Original article

An improved method for predicting CO_2 minimum miscibility pressure based on artificial neural network

Peng Dong, Xinwei Liao [ID]*, Zhiming Chen, Hongyang Chu

State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing 102249, P. R. China

Abstract:
The CO_2 enhanced oil recovery (EOR) method is widely used in actual oilfields. It is extremely important to accurately predict the CO_2 minimum miscibility pressure (MMP) for CO_2-EOR. At present, many studies about MMP prediction are based on empirical, experimental, or numerical simulation methods, but these methods have limitations in accuracy or computation efficiency. Therefore, more work needs to be done. In this work, with the results of the slim-tube experiment and the data expansion of the multiple mixing cell methods, an improved artificial neural network (ANN) model that predicts CO_2 MMP by the full composition of the crude oil and temperature is trained. To stabilize the neural network training process, L2 regularization and Dropout are used to address the issue of over-fitting in neural networks. Predicting results show that the ANN model with Dropout possesses higher prediction accuracy and stronger generalization ability. Then, based on the validation sample evaluation, the mean absolute percentage error and R-square of the ANN model are 6.99 and 0.948, respectively. Finally, the improved ANN model is tested by six samples obtained from slim-tube experiment results. The results indicate that the improved ANN model has extremely low time cost and high accuracy to predict CO_2 MMP, which is of great significance for CO_2-EOR.

1. Introduction

Excessive CO_2 emissions are key factors in global climate issues. In the oil industry, the main ways of CO_2 emissions reduction are CO_2 geosequestration and CO_2 flooding (Englezos and Lee, 2005; Lv et al., 2015; Chu et al., 2019). CO_2 flooding can not only sequester CO_2 but also enhance oil recovery, which is a win-win way and has become an effective tool for enhanced oil recovery (EOR). According to the mechanisms, CO_2 miscible flooding can achieve high-efficiency oil displacement, and the theoretical oil displacement efficiency can reach 100%. At present, the method for judging whether the reservoir can form the miscible flooding is mainly based on the minimum miscibility pressure (MMP) value. Therefore, accurate prediction of the MMP is particularly important.

The methods for determining the MMP of CO_2 are traditional experimental method, empirical formula method, and theoretical calculation method. Stalkup Jr (1983) proposed the slim-tube experiment, which is recognized as the most reliable and classical method for experimental determination. However, it takes a long time and it cannot obtain a large amount of data. Through experiments, many scholars have found that the

MMP is associated with the oil composition and the formation temperature, thus many prediction empirical formulas were developed. Commonly, the empirical formulas are Johnson and Pollin correlation (Johnson et al., 1981) and Yelling and Metcalfe correlation (Yellig and Metcalfe, 1980). Although the empirical formula method is simple and straightforward to solve MMP, it is based on the experiment of a specific oil field, and the scope of application is very limited. Stalkup Jr (1983) believed that no empirical relationship can be accurate enough for the final design. Due to the shortcomings of the empirical formula, many scholars began to study the MMP through theoretical calculations. Ahmed (1997) proposed using the equation of state to determine the MMP, based on the modified Peng-Robinson equation of state, and he established a miscible function to predict the MMP. Johns and Orr Jr (1996) used tie line analysis method to determine the MMP from a single injection of gas into a four-component system. Ahmadi and Johns (2011) developed this technique and proposed multiple mixing cell methods (MMCM) to determine the MMP of any component system, and they achieved higher prediction accuracy, but the calculation speed is slower.

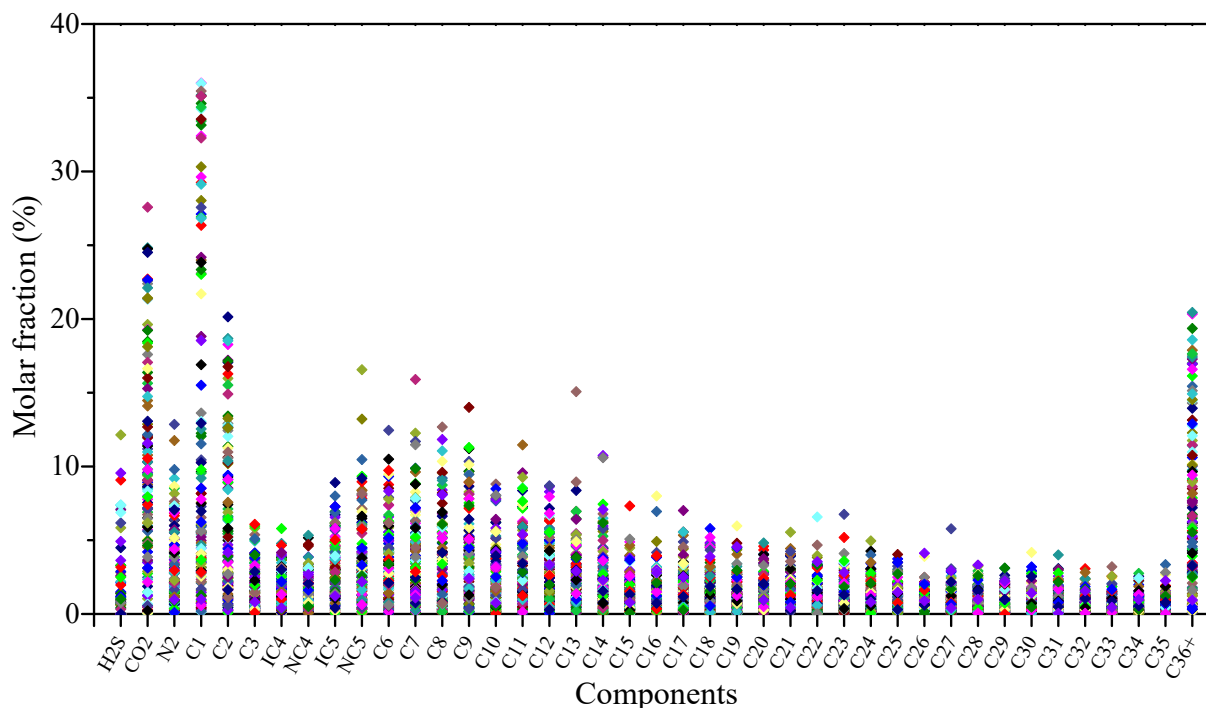In order to express the complex nonlinear relationship

