ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Estimation of CO_2-Brine Interfacial Tension Based on an Advanced Intelligent Algorithm Model: Application for Carbon Saline Aquifer Sequestration

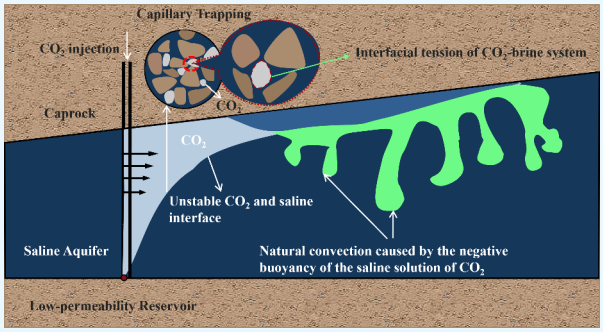
Meiheriayi Mutailipu,* Yande Yang, Kaishuai Zuo, Qingnan Xue, Qi Wang, Fusheng Xue, and Gang Wang

Cite This: *ACS Omega* 2024, 9, 37265−37277

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The emission reduction of the main greenhouse gas, CO_2, can be achieved via carbon capture, utilization, and storage (CCUS) technology. Geological carbon storage (GCS) projects, especially CO_2 storage in deep saline aquifers, are the most promising methods for meeting the net zero emission goal. The safety and efficiency of CO_2 saline aquifer storage are primarily controlled by structural and capillary trapping, which are significantly influenced by the interactions between fluid and solid phases in terms of the interfacial tension (IFT) between the injected CO_2 and brine at the reservoir site. In this study, a model based on the random forest (RF) model and the Bayesian optimization (BO) algorithm was developed to estimate the IFT between the pure and impure gas−brine binary systems for application to CO_2 saline aquifer sequestration. Then three heuristic algorithms were applied to validate the accuracy and efficiency of the established model. The results of this study indicate that among the four mixed models, the Bayesian optimized random forest model fits the experimental data with the smallest root-mean-square error (RMSE = 1.7705) and mean absolute percentage error (MAPE = 2.0687%) and a high coefficient of determination (R^2 = 0.9729). Then the IFT values predicted via this model were used as an input parameter to estimate the CO_2 sequestration capacity of saline aquifers at different depths in the Tarim Basin of Xinjiang, China. The burial depth had a limited influence on the CO_2 storage capacity.



1. INTRODUCTION

In recent years, reducing anthropogenic carbon emissions has aroused broad concern in countries around the world. Significant amounts of greenhouse gases have been emitted over the years, according to previous research data.[1,2] Significantly, China, whose energy system mainly relies on fossil fuels, emitted 10.76 GT of CO_2 and is currently the world's largest emitter of CO_2.[3,4] Thus, to address the global warming caused by anthropogenic carbon emissions, China proposed the carbon peak and carbon neutrality goal of the "3060" plan in 2020.[5−7]

There are several emission reduction methods, including energy structure transformation and energy use efficiency enhancement. Among these methods, the carbon capture, utilization, and storage (CCUS) technique is not only a key technology for reducing carbon emissions but also has significant commercial value compared with other technologies.[8−10] Geological CO_2 storage combined with enhanced water recovery (CO_2-EWR) technology, considered a newly developed carbon utilization technique, has the ability to safely store a significant amount of CO_2 without increasing the reservoir pressure, which can be utilized to facilitate reverse osmosis desalination of reservoir brine to produce drinkable water.[11,12] This solves the problem of an industry that suffers from water leakage, sometimes discharging wastewater to the neighborhood. Thus, this technology offers a dual solution to environmental and water scarcity challenges.[8,13] The geological storage and utilization of CO_2 has great potential both worldwide[11,14,15] and in China, and the Xinjiang region is suitable for CO_2 saline aquifer sequestration.[11] During this process, the thermodynamic properties of the gas−liquid phase, especially the interfacial tension (IFT) between the gas and liquid phases, play a crucial role in the flexibility of this scheme due to its direct influence on the efficiency of capillary trapping, which is one of the most critical trapping mechanisms in the CO_2 geological storage (CGS) process.[16−19]

Received: May 24, 2024
Revised: August 3, 2024
Accepted: August 7, 2024
Published: August 19, 2024

In the process of CO_2 geological storage, CO_2 is generally in a supercritical state. Supercritical CO_2 is lifted by buoyancy and settles beneath the cap. The increase in capillary pressure is caused by the numerous linked pore throats of varying sizes in the cap layer, as depicted in Figure 1. The breakthrough

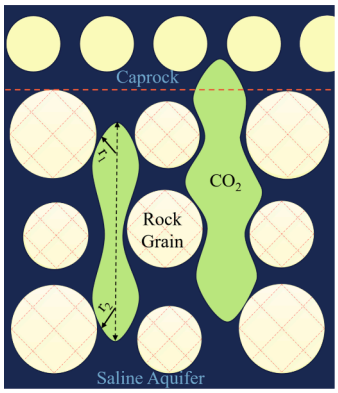


Figure 1. Comparative subsurface CO_2 flow.

pressure occurs when the pressure differential across the entire cap layer—that is, the pressure differential between the injected CO_2 and saline water—exceeds the capillary pressure of the pore throat. The nonwetting phase—in this case, CO_2—flows through the pore throat after the breakthrough pressure is reached, creating issues with CO_2 leakage. The consequences of capillary pressure breakthrough leakage are severe and occur far more quickly than leakage due to CO_2 diffusion. Consequently, one of the most critical metrics for evaluating the effectiveness of cap layer sealing is exceeding this threshold capillary pressure. The threshold capillary pressure can be represented by Laplace's law as follows:

$$Pce \approx \frac{2\sigma \cos \varphi}{R} \quad (1)$$

where P_{ce} is the threshold capillary pressure at the orifice throat; σ is the IFT between CO_2 and brine; φ is the contact angle between the caprock, CO_2, and brine system; and R is the size of the most prominent pore throat or microfracture in the caprock because the larger R is, the lower the breakthrough pressure. The orifice throats frequently connect to each other, leading to leakage when the maximum orifice throat is exceeded. It is essential to conduct necessary studies on the influence of temperature, pressure, and other factors on the IFT between the CO_2 and brine. These findings provide crucial guidance for improving the safety of CO_2 injection and increasing the capacity for storage.[20] In addition, these studies provide insights into the factors that affect the IFT, which will facilitate future research and advancements in the field.

Numerous academic researchers have conducted extensive investigations on the IFT in CO_2−brine systems utilizing a wide range of approaches, including experimental measurements,[21−27] molecular dynamics simulations,[28,29] empirical formulations,[30,31] density gradient models,[32] machine learning,[33−35] and various other techniques. Li et al.[36] conducted IFT measurements on a mixed salt system (0.864 NaCl and 0.136 KCl) using the pendant drop method and discovered a linear relationship between the IFT and the salt concentration. They also developed empirical equations with high accuracy and reached a good agreement with the experimental data. In our previous work, we[37] analyzed the relationships between the IFT and temperature, pressure, and salinity in detail via the pendant drop method. The results indicate that in the low-pressure phase, the IFT between CO_2 and brine decreases markedly with the increase in pressure. Upon reaching the high-pressure regime, the IFT enters a pseudoplateau where it becomes pressure-independent. In contrast, they fluctuate with temperature and increase with molality. Jerauld et al.[38] improved the empirical formula for the IFT in a CO_2−brine system. This empirical formula improved the correlation between the IFT and temperature, reducing the standard deviation to 3.4%. Iglauer et al.[29] utilized molecular dynamics to simulate IFT phenomena and discovered that their model aligned with the experimental data. However, it was not a reliable predictor at high pressures, exhibiting an error rate of up to 20%. As a result, the inclusion of a correction factor became necessary.

With the rapid development of computer science, machine learning has also experienced significant growth and is now being applied in various fields. Many scholars have recognized its effectiveness in multiple areas.[39] Ratnakar et al.[40] used machine learning to develop models for predicting the solubility of carbon dioxide in brine. They utilized two primary models, namely, a random forest and a decision tree. The final predictions of the models exhibited a relative error of 2−7% compared to the experimental observations. Liu et al.[41] constructed an optimized wavelet neural network to predict the IFT of a CO_2−brine system with a Model R^2 of 0.95 but used a small data set. Amooie et al.[42] conducted a machine learning prediction of IFT for a CO_2−brine system. However, the model they used was outdated. Amar[33] studied an IFT system, using genetic programming to improve the system.

Currently, machine learning has been widely used in various fields due to its high computing speed, strong adaptability, and fault tolerance.[43] A random forest model combined with the Bayesian optimization algorithm (BO-RF) was developed to estimate the IFTs between CO_2 and brine/water binary systems corresponding to the CGS conditions in this work. In regard to predictors, random forests outperform other conventional models and exhibit better stability. However, a disadvantage of random forests is that they erode over time, in contrast to the recently proposed hybrid algorithms. Here, a Bayesian optimization algorithm was applied to optimize the random forest hyperparameters to increase the accuracy of the prediction results. Then, the Tarim Basin in China's Xinjiang Province was used as a case study to investigate the influence of the IFT on the estimation of the storage capacity of the selected site. This study provides a high-precision model for interfacial tension acquisition in CO_2−brine systems, which may be helpful for CO_2 geological storage technology.

2. METHODOLOGY

2.1. Principles of Random Forests. Breiman[44] proposed random forest (RF), an integrated learning method based on several decision trees (weak learners). Machine learning applies the notion of integrated learning, which is often referred to as ensemble learning. It is an optimized combination of several different models.[45] Weak learners are the models that are employed in this combination, which enables more accurate predictions. The training sample set during training must be used to train these weak learners in a sequential manner. Then, these weak learner models are combined and used for predictions.

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

2.1.1. Regression Tree. In the prediction process of the regression tree model, objective evaluations and decisions are made along the branches of the tree based on the input eigenvalues until the leaf nodes are reached and the prediction is made. The associations are represented using a structure that resembles a tree, in which each inner node represents a feature, each branch represents a value obtained from that characteristic, and each terminal node represents the output of a prediction. The following steps are primarily involved in the creation of regression trees:

1. Selection of features: To obtain the best classification result for the split data set, the selection of the appropriate characteristics for the root node is carried out.

