Journal of Cleaner Production 546 (2026) 147737

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Multi-objective optimization of operation conditions by smart proxy model for geological CO_2 storage design in an over-pressured aquifer in the Ulleung Basin

Kyuhyun Kim, Seongjun You, Dayeon Kim, Jihoon Wang *

Department of Earth Resources and Environmental Engineering, Hanyang University, Seoul, 04763, South Korea

A B S T R A C T

This study aims to optimize the geological CO_2 storage (GCS) design in the D structure in the East Sea using a Smart Proxy Model (SPM). Analyzing the geomechanical stability of the overpressured aquifer, it was found that the pore pressure in the fault plane must be maintained below the current pressure value. Assuming that two existing wells were repurposed as an injection well and a relief well, Feature Engineering was adopted to derive new parameters for training the SPM with a reduced number of simulation runs. The optimization was conducted to maximize cumulative CO_2 injection and minimize brine production, considering several design parameters, including pre-relief period, the selection of pre-existing wells, and perforation intervals. The model exhibited high predictive accuracy in blind testing, with higher R^2 than 90%. 14 optimal solutions and the Pareto front were obtained by employing the SPM in the NSGA-II algorithm. As a result, the lowest cumulative CO_2 injection and brine production are 3.89 Mt and 5.12 Mt, respectively, while the largest CO_2 injection and brine production are 33.9 Mt and 59.7 Mt, respectively. In addition, it was found that the optimal perforation intervals strongly depend on both the operation conditions and reservoir characteristics. Therefore, the perforation intervals of the injection and relief wells should be optimized when a relief well design is incorporated into a GCS process. Although this study is specific to the D structure, the proposed ML-based optimization framework incorporating geomechanical risks can be adapted to other geological formations for GCS designs.

1. Introduction

According to the Intergovernmental Panel on Climate Change (IPCC) in 2018, achieving carbon neutrality by 2050 is imperative to limit the global temperature increase to 1.5 °C above pre-industrial levels (IPCC, 2022). The geological carbon storage (GCS) technology has been accepted as a key strategy to meet the goal by capturing CO_2 before it is emitted into the atmosphere by securely storing it in a geological structure. Among potential storage candidates, saline aquifers are particularly attractive due to their extensive geographical distribution and significant storage capacity (Global CCS Institute, 2023; Global institute, 2022). However, field-scale GCS applications targeting saline aquifers have faced practical challenges, including elevated pore pressures, uncertain leakage pathways, and higher capital expenditures.

To establish an optimal GCS design in a saline aquifer, it is crucial to perform geomechanical analysis to investigate the challenges, to provide solutions to prevent those, and to ultimately ensure the long-term reservoir stability. In general, the challenges can be prevented by limiting the injection pressure considering the maximum allowable pressure (Song et al., 2023). For a saline aquifer, which is initially filled with formation brine, operating a relief well can be a great tool to 'relieve' the pore fluid pressure by extracting the brine before or during the GCS process. When the optimization process is performed by the computational simulation, however, it requires massive computing power and time for repeated simulation runs to determine the optimized pre-relief periods, operating conditions, and the well locations.

Artificial intelligence techniques have been increasingly employed to optimize the operational conditions for GCS with proxy models that replicate reservoir simulation outcomes with reasonable accuracy. Musayev et al. developed an artificial neural network (ANN)-based proxy model integrated with a genetic algorithm to optimize the placement of CO_2 injection and brine production wells in the Pohang Basin (Musayev et al., 2023). developed an artificial neural network (ANN)-based proxy model integrated with a genetic algorithm to optimize the placement of CO_2 injection and brine production wells in the Pohang Basin. The proposed approach reduced the number of simulation runs by 80.7% compared to conventional numerical methods, while maintaining high prediction accuracy. Song and Wang proposed an ANN model to optimize the locations and operating conditions of relief wells for CO_2 storage in the Pohang Basin (Song and Wang, 2021). Considering geomechanical stability, the authors proposed a framework evaluating 10-, 20- and 30-year injection scenarios, and identified the

* Corresponding author. Dept. of Earth Resources and Environmental Engineering, Hanyang University, 105-2 Engineering Science and Technology Bldg., Hanyang University, 222 Wangsimni-ro, Seoul, 04763, South Korea.

E-mail addresses: skkkh95@hanyang.ac.kr (K. Kim), bag9543@hanyang.ac.kr (S. You), edr0309@hanyang.ac.kr (D. Kim), jihoonwang@hanyang.ac.kr (J. Wang).

https://doi.org/10.1016/j.jclepro.2026.147737
Received 23 September 2025; Received in revised form 29 January 2026; Accepted 3 February 2026
Available online 17 February 2026

optimal well location and injection rates. The results show that the proposed scenario achieved a 19.44% increase in the storage capacity with only 0.68% deviation from the full-physics reservoir simulation. Menad and Noureddine proposed a method to optimize a multi-objective problem of the CO_2-WAG process by using a multilayer perceptron-based proxy model combined with the NSGA-II algorithm (Menad and Noureddine, 2019). The authors indicated that the proposed approach gained the results more rapidly and efficiently. Leng et al. reviewed existing methodologies for estimating CO_2 storage capacity at a large scale, addressing discrepancies between theoretical predictions and actual field performance (Leng et al., 2024). The authors highlighted the emerging role of machine learning in integrating reservoir dynamics and uncertainties to enhance estimation accuracy. The study recommends future development of integrated ML-based tools for more reliable and scalable capacity assessments. Vo-Thanh et al. developed advanced machine learning models, including general regression neural networks, to accurately predict CO_2 residual and solubility trapping indexes in saline aquifers (Vo-Thanh et al., 2022). Utilizing over 1900 simulation samples, the models achieved exceptional accuracy, with R^2 values exceeding 0.999. The findings demonstrate the strong potential of machine learning to support reliable evaluation of CO_2 trapping mechanisms in a geological storage site.

However, running a number of reservoir simulations to train and verify an artificial intelligence technique still requires substantial computing resources, especially when the simulation model is complex, such as a geomechanics-fluid flow coupled simulation. Recently, Smart Proxy Model (SPM) has gained attention, which can be constructed much faster than the traditional proxy model by adopting a sampling method to obtain training data with fewer simulation runs (Motaei and Ganat, 2023).

An overpressured reservoir, which has a pore fluid pressure higher than the hydrostatic pressure, is not generally considered for a GCS candidate due to potential geomechanical risks. However, for countries such as South Korea, where the number of CO_2 storage candidates are few, a systematic analysis on the feasibility of utilizing an overpressured aquifer for GCS can provide valuable insights. In addition, the detailed geomechanical stability of existing faults and perforation intervals for both injection and relief wells is considered to optimize the GCS process, which has rarely been addressed in previous studies.

In this study, the multi-objective optimization was performed for the optimal GCS operation design considering geomechanical stability in the D structure, an overpressured saline aquifer in East Sea, South Korea. To establish a safe CO_2 injection strategy, it was assumed a relief well is operated to extract brine and to relieve the pore pressure. A relief well and a CO_2 injection well were designed to ensure that the pore pressure in an existing fault does not exceed its initial value. To ensure a safe and stable GCS process, a 50-year monitoring period after the injection was adopted to ensure that the fault pressure remains below the initial pressure. To identify the optimal operating conditions for both injection and relief wells, an SPM was constructed to predict the reservoir response under various CO_2 injection scenarios. The Nondominated Sorting Genetic Algorithm-II (NSGA) was employed to achieve the optimal solutions for the multi-objective problem, which aims at the maximum CO_2 storage and minimum brine production. The optimization considered several design parameters, including the selection of pre-existing wells, pre-relief period, and perforation intervals. Through this approach, an optimal GCS operation design was achieved by applying machine learning to reduce the number of simulation runs while simultaneously maintaining geomechanical stability.

2. Geological description

2.1. The D structure

The D structure is located east of the Donghae gas field (Fig. 1 (a) and (b)) and contains 2 existing wells (D-1 and D-2). According to analysis performed by the Korea National Oil Corporation, the D structure is a potential candidate for GCS, as it is a well-integrated geological structure with a low risk of CO_2 leakage. Although the reservoir is overpressured with the pore pressure gradient of 0.013 MPa/m, the structure is a viable storage candidate because there are few available structures for CO_2 storage in South Korea. Moreover, the overpressure state refers that the high-pressure fluid has been stored in the structure for a long period of time, implying that the structure is well-confined and the injected CO_2 can be securely stored.

The D structure has high permeability (up to 600 md) and porosity (up to 0.4) (Fig. 2 (a)and (b)). As shown in Fig. 2 (c), there are two wells (D-1 and D-2) and three identified faults in the D structure, which are labeled as F1, F2, and F3. F1, F2 and F3 are extended from 1924 to 2597 m, from 2177 to 2705 m, and from 2153 to 2739 m, respectively (red lines in Fig. 2 (c)). The distances from the D-1 well to F1, F2, and F3 are 268, 6440 and 6931 m, respectively. Although the locations and the depths of faults have been confirmed, their precise geometries including azimuth and dip remain undetermined. Consequently, the fault reactivation analyses were conducted with the worst-case scenario, i.e. the faults have the weakest geometries in terms of the potential fault reactivation. Moreover, the fault stability analysis was specifically performed on the closest F1 fault, where it is expected that the largest pressure change occurs.

2.2. Fault stability analysis

The fault stability evaluation incorporates properties that represent the weakest and most failure-prone conditions of the fault plane, providing the worst-case scenario for potential fault reactivation. To investigate the stability of the faults in terms of potential reactivation,

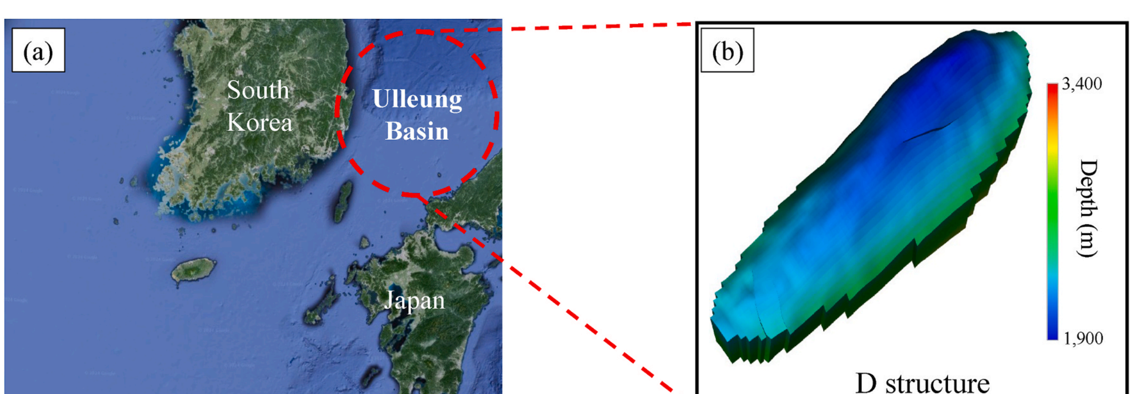

Fig. 1. (a) Location of the D structure in Ulleung Basin (b) Shape and depth of the top of the D structure.

