Petroleum Science 23 (2026) 4441-4460

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Coupled optimization framework of CO_2-WAG injection for enhanced oil recovery and storage: Integrating MLP with evolutionary and metaheuristic algorithms

Shahab Ud Din [a,b], Liang Xue [a,b,*], Fu-Long Ning [c], Dong-Dong Guo [d], Qin-Zhuo Liao [a], Hussan Zeb [e]

[a] State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[b] Department of Oil-Gas Field Development Engineering, College of Petroleum Engineering, China University of Petroleum (Beijing), Beijing, 102249, China
[c] Engineering Research Center of Rock-Soil Drilling & Excavation and Protection, Ministry of Education, China University of Geosciences, Wuhan, 430074, Hubei, China
[d] School of Earth and Environment, Anhui University of Science and Technology, Huainan, 232001, Anhui, China
[e] University of Rome Tor Vergata, Via Montpellier, 1-00133 Rome, Italy

A R T I C L E I N F O

Article history:
Received 1 June 2025
Received in revised form
16 October 2025
Accepted 18 January 2026
Available online 23 January 2026

Edited by Min Li

Keywords:
CO_2 sequestration
CO_2-WAG
Deep neural network
Coupled optimization
Reptile search algorithm

A B S T R A C T

This study introduces a machine-learning-driven workflow to optimize CO_2-WAG by considering both oil production and CO_2 sequestration as the output parameters. A compositional simulator is employed to create a baseline model by using data from the Bell Creek Oilfield. Computational models, specifically multi-layer perceptron (MLP) combined with evolutionary and metaheuristic algorithms, such as the genetic algorithm (GA), particle swarm optimization (PSO), and reptile search algorithm (RSA), are utilized to develop proxy models for predicting cumulative oil production and CO_2 storage. The suggested workflow integrates the optimization algorithms GA, PSO, and RSA with MLP-based proxy models (MLP-GA, MLP-PSO, and MLP-RSA) to enhance cumulative oil production and CO_2 storage. The findings demonstrated the reliability and efficient utilization of the proposed workflow for CO_2-WAG coupled optimization. All proposed models surpassed the base model; among them, MLP-RSA outperformed MLP-PSO and MLP-GA in terms of prediction accuracy and optimization performance. In comparison to the base case, the optimal scenarios increased the cumulative oil production by 10.65% and CO_2 storage by 24.72%. Among the optimization algorithms, RSA outperformed the others in terms of speed, processing, and yielded an incremental oil revenue of 15.31 million USD. The results of this study offered new insights into optimization techniques and decision-making processes for CO_2-WAG projects. Moreover, it provides engineers with a tool to conduct robust and effective optimization studies, avoiding tedious conventional optimization techniques.

1. Introduction

Fossil fuels, particularly oil, are still the primary energy source to meet growing demand. A large amount of oil is left behind during the production phase; specialized techniques are needed to increase recovery efficiency. However, fossil fuels also lead to CO_2 emissions, an immense challenge facing by the world. Carbon capture, utilization, and storage (CCUS) is a vital technology for coping with increasing CO_2 emissions (Bachu, 2001; Mac Dowell et al., 2017). CO_2-EOR is a CCUS method that can be used to address these challenges (Czernichowski et al., 2018; Dziejarski et al., 2023; Iskandar and Kurihara, 2022; Kolawole et al., 2021), not only to enhance oil recovery but also to sequester CO_2 in the subsurface (Dai et al., 2016, 2018).

The utilization of CO_2 injection to enhance oil recovery is because of its important and unique characteristics to improve fluid properties in subsurface conditions, such as the reduction of oil viscosity, oil swelling, hydrocarbon extraction, and the

* Corresponding author.
E-mail address: xueliang@cup.edu.cn (L. Xue).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.01.026
1995-8226/

reduction of interfacial tension (Lake et al., 2014; Huang et al., 2022; Jiang et al., 2019; Kong et al., 2019), that leads to enhance oil production. CO_2 typically exhibits a lower minimum miscibility pressure (MMP) with crude oil compared to other gases like nitrogen or methane. Thus, utilizing the injection of CO_2 to improve oil recovery can result in improved miscibility and effective oil recovery (Tang et al., 2021). However, continuous CO_2 injection has some challenges, such as gas overriding or fingering, which lead to early gas breakthrough. To address these challenges, another technique proposed as the CO_2-WAG method. This approach merges the attributes of gas flooding and water flooding and entails the injection of gas slug into the reservoir in conjunction with the alternating water injection. The macroscopic sweep efficiency is increased by water injection, and the microscopic displacement efficiency is improved by gas injection (Al-Bayati et al., 2018; Christensen et al., 2001; Karimaie et al., 2017; Khather et al., 2022; Kulkarni and Rao, 2005; Ren et al., 2023; Sun et al., 2021). Additionally, CO_2-WAG can address the issue of accelerated CO_2 flow and enhance gas-phase flow resistance (Pancholi et al., 2020; Ren and Duncan, 2021). It also enhanced the mobility ratio and decreased the resistance to water flow, which led to a significant improvement in recovery. According to previous reports, the WAG technique has been successfully employed in 80% of US fields and has immensely enhanced oil recovery (Sanchez, 1999).

Recently, the optimization of CO_2-WAG EOR technology has become a hot topic among academia and research organizations because of its potential to enhance oil recovery and sequester CO_2 (Rodrigues et al., 2019; Sanchez, 1999). Mostly, conventional optimization techniques have been used to optimize the CO_2-WAG project performance. Typically, compositional or numerical simulations are used to adjust parameters such as injection rates, WAG ratios, cycle lengths and bottom hole pressures, aiming to maximize oil recovery and CO_2 storage (Rodrigues et al., 2019; Kashkooli et al., 2022; Vo Thanh et al., 2020). For instance, numerical simulations were used for coupled optimization by dynamically adjusting well setting for both oil production and CO_2 storage (Kashkooli et al., 2022). A study conducted on the Jacksongburg-Stringtown Oilfield by employing compositional simulation deduces the importance of CO_2-WAG rates on oil recovery and CO_2 sequestration (Zhong et al., 2019). Another study investigated CO_2-WAG performance in tight reservoirs by employing numerical simulation to optimize operational parameters (Ghaderi et al., 2012). Similarly, a commercial numerical simulation module, such as CMG-CMOST was adopted to optimize CO_2 storage in heterogenous reservoirs (Vo Thanh et al., 2020). However, these conventional optimization techniques are laborious, time-consuming, and require high computational power. Fortunately, with the advent of advanced machine learning techniques and metaheuristic algorithms, oilfield complexities and uncertainties can be tackled, enhancing optimization efficiency and accuracy (He et al., 2021; Li et al., 2022; Sen et al., 2022; Wang et al., 2023; Xue et al., 2023a, 2023b, 2024). Machine learning techniques have penetrated every field and have various applications in the oil and gas industry.

To rapidly navigate the solution space to identify the global optimum solution, several researchers have developed machine-learning techniques with metaheuristic algorithms. Capitalizing on the predictive capabilities of machine learning to direct the search trajectory of metaheuristic algorithms enables rapid identification of optimal solutions, thereby yielding superior outcomes in the realm of parameter optimization. Mohagheghian et al. (2018) used an evolutionary algorithm to enhance the performance of the WAG technique in the Norne oilfield E segment. The objective function employed was NPV, with GA and PSO utilized as search strategies. The optimization process involved parameters such as the injection volumes, WAG ratio, length, bottom hole pressure, and hydrocarbon gas fraction. You et al. (2020) used a Gaussian-SVR with a Gaussian kernel to build proxy models and optimize its hyperparameters via Bayesian optimization. This model, combined with the MOPSO protocol, was employed to enhance the intricate CO_2-WAG technique. The optimization parameters included the injection rates, bottom-hole pressure, and WAG cycle duration. Jaber. (2022) used a GA technique-based proxy model to optimize key variables, such as the WAG ratio, size, length, and bottom hole pressure in the CO_2-WAG technique for the Subba-Nahr Umr reservoir. Some other studies utilized various input parameters and algorithms for CO_2-WAG optimization (Matthew et al., 2023; Nait Amar et al., 2021; Ng et al., 2021, 2023).

From the aforementioned studies, it is evident that prior studies on CO_2-WAG EOR optimization generally relied on comparatively small datasets and considered a single objective function, either CO_2 storage or oil production. Some studies used both CO_2 storage and cumulative oil production; however, they relied on a few input parameters. The absence of a large number of input parameters and extensive time-series data limits the full potential of machine learning techniques, and it is challenging to achieve global optimization results. Moreover, prior studies did not consider all trapping mechanisms. In contrast, this research study utilized a larger dataset comprising nine input parameters, two output parameters and incorporated all major CO_2 trapping mechanisms, including mineralization. Therefore, the research fills the gap and provides a more holistic approach that includes comprehensive time-series data, multiple trapping mechanisms and a large number of variables that will also enhance machine learning techniques predictive capability in achieving global optimization. Additionally, prior studies relied on GA and PSO for CO_2-WAG optimization problems. However, these algorithms have limitations such as premature convergence and stagnation when dealing with complex optimization problems.

To overcome these challenges, this study utilized a novel machine learning technique for CO_2-WAG coupled optimization based on the reptile search algorithm (RSA). To the best of our knowledge, this study utilized RSA for the first time for CO_2-WAG coupled optimization. In the absence of real field data, the study used numerical simulation by employing CMG-GEM to obtain a large dataset for the proposed machine learning study based on a conceptual model of the Bell Creek Oilfield. Furthermore, the study employed a robust deep neural network, such as a multi-layer perceptron (MLP), for the construction of proxy models, which considerably increases the computational speed of the simulated data. A novel technique is used to employ machine learning algorithms for hyperparameter tuning to obtain the best architecture. After the realization of the proxy models, the most widely used optimization algorithms, such as the GA, PSO, and the newly developed RSA, were used to obtain the optimal solution for the desired objective functions. This method considerably enhances the effectiveness of the optimization scheme and provides engineers with insights and a robust co-optimization scheme.

2. Methodology

2.1. Base model

This study considers the Bell Creek oil field for a CO_2-WAG optimization study, which is located in the Powder River Basin, Montana, and has been producing for more than 55 years and is now in the tertiary phase (Kurz et al., 2013). A representative model was constructed by utilizing the CMG-GEM module of the CMG simulator specialized for compositional studies, as shown in

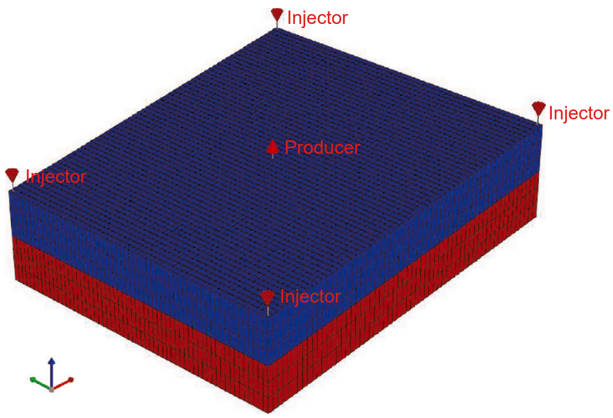

Fig. 1. Schematic of the reservoir permeability model.