Fig. 1. Component data distribution diagram.

of crude oil component and temperature to CO_2 MMP, a method of predicting CO_2 MMP by artificial neural network (ANN) was proposed. Huang et al. (2003) developed an ANN model with one hidden layer for predicting pure and impure CO_2 MMP. The pure CO_2 MMP is directly obtained by the ANN model, while the impure CO_2 MMP is corrected by the MMP factor. Edalat et al. (2007) constructed a more complex ANN model for predicting the CO_2 MMP in hydrocarbon gas injection, which mean absolute percentage error (*MAPE*) on the training set is 0.036 and on testing set is 6.9. Based on the ANN model, Dehghani et al. (2008) used the genetic algorithm to optimize the ANN weights. The results show that ANN with genetic algorithm has a stronger generalization ability for predicting MMP. Hassan et al. (2018) used temperature, C_2-C_6 mole percent, and C_7+ molar mass as sample characteristics to compare the ANN, generalized neural network, and radial basis function (RBF) for CO_2 MMP predictions. The results show that the RBF has the highest accuracy in predicting CO_2 MMP. Choubineh et al. (2019) used a classic 5-layer ANN to predict CO_2 MMP. The results show that the *MAPE* of MMP predicted by ANN with purelin activation function is 13.46%. According to previous studies, the input nodes of the neural network mostly is the plus components of crude oil. As the plus components cannot accurately represent the crude oil, the accuracy and generalization ability of the neural network model will be reduced.

In this paper, through the improved ANN model, this model can learn the behavior of the MMCM to predict the MMP of CO_2. Its accuracy is almost the same as the MMCM. What's more, the operation speed can be greatly improved, and it can be calculated in batches. Using ANN, the first step is to

obtain the data of a large number of sample components and their corresponding MMP. However, most samples are the plus fraction data of mole weight and lack full-component data. For this purpose, we use the Whitson split method (Whitson, 1983) to obtain the full component data and calculate the MMP of the crude oil sample by the MMCM. Then, the neural network models are trained by these data. Due to the powerful representation capabilities of the network, it is prone to over-fitting problems. To improve this situation, the L2 regularization (Ng, 2004) and Dropout methods (Srivastava, 2014) are adopted, respectively. Finally, by comparing the results with 6 sets of slim tube experiments, the improved ANN model reliability is verified.

2. Method

2.1 MMP sample obtained

Previous studies have shown that the MMP of CO_2 is related to crude oil composition and formation temperature (Yellig and Metcalfe, 1980; Johnson and Pollin, 1981; Enick et al., 1988). In this paper, the mole fraction of the crude oil component and the formation temperature data are taken as samples and the corresponding MMP data is used as the sample label for neural network learning. In general, to obtain a reliable neural network model, taking a large number of samples is necessary. Although MMP with mine guidance significance can be obtained through the slim tube experiment, it is impossible to obtain a large amount of experimental sample data due to time cost. Therefore, data augmentation is necessary. To this end, by consulting the literature (Yellig and Metcalfe, 1980; Johnson and Pollin, 1981; Alston et al.,
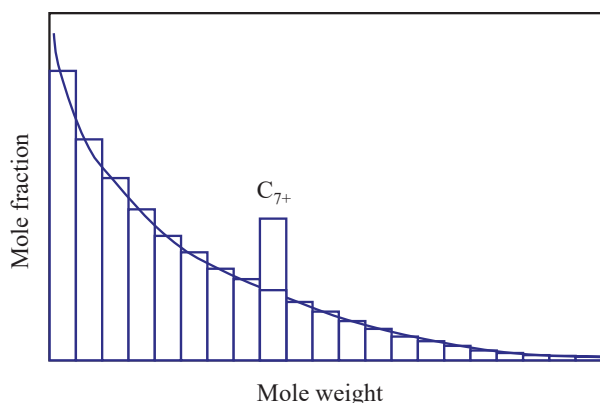
Fig. 2. Schematic of component splitting.

Fig. 3. Schematic of data set distributions.

1985; Enick et al., 1988), 122 samples containing oil sample components and formation temperature data were obtained. Since the sample component data is mostly given by the plus fraction data of mole weight, to obtain the full-component sample, it is necessary to split the component. In this paper, the plus fraction data of crude oil samples are divided into C_1, C_2,..., $C_{36}+$ using the Whitston splitting (Whitson, 1983), as shown in Fig. 1. In the Whiston splitting method, the mole fraction of a component is related to the molar weight and can be described by a three-parameter probability density function. This function divides the additive component into a number of single carbon number (SCN) (Riazi, 2007). As shown in Fig. 2, the original C_7+ component is split into SCN.

To obtain a reliable MMP, this paper applies MMCM to calculate MMP of a full-component sample. The MMCM method is developed based on the study of the tie line analysis method (Ahmadi et al., 2011), and it is the tie line method of any number of injected gas components. The tie line method is based on the theory of tie line analysis. The analytical solution is obtained by determining the algebra and geometry of a series of main tie lines. Once the main tie line is determined, the displacement process of multiple-contact miscible can be represented by the geometry of the main tie line, and the MMP can be derived from the geometry of the main tie line.