2. A finite number of subsets are created from the training data set, and the division of those subsets can be explained as follows: The jth eigenvector value is chosen for the input and output variables. For the purpose of categorizing characteristics and labeling points, the following two subsets are defined:

$$R_1(j, s) = \{x|x^{(j)} \leq s\} \tag{2}$$

$$R_2(j, s) = \{x|x^{(j)} > s\} \tag{3}$$

where $R_1(j, s)$ represents the left subset of the partition of the values s of the jth feature vector and $R_2(j, s)$ represents the right subset of the partition of the values s of the jth feature vector.

The best cutoff variable and cutoff point for resolution are selected as follows:

$$\min_{j,s}\left[\min \sum_{x_i \in R_1(j,s)} (y_i - y_{\cdot 1})^2 + \min \sum_{x_i \in R_2(j,s)} (y_i - y_{\cdot 2})^2\right] \tag{4}$$

$$y_{\cdot m} = \frac{1}{N_m} \sum_{x_i \in R_m(j,s)} y_i, \quad x \in R_m, \quad m = 1, 2 \tag{5}$$

Iterate over the variable j, search the intersection s for a fixed intersection variable j, choose the (j, s) that minimizes the above equation, and then divide the set into two subsets in turn. This process is repeated until the end conditions are met.

Regression trees operate on a simple, straightforward basis, which makes them easy to apply. However, they are sensitive to changes in the input data and prone to overfitting. Therefore, a regression tree can be optimized by applying integrated learning or pruning techniques to address the unpredictability problems of the data.

2.1.2. Random Forest. Regression trees are the fundamental basis of the random forest method, which integrates several regression trees to improve the accuracy and integrity of predictions, as illustrated in Figure 2. The construction of a random forest is as follows:

1. A data point n from the complete training data is chosen, where n is much smaller than N (the total training data set). This is achieved by selecting the input training data. The out-of-bag data are a subset of the complete training data set that cannot be chosen for analysis. Error estimates can be made using this particular subset.

2. At each segment node, m features are chosen from the whole feature set M to form a regression tree.

3. When building each regression tree, the splitting nodes are selected according to the lowest Gini index. Until the maximum depth of the tree is reached or all training samples within a node belong to the same class, the remaining nodes of the regression tree are built using the same splitting strategy.

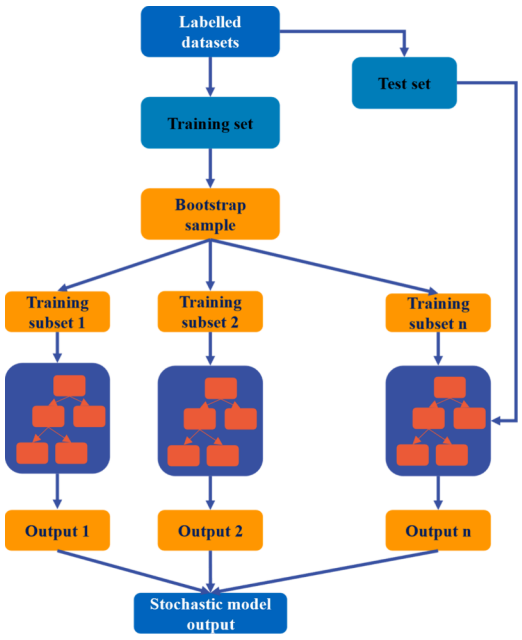


Figure 2. Random Forest flowchart.

4. Steps two and three are repeated multiple times. Each input corresponds to a regression tree. Then, a random forest model that can be used for predictive analysis will be established.

5. When the data to be predicted are entered, multiple regression trees can be used to make decisions at the same time, and the corresponding predicted values are obtained. The regression predicted value should be averaged from the predicted value of each regression tree to meet the final result.

The prediction process of the Random Forest is shown in eq 6 as follows:

$$h(x) = \frac{1}{n} \sum_{i=1}^{n} h_i(x) \tag{6}$$

Where n represents the number of decision trees, each T_i providing a prediction $h_i(x)$ for a given input x, and the final prediction of the Random Forest, denoted as $h(x)$, is the mean of all the predictions from the decision trees.

Following the above steps, a random forest can utilize the combination of results from multiple regression trees to make regression predictions. This approach can reduce the variance of the model and enhance its generalization ability.

2.2. Principles of Bayesian Optimization Algorithms.

To guarantee quick and accurate predictions while using a random forest for regression forecasting, it is critical to determine the model's optimal hyperparameters. In this study, the Bayesian optimization algorithm was used as a global optimization technique to determine the ideal hyperparameters for this model. Bayes theorem is the primary tool used in this method to identify the best results. To precisely fit the genuine goal function and determine the next evaluation position based on the fitted function, a probabilistic agent model was applied. This improves the search efficiency and makes it possible to find the ideal fit more quickly. The Bayesian optimization flowchart is shown in Figure 3. The two main parts of this

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

optimization approach are the collection function and the probabilistic agent model.

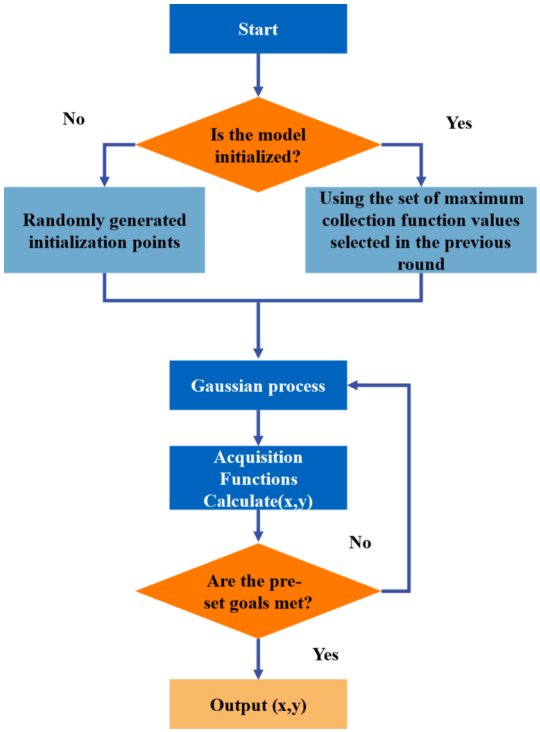


Figure 3. Flowchart of the Bayesian optimization algorithm.

Models of probabilistic agents can be roughly categorized as parametric or nonparametric. In this study, the collection function was subjected to the confidence boundary technique using a Gaussian process as the nonparametric model.

2.2.1. Probabilistic Agent Model. Regression, classification, and other fields require the inference of black-box functions based on Gaussian processes. Additionally, they have less tendency to "overfit". This model was used for this study. Generally, a neural network and a Gaussian process are related. In particular, a neural network with an infinite number of hidden layer units corresponds to a Gaussian process.[46] The basic model for the distribution of multivariate Gaussian probability is the Gaussian process. It consists of a mean function, m, and a covariance function, k, which refers to a semipositive definite.

In a Gaussian process, a finite set of random variables all follow a Gaussian joint distribution, assuming a prior distribution with a mean of zero:

$$p(f|X, \theta) = \mathrm{N}\left(0, \sum\right) \tag{7}$$

$$\sum_{i,j} = k(x_i, y_i) \tag{8}$$

where X is the training set, $\mathrm{X} = \{x_1, x_2,..., x_i\}f$ is the set of values of the unknown function f,$f = \{f(x_1), f(x_2),..., f(x_i)\}$; $\sum$ is the covariance matrix formed by $k(x, x')$; and θ is the hyperparameter.

The probability distribution is obtained by assuming the presence of noise that follows a Gaussian distribution, which is independent and identically distributed:

$$p(y|f) = \boldsymbol{N}(f, \sigma^2 I) \tag{9}$$

where y represents the set of observed values. The marginal likelihood distribution x is obtained based on the prior distribution and likelihood distribution equation:

$$p(y|X, \theta) = \int p(y|f)p(y|X, \theta)\,\mathrm{d}f = \boldsymbol{N}\left(0, \sum + \sigma^2 I\right) \tag{10}$$

Usually, by maximizing the marginal likelihood distribution to optimize hyperparameters, according to the properties of Gaussian processes, the following joint distribution exists:

$$\begin{bmatrix} y \\ f_* \end{bmatrix} : \boldsymbol{N}\left(0, \begin{bmatrix} \sum + \sigma^2 I & K_* \\ K_*^T & K_{**} \end{bmatrix}\right) \tag{11}$$

where f_* denotes the value of the prediction function and X_* represents the prediction input $K_*^T = \{k(x_1, X_*), k(x_2, X_*),..., k(x_i, X_*)\}$, $K_{**} = k(X_*, X_*)$

The predictive distribution derived from the joint distribution can be easily determined via the following equations:

$$p(f_*|X, y, X_*) = \boldsymbol{N}(\langle f_* \rangle, \mathrm{cov}(f_*)) \tag{12}$$

$$\langle f_* \rangle = K_*^T \left[\sum + \sigma^2 I\right]^{-1} y \tag{13}$$

$$\mathrm{cov}(f_*) = K_{**} - K_*^{\mathrm{T}} \left[\sum + \sigma^2 I\right]^{-1} K_* \tag{14}$$

where $\langle f_* \rangle$ denotes the predicted mean and $\boldsymbol{cov}(f_*)$ represents the expected covariance.

