Modeling solubility of carbon dioxide in reservoir brine via smart techniques: application to carbon dioxide storage

Mohammad-Ali Ahmadi1,2*
1Department of Petroleum Engineering, Ahwaz Faculty of Petroleum Engineering, Petroleum University of Technology, Ahwaz, Iran; 2IOR/EOR Research Institute, Tehran, Iran

Abstract
Nowadays, reduction of carbon dioxide (CO2) in the atmosphere via sequestration in deep saline aquifers is studied by numerous researchers. Solubility of CO2 in reservoir brine is a most important parameter in CO2 sequestration saline aquifers. Owing to the importance of this issue, in this paper different methods based on the concept of artificial intelligence, such as particle swarm optimization (PSO), artificial neural network (ANN) and hybrid approaches, are evolved to specify solubility of CO2 in brine at different conditions. The developed intelligent approaches are examined by comparing with precise actual data reported in previous publications. The results gained from the developed intelligent approaches are contrasted with the corresponding real CO2–brine solubility data. The average relative absolute error between the real and the corresponding model prediction data is found to be less than 1% for the hybrid PSO and genetic algorithm model.

Keywords: carbon dioxide reduction; CO2–brine solubility; neural network; CO2 sequestration

Received 4 December 2014; revised 30 May 2015; accepted 9 July 2015

1 INTRODUCTION
Since the early days of the industrial revolution, there have been noticeable, continuous and great efforts made to promote the standards of our lives. Although numerous conspicuous and perceptible steps of progress have clearly been implemented technologically so far, environmental and immediate concerns about increase in the average global surface temperature have repeatedly been voiced and addressed, which has originally been caused by the increase in greenhouse gas concentrations, emitted from different sources, in the atmosphere, particularly CO2 [1–16].

Hence, extending the plans related to the decrease of CO2 emissions and developing schemes to control the global warming phenomenon have necessarily turned into priorities for scientists. Regardless of some up-to-the-minute solutions that researchers and companies have come with such as ocean storing and surface mineral carbonation both of which have been considered as suitable alternatives to make the atmosphere as clean as it used to be in the past, based on the conspicuous obtained outputs and as experienced from the fields of oil and gas operations, deep waste disposal and eventually ground water protection, carbon capture storage (CCS) has noticeably been proved as a key, cutting-edge and vital technology that is crucial in modifying the effects of climate change [17–40].

Technically, the related processes of CCS in deep saline aquifers can economically be described as follows: in order to maximize the possible profits, the supposed volume of CO2 ought to be injected in the supercritical state, the thermodynamical conditions occurring elementally in pressure and temperature being 1070.378 psi and 87.98 °F, respectively. Through considering the given geothermal gradients, the depth of 800 m and more making the referral circumstances naturally fetch. In more detail, the density of supercritical CO2 is in the range of 15 605–46 816 lbm/ft3 and its viscosity is between 2 and 20% of water viscosity. Furthermore, the intended zone is supposed to have the industrial capability of handling the injection and storage parts of the considered projects along with possessing the inherent characteristic of preventing the upward escape of supercritical CO2 as a consequence of the buoyant force after its long passage, which it might result in a natural disaster [2, 41–50].

The injected plume of CO2 gets chronologically and mainly stored in the porous media of the aquifer through four different mechanisms, which are residual gas trapping, structural trapping, mineral trapping and solubility trapping. Adding hydrodynamic...
trapping, which is an assortment of all the other trapping mechanisms that take place concurrently at different scales of time, upgrades the number of storing mechanisms to five. Further information about hydrodynamic trapping and other mentioned mechanisms can be found in Refs [31, 51–56]. Bringing particularly and briefly solubility trapping to our attention, the injected plume of carbon dioxide, through dispersion, convection and diffusion, dissolves earlier in the present saltwater in which it forms. The rate of solubility trapping is mostly determined by the amount of carbon dioxide gas that comes into contact with the brine formed that is unsaturated with carbon dioxide. Therefore, owing to the economical aspect of the discussed issue, it can be inferred that the CO$_2$ solubility in brine is key, vital and crucial. To put it another way, to set the necessary drilling facilities and increase the wellhead utilization of injection, the storability of the situated fluid in the target formation should be the motivating factor [7, 48, 57, 58].