Table 2
Model characteristics.

Properties	Values
Depth	4400 ft
Reservoir thickness	60 ft
Dip angle	1° (northwest)
Porosity	Siltstone 0.15, Sandstone 0.25
Permeability	Siltstone 50, Sandstone 700 md
Reservoir pressure	16.55 MPa
Temperature	107 °F
Production bottom hole pressure	8.27 MPa
API gravity	37°
Salinity	5000 ppm
Oil viscosity	2.2 cP
Saturation pressure	6.38 MPa
MMP	9.73 MPa
Equation of state (EOS)	Peng Robinson
Injection and production	20 years
CO_2 storage	100 years

Fig. 1. The simulation model in a Cartesian grid system containing $59 \times 59 \times 6 = 20886$ grid blocks was developed. A dip angle of 1° was applied in the model to reflect the natural structural inclination of the formation. The model incorporates a five-spot injection pattern, a central well designed for production, and four injection wells in the corners for alternate water and CO_2 injection (Jin et al., 2018).

All the key reservoir parameters utilized in the model are listed in Table 1. The reservoir pressure and temperature conditions placed the model above a minimum miscibility pressure of 9.73 MPa. The reservoir temperature was recorded as 107 °F (Kurz et al., 2013). The upper siltstone has a porosity and permeability of 0.15 and 50 md, whereas the lower sandstone has a porosity and permeability of 0.25 and 700 md (adjusted for simplicity). The model considers all trapping mechanisms, such as structural, residual, solubility, and mineralization. For residual trapping, the study utilized a CMG-GEM in-built linear phase permeability hysteresis model ($S_{gr-max} = 0.30$) (Land, 1968). The details of the fluid model utilized, and relative permeability curves are given in Table 2 and Fig. 2, respectively. The PVT model comprised of seven components: CO_2, N_2–C_2, C_3–C_4, C_5–C_7, C_8–C_{13}, C_{14}–C_{24}, C_{25}–C_{36+}, which were lumped by lumping techniques. Peng-Robinson equation of state (EOS) was employed for model fine-tuning the fluid model utilizing the CMG Winprop module. To ensure that the model accurately captures reservoir fluid behavior, experimental data such as saturation pressure, swelling tests, constant composition expansion, and differential liberation test were matched at reservoir conditions.

The base model simulates a total of 20 years for cumulative oil production and 100 years for CO_2 storage, where CO_2-WAG injection (for 20 years) commences from the beginning as the model is assumed to be depleted ($S_w = 0.58$). Having a WAG-cycle size (total of 40 cycles) of 180 days (90 days each) and a WAG ratio of 1:1. The CO_2-WAG injection began with CO_2 injection followed by water injection, at an injection rate of 0.115×10^3 tons/day and 1730 bbl/day, respectively. The bottom hole pressure of the production well was 8.27 MPa, and the injection pressure limit was set at 20 MPa. The injection rates and production results of cumulative oil production and CO_2 storage are shown in Figs. 3 and 4 shows storage by different mechanisms.

Table 1
Fluid composition.

Components	Mole fraction
CO_2	0.0042
N_2	0.0019
CH_4	0.1909
C_2H_6	0.0033
C_3H–NC_4	0.0428
iC_5–C_7	0.1526
C_8–C_{13}	0.2860
C_{14}–C_{24}	0.1997
C_{25}–C_{36+}	0.1184

2.2. Database creation

To capture the intricacies of CO_2-WAG, this study thoroughly investigated nine input parameters: geological, fluid, injection, and production parameters. The geological parameters include the initial reservoir pressure, which is an important parameter to consider when conducting a CO_2-WAG study. The design parameters ensure the maintenance of the reservoir pressure. The fluid properties included oil saturation, water saturation, pH, and salinity. Water saturation has a significant effect on oil production; as water saturation increases, it decreases oil saturation and leads to lower oil production. High water saturation also mimics the advanced stage of production. Similarly, pH is also an important parameter when considering CO_2 sequestration (mineralization), which affects the mineralization reaction of CO_2 with rock minerals (Christensen et al., 2021; Delerce et al., 2023; Wan et al., 2017; Oskouei et al., 2012). The model considers the mineralization reactions of anorthite, calcite, and kaolinite. Another crucial factor affecting CO_2 dissolution is salinity, and various studies have demonstrated its importance (Forray et al., 2021; Kumar et al., 2020). This model achieves brine salinity by considering the molar mass concentrations of Na^+ and Cl^- ions in aqueous conditions. Similarly, the injection and production parameters include the bottom hole pressure (both injection and production wells) and gas and water injection rates. A commercial simulator, CMG-GEM was used to randomly generate cumulative oil and CO_2 storage for a dataset of 700 cases which is based on different input parameters such as pH, salinity, water saturation, oil saturation, average reservoir pressure, BHP production, BHP injection wells, water injection rate, and gas injection rate. The range of values for all parameters and their distribution are given in Table 3.

3. Machine learning techniques

3.1. Proxy models construction

Advancements in computational technology have broadened the possibilities of reservoir modeling. However, limited

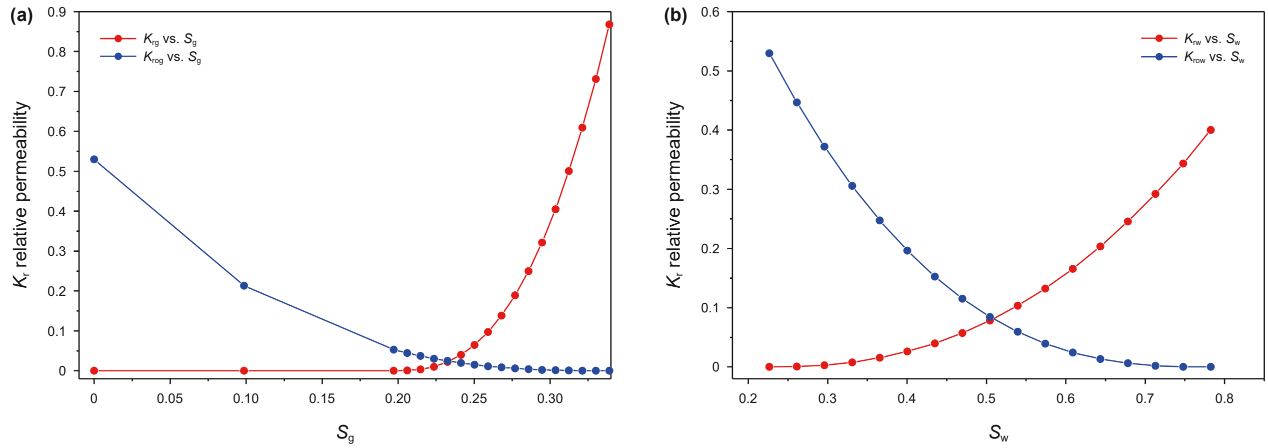

Fig. 2. Base model relative permeability curves. **(a)** Gas saturation (K_r vs. S_g) and **(b)** water saturation (K_r vs. S_w).

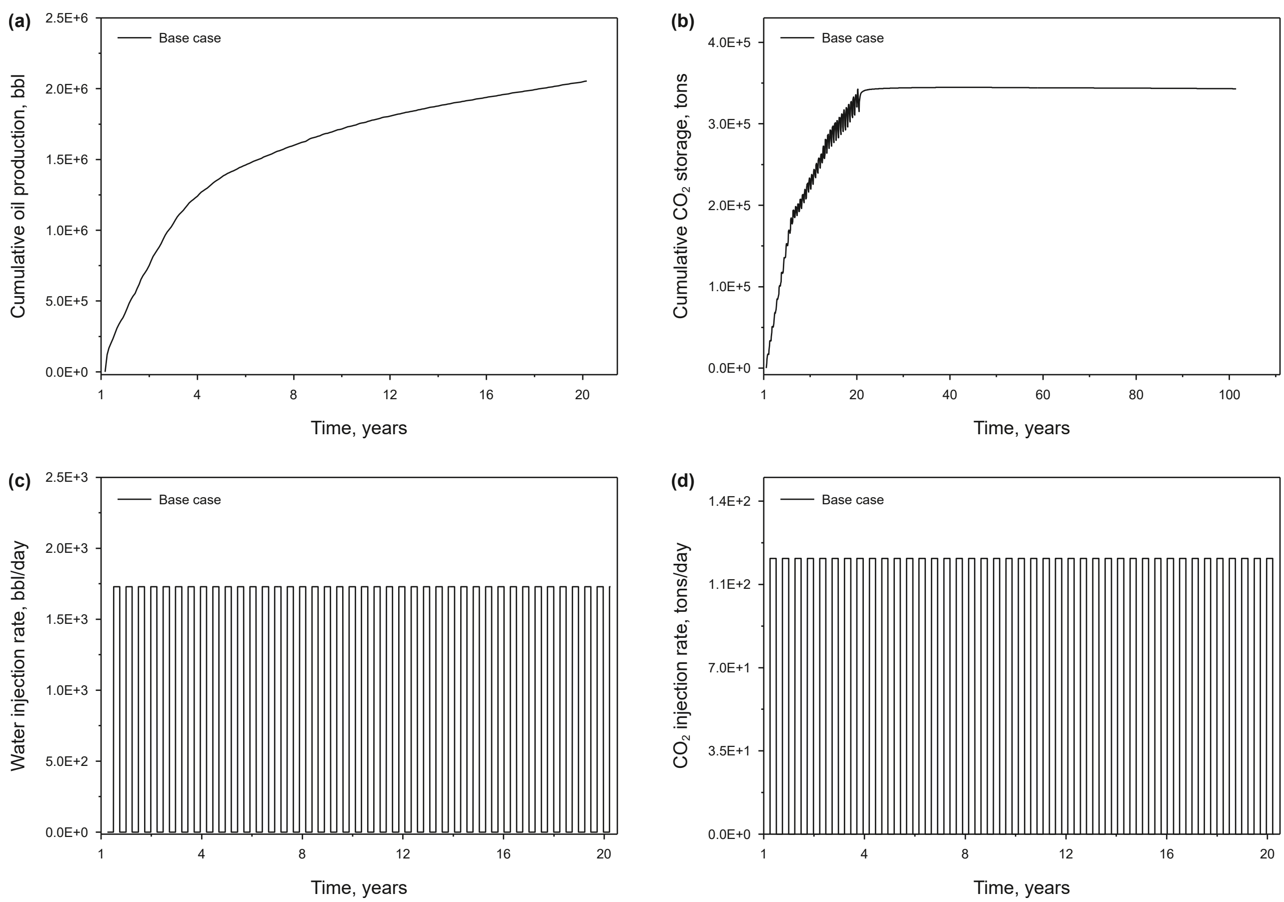

Fig. 3. Base model oil production, storage, and injection rates plots. **(a)** Cumulative oil production, **(b)** CO_2 storage, **(c)** water injection rate, and **(d)** CO_2 injection rate.

