Developing machine learning models to predict CO_2 trapping performance in deep saline aquifers

Hung Vo Thanh (vothanhhung1990@snu.ac.kr)
Seoul National University https://orcid.org/0000-0002-7094-9380
Kang-Kun Lee
Seoul National University

Research Article

Keywords: Geological Carbon Storage (GCS), CO2 storage, saline aquifers, Machine Learning

Posted Date: June 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-587644/v1

Abstract

Deep saline formations are considered as potential sites for geological carbon storage (GCS). To better understand the CO_2 trapping mechanism in saline aquifers, it is necessary to develop robust tools to evaluate CO_2 trapping efficiency. This paper introduces the application of Gaussian process regression (GPR), support vector machine (SVM), and random forest (RF) to predict CO_2 trapping efficiency in saline formations. First, the uncertainty variables, including geologic parameters, petrophysical properties, and other physical characteristics data were utilized to create a training dataset. A total of 101 reservoir simulation samples were then performed, and the residual trapping, solubility trapping, and cumulative CO_2 injection were collected. The predicted results indicate that three machine learning (ML) models that evaluate performance from high to low: GPR, SVM, and RF can be selected to predict the CO_2 trapping efficiency in deep saline formations. The GPR model has an excellent CO_2 trapping prediction efficiency with the highest correlation factor (R^2 = 0.992) and lowest root mean square error (RMSE = 0.00491). The accuracy and stability of the GPR models were verified for an actual reservoir in offshore Vietnam. The predictive models obtained a good agreement between the simulated field and the predicted trapping index. These findings indicate that the GPR ML models can support the numerical simulation as a robust predictive tool for estimating the performance of CO_2 trapping in the subsurface.

1. Introduction

Geological carbon storage (GCS) is a potential method for tackling greenhouse gas emissions and mitigating climate change (Bachu, 2018, 2000; Hitchon et al., 1999). CO_2 can be sequestered into many potential storage sites, such as deep saline aquifers, depleted oil and gas reservoirs, depleted unconventional formations, residual oil zones, and deep marine formations (Bachu, 2002; Edwards et al., 2015; Nguyen et al., 2020; Petvipusit et al., 2014; Ren and Duncan, 2019; Senel and Chugunov, 2012). Among these storage sites, deep saline formations have a largest storage capacity (Wang et al., 2016). The International Energy Agency (IEA) stated that saline aquifers could store a maximum of 10,000 Gigaton CO_2 (Javaheri et al., 2009).

Alcalde et al. (2018) presented excellent findings on CO_2 storage to secure climate change using a numerical program. These authors investigated the leakage of CO_2 stored in different storage formations that lead to GCS being identified as a favorable option for mitigating climate change.

Therefore, CO_2 storage in the subsurface aligns with the net-zero carbon emission strategy, globally. Numerous studies have investigated the potential of saline aquifers for CO_2 sequestration (Ajayi et al., 2019; Al-Khdheeawi et al., 2018a; Emad A. Al-Khdheeawi et al., 2017; Bachu, 2018; Fang et al., 2010; Khudaida and Das, 2020). Ajayi et al. (2019) demonstrated the feasibility of CO_2 sequestration in Dubai onshore The simulation results suggest that onshore saline reservoirs are excellent candidates for long-term CO_2 geosequestration. In addition, Kempka et al. (2013) explored long-term CO_2 trapping in the Ketzin pilot site (Germany). Their work demonstrated that 98.3% of the CO_2 injection was soluble in the

reservoir after 16,000 years of simulation. The latest research was conducted on CO_2 plume dynamics and storage efficiency in a fluvial reservoir by Zapata et al. (2020). These authors investigated the anisotropy, chemical reaction, and facies distribution effects on the CO_2 storage performance.

Furthermore, the feasibility of CO_2 geosequestration at the field scale could be classified into two categories. First, an aquifer's geological characteristics affects the potential of CO_2 trapping efficiency (Al-Khdheeawi et al., 2018a, 2018b; Emad A Al-Khdheeawi et al., 2017; Temitope et al., 2016a). Second, typical CO_2 trapping, including stratigraphic, capillary, dissolution, and mineral trapping, carries the risk associated with storage sites (Dai et al., 2016, 2014; Li and Zhang, 2014; Miocic et al., 2019; Nghiem et al., 2009b; Soltanian et al., 2017).

In terms of CO_2 trapping efficiency, Al-Khdheeawi et al. (2018b) proposed the Water Alternating Gas Process to enhance residual and solubility trapping in heterogeneous reservoirs. Recently, Vo Thanh et al. (2020a) developed a robust optimization workflow under geological uncertainty to improve the trapping efficiency in fluvial sandstone reservoirs. These scholars performed several simulation scenarios considering the petrophysical properties to predict CO_2 trapping efficiency in the subsurface.

Although modeling and simulation results can be employed to predict CO_2 flow in the subsurface, these results are used to analyze the amount and distribution of CO_2 under specific constraints for well optimization (Afanasyev et al., 2016; Kim et al., 2017; Lothe et al., 2014). However, reservoir simulations are complicated and time-consuming in real-field storage projects. Additionally, reservoir simulations require a large amount of subsurface data, such as horizontal seismic interpretations and petrophysical properties (Vo Thanh et al., 2020c). Machine learning (ML) techniques can overcome these challenges. ML and artificial intelligence can be adapted as practical tools for reservoir simulations (Ertekin and Sun, 2019; Lee, 2020).

Recently, Ahmadi et al. (2018) utilized the least square SVM to develop a robust proxy model for miscible CO_2 injection. Later, Al-Mudhafar (2019) applied multivariate adaptive regression splines and RF model proxy models to generate a metamodel for CO_2-enhanced oil recovery (EOR) in unconventional formation. Most recently, You et al. (2020) implemented ML-based optimization for CO_2-EOR and storage projects. Using the ML workflow, these authors could improve oil recovery, CO_2 storage, and net present value of EOR projects.

In addition, Zhong et al. (2015) used SVM to predict production performance in the Permian Basin Wolfcamp Shale wells. In contrast, Gaussian process regression (GPR) was used to predict the total organic carbon for shale gas reservoir characterization (Yu et al., 2016). These scholars found that prediction using GPR was accurate compared to the prediction results of TOC using an artificial neural network (ANN) and SVM. In the steam-assisted gravity drainage process, Kim and Shin (2020) used RF, multiple linear regression, and ANN methods to predict the shale barrier size and spatial location in

Canada. RF was also considered as a predictive model for evaluating CO_2-EOR and storage performance in residual oil zones (ROZs) (Chen and Pawar, 2018, 2019).

Furthermore, ML techniques have been widely applied to several subsurface problems, such as history matching (Hadi et al., 2016; He et al., 2016; Razak and Jafarpour, 2019) and geostatistical reservoir characterization (Al-Mudhafar, 2020, 2017).

In CO_2 sequestration applications, the least square SVM can generate high-accuracy prediction models for CO_2 solubility in brine (Ahmadi and Ahmadi, 2016). These authors stated that the average absolute deviation between the predictive models and experimental data was less than 0.1%. Later, Yan et al. (2020) successfully constructed predictive models for estimating permeability changes during CO_2 sequestration in coal seams using SVM and intelligent optimization methods.

As presented in the literature review, ML techniques have not been considered for developing robust and fast predictive models to estimate the CO_2 trapping efficiency in deep saline formations. Thus, this study proposes supervised ML algorithms for predictive construction models to predict the effectiveness of CO_2 trapping feasibility by reducing the computational time of the reservoir simulator in deep saline formations.

This study aims to propose ML-based prediction models for CO_2 trapping efficiency. These models can be further adapted as a robust and fast tool to investigate the feasibility of CO_2 geosequestration in the initial phase of geological CO_2 storage projects.

Three popular ML techniques, GPR, SVM, and RF, were used to predict the effectiveness of CO_2 trapping efficiency. The best ML technique was then employed in a real reservoir offshore Vietnam to evaluate the accuracy and stability of the proposed predictive model. To the best of our knowledge, this study is the first to adapt supervised ML techniques (GPR, SVM, RF) to develop predictive tools for CO_2 trapping in saline formation. Briefly, the main goals of our study were to:

- Develop predictive models based on ML techniques for the accurate prediction of CO_2 trapping efficiency in deep saline aquifers.
- Build fast predictive tools to evaluate trapping performance in CO_2 geosequestration.
- Evaluate the accuracy of the ML model using the real-field case from a previous study.

2. Methodology

ML techniques are powerful tools for developing predictive models. ML applies pattern recognition to guide reservoir performance developed on a computer (Mohaghegh, 2011). Furthermore, the predictive models based on ML could generate a rapid and accurate forecast in place of the reservoir simulator. In this study, ML algorithms were adapted to develop predictive models to evaluate the trapping efficiency in deep saline aquifers. The workflow for creating the predictive models is shown in **Fig. 1a**.

First step: CO_2 sequestration model construction. The 3D geological model was considered with a compositional modeling package (CMG-GEM) to simulate the CO_2 trapping mechanism. The specific equations of the simulation package are expressed as follows (Nghiem et al., 2004):

$$\sum_{j=o,g,w} \Delta T \frac{k_{ij}}{\mu_{ij}} \rho_j y_{ij} \Delta \phi_j (convection) + \sum_{j=o,g,w} \Delta A D_{ij} \Delta \rho_j y_{ij} (diffusion, dispersion) +$$

$$+ \quad r_i (reaction) + q_i (injection) = \frac{v}{\Delta t} (N_i^{n+1} - N_i^n)(accumulation) \qquad (10)$$

“Convection” represents the flow induced by the pressure difference; Darcy’s Law explains this mechanism. The second factor is the rate of diffusion in the liquid state (Kim et al., 2017). The interaction between the reaction and injection results in precipitation and solubility between mineral components and the formation of brine (Kim et al., 2017), are key points in the CO_2 sequestration process (Nghiem et al., 2010).

Furthermore, the petrophysical properties and other parameters were referenced from previous studies (Fang et al., 2010; House et al., 2003; Issautier et al., 2013; Jin et al., 2012; Lengler et al., 2010; Mediato et al., 2017; Shogenov et al., 2017; Singh et al., 2010). As illustrated in **Fig. 1b**, the simulation model included 2660 (19 × 28 ×5) grid cells. The porosity and permeability properties adapt from the PUNQS3 project (Gu et al., 2005). The data set is obtained from the website of Imperial College London (PUNQ S3, 2021).

Moreover, there is a large number of factors that influence CO_2 trapping performance in deep saline formations. These factors include saline formation, depth, reservoir thickness, petrophysical properties, residual gas saturation, compressibility, wettability, capillary pressure, injection methods, and anisotropy ratio (Al-Khdheeawi et al., 2018a; Dai et al., 2018; Silva and Ranjith, 2012). Among these factors, determining the most critical parameter is necessary. Therefore, several scholars have conducted a sensitivity analysis to rank the most influential factors in the CO_2 sequestration process (Abbaszadeh and Shariatipour, 2018; Gibson-Poole et al., 2006; Lee et al., 2010; Liu and Zhang, 2011). Based on these literature reviews, eight parameters, depth, porosity (Por), permeability (Perm), thickness (h), residual gas saturation (Sgr), salinity, injection rate, and ratio of vertical to horizontal permeability (Kv/Kh) were selected for this study. The geological and uncertainty parameters are presented in Table 1.

Table 1
The uncertainty factors considered for conducting the simulation jobs (Bachu, 2008; Beni et al., 2011; Bu Ali et al., 2011; Dai et al., 2017; Ezeanyim and Shariatipour, 2016; Fang et al., 2010; House et al., 2003; Jia et al., 2018; Singh et al., 2010; Song et al., 2020; Sung et al., 2014; Temitope et al., 2016b; Xiao et al., 2019)

Uncertainty factors	Minimum	Base case	Maximum	Units
Porosity	0.01	-	0.4	-
Permeability	0.01	-	2000	*mD*
Thickness	2	5	200	*M*
Depth	800	1000	3000	*M*
Residual Gas Saturation	0.1	0.4	0.5	-
Salinity	10000	120000	400000	ppm
Kv/Kh	100	1500	1500	-
CO_2 injection rate	39	2740	5480	*ton/day*

In this study, relative permeability properties were obtained from (Vo Thanh et al., 2020a). Moreover, the residual CO_2 trapping mechanism in this study was mainly based on land trapping models (Vo Thanh et al., 2020a). Therefore, the drainage and imbibition processes can be calculated during the simulation using Land's residual model (Land, 1968). Figure 1c depicts the relative permeability properties and the land trapping model used in this research.