The density of the brine saturated with CO$_2$ is another essential feature that is required to be thoroughly, carefully and technically considered. Since dissolution of CO$_2$ builds the brine density up and causes a convection flow, the differential density between brine with and without dissolved CO$_2$ is therefore one of the major factors that determine how CO$_2$–brine mixtures move downwards in the aqueous phase, because its weight is heavier than the brine by itself, and how fast this movement will be, which inclusively can reduce the mobility of the aquifer [59–62].

With the aim of doing investigations about two referred primary attributes of a CO$_2$–brine mixture, CO$_2$ solubility in NaCl brine and CO$_2$-saturated brine density, Yan et al. [13] planned initially an elaborate literature review and afterwards, they deduced that neither are all the models nor are all the data that have ever been suggested or generated appropriate to be referenced under the conditions of high pressure high temperature (HPHT). Next, they designed an experimental procedure to produce compatible outcomes with HPHT conditions. After that, they made a reasonable and logical comparison between the generated data of their laboratory test and existing models of commercial simulators [13].

Through this work, potential application of various intelligent approaches such as artificial neural network (ANN) that optimized by different algorithm optimization such as the genetic algorithm (GA) are suggested to estimate the solubility of CO$_2$ in saline water for carbon capture and sequestering purposes. Evolutionary algorithms are employed in this paper to specify the initial weights of the parameters used in an ANN. The intelligent approaches developed in this paper are evaluated by employing accurate actual data from Yan et al. [13]. Outputs gained from the aforementioned intelligent approaches are contrasted with the corresponding actual data and explicated in greater detail throughout this paper.

2 ARTIFICIAL NEURAL NETWORK

Briefly, an ANN actually is an elementary approach based on simulation of human brain procedures with the ability of learning and prediction with a mathematical structure in order to figure out the relationship between outputs and inputs of each system. This network is trained with the existing actual data during the learning processes and is implemented for elucidating the vague and ambiguous data [63, 64]. An ANN includes straightforward synchronous treating components that are known as nodes or neurons placed in layers. Generally an ANN has three layers: an input layer, a specified number of hidden layers and an output layer. The most common type of ANN in petroleum engineering is the multilayer perceptron (MLP), which is a back-propagation (BP) feed-forward network [65, 66]. MLP networks score over conventional approaches with respect to their ability to analyse related information and their much reduced development time [67–69].

In addition, the weight and biases must be optimized in this study, using a trial and error procedure in the training stage, as described in our previous investigation. Generally, the available experimental data were divided into two sets, namely training and testing sets. The training set was carried out to determine the appropriate network structure, while the testing set, which was not considered in the training stage, was implemented to examine the applicability of the proposed proper network to correlate the CO$_2$–brine solubility. The ANN employed various optimization approaches to specify optimum interconnection weights to train itself and approach the real data. The mean squared error (MSE) of the ANN is expressed as follows:

$$\text{MSE} = \frac{1}{2} \sum_{k=1}^{G} \sum_{j=1}^{m} [Y_j(k) - T_j(k)]^2.$$ 

Here $m$ stands for the number of output neurons, $G$ denotes the number of training data, $Y_j(k)$ stands for the expected outcome and $T_j(k)$ represents the real outcome. When the MSE approaches zero, the error of our developed network model is lowered.

3 HYBRID EVOLUTIONARY ALGORITHMS AND ANN

3.1 Genetic algorithm

The GA, which implements a direct route to Darwinian natural selection and genetics in biological systems, is a promising alternative to the conventional approaches. According to the Darwinian principle of ‘survival of the fittest’, the genetic algorithm (GA) can obtain the best route after a series of loop computations. The search process is composed of artificial mutation, cross-over and selection operators [70–73]. The methodology of the GA is depicted in Figure 1. Moreover, more details about the GA-ANN approach can be found in Ahmadi and Golshadi [73].