computing resources remain a challenge, particularly for uncertainty quantification and optimization workflows (Zubarev, 2009). To address this issue, the industry is turning to computationally efficient proxy models and high-performance computing techniques. The proxy model is succinctly characterized as a mathematical or statistical function designed to emulate the outcomes of intricate simulation models by employing a set of predetermined input parameters, also known as surrogate models. These models serve as proficient substitutes for more complex models, ensuring computational efficiency. Proxy modeling is considered highly effective for production optimization projects and can significantly reduce computation time and cost. In the oil and gas sector, different types of proxy models have been used for different optimization studies. This study constructed and employed two proxy models for cumulative oil production (for 20 years) and CO_2 storage (for 100 years). The main advantage of this method is its flexibility, which allows engineers to set any suitable time node as the optimization objective based on extensive statistical analysis. This ensures a robust workflow for co-optimization, enhancing the accuracy and reliability of reservoir simulation studies. This study

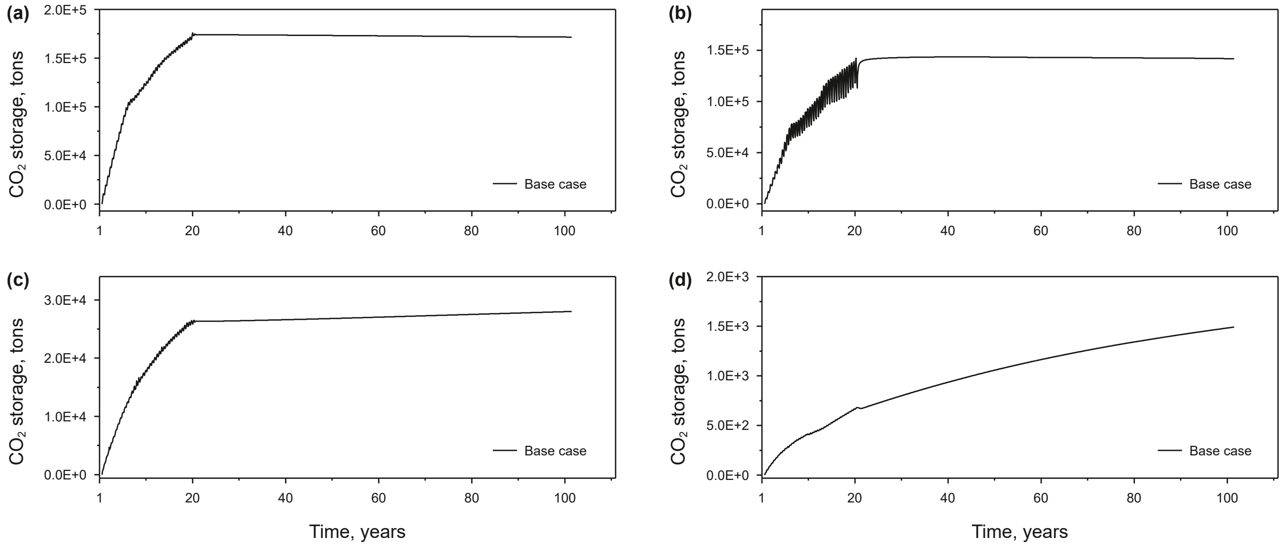

Fig. 4. Base case CO_2 trapping by different mechanisms. **(a)** Stratigraphic trapping, **(b)** residual trapping, **(c)** solubility trapping, and **(d)** mineralization trapping.

Table 3
Basic parameters and range for database creation.

Parameters	Min	Max
Water injection rate, $\times$ 10^4 bbl/day	0.52	2.2
Gas injection rate, $\times$ 10^3 tons/day	0.1	0.5
BHP production, MPa	5.5	9.6
BHP injection, MPa	10.3	22.7
Average pressure, MPa	9.65	16.5
S_o	0.2	0.9
S_w	0.2	0.8
pH	5.5	8.5
Salinity	2	200

employed a multi-layer perceptron (MLP) to construct two separate proxy models for cumulative oil production and CO_2 storage.

MLP is a type of artificial neural network (ANN) that includes more hidden layers than traditional ANN. A typical ANN model includes an input layer, hidden layer, and an output layer, where the MLP includes multiple hidden layers. The input layer is designed to receive and accommodate data of varying dimensions. The hidden layers are generally arranged in numerous tiers and execute fundamental computation by adding weights to the input data and transmitting it through an activation function, thereby generating intricate representations of the data (Goodfellow, 2016; Mohanasundaram et al., 2019; Subasi, 2020). The activation function creates nonlinearity, which allows the network to represent complex data patterns. These layers can have a number of neurons that learn the diverse characteristics and features of the input data to learn high-level abstractions. The output layer consolidates the information acquired from the hidden layers and employs a task-specific activation function, such as softmax for classification or linear activation for regression. This yields the final output, which is applicable for predictions or classifications contingent upon the specific problem being addressed. MLPs are proficient in discerning intricate patterns compared with traditional ANN networks.

This study utilized a sample set of nine input parameters and two output parameters for training and testing the proxy models. The dataset input parameters are oil saturation, water saturation, pH, salinity, average reservoir pressure, bottom hole pressure of the injection well, bottom hole pressure of the production well, water injection rates, and gas injection rates. The dataset output parameters are cumulative oil production and CO_2 storage. Initially, the data was carefully inspected for outliers. Subsequently, it was split, with 75% allocated for training and 25% allocated for testing. Feature scaling was then employed to normalize the data, ensuring that it fell within a specific range, typically from 0 to 1. This process ensures that the data is prepared and optimized for further analysis or modeling.

Upon completion of the initial MLP model, evolutionary algorithms, such as the RSA, PSO, and GA were employed to optimize the architecture of the proxy models. Using these algorithms, key parameters such as the number of layers, neurons, batch size, and epochs were optimized, additional details can be found in Emam et al. (2023), Nikbakht et al. (2021), and Qolomany et al. (2017) work. Fig. 5 shows the detailed workflow, and Table 4 summarizes the optimized design of the proxy models used in this study.

Following architecture optimization, the training of the proxy model commenced. To evaluate the precision of the predicted variable, cross-plots comparing the actual values to MLP predicted values were assessed. The effectiveness of the machine learning model and its overfitting and underfitting can be verified by the R^2 plots, mean square error (MSE), and mean absolute relative error (MARE). This can be calculated using the following Eqs. (1)–(3):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i,\mathrm{act}} - y_{i,\mathrm{pred}}\right)^2}{\sum_{i=1}^{n}\left(y_{i,\mathrm{act}} - \overline{y}_{i,\mathrm{act}}\right)^2} \tag{1}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_{i,\mathrm{act}} - y_{i,\mathrm{pred}}\right)^2 \tag{2}$$

$$\mathrm{MARE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_{i,\mathrm{act}} - y_{i,\mathrm{pred}}}{y_{i,\mathrm{act}}}\right| \tag{3}$$

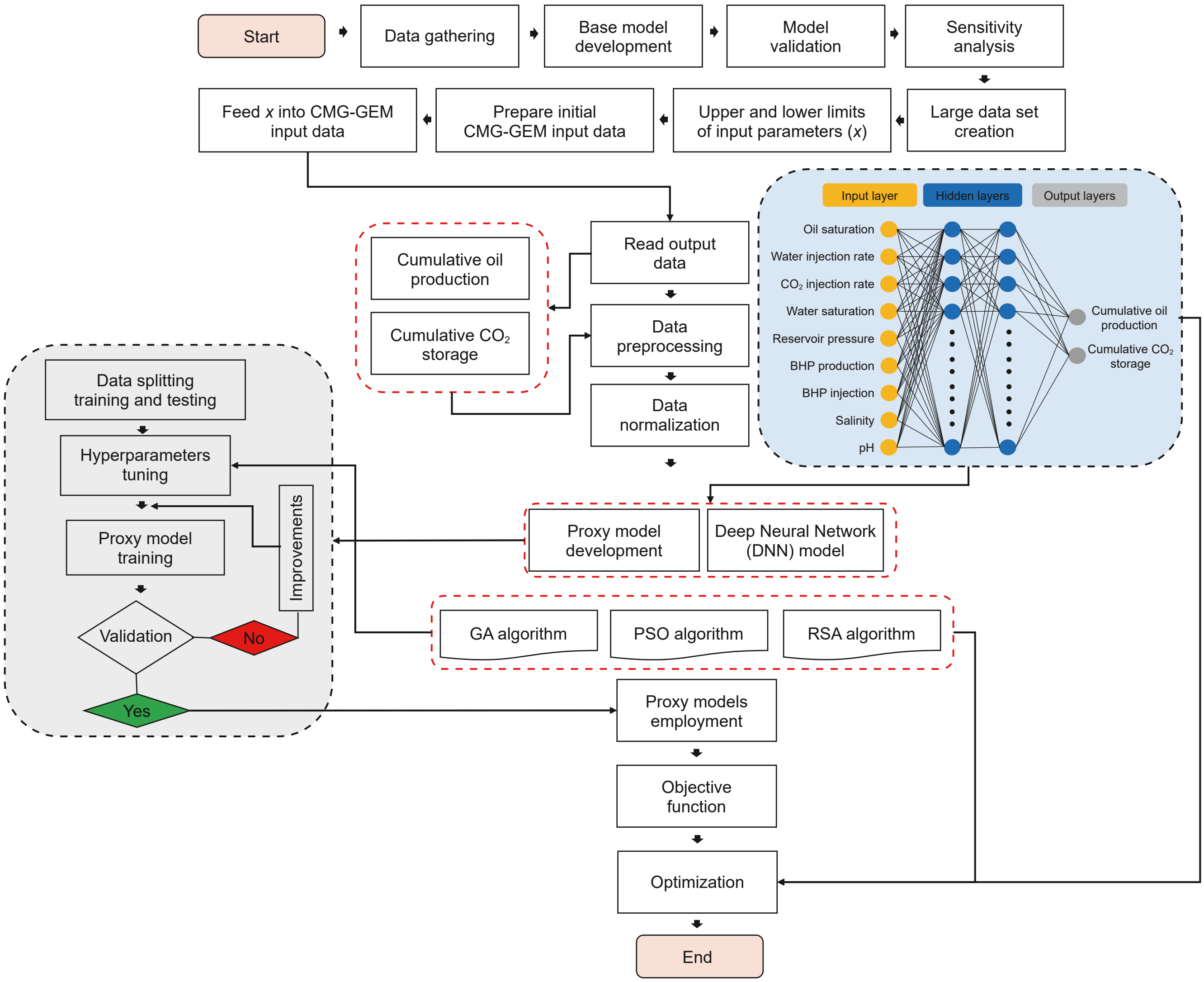

Fig. 5. The flowchart of the proposed optimization framework. Starting with base model development, it is used with deep neural networks, such as a multi-layer perceptron (MLP), to develop proxy models for the optimization step. GA, PSO, and RSA algorithms are utilized to get the optimal solutions for cumulative oil production and CO_2 storage.