To evaluate the efficiency of CO_2 trapping performance, three kinds of trapping indices need to be calculated (Nghiem et al., 2009a):

$$Residual\ Trapping\ Index\ (RTI) = \frac{Total\ mass\ of\ CO_2\ trapped\ as\ residual\ (ton)}{Total\ mass\ of\ CO_2\ injected\ (ton)} \quad (11)$$

$$Solubility\ Trapping\ Index\ (STI) = \frac{Total\ mass\ of\ CO_2\ dissolved\ in\ brine\ (ton)}{Total\ mass\ of\ CO_2\ injected\ (ton)} \quad (12)$$

$$Total\ Trapping\ Efficiency\ Index\ (TEI) = Residual\ Trapping\ Index\ (RTI) + \text{Solubility}\ Trapping\ Index\ (STI) \quad (13)$$

In order to investigate the uncertainty of reservoir heterogeneities, 101 geological realizations were considered during the CO_2 injection process. The Petrel package was automatically transferred to the compositional simulation module (CMG-GEM) to conduct the CO_2 sequestration process. Then, the simulation result was evaluated by the optimizer (CMOST-AI) before creating the following geological models and uncertainty variables. The new porosity and permeability models were generated in each new simulation job (Vo Thanh et al., 2020b). Figure 2 represents the process for integrating the Petrel into

CMOST to conduct simulations on different 101 geological realizations. Table 2 presents geological variables for creating realizations, consisting of azimuth, global seed number in the Petrel package. These variables would change the distribution of porosity and permeability for considering the geological uncertainty effect for machine learning models.

Table 2
The uncertainty limit of parameters for generation petrophysical realizations

Parameter	Lower bound	Base case	Upper bound
Azimuth Porosity	10	12	30
Azimuth Horizontal Permeability (Kh)	5	21	25
Azimuth Vertical Permeability (Kv)	30	36	60
Global seed number porosity	1000	1850	3000
Global seed number kh	650	1000	3000
Global seed number kv	20	10	50

Continuous CO_2 injection was applied for 10 years, followed by 490 years post-injection for this study. Figure 3 depicts the CO_2 saturation profiles for the 3D filter I-K directions after 10, 200, and 500 years. CO_2 is injected into the saline formation and rises to the reservoir's top due to the buoyancy mechanism (Nghiem et al., 2010).

Figure 3 shows the CO_2 saturation in the reservoir after the 10-year injection phase. The density of CO_2-saturated water increases because of CO_2 solubility in the aqueous state (Kim et al., 2017). The wet phase then migrates to the lower part of the saline aquifers as a CO_2 plume that activates different solubilities of CO_2 gas in the wet state. Here, the spread of the vaporous CO_2 developed in the top saline is more significant than that in the base supply because CO_2 moves vertically (Emad A. Al-Khdheeawi et al., 2017; Al-Khdheeawi et al., 2017) and is held by the seal rock. These simulation results can further explain the mechanism of CO_2 migration during saline formation.

Second step: Define the uncertainty variables. There are many uncertainties in deep saline aquifers because less observation data are available from these reservoirs due to limiting finances in CO_2 storage projects. Therefore, these uncertainties can be used to develop ML models to predict the CO_2 trapping efficiency during saline formation. The uncertainty parameters are listed in Table 1.

Third step: Latin hypercube sampling To generate the training dataset, 100 simulation experiments were performed using Latin hypercube sampling (LHS) utilizing CMOST, an artificial intelligence and ML tool in the CMG package (CMG, 2019).

This module can perform sensitivity assessment, history matching, optimization, and uncertainty analysis for simulation projects (CMG, 2019). The key point of this step is to employ LHS because it is

not dependent on the amount of training simulation jobs from the uncertainty parameters (Vo Thanh et al., 2020c).

Fourth step: Perform simulation jobs to gather inputs/outputs for the machine learning model. This procedure is vital for the development of ML models. The CMG-GEM module was used to perform 101 simulation experiments. For every simulation experiment, the uncertainty parameters and objective interests were gathered as the training dataset; residual trapping, solubility trapping, and cumulative CO_2 injection in a deep saline aquifer were the outcomes.

Furthermore, 4900 samples with elapsed time (10, 20, 30,…. 500 years) were prepared from 100 simulation experiments for training the ML models. In addition, 49 samples with the same elapsed time as the training stage from one simulation job was utilized to blind test ML models. These procedures are explained in detail in the final section.

Fifth step: Generate predictive models using Machine Learning Techniques The predictive models can estimate the relationship between the input variables and output functions depending on the training reservoir results. In addition, the developed predictive models were built for each objective interest. To demonstrate the robustness of developed predictive models, three popular and powerful supervised ML algorithms (GP, SVM, and RF) were employed in this study. The detail of ML techniques was introduced in Supplement Material. All the ML techniques were conducted in a MATLAB 2020b environment running on an Intel ® Core ™ i7 -8550U CPU with 16 GB RAM.

Regarding the performance criteria for each ML model, Stazio et al.(2019) suggested that the root mean square error (RMSE) and coefficient of determination (R^2) were used as two statistical indicators to assess the developed predictive models from ML techniques. These statistical indicators were computed using the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_{i,sam} - Y_{i,pred})^2}{\sum_{i=1}^{n}(Y_{i,sam} - \overline{Y}_{i,sam})^2} \qquad (14)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_{i,sam} - Y_{i,pred})^2} \qquad (15)$$

where, $Y_{i,sam}$ $Y_{i,pred}$ and $\overline{Y}_{i,sam}$ are the observed points of the reservoir simulation experiments, the outputs from the ML models, and the average of the simulation dataset, respectively.

Final step: Validation of predictive models. The 100 simulation jobs were used for calibration (training) and 10 fold cross-validation (Geisser, 1993). One simulation job was used for blind testing to evaluate the stability of the predictive models. Subsequently, the selected predictive models were employed in the field data from previous studies. This final step would ensure the application of predictive models in the actual

storage sites of the CO_2 trapping mechanism and other science disciplines. The MATLAB function of the developed training model for prediction in an existing reservoir is expressed as:

Result = trainedModel.predictFcn(X) where X is the data matrix.

3. Results

3.1. Generating Training Samples for Machine Learning Models

Figure 4 depicts the simulation performance of CO_2 residual trapping, CO_2 solubility trapping, and the amount of cumulative CO_2 injected. The time series of 101 simulation experiments is presented for three objective interests.

The CO_2 trapping index will be calculated to develop ML models from simulation results using equations (11), (12), and (13). First, the ML models (GPR, SVM, and RF) used 100 simulation experiments for training and 10-fold cross-validation. A single simulation experiment was then used to blind test the best machine model to verify the overfitting issue.

3.2. Developed Machine Learning Models

Figure 5 depicts the performance of the ML models for predicting the solubility trapping index(STI), residual trapping index (RTI), and total efficiency index (TEI). Generally, a higher correlation factor (R^2) and lower RMSE correspond to a higher predictive model accuracy. All the computed statistical indicators of the three ML models are listed in Table 3. According to the table, the three predictive models can generate a reasonable prediction for STI, RTI, and TEI. However, the performance of GPR models is markedly higher than that of other models developed utilizing RF, especially those applied to estimate the STI and RTI. The SVM model version is slightly less than the predictive GPR model performance but much higher than that of the EBT models. GPR models provide higher predictive accuracy than the SVM and RF models. Thus, GPR models are selected for predicting CO_2 trapping index, including RTI, STI, and TEI. The tuning hyper-parameter of three machine learning approaches for developing the

Table 3. *The statistical indicator of predicted and actual values for three different machine learning techniques*

Statistical Indicator	RTI		STI			TEI		
	GPR	SVM	GPR	SVM	RF	GPR	SVM	RF
R-squared	**0.999**	0.980	**0.999**	0.977	0.976	**0.990**	0.970	0.910
RMSE	**0.0031**	0.021	**0.007**	0.307	0.394	**0.014**	0.023	0.044

Although the GPR predictive models perform well in both indicators (RMSE and R^2), it is necessary to evaluate the selected GPR models using blind testing samples prior to the actual field applications. A total of 49 simulation time-elapsed CO_2 trapping indices were used to verify the developed predictive models. These blind testing samples were not utilized in training ML models. The blind data samples had a structure of 49 rows and nine columns.

Figure 6 highlights the blind testing results for RTI, STI, and TEI. The evaluated results fit well with the correlation factors 0.9306, 0.9562, and 0.9552 for RTI, STI, and TEI, respectively. Moreover, the RMSE of these trapping indices also reached the low values in **Figs. 6a, 6b, and 6c** for RTI, STI, and TEI, respectively. The blind testing results indicate that the predictive models based on GPR ML are stable and reasonable for further applications. From these blind testing results, the machine-predicted models would produce highly-accurate predicted results depending on the characteristics of geological data and other parameters. In this study, the field application of CO_2 sequestration is considered for validation from the development of predictive models. This validation will prove the practical applicability of the predictive models.

3.3. Application of the machine learning models by field application

We deployed predictive models created using GPR techniques in an actual field in the Cuu Long Basin, offshore Vietnam. The research area is located off southern Vietnam and consists of a fluvial Oligocene sandstone reservoir (Vo Thanh et al., 2019). Figure 7 illustrates the location and lithology of the actual field in offshore Vietnam.

The geological models were constructed based on geological and petrophysical data from Vo Thanh et al. (2020d). These models were used for simulating CO_2 trapping in this study. The field simulation performance was used to validate the selected GPR ML models by comparing the reservoir simulator and predictive models' outputs. Figure 8 shows the petrophysical model of the dynamic simulation data. This model consisted of 87,000 grid cells in the X, Y, and Z directions. The geological model was history-matched for dynamic simulation using an integrated modeling framework in a recent study (Vo Thanh and Sugai, 2021). Figure 9 presents the results of the history-matching models. The history-matching of reservoirs proved that the field simulated models could achieve good results for CO_2 trapping under actual conditions. The details of the simulation parameters are presented in Table 4.

Table 4
Reservoir characterization of the field study area

Variables	Values
Thickness (m)	100
Pressure at 2076 m depth (MPa)	20
Temperature (°C)	40
Vertical to horizontal permeability ratio	0.1
Salinity (ppm)	40000
Injection depth (m)	2157
Maximum bottom pressure (MPa)	32
Residual Gas Saturation	0.3
Porosity	0.025–0.258
Permeability mD	0.1–249

In this study, a total of one million tons of CO_2 was injected into the reservoir for 10 years, followed by a shut-in stage for 290 years of the monitoring period. The trapping index was then calculated to evaluate the effectiveness of the injected CO_2. Figure 10 illustrates the CO_2 saturation in the fluvial sandstone reservoir at the end of the simulation period.

As shown in this figure, CO_2 migrated into the sandstone reservoir through the sand channel distribution. The effect of facies distribution on CO_2 plume migration was explored by Vo Thanh et al. (2018). The present study found that the CO_2 flow followed the direction of the sand channel in the subsurface. After the completion of the simulation of CO_2 trapping at the field scale, the results of the trapping index of the simulated and GPR ML models would be compared for a better vision of our predictive models. The performance of the GPR models in the fluvial sandstone reservoir of Nam Vang is depicted in **Fig. 11**. The fluvial sandstone reservoir input parameters were considered to validate the GPR models using the R^2 and RMSE.

Figure 11 shows the validation results of the GPR models in the fluvial sandstone aquifer. A comparison of the predicted results and the simulation data for better evaluation was applied to the developed ML models. The GPR models achieved the prediction results with an R-squared value of 0.999 and an RMSE of 0.079 for STI. Regarding the RTI, the R-squared value was 0.9935, and the RMSE was 0.039. In TEI, the GPR models provided excellent performance matching between the target and actual values with an R-square of 0.992 and an RMSE of 0.084. The validation results showed that the GPR models were accurately predicted for 99.90%, 99.35%, and 99.2% of the STI, RTI, and TEI field simulated effects. Table 5 shows the results of the GPR models for the predicted CO_2 trapping efficiency in fluvial sandstone aquifers.