In the ANN, to train, improve and evaluate the neural network model for predicting MMP, it is necessary to divide the crude oil component and temperature data sets into training set, validation set, and test set. The primary role of the validation set is to improve and select the final model that predicts MMP. The test set provides an unbiased estimation of the model for predicting MMP and verifies the generalization ability of the model. In this paper, 100 components of the crude oil component and temperature sample obtained by the MMCM are the training set, and the remaining 22 samples are the validation set. Then, a sample out of order operation, named sample shuffle, is needed before the training of the ANN. It proves that shuffle can not only improve the robustness of the model but also avoid bias in the model. Finally, 6 sets of samples obtained through the slim tube experiment are used as test sets. Fig. 3 is a schematic diagram of data set distributions.
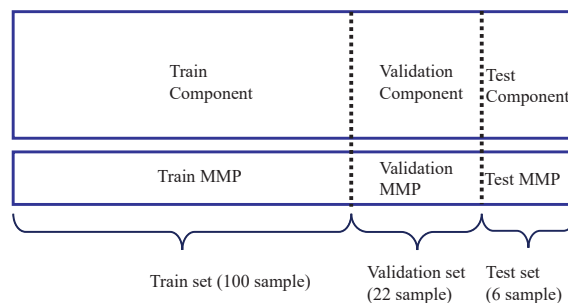
2.2 Model building

The ANN model has strong nonlinear representation ability, and its basic unit is a neuron. Through the design of different numbers of neurons and different layers, the mapping relationship of different complexity levels can be characterized. This paper sets up a 4-layer ANN structure: The input layer consists of oil composition and temperature, the two hidden layers respectively contain 30 neurons and 12 neurons, and the output layer represents the predicted MMP, as shown in Fig. 4. The total number of weights and biases of the neural network is 1,675, and their number in each layer is shown in Table 1. In order to complete the construction of the neural network, we need to specify an activation function for the hidden layer. Its powerful nonlinear representation ability is due to the nonlinear activation function that neurons have. Different activation functions apply to different problem types. In this paper, since the output layer data is a floating-point data structure, the output range of the *Relu* activation function (Glorot et al., 2011) is $[0, +\infty]$, the *Relu* activation function is used, as shown in Eq. (1).

$$g(z) = \max(0, x) \tag{1}$$

Next, we want to train the network so that the predicted MMP is as close as possible to the target MMP. When the mean square error between the MMP and the expected MMP predicted by the model is minimal, we get the optimal ANN model. The mean square error is defined as the loss function L of the model, as shown in Eqs. (2)~(5). Our goal is to adjust the weight and biases, as shown in Eq. (5), to minimize the L function.

$$L(\hat{y}, y) = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 \tag{2}$$

$$\hat{y} = a^{[3]} \tag{3}$$

$$a^{[l]} = g\left(z^{[l]}\right) \tag{4}$$

$$z_j^{[l+1]} = w_j^{[l+1]} a^{[l]} + b_j^{[l+1]} \tag{5}$$

where $\hat{y}$ represents the MMP obtained by ANN; y represents the target MMP; z represents the inputs of a layer; a represents the outputs of a layer ($a^{[0]} = x$); m represents the sample size;
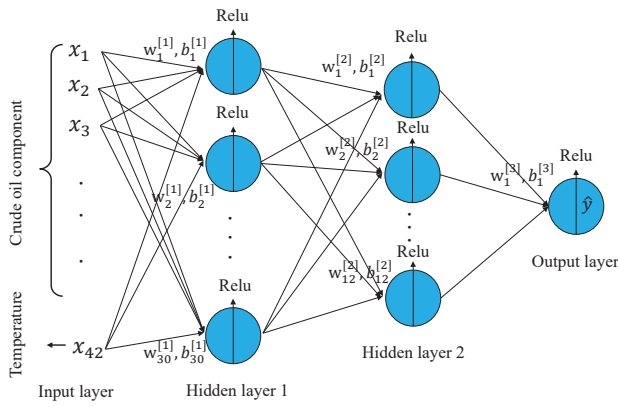
Fig. 4. Network structure.

<center>**Table 1.** Network weights.</center>

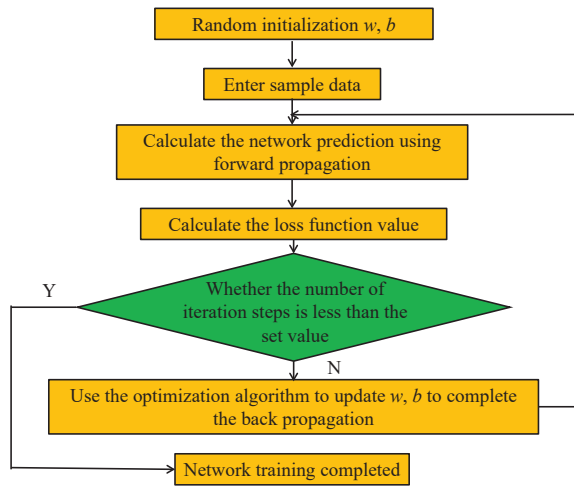| Layer	Output shape	Number of weights and biases
Input layer	42	0
Hidden layer 1	30	1,290
Hidden layer 2	12	372
Output layer	1	13

Fig. 5. Flow chart of neural network calculation.

w and b represents the weights and biases respectively; the superscript $[l]$ indicates the lth layer; the subscript i represents the i-th sample; the subscript j represents the j-th neuron in a layer.