Table 4
MLP-GA, MLP-PSO, and MLP-RSA optimized models parameter architecture.

Parameters	Optimal parameters		
	GA	PSO	RSA
Number of layers	3	3	4
Neurons	600, 400, 200	442, 456, 451	88, 507, 413, 451
Activation function	ReLU	ReLU	ReLU
Epochs	150	180	190
Batch size	128	68	128

where $y_{i,\text{act}}$ represents the actual values used for training purposes and $y_{i,\text{pred}}$ represent the values predicted by the proxy model.

Both proxy models, such as cumulative oil production and cumulative storage, are fine-tuned by the RSA, PSO, and GA algorithms to obtain the best representative proxy models, as shown in Fig. 6. The best representative models were based on high R^2 values and lower MSE and MARE values. The MSE of the loss function plays an important role in evaluating the accuracy and precision of the model while preventing model overfitting. Once the proxy models were accepted, they were integrated for the optimization step to predict the output parameters.

3.2. Objective function

For the co-optimization, the study utilized two objective functions, cumulative oil and CO_2 storage, as both are dynamic properties, so a dynamic weight coefficient is set as follows:

3.2.1. Cumulative oil production

In CO_2-WAG EOR, the main goal is to enhance the long-term cumulative oil production. The study utilized dimensionless oil production to eliminate the impact of the unit on the coupled optimization and can be represented as follows (Xue et al., 2025):

$$N_{Pi}^{t} = \frac{P_i^t - \frac{1}{N}\sum_{i}^{N} P_i^t}{\sqrt{\frac{1}{N-1}\sum_{i}^{N}\left(P_i^t - \frac{1}{N}\sum_{i}^{N} P_i^t\right)^2}} \tag{4}$$

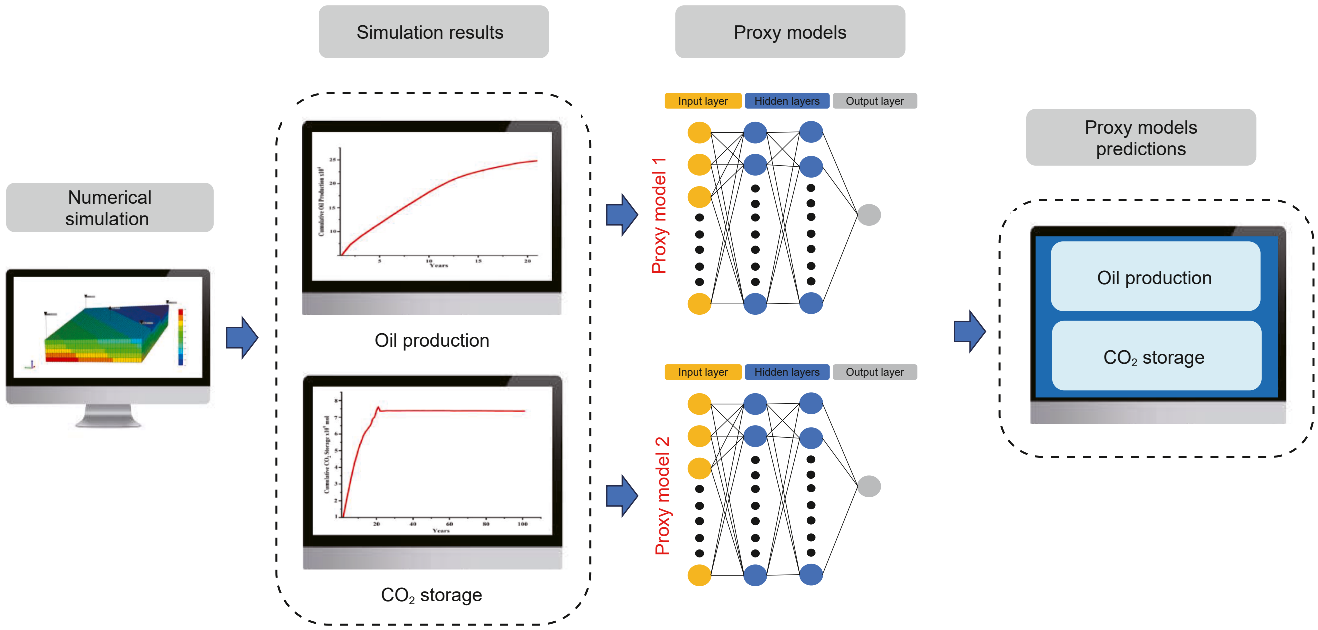

Fig. 6. The training method employed in this study. Proxy 1 is used for cumulative oil production, and Proxy 2 for cumulative CO_2 storage. Both are then utilized by the optimization algorithms GA, PSO, and RSA for the optimization process.

where N_{Pi}^t denotes dimensionless cumulative oil production (of $i-$th sample at time step t), P_i^t denotes the cumulative oil production (of $i-$th sample at time step t), N represents the sample number, and t is the time step.

3.2.2. CO_2 sequestration

In CO_2-WAG EOR, another important parameter is CO_2 storage. The dimensionless CO_2 sequestration formula is given as:

$$N_{Si}^t = \frac{S_i^t - \frac{1}{N}\sum_i^N S_i^t}{\sqrt{\frac{1}{N-1}\sum_i^N\left(S_i^t - \frac{1}{N}\sum_i^N S_i^t\right)^2}} \tag{5}$$

where N_{Si}^t denotes dimensionless CO_2 storage (of $i-$th sample at time step t), S_i^t denotes CO_2 storage (of $i-$th sample at time step t), N is the sum of the samples, and t is the time step.

3.2.3. Coupled optimization

The study utilized both cumulative oil production and CO_2 storage for the optimized objective function, as both objective functions are dynamic, so a dynamic weight coefficient will be employed, and the objective formula for both functions is given as follows (De Weck and Kim, 2004):

$$F = w \times N_S^{t1} + (1-w) \times N_P^{t2} \tag{6}$$

where F is the objective function, w is the weight coefficient ($w = 0.5$, to give equal importance to both objective functions), N_S^{t1} and N_P^{t2} denote the dimensionless CO_2 storage and cumulative oil production at the $i-$th time step t, respectively.

3.3. Optimization

For the optimization, this study utilized three algorithms: the GA, PSO, and RSA. A detailed introduction of the RSA algorithm is presented here, and the details of the GA and PSO algorithms can be found in the supplementary data (Section 2).

3.3.1. Reptile search algorithm (RSA)

The Reptile search algorithm (RSA) introduced by Abualigah et al. (2022), is a nature-inspired metaheuristic algorithm that imitates the foraging behavior of reptiles. Hunting behavior of crocodiles, such as the encircling and hunting phases. The encircling phase is used for exploration (global search), and the hunting phase for exploitation (local search), as depicted in Fig. 7. It combines exploration and exploitation using adaptive position updates. The first step of the RSA optimization process starts with initialization by generating a set of random solutions (X), as given in Eq. (7), where the best solution obtained in each iteration is considered as an optimum solution.

$$\mathbf{X} = \begin{bmatrix} X_{1,1} & \cdots & \cdots & X_{1,n-1} & X_{1,n} \\ X_{2,1} & \cdots & \cdots & X_{2,n-1} & X_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{N,1} & \cdots & \cdots & X_{N,n-1} & X_{N,n} \end{bmatrix} \tag{7}$$

where, $X_{i,j}$ represents the solution of $i-$th candidate at the $j-$th position and can be calculated using Eq. (8). N and n represent the total candidate solutions and dimensions of the given problem, respectively.

$$X_{i,j} = \text{rand} \times (\text{UB} - \text{LB}) + \text{LB}, j = 1, 2, \ldots\ldots, n \tag{8}$$

where, rand denotes any random value, UB and LB represent the upper and lower limits, respectively.

The second step is exploration or encircling, where crocodiles encircle prey in the global search space. This exploration of the global search space is done by two unique walking styles, such as high walk $\left(t \le \frac{T}{4}\right)$ and abdominal walk $\left(\left(t \le \frac{T}{2}, t > \frac{T}{4}\right)\right)$ These

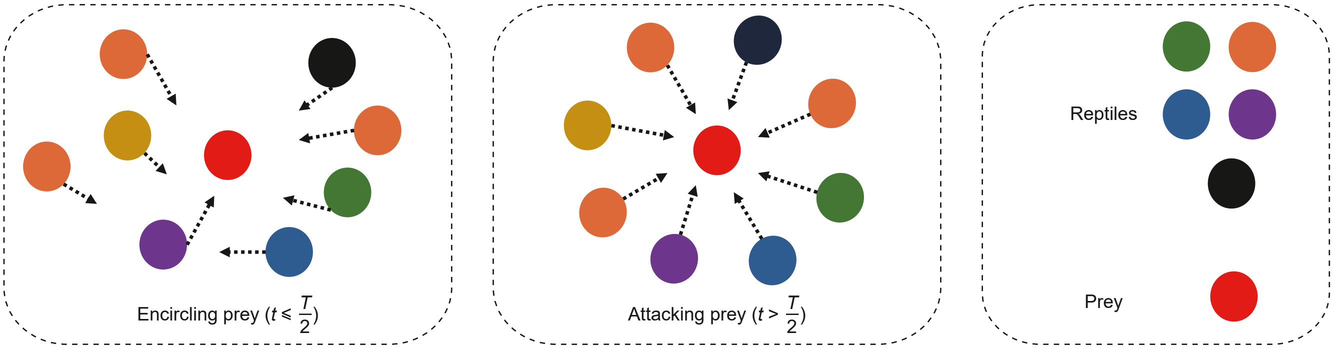

Fig. 7. Schematic diagram of reptile swarm algorithm (RSA) modified from Abualigah et al., (2022).

