SPE-218906-MS

Parametric Study of CO_2 Sequestration in Deep Saline Aquifers Using Data-Driven Models

M. I. Khan and A. Khanal, Jasper Department of Chemical Engineering, Tyler, Texas, USA

This paper was prepared for presentation at the SPE Western Regional Meeting held in Palo Alto, California, USA, 16 - 18 April 2024.

Abstract

Large-scale geo-sequestration of anthropogenic carbon dioxide (CO_2) is one of the most promising methods to mitigate the effects of climate change without significant stress on the current energy infrastructure. However, the successful implementation of CO_2 sequestration projects in suitable geological formations, such as deep saline aquifers and depleted hydrocarbon reservoirs, is contingent upon the optimal selection of decision parameters constrained by several key uncertainty parameters. This study performs an in-depth parametric analysis of different CO_2 injection scenarios (water-alternating gas, continuous, intermittent) for aquifers with varying petrophysical properties. The petrophysical properties evaluated in this study include aquifer permeability, porosity, relative permeability, critical gas saturation, and others. Based on the extensive data collected from the literature, we generated a large set of simulated data for different operating conditions and geological settings, which is used to formulate a proxy model using different machine learning methods. The injection is run for 25 years with 275 years of post-injection monitoring. The results demonstrated the effectiveness of the machine learning models in predicting the CO_2 trapping mechanism with a negligible prediction error while ensuring a low computational time. Each model demonstrated acceptable accuracy (R^2 >0.93), with the XGBoost model showing the best accuracy with an R^2 value of 0.999, 0.995, and 0.985 for predicting the dissolved, trapped, and mobile phase CO_2. Finally, a feature importance analysis is conducted to understand the effect of different petrophysical properties on CO_2 trapping mechanisms. The WAG process exhibited a higher CO_2 dissolution than the continuous or intermittent CO_2 injection process. The porosity and permeability are the most influential features for predicting the fate of the injected CO_2. The results from this study show that the data-driven proxy models can be used as a computationally efficient alternative to optimize CO_2 sequestration operations in deep saline aquifers effectively.

Keywords: CO_2 sequestration, Saline aquifers, Machine learning, Reservoir simulation, WAG process

Introduction

The increasing levels of carbon dioxide (CO_2) emission into the atmosphere from human activities such as burning fossil fuels and deforestation have raised serious concerns about their impact on the environment

and climate. This increased emission has led to global warming, causing significant changes in weather patterns, melting polar ice caps, and rising sea levels. To mitigate these impacts, there is a growing need to find ways to capture and sequester anthropogenic CO_2. One such approach is CO_2 sequestration, where CO_2 is captured from the atmosphere and industrial processes and then injected into geological formations such as depleted oil, deep saline aquifers, gas reservoirs, and unmineable coal beds (Klara et al., 2003). The International Energy Agency (IEA) predicts that by 2030, over 220 Mt of CO_2 will be captured annually (IEA, 2022), and the deep saline aquifers have the most significant storage potential for larger-scale migration of the captured CO_2 (Wei et al., 2022; Yu et al., 2021). However, despite offering significant storage potential, implementing carbon capture and storage (CCS) projects in deep saline aquifers is challenging due to the various technical and economic uncertainties. These challenges include issues related to site characterization, leakage, monitoring, and economic feasibility (Bachu, 2008; Lucier & Zoback, 2008). Therefore, properly investigating the site geology and optimizing the relevant parameters are important for successfully implementing CO_2 sequestration projects in deep saline aquifers.

Numerous experimental and numerical studies have evaluated different aspects of CO_2 storage in deep saline aquifers. The experimental studies conducted on CO_2 sequestration in deep saline aquifers are focused on exploring various aspects of the sequestration process, such as investigating different trapping mechanisms (dissolution, mineralization, and residual), determining storage capacity, enhancing the trapping capacity, and evaluating other relevant mechanisms (Bachu, 2002). However, experimental studies on deep saline aquifers for CO_2 sequestration are limited by several factors, including the high-cost experiments, safety concerns associated with the large amount of CO_2, and the limitations of techniques used to monitor and measure the storage amount of CO_2. Additionally, the small scale of the experiments can lead to difficulties in extrapolating results to a larger scale and accurately predicting the behavior of CO_2 storage in the subsurface over long periods (Liu et al., 2019; Oloruntobi & LaForce, 2009).

Physics-based numerical simulation is a widely used approach to study the CO_2 sequestration process in deep saline aquifers. This approach is particularly useful as it offers an economical and non-invasive solution to assess the CO_2 trapping mechanisms in deep saline aquifers, and it can also be used to complement and validate the findings from experimental studies. Kumar et al. (Kumar et al., 2004) conducted a compositional reservoir simulation of a typical CO_2 sequestration project in a deep saline aquifer to comprehend and quantify the sequestrated CO_2 in different mechanisms. They altered the geological parameters of the base case model to present that aqueous and mineral trapping is the most efficient mechanism for CO_2 storage. Based on the cases of their study, they concluded that the capacity of CO_2 storage by residual trapping is higher than mineral trapping. Rasheed et al.(Rasheed et al., 2020) concluded that low to medium-level heterogeneous aquifers with good porosity (>20%) are suitable for CO_2 storage. The understanding of the detailed mechanism of subsurface sequestration is limited due to the scarcity of information about the heterogeneity and the geometry of the aquifer selected for the process (Bachu et al., 2007). The outcome of CO_2 sequestration is contingent upon various factors such as the injection rate, pressure, strategy, and the design of the injection well. In their study, Calabrese et al. (Calabrese et al., 2005) aimed to understand the physical and chemical processes that occur during CO_2 sequestration in a depleted gas reservoir located in northern Italy. They concluded that an optimum injection rate must be obtained to maximize the storage capacity. At a higher injection rate, the storage capacity is reduced as the gas channels through high permeability streaks, and with a lower injection rate, denser CO_2 sinks to the bottom of the gas zone and dissolves into the aquifers. The researchers also concluded that factors such as molecular diffusion, dispersion, and geochemistry is less important in assessing CO_2 storage. Water alternating gas (WAG) has been found to be effective in increasing the aquifer's storage capacity and improving the efficiency of the CO_2 sequestration process. Pan et al. (Pan et al., 2016) compared the WAG process with continuous CO_2 injection schemes in their investigation. They concluded that in WAG process, the water helps to create pathways (increased effective porosity by 2.7% and permeability by 8.4%) for the CO_2 to flow into the

aquifer, and the pressure changes caused by the alternation of the water and CO_2 help to keep the CO_2 in place within the aquifer. Al-Khdheeawi et al. (Al-Khdheeawi et al., 2018a) also concluded that the WAG process improves CO_2 storage efficiency and reduces the vertical CO_2 leakage risk. Several other investigations have explored the different aspects of CO_2 storage in deep saline aquifers, including the effect of uncertain geological and decision parameters of the storage capacity (Khanal & Shahriar, 2022), trapping efficiency, and risk associated with leakage of the storage (Al-Khdheeawi et al., 2018b), application of nanoparticles for quick mixing of CO_2 in brine (Singh et al., 2012).