These results were compared to the trapping efficiency calculated using the field simulation data. As shown in **Fig. 11** and Table 5, these results suggest that predictive models based on the GPR ML technique can accurately predict CO_2 trapping efficiency in the field scale of deep saline formations. Also, the predicted objective storage sites should have a similar geological setting to developing ML models. To recap, the developed models will achieve different accuracy predicted results. In particular, the predicted samples are outside the range of creating models. The anticipated results may not be reliable for decision-making.

Table 5

The comparison between field-scale simulation and the GPR model in Nam Vang field, Cuu Long Basin.

	Reservoir Simulation			GPR model		
Year	RTI	STI	TEI	RTI	STI	TEI
30	0.364	0.131	0.494	0.218	0.180	0.398
40	0.408	0.134	0.541	0.289	0.185	0.474
50	0.436	0.136	0.571	0.330	0.189	0.519
60	0.457	0.137	0.594	0.367	0.194	0.561
70	0.473	0.139	0.611	0.400	0.198	0.598
80	0.486	0.140	0.626	0.429	0.202	0.631
90	0.497	0.142	0.639	0.455	0.205	0.660
100	0.507	0.143	0.650	0.476	0.209	0.685
110	0.516	0.144	0.660	0.495	0.212	0.707
120	0.521	0.145	0.667	0.510	0.216	0.725
130	0.526	0.147	0.673	0.522	0.219	0.741
140	0.530	0.148	0.678	0.531	0.222	0.753
150	0.532	0.149	0.681	0.538	0.225	0.763
160	0.536	0.150	0.686	0.542	0.228	0.769
170	0.538	0.151	0.689	0.543	0.231	0.774
180	0.541	0.152	0.692	0.547	0.234	0.781
190	0.542	0.153	0.694	0.549	0.236	0.785
200	0.543	0.154	0.696	0.548	0.239	0.787
210	0.543	0.155	0.698	0.548	0.242	0.790
220	0.544	0.156	0.699	0.556	0.245	0.800
230	0.544	0.157	0.701	0.556	0.247	0.803
240	0.544	0.158	0.702	0.556	0.250	0.806
250	0.545	0.159	0.703	0.559	0.253	0.811
260	0.545	0.160	0.704	0.560	0.255	0.815
270	0.545	0.161	0.705	0.561	0.258	0.819

	Reservoir Simulation			GPR model		
280	0.544	0.162	0.706	0.562	0.261	0.823
290	0.544	0.163	0.707	0.563	0.264	0.826
300	0.544	0.164	0.708	0.563	0.267	0.830

3.4. Comparison of proposed models and previous study

The field-scale simulation and predicted results demonstrated the excellent performance of the GPR ML models. A comparative analysis was conducted to evaluate the predictive performance of the currently developed models. Kim et al. (2017) used an ANN to predict CO_2 trapping efficiency in saline aquifers. Therefore, they completed a more comprehensive study to understand the developed models better.

The RMSE and R-squared indicators were used to evaluate the models by graphical analysis. Equations (14) and (15) explain the details of the calculation procedures for the two statistical indicators.

Figure 12a illustrates the comparative results of the correlation factor between the proposed GPR models and those of Kim et al. (2017). The figure indicates that the models developed in this study presented a higher correlation factor for three types of CO_2 trapping indices, RTI, STI, and TEI (0.9935, 0.999, and 0.992, respectively). The other ANN models had lower correlation factors for the RTI, STI, and TEI: 0.974, 0.978, and 0.964. The improved accuracy of the developed GPR models proved that these ML models were successfully trained.

Figure 12b shows another statistical indicator that demonstrates the robustness of the proposed GPR models. In the figure, the proposed ML models in this research can obtain smaller values for RMSE to evaluate the CO_2 trapping efficiency. The comparative results show that the developed ML models were better than the available developed predictive models.

4. Discussions

Our findings demonstrate the necessity of developing ML models to generate a robust predictive method for evaluating the CO_2 trapping performance in deep saline aquifers. We proved that the ML models could accurately predict CO_2 storage efficiency by comparing the simulation results for a fluvial aquifer in the Cuu Long Basin, offshore Vietnam. The developed models obtained excellent predictive performance because the uncertainty variables were carefully selected for training and validating the ML models. Also, the blind testing performance is crucial for validating the accuracy and stability of ML models. Numerous studies have developed ML methods that ignore the blind testing process. Therefore, a blind testing process should be included to build predictive models based on ML.

Notably, each developed ML model was considered in a specific range; the ML models are only used within the capacities of the uncertainty variables mentioned in this research. This study also emphasized saline aquifers for the application of the developed predictive models to saline formation. However, the workflow of this research can be easily adapted to other storage formations. The uncertainty parameters can be modified according to specific storage formations.

Besides, the range of uncertainty variables used in developing predictive models should be carefully considered. The specific training samples of ML models should be obtained from literature reviews or previous studies. ML predictive models cannot provide excellent performance if the models are employed on the samples inadequate for the proposed uncertainty bounds.

Our study found that GPR created excellent CO_2 trapping efficiency prediction compared with other ML methods such as SVM and RF. In terms of the correlation coefficient (R^2), the GPR models could predict the CO_2 trapping index in actual field applications with higher R^2 values. We subjected GPR models to blind testing before employing them in the field scale of saline aquifers. Previous studies have not focused on blind testing for the development of predictive models.

In conclusion, our work presents an innovative flowchart for “construction,” a robust and highly accurate ML predictive model. To validate the proposed method, it is suggested hitherto for other CO_2 geo-storage “formations” such as oil and gas reservoirs, ROZs, and fractured reservoirs. The limitations of ML models apply to other geological formations. However, the methodology of this study can be easily developed to address further engineering and science problems.

5. Conclusions

This study used ML models to predict the CO_2 trapping index of deep saline aquifers. It found applicable ML models for predicting RTI, STI, and TEI in the Cuu Long Basin reservoir, offshore Vietnam. The highlights of this study are as follows.

- Reservoir simulation of the CO_2 trapping mechanism in deep saline formation was performed to create a training database to develop predictive ML models. In this study, 101 simulation experiments were completed to gather objective interests, including residual trapping, solubility trapping, and cumulative CO_2 injection in the trapping reservoir model, for the training database.
- Predictive results confirmed that GPR produced the best predictive performance compared to SVM and RF. Also, the predictive ML models developed by RF achieved the worst predictive accuracy. The GPR model was identified as a more suitable method for creating predictive models for CO_2 trapping efficiency in deep saline aquifers. This study used only 100 numerical simulation jobs with 49 elapsed times to develop highly accurate predictive models.
- With regard to actual field application, GPR ML models were adopted for a fluvial sandstone reservoir in offshore Vietnam. The findings of this study indicated that the GPR models could accurately

predict CO_2 trapping in an actual field. By comparing the trapping efficiency, the predictive performances obtained excellent fitting with objective values and accuracy between the GPR models; a field-scale simulation result was achieved with a high correlation factor (R^2) of more than 0.95, and a low RMSE of 0.016 for TEI. To the best of our knowledge, the CO_2 trapping efficiency has not yet been explored in deep saline aquifers using predictive models based on ML.

- The developed GPR ML models can predict the CO_2 trapping index with high accuracy in deep saline aquifers. Our study recommends that the GPR models be reproduced and adapted in future carbon capture utilization and storage (CCUS) tasks, such as oil recovery prediction and storage performance in the CO_2-EOR project. The proposed ML predictive workflow will also serve as a reference for future work, especially the machine techniques that could be coupled with a reservoir simulator to reduce computational time.
- The proposed ML models can be coupled with a commercial reservoir simulator to improve their ability and accuracy for predicting CO_2 trapping at storage sites.
- The proposed ML models demonstrated robustness compared to a previous study. This was inferred using R^2 and RMSE analysis. The ML models offered in this study have a larger R^2 and smaller RMSE in three prediction trapping efficiencies (RTI, STI, and TEI). Besides, this study demonstrated the power of the GPR models for predictive modeling. The GPR models had a better predictive performance than the ANN models.
- Predictive models based on ML can be applied to tackle climate change by predicting CO_2 trapping efficiency deep underground.

However, there are some limitations in developing predictive models:

- The developed models could only be employed in deep saline aquifers and geological characteristics similar to the simulation model design expressed in this study.
- The predictive ML models were applied within the bounds of the uncertainty factors identified in this study.

Declarations

Acknowledgments

The authors would like to express their gratitude to the BrainKorea21 FOUR Postdoctoral Fellowship for financial support. Field application study was partly supported by the National Research Foundation of Korea (NRF) grand (No. 2017R1A2B3002119) through K.-K. Lee.

Competing Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Abbaszadeh, M., Shariatipour, S., 2018. Investigating the Impact of Reservoir Properties and Injection Parameters on Carbon Dioxide Dissolution in Saline Aquifers. Fluids 3, 76. https://doi.org/10.3390/fluids3040076
2. Afanasyev, A., Kempka, T., Kühn, M., Melnik, O., 2016. Validation of the MUFITS Reservoir Simulator Against Standard CO2 Storage Benchmarks and History-matched Models of the Ketzin Pilot Site. Energy Procedia 97, 395–402. https://doi.org/10.1016/j.egypro.2016.10.032
3. Ahmadi, M.A., Ahmadi, A., 2016. Applying a sophisticated approach to predict CO2 solubility in brines: Application to CO2 sequestration. Int. J. Low-Carbon Technol. 11, 325–332. https://doi.org/10.1093/ijlct/ctu034
4. Ahmadi, M.A., Zendehboudi, S., James, L.A., 2018. Developing a robust proxy model of CO2 injection: Coupling Box–Behnken design and a connectionist method. Fuel 215, 904–914. https://doi.org/10.1016/j.fuel.2017.11.030
5. Ajayi, T., Awolayo, A., Gomes, J.S., Parra, H., Hu, J., 2019. Large scale modeling and assessment of the feasibility of CO2 storage onshore Abu Dhabi. Energy 185, 653–670. https://doi.org/10.1016/j.energy.2019.07.052
6. Al-Khdheeawi, E.A., Vialle, S., Barifcani, A., Sarmadivaleh, M., Iglauer, S., 2018a. Effect of wettability heterogeneity and reservoir temperature on CO2 storage efficiency in deep saline aquifers. Int. J. Greenh. Gas Control 68, 216–229. https://doi.org/10.1016/j.ijggc.2017.11.016
7. Al-Khdheeawi, E.A., Vialle, S., Barifcani, A., Sarmadivaleh, M., Iglauer, S., 2018b. Enhancement of CO2 trapping efficiency in heterogeneous reservoirs by water-alternating gas injection. Greenh. Gases Sci. Technol. 12, 1–12. https://doi.org/10.1002/ghg.1805
8. Al-Khdheeawi, Emad A., Vialle, S., Barifcani, A., Sarmadivaleh, M., Iglauer, S., 2017. Effect of brine salinity on CO2 plume migration and trapping capacity in deep saline aquifers. APPEA J. 57, 100. https://doi.org/10.1071/aj16248
9. Al-Khdheeawi, Emad A, Vialle, S., Barifcani, A., Sarmadivaleh, M., Iglauer, S., 2017. Effect of brine salinity on CO2 plume migration and trapping capacity in deep saline aquifers. APPEA J. 57, 100–109.
10. Al-mudhafar, W.J., 2019. Polynomial and nonparametric regressions for efficient predictive proxy metamodeling: Application through the CO2-EOR in shale oil reservoirs. J. Nat. Gas Sci. Eng. 72, 103038. https://doi.org/10.1016/j.jngse.2019.103038
11. Al-Mudhafar, W.J., 2020. Integrating machine learning and data analytics for geostatistical characterization of clastic reservoirs. J. Pet. Sci. Eng. 195, 107837. https://doi.org/10.1016/j.petrol.2020.107837
12. Al-Mudhafar, W.J., 2017. Integrating well log interpretations for lithofacies classification and permeability modeling through advanced machine learning algorithms. J. Pet. Explor. Prod. Technol. 7, 1023–1033. https://doi.org/10.1007/s13202-017-0360-0
13. Al-Khdheeawi, E.A., Vialle, S., Barifcani, A., Sarmadivaleh, M., Iglauer, S., 2017. Influence of CO2-wettability on CO2 migration and trapping capacity in deep saline aquifers. Greenh. Gases Sci.