3.2 Particle swarm optimization

The particle swarm optimization (PSO) algorithm inspired the behavior of social organisms such as flocks of birds. As with
3. Hybrid GA and PSO

Even though the implementation of the GA has been effectively accomplished in an extensive range of engineering and design issues [72, 73, 81], it is still a time-consuming process if carried out for wide-ranging optimizations that need numerous function assessments for convergence. Consequently, to overcome the aforementioned problem (e.g., costs and high computation efforts), there is great value in linking the GA and PSO into one approach for the purpose of profit from the suitable features and penetrating attributes of both approaches in estimating chosen factors. In this paper, the hybrid of the GA and PSO called ‘HGAPSO’, initially suggested by Juang (2004), is implemented to estimate the CO$_2$—brine solubility. Figure 3 illustrates the box chart of the proposed hybrid GA and PSO method [81].

3.4 Imperialist competitive algorithm

A newly introduced evolutionary algorithm based on the human being’s socio-political evolution was developed firstly by Atashpaz-gargari et al. (2007) and called the imperialist competitive algorithm (ICA). As with other evolutionary optimization algorithms, the ICA launches with initial populations named ‘countries’. In the ICA approach, two types of countries exist as follows: imperialist and colonies which together form empires. In the imperialistic competition process, imperialists make attempts to gain more colonies. As a result, during the competition, the potent imperialists will augment their power and the weak ones will decrease in power. At the end of this aforementioned approach, the most potent imperialist remains in the world and the rest of the countries become colonies of this sole empire [81–83]. The application processes of the suggested corresponding approach based on the ICA are depicted in Figure 4 [81–83].
4 RESULTS AND DISCUSSION

4.1 ANN output results

Through this work, three independent variables such as \( P \) (pressure), \( T \) (temperature) and NaCl molality were implemented as inputs of the constructed network model to estimate the \( \text{CO}_2 \)-brine solubility. Based on sensitivity analysis, the high performance network construction was 3-7-1 (see Figure 5). The constructed approach trained with the BP method by implementing the Levenberg–Marquardt algorithm to specify the \( \text{CO}_2 \)-brine solubility by using the three aforesaid inputs.

To show the accuracy of the introduced network approach and further hybrid models, 32 data points were discriminated for network training and the remaining 22 samples were put aside to be undertaken for testing and validating the network's precision and effectiveness. In addition, to compare output results of various introduced intelligent approaches, the robust and routinely implemented performance criteria should be carried out while different predictive models are implemented. To tackle the addressed issue, two statistical indexes that are routinely implemented as performance criteria in different engineering approaches were selected to quantify the performance of each intelligent model. The crucial point to be noted here is that a learning coefficient of 0.71 and a momentum correction factor of 0.001 were used for the back-propagation training algorithm. As demonstrated in Figure 6, neural network outputs in contrast with corresponding experimental data do not have good agreement, while some experimental points were well estimated. As previously mentioned, to quantify the performance of each model, the correlation coefficient was determined and based on this criterion, as illustrated in Figure 7, the neural network model has unsatisfactory performance due to an intermediate correlation coefficient which is lower than 0.8.

4.2 Hybrid evolutionary algorithms and ANN output results

As earlier mentioned in this paper, four robust popular optimization algorithms: the genetic algorithm, ICA, particle swarm optimization and a hybrid of them were implemented through this research to achieve enhanced performance and precision of the introduced network model by means of deciding connection weights of the neural network. To defeat the referred obstacle of this section, the main goal of each optimization algorithm is to approach the defined fitness or achieve the objective function to its relevant optimum value; in this research the MSE was implemented as an objective function of optimization algorithms and the critical point to be noted here is that every weight in the neural network must be between \(-1\) and \(+1\).
Figure 3. Flow chart of the hybrid genetic algorithm and particle swarm optimization process [81].
To examine the effectiveness and robustness of the evolutionary algorithms to improve the performance of the ANN, a BP-ANN was implemented with the same data assortment as that in the ANN approach. For every item, 25 runs with different randomly generated populations were carried out. The hybrid genetic algorithm and particle swarm optimization-artificial neural network (HGAPSO-ANN) algorithm was run considering a population size of 100.

Various cross-over possibilities are used to figure out the optimal cross-over probability. To unravel this issue, cross-over possibilities are used to figure out the optimal cross-over probability. To unravel this issue, cross-over
Figure 5. Architecture of the three-layer ANN.

Figure 6. Measured vs. predicted CO$_2$ solubility (BP-ANN): (a) training (b) testing.