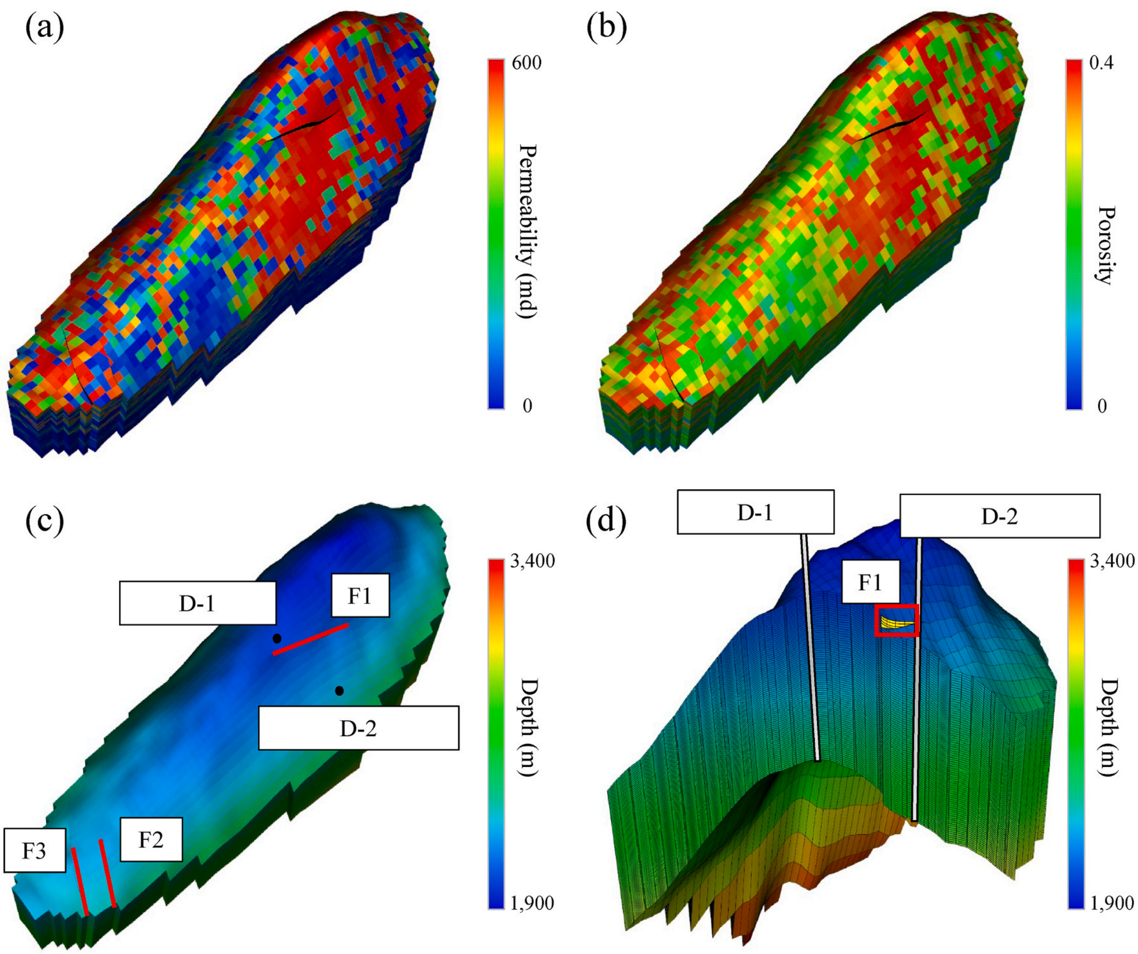

Fig. 2. (a) Permeability (b) Porosity of the D structure (c) Existing wells (D-1, 2) and faults (F1, F2, F3) in the D structure (d) Cross-sectional views for both wells.

the Mohr-Coulomb (MC) criterion was incorporated. The maximum and minimum x-intercepts of the Mohr circle correspond to the effective maximum and minimum principal stresses, respectively. As the effective normal stress acting on the fault increases, the shear stress required to induce fault slip also increases. A linear relationship exists between shear stress and effective normal stress, known as the Coulomb friction failure criterion, can be written as follows,

$$\tau = C + \mu \sigma_n' \qquad \text{Equation 1}$$

where C is the cohesion of the fault plane, and μ is the friction coefficient, which are material properties of the fault plane. σ_n' is the effective normal stress on the fault plane, which is the difference between the normal stress and the pore pressure, $\sigma_n - P_p$. It was assumed that the cohesion of the fault is zero and the friction coefficient is 0.6 for design purposes, reflecting the worst-case scenario.

The stability of the fault is evaluated by how close the stress state on the fault plane is to the stress condition required for the fault reac-

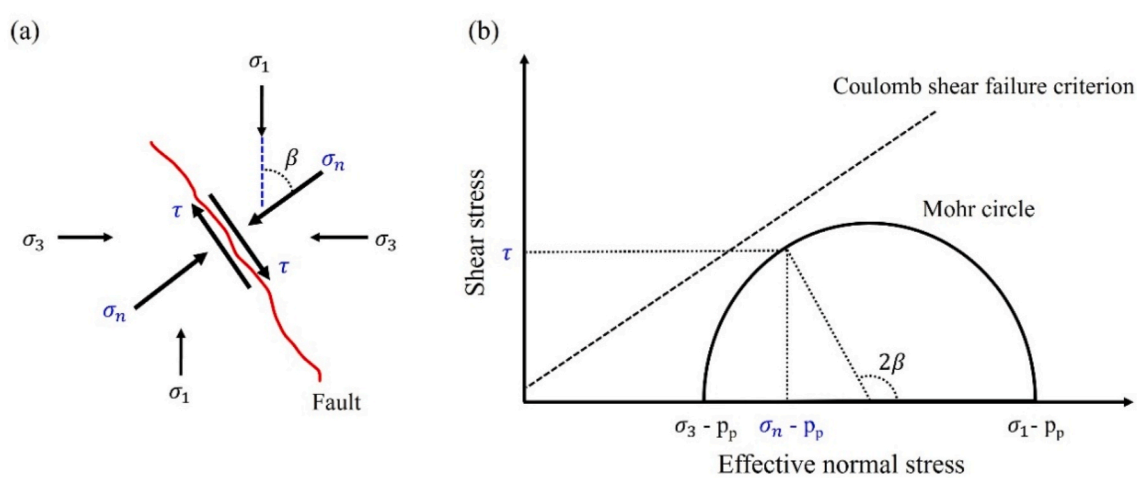

Fig. 3. (a) Stress state of a fault plane (b) Mohr Circle and Coulomb shear failure criterion.

tivation. The normal and shear stresses on the fault plane are determined by the magnitudes of the stress field around the fault, as well as the orientation of the fault relative to the stress field. In a two-dimensional stress regime, if the maximum and minimum in-situ principal stresses (σ_1 and σ_3) and the angle β between the σ_1 and σ_n are known, the normal stress (σ_n) and shear stress (τ) on the fault can be read directly from the Mohr circle (Fig. 3). Thus, under a given principal stress field, the fault's orientation significantly influences the magnitudes of σ_n and τ acting upon it.

When CO_2 is injected into a reservoir, the top layers are generally considered the most vulnerable. When the pressure gradient and the stress gradient within the reservoir are constant (Fig. 4 (a)), the Mohr circle becomes smaller and shifts further to the left as one moves upward, indicating lower fault stability at shallower depths. When comparing the Mohr circles at the top and bottom of the fault (Fig. 4 (b)), the circle at the top of fault is smaller and lies closer to the Coulomb shear failure envelope. This indicates that, during CO_2 injection, the Mohr circle at the top of the fault is more likely to reach the Coulomb shear failure criterion first. Consequently, the analysis focuses on the top layer of the fault.

Based on the leak-off test (LOT) and the image log analysis, it was found that the D structure is located within the strike-slip regime with the maximum horizontal, vertical and minimum horizontal principal stress gradients of 0.030, 0.021, and 0.016 MPa/m, respectively. In addition, the directions of the maximum and minimum principal stresses are N86°E and N4°W. The results of the LOT indicate that the fracture pressure gradient in the target formation is 0.016 MPa/m. The pore pressure gradient was measured as 0.013 MPa/m. The uppermost layer of F1 is located at a depth of 1930 m as shown in Fig. 5 (a), with the maximum and minimum principal stresses of 57.49 MPa and 31.24 MPa, respectively. Although the fault stability analysis based on the parameters revealed that the F1 within the D structure is expected to be geomechanically unstable, as the Coulomb shear failure criterion intersects the Mohr circle (Fig. 5 (b)), the fault stays stable currently. There are a few reasons that the fault is stable, even though the analysis predicts it is not. At first, this is because the actual fault geometry might not be the weakest for the potential reactivation. The current stress state of the fault plane is determined by the in-situ stress distribution, and the fault geometry might lie within the dotted area in Fig. 6 (a). In addition, although the friction coefficient of the fault plane was assumed as 0.6, the lower bound to represent a worst-case scenario, the actual value may exceed 0.6, which is commonly in the range from approximately 0.6 to 1.0 (Byerlee, 1978). For comparison, the Coulomb shear failure envelopes corresponding to the friction coefficients of 0.6 and 1.0 are also shown in Fig. 6 (b). Similarly, the cohesion of the fault plane can be a positive value. Accordingly, Fig. 6 (c) compares the Coulomb shear failure envelopes corresponding to the cohesion values of 0 and 1.0 MPa. Moreover, the MC criterion used for the analysis underestimates the fault plane strength by ignoring the effect of the intermediate principal stress. Therefore, the current fault stability is primarily due to the worst-case assumptions for the design purpose and the criterion that underestimates the strength of the plane. However, the fault can be unstable if the pore pressure increases by the CO_2 injection and reaches the value that induces the fault reactivation. Consequently, in order to maintain the fault stability, the pore pressure at the faults has to be maintained same as or below the current pore pressure value. Therefore, it was assumed that a relief well operates to extract brine and to reduce the pore pressure.

3. Methodology

3.1. An overview of the modeling framework: smart proxy model and genetic algorithm

3.1.1. Optimization strategy for GCS design in the D structure

To establish the optimal GCS design in the D structure, design parameters for optimization were determined, i.e. pre-relief period, well type, and well perforation intervals. At first, the pre-relief period was selected as a design parameter, as it is required to extract a specific amount of formation brine before the CO_2 injection and to prevent pressure buildup over the initial pore fluid pressure. In addition, either of the existing two wells in the field, D-1 and D-2, was purposed as a CO_2 injection well or a relief well. On the other hand, although it has not been taken into account as an optimized parameter during a GCS process, the perforation interval of both injection and relief wells is also an important design parameter. This is because when the injection and relief wells operate at the lower and upper portions of the reservoir, respectively, the CO_2 storage would be efficient with the buoyancy effect of the injected CO_2. However, it might also cause the CO_2 breakthrough to the relief well. Since the CO_2 breakthrough would require additional operational processes and continue to occur due to the low viscosity of the CO_2, there is a previousness that terminates the operation once it is detected (Song and Wang, 2021).

Consequently, the multi-objective optimization was conducted to minimize cumulative brine production and to maximize the cumulative CO_2 injection over an 80-year period (including 30-year injection and 50-year monitoring). Due to the high computational cost associated with reservoir simulations, a SPM combined with NSGA-II was adopted to efficiently solve the optimization problem with the selected design parameters.

3.1.2. Smart Proxy Model (SPM)

Traditional Proxy Model (TPM) use statistical methods based on results from numerical simulation to identify the complex, non-linear connections between model parameters and output data. If the proxy model is developed appropriately, its outcome efficiently represents that

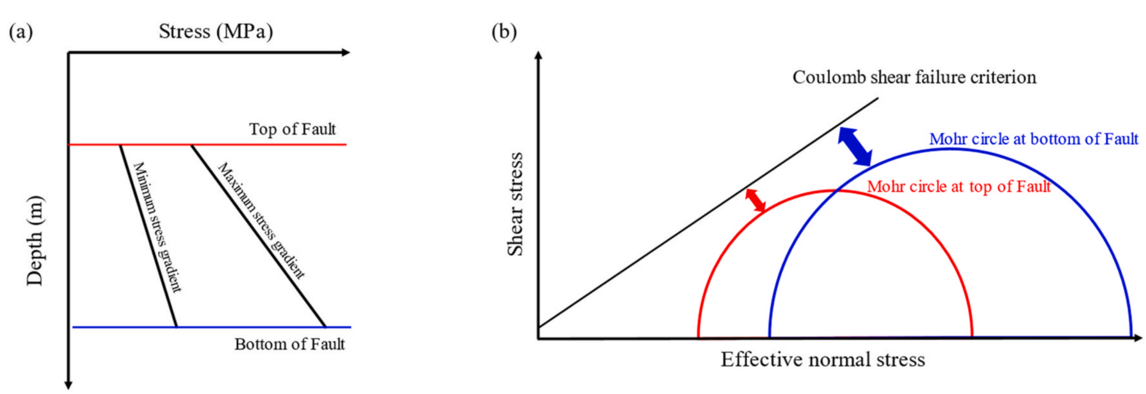

Fig. 4. Comparison of the Mohr circles at the top and bottom of a fault. (a) Maximum and minimum stress gradient along the fault. (b) Stress states at the top and bottom of the fault under the corresponding stress gradients.