Technol. 7, 352–368. https://doi.org/10.1002/ghg

14. Alcalde, J., Flude, S., Wilkinson, M., Johnson, G., Edlmann, K., Bond, C.E., Scott, V., Gilfillan, S.M.V., Ogaya, X., Stuart Haszeldine, R., 2018. Estimating geological CO2 storage security to deliver on climate mitigation. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-04423-1
15. Bachu, S., 2018. Identification of oil reservoirs suitable for CO2-EOR and CO2 storage (CCUS) using reserves databases, with application to Alberta, Canada. Int. J. Greenh. Gas Control 44, 152–165. https://doi.org/10.1016/j.ijggc.2015.11.013
16. Bachu, S., 2008. CO2storage in geological media: Role, means, status and barriers to deployment. Prog. Energy Combust. Sci. 34, 254–273. https://doi.org/10.1016/j.pecs.2007.10.001
17. 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 Convers. Manag. 43, 87–102. https://doi.org/10.1016/S0196-8904(01)00009-7
18. Bachu, S., 2000. Sequestration of CO2 in geological media: Criteria and approach for site selection in response to climate change. Energy Convers. Manag. 41, 953–970. https://doi.org/10.1016/S0196-8904(99)00149-1
19. Beni, A.N., Clauser, C., Erlstrom, M., 2011. System analysis of underground CO2 storage by numerical modeling for a case study in Malmo. Am. J. Sci. 311, 335–368. https://doi.org/10.2475/04.2011.03
20. Bu Ali, A.T., Alali, M., Alzaid, M., Bunch, M., Menacherry, S., 2011. Numerical simulation model of geological storage of carbon dioxide in onshore Gippsland Basin, Victoria. Soc. Pet. Eng. – 9th Eur. Form. Damage Conf. 2011 2, 796–812. https://doi.org/10.2118/144075-ms
21. Chen, B., Pawar, R., 2018. Capacity assessment of CO 2 storage and enhanced oil recovery in residual oil zones. J. Pet. Sci. Eng. 2018-Septe, 106342. https://doi.org/10.1016/j.petrol.2019.106342
22. Chen, B., Pawar, R.J., 2019. Characterization of CO2 storage and enhanced oil recovery in residual oil zones. Energy 183, 291–304. https://doi.org/10.1016/j.energy.2019.06.142
23. CMG, 2019. Manual of Computer Modelling Group's Software. Calgary, Cananda.
24. Dai, Z., Middleton, R., Viswanathan, H., Fessenden-Rahn, J., Bauman, J., Pawar, R., Lee, S.-Y., McPherson, B., 2014. An Integrated Framework for Optimizing CO_2 Sequestration and Enhanced Oil Recovery. Environ. Sci. Technol. Lett. 1, 49–54. https://doi.org/10.1021/ez4001033
25. Dai, Z., Viswanathan, H., Middleton, R., Pan, F., Ampomah, W., Yang, C., Jia, W., Xiao, T., Lee, S.Y., McPherson, B., Balch, R., Grigg, R., White, M., 2016. CO2 Accounting and Risk Analysis for CO2 Sequestration at Enhanced Oil Recovery Sites. Environ. Sci. Technol. 50, 7546–7554. https://doi.org/10.1021/acs.est.6b01744
26. Dai, Z., Zhang, Y., Bielicki, J., Amooie, M.A., Zhang, M., Yang, C., Zou, Y., Ampomah, W., Xiao, T., Jia, W., Middleton, R., Zhang, W., Sun, Y., Moortgat, J., Soltanian, M.R., Stauffer, P., 2018. Heterogeneity-assisted carbon dioxide storage in marine sediments. Appl. Energy 225, 876–883. https://doi.org/10.1016/j.apenergy.2018.05.038

27. Dai, Z., Zhang, Y., Stauffer, P., Xiao, T., Zhang, M., Ampomah, W., Yang, C., Zhou, Y., Ding, M., Middleton, R., Soltanian, M.R., Bielicki, J.M., 2017. Injectivity Evaluation for Offshore CO2Sequestration in Marine Sediments. Energy Procedia 114, 2921–2932. https://doi.org/10.1016/j.egypro.2017.03.1420
28. Edwards, R.W.J., Celia, M.A., Bandilla, K.W., Doster, F., Kanno, C.M., 2015. A Model To Estimate Carbon Dioxide Injectivity and Storage Capacity for Geological Sequestration in Shale Gas Wells. Environ. Sci. Technol. 49, 9222–9229. https://doi.org/10.1021/acs.est.5b01982
29. Ertekin, T., Sun, Q., 2019. Artificial Intelligence Applications in Reservoir Engineering: A Status Check. Energies 12, 2897. https://doi.org/10.3390/en12152897
30. Ezeanyim, P.O., Shariatipour, S.M., 2016. Numerical and optimisation studies of CO2 dissolution and trapping efficiency in saline aquifers. 78th EAGE Conf. Exhib. 2016 Effic. Use Technol. - Unlocking Potential. https://doi.org/10.3997/2214-4609.201601075
31. Fang, Y., Baojun, B., Dazhen, T., Dunn-norman, S., Wronkiewicz, D., 2010. Characteristics of CO 2 sequestration in saline aquifers. Pet. Sci. 7, 83–92. https://doi.org/10.1007/s12182-010-0010-3
32. Geisser, S., 1993. Predictive Inference: An Introduction. CRC Press.
33. Gibson-Poole, C.M., Svendsen, L., Underschul-, J.T., Ennis-King, J., Ruth, P.J. van, Nelson, E.J., Daniel, R.F., Cinar, Y., 2006. Gippsland Basin Geosequestration: a Potential Solution for the Latrobe Valley GIPPSLAND BASIN GEOSeQUESTRATION. APPEA J. https://doi.org/10.1071/AJ05024
34. Gu, Y., Oliver, D.S., Oklahoma, U., 2005. History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter 217–224.
35. Hadi, S., Zargar, G., Baharimoghadam, M., 2016. Fractured reservoir history matching improved based on artificial intelligent 2, 344–360. https://doi.org/10.1016/j.petlm.2016.09.001
36. He, J., Xie, J., Wen, X.H., Chen, W., 2016. An alternative proxy for history matching using proxy-for-data approach and reduced order modeling. J. Pet. Sci. Eng. 146, 392–399. https://doi.org/10.1016/j.petrol.2016.05.026
37. Hitchon, B., Gunter, W.D., Gentzis, T., Bailey, R.T., 1999. Sedimentary basins and greenhouse gases: A serendipitous association. Energy Convers. Manag. 40, 825–843. https://doi.org/10.1016/S0196-8904(98)00146-0
38. House, N.J., Faulder, D.D., Olson, G.L., Fanchi, J.R., 2003. Simulation Study of CO2 Sequestration in a North Sea Formation, in: SPE/EPA/DOE Exploration and Production Environmental Conference. San Antonio, Texas, U.S.A., 10–12 March 2003, pp. 1–7.
39. Issautier, B.Î., Fillacier, S., Le Gallo, Y., Audigane, P., Chiaberge, C., Viseur, S., 2013. Modelling of CO2 injection in fluvial sedimentary heterogeneous reservoirs to assess the impact of geological heterogeneities on CO2 storage capacity and performance. Energy Procedia 37, 5181–5190. https://doi.org/10.1016/j.egypro.2013.06.434
40. Javaheri, M., Abedi, J., Hassanzadeh, H., 2009. Onset of Convection in CO 2 Sequestration in Deep Inclined Saline Aquifers. J. Can. Pet. Technol. 48.

41. Jia, W., McPherson, B., Pan, F., Dai, Z., Xiao, T., 2018. Uncertainty quantification of CO2storage using Bayesian model averaging and polynomial chaos expansion. Int. J. Greenh. Gas Control 71, 104–115. https://doi.org/10.1016/j.ijggc.2018.02.015
42. Jin, M., Pickup, G., Mackay, E., Todd, A., Sohrabi, M., Monaghan, A., Naylor, M., 2012. Static and Dynamic Estimates of CO2-Storage Capacity in Two Saline Formations in the UK. SPE J. 17, 1108–1118. https://doi.org/10.2118/131609-PA
43. Kempka, T., Klein, E., De Lucia, M., Tillner, E., Kühn, M., 2013. Assessment of long-term CO2trapping mechanisms at the Ketzin pilot site (Germany) by coupled numerical modelling. Energy Procedia 37, 5419–5426. https://doi.org/10.1016/j.egypro.2013.06.460
44. Khudaida, K.J., Das, D.B., 2020. A Numerical Analysis of the Effects of Supercritical CO2 Injection on CO2 Storage Capacities of Geological Formations. Clean Technol. 2, 333–364. https://doi.org/10.3390/cleantechnol2030021
45. Kim, M., Shin, H., 2020. Machine learning-based prediction of the shale barrier size and spatial location using key features of SAGD production curves. J. Pet. Sci. Eng. 191, 107205. https://doi.org/10.1016/j.petrol.2020.107205
46. Kim, Y., Jang, H., Kim, J., Lee, J., 2017. Prediction of storage efficiency on CO2 sequestration in deep saline aquifers using artificial neural network. Appl. Energy 185, 916–928. https://doi.org/10.1016/j.apenergy.2016.10.012
47. Land, C.S., 1968. Calculation of Imbibition Relative Permeability for Two- and Three-Phase Flow From Rock Properties. Soc. Pet. Eng. J. 8, 149–156. https://doi.org/10.2118/1942-pa
48. Lee, J.H., Park, Y.C., Sung, W.M., Lee, Y.S., Park, Y.C., Sung, W.M., Simulation, Y.S.L.A., 2010. Environmental Effects A Simulation of a Trap Mechanism for the Sequestration of CO 2 into Gorae V Aquifer, Korea A Simulation of a Trap Mechanism for the Sequestration of CO 2 into Gorae V Aquifer, Korea 7036. https://doi.org/10.1080/15567030903436822
49. Lee, K.J., 2020. Data-Driven models to predict hydrocarbon production from unconventional reservoirs by thermal recovery. J. Energy Resour. Technol. Trans. ASME 142, 1–17. https://doi.org/10.1115/1.4047309
50. Lengler, U., De Lucia, M., Kühn, M., 2010. The impact of heterogeneity on the distribution of CO2: Numerical simulation of CO2storage at Ketzin. Int. J. Greenh. Gas Control 4, 1016–1025. https://doi.org/10.1016/j.ijggc.2010.07.004
51. Li, S., Zhang, Y., 2014. Model complexity in carbon sequestration: A design of experiment and response surface uncertainty analysis. Int. J. Greenh. Gas Control 22, 123–138. https://doi.org/10.1016/j.ijggc.2013.12.007
52. Liu, B., Zhang, Y., 2011. CO2 modeling in a deep saline aquifer: A predictive uncertainty analysis using design of experiment. Environ. Sci. Technol. 45, 3504–3510. https://doi.org/10.1021/es103187b
53. Lothe, A.E., Emmel, B., Grøver, A., Bergmo, P.E., 2014. CO2 storage modelling and capacity estimation for the Trøndelag Platform, offshore Norway - using a basin modelling approach. Energy Procedia