In order to minimize the L function value of the predicted MMP model, iteratively updating the weights w and b are required, as shown in Fig. 5. An optimization algorithm needs to be used when updating w and b. In each of the optimization algorithms, the derivative of L for each w and b is required. Then, the w and b are derived and iteratively updated to obtain the minimum value of L. This process completes the backpropagation of the neural network, as shown in Fig. 5. In the process of backpropagation, there are many optimization
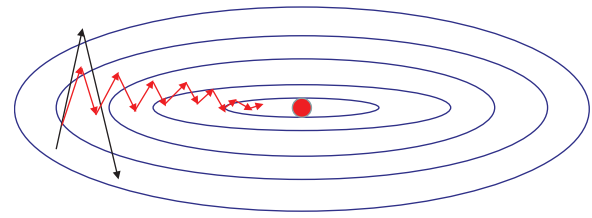

Fig. 6. Schematic diagram of the advantages of Adam algorithm. It can make the iterative process more stable and faster, as shown by the red line.
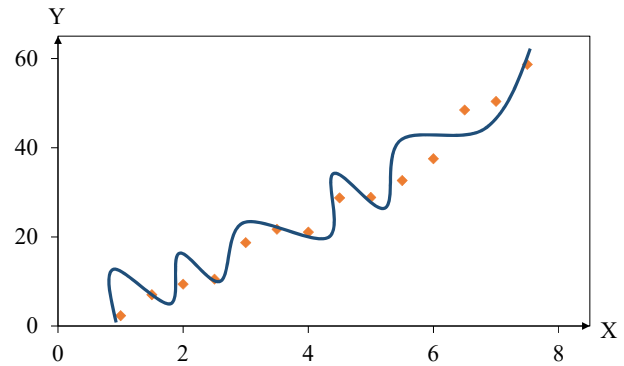

Fig. 7. Schematic of over-fitting.

algorithms for updating w and b, such as momentum gradient descent method (Qian, 1999), RMSprop algorithm (Sutton, 1986), and Adam algorithm (Kingma and Ba, 2014). The Adam optimization algorithm combines the advantages of both the Momentum algorithm and the RMSprop algorithm. When the weights are updated, the differential of the weights of the previous iterative step takes an exponentially weighted average to make the differentiation of the partial weights in the iteration without a large swing. It can eliminate a part of the swing, as shown in Fig. 6, and a larger learning rate can be allowed to speed up the training. This paper uses the Adam algorithm to update the w and b. The w and b of the training completion can be obtained by the set iteration termination condition.

There is an overfitting problem in training neural networks for predicting MMP. When this problem occurs, the bias of the prediction result in the training set is small, but it will produce a large bias in the actual prediction. The schematic of the model over-fitting is shown in Fig. 7. The reason for this is the powerful representation of the network. At this point, there are two solutions, one is algorithm optimization and the other is getting more data. Extending the experimental data of MMP is a very reliable method, but the cost of obtaining more experimental data is high, and algorithm optimization usually helps to avoid over-fitting or reduce the network error of predicting MMP.

The L2 regularization and Dropout methods are commonly used for algorithm optimization. The L2 regularization pass adds an F-norm of the network weight after the loss function, to punish the network weights and make the value smaller. As Eq. (6):
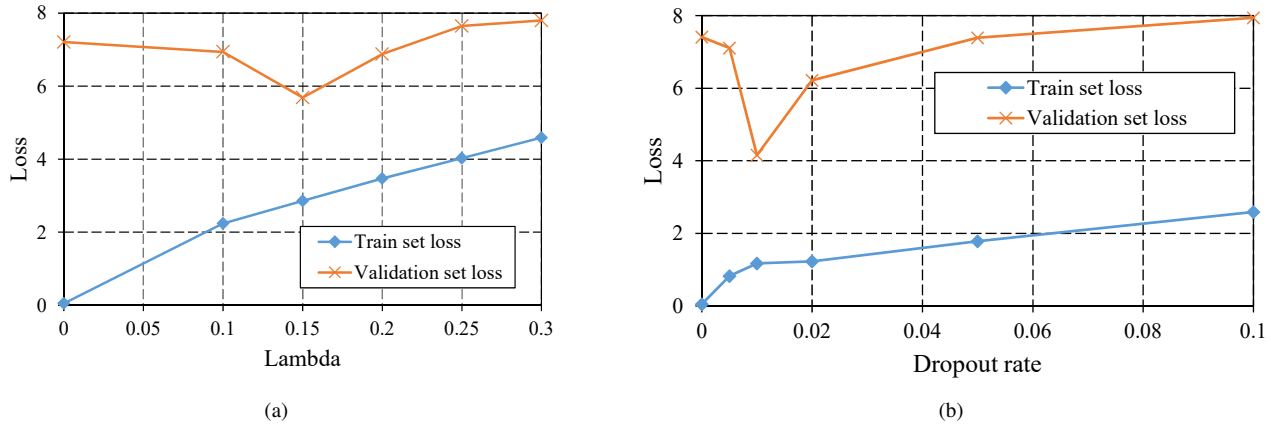
(a)

(b)

Fig. 8. Parameter optimization comparison chart. (a) Influence of lambda on loss function when using the L2 regularization method; (b) Influence of dropout rate on loss function when using the dropout method.

Table 2. Lambda value optimization.