Table 1. Statistical Table of the Evaluation Indicators for Each Type of Model

Source	System	m/(mol·kg^{-1})	T/K	p/MPa	Number
Aggelopoulos, 2010[22]	pure CO_2-brine	0.045−2.7	300−373	4.9−25	103
Aggelopoulos, 2011[23]	pure CO_2-brine	0.045−1.5	300−373	5−25	95
Bachu et al, 2009[49]	pure CO_2-brine	0.117−4.087	293−398	2−27	280
Chalbaud, 2009[24]	pure CO_2-brine	0.085−2.75	300−373	4.8−25.8	107
Li, 2012a[36]	pure CO_2-brine	0.98−4.95	298−448	2−50	336
Li, 2012b[25]	pure CO_2-brine	0.98−5	343−423	2−50	232
Mutailipu, 2019[37]	pure CO_2-brine	1.05−4.9	298−373	2.98−15	160
Liu,2017[50]	pure CO_2-brine	0.98−1.98	298−423	2−69	42
Pereira,2017[51]	pure CO_2-brine	0.2−1.8	300−353	3−12	152
Ren, 2000[26]	impure CO_2-pure water	0.0	298−373	1−30	90
Yan, 2001[52]	impure CO_2-pure water	0.0	298−373	1−30	120

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

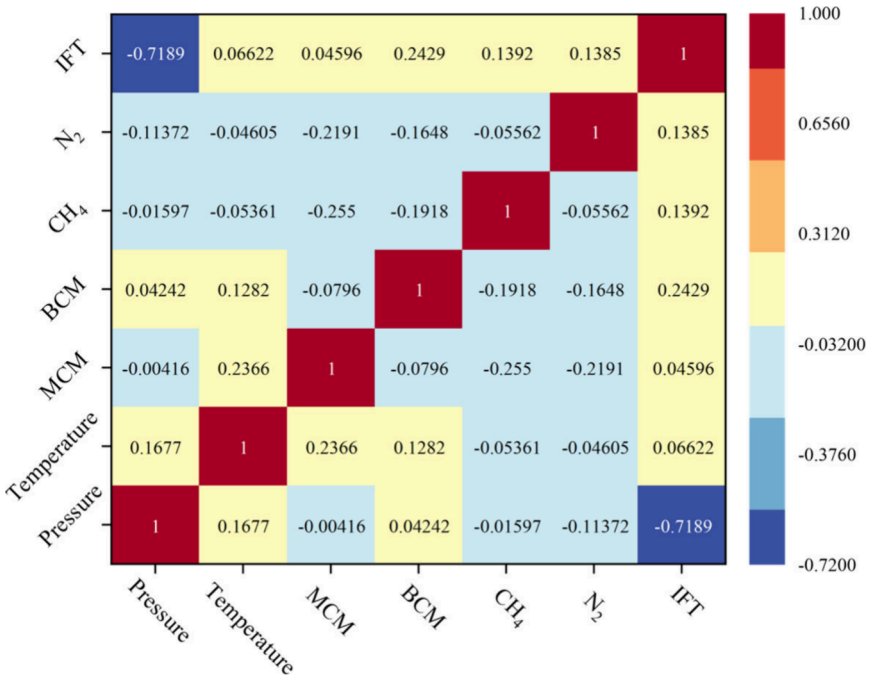


Figure 4. Correlation analysis between impact factors.

2.2.2. Collection Function. The collection function is primarily used in the process of Bayesian optimization to determine the next most likely point of evaluation to achieve optimal model performance. A confidence-bounding strategy known as GP-UCB, proposed by Srinivas et al.[47] for use with Gaussian processes, was applied in this study.

When maximizing the objective function, the acquisition function of the upper confidence bound (UCB) strategy was calculated via the following equations:

$$\alpha_t(x; D_{1:t}) = \mu_t(x) + \sqrt{\beta_t}\sigma_t(x) \tag{15}$$

where the β parameter balances the expectation and variance.

3. MODEL ESTABLISHMENT

3.1. Collecting and Preprocessing Data. The experimental data used in the training of the model for the CO_2 and H_2O/brine binary system are detailed in Table 1. The assembled database contained 1507 IFT data points of CO_2 and brine systems and 210 IFT data points of impure CO_2 mixed with CH_4 or N_2 and brine systems, with a total of 1717 points covering temperatures ranging from 293.15 to 448.15 K, with pressures up to approximately 69 MPa and salinities between 0 and 5.0 mol·kg^{-1}. According to previous research, the IFT in CO_2−brine systems is primarily influenced by pressure, temperature, the molar concentration of brine, and the molar fractions of methane and nitrogen.[48] Correlation analysis of the data can clarify the relationships between variables, and this study adopted Pearson correlation analysis. The correlation between each influencing factor and the IFT is illustrated in Figure 4. The subsequent section includes an examination of the impact of every variable on the IFT.

There are two primary phases to the effects of pressure. In the low-pressure phase (p < 10 MPa), IFT decreases significantly with the increase in pressure. This is considered the initial stage of IFT variation concerning pressure. The primary reason for this phase is the direct correlation between IFT and the density difference. As pressure increases, the density of CO_2 also increases, reducing the density difference between the two phases, thereby decreasing the IFT of CO_2-brine. However, once the critical pressure point is reached, known as the "pseudoplateau", IFT tends to stabilize, entering the second phase where the effect of pressure on IFT is no longer significant.[53,54] Within the range of low temperature and high pressure, the IFT between CO_2 and brine increases with the rise in temperature. However, at high temperatures (T > 343 K) and low pressures (p < 5 MPa), the IFT between carbon dioxide and brine decreases as the temperature rises. This phenomenon can be explained by the theory of Gibbs−Duhem surface energy, which posits that as the temperature increases, the kinetic energy of molecules at the interface also increases. This leads to more intense intermolecular interactions, thereby causing an increase in IFT.[37,55] The correlation between the salt molar concentration and the outcome is shown in Figure 4. As the salinity of the brine increases, the IFT also increases accordingly. This occurs because the high salinity enhances the intermolecular forces between water molecules, requiring more energy to expand the droplet's surface, thus increasing the IFT. The impact of brines containing divalent cations on IFT is mainly more pronounced.[56,57] For impurity gases, namely, CH_4 and N_2, the correlation is second only to that of pressure and bivalent cations. As the content of impurity gases increases, IFT also rises, which is highly beneficial for carbon dioxide storage in saline aquifers.[58] This must be taken into account by the geological sequestration of CO_2.

There are some outliers in the data collected from the experiment for measurement reasons. These outliers can impact the authenticity and objectivity of the data, which in turn can affect the fitting and generalization effect of the subsequent prediction model. The data in Table 1 were processed using the isolated forest detection approach in this study. This method mainly uses the characteristics of a high isolation degree of outliers for data screening. It calculates the isolation degree of data points by constructing a binary tree

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

and determines whether they are outliers according to the isolation degree. More remote data points are divided earlier in the division process. Hence, their average depth is smaller.[59]

The outlier score for sample x can be calculated via the following equation:

$$\text{Score}(x) = 2^{-E[h(x)]/c(N)} \tag{16}$$

where $E[h(x)]$ is the path length of sample x, which is calculated as follows:

$$E[h(x)] = \frac{1}{t}\sum_{i=1}^{t} h_j(x) \tag{17}$$

where $c(N)$ is the average path length of the tree given N data samples and is used to normalize x:

$$c(N) = \begin{cases} 2H(N-1) - 2(N-1)/N & N > 2 \\ 1 & N = 2 \\ 0 & N < 2 \end{cases} \tag{18}$$

where $H(X)$ is the harmonic number according to the following equation and Υ is Euler's constant, approximately 0.577215:

$$H(X) = \ln X + \Upsilon \tag{19}$$

According to the above steps, the larger the anomaly score is, the greater the anomaly degree of the sample points. In this study, 76 sets of outlier data were eliminated after detection using isolated forests.

3.2. BO-RF Modeling. The flowchart for the BO-RF model is shown in Figure 5. The data set in this study was divided into an 8:1:1 ratio between a training set, a validation set, and a test set. The main procedure is to use the training set for model training and a hyperparameter search. A validation set was used to achieve the predetermined goal of improving the model performance. Then, a pertinent prediction of the test set based on the optimized model was carried out. In this study, a heuristic algorithm was developed to optimize the random forest (RF) model for predicting the CO_2−water/brine IFT with high integrity. These heuristic algorithms include the Sparrow Search Algorithm (SSA),[60] the Particle Swarm Optimization algorithm (PSO),[61] and the Improved Gray Wolf Optimization algorithm (IGWO),[62] which are described in detail below.

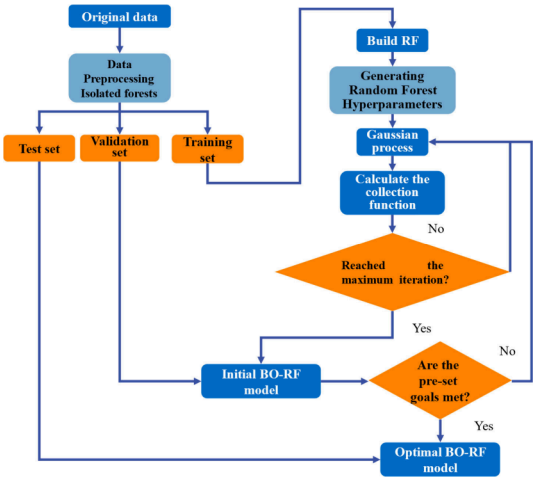


Figure 5. Flowchart of the BO-RF model.

3.3. Iterative Optimization Process. The optimization search procedure for each model is illustrated in Figure 6. The heuristic approach iterates more slowly and virtually levels off after the 15th step to find the ideal answer, as shown in Figure 6. On the other hand, the perfect solution is roughly reached by the Bayesian optimization process in the eighth phase. The superiority of the Bayesian optimization algorithm over the heuristic method in identifying an optimal solution indicates its inherent qualities. It is also demonstrated that the Bayesian optimization method outperforms the heuristic approach in terms of efficiency. Furthermore, the Bayesian optimization algorithm yields the optimal hyperparameter for the tree as 300, the minimum leaf node for the tree as 1, and the number of selected features as 4.