Figure 7. $R^2$ for the BP-ANN model: (a) training (b) testing.
possibilities in ranges of 0.5–0.92 are implemented in this paper. Accuracy sensitivity analysis according to suggested statistical indexes illustrated that the convergence rate decreases as the cross-over probability rises, while best performance was gained for a cross-over probability of 0.92. In addition, the same route was implemented to exhibit sensitivity of genetic accuracy as a function of mutation probabilities. To assess this end, various mutation rates in the range of 0.0001–0.05 such as 0.05, 0.025, 0.0125, 0.001 and 0.0001 are selected and evaluated. One should keep in mind that poorer answers and fast convergence is a feature of low mutation possibilities. Conversely, as the mutation rate rises, it gives better answers; however, it prevents a high degree of convergence. Based on the statistical accuracy norms employed in this work, the uniform mutation possibility and the uniform cross-over possibility are set to 0.025 and 0.92, correspondingly.

Through implementation of the ICA to determine optimum connection weights of the ANN, based on the sensitivity analysis performed, the number of colonies and the imperialists considered are 40 and 4, correspondingly, and the parameter $\beta$ was considered to be 2.

The input parameters of an ANN system are commonly of various types with various orders of magnitudes, such as pressure ($P$) and temperature ($T$) in this paper. Therefore, it is mandatory to normalize the output and input parameters according to the scales of implemented data as they fall within a specific scale. It is worth mentioning that the legend of the vertical axis in Figures 8–15 exhibits normalized CO$_2$–brine solubility, which was calculated using the following equation:

$$\text{Normalized CO}_2 = \frac{2(\text{CO}_2 - \text{CO}_{2\min})}{(\text{CO}_{2\max} - \text{CO}_{2\min})} - 1,$$

where $\text{CO}_{2\min}$ and $\text{CO}_{2\max}$ are the minimum and maximum solubility of CO$_2$–brine, respectively, among the data implemented in this work.

Clearly, the outputs of the HGAPSO model, in contrast to the results of the GA, PSO and ICA, in comparison with corresponding experimental data demonstrated in Figures 8–11 are in a good conformity with the real CO$_2$–brine solubility data. This indicates that training the neural network model by HGAPSO (Figure 10) leads to better results than the GA, PSO, ICA and the BP approach (Figures 6, 8–11). As earlier mentioned in this research, two different statistical criteria such as the MSE and the correlation coefficient ($R^2$) are introduced to quantify the robustness and effectiveness of various intelligent approaches [84, 85]. Based on introduced performance indexes, the performance criteria are MSE = 0.047543 and $R^2 = 0.97$ for HGAPSO in comparison with MSE = 0.26151 and $R^2 = 0.4349$ for the BP-ANN, MSE = 0.19723 and $R^2 = 0.8444$ for the Whitson model, MSE = 0.11936 and $R^2 = 0.8622$ for the genetic approach, MSE = 0.11055 and $R^2 = 0.879$ for the particle swarm

![Figure 8](https://example.com/fig8.png)  
*Figure 8. Measured vs. predicted CO$_2$ solubility (GA-ANN): (a) training (b) testing.*

![Figure 9](https://example.com/fig9.png)  
*Figure 9. Measured vs. predicted CO$_2$ solubility (PSO-ANN): (a) training (b) testing.*
Figure 10. Measured vs. predicted CO₂ solubility (HGAPSO-ANN): (a) training (b) testing.

Figure 11. Measured vs. predicted CO₂ solubility (ICA-ANN): (a) training (b) testing.

Figure 12. R² for the GA-ANN model: (a) training (b) testing.

Figure 13. R² for the PSO-ANN model: (a) training (b) testing.
optimization model and MSE = 0.10572 and $R^2 = 0.8984$ for the imperialist approach; this confirms the high degree of accuracy of HGAPSO (Figures 12 and 16). Based on statistical approaches for determination of correlation coefficients of the model, three ranges of correlation coefficients were introduced as follows [84, 86]:

1) Higher than 0.9 indicates high performance.
2) Between 0.8 and 0.9 indicates good performance of the approach.
3) Less than 0.8 indicates unsatisfactory precision.