63, 3648–3657. https://doi.org/10.1016/j.egypro.2014.11.394

54. Mediato, J.F., García-Crespo, J., Izquierdo, E., García-Lobón, J.L., Ayala, C., Pueyo, E.L., Molinero, R., 2017. Three-dimensional Reconstruction of the Caspe Geological Structure (Spain) for Evaluation as a Potential CO2Storage Site. Energy Procedia 114, 4486–4493. https://doi.org/10.1016/j.egypro.2017.03.1608
55. Miocic, J.M., Gilfillan, S.M.V., Frank, N., Schroeder-Ritzrau, A., Burnside, N.M., Haszeldine, R.S., 2019. 420,000 Year Assessment of Fault Leakage Rates Shows Geological Carbon Storage Is Secure. Sci. Rep. 9, 1–9. https://doi.org/10.1038/s41598-018-36974-0
56. Mohaghegh, S.D., 2011. Reservoir simulation and modeling based on pattern recognition, in: SPE Digital Energy Conference and Exhibition 201. The Woodlands, Texas, USA, 19–21 April. https://doi.org/10.2118/143179-ms
57. Nghiem, L., Sammon, P., Grabenstetter, J., Ohkuma, H., 2004. Current Models for Transport of Reactive Multicomponent Systems.
58. Nghiem, L., Shrivastava, V., Kohse, B., Hassam, M., Yang, C., 2010. Simulation and optimization of trapping processes for CO2 storage in saline aquifers. J. Can. Pet. Technol. 49, 15–22. https://doi.org/10.2118/139429-PA
59. Nghiem, L., Yang, C., Shrivastava, V., Kohse, B., Hassam, M., Card, C., 2009a. Risk mitigation through the optimization of residual gas and solubility trapping for CO2storage in saline aquifers. Energy Procedia 1, 3015–3022. https://doi.org/10.1016/j.egypro.2009.02.079
60. Nghiem, L., Yang, C., Shrivatava, V., Kohse, B., Hassam, M., Chen, D., Card, C., 2009b. Optimization of Residual Gas and Solubility Trapping for CO 2 Storage in Saline Aquifers. Soc. Pet. Eng. 1–9. https://doi.org/10.1016/j.egypro.2009.02.079
61. Nguyen, N., Le, H., Sugai, Y., Sasaki, K., 2020. Investigation of Stability of CO 2 Microbubbles — Colloidal Gas Aphrons for Enhanced Oil Recovery Using Definitive Screening Design. Colloids and interfaces 4.
62. Petvipusit, K.R., Elsheikh, A.H., Laforce, T.C., King, P.R., Blunt, M.J., 2014. Robust optimisation of CO2 sequestration strategies under geological uncertainty using adaptive sparse grid surrogates. Comput. Geosci. 18, 763–778. https://doi.org/10.1007/s10596-014-9425-z
63. Razak, S.M., Jafarpour, B., 2019. Supervised Machine-Learning for History Matching: Learning the Inverse Mapping from Low-Rank Data and Model Representations, in: 81st EAGE Conference & Exhibition. 3–6 June 2019, London, UK, pp. 1–5.
64. Ren, B., Duncan, I.J., 2019. Reservoir simulation of carbon storage associated with CO2 EOR in residual oil zones, San Andres formation of West Texas, Permian Basin, USA. Energy 167, 391–401. https://doi.org/10.1016/j.energy.2018.11.007
65. Senel, O., Chugunov, N., 2012. CO2 injection in a saline formation: How do additional data change uncertainties in our reservoir simulation predictions? Carbon Manag. Technol. Conf. 2012, Febr. 7, 2012 - Febr. 9 2, 778–788.

66. Shogenov, K., Forlin, E., Shogenova, A., 2017. 3D Geological and Petrophysical Numerical Models of E6 Structure for CO2 Storage in the Baltic Sea. Energy Procedia 114, 3564–3571. https://doi.org/10.1016/j.egypro.2017.03.1486
67. Silva, P.N.K. De, Ranjith, P.G., 2012. Review article A study of methodologies for CO 2 storage capacity estimation of saline aquifers. Fuel 93, 13–27. https://doi.org/10.1016/j.fuel.2011.07.004
68. Singh, V., Asa, S., Cavanagh, A., Hansen, H., Nazarian, B., Iding, M., Ringrose, P., Asa, S., 2010. Reservoir Modeling of CO 2 Plume Behavior Calibrated Against Monitoring Data From Sleipner, Norway, in: SPE Annual Technical Conference and Exhibition. Florence, Italy, 19–22 September 2010.
69. Soltanian, M.R., Amooie, M.A., Gershenzon, N., Dai, Z., Ritzi, R., Xiong, F., Cole, D., Moortgat, J., 2017. Dissolution Trapping of Carbon Dioxide in Heterogeneous Aquifers. Environ. Sci. Technol. 51, 7732–7741. https://doi.org/10.1021/acs.est.7b01540
70. 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. Int. J. Greenh. Gas Control 98, 103042. https://doi.org/10.1016/j.ijggc.2020.103042
71. Stazio, A., Victores, J.G., Estevez, D., Balaguer, C., 2019. A study on machine vision techniques for the inspection of health personnels' protective suits for the treatment of patients in extreme isolation. Electron. 8. https://doi.org/10.3390/electronics8070743
72. Sung, R.T., Li, M.H., Dong, J.J., Lin, A.T.S., Hsu, S.K., Wang, C.Y., Yang, C.N., 2014. Numerical assessment of CO2 geological sequestration in sloping and layered heterogeneous formations: A case study from Taiwan. Int. J. Greenh. Gas Control 20, 168–179. https://doi.org/10.1016/j.ijggc.2013.11.003
73. Temitope, A., Gomes, J.S., Al Kobaisi, M., Hu, J., 2016a. Characterization and quantification of the CO2 sequestration potential of a carbonate aquifer in Falaha syncline, onshore Abu Dhabi, in: Abu Dhabi International Petroleum Exhibition and Conference. Dhabi, UAE, 7–10 November. https://doi.org/10.2118/183442-ms
74. Temitope, A., Sulemana, N., Gomes, J.S., Oppong, R., Russian, G., 2016b. Statistical Uncertainty Analysis and Design Optimization of CO2 Trapping Mechanisms in Saline Aquifers, in: International Petroleum Technology Conference.
75. Vo Thanh, H., Sugai, Y., 2021. Integrated modelling framework for enhancement history matching in fluvial channel sandstone reservoirs. Upstream Oil Gas Technol. 100027. https://doi.org/10.1016/j.upstre.2020.100027
76. Vo Thanh, H., Sugai, Y., Nguele, R., Sasaki, K., 2020a. Robust optimization of CO2 Sequestration through a water alternating gas process under geological uncertainties in Cuu Long Basin, Vietnam. J. Nat. Gas Sci. Eng. https://doi.org/10.1016/j.jngse.2020.103208
77. Vo Thanh, H., Sugai, Y., Nguele, R., Sasaki, K., 2020b. Robust optimization of CO2 sequestration through a water alternating gas process under geological uncertainties in Cuu Long Basin, Vietnam. J. Nat. Gas Sci. Eng. 76. https://doi.org/10.1016/j.jngse.2020.103208

78. Vo Thanh, H., Sugai, Y., Nguele, R., Sasaki, K., 2019. Integrated work flow 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 90, 102826. https://doi.org/10.1016/j.ijggc.2019.102826
79. Vo Thanh, H., Sugai, Y., Sasaki, K., 2020c. Application of artificial neural network for predicting the performance of CO2 enhanced oil recovery and storage in residual oil zones. Sci. Rep. 10, 18204. https://doi.org/10.1038/s41598-020-73931-2
80. Vo Thanh, H., Sugai, Y., Sasaki, K., 2020d. Impact of a new geological modelling method on the enhancement of the CO2 storage assessment of E sequence of Nam Vang field, offshore Vietnam. Energy Sources, Part A Recover. Util. Environ. Eff. 42, 1499–1512. https://doi.org/10.1080/15567036.2019.1604865
81. Vo Thanh, H., Sugai, Y., Sasaki, K., 2018. AN OBJECT-BASE MODELING AND SIMULATION OF CO 2 PLUME DYNAMIC IN SALINE FORMATION IN NAM VANG FIELD, CUU LONG, in: 24th Formation Evaluation Symposium of Japan. Chiba, Japan.
82. Wang, Z., Wang, J., Lan, C., He, I., Ko, V., Ryan, D., Wigston, A., 2016. A study on the impact of SO2 on CO2 injectivity for CO2 storage in a Canadian saline aquifer. Appl. Energy 184, 329–336. https://doi.org/10.1016/j.apenergy.2016.09.067
83. Xiao, T., McPherson, B., Esser, R., Jia, W., Moodie, N., Chu, S., Lee, S.Y., 2019. Forecasting commercial-scale CO 2 storage capacity in deep saline reservoirs: Case study of Buzzard's bench, Central Utah. Comput. Geosci. 126, 41–51. https://doi.org/10.1016/j.cageo.2018.12.006
84. Yan, H., Zhang, J., Rahman, S.S., Zhou, N., Suo, Y., 2020. Predicting permeability changes with injecting CO 2 in coal seams during CO 2 geological sequestration: A comparative study among six SVM-based hybrid models. Sci. Total Environ. 705, 135941. https://doi.org/10.1016/j.scitotenv.2019.135941
85. 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. J. Clean. Prod. 260, 120866. https://doi.org/10.1016/j.jclepro.2020.120866
86. Yu, H., Wang, Z., Rezaee, R., Zhang, Y., Xiao, L., Academy, C., Wang, X., Zhang, L., 2016. The Gaussian Process Regression for TOC Estimation Using Wireline Logs in Shale Gas Reservoirs.
87. Zapata, Y., Kristensen, M.R., Huerta, N., Brown, C., Kabir, C.S., Reza, Z., 2020. CO2 geological storage: Critical insights on plume dynamics and storage efficiency during long-term injection and post-injection periods. J. Nat. Gas Sci. Eng. 83, 103542. https://doi.org/10.1016/j.jngse.2020.103542
88. Zhong, M., Hughes, B., Schuetter, J., Mishra, S., Memorial, B., Lafollette, R.F., Hughes, B., 2015. Do Data Mining Methods Matter ?: A Wolfcamp " Shale " Case Study, in: SPE Hydraulic Fracturing Technology Conference. Woodlands, Texas, USA, 3–5 February 2015. This.

Figures

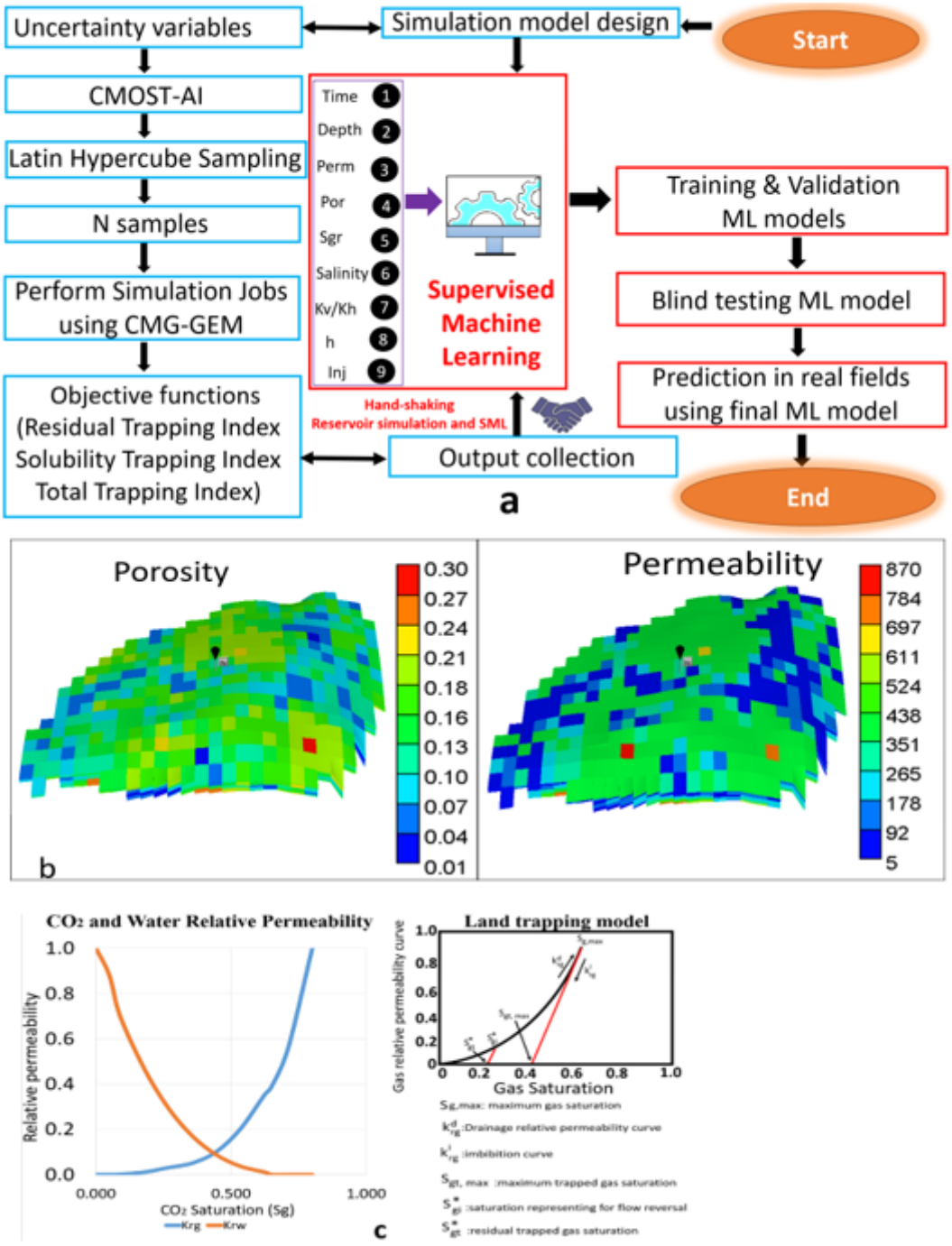

Figure 1

Flow chart of predictive models based machine learning workflow, geological model, and relative permeability curves. a. The workflow connected between the reservoir simulator and machine learning models b. The porosity model and permeability from PUNQS3 project was adapted for this work c. The relative permeability curves and Land trapping models used for CO2 trapping simulation.