4. RESULTS AND DISCUSSION

4.1. Analysis of the Predictive Results of the Model. The prediction results of each model are shown in Figure 7. The closer the data point is to the 45° reference line, the greater the fitting degree between the predicted and experimental values. The figure shows that the BO-RF model has the best fitting degree compared with the other three models. Figure 8 more intuitively shows the prediction results of the four models under the same conditions, and it can be seen from the figure that the BO-RF model has a high prediction accuracy. Figure 8's experimental data is sourced from Reference.[37]

The distributions of the differences between the experimental and predicted values are displayed in Figure 9. The results indicate that the machine learning prediction has a better fit and accuracy, with the majority of the predicted data falling within ±1.5 $mN \cdot m^{-1}$ and the remaining data being very sparse. The primary comparison of the expected and experimental values in Figure 9(B) demonstrates the remarkable accuracy of the model suggested in this study.

4.2. Indicators for Model Evaluation. To validate the accuracy of the model proposed in this study, several statistical parameters were calculated. The equations for each evaluation indicator are shown in Table 2.

The statistical parameters of each model were calculated, and the results of each statistical parameter are shown in Table 3. Overall, each model tested had good predictive performance, with the BO-RF model outperforming all other predictive models, having the highest R^2 (0.9729) and the lowest RMSE (1.7705) and MAPE (2.0687%). The model predictions exhibit a strong fit with the observed values.

4.3. Comparison with Other Models. This study compared several high-performance models with the BO-RF model to demonstrate its accuracy. It was compared with prediction methods based on genetic programming (GP),[33] the group method of data handling (GMDH),[42] gene expression programming (GEP),[63] optimizing the WNN model(I-WNN),[41] and the correlation formula derived through GPTIPS.[54] The RMSE and R^2 values were chosen as the evaluation indices in this study, as shown in Table 4. It can be concluded that the BO-RF model performs better than the other methods, with an RMSE of 1.7705 and an R^2 of 0.9729.

4.4. Correlation Validation. In this section, the relationship between various influencing factors and IFT is depicted using our model. The four panels in Figure 10 illustrate the

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

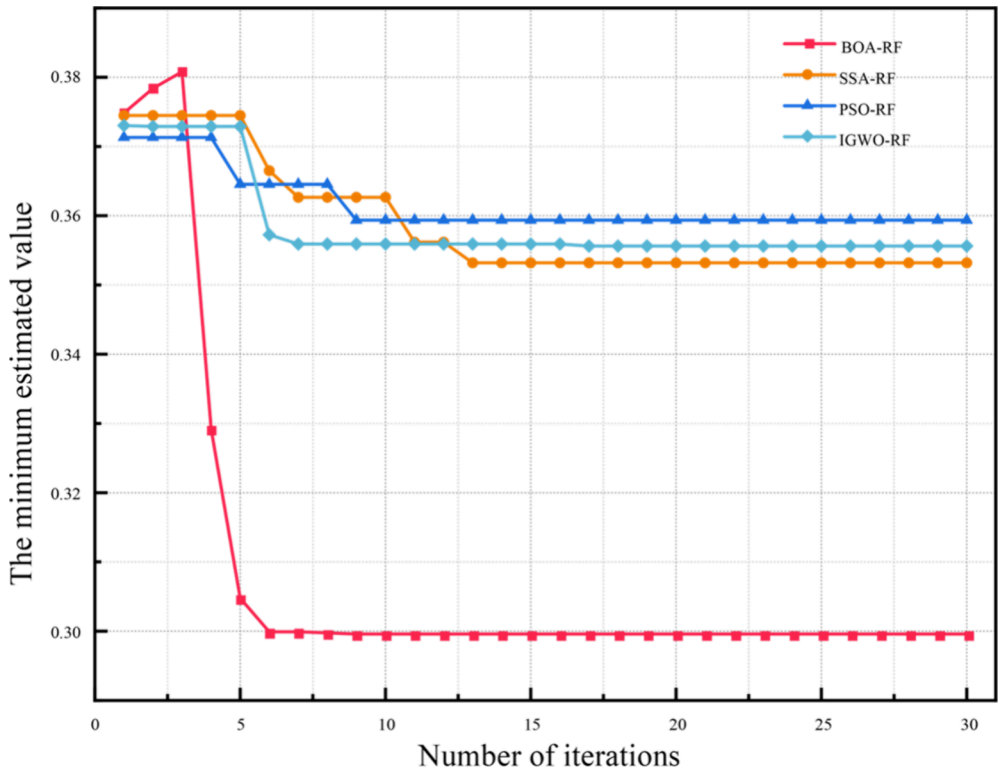


Figure 6. Iterative optimization process.

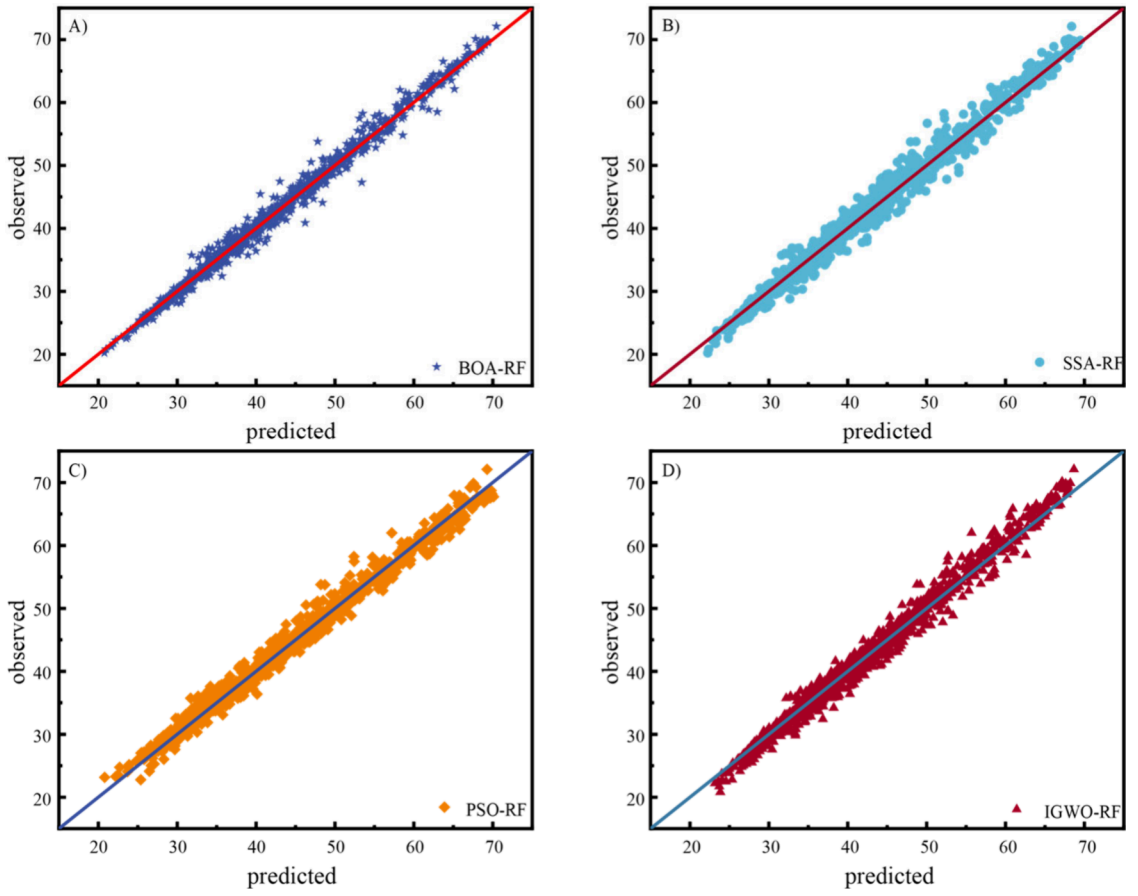


Figure 7. Comparison of results obtained from various types of models with actual data (A: BO-RF, B: SSA-RF, C: PSO-RF, D: IGWO-RF).

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

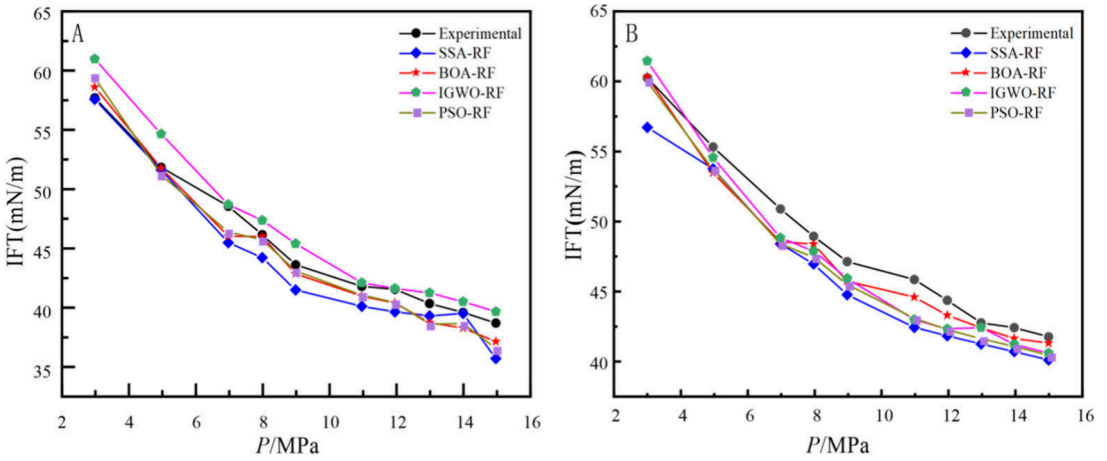


Figure 8. Comparison of four model predictions under isothermal conditions (A: T = 373.2 K, m = 1.98 mol/kg; B: T = 373.1 K, m = 4.9 mol/kg)