Despite their versatility, numerical reservoir simulations can be difficult, time-consuming, and computationally intensive. Processing the simulation data for sensitivity analysis and optimization can be time-consuming, and each new simulation requires a significant amount of complex geological data. While a well-validated reservoir simulation model is necessary for mature projects and requires a lot of geological data, more accurate, reliable, and quicker results are often needed for initial screening and evaluation studies.

Machine learning (ML) models can be adapted as a practical tool for CCS projects. ML algorithms are able to learn from historical data, make predictions, and perform optimizations more efficiently and with greater speed than traditional simulation methods. For instance, Song et al. (Song et al., 2020) used an artificial neural network (ANN) to forecast the Capacity for CO_2 storage by considering a synthetic model to generate the training dataset. Kim et al. (Kim et al., 2017) also used ANN to predict the storage efficiency of CO_2 sequestration in deep saline aquifers. They generated 150 different reservoir models by altering properties such as permeability, porosity, residual oil saturation, thickness, and depth, which were then utilized to create a proxy model based on an artificial neural network (ANN). Menad et al. (Menad et al., 2019) utilized advanced machine learning models, such as multilayer perceptron (MLP) and radial basis function neural network (RBFNN), to predict the solubility of CO_2 in brine for monitoring CO_2 sequestration in saline aquifers. They applied different algorithms to optimize the models and concluded that RBFNN optimized with artificial bee colony (ABC) algorithm is the most reliable one for predicting the solubility of CO_2 in brine. Recently Safaei-Farouji et al. (Safaei-Farouji et al., 2022) worked on predicting the carbon trapping efficiency in storage formations by using four different ML algorithms: neuro-fuzzy inference system (ANFIS), extra tree (ET), random forest (RF) and radial basis function (RBF). They extracted 1868 data points of residual and solubility trapping of CO_2 from the published literature for training, testing, and validation. They concluded that RF provides better accuracy in predicting residual and solubility trapping with the R^2 value of 0.995 and 0.965, respectively. They also concluded that depth and post-injection time are the most influential factors for residual and solubility trapping performance.

In summary, this article investigates the potential effect of different injection scenarios for CO_2 trapping and compared the effectiveness of the WAG process over continuous and intermittent CO_2 injection. This work will use various machine learning algorithms to optimize the ultimate storage capacity of CO_2 in deep saline aquifers. The paper is structured as follows: First, detailed numerical simulations with varying conditions are conducted using a 3D model to evaluate various aspects of CO_2 sequestration in a deep saline aquifer. A finite reservoir is considered for the base case model where CO_2 is injected for 25 years, after which the reservoir will be monitored for a further period of 275 years to evaluate the efficiency of different trapping mechanisms (dissolution, residual, and structural trapping). After that, numerous iterations are conducted by changing operating conditions and reservoir properties. Finally, the generated data is used to train the XGBoost, RF, and ANN models for prediction. The proposed methodology can be utilized to predict CO_2 sequestration efficiency without extensive and time-consuming compositional reservoir simulations.

Materials and Methods

Reservoir Model

The CO_2 sequestration process is simulated in the GEM reservoir simulation package by the computer modeling group (CMG-v2023.40). This model considers a multi-component system of brine and CO_2 through component transport equations and thermodynamic equations between the gas and aqueous phases. In this simulation, the foundation for the mathematical models is established through the application of mass balance equations for every key component or phase present in the system (Celia et al., 2015). The accumulation of CO_2 in a closed system with α phases and i components can be calculated using the governing equation, which takes into account the contributions from convective mass transfer, diffusive/dispersive mass transfer, consumption due to chemical reactions, and the mass of the injected gas. The governing equation is-

$$\sum_{\alpha}\frac{\partial}{\partial t}\rho_{\alpha}\phi s_{\alpha}m_{\alpha i}=\sum_{\alpha}\nabla\cdot\left(\rho_{\alpha}u_{\alpha}m_{\alpha i}+j_{\alpha}\right)+r^{i}+\psi_{\alpha}{}^{i} \quad (1)$$

Where, ρ_{α} is the density of the phase α, ϕ is the porosity, s_{α} is the saturation of the fluid phase α, $m_{\alpha i}$ is the mass fraction of the i component in α phase, $\boldsymbol{u}_{\alpha}$ is the Darcy flux for phase α, $\boldsymbol{j}_{\alpha}$ is the nonadvective flux of the i component in α phase and $\psi_{\alpha i}$ represents external sources or sinks of mass of component i in α phase.

A synthetic single-well CO_2 injection model shown in Fig.1 is used for the base case model in this study. This study employs a limited reservoir size to reduce computational time, with a potential for scaling to accommodate larger reservoirs (Khanal & Weijermars, 2019). The homogeneous model uses uniform porosity and permeability and is 500 m × 500 m × 125 m with a uniform grid system of 31,250 cells divided into 50 × 25 × 25 grids. The total pore and block volumes are calculated as 4.687×10^6 m^3 and 3.125×10^7 m^3, respectively. 300 m^3/day of CO_2 is injected at a depth range of 2365 m – 2375 m through the injector placed at the corner of the model. The perforation of the well was performed on the 23rd to the 25th layer and covered a length of 10 meters. The CO_2 is injected into the aquifer for 25 years, followed by a 275-year post-injection period. The geological properties, including input parameters of the saline aquifer model, are presented in Table 1.

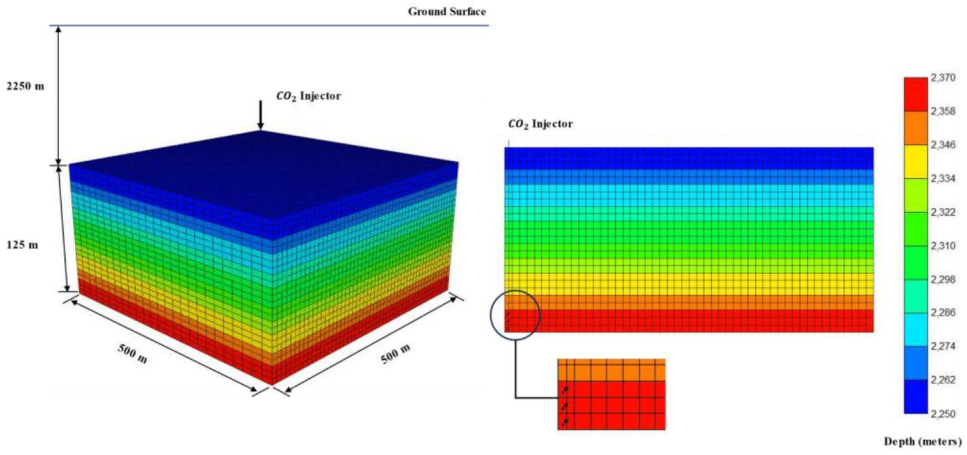

Figure 1—A visual representation of the deep saline aquifer for CO_2 storage is shown through a 3D view on the left and a cross-section view on the right.

