J. of Supercritical Fluids 107 (2016) 31–37

Contents lists available at ScienceDirect

The Journal of Supercritical Fluids

journal homepage: www.elsevier.com/locate/supflu

Estimation of CO_2–brine interfacial tension using an artificial neural network

Jiyuan Zhang [a], Qihong Feng [a,*], Shuhua Wang [b], Xianmin Zhang [a], Shoulei Wang [c]

[a] School of Petroleum Engineering, China University of Petroleum (East China), No. 66 Changjiang West Road, Economic & Technological Development Zone, Qingdao, China
[b] Schulich School of Engineering, University of Calgary, Calgary, Canada
[c] Research Institute of China National Offshore Oil Corporation, Beijing, China

ARTICLE INFO

Article history:
Received 18 April 2015
Received in revised form 14 August 2015
Accepted 14 August 2015
Available online 19 August 2015

Keywords:
CO_2–brine systems
Interfacial tension
Artificial neural network
Estimation
Empirical correlations

ABSTRACT

Experimental determination of CO_2–brine interfacial tension (IFT) usually requires expensive apparatus and sophisticated interpretation procedure and is time-consuming. Hence, it is of practical importance to develop an accurate and reliable model for determining the CO_2–brine IFT. This paper presents the use of feed forward artificial neural network (ANN) to accurately estimate CO_2–brine IFT based on a database acquired from previous literature. The database consists of a total of 1716 CO_2–brine IFT datasets that cover relatively large ranges of pressure (0.1–60.05 MPa), temperature (5.25–175 °C), total salinity (0–5 mol kg^{-1}) and mole fractions (0–80%) of impure components. Six independent variables were considered to develop the IFT estimation model: pressure, temperature, monovalent cation (Na^+ and K^+) molality, bivalent cation (Ca^{2+} and Mg^{2+}) molality in brine, and mole fractions of N_2 and CH_4 in injected CO_2 streams. The ANN topology was optimized by trial-and-error in order to enhance its capability of generalization and the optimal one was determined to be 6-10-20-1 (10 and 20 neurons in the first and second hidden layers, respectively). The accuracy of the proposed ANN model was highlighted by four evaluation matrices, namely mean absolute error (MAE), mean absolute relative error (MARE), mean squared error (MSE), and determination coefficient (R^2) between the measured and estimated IFT. The ANN model was further compared against four empirical IFT correlations developed in previous studies. It was observed that the ANN model outperforms significantly the empirical correlations and provides the most accurate IFT reproduction with respect to pure CO_2–pure water, pure CO_2–brine and impure CO_2 systems.

1. Introduction

Injection and storage of CO_2 into underground saline aquifers or depleted hydrocarbon reservoirs is one of the most promising options for reducing CO_2 release into atmosphere and slow global temperature rise [1–5]. Storage safety of CO_2 in these geological sequestration operations is by far most threatened by the capillary breakthrough due to excessive CO_2 injection [4], which is predominantly controlled by the interfacial tension (IFT) between CO_2 and host fluids. Therefore, a sound understanding of the IFT behavior between CO_2 and host fluid is of vital importance in designing a CO_2 injection and storage project that can maximize injected volume while ensuring the goal of no significant CO_2 leakage [6].

Numerous studies have been conducted to obtain CO_2–water IFT data by carrying out laboratory measurements [5,7–19]. Among these experimental studies, pendant drop and capillary rise methods are two techniques commonly used for measuring CO_2–brine IFT at high pressures and elevated temperatures [14]. However, laboratory measurement is usually time-consuming and requires an expensive experimental apparatus and sophisticated interpretation procedures. As a comparison, empirical correlation provides an alternative way to fast estimate CO_2-IFT given certain conditions (e.g., pressure, temperature, salinity, etc.). A series of empirical correlations (Table 1) have been proposed using the least square method to estimate CO_2–water IFT in pure CO_2–pure water [9,11,14,20] and pure CO_2–brine [5,8,18] systems. More recently, Li et al. [21] developed a CO_2–brine IFT model using the alternating conditional expectation algorithm. To the knowledge of

* Corresponding author. Tel.: +86 532 86981229.
E-mail addresses: zhjy221@126.com (J. Zhang), fengqihong@126.com (Q. Feng), shuhwang@ucalgary.ca (S. Wang).

http://dx.doi.org/10.1016/j.supflu.2015.08.010
0896-8446/

Table 1
Summary of previous empirical correlations for estimating interfacial tension of CO_2/water systems.

Author	Reference	Year	Correlation	System	Temperature range (°C)	Pressure range (MPa)
Massoudi and King	[20]	1974	$IFT = 71.98 - 0.7789p + 0.00543p^2 - 0.000042p^3$ $IFT = 27.514\left[1 - \exp\left(-35.25\sqrt{dd}\right)\right] + 31.916dd - 91.016 \times dd^2 + 103.233dd^3 + 4.513\exp[351.903(dd - 0.9958)]$	Pure CO_2–pure water	25	0.1–6.2
Hebach et al.[a]	[9]	2002		Pure CO_2–pure water	5.25–60.15	0.1–20.03
Bennion and Bachu[b]	[13]	2008	$IFT = A \times p^{-B}$	Pure CO_2–brine	41–125	2.0–27.0
Chalbaud et al.	[10]	2009	$IFT = 26 + 1.255y_{NaCl} + \left(82\Delta\rho/44010\right)^{4.718} T_r^{1.0243}$	Pure CO_2–brine	27–100	4.8–25.8
Georgiadis et al.[c]	[14]	2010	The isotherms at 25, 40, and 60 °C:$IFT = c_1 - d_1 p$ The isotherms at 70 and 101 °C: $IFT = (c_2 d_2 + mp^n)/(d_2 + p^n)$	Pure CO_2–pure water	25, 40, 60, 70, 101	1–60.05
Li et al.[d]	[18]	2012	$IFT = A_0 m^+ + B_0$	Pure CO_2–brine	25–175	2–50
Li et al.[e]	[21]	2014	$IFT = 40.59259 + 11.41295S - 0.81418S^2 + 0.21986S^3$	Impure CO_2–brine	5.25–175	0.1–60.05

n.a. is not applicable.