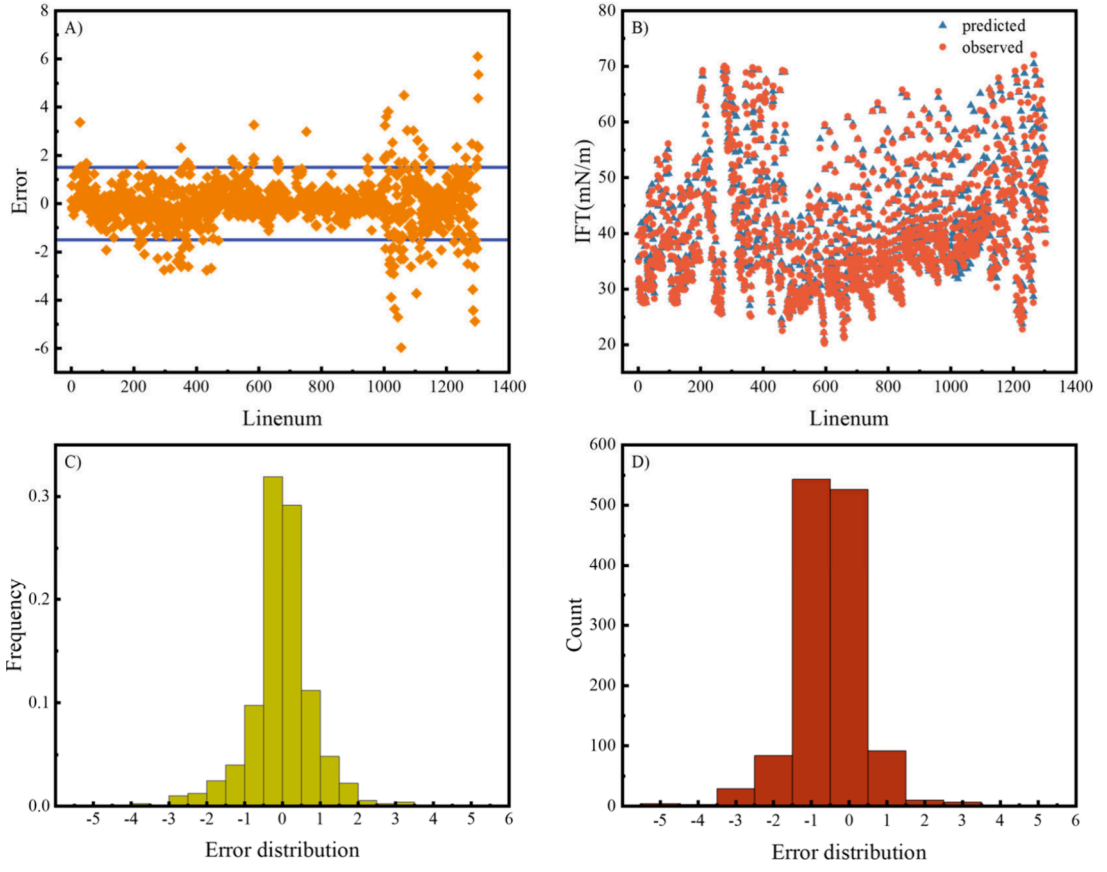


Figure 9. Data statistics of the BO-RF model. A: difference distribution between the experimental and predicted values, B: comparison between the experimental and predicted values, C: frequency plot of the difference distribution, D: statistics of the number of difference distributions.

Table 2. Presentation of Evaluation Indicators

Evaluation indicators	Expression
Mean absolute percentage error (MAPE)	$MAPE = \frac{100}{N}\sum_{t=1}^{N}\left\lvert\frac{observed_t - predicted_t}{observed_t}\right\rvert$
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(observed_t - predicted_t)^2}$
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2}$

trend of changes in IFT in relation to pressure, temperature, ionic concentration, and impurity gases, respectively. The predicted IFT by the model aligns with the trends of change in various influencing factors as discussed previously, thereby substantiating the accuracy of the model developed in this study.

5. IMPLICATIONS FOR CARBON STORAGE

The primary goal of CO_2-EWR technology is to achieve enhanced water recovery and safe CO_2 geological storage, where site selection determines the flexibility of this scheme. Considerable numbers of sedimentary basins can be found in

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

Table 3. Statistical Table of the Evaluation Indicators for Each Type of Model

	pure CO_2–brine			impure CO_2–pure water			total		
	RMSE	MAPE	R^2	RMSE	MAPE	R^2	RMSE	MAPE	R^2
PSO-RF	1.9555	2.6713	0.9651	1.6486	2.7187	0.9828	1.9160	2.6778	0.9674
IGWO-RF	1.9164	2.6676	0.9657	1.7937	2.9110	0.9786	1.9000	2.7012	0.9691
SSA-RF	1.9283	2.6625	0.9656	1.7149	2.7396	0.9792	1.9003	2.6731	0.9693
BO-RF	1.7666	1.9902	0.9709	1.7943	2.5584	0.9755	1.7705	2.0687	0.9729

Table 4. Comparison of the Prediction Performance of This Model with That of Other Explicit Models

Model	RMSE	R^2
Kamari et al. (2017)[63]	7.5016	
Amooie et al. (2019)[42]	3.81	
Amar et al. (2021)[33]	3.3	0.951
Liu et al. (2021)[41]	2.717	0.9560
Mouallem et al. (2024)[54]	4.28	0.886
This study	1.7705	0.9729

China, both on land and on the continental shelf. These basins have a wide distribution area and considerable sediment thickness, with the widespread occurrence of brine layers suitable for CO_2 sequestration, especially in the Xinjiang region.[11] According to previous studies, there are four primary forms of sequestration: structural, residual (capillary), solubility, and mineralization trapping.[16−19] The widespread existence of saline aquifers is suitable for CO_2-EWR projects, and assessing the storage capacity of injection sites is crucial. This study evaluated capillary trapping in the Tarim Basin in Xinjiang.

During the geological storage of CO_2, as mentioned in introuduction section, leakage occurs when the pressure difference between CO_2 and the saline system exceeds the breakthrough pressure. Due to the buoyancy effect of supercritical CO_2, buoyancy becomes the driving force for CO_2 leakage. As mentioned in eq 1, leakage occurs when the pressure difference between CO_2 and the saline system exceeds the breakthrough pressure. Due to the buoyancy effect of supercritical CO_2, buoyancy becomes the driving force for CO_2 leakage. To guarantee the airtightness and safety of the stored

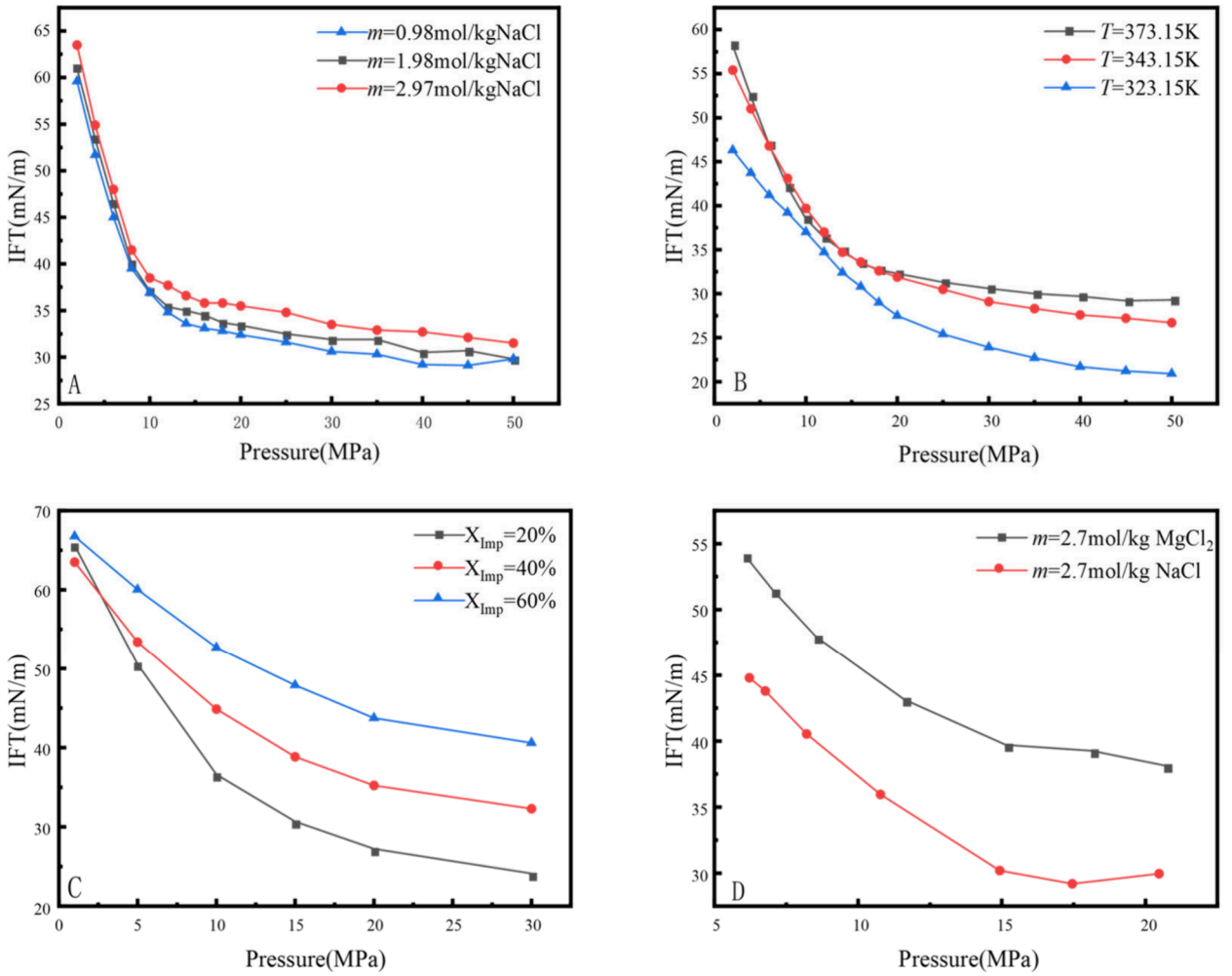


Figure 10. Prediction of IFT Influencing factors A: Salinity at T = 323.15 K with no impurities, B: Temperature by keeping salinity 0.98 mol/kg with no impurities, C: Presence of Impurities at T = 333.15 K, and D: Salt type at T = 344.15 K with no impurities.