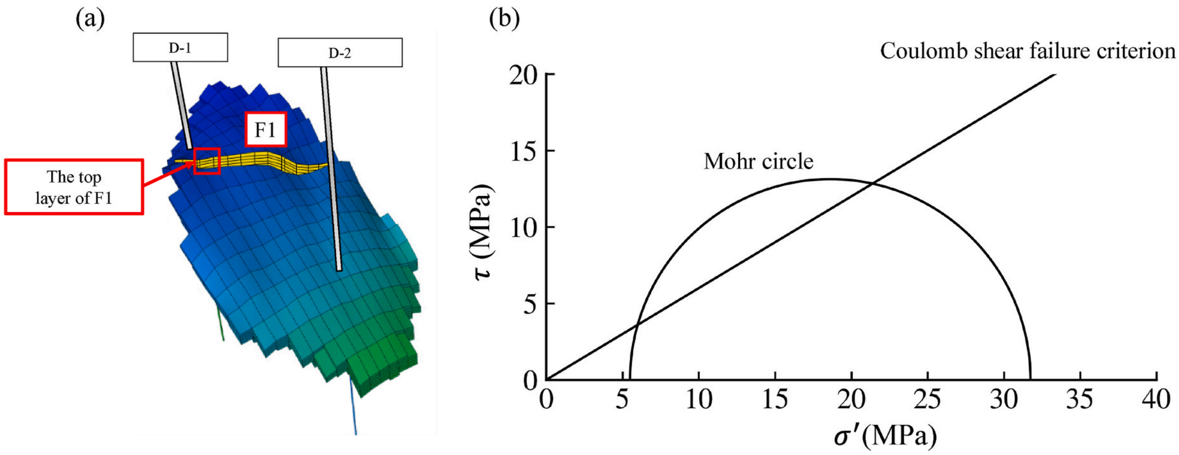

Fig. 5. (a) The top layer of F1 location (b) Mohr circle showing the stress state at the top layer of F1 under worst case scenario.

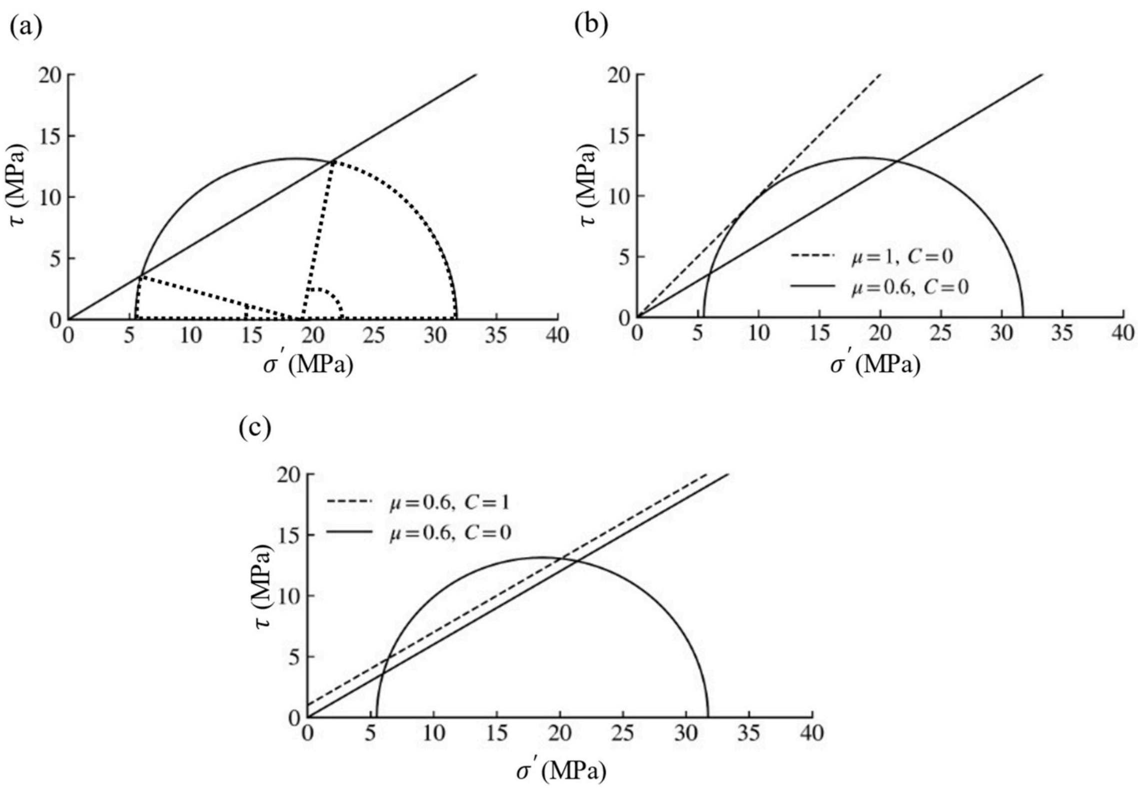

Fig. 6. Effects of fault geometry, friction coefficient, and cohesion on fault reactivation according to the Mohr–Coulomb failure criterion (a) Range of angles between the principal stress and the fault (b) Coulomb failure envelopes for friction coefficients μ = 0.6 and 1.0 (c) Coulomb failure envelopes for cohesion values C = 0 and 1.0 MPa.

of the original model by capturing the characteristics of the original model. However, it requires a number of simulation runs to accurately capture the reservoir behavior. Moreover, if the model is large and complex, the computational power and time required to generate the training data would be massive.

SPM, proposed by Mohaghegh (2018), includes an additional step of extracting important features from simulation data, and significantly reduces the number of simulation runs required to replicate the original numerical model. The feature engineering process is a key component that generates new static and dynamic derived variables from simulation results. The additional variables help capture hidden patterns that cannot be detected using the original dataset alone, allowing the SPM to achieve higher accuracy than the TPM. The method has been widely used for various applications, such as history matching analysis, operational monitoring, and well operation condition optimization (Bahrami and James, 2023; Mohaghegh et al., 2009; Vida et al., 2019). SPM in the field of reservoir engineering can be categorized into the well-based SPM and grid-based SPM, depending on the objective of the problem. The well-based SPM is typically employed in history matching and production optimization, as it focuses on well-level predictions, while the grid-based SPM is used for grid-scale evaluations, such as pressure distribution forecasting (Mohaghegh et al., 2009).

As illustrated in Fig. 7, proxy models can be classified into multi-fidelity models, reduced order models, TPM, and SPM (Bahrami, Moghaddam, 2022). According to Bahrami and Moghaddam, multi-fidelity models and Reduced-Order Models (ROM) are physics-based proxy

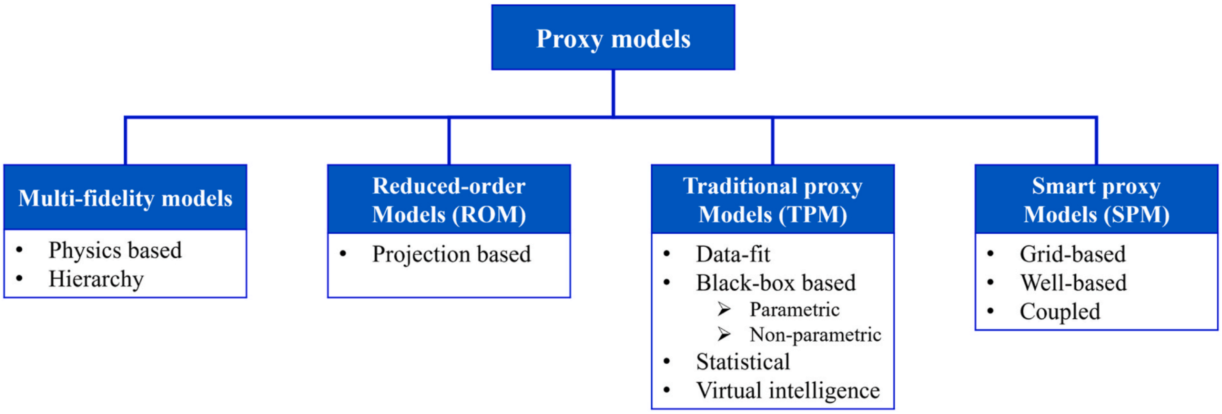

Fig. 7. Proxy model categories—MFM, ROM, TPM, and SPM.

models that improve computational efficiency by simplifying or reducing the dimensionality of numerical simulations. These methods can reduce computation cost compared to full-physics simulations, but they involve a trade-off between accuracy and efficiency.

In contrast, TPM and SPM are data-driven approaches that do not solve the governing physical equations directly, but instead learn relationships between input and output data from numerical simulation results. Although TPM requires computationally inexpensive processes, its regression-driven nature limits its ability to capture key nonlinear and dynamic reservoir behaviors, such as pressure changes and multiphase flow interactions. Moreover, achieving reliable TPM performance typically requires a large number of simulation runs. In contrast, SPM generates additional learning data through feature engineering and a tiering system, and enables the model to learn temporal and spatial reservoir patterns more explicitly. This structure allows SPM to represent complex pressure diffusion behaviors and operational changes with better accuracy. With the enhanced learning mechanisms, SPM typically achieves high fidelity with a smaller number of simulation iterations compared to other proxy models, making it a more efficient option for reservoir reaction forecasting (Bahrami, Moghaddam, 2022).

A well-based SPM was constructed to predict cumulative CO_2 injection and cumulative brine production to optimize the well operation conditions in GCS design in the overpressured D structure. The workflow for developing SPM is illustrated in Fig. 8. The initial step includes defining the objectives of the model and selecting the design parameters with their respective ranges. The Latin Hypercube Sampling (LHS) method, which divides design parameters into spatially equal intervals, was adopted to capture spatially diverse information while minimizing the number of simulation runs (Bahrami and James, 2023).

Upon completion of all the simulation runs, an enhanced dataset was constructed with additional features derived from the numerical simulations, which enable the development of an SPM capable of capturing complex reservoir behavior under various scenarios. This feature extraction step is essential for identifying hidden patterns within the data, which traditional proxy models often miss. The final stage of the workflow involves selecting an appropriate machine learning algorithm for SPM construction.

A 1D Convolutional Neural Network (1D-CNN) is recognized for its capability to capture local patterns within sequential data, and is particularly efficient for predicting reservoir responses under various operational conditions, such as injection and production rates, and bottomhole pressure (Wang, 2022). 1D-CNN offers stable predictive performance with relatively low computational cost in time-series applications, such as stock-price forecasting and signal processing. Since the input variables in this study exhibit clear local temporal patterns, such as cumulative CO_2 injection, and cumulative brine production, a 1D-CNN is well suited to learn these behaviors effectively. In addition, a 1D-CNN tends to provide faster training and prediction speeds compared to recurrent architectures such as LSTM and GRU, which can be beneficial for repeated evaluations when NSGA-II is incorporated for multi-objective optimization (Lee et al., 2025). A 1D-CNN model was adopted to construct the SPM. After developing the 1D-CNN-based SPM, its performance and reliability were evaluated by blind tests, which include additional simulation runs. In addition, the validation step was conducted to ensure the model's accuracy in predicting the reservoir behavior under various conditions.

3.1.3. Genetic algorithm

An optimal GCS operation design must account for brine production, CO_2 injectivity, and geomechanical stability to ensure safe and efficient process. Excessive brine production reduces economic feasibility, while pressure buildup caused by excessive CO_2 injection can lead to tensile failure near the injection well or reactivation of existing faults (Song et al., 2023). The optimization was performed to maximize the cumulative CO_2 injection and to minimize the cumulative brine production with the geomechanical stability. Therefore, a multi-objective

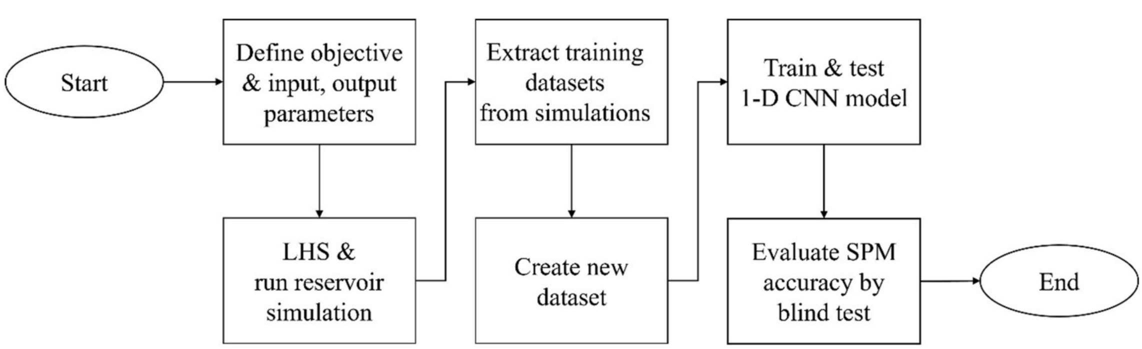

Fig. 8. Workflow of SPM development.

optimization approach was adopted to determine the optimal GCS operation strategy. Genetic Algorithm (GA) is a widely used meta-heuristic optimization technique that mimics principles of natural evolution (Deb et al., 2002). It explores multiple solutions simultaneously, which is particularly beneficial in multi-objective optimization as it effectively balances conflicting objectives (Srinivas and Deb, 1994). This study adopted Nondominated Sorting Genetic Algorithm-II (NSGA-II) proposed by Deb et al., which applies the mechanisms of nondominated sorting and crowding distance to maintain computational efficiency while generating a well-distributed Pareto front as illustrated in Fig. 9 (Deb et al., 2002). The Pareto front is a set of solutions in a multi-objective optimization problem where no solution can be improved on an objective without making the other objective worse, as shown in Fig. 10 (Joshua et al., 2023).