Table 1—Saline aquifer system input parameters.

Aquifer Properties	Values
Reservoir temperature (°C)	50
Max. bottom hole pressure (KPa)	44,500
Grid number	31,250 (50×25×25)
Length (m)	500
Width (m)	500
Depth at the top (m)	225,0
Thickness (m)	125
Permeability (mD)	200
Porosity	0.15
SC_{gas}/PV	0.2

*SC = surface condition, PV = pore volume

Relative permeability is an important parameter and is used to determine the effective permeability of CO_2 and brine. The relative permeability curve shown in Fig. 2 has been sourced from Khanal et al. (Khanal et al., 2024) for the reservoir model used in this study. The thermodynamic equilibrium between the gas and the aqueous phase determines the solubility of CO_2 in the aqueous phase. Peng Robinson's Equation of State (PR EOS) and Henry's law (Ali et al., 2014) are used to calculate the gas phase's fugacity and the components dissolved in the aqueous phase, respectively. The Harvey correlation is used to calculate Henry's law constant at varying temperatures and pressures (Ali et al., 2014).

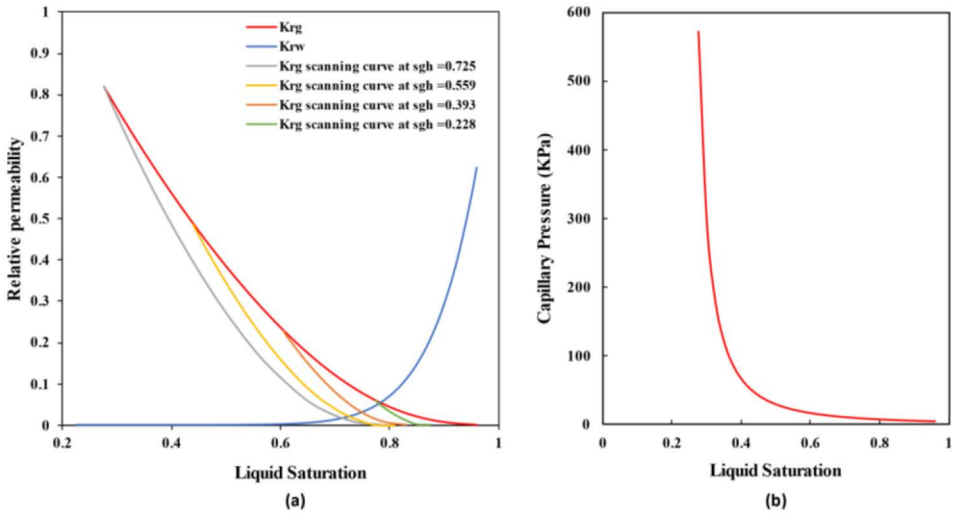

Figure 2—(a) Relative permeability and scanning curve for drainage and imbibition and (b) Capillary pressure curve (Khanal et al., 2024).

Workflow for Developing Machine Learning Models

The data generated from the simulation model is prepared for the testing, training, and validation of the model. Then, feature scaling is applied to transform the input and target data into the same scale. In this work, data normalization has been done into a standard range of -1 to 1 by using the following formula-

$$x' = 2\left(\frac{x - min(x)}{max(x) - min(x)}\right) - 1 \quad (2)$$

Where x represents the original value of the given parameter, x′ represents the scaled value of x, min(x) is the minimum value of x, and max(x) represents the maximum value of x.

The phase then begins by subdividing the dataset into training, testing, and validation. A K-fold (K=5) cross-validation method is applied during the selection of training and testing. Next, the multilayer perceptron (MLP), Random Forest (RF), and Extreme gradient boosting (XGBoost) machine learning models are applied to the target dataset. Then, finally, the performance of each model is analyzed using various statistical indexes described below-

$$\text{Coefficient of determination } (R^2) \quad : \quad R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i_{sim}} - y_{i_{pred}}\right)^2}{\sum_{i=1}^{n}\left(y_{i_{pred}} - \bar{y}\right)^2} \tag{3}$$

$$\text{Mean Squared Error } (\text{MSE}) \quad : \quad MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{i_{sim}} - y_{i_{pred}}\right| \tag{4}$$

$$\text{Mean Squared Error } (\text{MSE}) \quad : \quad MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_{i_{sim}} - y_{i_{pred}}\right)^2 \tag{5}$$

$$\text{Root Mean Squared Error } (\text{RMSE}) \quad : \quad RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{i_{sim}} - y_{i_{pred}}\right)^2} \tag{6}$$

Where, $\bar{y}$ indicates the average value, $y_{i_{sim}}$ and $y_{i_{pred}}$ denotes the simulated and predicted values, respectively. Fig.3. Illustrates the general workflow of applying ML models.

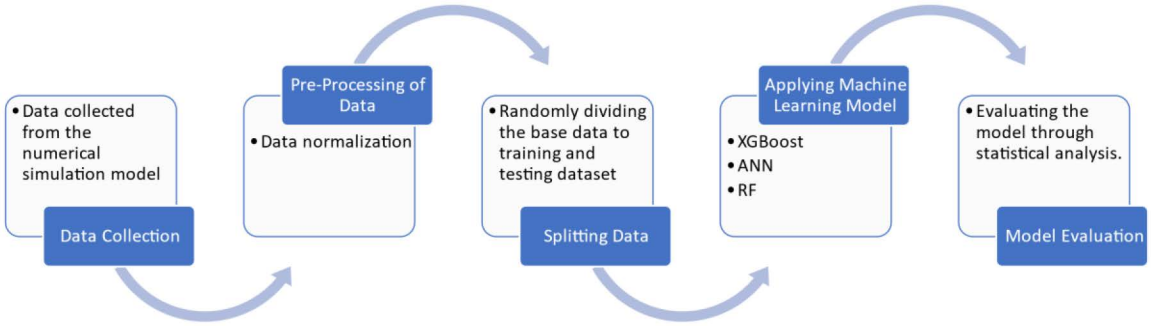

Figure 3—General workflow of applying machine learning algorithms for predicting the CO_2 trapping mechanisms in the deep saline aquifer.