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

CO_2, its buoyancy must not exceed the capillary pressure of the largest pore. Assuming that the height of the reservoir occupied by CO_2 stored in the sedimentary layer is h, the magnitude of the buoyancy force acting on CO_2 is $(\rho_{water}-\rho_{CO_2})gh$. If the variation in the contact angle with temperature and pressure is ignored, the threshold capillary pressure is $P_{ce} = 2\sigma/R$. When the maximum height of CO_2 storage is denoted as H, then:[64]

$$H = \frac{2\sigma}{(\rho_{water} - \rho_{CO_2})gR} \tag{20}$$

where σ is the IFT between CO_2 and brine; ρ is the density, with the subscripts water and CO_2 being the density of brine and carbon dioxide, respectively; g is the acceleration due to gravity; and R is the size of the largest pore or small seam in the cap layer. Here, due to the lack of contact angle data, we assume that the brine is an excellent wetting fluid; therefore, the contact angle is not considered in calculating the CO_2 storage capacity in the saline aquifer.[55,64,65]

Thus, the mass of CO_2 per unit of reservoir area can be expressed as:[42]

$$MCO_2 = \frac{2\sigma\rho CO_2\phi(1 - S_{swirr})}{(\rho_{water} - \rho_{CO_2})gR} \tag{21}$$

where x is the porosity of the brine layer in the sedimentary basin and y is the residual water saturation.

Recently, a number of approaches have been proposed for evaluating CO_2 sequestration capacity of the reservoir. The US Department of Energy[66] and CSLF have suggested evaluation methods; however, these methods are assessed on a macroscopic level, and intermolecular interactions are not taken into consideration. Thus, the Xinjiang region of China's Tarim Basin was selected as the injection site, and eq 21 was applied to estimate the CO_2 sequestration capacity of this basin at various depths in this study. The BO-RF model presented in this study was used to determine the IFT for the CO_2−brine system in eq 21, while the density of the fluids can be obtained via the NIST Web site (http://www.ap1700.com/). The values of the remaining parameters are detailed in Table 5.

Table 5. Values of Relevant Parameters

Parameters	R(m)	g(N/kg)	S_{swirr}	φ
numerical value	10^{-7}	9.8	0.1	0.15

Due to the variation in the concentration of saltwater in the selected basin, concentrations of 0.98 mol/kg and 2.97 mol/kg were used to simulate the influence of salinity on the estimation of the CO_2 storage capacity. In contrast to other scholars,[55,65,67] this study considered the seawater concentration and assessed sequestration between depths of 800 and 5000 m. The basin has a greater CO_2 storage capacity when CO_2 is injected approximately 1,000−2,000 m below the surface. The same observation was made by Chiquet et al.[64] The amount of sequestered CO_2 increases, although at a prolonged rate, until the depth of sequestration approaches 2500 m, as illustrated in Figure 11. R is a constant in this study. However, with increasing injection depth, the CO_2 sequestered increases in total. It can also be seen from Figure 11 that the higher the concentration in the brine layer is, the more CO_2 can be sequestered in the reservoir. As illustrated in Figure 10(B), IFT between CO_2 and brine increases with rising temperature, resulting in the increase of the CO_2 sequestered capacity of the reservoir. The burial depth had a limited influence on the CO_2 storage capacity. This may be because at deeper storage altitudes, the density differences between reservoir brine and CO_2, where it is in a supercritical state, are lower, resulting in a smaller change in IFT, which ultimately leads to a little change in CO_2 sequestration capacity of the reservoir. On the other hand, this is also related to wettability. Although brine is assumed to be an excellent wetting fluid in the calculation of CO_2 storage capacity in saline aquifers, thereby not considering the impact of the wettability remains an important parameter for CO_2 geological sequestration. It determines the affinity between CO_2 and the reservoir rock fluids, affecting the distribution and migration of CO_2 within the reservoir. Under water-wet conditions, the rock surface tends to bind more readily with water molecules, while CO_2 is more likely to form a continuous nonwetting phase, which aids in capturing and storing CO_2.[68] Temperature, pressure, salinity, salt type, and impurities gases impact wettability. Additionally, wettability influences the solubility of CO_2 in reservoir water, thereby affecting the efficiency of dissolution trapping.[69,70] IFT and wettability are significant factors affecting the quantity and safety of CO_2 storage in saline aquifers and should be given considerable attention during the estimation of the the CO_2 sequestered capacity of the reservoir.

6. CONCLUSIONS

In this study, a database of 1717 sets of experimental values was established, covering the temperature range of 293.15 to 448.15 K, a pressure of up to 69 MPa, and a concentration of salinities from 0 to 5.0 mol·kg^{-1}, basically covering the geological environment conditions of the storage of CO_2 saline aquifers. The database was divided into a training set, a verification set, and a test set, accounting for 8:1:1, respectively. In this study, three combined models, the SSA-RF, PSO-RF, and IGWO-RF models, were established for comparison with the BO-RF model. In addition, the IFT was used as an input parameter to evaluate the carbon sequestration potential of the Tarim Basin in China. On the basis of results, the following conclusions can be drawn:

1. The precision of the BO-RF model made in this study is nicer than other heuristic algorithms and displays good results in foretelling the IFT of carbon dioxide in brine. The BO-RF model shows the lowest RMSE (1.7705) and MAPE (2.0687%), and the highest R^2 value (0.9729). This study also selected other excellent ML models for comparison. The results revealed that the BO-RF model has the best performance.

2. The model was statistically sound since the use of isolated forests assisted in identifying outliers, which accounted for just 4% of the total data set. At the same time, the analysis in this study shows the critical factors that have an influence on the magnitude of the IFT in the CO_2-brine system, and these variables have been ranked in order of importance: pressure > bivalent cation molality > mole fraction of CH_4 > mole fraction of N_2 > temperature > monovalent cation molality. It was found that pressure had the greatest effect, and the molar concentration of monovalent cations had the least effect on the IFT.

3. With increasing injection depth, the CO_2 concentration also decreases, and eventually, the CO_2 sequestered capacity of

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

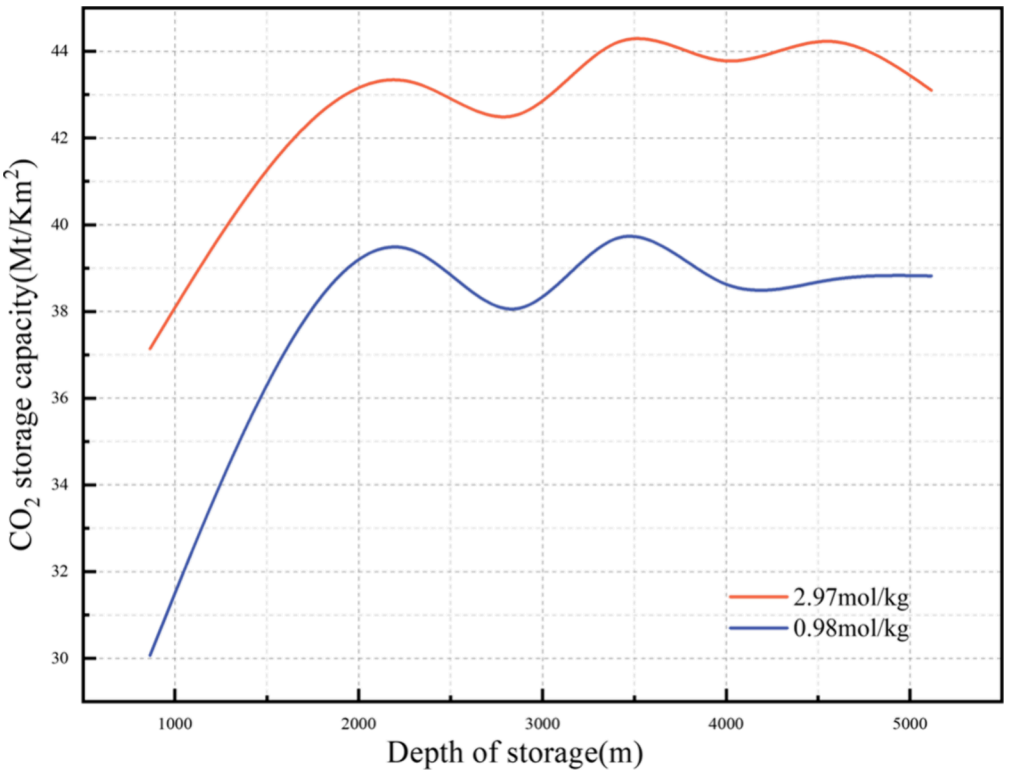


Figure 11. Variation curve of CO_2 sequestration with burial depth per unit basin area.

the reservoir CO_2 concentration increases in total. However, the burial depth had a limited influence on the CO_2 storage capacity due to the phase change of the CO_2. It has also been found that if the salt content of the water increases, the carbon storage potential will also increase. However, the cost of CO_2 storage in this project increased with increasing injection depth, which led to a decrease in its feasibility. Therefore, the storage depth should be comprehensively considered in the site selection process, and the physical properties of reservoir fluids, especially the salinity, should be sampled and analyzed to obtain the injection site and storage depth that are most suitable for carbon storage projects.