The workflow of NSGA-II begins with generating an initial random population, which is evaluated based on a fitness function (Verma et al., 2021). The selection mechanisms, such as the tournament selection and roulette wheel selection, identify high-performing solutions for reproduction. The crossover and mutation operations introduce diversity, improving search efficiency and preventing local optima. The next generation is formed by replacement strategies, often incorporating elitism to retain the best solution. This process iterates until a termination criterion is met, such as convergence or a predefined generation limit. NSGA-II further enhances the workflow by ensuring a well-distributed Pareto front through non-dominated sorting and crowding distance mechanisms (Deb et al., 2002). The combination of proxy modeling and NSGA-II has been proven effective in CO_2-EOR and CCS due to its advantage in reducing computational costs while efficiently managing multi-objective optimization with conflicting objectives (Menad and Noureddine, 2019; Na et al., 2024; Sun et al., 2021).

3.2. Data generation and preparation

It was assumed that the pre-existing drilling wells were repurposed as an injection well and a relief well. The injection well was operated with the bottomhole pressure of 90% of the fracture pressure at the uppermost perforation interval in accordance with the U.S. Class VI well regulations (US EPA, 2012). In addition, considering the overpressured state of the reservoir, a relief well operates for a specific period before the injection process. For the relief well, operational constraints were imposed to maintain the reservoir pressure above the hydrostatic pressure. After 30 years of process, the pore pressure at the top of F1 was monitored for an additional 50 years. Additionally, the process

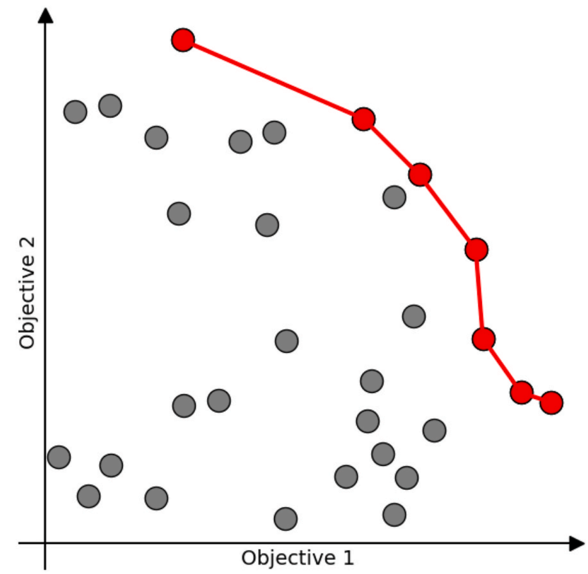

Fig. 10. Example Pareto front with two objectives. The gray points represent non-dominated solutions. The red points at the Pareto front represent the best solutions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

terminates immediately if the injected CO_2 is detected in the relief well or the pore pressure at the top of F1 exceeds 25.8 MPa. During the reservoir simulation process, it was found that approximately 0.1-0.4 kg/day of native CO_2 is produced, even before the injected CO_2 reaches the relief well. Once the injected CO_2 plume reaches the relief well, the CO_2 production rate increases to several tens of kg/day or higher, indicating a clear breakthrough phenomenon. Therefore, in this study, the breakthrough was defined as the condition in which CO_2 production greater than 1 kg/day is detected at the relief well. The fault instability case was defined as the condition in which the pore pressure at the top of F1 exceeds the initial pressure of 25.8 MPa.

The optimization process aims to maximize the cumulative CO_2 injection and to minimize the cumulative brine production for 30 years. The pre-relief period, the perforated interval of each well, and the well type were selected as design parameters. Once established, the design parameters remained constant throughout the operation for each scenario. Table 1 presents the types and ranges of the design parameters

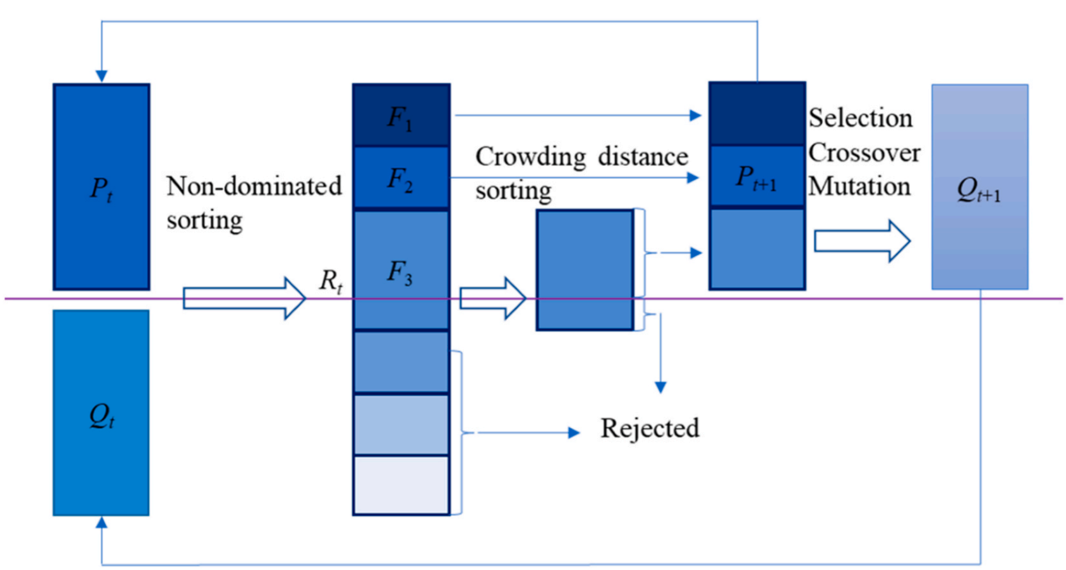

Fig. 9. A schematic representation of the NSGA-II workflow.

Table 1
Types and ranges of design parameters.

Parameter	Min	Max
Pre-relief period (month)	3	36
Well perforation grids (grid cell)	1	80
Well type	Injection well or relief well	

employed in the simulation. The main advantage of SPM is that it can achieve high predictive performance with far fewer simulation iterations compared to TPM. According to Bahrami and Moghaddam, 2022; Na et al., 2024; Vida et al., 2019, SPM can be constructed with a small number of simulation runs, as the previous studies demonstrated that successful development of SPM can be achieved with only 3, 13, or 16 simulations. In this study, a total of 60 scenarios were generated using the LHS method to construct the training dataset, aiming to more accurately capture the pressure and flow behavior of the reservoir according to the perforation interval. All simulations were conducted using the reservoir simulator, GEM by Computer Modeling Group (CMG).

To reduce the required simulation runs for constructing the SPM, additional parameters were extracted from the simulation results (Table 2). The parameters used for constructing SPM were classified as static and dynamic parameters. The static parameters are constant during an operation, such as well operating conditions. On the other hand, the dynamic parameters include the cumulative CO_2 injection and brine production, and the pore pressure at the top of F1, which vary during the operation. To analyze new patterns in the generated reservoir simulation dataset, the tiering system, proposed by Bahrami et al. and Vida et al. (Bahrami and James, 2023; Vida et al., 2019), was incorporated into the porosity and permeability. The tiering system was implemented at the grid level, where the tiers was determined by its relative distance from the well. According to Vida et al., the tiering system enables the SPM to effectively capture grid-level flow behavior and spatial structure, and to improve the learning efficiency. Based on the well perforation interval of each well, the porosity and permeability values were calculated by averaging the values of the grids from Tier 1 and Tier 2, where Tier 1 represents the perforation interval grids and Tier 2 includes the grids surrounding the perforated interval (Fig. 11). The dynamic parameters were extracted daily and incorporated into the training process, while the predictive model was designed to estimate the cumulative CO_2 injection and brine production, the F1 pore pressure, and the potential CO_2 breakthrough.

Table 2
Static and dynamic parameters extracted from simulation result.

Parameter		Details
Static	Pre-relief period	Designed relief period before CO_2 injection
	Well type	Purpose of wells: Injection well or relief well
	Well perforation grids	Designed perforated grids for both wells
	Perforation interval	Vertical depth of well perforation grids
	Porosity (Tier 1)	Average porosity of well perforated grids
	Porosity (Tier 2)	Average porosity of well perforated grids and surrounding grids
	Permeability (Tier 1)	Average permeability of well perforated grids
	Permeability (Tier 2)	Average permeability of well perforated grids and surrounding grids
	CO_2 injection pressure	90% of fracture pressure at injection well
Dynamic	Cumulative CO_2 injection	Field cumulative CO_2 injection mass
	Cumulative brine production	Field cumulative brine production mass
	F1 pore pressure	Pore pressure at the uppermost layer of F1
	Potential CO_2 breakthrough	CO_2 production rate at relief well

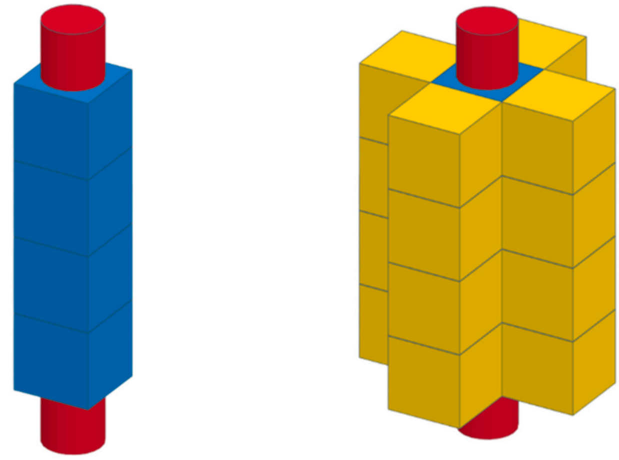

Fig. 11. Tiering system around the well perforation for porosity and permeability data extraction. Tier 1 represents the perforation interval grids, and Tier 2 includes the grids surrounding the perforated interval. The red cylinder represents the wellbore. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.3. SPM construction/validation

The designed SPM operates using the autoregressive forecasting method, leveraging historical data to predict future values by repeatedly using its previous predictions as inputs for the next step, specifically utilizing the predicted values from data in the preceding 3 years to generate predictions in the next time step. This method is particularly effective in capturing temporal patterns and autocorrelations in time-series data. Fig. 12 illustrates the operational mechanism of the autoregressive forecasting approach. Due to its straightforward adaptability, the approach is widely utilized over various scientific and engineering disciplines, including structural mechanics, materials modeling, fluid dynamics, and general dynamical systems (Farooq et al., 2024). The SPM was designed to iteratively generate the next time step data using the previous three years of data.

To assess the accuracy of the SPM predictions, the R^2 was used as the performance metric, R^2. The R^2 can be calculated using Equation (2).

$$R^2 = \left(1 - \frac{\sum_{i=1}^{n}\left(y_{Act,i} - \hat{y}_{Pred,i}\right)^2}{\sum_{i=1}^{n}\left(y_{Act,i} - \bar{y}_{Act}\right)^2}\right) \quad \text{Equation 2}$$

Where, $y_{Act,i}$ is the actual value, $\hat{y}_{Pred,i}$ is the predicted value, and $\bar{y}_{Act}$ is the average of the actual values. As R^2 approaches 1, it indicates that the predicted values closely align with the actual trend.

The accuracy of the SPM model was further evaluated by blind testing, comparing its predictions against additional numerical simulation results to validate the model's predictive reliability. For a blind test, the cumulative CO_2 injection, cumulative brine production, potential CO_2 breakthrough, and F1 pore pressure were obtained annually during 30 years of CO_2 injection and 50 years of monitoring by the additional reservoir simulation runs. Consequently, each scenario contained 80 verification points for comparison. Fig. 13 (a), (b), (c) and (d) illustrate the blind test results, which compared the numerical reservoir model results (blue curve) and the SPM predictions (red curve) for the cumulative CO_2 injection, cumulative brine production, F1 pore pressure and the potential CO_2 breakthrough, respectively.