Figures 11–16 illustrate the extent of the fit between the real and estimated CO$_2$–brine solubility values by the GA-ANN, PSO-ANN, HGAPSO-ANN, ICA-ANN and Whitson approaches in terms of a scatter plot, correspondingly. Evidently, the HGAPSO-ANN provides results in good conformity with real measured data. The estimations which follow actual experimental values should fall on the line $Y = X$. Almost all of the data lie on this line, which endorses the precision of the HGAPSO-ANN model. It can be deduced that the Hybrid PSO and GA model has the ability of preventing being trapped in local optima for estimating CO$_2$–brine solubility, since this model has both local and global penetrating capabilities.

Performance plots based on the MSE for the proposed hybrid of evolutionary approaches and the ANN model are illustrated in Figures 17–21, respectively. It can be deduced that the convergence of the hybrid genetic algorithm and particle swarm optimization (HGAPSO-ANN) is considerably faster than other approaches such as the GA, ICA, PSO and the conventional algorithm (BP-ANN) (Figures 17, 18, 20 and 21).

Figure 22 illustrates the CO$_2$ solubility change against relevant pressure when $T = 413K$, for all the employed models such as the PSO-ANN, ICA-ANN and GA-ANN. As can be clearly seen
Figure 17. Performance plot for the proposed GA-ANN model.

Figure 18. Performance plot for the proposed PSO-ANN model.

Figure 19. Performance plot for the proposed HGAPSO-ANN model.

Figure 20. Performance plot for the proposed ICA-ANN model.

Figure 21. Performance plot for the proposed BP-ANN model.

Figure 22. Comparison between experimental and estimated CO₂ solubilities against relevant pressure when T = 413K.
from Figure 22, the HGAPSO-ANN model has high precision and integrity in comparison with the Whitson model.

The MSE and $R^2$ values for the six different approaches carried out to validate the developed intelligent approaches are demonstrated in Table 1. As reported in Table 1, the performance of the HGAPSO-ANN based on two introduced performance indexes is superior to other models such as the Whitson model, while hybrid methods have high accuracy and robustness in comparison with the back-propagation algorithm. It can be deduced that the HGAPSO-ANN model explicated in this paper offers excellent accuracy in both global optima attainment and convergence rate.

### Table 1. Performances of the various intelligent approaches vs. Whitson model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>ICA-ANN</th>
<th>BP-ANN</th>
<th>Whitson</th>
<th>GA-ANN</th>
<th>PSO-ANN</th>
<th>HGAPSO-ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.10572</td>
<td>0.26151</td>
<td>0.19723</td>
<td>0.11936</td>
<td>0.11055</td>
<td>0.047543</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.8984</td>
<td>0.4349</td>
<td>0.8444</td>
<td>0.8622</td>
<td>0.879</td>
<td>0.97</td>
</tr>
</tbody>
</table>

5 CONCLUSIONS

Various intelligent approaches such as hybrid evolutionary algorithms and an artificial neural network algorithm were developed and demonstrated in detail through this work to estimate CO$_2$–brine solubility, which efficiently combines the local and global penetrating attributes of introduced intelligent approaches for CO$_2$–brine solubility prediction. Real measured data from the previous works [35] are used to assess the robustness and precision of the proposed predictive intelligent approaches.

The following deductions are drawn from the results of this paper:

1. The estimation accuracy of the developed approach is outstanding compared with that of conventional BP-ANN and Whitson approaches.
2. The Hybrid PSO and GA model has the ability of preventing being trapped in local optima for estimating CO$_2$–brine solubility, since this new employed model has both local and global penetrating capabilities.
3. The BP-ANN and Whitson models are not suitable for predicting CO$_2$–brine solubility; however, there is good conformity between relevant real measured data and model results while implementing the HGAPSO-ANN model.
4. The proposed CO$_2$–brine solubility estimation approaches, especially the HGAPSO model, can be combined with current CO$_2$ sequestration modeling softwares to improve robustness and effectiveness by increasing their estimation and modeling abilities and reducing doubt.
5. The proposed neural network structure was specified manually. An auxiliary approach is necessary to be developed to utilize the evolutionary algorithms for ANN architecture optimization.
6. When considering the neural network model for estimation of CO$_2$–brine solubility, specifying the optimum ANN architecture poses a big question that needs further investigations.

REFERENCES