[a] $dd = \left((\rho_w - \rho_{corr})/1000\right)^2$, where $\rho_{corr} = \rho_g + 0.00022(30.85 - T)p^{-1.9085}$ for 25 kg m^{-3} < ρ_g < 250 kg m^{-3} and $\rho_{corr} = \rho_g$ in all other cases.

[b] For T = 25 °C, A = 78.819 and B = 0.4168. For T = 41 °C, A = 90.959 and B = 0.5543. For T = 60 °C, A = 82.198, B = 0.3508.

[c] c_1, d_1, c_2, d_2, n are fitting parameters which can be found in Ref. [14].

[d] $A_0 = 0.45101 + 6.2018 \times 10^{-3}p + 3.3654 \times 10^{-3}(T + 273.15)$. $B_0 = -41.203 + 0.43549(T + 273.15) - 7.2502 \times 10^{-4}(T + 273.15)^2 + [-538.9 + 2.0305(T + 273.15)]p^{-1} + [3831.2 - 11.694(T + 273.15)]\,p^{-2} + [-5165.9 + 15.073(T + 273.15)]\,p^{-3}$.

[e] $S = \sum_{i=1}^{6} \varphi_i^*(X_i) = \sum_{i=1}^{6}\sum_{j=0}^{7} a_{ij}x_i^j$, where aij is fitting parameter and xi is independent variable that can be found in Ref. [21].

the authors, Li et al.'s model is currently the only empirical correlation that is applicable of predicting IFT in impure CO_2 and water/brine systems. The overall estimation accuracy of Li et al.'s model, though improved significantly compared with its previous ones, results in significant errors especially for data points with IFT above 60 mN m^{-1}. Yan et al. [8] presented the use of linear gradient theory (LGT) to estimate the IFT between gas mixtures and water. Unfortunately, LGT was found to exhibit poor estimation accuracy when gas mixtures contain high concentrations of CO_2 [21].

Having addressed these problems, there is an important need to develop a comprehensive and reliable CO_2–brine IFT estimation model that can simultaneously consider both impurity in CO_2 phase and salt concentration in water phase. ANNs are powerful mathematical tools which are capable of providing a reliable model of a complex and nonlinear system between inputs and outputs through mimicking the biological neural network behavior. To date, ANNs have seen enormous applications in various engineering fields [22–34]. This paper presents the use of an ANN to accurately estimate CO_2–brine IFT based on experimental data acquired from previous literature reports.

2. Methodology

2.1. Experimental database

In this study, a total of 1716 experimental CO_2–brine IFT data sets (see Appendix A) were acquired from previous studies [4,5,7–19,35] and utilized for developing the ANN model. These data were obtained from pendant drop experiments and consist of both pure and impure CO_2. In impure CO_2 streams, the impure component is either CH_4 or N_2, i.e., impure CO_2 streams are either CO_2/CH_4 or CO_2/N_2 binary mixtures. Brines used for experiments were synthetized by adding one of the following solutes into distilled water: NaCl, KCl, Na_2SO_4, $MgCl_2$ and $CaCl_2$. Previous studies [18,19] have revealed that the effect of cations on CO_2–brine IFT depends mainly on their valency and the effect of different cations with identical valency on IFT is closely analogous. Therefore, in this database, brine type and salinity are characterized by the monovalent cation (Na^+ and K^+) and bivalent cation (Ca^{2+} and Mg^{2+}) molalities and pure water are represented by zero values of both monovalent cation and bivalent cation molalities.

2.2. Basics of ANN

ANNs are powerful mathematical tools that are capable of modeling the complex and nonlinear relationship between inputs and outputs through mimicking the biological neural network behavior. ANN is able to identify and learn the correlated patterns from the training data set between input data set and corresponding output data set. The most basic and commonly used ANN is the multilayer perception (MLP) network [28] that includes input, hidden, and output layers (Fig. 1a). A neural network is composed of a number of highly interconnected and parallel computational elements called neurons or nodes. Each neuron consists of a bias, a transfer function and weights (Fig. 1b) that connect it to other neurons. The operation of a single neuron is dependent on the weighted sum of the incoming signals and a bias term, fed through a transfer function, resulting in an output [34]. The mathematical model of the ith neuron can be described as:

$$\mathrm{Out}_i = f\left(\sum_{j=1}^{n} w_{ij}x_j + b_i\right) \quad (1)$$

where w is the synaptic weight; b is bias; f is transfer function; x is the incoming signals; Out is the output; n is the number of neurons that connect to the ith neuron.

The process of finding suitable weights and biases to correlate output with input variables is called "training". Two phases namely feed forward pass and backward pass constitute the training process. In the feed forward pass, the processing of information is propagated from the input layer to the output layer [27]. During the backward process, the error between the predicted targets by the feed forward process and actual ones are calculated and then propagated back to the input layer to update the weights. The training process is in effect to adjust the weights and biases step by step to