Result and Discussion

Comparison between Continuous, Intermittent, and WAG Injection Processes

This section analyzes continuous and intermittent WAG injection processes to assess the CO_2-trapping efficiency. For all the processes, an equal amount of CO_2 is injected for 25 years and then monitored for 275 years. The comparison between continuous, intermittent, and WAG processes demonstrated that a higher amount of CO_2 dissolved in the WAG process. Fig.4 presents the extent of CO_2 trapped by dissolution, residual, and structural trapping over the simulation period (300 years).

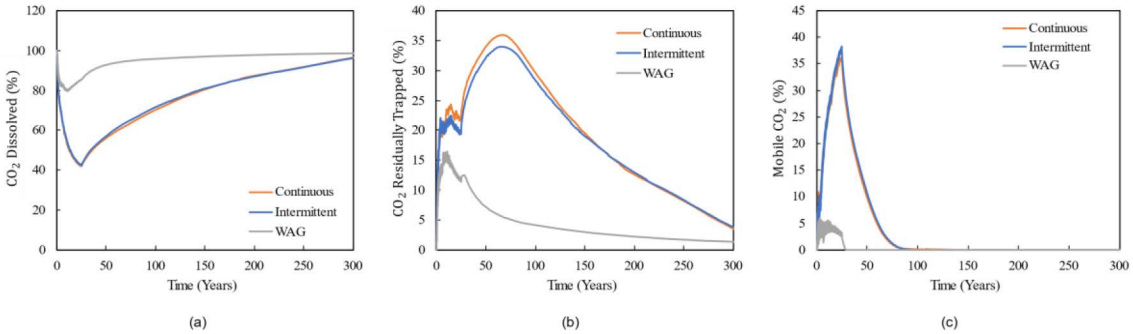

Figure 4—a) Dissolved, (b) Residually trapped, and (c) Mobile phase CO_2 in Continuous, Intermittent, and WAG injection process after 300 years of period.

After the injection period (25 years), ~85% of the injected CO_2 is dissolved in the WAG process, whereas in the intermittent and continuous injection process, ~42% of the injected CO_2 is dissolved (Fig. 5). In the WAG process, the larger CO_2-water contact surface area increases the lateral spreading of CO_2 (Doughty, 2010); hence, more CO_2 comes into contact with the fresh brine and get dissolved. During the initial period, it is observed that more CO_2 tends to be trapped residually in the intermittent injection process. However, no significant differences in dissolution trapping are observed for the intermittent and continuous injection process. The reason behind the initial differences in residual trapping for intermittent and continuous processes is due to the time lag between imbibition and maximum CO_2 gas saturation achieved before imbibition starts. As no significant imbibition is expected to happen before the injection stops (Doughty, 2007), the intermittent injection process experiences more imbibition in the initial period. After several imbibition and drainage cycles, the maximum saturation of CO_2 is reduced in the intermittent injection process. Therefore, less residual trapping is observed after ~9 years in intermittent injection compared to the continuous injection process (Fig.5).

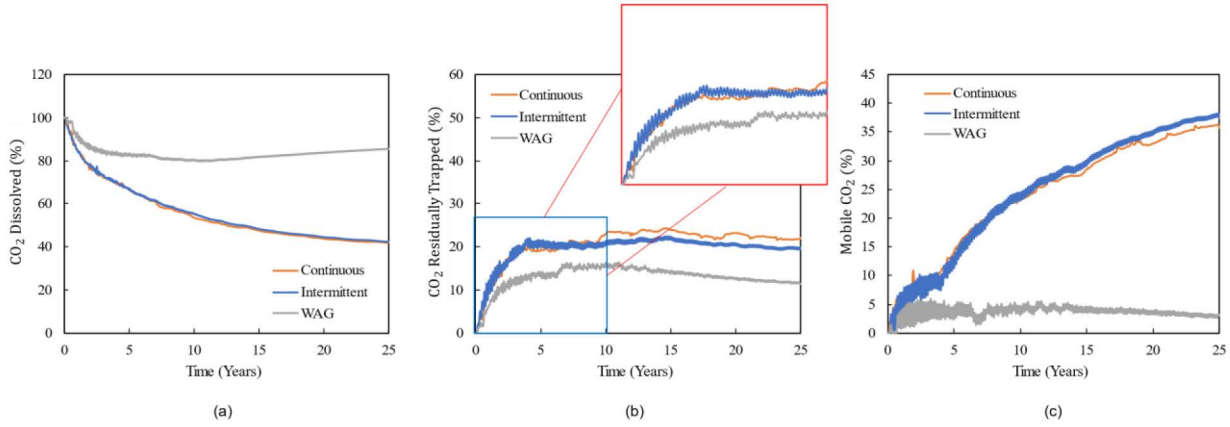

Figure 5—(a) Dissolved, (b) Residually trapped, and (c) Mobile CO_2 in Continuous, Intermittent, and WAG injection process after 25 year of period. The red box in (b) is the zoomed section of the blue box.

Effects of Wag Ratio

The WAG ratio significantly affects the trapping efficiency of CO_2 in the reservoir. This study uses WAG ratios of 1:1, 1:2, 1:3, and 2:1 to examine the effect on CO_2 trapping efficiency (Fig. 6a-b). For all the cases, an equal amount of CO_2 is injected over a period of 25 years. It is observed that increasing the CO_2 injection cycle time increases the residual trapping and mobile phase of CO_2. On the other hand, increasing the water injection cycle time leads to more dissolution trapping, as more water would be available to dissolve the injected gas. The findings are important as they suggest that CO_2 injection for a longer period of time

in the WAG process could increase the mobile phase CO_2, which may increase the risk for CO_2 leakage through caprock, whereas increasing the water injection cycle time will lead to a lower mobile phase CO_2 and increase the dissolution trapping.