Moreover, to verify the robustness of the model under operational constraints, an additional blind test was conducted when the operation was terminated before the 30-year injection was finished due to either the injected CO_2 production or the pore pressure at F1 exceeding the

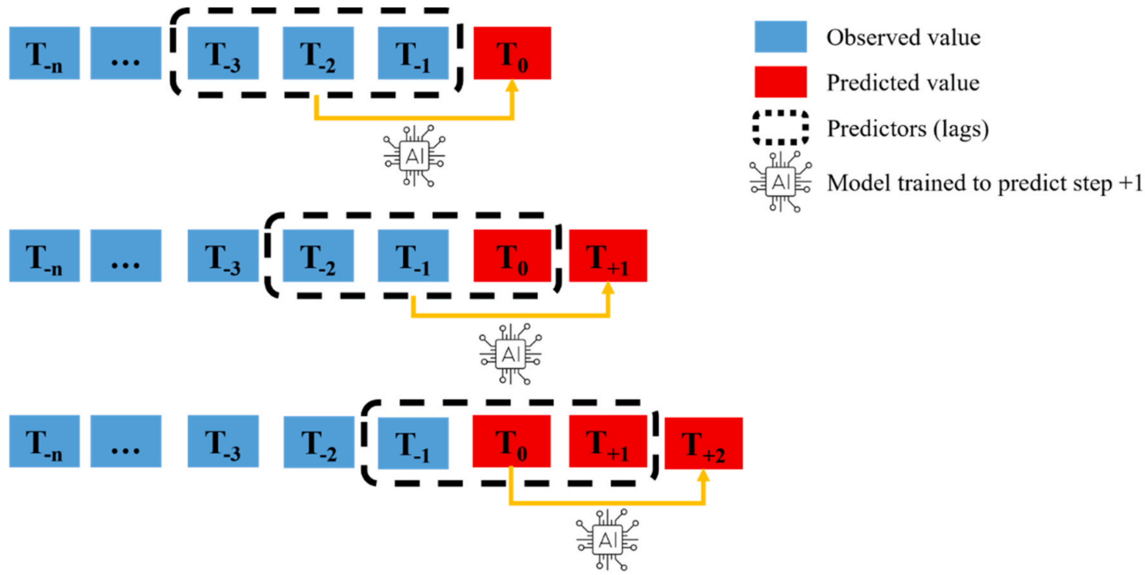

Fig. 12. Demonstration of autoregressive forecasting. The blue boxes represent the available data, while the red boxes indicate the data predicted by the AI model. Using the AI-predicted data, future states are iteratively forecasted. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

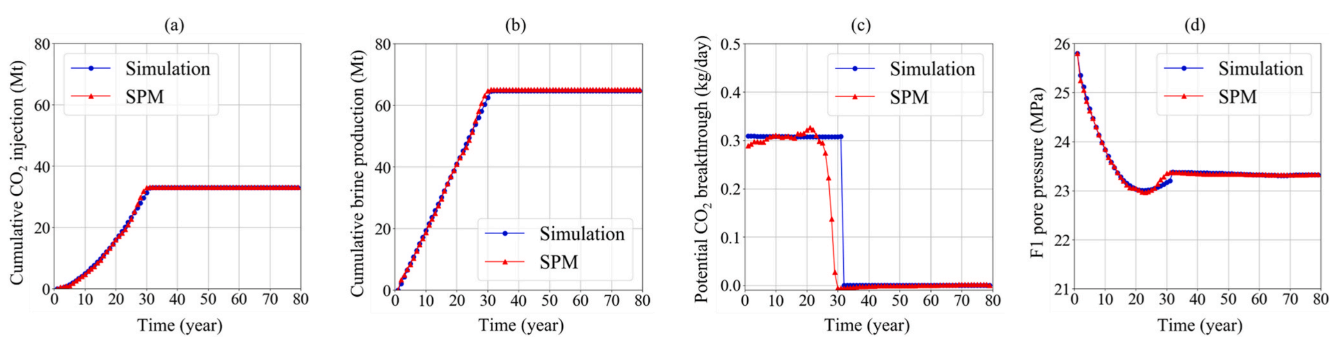

Fig. 13. Results of the blind test for the 30-year operation without early termination. The blue curves represent the reservoir simulation outputs, and the red curves represent the SPM predictions. (a) Cumulative CO_2 injection (Mt), (b) Cumulative brine production (Mt), (c) Potential CO_2 breakthrough (kg/day) and (d) F1 pore pressure (MPa). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

initial pore pressure. The results demonstrated consistent predictive accuracy even for the early-terminated scenarios. Fig. 14 (a), (b), (c), and (d) illustrate the blind test results for CO_2 breakthrough case, while Fig. 14 (e), (f), (g), and (h) present the results for the pore pressure at F1 exceeding the initial pressure.

The coefficient of determination (R^2) values for the key performance indicators are described in Table 3, which represent the average results obtained from the three blind tests conducted for each of the three cases: the stable 30-year injection case, the CO_2 breakthrough case, and the fault instability case. As the results show, the R^2 values of all of the cases are higher than 90%. In addition, for the stable 30-year injection case, in which the CO_2 injection is performed for 30 years stably, the R^2 values are higher than 93.8%. In conclusion, the blind test results confirm the reliability of the SPM model in reproducing simulation outputs without overfitting, ensuring accurate and efficient reservoir performance assessment.

3.4. Multi objective optimization

The developed SPM mimics reservoir responses and was utilized with NSGA-II to perform multi-objective optimization for the optimal GCS design in the D structure. The cumulative CO_2 injection and cumulative brine production at the end of 30 years of injection were set as the optimization objectives. Each objective was normalized and incorporated into the NSGA-II process. When the early termination occurred due to the CO_2 breakthrough or the fault instability, the cases were considered unacceptable and were automatically removed from the NSGA-II searching space. As a result, only the stable 30-year injection cases were used to construct the Pareto front.

Fig. 15 illustrates the overall workflow of the NSGA-II process. The population size was defined as 200 with the generation size of 100. In addition, the crossover rate was 0.5, and the mutation rate adaptively adjusted from 0.2 to 0.05. The parent individuals were selected using the tournament selection method. In each generation, 20 elite individuals were preserved, while 50 random immigrants were added to maintain the population diversity. Table 4 presents the parameters adopted for the NSGA-II.

4. Results

4.1. Pareto front for the multi-objective problem

200 random samples were generated by the SPM to identify the cumulative CO_2 injection and brine production under various operating conditions. Fig. 16 (a) illustrates the random case results in the domain of the cumulative CO_2 injection and brine production. All of the cases

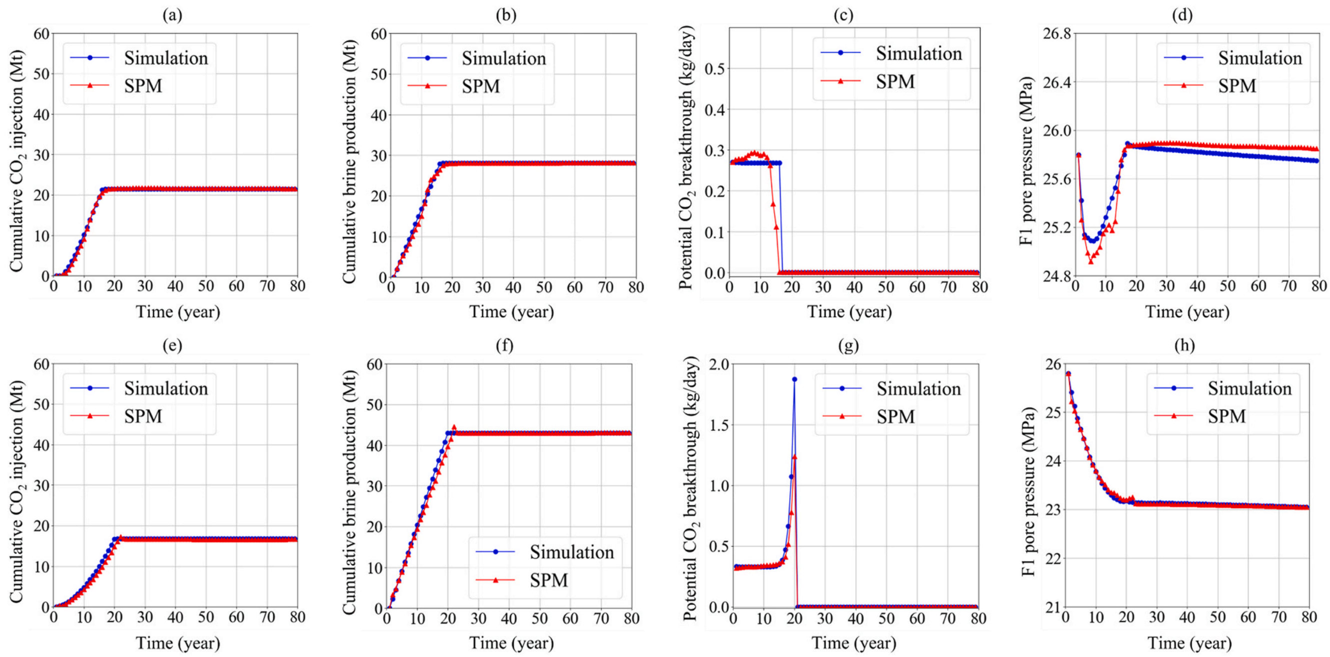

Fig. 14. Results of the blind test for early termination. The blue curves represent the reservoir simulation outputs, and the red curves represent the SPM predictions. (a), (e) Cumulative CO_2 injection (Mt), (b), (f) Cumulative brine production (Mt), (c), (g) Potential CO_2 breakthrough (kg/day) and (d), (h) F1 pore pressure (MPa). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Results of R^2(%) accuracy assessment for blind tests.

	Stable 30-year injection case	CO_2 breakthrough case	Fault instability case
Cumulative CO_2 injection	99.4	98.0	99.6
Cumulative brine production	99.6	99.3	99.3
Potential CO_2 breakthrough	97.4	91.3	90.4
F1 pore pressure	93.8	99.5	96.4

are classified into three categories: i) the stable 30-year injection was achieved (stable 30-year injection case, blue dots), ii) the injected CO_2 was produced (CO_2 produced case, yellow dots), and iii) the fault instability is expected (fault instability case, purple dots). When the stable 30-year injection is achieved, a trade-off relationship between the cumulative CO_2 injection and cumulative brine production is observed. The CO_2-produced cases exhibit high CO_2 injection and brine production, clustering in the lower-right region of the plot. Therefore, the results indicate that even though large brine production would allow to inject large amount of CO_2, it may lead to CO_2 breakthrough to the relief well. On the other hand, the fault instability cases show that relatively low brine production to the CO_2 injection may lead to excessive pressure buildup, which can result in fault instability.

Using NSGA-II, 14 optimal cases with the stable CO_2 injection for 30 years were selected for the Pareto front curve construction as shown in Fig. 16 (b). The blue dots in Fig. 16 (b) indicate the stable CO_2 injection cases, and the red curve depicts the Pareto front constructed with the optimal cases, which have the largest CO_2 injection with given cumulative brine production.

4.2. Analysis of the optimal cases

Fig. 17 and Table 5 present the results of the 14 optimal cases, the cumulative CO_2 injection, cumulative brine production, ratio of cumulative injected CO_2 to produced brine, and the final F1 pore pressure, defined as the pore pressure at the uppermost layer of F1 at the end of the monitoring phase. Case 1 has the lowest cumulative CO_2 injection mass and brine production, 3.89 Mt and 5.12 Mt, respectively, while Case 14 shows the largest cumulative CO_2 injection and brine production of 33.9 Mt and 59.7 Mt. In addition, a clear trade-off trend was observed, as the produced brine mass increased proportionally with the cumulative CO_2 injection for the stable CO_2 injection process for 30 years. From the perspective of the cumulative CO_2/brine ratio (green solid curve in Fig. 17), the values ranged approximately from 0.55 to 0.76. Consequently, it was found that the mass of the brine required to extracted from the reservoir is always larger than that of the injected CO_2 to perform the stable CO_2 storage for the overpressure D structure. Cases 4, 12, 13, and 14 exhibit CO_2/brine ratios below 0.6, indicating relatively low cumulative CO_2 injection compared to the produced brine mass, while the other cases have ratios higher than 0.7. The final pore pressure at the uppermost layer of F1 (black dashed curve in Fig. 17) tends to approach the maximum allowable pressure of 25.8 MPa (red dashed line in Fig. 17) as the CO_2/brine ratio increases. Moreover, it indicates that the cumulative CO_2/brine ratio strongly correlates with the final pressure in the fault. Therefore, the observation implies that the fault stability can be achieved if sufficient brine is removed from the reservoir.

