H.K. An experimental study and genetic algorithm (GA) correlation to explore the effect of nC_5 on impure CO_2 minimum miscibility pressure (MMP). Paper SPE 101036 Presented at the SPE Asia Pacific Oil & Gas Conference and Exhibition, Adelaide, Australia, 11-13 September, 2006.

Bon, J., Sarma, H.K., Theophilos, A.M. An investigation of minimum miscibility pressure for CO_2-rich injection gases with pentanes-plus fraction. Paper SPE 94357 Presented at the SPE International Improved Oil Recovery Conference, Kuala Lumpur, Malaysia, 5-6 December, 2005.

Cardenas, R.L., Alston, R.B., Nute, A.J., et al. Laboratory design of a gravity-stable miscible CO_2 process. J. Pet. Technol. 1984, 36(1): 111-118.

Choubineh, A., Mousavi, S.R., Vafaee Ayouri, M., et al. Estimation of the CO_2-oil minimum miscibility pressure for enhanced oil recovery. Pet. Sci. Technol. 2016, 34(22): 1847-1854.

Cobb, L., Goodrich, J. Target reservoirs for CO_2 miscible flooding. Houston, Texas, USA, National Energy Technology Laboratory, 1981.

Cronquist, C. Carbon dioxide dynamic displacement with light reservoir oils. Paper Presented at SPE Annual U.S. DOE Symposium, Tulsa, USA, 1978.

Dehghani, S.A.M., Sefti, M.V., Ameri, A., et al. Minimum miscibility pressure prediction based on a hybrid neural genetic algorithm. Chem. Eng. Res. Des. 2008, 82(2): 173-185.

Dicharry, R.M., Perryman, T.L., Ronquille, J.D. Evaluation and design of a CO_2 miscible flood project-SACROC unit, Kelly-Snyder field. J. Pet. Technol. 1973, 25(11): 1309-1318.

Dong, M. Minimum miscibility pressure (MMP) studies, in the technical report: Potential of greenhouse storage and utilization through enhanced oil recovery. Petroleum Technology Research Centre, Saskatchewan Research

Council, Regina, Saskatchewan, Canada, 1999.

Dong, M., Huang, S.S., Srivastava, R. A Laboratory study on near-miscible CO_2 injection in Steelman reservoir. J. Can. Pet. Technol. 2001, 40(2): 53-61.

Dorshow, B.R. The simultaneous measurement of interfacial tension and oil viscosity at reservoir conditions for Prudhoe Bay Fluids by surface laser light scattering spectroscopy. SPE Adv. Technol. Ser. 1995, 3(1): 120-128.

El-Sharkawy. Measuring minimum miscibility pressure slim-tube or rising-bubble method? Paper SPE 24114 Presented at the SPE/DOE Enhanced Oil Recovery Symposium, Tulsa, Oklahoma, 22-24 April, 1992.

Emera, M.K., Sarma, H.K. Use of genetic algorithm to estimate CO_2-oil minimum miscibility pressure-A key parameter in design of CO_2 miscible flood. J. Pet. Sci. Eng. 2005, 46(1-2): 37-52.

Fazlali, A., Nikookar, M., Mohammadi, A.H. Computational procedure for determination of minimum miscibility pressure of reservoir oil. Fuel 2013, 106: 707-711.

Firoozabadi, A., Aziz, K. Analysis and correlation of nitrogen and lean gas miscibility pressure. SPE Reserv. Eng. 1986, 1(6): 575-582.

Frimodig, J.P., Reese, N.A., Williams, C.A. Carbon dioxide flooding evaluation of high-pour-point, Paraffinic Red Wash reservoir oil. SPE J. 1983, 23(4): 587-594.

Gardner, J.W., Orr, F.M., Patel, P.D. The effect of phase behavior on CO_2 flood displacement efficiency. J. Pet. Technol. 1981, 33(11): 2067-2081.

Glaso, O. Generalized minimum miscibility pressure correlation. SPE J. 1985, 25(6): 927-934.

Goh, A.T.C. Back-propagation neural networks for modeling complex systems. Artif. Intell. Eng. 1995, 9(3): 143-151.

Graue, D.L., Zana, E.T. Study of a possible CO_2 flood in Rangely field. J. Pet. Technol. 1981, 33(7): 1312-1318.