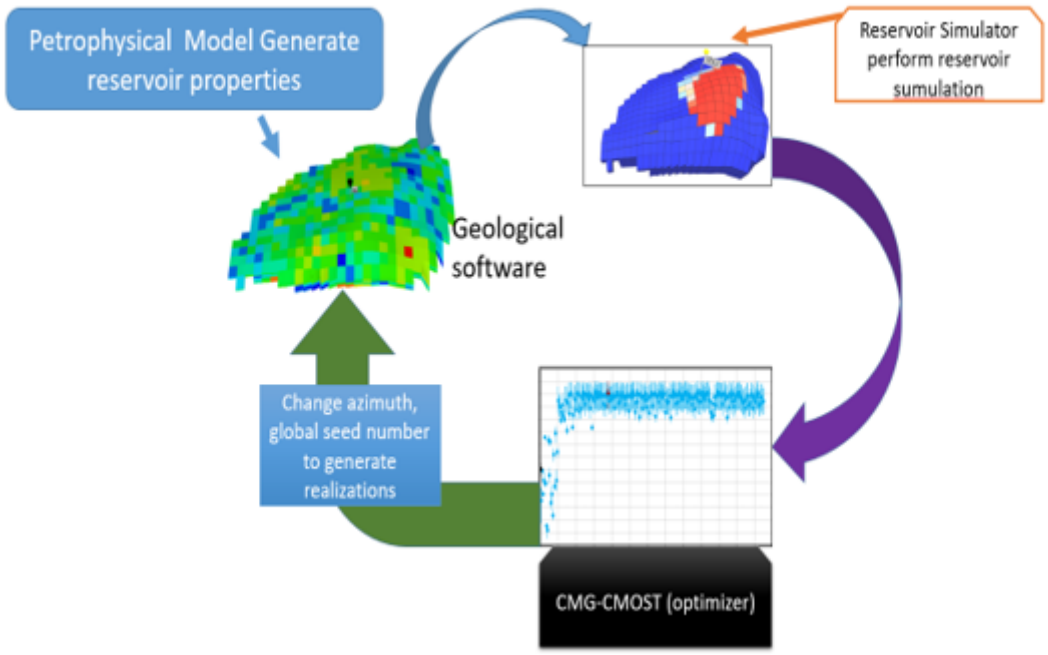

Figure 2

The integrated process Petrel and CMOST optimizer for considering geological realizations to generate the training samples

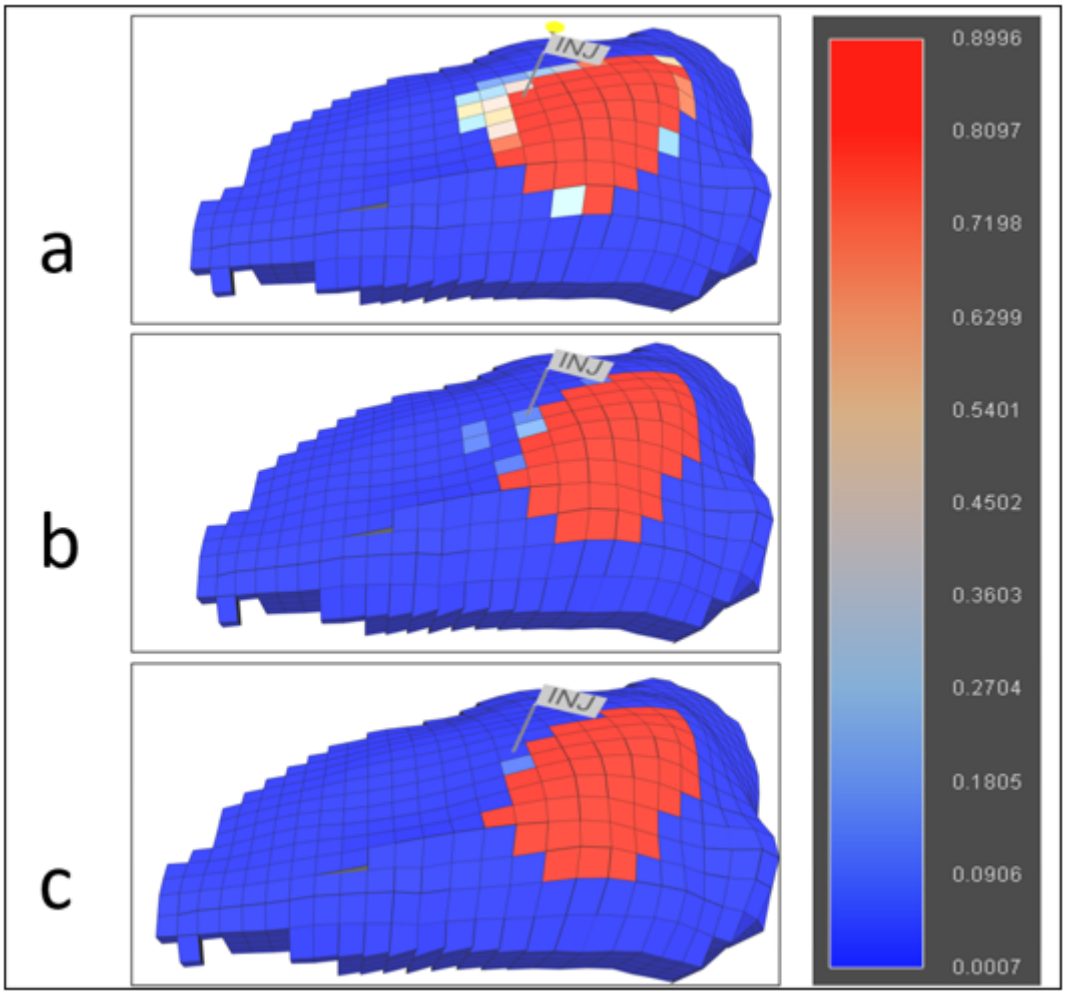

Figure 3

CO2 saturation over time. a) CO2 saturation after 10 years injection; b) CO2 saturation after 200 years injection, CO2 saturation after 500 years injections

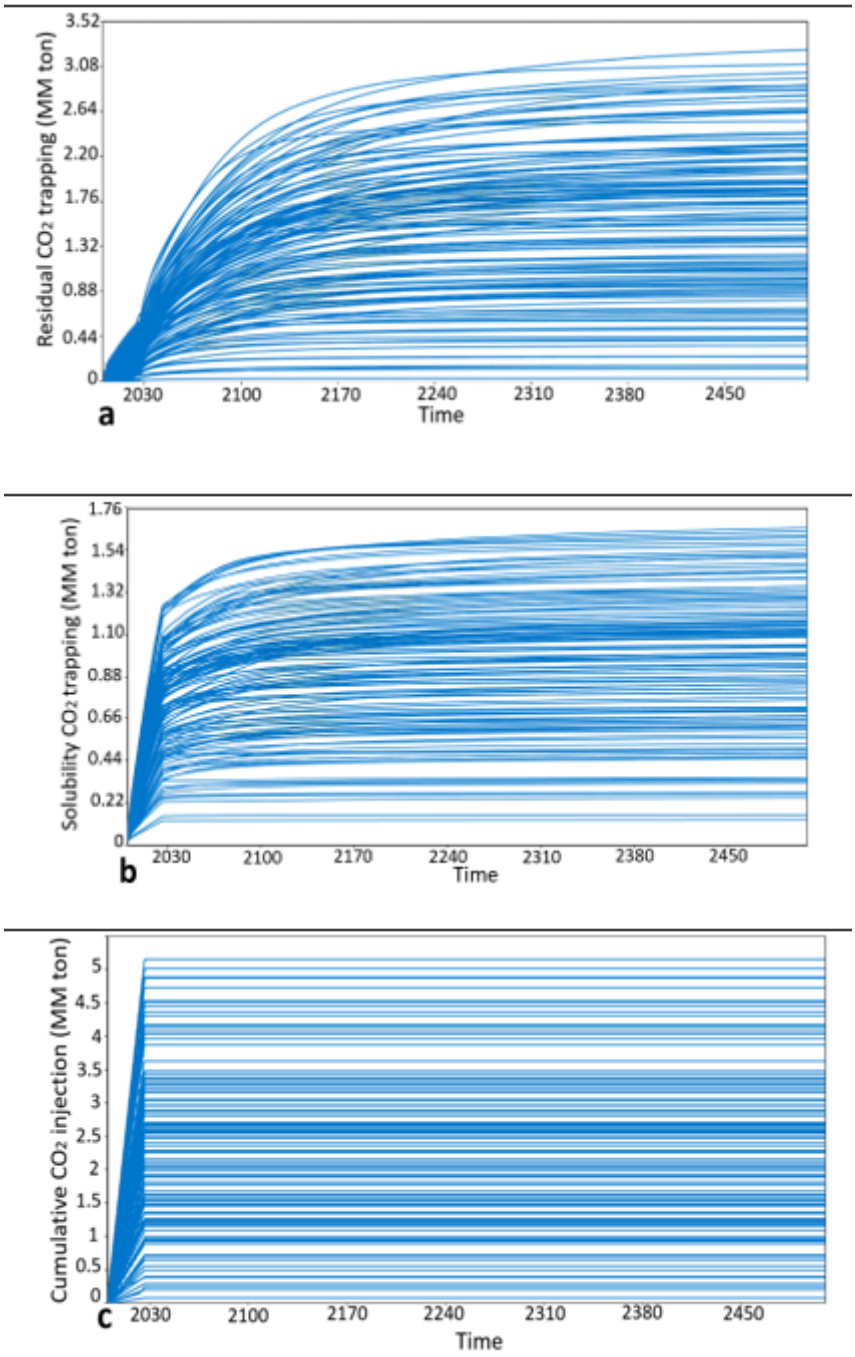

Figure 4

Time series of CO2 injection in the reservoir simulation jobs produced for developing machine learning models. a. Time series of residual CO2 trapping. b. Time series of solubility CO2 trapping. c. Time series of cumulative CO2 trapping