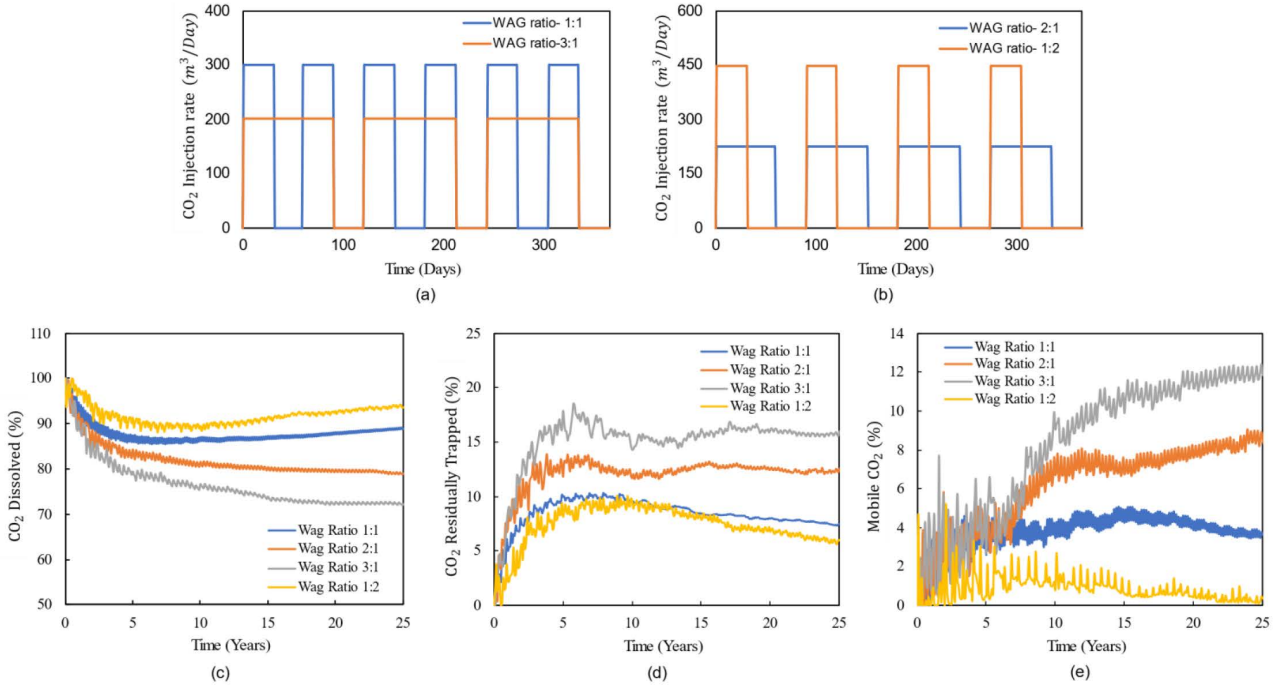

Figure 6—(a) and (b) Presents the WAG ratio for one year used in this study, and (c), (d), and (e) are the dissolved, residually trapped, and mobile phase CO_2 in different WAG ratios after 25 years of period.

Results of Proxy Models

Data Generation for Proxy Model. The WAG process with a ratio of 1:1 is selected as the base case model for developing the ML model. We generated numerous iterations of the base case by changing different uncertain properties of the reservoir. To represent different geological storage, nine parameters were varied. The range of values for the selected parameters is based on previous literature (Krevor et al., 2012; Medina et al., 2011; Michael et al., 2010; Vo Thanh & Lee, 2022; Zhao et al., 2010). Table 2 Represents the statistical overview of the input parameters considered in this study. In total, 90 geological realizations are considered during the WAG process to investigate the uncertainty of reservoir heterogeneities. For each case after five years of interval, the amount of CO_2 trapped by dissolution, residual, and structural is considered as one data point. In total, 5,551 data points are generated from the simulations for training, testing, and validating the ML model. In order to avoid bias in selecting the training and testing dataset, this study employs a five-fold cross-validation method.

Table 2—Statistical overview of the input parameters.

Parameter	Range	Reference
Porosity	0.10 - 0.30	(Medina et al., 2011; Michael et al., 2010)
Permeability (mD)	20 - 1000	(Medina et al., 2011; Michael et al., 2010)
Permeability ratio, k_v/k_h)	0.01 - 1	(Zhao et al., 2010)
Maximum residual gas saturation	0.1-0.5	(Vo Thanh & Lee, 2022)
Krg at Connate liquid	0.5-0.9	(Krevor et al., 2012)
Exponent for calculating Krw	5-17	(Krevor et al., 2012)

Parameter	Range	Reference
Exponent for calculating Krg	1-4	(Krevor et al., 2012)
Endpoint Saturation: Critical gas	0.02-0.2	(Krevor et al., 2012)
Endpoint Saturation: Critical water	0.2-0.4	(Krevor et al., 2012)

Results of Machine Learning Models. In this study, three different ML models (RF, MLP, and XGBoost) are applied to predict the fate of the CO_2 inside the reservoir. The capacity to capture nonlinear relationships and the ability to handle high-dimensional and heterogeneous data made these models well-recognized in CCS projects (Khanal & Shahriar, 2022; You et al., 2020). However, the hyperparameters of the ML models need to be tuned to improve the prediction accuracy. This study uses the RandomSearchCV optimization technique to identify the optimal set of hyperparameters. As it can effectively explore the hyperparameter space by randomly sampling values from redefined distributions, especially when dealing with large search space of hyperparameters. Moreover, it is faster than other hyperparameter tuning techniques. Table 3 summarizes the hyperparameters and their optimal values for RF, MLP, and XGBoost models.

Table 3—Hyperparameters and their optimal values for the ML models used in this study.

Models	Prediction Target	Parameters
XGBoost	CO_2 Dissolved	{'colsample_bytree': 0.84, 'learning_rate': 0.189, 'max_depth': 5, 'n_estimators': 177, 'subsample': 0.96}
	CO_2 Trapped	{'colsample_bytree': 0.83, 'learning_rate': 0.108, 'max_depth': 8, 'n_estimators': 117, 'subsample': 0.82}
	CO_2 Mobile	{'colsample_bytree': 0.91, 'learning_rate': 0.112, 'max_depth': 8, 'n_estimators': 168, 'subsample': 0.85}
MLP	CO_2 Dissolved	{'solver': 'lbfgs,' 'hidden_layer_sizes': (64, 32), 'alpha': 0.1, 'activation': 'relu'}
	CO_2 Trapped	
	CO_2 Mobile	
RF	CO_2 Dissolved	{'bootstrap': True, 'max_depth': 21, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 1281}
	CO_2 Trapped	
	CO_2 Mobile	

Figure 7 shows the performance of ML models for predicting the dissolved, trapped, and mobile phase CO_2. A higher correlation factor (R^2) and lower evaluation matrices generally correspond to higher predictive model accuracy. For a better understanding of the prediction accuracy, the statistical accuracy of each model is detailed in Table 4. For all the models, the R^2 values are satisfactory. However, the XGBoost model shows the best accuracy with R^2 values of 0.999, 0.995, and 0.985 for predicting the dissolved, trapped, and mobile phase CO_2, respectively. For all the cases, the R^2 value for mobile phase CO_2 is lower compared to the dissolved and trapped CO_2. As time goes on, the mobile phase CO_2 approaches zero due to dissolution and residual trapping. Since all the data for dissolved, trapped, and mobile phase CO_2 are recorded after every five-year interval, time is strongly correlated with the amount of CO_2 trapped by these mechanisms. But as the mobile phase CO_2 diminishes to zero sooner than the dissolved and trapped phase CO_2, less data is available to train the machine learning models, which leads to a lower R^2 value.