The operating conditions of the 14 optimal cases were categorized into the pre-relief period, well type, and perforation interval. While most cases required more than 24 months of pre-relief period, Cases 7 and 11 show the shorter pre-relief periods of 6 and 7 months, respectively. Therefore, the latter would be economically beneficial in terms of the operational cost before the CO_2 injection process.

The type of wells and their perforation intervals for the optimal cases are described in Fig. 18. The yellow and green bars respectively represent the perforation intervals of D-1 and D-2. The red and blue arrows indicate that the well is adopted as a CO_2 injection or a brine production well, respectively. The results show that D-1 and D-2 wells were chosen as the injection well for 9 and 5 cases, respectively. The perforation intervals of D-1 for the optimal cases are shallower than that of D-2. This is primarily because the penetration depth in the reservoir of D-1 is shallower (1915–2403 m) than that of D-2 (2190–2753 m) (Fig. 2). In

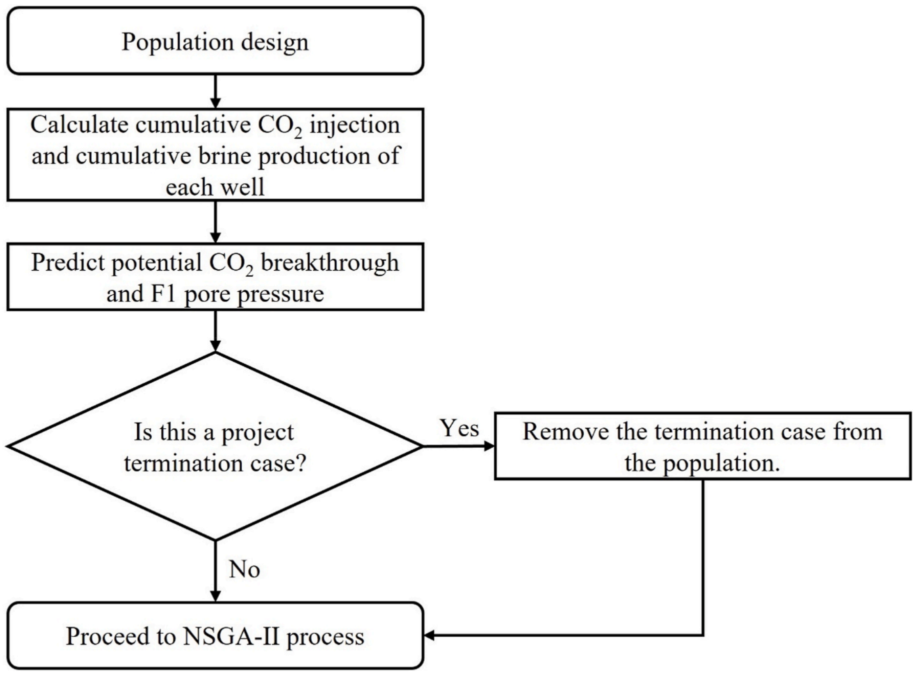

Fig. 15. Workflow chart for NSGA-II calculation process.

Table 4
Parameter settings of NSGA-II.

Parameter	Details
Population Size	200
Generations	100
Selection Method	Tournament selection
Crossover Method	Blended crossover
Crossover Rate	0.5
Mutation Method	Random selection within a range
Mutation Rate	0.2 to 0.05 by generation
Elite Count	20
Random Immigrants	50

addition, D-1 shows a specific trend in the perforation interval. When the well is chosen as the injection well, the perforations are clustered in the upper layers (2000–2200 m). Otherwise, if it is adopted as a relief well, they are clustered in the lower layers (2200–2400 m). However, there is no clear trend in the perforation intervals of D-2.

There are a lot of interesting points for the optimal completion intervals. First of all, based on our basic insight, the perforation intervals for both injection and relief wells should be apart in depth to prevent the early CO_2 breakthrough. This is also because we normally assume the horizontal permeability is much higher than the vertical permeability. Therefore, it was found that there are cases that the perforation intervals of both wells are quite distanced in depth. For example, half of the cases, Cases 1, 4, 7, 9, 10, 13, 14, show that the perforation intervals do not overlap at all. However, the other half, Cases 2, 3, 5, 6, 8, 11, 12, show that the intervals are partially overlapped. This implies that although the locations of the wells are fixed, the optimal perforation intervals

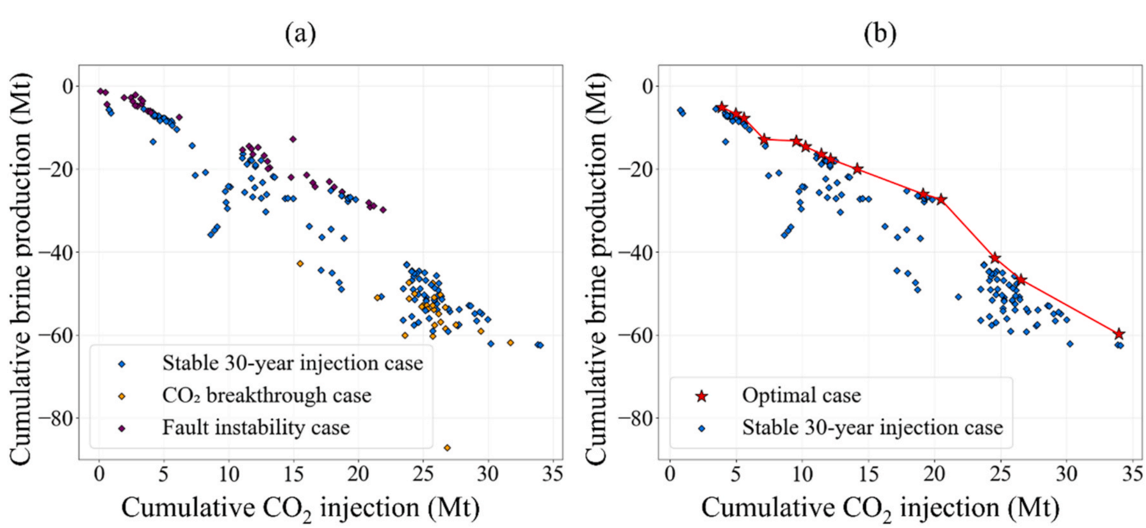

Fig. 16. (a) The cumulative brine production versus cumulative CO_2 injection for 200 random samples generated by SPM. Blue dots represent stable 30-year injections case. Yellow and purple dots represent cases terminated due to CO_2 breakthrough and fault instability, respectively. (b) The Pareto front curve determined by NSGA-II with 14 optimal cases marked by red stars. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

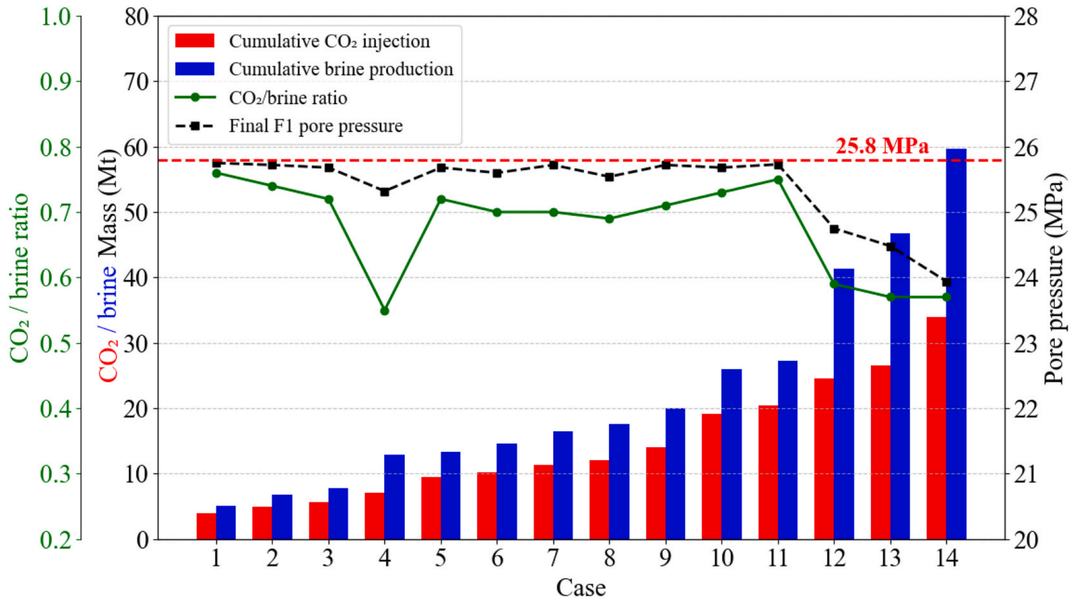

Fig. 17. Results of the 14 optimal cases: Red and blue bars represent the cumulative CO_2 injection and brine production. The green solid curve and black dashed curve respectively indicate the cumulative CO_2/brine ratio and the final F1 pore pressure. The red dashed line specifies the maximum allowable pore pressure. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 5
Results and operation conditions for 14 optimum cases.

Case	Results						Operation conditions			
	Cumulative CO_2 injection (Mt)	Cumulative brine production (Mt)	Average CO_2 injection rate (Mt/year)	Average brine production rate (Mt/year)	CO_2/ brine ratio	Final F1 pore pressure (MPa)	Pre-relief period (month)	Injection well	D-1 perforation interval (m)	D-2 perforation interval (m)
1	3.89	5.12	0.14	0.16	0.76	25.75	24	D-2	2211~2403	2474~2525
2	4.96	6.73	0.18	0.21	0.74	25.72	26	D-2	2211~2403	2302~2582
3	5.58	7.77	0.20	0.24	0.72	25.68	27	D-2	2211~2403	2302~2639
4	7.12	12.9	0.26	0.47	0.55	25.32	25	D-1	1964~2057	2474~2525
5	9.51	13.3	0.33	0.43	0.72	25.68	13	D-2	2310~2354	2302~2696
6	10.2	14.6	0.37	0.45	0.70	25.60	26	D-2	2211~2403	2245~2696
7	11.4	16.4	0.39	0.54	0.70	25.72	6	D-1	2013~2206	2474~2753
8	12.1	17.6	0.45	0.54	0.69	25.54	34	D-2	2211~2403	2302~2410
9	14.1	20.0	0.51	0.62	0.71	25.72	25	D-1	2063~2206	2474~2753
10	19.1	26.0	0.69	0.81	0.73	25.68	26	D-2	2063~2255	2302~2410
11	20.5	27.3	0.70	0.89	0.75	25.73	7	D-2	2260~2354	2302~2410
12	24.5	41.4	0.88	1.29	0.59	24.75	24	D-1	2063~2403	2359~2696
13	26.5	46.7	0.95	1.46	0.57	24.48	25	D-1	2063~2255	2359~2410
14	33.9	59.7	1.22	1.86	0.57	23.94	26	D-2	2211~2403	2417~2467

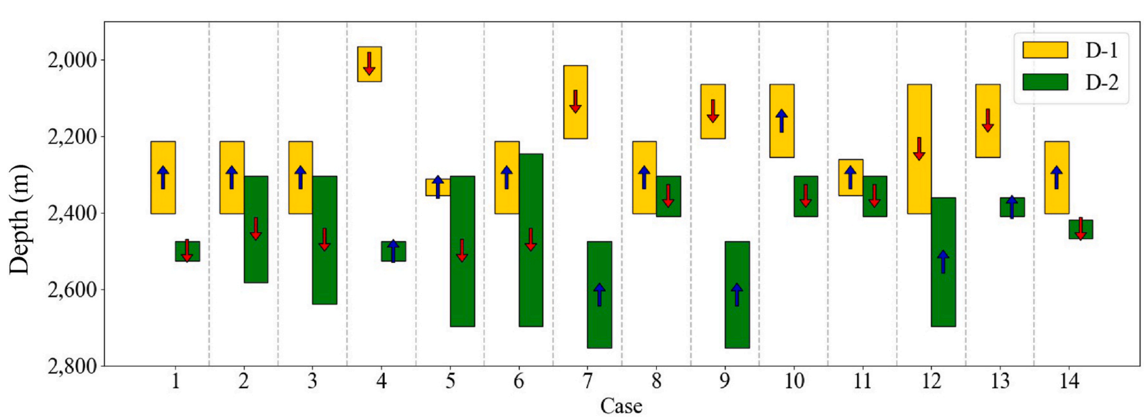

Fig. 18. Purpose of well and perforation interval by optimal case: The yellow bars represent the perforation intervals of D-1, and the green bars represent that of D-2. The red arrows indicate the CO_2 injection, while the blue arrows represent the brine production. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

strongly depend on the other operation conditions, such as pre-relief period. Moreover, although the optimal cases are arranged in order of the cumulative CO_2 injection mass, no specific trend in the perforation intervals is observed. Consequently, it was concluded that the perforation intervals, the depths the CO_2 is injected and the brine is extracted, have to be optimized when a relief well is operated.