Lambda	Training set loss	MAPE (%)	Validation set loss	MAPE(%)
0	0.049	0.689	7.41	10.21
0.1	2.24	2.7823	6.94	10.13
0.15	2.86	3.59	5.69	8.288
0.2	3.47	4.162	6.88	9.09
0.25	4.028	4.583	7.65	9.6
0.3	4.59	5.5153	7.8	9.47

$$L\left(w^{[l]}, b^{[l]}, \ldots, w^{[L]}, b^{[L]}\right) = \frac{1}{2m}\sum_{i=l}^{m}(\hat{y}_i - y_i)^2 + \frac{\lambda}{2m}\sum_{l=1}^{L}\left\|w^{[l]}\right\|_F^2 \tag{6}$$

where:

$$\left\|w^{[l]}\right\|_F^2 = \sum_{i=1}^{n^{[l-1]}}\sum_{j=1}^{n^{[l])}}\left(w_{ij}^{[l]}\right)^2$$

Fig. 9. Dropout schematic.

In L2 regularization, λ is the key super-parameter to be optimized. It controls the degree of weight decay of the neural network. As λ increases, the degree of weight decay increases and the over-fitting decreases. However, excessive λ may lead to under-fitting. By optimizing the value of λ, the over-fitting problem of the ANN model for predicting CO_2 MMP is avoided. As shown in Fig. 8(a), the ANN model is trained separately under different λ values. Then the loss function is compared in the training set and the validation set. As the value of the λ increases, the loss of the training set increases, indicating that the L2 algorithm increases the weight decay for the model. For the validation set, the loss function value first decreases and then increases. When the λ is less than 0.15, the decrease of the loss function value indicates that the overfitting degree of the model is reduced. When the λ is greater than 0.15, the increase of loss indicates that the model has an under-fitting phenomenon. Therefore, 0.15 is the optimal value for λ. Finally, we use MAPE to visually understand the error range
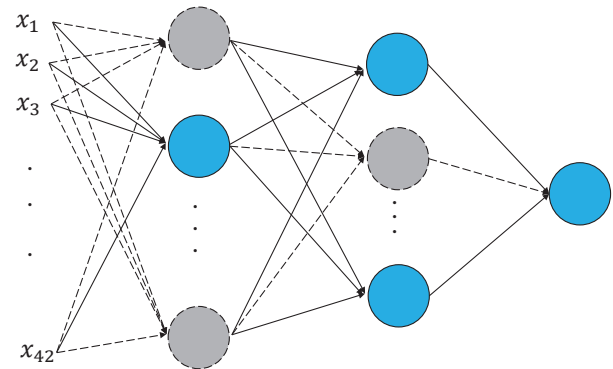
of the model prediction, as shown in Eq. (7).

$$MAPE = \frac{1}{m}\sum_{i}^{m}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{7}$$

As shown in Table 2, the *MAPE* value on the validation set indicates that the model's prediction error is around 8.3% when the λ is 0.15.

Dropout is a completely different technology in regularization. Unlike the L2 regularization, Dropout does not modify the loss function but modifies the network itself. Dropout solves the problem of overfitting and gradient vanishing of neural networks by means of randomly inactivating neurons.

Table 3. Dropout rate optimization.

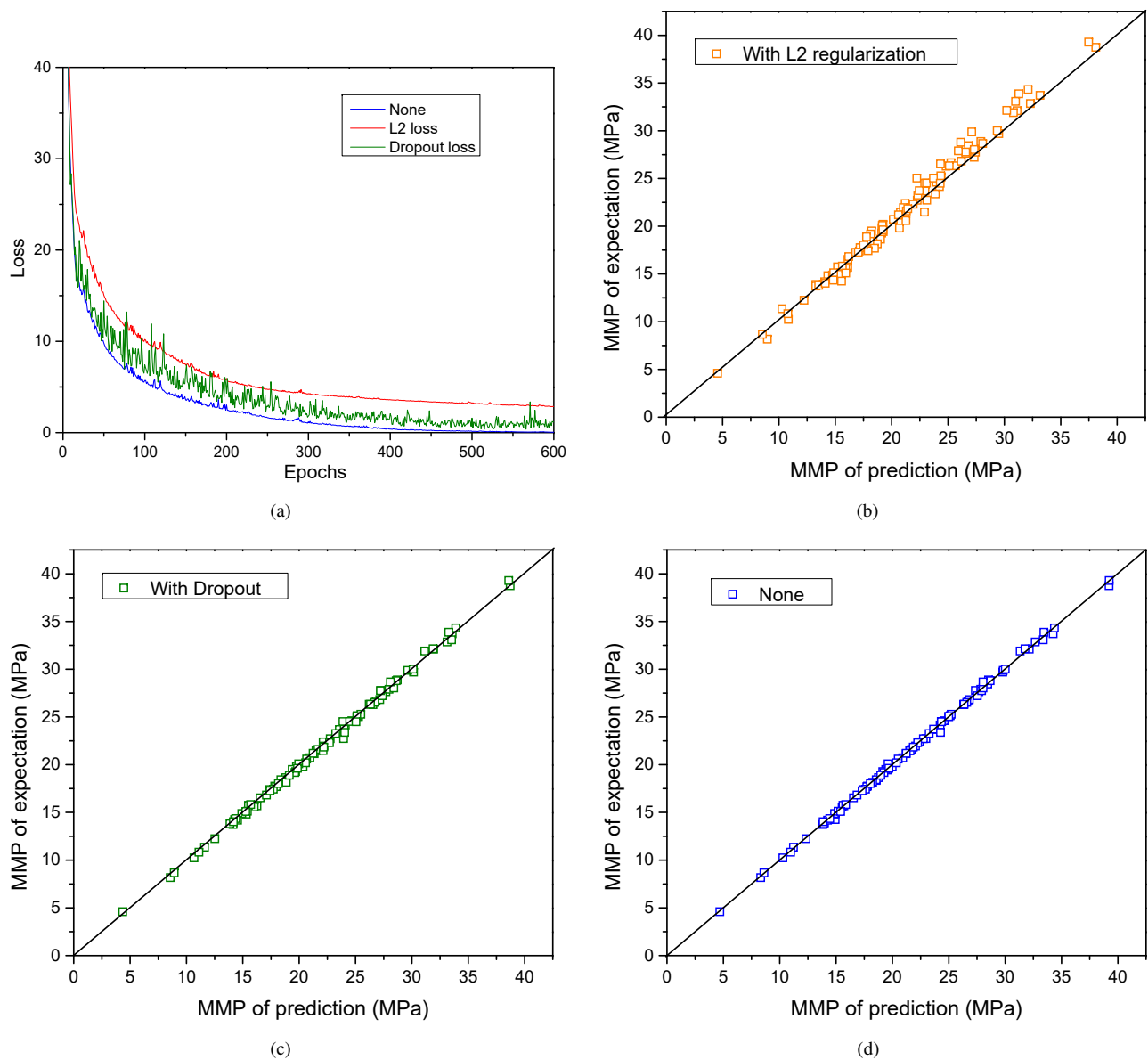| Dropout rate	Training set loss	*MAPE* (%)	Validation set loss	*MAPE* (%)
0	0.049	0.689	7.41	10.21
0.005	0.8252	1.176	7.1085	9.447
0.01	1.17	1.327	4.156	6.99
0.02	1.226	1.624	6.22	10.59
0.05	1.783	2.149	7.39	12.8
0.1	2.59	2.545	7.94	14.4