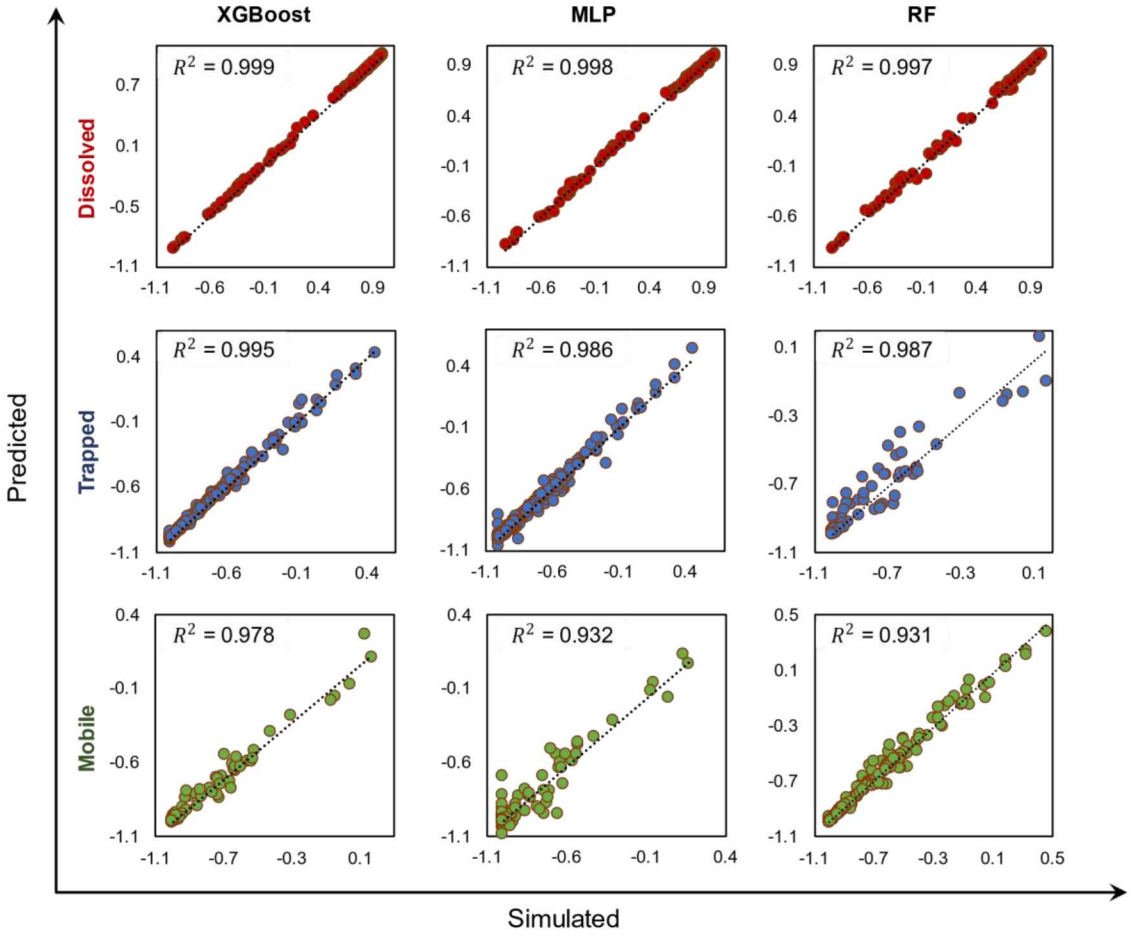

Figure 7—Performance of predictive models obtained from XGBoost, RF, and MLP models.

Table 4—Statistical accuracy of the proposed machine learning models in predicting dissolved, trapped, and mobile phase CO_2.

Models	Prediction Target	Evaluation Metrics			
		R^2	MAE	MSE	RMSE
XGBoost	CO_2 Dissolved	0.99962	0.00306	0.00004	0.00620
	CO_2 Trapped	0.99514	0.00667	0.00025	0.01569
	CO_2 Mobile	0.97876	0.00456	0.00033	0.01823
MLP	CO_2 Dissolved	0.99862	0.00633	0.00014	0.01174
	CO_2 Trapped	0.98672	0.01409	0.00067	0.02594
	CO_2 Mobile	0.93292	0.00977	0.00105	0.03240
RF	CO_2 Dissolved	0.99799	0.00440	0.00019	0.01414
	CO_2 Trapped	0.98773	0.00937	0.00062	0.02494
	CO_2 Mobile	0.93123	0.00839	0.00108	0.03281

Feature Importance Analysis. We investigated the effect of input parameters on the fate of CO_2 inside the reservoir by calculating the feature importance plot for the machine learning models evaluated in this study (XGBoost and RF). In the feature analysis of RF and XGBoost, there must be a strong relationship of time with dissolved, mobile, and trapped CO_2, as with time, the fate of the CO_2 changes within the reservoir. However, we excluded time in the feature analysis to underscore the importance of other parameters in relation to the different phases of CO_2. Fig.8 shows that for both models, porosity is the most important

parameter for predicting the dissolved CO_2, which means dissolution trapping of CO_2 is greatly related to porosity. If the reservoir porosity is low, the pressure within the reservoir will increase with CO_2 injection. As the pressure-dependent Henry's constant controls the gas dissolution, the dissolution of CO_2 will also increase (Khanal et al., 2024). For the trapped CO_2, the most important feature is the vertical permeability, with a score of 0.60 in RF and 0.33 in the XGBoost model. This indicates that compared to the horizontal permeability, the vertical permeability of the reservoir significantly influences the trapping of CO_2. In the case of predicting the mobile phase CO_2, vertical permeability is again the most crucial feature, with an importance factor of 0.62, and the second most important feature is horizontal permeability in the RF model. This is because, with high permeability, more brine can flow towards the perforation zone and will come in contact with the fresh brine; hence, more CO_2 will dissolve. We can also relate this with the feature importance plot, which shows that the second most important feature for predicting dissolved CO_2 is the vertical permeability of the reservoir.