walking styles identify the number of iterations and can be expressed by Eqs. (9) and (10), respectively:

$$X_{i,j}(t+1) = \text{Best}_j(t) \times \text{ES}(t) \times \text{rand}, \text{if } t \leq \frac{T}{2} \text{ and } t > \frac{T}{4} \tag{9}$$

$$X_{i,j}(t+1) = \text{Best}_j(t) \times \eta_{(i,j)}(t) \times \beta - R_{(i,j)}(t) \times \text{rand}, \text{if } t \leq \frac{T}{4} \tag{10}$$

where, $\text{Best}_j(t)$ denotes the best position among the population, t is iteration, $X_{i,j}(t+1)$ represents the position in the next iteration, T stands for total iterations, $\eta_{(i,j)}$ is the hunting factor and can be calculated by using Eq. (11), β is the exploration accuracy having a constant value of 0.1. $R_{(i,j)}$ is used for search area reduction as given in Eq. (12), where r_2 is a random integer between $[1,N]$, $X_{(i,j)}$ shows a random position of the $i-$ th solution, N is the candidate solution, ES is the probability ratio that falls between 2 and -2, and can be calculated as:

$$\eta_{(i,j)} = \text{Best}_j(t) \times P_{(i,j)} \tag{11}$$

$$R_{(i,j)} = \frac{\text{Best}_j(t) - x_{(r_2,j)}}{\text{Best}_j(t) + \in} \tag{12}$$

$$\text{ES}(t) = 2 \times r_3 \times \left(1 - \frac{1}{T}\right) \tag{13}$$

where r_2 represent a random integer $[1,N]$ and r_3 between $[-1,1]$. $P_{(i,j)}$ is the percentage difference between the current and optimal solutions and can be calculated by using Eqs. (14) and (15), respectively:

$$P_{(i,j)} = \alpha + \frac{x_{(i,j)} - M_{(x_i)}}{\text{Best}_j(t) \times \left(\text{UB}_{(j)} - \text{LB}_{(j)}\right) + \in} \tag{14}$$

$$M(x_i) = \frac{1}{N}\sum_{j=1}^{M} X_{(i,j)} \tag{15}$$

where $M_{(x_i)}$ represent the average position and α is a sensitivity parameter with a value of 0.1.

The third step is the exploitation or hunting phase, where the crocodiles hunt the prey, and can be performed by using two strategies. Such as a coordinated and cooperation strategy. A coordinated strategy can be used when $t \leq 3\frac{T}{4}$ and $t > \frac{T}{2}$, whereas a cooperation strategy can be employed when $t \leq T$ and $t > 3\frac{T}{4}$, and the exploitation phase is represented by Eqs. (16) and (17), respectively.

$$X_{i,j}(t+1) = \text{Best}_j(t) \times P_{(i,j)}(t) \times \text{rand}, \text{if } t \leq 3\frac{T}{4} \text{ and } t > \frac{T}{2} \tag{16}$$

$$X_{i,j}(t+1) = \text{Best}_j(t) - \eta_{(i,j)}(t) \times \in - R_{(i,j)}(t) \times \text{rand}, \text{if } t \leq T \text{ and } t > 3\frac{T}{4} \tag{17}$$

RSA can achieve competitive results in optimizing multimodal and high-dimensional functions. This study used a population size of 50, exploration factor of 0.1, and exploitation factor of 0.5 to optimize solution exploration and exploitation while maintaining computational efficiency. The main advantage of RSA is its unique strategy that changes between encircling (where it does diverse exploration) and hunting (also known as targeted exploitation), enabling it to maintain a balance between local and global search. This unique strategy increases diversity and also reduces the risk of stagnation or trapping in the local optima, leading to better convergence and accuracy in complex optimization problems. More details can be found in Abualigah et al. (2022).

4. Results and discussion

4.1. Feature importance

The random forest algorithm is frequently employed for feature analysis in machine learning, including CO_2-WAG problems. This method can manage complex data relationships and offers comprehensive insights into the importance of features (Chen et al., 2020; Lazreg et al., 2020). It also identifies key variables that play a vital role in influencing predictive outcomes. This feature ranking technique is widely used for both classification and regression problems (Jaiswal and Samikannu, 2017). This study also employed this technique to understand the impact of input features on output parameters. The results revealed that the input features had a significant influence on the cumulative oil production and CO_2 storage. This understanding of the correlation is of paramount importance for CO_2-WAG optimization because it identifies the most crucial parameter affecting the outputs. Furthermore, insights from this analysis can inform decision-making processes in reservoir management and potentially contribute to the development of accurate models in various geological settings. In the selected input parameters, such as oil saturation, water saturation, salinity, and pH, are not controlled variables, where injection rates (gas and water) are controlled variables, and pressure, such as reservoir pressure and BHP (injection and production), can be indirectly influenced by injection fluids and rates. Among the input parameters, oil saturation, water saturation, water injection rate, and gas injection rate have significant effects on the output parameters with relative importance values of 0.40, 0.35, 0.14, and 0.13, respectively, as shown in Fig. 8. Similar effects of oil saturation and pressure on oil production and CO_2 storage are also reported in prior studies (Aliyuda et al., 2020; Shen et al., 2023). Similarly, the BHP Injection, BHP production, and average reservoir pressure follow with relative importance values of less than 0.05, which is consistent with the sensitivity analysis results (Fig. S2). Indicating weak relative importance, such results were also reported by (Gao et al., 2023), where the gas injection rate had a higher impact on cumulative oil production than BHP and water saturation. Furthermore, salinity and pH had the lowest relative importance values, which translates to the least effect on cumulative oil and gas storage in the presence of other parameters. Even some parameters have the least effect on the objective parameters, but to ensure the accuracy of the regression model, we retain all parameters. As mentioned earlier, among the variables in the CO_2-WAG process, some are controlled, while others are not controlled; thus, in the optimization framework, the controlled variables are utilized for the coupled optimization. In a CO_2-WAG project, the controlled variables must be carefully designed to get optimum results as they directly influence oil production and CO_2 storage performance, where adequate CO_2 and water injection will result in maximum oil recovery, CO_2 storage, and will also help to maintain reservoir pressure (Jarrell et al., 2002).

4.2. MLP-GA integration

MLP-GA was used to enhance the designed neural network architecture by utilizing GA for hyperparameter tuning. Initially, separate proxy models, cumulative oil, and CO_2 storage were developed to optimize both proxy models, GA algorithm was utilized to select the optimal MLP architecture for both proxy models. The model accuracy and precision were evaluated by mean square error plots, Fig. 9(a) and (b) show the MSE loss function for Proxy 1 and Proxy 2, respectively. The figure shows how quickly both proxy models learn and converge, without overfitting. Fig. 10(a) and (d) depicts the cross-validation plots of the predicted versus actual values of the cumulative oil and CO_2 storage, respectively. The plots reveal that most of the data points are closely aligned with the diagonal line, indicating that the model predictions accurately match the real values.

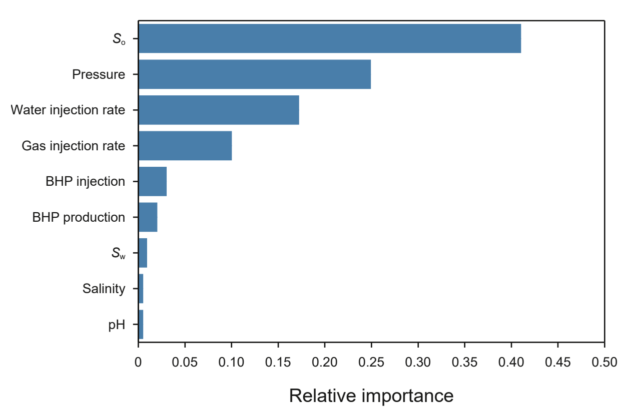

Fig. 8. Random forest algorithm is employed to rank the relative importance of the input parameters on the objective function, cumulative oil production, and CO_2 storage.

The data shows a 45-degree line between the predicted and actual values, which represents the reliability of the proxy models. Blind testing is an important step in validating a model on unseen data. For this purpose, 25% of the data were utilized, and the representative R^2 values for both the proxy models are 0.99 and 0.97, respectively. Furthermore, Fig. 10(b, c) and (e, f) show the MARE between the actual and predicted values of the cumulative oil and CO_2 storage, falling between 0 and 1.2. Based on the MSE plots, R^2, and absolute relative error plot, it is evident that the MLP-GA model demonstrates strong and robust performance for both proxies. The MLP-GA enhanced the neural network architecture and reliability of the proxy models. Now, both the proxies will be integrated and optimized with GA to optimize the objective function.

Fig. 11(a, b) depicts the cumulative oil production and CO_2 storage (Fig. S5 shows CO_2 storage by different mechanisms) for both the base case and GA-optimized case. In 20 years, the base case cumulative oil production is 2.05×10^6 bbl, whereas the GA base optimized case is 2.09×10^6 bbl. This indicates that GA-optimized case increased the oil recovery by 2.15%, where the cumulative CO_2 storage of the base case and GA-based optimized case were 3.43×10^5 and 3.59×10^5 metric tons, respectively. An increment of 4.72% was noticed in the CO_2 storage compared to the base case. Furthermore, the GA-optimized case resulted in a higher oil revenue of 146.84 million USD in comparison to the base case of 143.76 million USD (average price of oil was considered USD 70/bbl), the details are given in Table 5.

Furthermore, Fig. 11(c, d) shows the CO_2 injection and water injection plots. Based on the prediction, both the CO_2 injection rate and water injection rate were increased from the base model, such as 0.115×10^3–0.120×10^3 metric tons/day and 1730–1761 bbl/day, respectively. This increase in the injection profile suggested by MLP-GA can be beneficial twofold: First, this increase in gas injection rate improves microscopic sweep efficiency by increasing the residual oil displacement at the pore scale. CO_2 in miscible or in near-miscible condition reduces interfacial tension and oil viscosity, which leads to improved oil mobility and enhances the miscibility front to contact with zones that were previously bypassed. It also aids in larger amounts of CO_2 storage where more of the CO_2 is injected deeper and can be sequestered by residual and solubility, and other trapping mechanisms. Secondly, the higher water injection rate can achieve better macro sweep efficiency, by pushing the front uniformly and a large volume of oil by sweeping a large volume of the reservoir. Water acts as a piston by supporting the reservoir pressure and helps in preventing large gas breakthroughs, which is one of the challenges faced during the CO_2-EOR process. Maintaining reservoir pressure above the MMP, making sure CO_2 is in a miscible state, and enhancing its interaction with oil. Furthermore, this cyclic nature of the CO_2-WAG technique gives better mobility control, balances the oil and gas mobility ratio, and avoids viscous fingering and gas channeling, which is another inherent problem faced during the CO_2-EOR technique. This unique synergy between water and gas leads to better reservoir conformance and volumetric displacement, as shown in Fig. 19. These results are in line with prior work that optimized injection profiles to enhance both oil recovery and CO_2 storage (Christensen et al., 2001; Kulkarni and Rao, 2005; Lake et al., 2014).