Fig. 10. Forecast result comparison chart on the training set. (a) Model training process with different improvement methods; (b) The prediction result when the L2 regularization method is adopted; (c) The prediction result when the Dropout method is adopted; (d) The prediction result without any improvement method.

Table 4. Comparison of different optimization methods in different data sets.

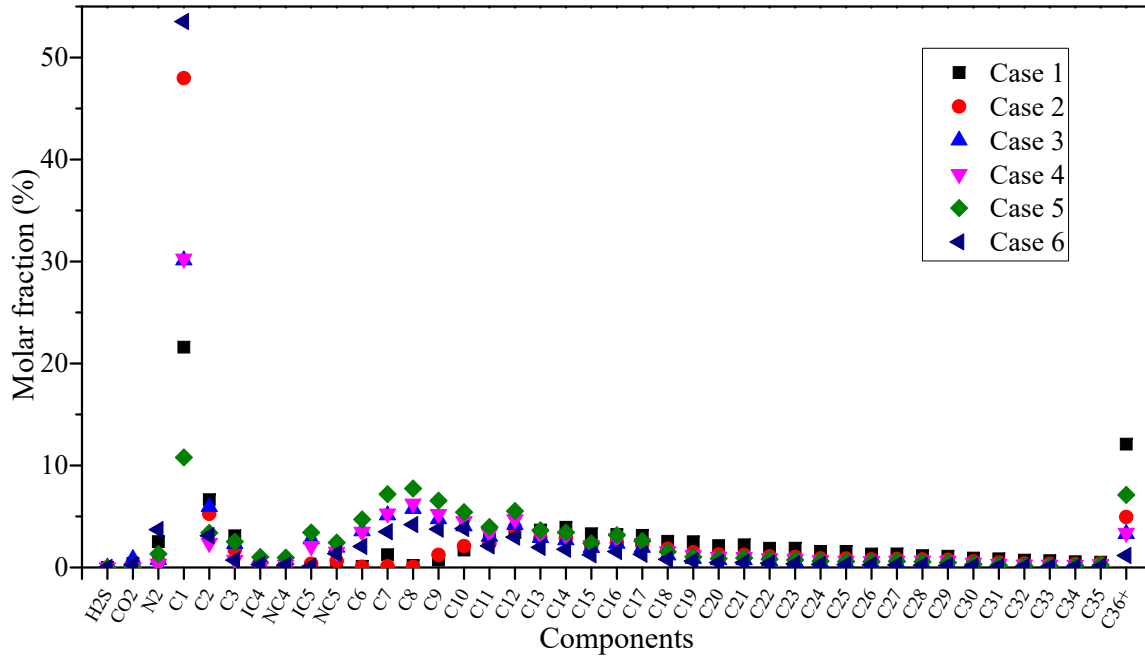| Method	Training set			Validation set		
	Loss	$MAPE$ (%)	R^2	Loss	$MAPE$(%)	R^2
None	0.05	0.649	0.999	7.21	10.21	0.840
L2	2.86	3.59	0.977	5.69	8.29	0.929
Dropout	1.17	1.33	0.998	4.16	6.99	0.948

Fig. 11. Case sample.

By randomly setting the weight of some hidden layer nodes to zero, the neurons are inactivated, as shown in Fig. 9. As the nodes affected by zeroing at each iteration are different, the "importance" of each node is balanced.

After using dropout, Eq. (5) will be rewritten as follows:

$$z_j^{[l+1]} = w_j^{[l+1]} \tilde{a}^{[l]} + b_j^{[l+1]} \tag{8}$$

$$\tilde{a}^{[l]} = r^{[l]} * a^{[l]} \tag{9}$$

$$r_j^{[l]} \sim \text{Bernoulli}(p) \tag{10}$$

where r represents a random number obeying the Bernoulli distribution (Hogg et al., 2005). The symbol $*$ represents an element-wise product.