AUTHOR INFORMATION

Corresponding Author

Meiheriayi Mutailipu − *Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education and School of Electrical Engineering, Xinjiang University, Urumqi 830017, China;* orcid.org/0009-0002-8920-1687; Email: mhriay@xju.edu.cn

Authors

Yande Yang − *Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education and School of Electrical Engineering, Xinjiang University, Urumqi 830017, China*

Kaishuai Zuo − *Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education and School of Electrical Engineering, Xinjiang University, Urumqi 830017, China*

Qingnan Xue − *Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education and School of Electrical Engineering, Xinjiang University, Urumqi 830017, China*

Qi Wang − *Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education and School of Electrical Engineering, Xinjiang University, Urumqi 830017, China*

Fusheng Xue − *Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education and School of Electrical Engineering, Xinjiang University, Urumqi 830017, China*

Gang Wang − *Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education and School of Electrical Engineering, Xinjiang University, Urumqi 830017, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c04888

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the funding from National Natural Science Foundation of China (grant number 52366010), Special Project for key Research and Development Program of Xinjiang Autonomous Region (grant number 2022B01033-2), Central guide the development of local technology specific fund (grant number ZYYD2022C16), Basic scientific research project of universities in Xinjiang Autonomous Region-Education & Training (grant number XJEDU2023P026), and Xinjiang Autonomous Region-Outstanding Young Talent Program-Youth Talent Support Project (grant number 2023TSYCQNTJ0035).

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

REFERENCES

(1) Awaworyi Churchill, S.; Inekwe, J.; Smyth, R.; Zhang, X. R&D intensity and carbon emissions in the G7:1870−2014. *Energy Econ* **2019**, *80*, 30−37.

(2) Liu, Y.; Huang, L.; Onstein, E. How do age structure and urban form influence household CO_2 emissions in road transport? Evidence from municipalities in Norway in 2009, 2011 and 2013. *J. Cleaner Prod* **2020**, *265*, No. 121771.

(3) Duan, H.; He, C.; Pu, S. A new circular neural grey model and its application to CO_2 emissions in China. *J. CLEAN PROD* **2024**, *445*, No. 141318.

(4) Cao, Q.; Kang, W.; Sajid, M. J.; Cao, M. Research on the optimization of carbon abatement efficiency in China on the basis of task allocation. *J. CLEAN PROD* **2021**, *299*, No. 126912.

(5) Guo, C.; Zhou, Z.; Liu, X.; Liu, X.; Meng, J.; Dai, H. The unintended dilemma of China's target-based carbon neutrality policy and provincial economic inequality. *Energy Econ* **2023**, *126*, No. 107002.

(6) Guan, D.; Meng, J.; Reiner, D. M.; Zhang, N.; Shan, Y.; Mi, Z.; Shao, S.; Liu, Z.; Zhang, Q.; Davis, S. J. Structural decline in China's CO_2 emissions through transitions in industry and energy systems. *NAT GEOSCI* **2018**, *11* (8), 551−555.

(7) Wang, X.; Liu, Y.; Bi, J.; Liu, M. New challenges of the Belt and Road Initiative under China's″ 3060″ carbon target. *J. Clean. Prod.* **2022**, *376*, No. 134180.

(8) Li, Q.; Chen, Z. A.; Zhang, J. T.; Liu, L. C.; Li, X. C.; Jia, L. Positioning and revision of CCUS technology development in China. *Int. J. Greenh. Gas Control* **2016**, *46*, 282−293.

(9) Wang, N.; Akimoto, K.; Nemet, G. F. What went wrong? Learning from three decades of carbon capture, utilization and sequestration (CCUS) pilot and demonstration projects. *ENERG POLICY* **2021**, *158*, No. 112546.

(10) Razak, A. A. A.; Saaid, I. M.; Yusof, M. A. M.; Husein, N.; Zaidin, M. F.; Mohamad Sabil, K. Physical and chemical effect of impurities in carbon capture, utilisation and storage. *J. Pet Explor Prod Technol.* **2023**, *13* (5), 1235−1246.

(11) Li, Q.; Wei, Y. N.; Liu, G.; Shi, H. CO_2-EWR: a cleaner solution for coal chemical industry in China. *J. CLEAN PROD* **2015**, *103*, 330−337.

(12) Edouard, M. N.; Okere, C. J.; Ejike, C.; Dong, P.; Suliman, M. A. Comparative numerical study on the co-optimization of CO_2 storage and utilization in EOR, EGR, and EWR: Implications for CCUS project development. *APPL ENERG* **2023**, *347*, No. 121448.

(13) Li, X.; Wei, N.; Jiao, Z.; Liu, S.; Dahowski, R. Cost curve of large-scale deployment of CO_2-enhanced water recovery technology in modern coal chemical industries in China. *Int. J. Greenh. Gas Control* **2019**, *81*, 66−82.

(14) Eyitayo, S. I.; Okere, C. J.; Hussain, A.; Gamadi, T.; Watson, M. C. Synergistic sustainability: Future potential of integrating produced water and CO_2 for enhanced carbon capture, utilization, and storage (CCUS). *ENERG CONVERS MANAGE* **2024**, *351*, No. 119713.

(15) Yusuf, M.; Ibrahim, H. A comprehensive review on recent trends in carbon capture, utilization, and storage techniques. *J. Environ* **2023**, *11* (6), No. 111393.

(16) Zhang, C.; Wang, M. A critical review of breakthrough pressure for tight rocks and relevant factors. *J. Nat. Gas Sci. Eng.* **2022**, *100*, No. 104456.

(17) Zhang, C.; Wang, M. CO_2/brine interfacial tension for geological CO_2 storage: A systematic review. *J. Pet. Sci. Eng.* **2023**, *220*, No. 111154.

(18) Mouallem, J.; Arif, M.; Raza, A.; Glatz, G.; Rahman, M. M.; Mahmoud, M.; Iglauer, S. Critical review and meta-analysis of the interfacial tension of CO_2-brine and H_2-brine systems: Implications for CO_2 and H_2 geo-storage. *Fuel* **2024**, *356*, No. 129575.

(19) Raza, A.; Rezaee, R.; Bing, C. H.; Gholami, R.; Hamid, M. A.; Nagarajan, R. Carbon dioxide storage in subsurface geologic medium: A review on capillary trapping mechanism. *J. Pet* **2016**, *25* (3), 367−373.

(20) Ali, F.; Negash, B. M.; Ridha, S.; Abdulelah, H. A review on the interfacial properties of caprock/CO_2/brine system-implications for structural integrity of deep saline aquifers during geological carbon storage. *Earth Sci. Rev.* **2023**, *247*, No. 104600.

(21) Bikkina, P. K.; Shoham, O.; Uppaluri, R. Equilibrated Interfacial Tension Data of the CO_2−Water System at High Pressures and Moderate Temperatures. *J. Chem. Eng. Data* **2011**, *56* (10), 3725−3733.

(22) Aggelopoulos, C. A.; Robin, M.; Perfetti, E.; Vizika, O. CO_2/$CaCl_2$ solution interfacial tensions under CO_2 geological storage conditions: Influence of cation valence on interfacial tension. *Adv. Water Resour* **2010**, *33* (6), 691−697.

(23) Aggelopoulos, C.; Robin, M.; Vizika, O. Interfacial tension between CO_2 and brine (NaCl+ $CaCl_2$) at elevated pressures and temperatures: The additive effect of different salts. *Adv. Water Resour.* **2011**, *34* (4), 505−511.

(24) Chalbaud, C.; Robin, M.; Lombard, J.; Martin, F.; Egermann, P.; Bertin, H. Interfacial tension measurements and wettability evaluation for geological CO_2 storage. *Adv. Water Resour.* **2009**, *32* (1), 98−109.

(25) Li, X. S.; Boek, E. S.; Maitland, G. C.; Trusler, J. P. M. Interfacial Tension of (Brines+CO_2): $CaCl_2$(aq), $MgCl_2$(aq), and Na_2SO_4(aq) at Temperatures between (343 and 423) K, Pressures between (2 and 50) MPa, and Molalities of (0.5 to 5) mol.kg^{-1}. *J. Chem. Eng. Data* **2012**, *57* (5), 1369−1375.

(26) Ren, Q.-Y.; Chen, G.-J.; Yan, W.; Guo, T.-M. Interfacial Tension of (CO_2 + CH_4) + Water from 298 to 373 K and Pressures up to 30 MPa. *J. Chem. Eng. Data* **2000**, *45* (4), 610−612.

(27) Li, Z.; Wang, S.; Li, S.; Liu, W.; Li, B.; Lv, Q. C. Accurate Determination of the CO_2−Brine Interfacial Tension Using Graphical Alternating Conditional Expectation. *Energy Fuels* **2014**, *28* (1), 624−635.

(28) Shiga, M.; Morishita, T.; Sorai, M. Interfacial tension of carbon dioxide - water under conditions of CO_2 geological storage and enhanced geothermal systems: A molecular dynamics study on the effect of temperature. *Fuel* **2023**, *337*, No. 127219.

(29) Iglauer, S.; Mathew, M. S.; Bresme, F. Molecular dynamics computations of brine-CO_2 interfacial tensions and brine-CO_2-quartz contact angles and their effects on structural and residual trapping mechanisms in carbon geo-sequestration. *J. COLLOID INTERF SCI* **2012**, *386* (1), 405−414.

(30) Wang, L.; Yang, J.; Mendenhall, J. D.; Cristancho, D. E.; Dowdle, J. R. An interfacial statistical associating fluid theory (iSAFT) approach for surface/interfacial tension predictions. *Fluid Phase Equilib.* **2018**, *476* (Pta2), 193−201.

(31) Mutailipu, M.; Song, Y.; Yao, Q.; Liu, Y.; Trusler, J. M. Solubility and interfacial tension models for CO_2−brine systems under CO_2 geological storage conditions. *Fuel* **2024**, *357*, No. 129712.

(32) Chow, Y. F.; Eriksen, D. K.; Galindo, A.; Haslam, A. J.; Jackson, G.; Maitland, G. C.; Trusler, J. M. J. F. P. E. Interfacial tensions of systems comprising water, carbon dioxide and diluent gases at high pressures: Experimental measurements and modelling with SAFT-VR Mie and square-gradient theory. *Fluid Phase Equilib.* **2016**, *407*, 159−176.

(33) Nait Amar, M. Towards improved genetic programming based-correlations for predicting the interfacial tension of the systems pure/impure CO_2-brine. *J. Taiwan Inst Chem. Eng.* **2021**, *127*, 186−196.