4.3. MLP-PSO integration

Another model for the optimization process is MLP-PSO, where the MLP is coupled with the PSO algorithm. Fig. 12(a, b) depicts the loss vs. epoch plots for cumulative oil production and CO_2 storage, both of which represent the proxy model accuracy and precision,

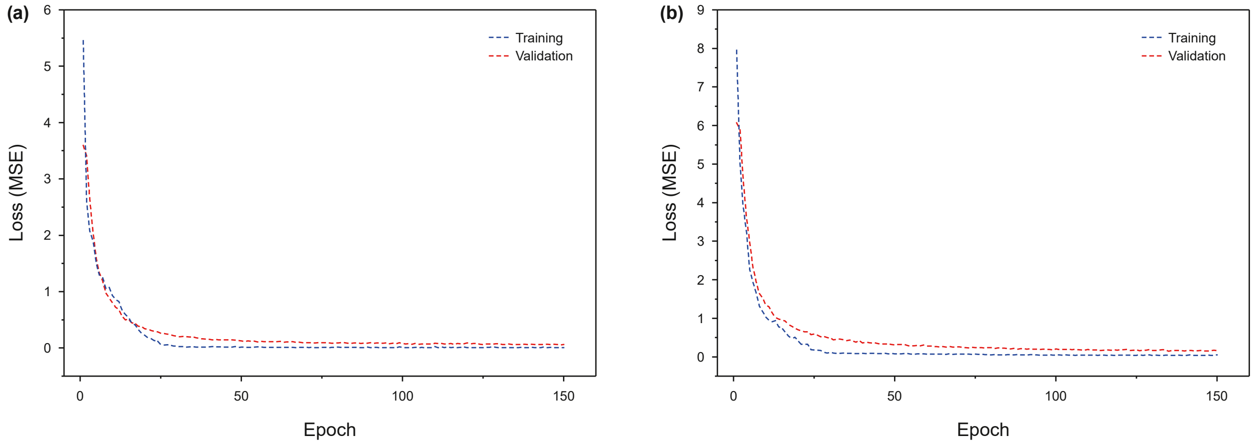

Fig. 9. GA model loss vs. epochs plots. **(a)** Cumulative oil production Proxy 1, **(b)** CO_2 storage Proxy 2.

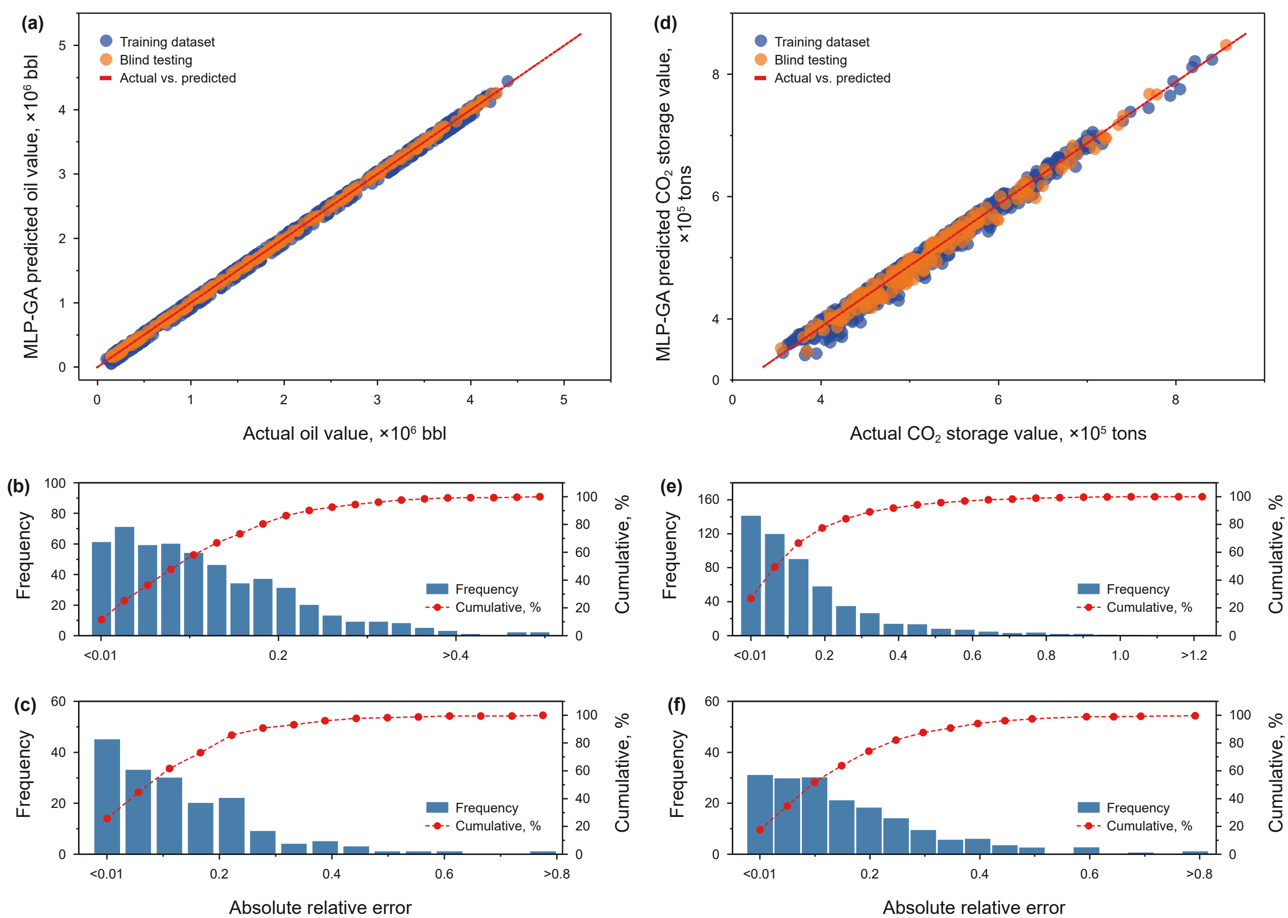

Fig. 10. MLP-GA model actual vs. prediction plots, and absolute relative distribution. **(a)** Proxy 1 training and blind testing actual vs. prediction plot, **(b, c)** absolute relative error plots for Proxy 1 training and validation data. **(d)** Proxy 2 training and blind testing actual vs. prediction plot, **(e, f)** absolute relative error plots for Proxy 1 training and validation data.

Table 5
Comparison between base case and proposed workflows (MLP-GA, MLP-PSO, and MLP-RSA).

Results	Base case	MLP-GA case	MLP-PSO case	MLP-RSA case	Difference (MLP-RSA to MLP-PSO)
Cumulative oil production, $\times 10^6$ bbl	2.05	2.09	2.15	2.27	5.74%
Cumulative CO_2 storage, $\times 10^5$ tons	3.43	3.59	3.95	4.28	8.10%
Oil revenue, million USD bbl $\times$ oil price	143.76	146.84	150.44	159.07	8.62
Objective function ($w = 0.5$)	0.88	0.91	1.007	1.088	8.04%

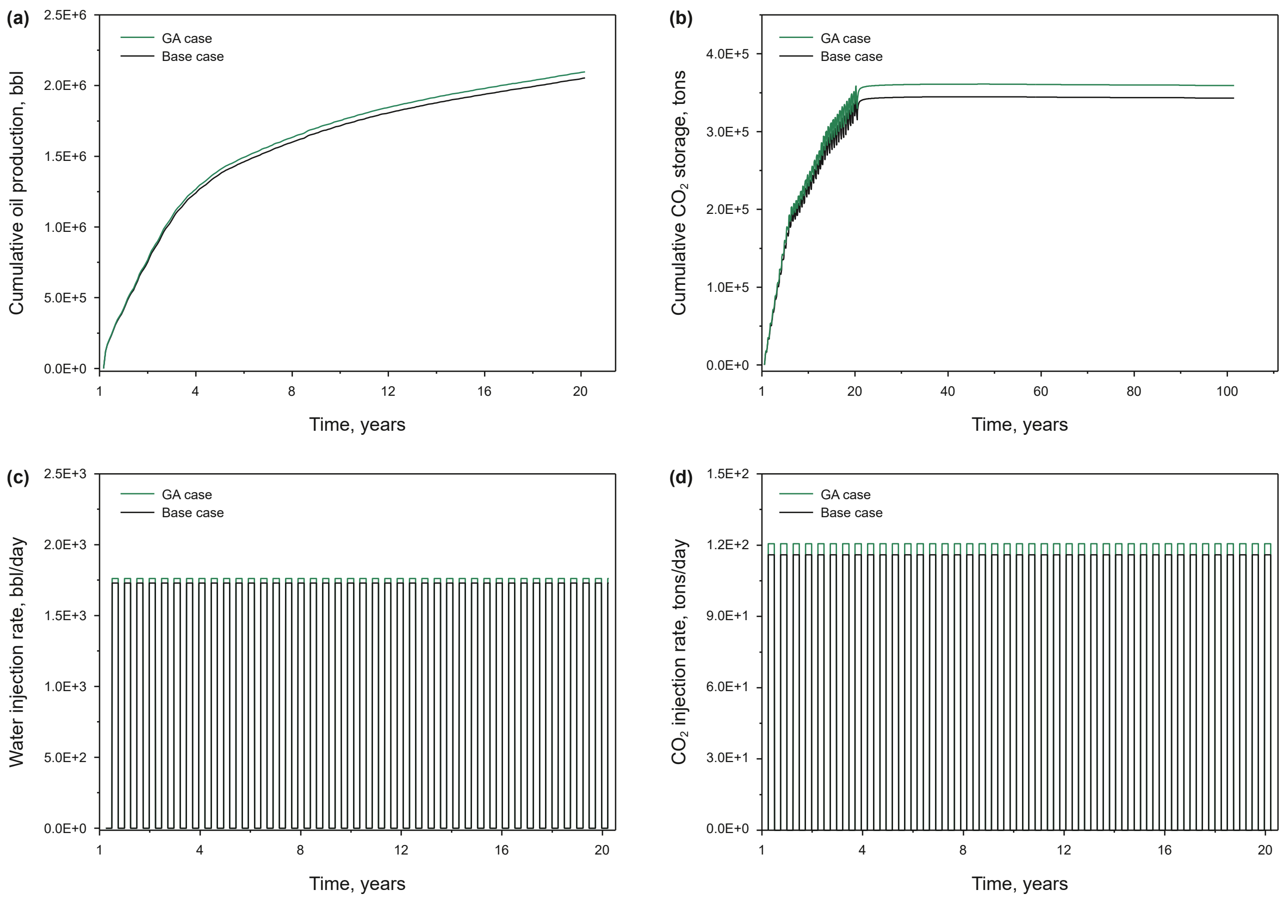

Fig. 11. GA optimal model oil production, storage, and injection rates plots. **(a)** Cumulative oil production, **(b)** CO_2 storage, **(c)** water injection rate, and **(d)** CO_2 injection rate.