Green, D.W., Willhite, G.P. Enhanced oil recovery. Texas, USA, Society of Petroleum Engineers Text Book Series, 1998.

Hand, J.L., Pinczewski, W.V. Interpretation of swelling/extraction tests. SPE Reserv. Eng. 1990, 5(4): 595-600.

Hanssen, J.E. Nitrogen as a low-cost replacement for natural gas reinfection offshore. Paper SPE 17709 Presented at the SPE Gas Technology Symposium, Dallas, Texas, USA, 13-15 June, 1988.

Harmon, R.A., Grigg, R.B. Vapor-density measurement for estimating minimum miscibility pressure. SPE Reserv. Eng. 1998, 3(4): 1215-1220.

Holm, L.W., Josendal, V.A. Mechanisms of oil displacement by carbon dioxide. SPE Reserv. Eng. 1974, 26(12): 1427-1438.

Huang, Y., Huang, G., Dong, M., et al. Development of an artificial neural network model for predicting minimum miscibility pressure in CO_2 flooding. J. Pet. Sci. Eng. 2003, 37(1-2): 83-95.

Hudgins, D.A., Liave, F., Chung, F. Nitrogen miscible displacement of light crude oil: A laboratory study. SPE Reserv. Eng. 1990, 5(1): 100-106.

Jacobson, H.A. Acid gases and their contribution to miscibility. J. Can. Pet. Technol. 1972, 11(2): 56-59.

Jaubert, J.N., Avaullee, L., Souvay, J.F. A crude oil data bank containing more than 5000 PVT and gas injection data. J. Pet. Sci. Eng. 2002, 34(1-4): 65-107.

Krenker, A., Bester, J., Kos, A. Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech, 2011.

Lee, J.I. Effectiveness of carbon dioxide displacement under miscible and immiscible conditions. Paper PRI-7910/CE-02239 Presented at the Petroleum Recovery Institute, Calgary, Alberta, Canada, 1979.

Nezhad, A.B., Mousavi, S.M., Aghahoseini, S. Development of an artificial neural network model to predict CO_2 minimum miscibility pressure. Nafta 2011, 62(3-4): 105-108.

Orr, F.M. Theory of Gas Injection Processes. Copenhagen, Denmark, Tie-Line Publications, 2007.

Rao, D.N. A new technique of vanishing interfacial tension for miscibility determination. Fluid Phase Equilib. 1997, 139(1-2): 311-324.

Rao, D.N., Lee, J.I. Miscibility evaluation for Terra Nova offshore field. Paper PETSOC-2000-082 Presented at the Canadian International Petroleum Conference, Calgary, Alberta, Canada, 4-8 June, 2000.

Rathmell, J.J., Stalkup, F.L., Hassinger, R.C. A laboratory investigation of miscible displacement by CO_2. Paper SPE 3483 Presented at SPE Annual Fall Meeting, New Orleans, Louisiana, USA, 3-6 October, 1971.

Ren, B., Xu, Y., Ren, S., et al. Laboratory assessment and field pilot of near miscible CO_2 injection for IOR and storage in a tight oil reservoir of Shengli oilfield China. Paper SPE 144108 Presented at the SPE Enhanced Oil Recovery Conference, Kuala Lumpur, Malaysia, 9-21 July, 2011.

Sebastian, H.M., Lawrence, D.D. Nitrogen minimum miscibility pressures. Paper SPE 24134 Presented at the SPE/DOE Enhanced Oil Recovery Symposium, Tulsa, Oklahoma, USA, 22-24 April, 1992.

Sebastian, H.M., Wenger, R.S., Renner, T.A. Correlation of minimum miscibility pressure for impure CO_2 streams. J. Pet. Technol. 1985, 37(11): 2076-2082.

Shokir, E.M.E.M. CO_2-oil minimum miscibility pressure model for impure and pure CO_2 streams. J. Pet. Sci. Eng. 2007, 58(1-2): 173-185.

Shokrollahi, A., Arabloo, M., Gharagheizi, F., et al. Intelligent model for prediction of CO_2-reservoir oil minimum miscibility pressure. Fuel 2013, 112: 375-384.