(34) Zhang, J.; Feng, Q.; Wang, S.; Zhang, X.; Wang, S. Estimation of CO_2−brine interfacial tension using an artificial neural network. *J. Supercrit. Fluids* **2016**, *107* (4), 31−37.

(35) Rashid, S.; Harimi, B.; Hamidpour, E. Prediction of CO_2-Brine interfacial tension using a rigorous approach. *J. Nat. Gas Sci. Eng.* **2017**, *45*, 108−117.

(36) Li, X. S.; Boek, E.; Maitland, G. C.; Trusler, J. P. M. Interfacial Tension of (Brines + CO_2): (0.864 NaCl+0.136 KCl) at Temperatures between (298 and 448) K, Pressures between (2 and 50) MPa, and Total Molalities of (1 to 5) mol/kg. *J. Chem. Eng. Data* **2012**, *57* (4), 1078−1088.

https://doi.org/10.1021/acsomega.4c04888
ACS Omega 2024, 9, 37265−37277

(37) Mutailipu, M.; Liu, Y.; Jiang, L.; Zhang, Y. Measurement and estimation of CO_2−brine interfacial tension and rock wettability under CO_2 sub-and super-critical conditions. *J. COLLOID INTERF SCI* **2019**, *534*, 605−617.
(38) Jerauld, G. R.; Kazemi, A. An improved simple correlation for accurate estimation of CO_2-Brine interfacial tension at reservoir conditions. *J. Pet. Sci. Eng.* **2022**, *208*, No. 109537.
(39) Guo, W.; Gao, Z.; Guo, H.; Cao, W. Hydrogeochemical and sediment parameters improve predication accuracy of arsenic-prone groundwater in random forest machine-learning models. *SCI TOTAL ENVIRON* **2023**, *897*, No. 165511.
(40) Ratnakar, R. R.; Chaubey, V.; Dindoruk, B. A novel computational strategy to estimate CO_2 solubility in brine solutions for CCUS applications. *APPL ENERG* **2023**, *342*, No. 121134.
(41) Liu, X.; Mutailipu, M.; Zhao, J.; Liu, Y. Comparative analysis of four neural network models on the estimation of CO_2−brine interfacial tension. *ACS omega* **2021**, *6* (6), 4282−4288.
(42) Amooie, A.; Hemmati-Sarapardeh, A.; Karan, K.; Husein, M. M.; Dabir, B.; et al. Data-Driven Modeling of Interfacial Tension in Impure CO_2-Brine Systems with Implications for Geological Carbon Storage. *Int. J. Greenh. Gas Control* **2019**, *90*, No. 102811.
(43) Mutailipu, M.; Xue, Q.; Li, T.; Yang, Y.; Xue, F. Thermodynamic Properties of a Gas−Liquid−Solid System during the CO_2 Geological Storage and Utilization Process: A Review. *Energies* **2023**, *16* (21), 7374.
(44) Breiman. Random forests. *MACH LEARN* **2001**, *45* (1), 5−32.
(45) Mohammed, A.; Kora, R. A comprehensive review on ensemble deep learning: Opportunities and challenges. *J. King Saud Univ Sci.* **2023**, *35* (2), 757−774.
(46) Rasmussen, C. E.; Nickisch, H. Gaussian processes for machine learning (GPML) toolbox. *J. MACH LEARN RES* **2010**, *11*, 3011−3015.
(47) Srinivas, N.; Krause, A.; Kakade, S. M.; Seeger, M. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. *arXiv.org* **2009**. DOI: 10.48550/arXiv.0912.3995.
(48) Yekeen, N.; Padmanabhan, E.; Abdulelah, H.; Irfan, S. A.; Okunade, O. A.; Khan, J. A.; Negash, B. M. CO_2/brine interfacial tension and rock wettability at reservoir conditions: A critical review of previous studies and case study of black shale from Malaysian formation. *J. PETROL SCI ENG* **2021**, *196*, No. 107673.
(49) Bachu, S.; Bennion, D. B. Interfacial tension between CO_2, freshwater, and brine in the range of pressure from (2 to 27) MPa, temperature from (20 to 125) C, and water salinity from (0 to 334 000) mg· L−1. *J. Chem. Eng. Data* **2009**, *54* (3), 765−775.
(50) Pereira, L. M.; Chapoy, A.; Burgass, R.; Oliveira, M. B.; Coutinho, J. A.; Tohidi, B. Study of the impact of high temperatures and pressures on the equilibrium densities and interfacial tension of the carbon dioxide/water system. *J. Chem. Thermodyn* **2016**, *93*, 404−415.
(51) Liu, Y.; Tang, J.; Wang, M.; Wang, Q.; Tong, J.; Zhao, J.; Song, Y. Measurement of interfacial tension of CO_2 and NaCl aqueous solution over wide temperature, pressure, and salinity ranges. *J. Chem. Eng. Data* **2017**, *62* (3), 1036−1046.
(52) Yan, W.; Zhao, G.-Y.; Chen, G.-J.; Guo, T.-M. Interfacial tension of (methane+ nitrogen)+ water and (carbon dioxide+ nitrogen)+ water systems. *J. Chem. Eng. Data* **2001**, *46* (6), 1544−1548.
(53) Zhao, L.; Ji, J.; Tao, L.; Lin, S. Ionic effects on supercritical CO_2−brine interfacial tensions: Molecular dynamics simulations and a universal correlation with ionic strength, temperature, and pressure. *Langmuir* **2016**, *32* (36), 9188−9196.
(54) Mouallem, J.; Raza, A.; Glatz, G.; Mahmoud, M.; Arif, M. Estimation of CO_2-brine interfacial tension using machine learning: implications for CO_2 geo-storage. *J. Mol. Liq.* **2024**, *393*, No. 123672.
(55) Shen, B.; Yang, S.; Hu, J.; Gao, Y.; Xu, H.; Gao, X.; Chen, H. Application of Heterogeneous Ensemble Learning for CO_2−Brine Interfacial Tension Prediction: Implications for CO_2 Storage. *ENERG FUEL* **2024**, *38* (5), 4401−4416.
(56) Mouallem, J.; Arif, M.; Isah, A.; Raza, A.; Rahman, M. M.; Mahmoud, M.; Kamal, M. S. Effect of formation brine on interfacial interaction: Implications for CO_2 storage. *Fuel* **2024**, *371*, No. 131986.
(57) Mutailipu, M.; Jiang, L.; Fu, J.; Wang, Z.; Yu, T.; Lu, Z.; Liu, Y. Effects of Na^+, K^+, Ca^{2+}, and Mg^{2+} cations on CO_2−brine interfacial tension under offshore storage conditions. *GREENH GASES* **2018**, *8* (4), 762−780.
(58) Jafari Raad, S. M.; Hassanzadeh, H. Does impure CO_2 impede or accelerate the onset of convective mixing in geological storage? *Int. J. Greenh. Gas Control* **2016**, *54*, 250−257.
(59) Lesouple, J.; Baudoin, C.; Spigai, M.; Tourneret, J.-Y. Generalized isolation forest for anomaly detection. *Pattern Recognit* **2021**, *149*, 109−119.
(60) Xue, J.; Shen, B. A novel swarm intelligence optimization approach: sparrow search algorithm. *SYST SCI CONTROL ENG* **2020**, *8* (1), 22−34.
(61) Zhan, Z. H.; Zhang, J.; Li, Y.; Chung, H. S. H. Adaptive Particle Swarm Optimization. *IEEE Trans Cybern* **2009**, *39* (6), 1362−1381.
(62) Nadimi-Shahraki, M. H.; Taghian, S.; Mirjalili, S. An improved grey wolf optimizer for solving engineering problems. *Expert Syst. Appl.* **2021**, *166*, No. 113917.
(63) Kamari, A.; Pournik, M.; Rostami, A.; Amirlatifi, A.; Mohammadi, A. H. Characterizing the CO_2-brine interfacial tension (IFT) using robust modeling approaches: A comparative study. *J. MOL LIQ* **2017**, *246*, 32−38.
(64) Chiquet, P.; Daridon, J. L.; Broseta, D.; Thibeau, S. CO_2/water interfacial tensions under pressure and temperature conditions of CO_2 geological storage. *ENERG CONVERS MANAGE* **2007**, *48* (3), 736−744.
(65) Safaei-Farouji, M.; Vo Thanh, H.; Sheini Dashtgoli, D.; Yasin, Q.; Radwan, A. E.; Ashraf, U.; Lee, K.-K. Application of robust intelligent schemes for accurate modelling interfacial tension of CO_2 brine systems: Implications for structural CO_2 trapping. *Fuel* **2022**, *319*, No. 123821.
(66) Goodman, A.; Hakala, A.; Bromhal, G.; Deel, D.; Rodosta, T.; Frailey, S.; Small, M.; Allen, D.; Romanov, V.; Fazio, J. U.S.; et al. DOE methodology for the development of geologic storage potential for carbon dioxide at the national and regional scale. *Int. J. Greenh. Gas Control* **2011**, *5* (4), 952−965.
(67) Vo Thanh, H.; Sugai, Y.; Nguele, R.; Sasaki, K. Integrated workflow in 3D geological model construction for evaluation of CO_2 storage capacity of a fractured basement reservoir in Cuu Long Basin. *Vietnam. Int. J. Greenh. Gas Control* **2019**, *90*, No. 102826.
(68) Wang, X.; Li, S.; Tong, B.; Jiang, L.; Lv, P.; Zhang, Y.; Liu, Y.; Song, Y. Multiscale wettability characterization under CO_2 geological storage conditions: A review. *RENEW SUST ENERG REV* **2024**, *189*, No. 113956.
(69) Awad, M. M.; Espinoza, D. N. Mudrock wettability at pressure and temperature conditions for CO_2 geological storage. *Int. J. Greenh. Gas Control* **2024**, *135*, No. 104160.
(70) Ali, F.; Negash, B. M.; Ridha, S.; Abdulelah, H. A review on the interfacial properties of caprock/CO_2/brine system-implications for structural integrity of deep saline aquifers during geological carbon storage. *Earth Sci. Rev.* **2023**, *247*, No. 104600.