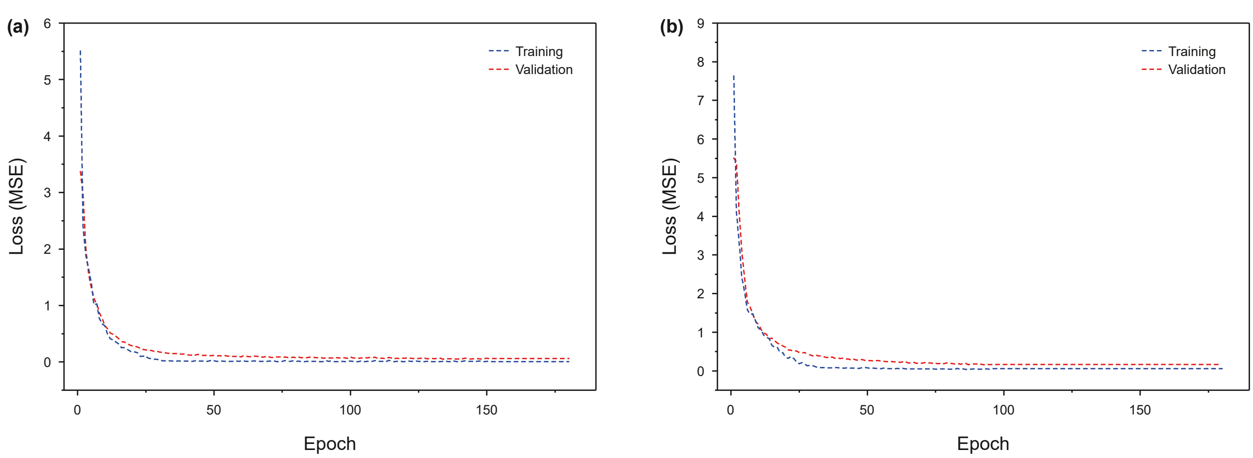

Fig. 12. PSO model loss vs. epochs plots. **(a)** Cumulative oil production Proxy 1, **(b)** CO_2 storage Proxy 2.

and the plots also represent the quick learning and convergence of the proxy models without overfitting. Fig. 13(a, d) shows the cross plots of the blind testing, showing how effectively the cumulative oil production and CO_2 storage proxy models predicted the actual values on the unseen data. Both the actual and predicted data points effectively aligned across the diagonal, validating the model accuracy and prediction ability. The representative R^2 values of both cumulative oil production and CO_2 storage were 0.99 and 0.97, respectively. Similarly, Fig. 13(b, c) and (e, f) show that the MARE for both the cumulative oil production and CO_2 storage models lies between 0 and 1.2. The results show that the model effectively predicts the actual values of both cumulative oil production and CO_2 storage, indicating that both proxy models can be coupled to perform the optimization process.

After coupling both proxy models, the PSO algorithm was employed to optimize the desired objective function parameters. Fig. 14(a, b) depicts the cumulative oil production and CO_2 storage (Fig. S6 shows CO_2 storage by different mechanisms) of the MLP-PSO optimized case with the base case. The cumulative oil production predicted by the base case and MLP-PSO optimized

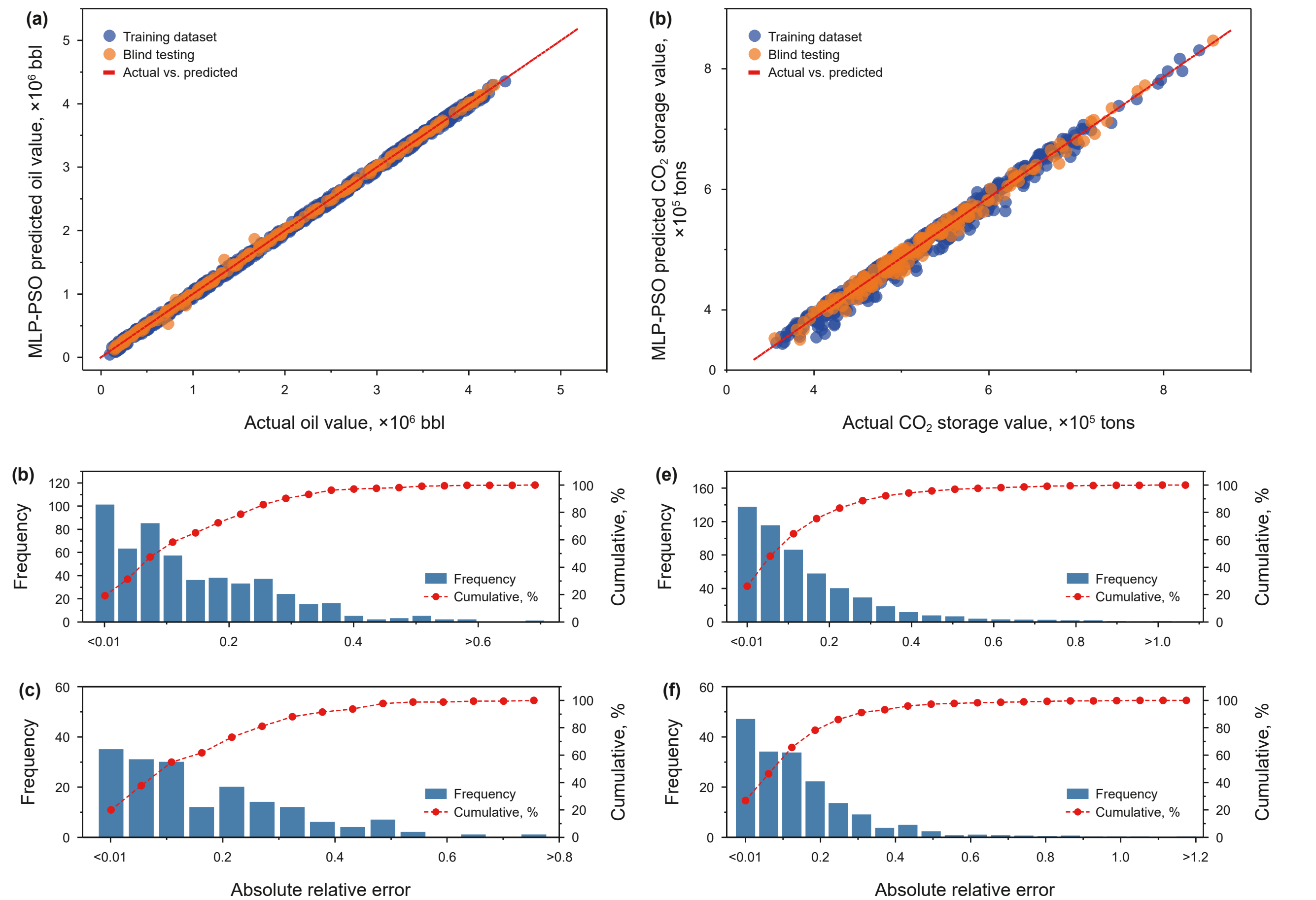

Fig. 13. MLP-PSO model actual vs. prediction plots, and absolute relative distribution. **(a)** Proxy 1 training and blind testing actual vs. prediction plot, **(b, c)** absolute relative error plots for Proxy 1 training and validation data. **(d)** Proxy 2 training and blind testing actual vs. prediction plot, **(e, f)** absolute relative error plots for Proxy 1 training and validation data.

case is 2.05×10^6 and 2.15×10^6 bbls, respectively, resulting in an increment of 4.64%. Whereas the CO_2 storage prediction of the base model and PSO optimized case are 3.43×10^5 and 3.95×10^5 metric tons respectively, indicating an increase of 15.4%, as detailed in Table 5. Furthermore, the PSO-optimized case resulted in a higher oil revenue of 150.44 million USD in comparison to the base model of 143.76 million USD. The given results deduced that the optimization process was successful and improved all vital performance indicators. Furthermore, Fig. 14(c, d) depicts the water and gas injection profiles of the base case and PSO-optimized case. Both the water and gas injection rates of the optimized case were higher than those of the base case, resulting in enhanced oil recovery, higher CO_2 storage, and maintenance of the reservoir pressure above the MMP, as shown in Fig. S10.

4.4. MLP-RSA integration

The MLP-RSA model offers a novel approach to optimization by coupling MLP with RSA, where the MLP is coupled with a newly developed RSA algorithm. The MLP architecture was optimized using the RSA algorithm for both proxy models. Fig. 15(a) and (b) illustrate the loss vs. epoch plots for these models, demonstrating the proxy models accuracy and precision, rapid learning, and convergence without overfitting. The effectiveness of both proxy models in forecasting the actual values for unseen data is illustrated through blind testing cross-plots, as shown in Fig. 16(a) and (d). The close alignment of the actual and predicted data points along the diagonal confirmed the accuracy and predictive power of the models. The representative R^2 values for both proxy models are 0.99 and 0.97, respectively. Fig. 16(b, c) and (e, f) also demonstrates that the MARE for both models is within the range of 0–1, emphasizing their proficiency in predicting actual values and their suitability to couple for the optimization process.

After coupling both proxy models, the RSA algorithm was employed to optimize the objective function parameters, total oil production, and CO_2 storage amount. Fig. 17(a, b) compares the cumulative oil production and CO_2 storage (Fig. 18 shows CO_2 storage by different mechanisms) of the MLP-RSA-optimized case with those of the base case. The optimized scenario resulted in a cumulative oil production of 2.27×10^5 bbl, an increase of 10.65% from the base case. Likewise, CO_2 storage predictions increased from 3.43×10^5 metric tons in the base case to 4.28×10^5 metric tons in the RSA-optimized case, an increase of 24.72%, as illustrated in Table 6. From an economic analysis perspective, the RSA-optimized case resulted in oil revenue of 159.07 million USD, an increment of 15.31 million USD from the base case (143.76 million USD). These outcomes validate the effectiveness of the optimization process for improving the performance indicators. Fig. 17(c, d) shows the water and gas injection profiles for both base cases, with the optimized scenario showing higher injection rates. This