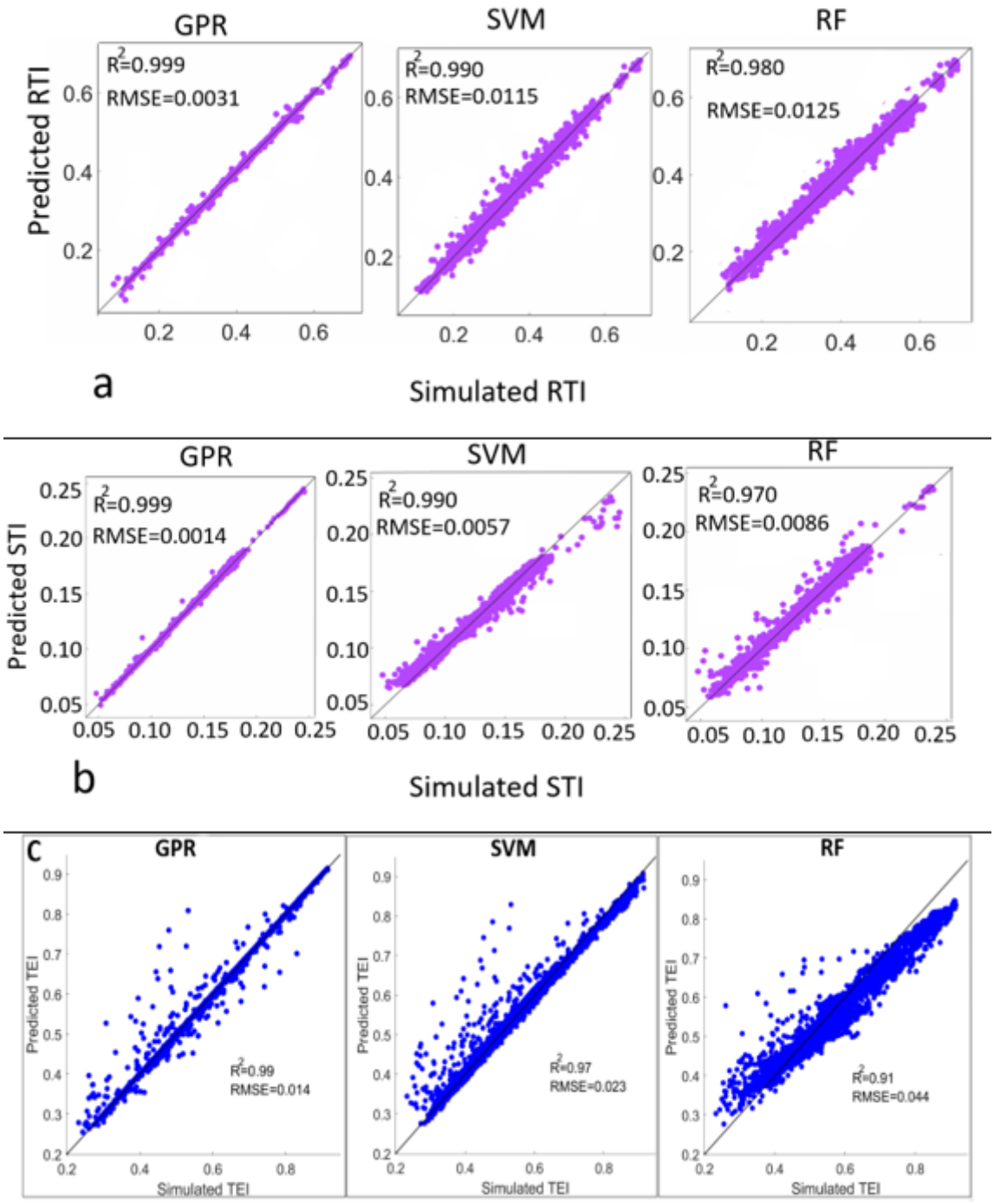

Figure 5

Performance of predictive models obtained from GPR, SVM, and RF. a. Quality of predictive RTI using three machine learning models; b. quality of predictive STI from three machine learning models; c. performance of predictive TEI from GPR, SVM, and RF.

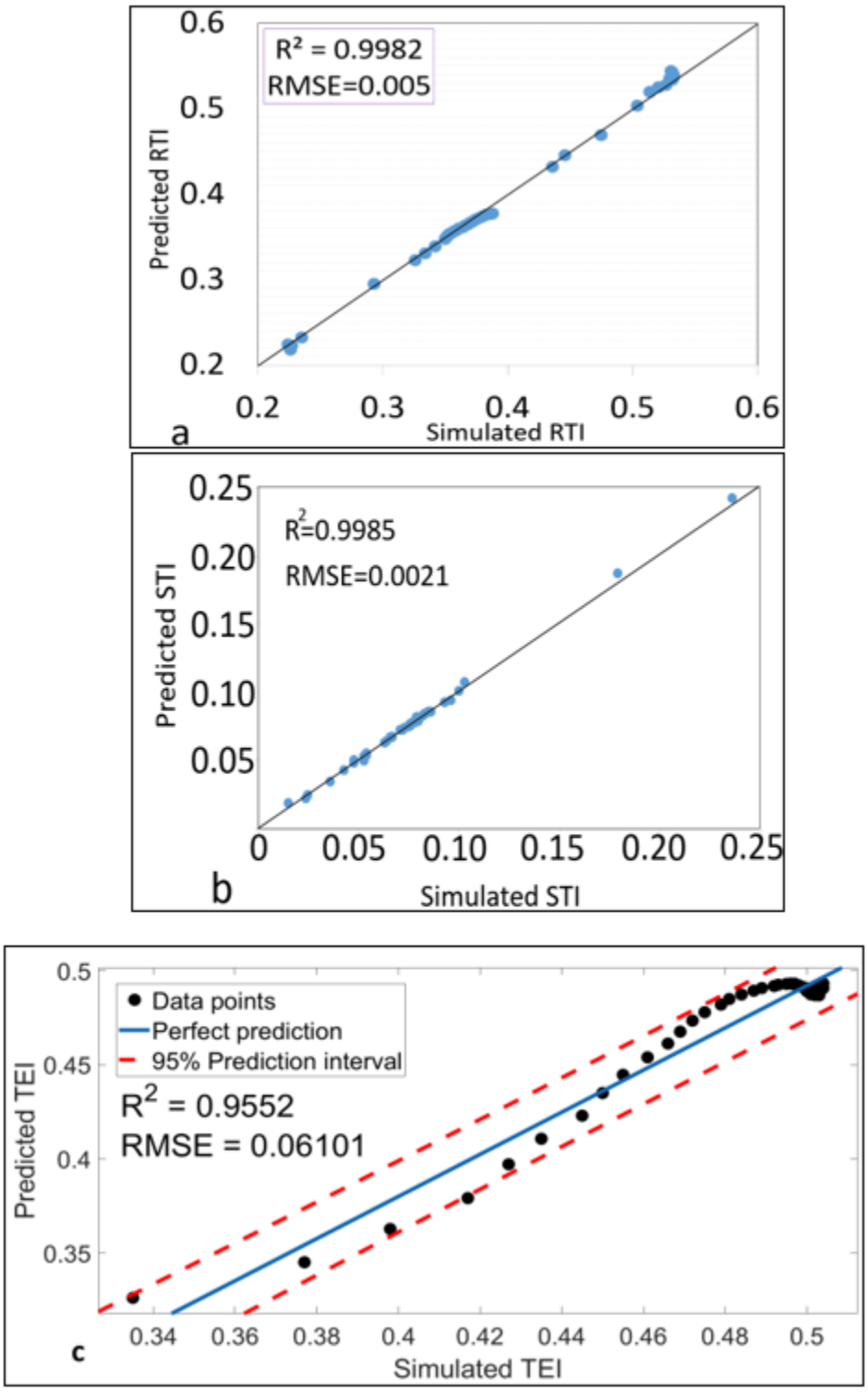

Figure 6

Blind testing results of GPR models for predicting CO2 trapping efficiency. a. Blind testing of RTI; b. Excellent blind testing performance of STI; c. more significant 95% R2 for TEI

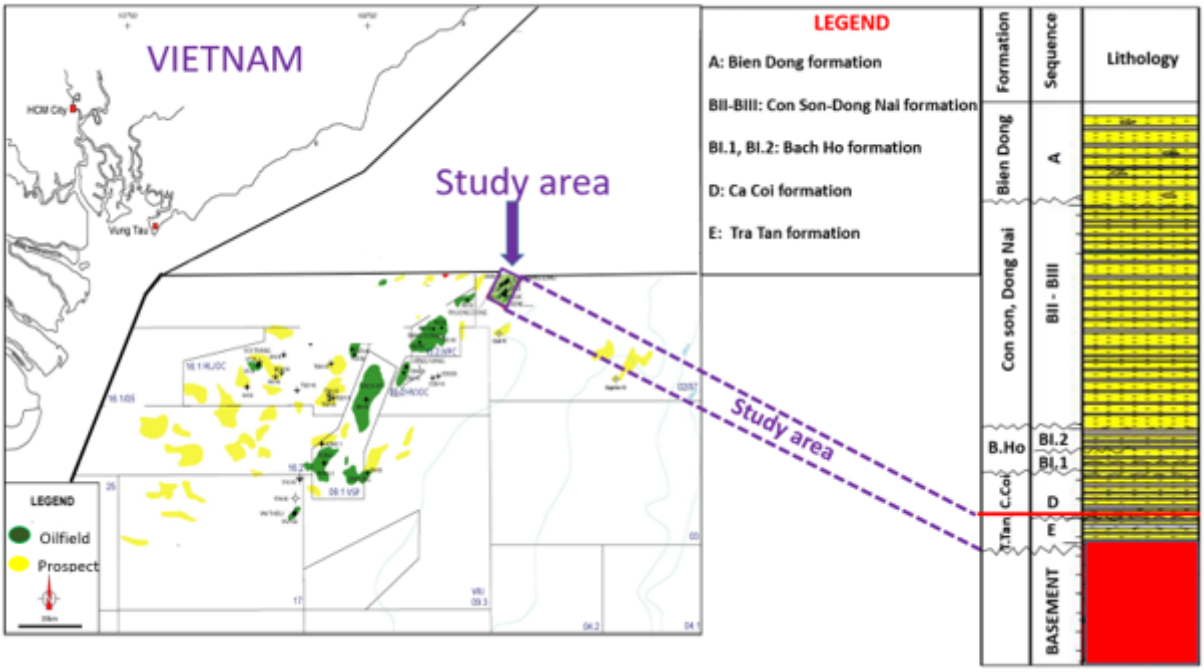

Figure 7

Location and lithology of study area modified after Vo Thanh et al. (2020d). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

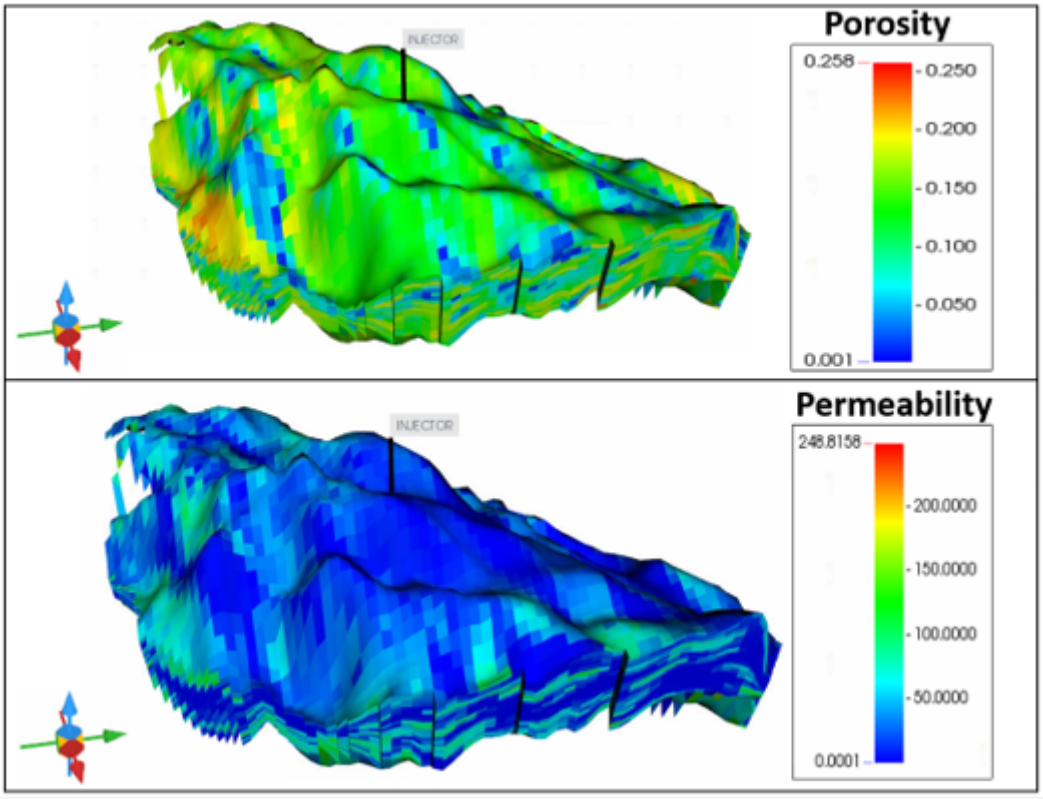

Figure 8

The porosity and permeability models of fluvial channel saline aquifers to compare the prediction values and field scale simulation. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