Srivastava, R.K., Huang, S.S. New interpretation technique for determining minimum miscibility pressure by rising bubble apparatus for enriched-gas drives. Paper SPE 39566 Presented at the SPE India Oil and Gas Conference and Exhibition, New Delhi, India, 17-19 February, 1998.

Srivastava, R.K., Huang, S.S., Dong, M. Laboratory investigation of Weyburn CO_2 miscible flooding. Paper PETSOC-97-154 Presented at the Technical Meeting/Petroleum Conference of The South Saskatchewan Section, Regina, 19-22 October, 1997.

Stalkup, F.I. Status of miscible displacement. J. Pet. Technol. 1983, 35(4): 815-826.

Tarybakhsh, M.R., Assareh, M., Sadeghi, M.T., et al. Improved minimum miscibility pressure prediction for gas injection process in petroleum reservoir. Nat. Resour. Res. 2018, 27(4): 517-529.

Thomas, F.B., Zhou, X.L., Bennion, D.B., et al. A comparative study of RBA, Px, multicontact and slim tube results. J. Can. Pet. Technol. 1994, 33(2): 17-26.

Whitson, C.H., Michelsen, M.L. The negative flash. Fluid Phase Equilib. 1989, 53: 51-72.

Yellig, W.F., Metcalfe, R.S. Determination and prediction of CO_2 minimum miscibility pressure. J. Pet. Technol. 1980, 32(1): 160-180.

Yuan, H. Application of miscibility calculations to gas floods. USA, The University of Texas at Austin, 2003.

Yuan, H., Johns, R.T. Simplified method for calculation of minimum miscibility pressure or enrichment. SPE J. 2005, 10(4): 416-425.

Yuan, H., Johns, R.T., Egwuenu, A.M., et al. Improved MMP correlation for CO_2 floods using analytical theory. SPE

Reserv. Eval. Eng. 2005, 8(5): 418-425.

Yurkiw, F.J., Flock, D.L. A comparative investigation of minimum miscibility pressure correlations for enhanced oil recovery. J. Can. Pet. Technol. 1994, 33(8): 35-41.

Zarenezhad, B., Aminian, A. An artificial neural network model for design of wellhead chokes in gas condensate production fields. Pet. Sci. Technol. 2011, 29(6): 579-587.

Zhang, K., Jia, N., Zeng, F. Application of predicted bubble-rising velocities for estimating the minimum miscibility pressures of the light crude oil-CO_2 systems with the rising bubble apparatus. Fuel 2018, 220: 412-419.

Zhong, Z., Carr, T.R. Application of mixed kernels function (MKF) based support vector regression model (SVR) for CO_2-Reservoir oil minimum miscibility pressure prediction. Fuel 2016, 184: 590-603.

Zick, A.A. A combined condensing/vaporizing mechanism in the displacement of oil by enriched gases. Paper SPE 15493 Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA, 5-8 October, 1986.

Appendix A: Dataset description

The complete dataset compiled for and used in this study consists of 251 published data records. The details of each record are included in the supplementary file comprising: The sources reference; measured values of each of the four input variables; the percentage components making up the injected gas composition, and the minimum miscible pressure (MMP) constituting the measured value of the dependent variable. The table below summarizes the dataset highlighting the wide range of MMP it covers.

Source reference number	Number of data records	MMP/Psia
3	12	1,164.5-2,364.5
28	19	1,375-4,390
30	6	4,850-6,850
35	4	4,814.5-6,314.5
36	3	5,114.5-6,864.5
37	5	1,750-2,650
38	1	4,090
39	4	2,462.7-3,063.2
40	2	2,219-2,393
41	10	1,305.3-2,320.6
42	3	4,675- 5,600
43	3	3,350-3,502
44	2	1,594-2,250
45	2	1,540-2,580
46	1	3,400
47	1	4,650
48	1	1,000
49	30	948-2,700
50	3	1,500-5,000
51	9	1,740-3,974
52	9	2,135-2,790
53	36	1,283-2,705
54	4	1,189.5-1,829.5
55	16	2,340-5,226
56	3	1,850-2,750
57	6	2,030.5-9,115.6
58	4	1,708-3,190
59	6	1,889.5-5,009.5
60	17	3,754-7,000
61	10	3,582.4-7,977
62	4	3,916-4,684.7
63	15	3,205-5,496