As seen from Eqs. (8)~(10), neurons will be randomly activated. Although each node of the neural network will still contribute content, there is no situation where a few high-weight nodes completely control the output result. Through the Dropout, the structural risk of the network is reduced and the computational overhead of the neural network goes down. The degree of overfitting can be controlled by the Dropout rate,

probability p in Eq. (10), which represents the proportion of disconnected neurons. By setting different rates, the loss of the training set and validation set is calculated, and the result is shown in Fig. 8(b). It can be seen from Fig. 8(b) that when the Dropout rate is 0.01, the model effectively avoids the problem of overfitting. Similarly, we calculate the $MAPE$ value of the model. It can be seen from Table 3 that when the Dropout rate is 0.01, the model's prediction error is around 7%. By comparing two different methods of suppressing overfitting, the Dropout has less loss in the training set and validation set, indicating that the Dropout method is a better choice. In practical applications, the over-fitting reduction will improve the accuracy of the model for predicting CO_2 MMP.

3. Comparison of MMP prediction results

The training was carried out in 600 epochs. The initialization method of the ANN's weight uses Xavier_normal proposed by Glorot and Bengio (2010). Fig. 10(a) shows the training process of the model using different improved methods. As shown in Fig. 10(a), the training error of Dropout is lower than that of L2 in the training process, and the loss

Table 5. Case test results.

| Method\CO$_2$ MMP (MPa)	1 (96 °C)	2 (75 °C)	3 (60 °C)	4 (62 °C)	5 (89 °C)	6 (74 °C)	Avg. time cost
Slim-tube experiment	28.8	34.3	21.5	18.8	17.6	28.6	3-6 days
MMCM	23.9	31.3	18.5	16.9	16.1	24.9	1-200 seconds
Improved ANN	22.2	30.8	19.2	19.2	19.5	29.5	<1 second

of the model without the improved method is the smallest. In the training process, the loss of the Dropout fluctuates greatly, which is caused by the continuous changes in the network structure. When the ANN model is not improved, the loss function value in the training set is 0.049 and the average error rate is 0.64%. By comparing the expected CO$_2$ MMP with the CO$_2$ MMP predicted by the ANN model without improved methods, the scatter plot of the relationship between the two is shown in Fig. 10(d). It shows that the network has a very good fitting effect on the expected CO$_2$ MMP. After the prediction of CO$_2$ MMP on the validation set, the loss function value of CO$_2$ MMP predicted by the ANN model reached 7.41, and the average error rate was 10.21%. Training and the validation results indicate that the model has an overfitting problem.

After avoiding overfitting, the predicted effect of the model in the training set is shown in Figs. 10(b) and (c). The results show that both L2 and Dropout increase the error of the training set, and the error in the validation set is reduced, as shown in Fig. 11. The figure also indicates that the model with Dropout has the smallest prediction error in the validation set. This shows that the generalization ability of the model is enhanced by the regularization method. Table 4 is the evaluation results of the model improvement effect when different methods of suppressing overfitting are used. In order to evaluate the strength of the linear relationship between the predicted MMP and the target MMP of the different methods, the coefficient of determination R^2 is selected as the evaluation index, as in Eq. (11).

$$R^2 = 1 - \frac{\sum_i^m (\hat{y}_i - y_i)^2}{\sum_i^m (y_i - \bar{y})^2} \tag{11}$$

where $\bar{y}$ represents the mean of the label data. The closer R^2 is to 1, the stronger the linear relationship between the predicted value and the target value, that is, the closer the predicted value is to the target value. The smaller the R^2 is, the weaker the linear relationship between the predicted value and the target value is, and the worse the prediction effect of the model is. The R^2 values of the ANN model before and after the improvement in different data sets are shown in Table 4.

Table 4 compares the pre- and post-improvement errors of the model on different datasets, where the error of the validation set demonstrates the generalization capabilities of the model. It can be seen from Table 4 that after the model is improved by Dropout, the prediction *MAPE* is lowest and the R^2 score is highest, which indicates the generalization ability of the model is best.
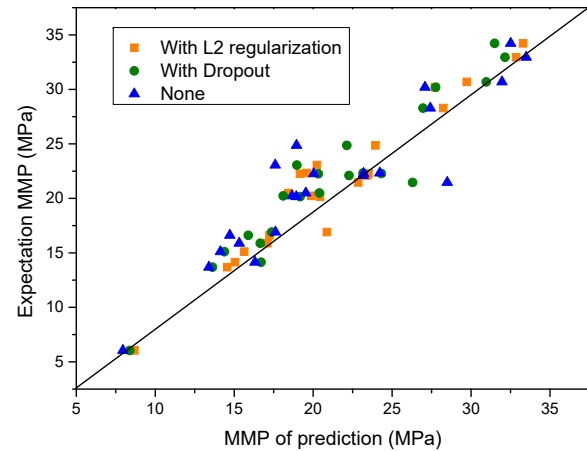
Fig. 12. Forecast result comparison chart on the validation set.

Finally, based on the actual oil samples, we verify the prediction effect of the ANN method. Fig. 12 shows the distribution of component data for six actual oil samples from the Junggar Basin. Table 5 compares the efficiency and accuracy of 3 various calculating CO$_2$ MMP methods on actual oil samples. The results show that the ANN method has a good effect on practical applications, and the calculation speed is obviously improved.