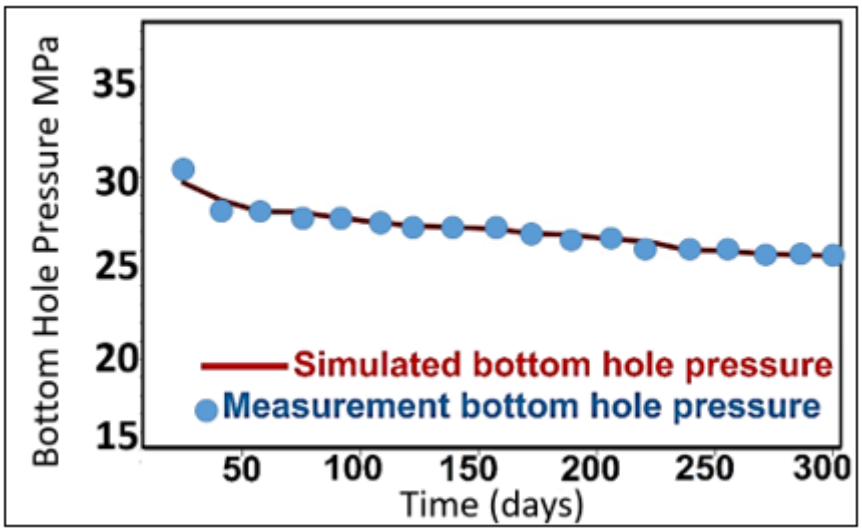

Figure 9

History matching model of bottom hole pressure

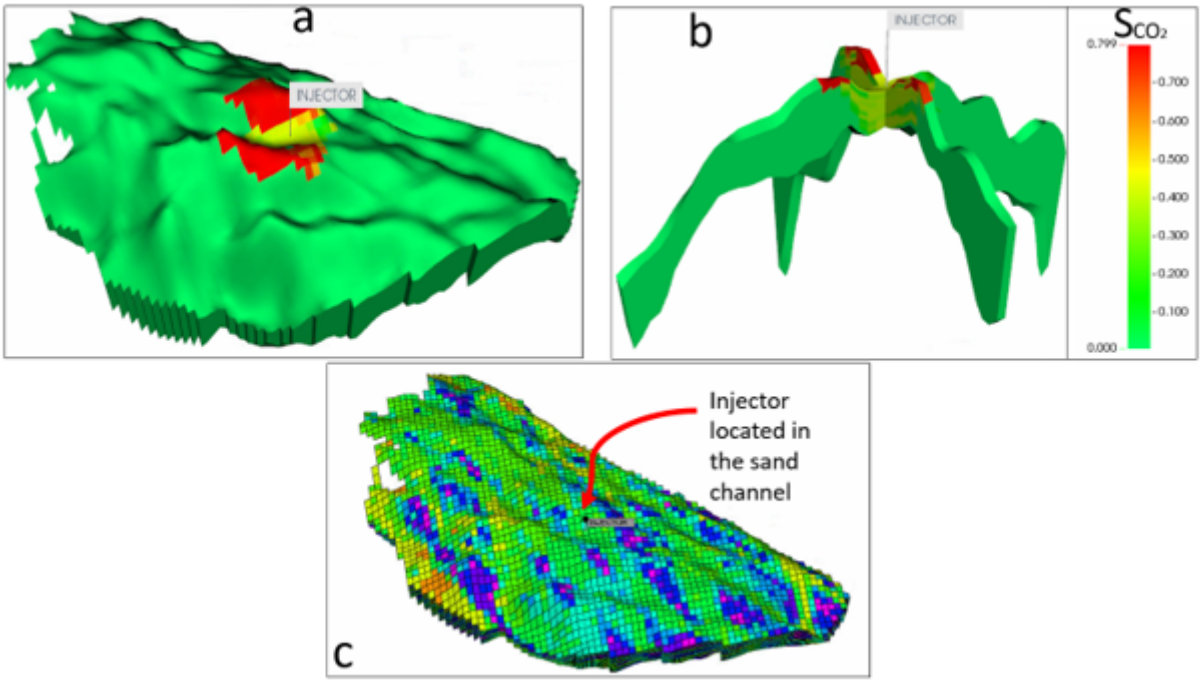

Figure 10

CO2 saturation performance at the end of the simulation period. The distribution of CO2 saturation in a fluvial sandstone aquifer. a. 3D gas saturation model at the end of simulation period; b. the cross-section to represent the CO2 flow in fluvial system; c. the injector is located in the fluvial sandstone system. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

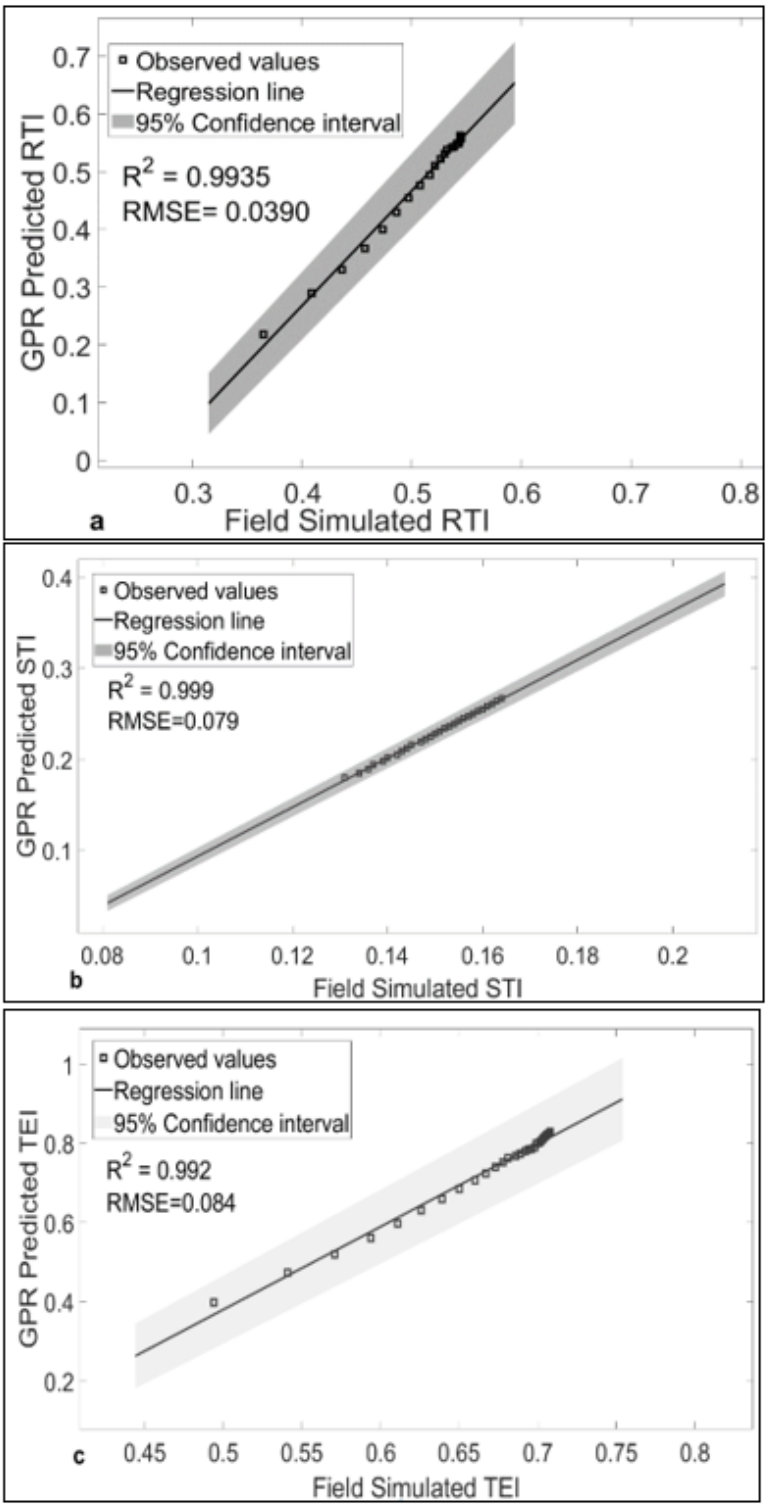

Figure 11

Validation results between GPR models and field simulated CO2 trapping efficiency. a. Excellent agreement between GPR predicted and field simulated RTI; b. remarkable correlation between predicted and field simulated STI using GPR based machine learning predictive models; c. reasonable predicted TEI using GPR predictive models.

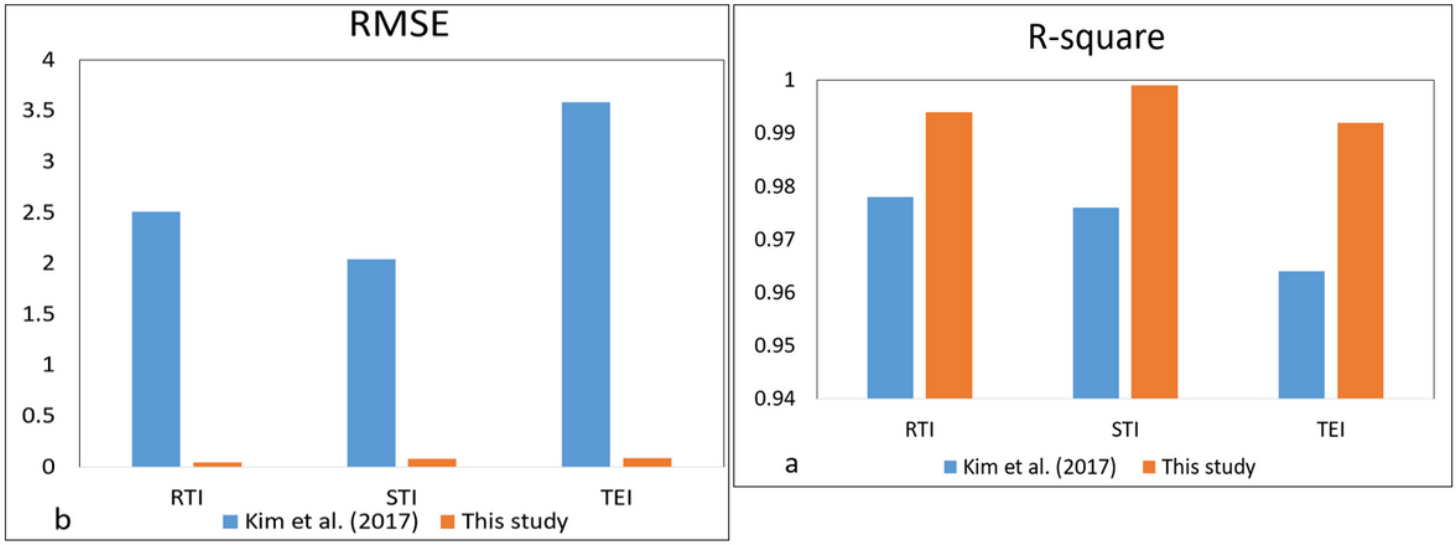

Figure 12

The comparison performance between this study and previous model in terms of R-squared and RMSE. a. Present study showed a larger R-squared value than the previous study, in three trapping efficiency (RTI, STI, and TEI); b. RMSE of present study is smaller than the RMSE value of previous research.