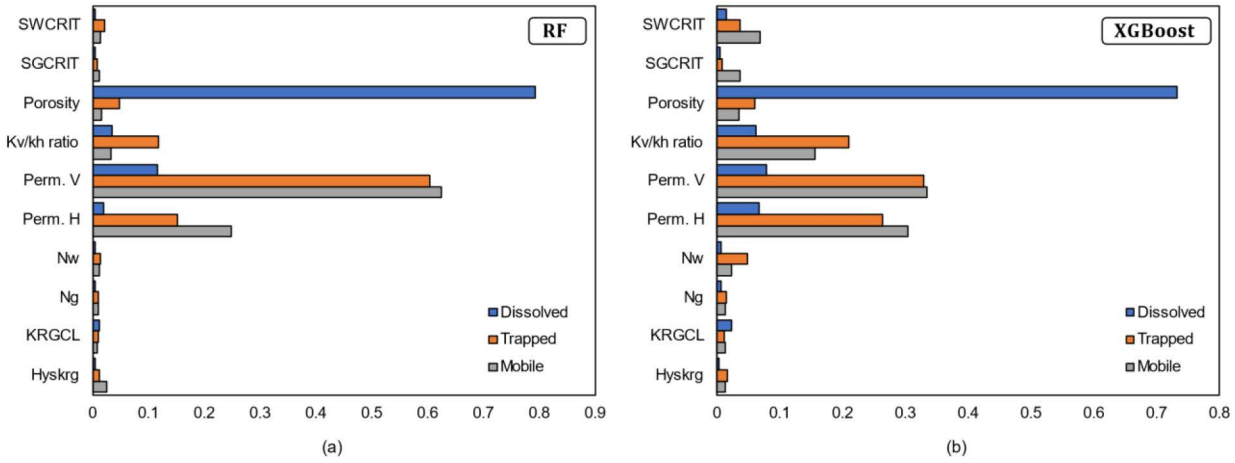

Figure 8—Feature importance plot for dissolved, residual, and mobile phase CO_2 for (a) RF and (b) XGBoost model. SWCRIT: Critical water saturation, SGCRIT: Critical gas saturation, Perm. V: Vertical permeability, Perm. H: Horizontal permeability, KRGCL: Krg at Connate Liquid, Hyskrg: Hysteresis residual gas saturation, Nw: Exponent for calculating Krw and Ng: Exponent for calculating Krg.

Conclusions

This study investigated different CO_2 injection schemes in underground CO_2 storage projects. A proxy model for a deep saline aquifer is developed using the machine learning model for predicting the CO_2 trapping scenarios. Moreover, a feature importance analysis has been conducted to understand the effect of different parameters on dissolved, trapped, and mobile phase CO_2. The following conclusion can be drawn from the findings of this work:

1. The WAG process is more suitable for safe CO_2 trapping in deep saline aquifers (Fig.5). In the WAG process, more CO_2 comes in contact with the brine solution; hence, the dissolution rate is higher compared to the intermittent and continuous injection process. However, further investigation on the effect of injection rate, injection position, and economic analysis is required before concluding which injection process is optimal for CCS projects.
2. Increasing the water injection cycle time in the WAG process is more suitable for CO_2 storage through dissolution trapping (Fig.6).
3. A proxy-based reservoir model can be developed using the ML models to accurately predict the CO_2 trapping mechanisms, which offers saving in both time and resources.
4. Three different ML models (XGBoost, RF, and MLP) are used to predict the dissolved, trapped, and mobile CO_2 scenarios. The R^2, MAE, MSE, and RMSE values are satisfactory for all the models.

However, the XGBoost model shows better prediction accuracy than other models, with an R^2 value of 0.999, 0.995, and 0.985 for predicting the dissolved, trapped, and mobile phase CO_2, respectively (Table 4).

5. Porosity is the most important feature for predicting dissolved CO_2, while vertical permeability is the most crucial parameter for predicting trapped and mobile phase CO_2 (Fig.8).

Acknowledgments

This material is based upon work partly supported by the National Science Foundation Award under CBET-2245484. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Al-Khdheeawi, E. A., Vialle, S., Barifcani, A., Sarmadivaleh, M., & Iglauer, S. (2018a). Impact of injected water salinity on CO2 storage efficiency in homogenous reservoirs. *The APPEA Journal*, **58**(1), 44. https://doi.org/10.1071/AJ17041

Al-Khdheeawi, E. A., Vialle, S., Barifcani, A., Sarmadivaleh, M., & Iglauer, S. (2018b, March 20). Impact of Injection Scenario on CO2 Leakage and CO2 Trapping Capacity in Homogeneous Reservoirs. *Day 4 Fri, March 23, 2018*. https://doi.org/10.4043/28262-MS

Ali, E., Hadj-Kali, M. K., Mulyono, S., Alnashef, I., Fakeeha, A., Mjalli, F., & Hayyan, A. (2014). Solubility of CO2 in deep eutectic solvents: Experiments and modelling using the Peng–Robinson equation of state. *Chemical Engineering Research and Design*, **92**(10), 1898–1906. https://doi.org/10.1016/j.cherd.2014.02.004

Bachu, S. (2002). Sequestration of CO2 in geological media in response to climate change: road map for site selection using the transform of the geological space into the CO2 phase space. *Energy Conversion and Management*, **43**(1), 87–102. https://doi.org/10.1016/S0196-8904(01)00009-7

Bachu, S. (2008). CO2 storage in geological media: Role, means, status and barriers to deployment. *Progress in Energy and Combustion Science*, **34**(2), 254–273. https://doi.org/10.1016/j.pecs.2007.10.001

Bachu, S., Bonijoly, D., Bradshaw, J., Burruss, R., Holloway, S., Christensen, N. P., & Mathiassen, O. M. (2007). CO2 storage capacity estimation: Methodology and gaps. *International Journal of Greenhouse Gas Control*, **1**(4), 430–443. https://doi.org/10.1016/S1750-5836(07)00086-2

Calabrese, M., Masserano, F., & Blunt, M. J. (2005, October 9). Simulation of Physical-Chemical Processes During Carbon Dioxide Sequestration in Geological Structures. *All Days*. https://doi.org/10.2118/95820-MS

Celia, M. A., Bachu, S., Nordbotten, J. M., & Bandilla, K. W. (2015). Status of CO 2 storage in deep saline aquifers with emphasis on modeling approaches and practical simulations. *Water Resources Research*, **51**(9), 6846–6892. https://doi.org/10.1002/2015WR017609

Doughty, C. (2007). Modeling geologic storage of carbon dioxide: Comparison of non-hysteretic and hysteretic characteristic curves. *Energy Conversion and Management*, **48**(6), 1768–1781. https://doi.org/10.1016/j.enconman.2007.01.022

Doughty, C. (2010). Investigation of CO2 Plume Behavior for a Large-Scale Pilot Test of Geologic Carbon Storage in a Saline Formation. *Transport in Porous Media*, **82**(1), 49–76. https://doi.org/10.1007/s11242-009-9396-z

IEA. (2022). Capacity of large-scale CO2 capture projects, current and planned vs. the Net Zero Scenario, 2020-2030. IEA, Paris. https://www.iea.org/data-and-statistics/charts/capacity-of-large-scale-co2-capture-projects-current-and-planned-vs-the-net-zero-scenario-2020-2030

Khanal, A., Irfan Khan, M., & Fahim Shahriar, M. (2024). Comprehensive parametric study of CO2 sequestration in deep saline aquifers. *Chemical Engineering Science*, **287**, 119734. https://doi.org/10.1016/j.ces.2024.119734