4.3. Integrated analysis

Summarizing the results, it was found that the minimum and maximum cumulative CO_2 injection were found in Case 1(3.89 Mt) and 14(33.9 Mt), respectively. In addition, to perform the safe and stable GCS process in the overpressured aquifer, it was found that more brine has to be removed from the reservoir than the injected CO_2, as the mass ratios of the injected CO_2 to the produced brine are in the range between 0.55 and 0.76, and always lower than 1. The ratios also correspond well to the final pore pressure at the top of F1, i.e., the higher the ratio, the lower the final pore pressure at F1. On the other hand, the pre-relief periods of longer than 24 months were required for 11 cases, except for Cases 5, 7, and 11. Especially, only 6 or 7 months of brine extractions need to be performed for Cases 7 and 11 before the CO_2 injection process.

Moreover, the characteristics of the perforation intervals for both wells are quite distinctive. According to the optimized results, it was found that half of the optimal cases show that the intervals of both wells are partially overlapped, while the other half show that there are no overlapped portions in the interval at all. In addition, there is no specific trend in the perforation intervals with the cumulative CO_2 injection or brine production. Consequently, it was concluded that the perforation interval of the injection and relief wells must be optimized when a relief well design is incorporated for a GCS process.

5. Discussion

Optimal CO_2 injection scenarios were derived for the overpressured the D structure, considering the stability of the existing fault. A multi-objective optimization was performed using an SPM coupled with NSGA-II to maximize the CO_2 injection while minimizing the brine production. The optimization workflow and results presented are specific to the D structure. However, when complemented by the discussions and considerations presented in this study, the geomechanical considerations and the ML-based optimization approach can be adapted to guide CO_2 injection design at other geological sites under different site-specific conditions.

5.1. Completion interval for injection and relief well during GCS operation

When a relief well is operated for a GCS process, the optimal completion intervals for both injection and relief wells are always in question. At a glance, injecting CO_2 in the upper portion of the reservoir while extracting brine from the lower portion would make the CO_2 diffusion in the reservoir more efficient, as the injected CO_2 may migrate upward easily with the buoyancy effect. However, the design would facilitate the CO_2 breakthrough in the relief well, which results in the unexpected termination of the GCS process. This is a basic motivation of this study, to optimize the perforation interval of both injection and relief wells. Although the intervals need to be optimized, related previous studies are hardly available.

In order to analyze the effect of the perforation intervals for both wells, a comparison was performed for Cases 1 and 14. Both adopted D-2 and D-1 as the injection and relief wells, respectively. In addition, both have the same perforation intervals for the relief well of 2211–2403 m. The pre-relief periods only have a difference of 2 months (Case 1 has 24 and Case 14 has 26 months). Although the only distinctive difference between the cases is the perforation interval of the injection well (2474-2525 m for Case 1 and 2417-2467 m for Case 14), the results from the cases have significant differences. Case 1 shows the minimum cumulative CO_2 injection of 3.89 Mt and the brine production of 5.12 Mt, while Case 14 has the maximum values of 33.9 and 59.7 Mt. In addition, the ratios of the injected CO_2 to the produced brine for Cases 1 and 14 are 0.76 and 0.57, respectively.

Although unknown patterns may affect the optimized design, it is expected that the permeability heterogeneity is one of the primary reasons for the discrepancy in the results. As shown in Fig. 19, the near-well permeability of the completion intervals of Case 1, 74.4 mD, is higher than that of Case 14, 21.1 mD. Due to the high permeability of the perforation interval of Case 1, it was found that the relief restricted the brine production rate to prevent the CO_2 breakthrough (Fig. 17). On the other hand, the relief well of Case 14 can maintain the high brine production rate with a lower risk of CO_2 breakthrough. As a result, when the injection was stopped, the average pressure values at the perforation interval were 25.1 MPa for Case 1 and 22.5 MPa for Case 14. The higher brine production rate in Case 14 is further reflected in the lower CO_2 to brine ratio compared to Case 1. The CO_2 saturation distribution supports the observations. Fig. 20 describes the CO_2 saturation distribution for (a) Cases 1 and (b) 14, before injection (left), at the end of the injection period (30 years; middle), and at the end of the monitoring period (80 years from the beginning of injection; right). At the end of the injection period, the CO_2 plume is laterally restricted, and it does not reach the relief well. In addition, during the 50-year monitoring phase, the CO_2 plume has moved upward by buoyancy. Therefore, it was concluded that the operating conditions and the subsequent outcome were optimized to prevent the CO_2 breakthrough with the given perforation intervals.

When a GCS design is performed, the locations of the injection and relief wells are normally optimized with the full completion assumption, i.e., the entire penetrated intervals are perforated (Song and Wang, 2021). However, according to the results, it was found that although the locations of the wells are fixed, the optimized perforated intervals vary significantly with operating conditions and the reservoir characteristics. Therefore, if the well locations are not fixed, the perforation intervals would yield more differences in results, and need to be considered during the optimization process.

In the D structure, the perforation interval was identified as a key operating condition that strongly influences GCS design. In this study, the underlying cause was examined by focusing on permeability

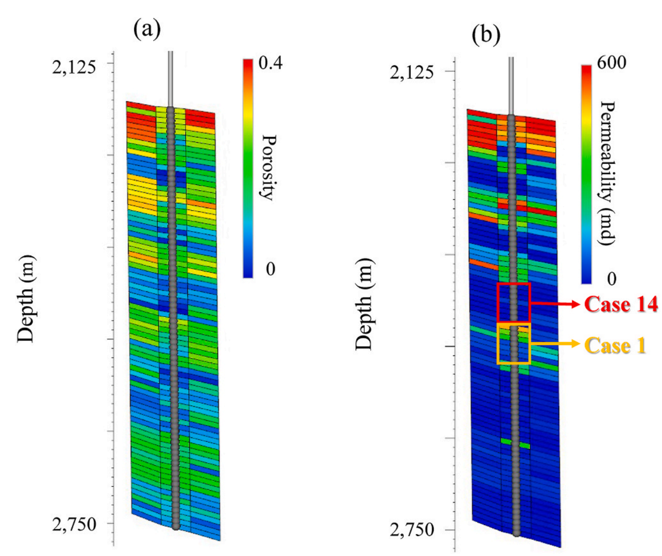

Fig. 19. Porosity and permeability distribution around the perforated grids of D-2, Case 1(yellow) and Case 14(red). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

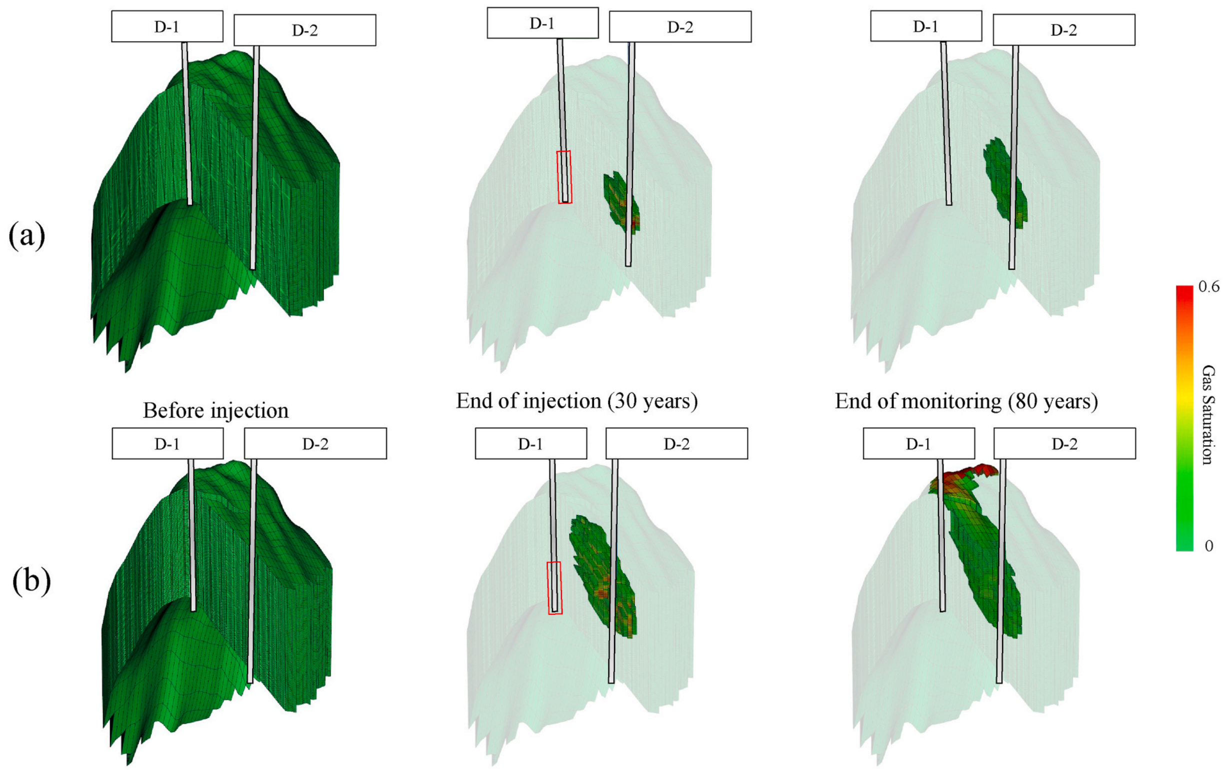

Fig. 20. CO_2 saturation distribution for (a) Cases 1 and (b) 14, before injection (left), at the end of the injection period (30 years; middle), and at the end of the monitoring period (80 years from the beginning of injection; right). The figures are cross-sectional view at the well locations.

heterogeneity; however, additional factors may also affect optimal GCS design. Furthermore, at other storage sites, different operating conditions-such as well location or injection rate-may play a more dominant role. Therefore, future studies should focus on reservoir characteristics and operating parameters that influence the optimization process for each geological storage site.

5.2. Geomechanical stability for GCS

The geomechanical analysis conducted in this study was based on the worst-case scenario for design purposes. For example, it was assumed that the fault plane has the cohesion of zero, the friction coefficient of 0.6, and the geometry most prone to the reactivation. In addition, the analysis focuses on the uppermost portion of the fault where the stress state is closest to the failure envelope. However, it was found that the actual frictional and geometrical characteristics of the fault are expected to be different from the conservative assumptions, as the fault currently remains stable, unlike the results. Therefore, the maximum allowable pressure of the initial pore pressure was selected to perform the optimization process.

The model used in this study was constructed from detailed field-scale measurement as natural gas exploration processes were performed at the reservoir, including seismic exploration, appraisal well drilling, well logging, and core analysis. Therefore, it was concluded that the model is quite reliable. However, reservoir model and characteristics may not be available for some cases. This is because it requires extensive cost to obtain detailed field-scale characteristics for a GCS project. When the model is unavailable, optimization process should perform sensitivity analysis with wide range variables.

According to requirement of the Area of Review (AoR) analysis designated in the Class VI well regulation, it must be re-evaluated at least every five years throughout the project lifetime, including the post-injection site-care period. Since the migration behavior of the injected CO_2 can differ from the initial predictions, consistent monitoring and periodic reporting are required. For example, reservoir and fault heterogeneity can cause unpredicted regional CO_2 accumulation and potential leakage to upper layers. Such heterogeneity can modify multiphase flow behavior and CO_2 plume migration, particularly in low-permeability or structurally complex reservoirs. In these settings, pore-scale trapping, preferential flow paths, and buoyancy-driven redistribution can lead to localized pressure buildup and CO_2 accumulation beyond the initially predicted AoR (Liu et al., 2025; Liu et al., 2025). In addition, pressure and stress alterations by fluid injections directly influence the effective normal and shear stresses on the fault, and possibility of the fault slip. The magnitude and distribution of these perturbations depend on operational variables such as the injection rate, injection period, and perforation intervals. Therefore, the CO_2 injection design must be continuously re-evaluated and optimized-not only during the injection phase but also throughout the post-injection period-to establish a safe and stable GCS strategy.