4. Conclusion

Based on the MMCM, slim-tube experiment, and ANN, the 122 sets of crude oil samples are split into SCN. Then, the MMP is calculated by MMCM for samples. The MMP model of the CO$_2$-crude oil system and formation temperature is established by ANN. Then, by improving the ANN model, a series of methods are used to overcome the problem of overfitting. Finally, the experimental data are used to test the performance of the model. The main conclusions are as follows:

1) The CO$_2$ MMP prediction model established by the improved ANN model has a generalized error rate of 6.99% and R^2 of 0.948, which indicates that it has certain accuracy. The actual case proves that the ANN model can be used to predict the CO$_2$ MMP in different oil composition at different formation temperatures

2) In this paper, the effects of L2 regularization and Dropout methods on reducing the overfitting of the ANN model are compared. The results show that for the ANN model of predicting CO$_2$ MMP, Dropout is a more effective

means of avoiding overfitting. Through Dropout, the generalization capabilities of the model are enhanced. In practical applications, the accuracy of CO_2 MMP is improved. For the calculation speed, the ANN model is significantly faster than the experiment and MMCM. It shows that the ANN model has the characteristics of fast calculation speed under the condition of certain precision.

Nomenclature

Abbreviations

EOR = enhanced oil recovery
MMP = minimum miscibility pressure
ANN = artificial neural network
$MMCM$ = multiple mixing cell methods
$MAPE$ = mean absolute percentage error
RBF = radial basis function
SCN = single carbon number

Variables

L = number of layers
x = sample matrix
y = real sample label vector
$\hat{y}$ = predictive sample label vector
$\bar{y}$ = mean of the label data
a = output of a layer
z = input of a layer
m = number of training samples
n = number of neurons in a layer
w = network weight
b = network bias
r = random number obeying the bernoulli distribution
λ = L2 regularization parameter
i = i-th sample
j = j-th neuron in a layer
l = l-th layer

Acknowledgments

References

Ahmadi, K., Johns, R.T., Mogensen, K., et al. Limitations of current method-of-characteristics (MOC) methods using shock-jump approximations to predict MMPs for complex gas/oil displacements. SPE J. 2011, 16(4): 743-750.

Ahmed, T. A generalized methodology for minimum miscibility pressure. Paper SPE 39034 presented at the Latin American and Caribbean Petroleum Engineering Conference, Rio de Janeiro, Brazil, 30 August-3 September, 1997.

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.

Choubineh, A., Helalizadeh, A., Wood, D.A. 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. Adv. Geo-Energy Res. 2019, 3(1): 52-66.

Chu, H., Liao, X., Chen, Z., et al. A new methodology to assess the maximum CO_2 geosequestration capacity of shale reservoirs with SRV based on wellbore pressure. J. CO_2 Util. 2019, 34: 239-255.

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, 86(2): 173-185.

Edalat, M., Dinarvand, N., Shariatpanahi, S.F. Development of a new artificial neural network model for predicting minimum miscibility pressure in hydrocarbon gas injection. Paper SPE 105407 presented at the SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, 11-14 March, 2007.

Englezos, P., Lee, J.D. Gas hydrates: A cleaner source of energy and opportunity for innovative technologies. Korean J. Chem. Eng. 2005, 22(5): 671-681.

Enick, R.M., Holder, G.D., Morsi, B.I. A thermodynamic correlation for the minimum miscibility pressure in CO_2 flooding of petroleum reservoirs. Reserv. Eng. 1988, 3(1): 81-92.

Glorot, X., Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. Proc. Mach. Learn. Res. 2010, 9: 249-256.

Glorot, X., Bordes, A., Bengio, Y. Deep sparse rectifier neural networks. Proc. Mach. Learn. Res. 2011, 15: 315-323.

Hassan, A., Elkatatny, S., Mahmoud, M., et al. A new approach to characterize CO_2 flooding utilizing artificial intelligence techniques. Paper SPE 192252 presented at the SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, Dammam, Saudi Arabia, 23-26 April, 2018.

Hogg, R.V., McKean, J., Craig, A.T. Introduction to Mathematical Statistics. New York, USA, Pearson Education, 2005.

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): 83-95.

Johns, R.T., Orr Jr, F.M. Miscible gas displacement of multicomponent oils. SPE J. 1996, 1(1): 39-50.

Johnson, J.P., Pollin, J.S. Measurement and correlation of CO_2 miscibility pressures. Paper SPE 9790 presented at the SPE/DOE Enhanced Oil Recovery Symposium, Tulsa, Oklahoma, 5-8 April, 1981.

Kingma, D.P., Ba, J. Adam: A method for stochastic optimization. Paper arXiv:1412.6980 presented at the 3rd

International Conference for Learning Representations, San Diego, 2015.

Lv, G., Li, Q., Wang, S., et al. Key techniques of reservoir engineering and injection-production process for CO_2 flooding in China's SINOPEC Shengli Oilfield. J. CO_2 Util. 2015, 11: 31-40.

Ng, A.Y. Feature selection, L1 vs. L2 regularization, and rotational invariance. Paper presented at ICML' 04 Proceedings of the Twenty-first International Conference on Machine Learning, Banff, Alberta, Canada, 4-8 July, 2004.

Qian, N. On the momentum term in gradient descent learning algorithms. Neural Netw. 1999, 12(1): 145-151.

Riazi, M. Characterization and Properties of Petroleum Fractions. West Conshohocken, USA, ASTM International, 2007.

Srivastava, N., Hinton, G., Krizhevsky, A., et al. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 2014, 15(1): 1929-1958.

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

Sutton, R. Two problems with back propagation and other steepest descent learning procedures for networks. Paper presented at Proceedings of the Eighth Annual Conference of the Cognitive Science Society, 1986.

Whitson, C.H. Characterizing hydrocarbon plus fractions. SPE J. 1983, 23(4): 683-694.

Yellig, W.F., Metcalfe, R.S. Determination and prediction of CO_2 minimum miscibility pressures (includes associated paper 8876). J. Pet. Technol. 1980, 32(1): 160-168.