Khanal, A., & Shahriar, M. F. (2022). Physics-Based Proxy Modeling of CO2 Sequestration in Deep Saline Aquifers. *Energies*, **15**(12), 4350. https://doi.org/10.3390/en15124350

Khanal, A., & Weijermars, R. (2019). Pressure depletion and drained rock volume near hydraulically fractured parent and child wells. *Journal of Petroleum Science and Engineering*, **172**, 607–626. https://doi.org/10.1016/j.petrol.2018.09.070

Kim, Y., Jang, H., Kim, J., & Lee, J. (2017). Prediction of storage efficiency on CO2 sequestration in deep saline aquifers using artificial neural network. *Applied Energy*, **185**, 916–928. https://doi.org/10.1016/j.apenergy.2016.10.012

Klara, S. M., Srivastava, R. D., & McIlvried, H. G. (2003). Integrated collaborative technology development program for CO2 sequestration in geologic formations—United States Department of Energy R&D. *Energy Conversion and Management*, **44**(17), 2699–2712. https://doi.org/10.1016/S0196-8904(03)00042-6

Krevor, S. C. M., Pini, R., Zuo, L., & Benson, S. M. (2012). Relative permeability and trapping of CO2 and water in sandstone rocks at reservoir conditions. *Water Resources Research*, **48**(2). https://doi.org/10.1029/2011WR010859

Kumar, A., Noh, M., Pope, G. A., Sepehrnoori, K., Bryant, S., & Lake, L. W. (2004, April 17). Reservoir Simulation of CO2 Storage in Deep Saline Aquifers. *All Days*. https://doi.org/10.2118/89343-MS

Liu, B., Zhao, F., Xu, J., & Qi, Y. (2019). Experimental Investigation and Numerical Simulation of CO2–Brine–Rock Interactions during CO2 Sequestration in a Deep Saline Aquifer. *Sustainability*, **11**(2), 317. https://doi.org/10.3390/su11020317

Lucier, A., & Zoback, M. (2008). Assessing the economic feasibility of regional deep saline aquifer CO2 injection and storage: A geomechanics-based workflow applied to the Rose Run sandstone in Eastern Ohio, USA. *International Journal of Greenhouse Gas Control*, **2**(2), 230–247. https://doi.org/10.1016/j.ijggc.2007.12.002

Medina, C. R., Rupp, J. A., & Barnes, D. A. (2011). Effects of reduction in porosity and permeability with depth on storage capacity and injectivity in deep saline aquifers: A case study from the Mount Simon Sandstone aquifer. *International Journal of Greenhouse Gas Control*, **5**(1), 146–156. https://doi.org/10.1016/j.ijggc.2010.03.001

Menad, N. A., Hemmati-Sarapardeh, A., Varamesh, A., & Shamshirband, S. (2019). Predicting solubility of CO2 in brine by advanced machine learning systems: Application to carbon capture and sequestration. *Journal of CO2 Utilization*, **33**, 83–95. https://doi.org/10.1016/j.jcou.2019.05.009

Michael, K., Golab, A., Shulakova, V., Ennis-King, J., Allinson, G., Sharma, S., & Aiken, T. (2010). Geological storage of CO2 in saline aquifers—A review of the experience from existing storage operations. *International Journal of Greenhouse Gas Control*, **4**(4), 659–667. https://doi.org/10.1016/j.ijggc.2009.12.011

Oloruntobi, O. S., & LaForce, T. (2009, June 8). Effect of Aquifer Heterogeneity on CO2 Sequestration. *All Days*. https://doi.org/10.2118/121776-MS

Pan, F., McPherson, B. J., Esser, R., Xiao, T., Appold, M. S., Jia, W., & Moodie, N. (2016). Forecasting evolution of formation water chemistry and long-term mineral alteration for GCS in a typical clastic reservoir of the Southwestern United States. *International Journal of Greenhouse Gas Control*, **54**, 524–537. https://doi.org/10.1016/j.ijggc.2016.07.035

Rasheed, Z., Raza, A., Gholami, R., Rabiei, M., Ismail, A., & Rasouli, V. (2020). A numerical study to assess the effect of heterogeneity on CO2 storage potential of saline aquifers. *Energy Geoscience*, **1**(1–2), 20–27. https://doi.org/10.1016/j.engeos.2020.03.002

Safaei-Farouji, M., Vo Thanh, H., Dai, Z., Mehbodniya, A., Rahimi, M., Ashraf, U., & Radwan, A. E. (2022). Exploring the power of machine learning to predict carbon dioxide trapping efficiency in saline aquifers for carbon geological storage project. *Journal of Cleaner Production*, **372**, 133778. https://doi.org/10.1016/j.jclepro.2022.133778

Singh, H., Hosseini, S. A., & Javadpour, F. (2012, June 12). Enhanced CO2 Storage In Deep Saline Aquifers By Nanoparticles: Numerical Simulation Results. *All Days*. https://doi.org/10.2118/156983-MS

Song, Y., Sung, W., Jang, Y., & Jung, W. (2020). Application of an artificial neural network in predicting the effectiveness of trapping mechanisms on CO2 sequestration in saline aquifers. *International Journal of Greenhouse Gas Control*, **98**, 103042. https://doi.org/10.1016/j.ijggc.2020.103042

Vo Thanh, H., & Lee, K.-K. (2022). Application of machine learning to predict CO2 trapping performance in deep saline aquifers. *Energy*, **239**, 122457. https://doi.org/10.1016/j.energy.2021.122457

Wei, N., Li, X., Jiao, Z., Stauffer, P. H., Liu, S., Ellett, K., & Middleton, R. S. (2022). A Hierarchical Framework for CO2 Storage Capacity in Deep Saline Aquifer Formations. *Frontiers in Earth Science*, 9. https://doi.org/10.3389/feart.2021.777323

You, J., Ampomah, W., Sun, Q., Kutsienyo, E. J., Balch, R. S., Dai, Z., Cather, M., & Zhang, X. (2020). Machine learning based co-optimization of carbon dioxide sequestration and oil recovery in CO2-EOR project. *Journal of Cleaner Production*, **260**, 120866. https://doi.org/10.1016/j.jclepro.2020.120866

Yu, Y., Yang, G., Cheng, F., & Yang, S. (2021). Effects of impurities N2 and O2 on CO2 storage efficiency and costs in deep saline aquifers. *Journal of Hydrology*, **597**, 126187. https://doi.org/10.1016/j.jhydrol.2021.126187

Zhao, H., Liao, X., Chen, Y., & Zhao, X. (2010). Sensitivity analysis of CO2 sequestration in saline aquifers. *Petroleum Science*, **7**(3), 372–378. https://doi.org/10.1007/s12182-010-0080-2