5.3. Data-driven proxy modeling for GCS

SPM was adopted to optimize the geological CO_2 storage (GCS) design in the D structure in the East Sea. The developed SPM demonstrated high predictive accuracy with fewer simulation runs. The approach can be efficiently incorporated, especially for complex reservoir models that require substantial computing power and time to generate the dataset for traditional proxy model development.

The SPM, specifically tailored to the D structure, effectively utilizes the existing exploration wells in this area. The well locations were fixed as the existing wells are assumed to be used as an injection and a relief well. In addition, the well types can be switched in the simulation to generate training data. As a result, the well-based SPM was designed as a structure-specific model, focusing on reservoir responses around the wells. For more practical applications, if drilling an additional well is planned, the well location optimization needs to be taken into account. Moreover, in reservoirs with multiple faults or complex structural features, the identification of specific weak points can be challenging, in which case an evaluation of pressure distribution changes across the entire reservoir may be required. A 3D grid-based SPM can be adopted for the application, enabling the prediction of pressure and saturation changes at the grid level and providing a more comprehensive description of the entire reservoir (Amiri et al., 2024).

The SPM developed in this study operates as a point-estimate predictor using an autoregressive forecasting scheme, and its effectiveness has been demonstrated in previous research (Vida et al., 2019). However, such autoregressive prediction may lead to the accumulation of prediction errors over time. As a complementary approach, incorporating a Monte Carlo dropout–based uncertainty band can be considered, allowing the SPM to reflect predictive uncertainty (Yarin Gal and Zoubin Ghahramani, 2016).

The proposed SPM based on a 1D-CNN shows strong performance for well-based time-series data with limited training resources. As future work, adopting Physics-Informed Neural Networks (PINNs) can be suggested. It may offer a promising direction for grid-scale predictions, as they can integrate known governing equations and potentially reduce the reliance on extensive reservoir simulations.

6. Conclusion

In this study, CO_2 injection strategies for the overpressured D structure in the East Sea, which contains existing faults, were optimized using an SPM. The optimization objectives were to maximize the cumulative CO_2 storage, minimize the brine production, while maintaining the stability of existing faults, using NSGA-II. The key findings are as follows:

1) Since the D structure is an overpressured aquifer with existing faults, a geomechanical analysis was performed to assess the stability of the existing faults during the GCS process, based on the worst-case scenario for design purposes. It was concluded that the fault is currently stable the actual frictional and geometrical characteristics of the fault are expected to be different from the conservative assumptions. Therefore, it was concluded that the pore pressure at the top of the fault needs to be maintained below the current pressure level, and that a pre-relief period needs to be adopted to reduce the pore pressure before the CO_2 injection.
2) It was assumed that two existing wells in the structure were repurposed as an injection well and a relief well. The injection pressure was maintained lower than 90% of the fracturing pressure. The operation was terminated if CO_2 was detected at the relief well or if the pore pressure at the top of F1 reached its initial value. To construct the optimal injection scenario, designed parameters were considered, such as the pre-relief period, the purposes of the wells, and the perforation interval.
3) A total of 60 reservoir simulations selected using LHS were performed with the software GEM, for 30 years of injection and 50 years of monitoring. Feature Engineering was applied to derive new parameters for training the SPM, with a reduced number of simulation runs. The model exhibited high predictive accuracy in blind testing, achieving average R^2 values of 99.0% for cumulative CO_2 injection, 99.4% for cumulative brine production, 96.6% for pore pressure at the top of F1 and 93.0% for CO_2 production rate.
4) The constructed SPM was employed in the NSGA-II algorithm with a population size of 200 over 100 generations to identify 14 optimal solutions and the Pareto front. As a result, the lowest CO_2 injection mass and brine production are 3.89 Mt and 5.12 Mt, respectively, while the largest cumulative CO_2 injection and brine production are 33.9 Mt and 59.7 Mt, respectively. In addition, the cumulatively produced brine mass increased proportionally with the cumulative CO_2 injection, with the cumulative CO_2/brine ratio between 0.55 and 0.76.
5) It was found that there is no specific trend in the perforation intervals with the cumulative CO_2 injection and brine production. Half of the optimal cases show that the intervals of both wells are partially overlapped, while the other half shows that there are no overlapped portions in the interval at all. In addition, it was found that the optimal intervals strongly depend on both the operation conditions and reservoir characteristics. Consequently, it was concluded that the perforation interval of the injection and relief wells must be optimized when a relief well design is incorporated for a GCS process.
6) The importance of operational conditions, such as the perforation interval, may vary depending on the characteristics of the geological structure. Therefore, these operational parameters should be carefully considered in the optimization process during CO_2 injection design, and the evaluation of geomechanical risks, such as the potential instability of the existing faults, should also be adapted to the specific geological setting. In this study, the initial pressure was imposed as a constraint because the D structure is overpressured. However, for other geological structures, it may be more appropriate to define constraints based on the fault reactivation pressure. In addition, monitoring of pressure and stress not only at localized points but across the entire fault system is necessary to ensure geomechanical stability.
7) Although this study focuses on the optimization of GCS injection scenarios for the D structure, the proposed ML-based optimization framework that explicitly incorporates geomechanical considerations can serve as a useful reference for deriving optimal CO_2 injection strategies in other geological formations.

CRediT authorship contribution statement

Kyuhyun Kim: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Seongjun You:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation. **Dayeon Kim:** Visualization, Validation, Methodology. **Jihoon Wang:** Writing – review & editing, Writing – original draft, Validation, Investigation, Conceptualization.

Declaration of 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.

Acknowledgments

This work was supported by the International Cooperation of the Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Korea government(Ministry of Trade, Industry and Resources, MOTIR) (RS-2025-14383289).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclepro.2026.147737.

Data availability

The data that has been used is confidential.

References

Amiri, B., Jahanbani Ghahfarokhi, A., Rocca, V., Ng, C.S.W., 2024. Optimization of offshore saline aquifer CO2 storage in smeaheia using surrogate reservoir models. Algorithms 17. https://doi.org/10.3390/a17100452.

Bahrami, P., James, L.A., 2023. Screening of waterflooding using smart proxy model coupled with deep convolutional neural network. J. Pet. Sci. Eng. 221, 111300. https://doi.org/10.1016/j.petrol.2022.111300.

Bahrami, Moghaddam, A.J., 2022. A review of proxy modeling highlighting applications for Reservoir Engineering. Geotherm. Power Plants 53–83. https://doi.org/10.1016/b978-185617474-9/50035-2.

Byerlee, J., 1978. Friction of rocks. Pure Appl. Geophys. 116, 615–626.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. 6, 182–197. https://doi.org/10.1109/4235.996017.

Farooq, J., Rafiq, D., Vlachas, P.R., Bazaz, M.A., 2024. RefreshNet: learning multiscale dynamics through hierarchical refreshing. Nonlinear Dyn. 112, 14479–14496. https://doi.org/10.1007/s11071-024-09813-3.

Gal, Yarin, Ghahramani, Zoubin, 2016. Dropout as a bayesian approximation: representing model uncertainty in deep learning. Proc. 33rd Int. Conf. Mach. Learn. 1050–1059. https://doi.org/10.1109/TKDE.2015.2507132.

Global CCS Institute, 2023. Global Status of CCS 2023: Scaling up Through 2030, pp. 1–98.

Global institute, 2022. Global Status of CCS 2022. Glob. CCS Inst.

IPCC, 2022. Global Warming of 1.5°C. Cambridge University Press. https://doi.org/10.1017/9781009157940.

Joshua, Schrier, Norquist, A.J., Buonassisi, Tonio, Brgoch, Jakoah, 2023. In pursuit of the exceptional: research directions for machine learning in chemical and materials science. J. Am. Chem. Soc. 145, 21699–21716. https://doi.org/10.1021/jacs.3c04783.

Lee, T.H., Shair, E.F., Abdullah, A.R., Rahman, K.A., Ali, N.M., Saharuddin, N.Z., Nazmi, N., 2025. Comparative analysis of 1D – CNN, GRU, and LSTM for classifying step duration in elderly and adolescents using computer vision. Int. J. Robot. Control Syst. 5, 426–439. https://doi.org/10.31763/ijrcs.v5i1.1588.

Leng, J., Bump, A., Hosseini, S.A., Meckel, T.A., Wang, Z., Wang, H., 2024. A comprehensive review of efficient capacity estimation for large-scale CO2 geological storage. Gas Sci. Eng. 126, 205339. https://doi.org/10.1016/j.jgsce.2024.205339.

Liu, X., Chen, H., Cheng, W., Li, Y., Zhao, Y., Zhu, Y., Zeng, H., 2025a. Occurrence states and transport behavior of crude oil in different permeability oil reservoirs during depletion development. Geoenergy Sci. Eng. 252, 213944. https://doi.org/10.1016/j.geoen.2025.213944.

Liu, X., Chen, H., Li, Y., Zhu, Y., Liao, H., Zhao, Q., Zhou, X., Zeng, H., 2025b. Oil production characteristics and CO2 storage mechanisms of CO2 flooding in ultra-low permeability sandstone oil reservoirs. Petrol. Explor. Dev. 52, 196–207. https://doi.org/10.1016/S1876-3804(25)60014-0.

Menad, N.A., Noureddine, Z., 2019. An efficient methodology for multi-objective optimization of water alternating CO2 EOR process. J. Taiwan Inst. Chem. Eng. 99, 154–165. https://doi.org/10.1016/j.jtice.2019.03.016.

Mohaghegh, S.D., 2018. Data-driven analytics for the geological storage of CO_2. CRC Press, Boca Raton. https://doi.org/10.1201/b21913.

Mohaghegh, S.D., Modavi, A., Hafez, H., Haajizadeh, M., Guruswamy, S., 2009. Development of Surrogate Reservoir Model (SRM) for fast track analysis of a complex reservoir. Int. J. Oil Gas Coal Technol. 2, 2–23. https://doi.org/10.1504/IJOGCT.2009.023627.

Motaei, E., Ganat, T., 2023. Smart proxy models art and future directions in the oil and gas industry: a review. Geoenergy Sci. Eng. 227, 211918. https://doi.org/10.1016/j.geoen.2023.211918.

Musayev, K., Shin, H., Nguyen-Le, V., 2023. Optimization of CO2 injection and brine production well placement using a genetic algorithm and artificial neural network-based proxy model. Int. J. Greenh. Gas Control 127, 103915. https://doi.org/10.1016/j.ijggc.2023.103915.

Na, Y., Song, Y., Ampomah, W., Wang, J., 2024. Application of smart proxy model for multi-optimization of CO2-EOR design in farnsworth unit with production-injection well conversion. Energy Rep. 12, 3196–3206. https://doi.org/10.1016/j.egyr.2024.09.006.

Song, Y., Wang, J., 2021. Optimization of relief well design using artificial neural network during geological CO2 storage in Pohang Basin, South Korea. Appl. Sci. 11. https://doi.org/10.3390/app11156996.

Song, Y., Jun, S., Na, Y., Kim, K., Jang, Y., Wang, J., 2023. Geomechanical challenges during geological CO2 storage: a review. Chem. Eng. J. 456, 140968. https://doi.org/10.1016/j.cej.2022.140968.

Srinivas, N., Deb, K., 1994. Muiltiobjective optimization using nondominated sorting in genetic algorithms. Evol. Comput. 2, 221–248. https://doi.org/10.1162/evco.1994.2.3.221.

Sun, Q., Ampomah, W., You, J., Cather, M., Balch, R., 2021. Practical co2—wag field operational designs using hybrid numerical-machine-learning approaches. Energies 14. https://doi.org/10.3390/en14041055.

US EPA, 2012. Class VI Well Construction Guidance.

Verma, S., Pant, M., Snasel, V., 2021. A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems. IEEE Access 9, 57757–57791. https://doi.org/10.1109/ACCESS.2021.3070634.

Vida, G., Shahab, M.D., Mohammad, M., 2019. Smart proxy modeling of SACROC CO2-EOR. Fluids 4. https://doi.org/10.3390/fluids4020085.

Vo-Thanh, H., Amar, M.N., Lee, K.K., 2022. Robust machine learning models of carbon dioxide trapping indexes at geological storage sites. Fuel 316, 123391. https://doi.org/10.1016/j.fuel.2022.123391.

Wang, M., 2022. Prediction of the technology company's stock price through the deep learning method. Open J. Model. Simulat. 10, 428–440. https://doi.org/10.4236/ojmsi.2022.104024.