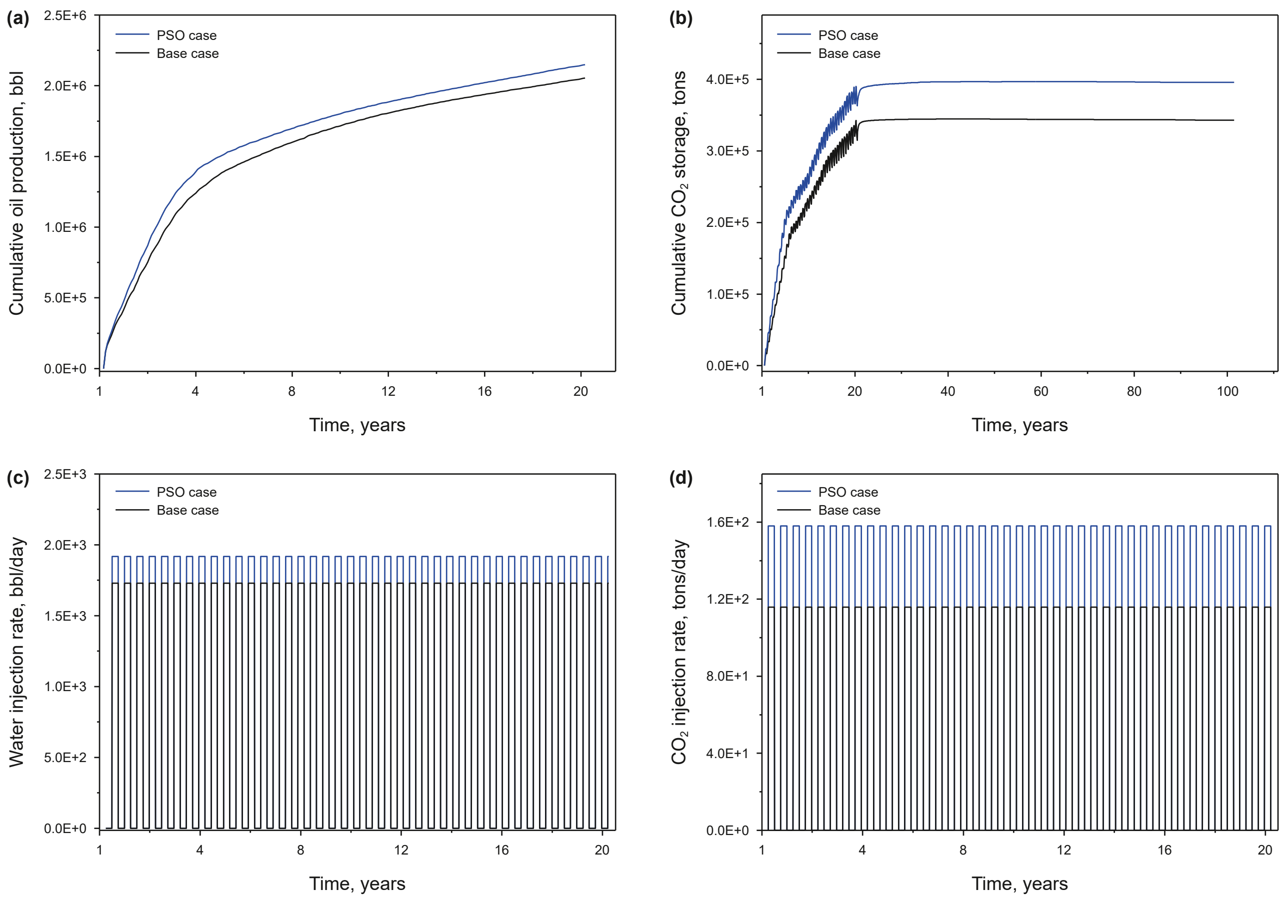

Fig. 14. PSO optimal model oil production, storage, and injection rates plots. **(a)** Cumulative oil production, **(b)** CO_2 storage, **(c)** water injection rate, and **(d)** CO_2 injection rate.

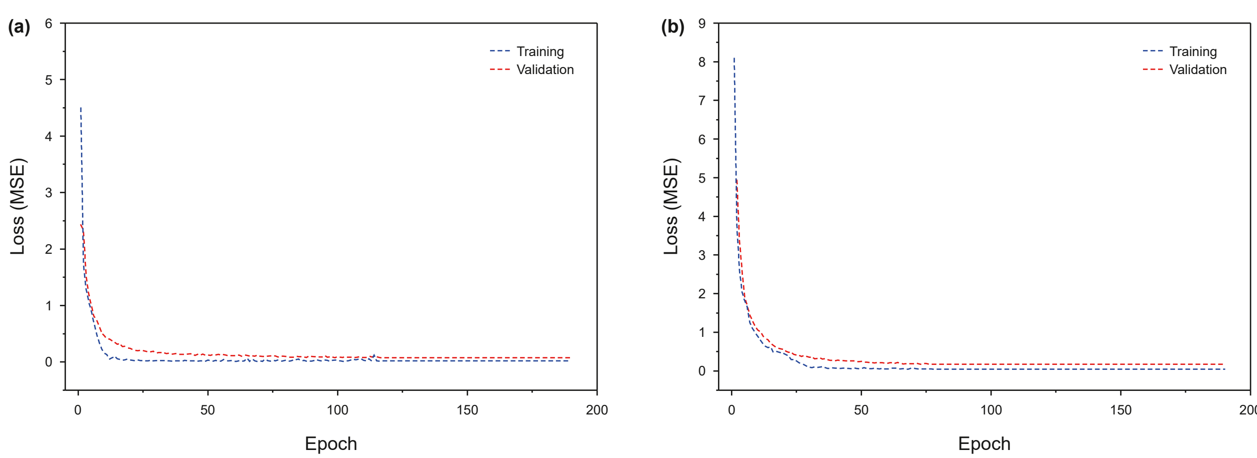

Fig. 15. RSA model loss vs. epochs plots. **(a)** Cumulative oil production Proxy 1, **(b)** CO_2 storage Proxy 2.

increase in injection rate led to improved oil recovery, greater CO_2 storage, and maintenance of the reservoir pressure above the MMP, as shown in Fig. S10.

4.5. Comparative analysis

This study utilizes a novel technique that employs GA, PSO, and RSA to determine the best input parameters to optimize the desired objective function. All three algorithms use different approaches to search for the optimal objective function value.

4.5.1. Objective function optimization

A comparative assessment of the three machine learning models, MLP-GA, MLP-PSO, and MLP-RSA, is presented in Table 5, focusing on their ability to predict cumulative oil production and CO_2 storage. All models utilized the same objective function. The analysis revealed that MLP-RSA achieved better predictions compared to MLP-PSO and MLP-GA, as shown in Fig. 20(a and b). This advantage can be attributed to the higher objective function value obtained by RSA 1.088, where PSO and GA resulted in objective values of 1.007 and 0.91, respectively, as illustrated in

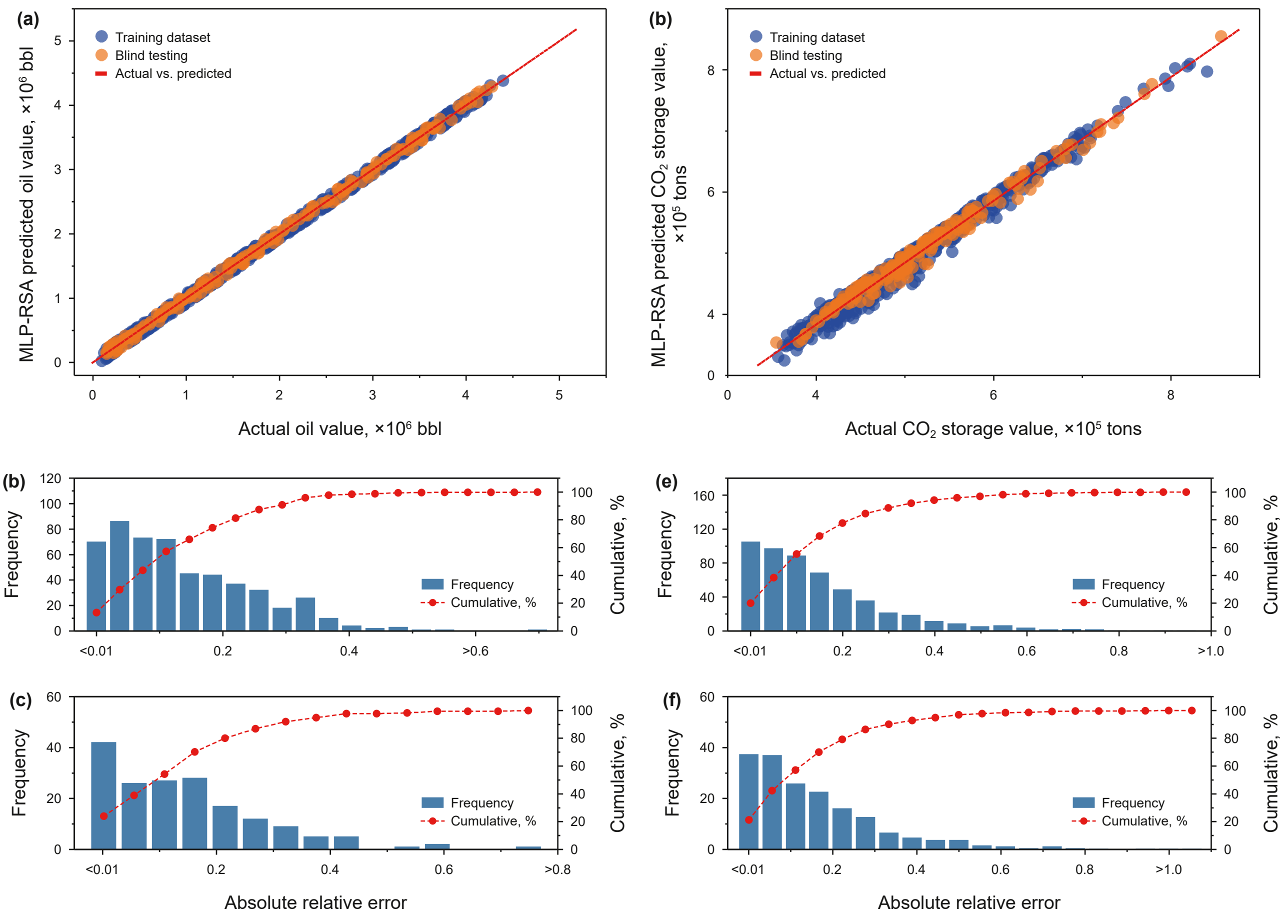

Fig. 16. MLP-RSA model actual vs. prediction plots, and absolute relative distribution. **(a)** Proxy 1 training and blind testing actual vs. prediction plot, **(b, c)** absolute relative error plots for Proxy 1 training and validation data. **(d)** Proxy 2 training and blind testing actual vs. prediction plot, **(e, f)** absolute relative error plots for Proxy 1 training and validation data.

Fig. 19. The optimized parameters proposed by RSA, particularly the water injection rates, were generally higher than those suggested by PSO and GA. The GA recommended slightly lower injection rates for both CO_2 and water compared to the PSO, as shown in Fig. 20(c and d). Interestingly, all the algorithms advocated for increased injection rates during the optimization process. This higher objective function value can be attributed to the intrinsic properties of RSA that it very effectively balances the exploration and exploitation stages; it also has the ability to escape the local optima successfully, and it can control the step size. Based on this mechanism, we presume that it helps in maintaining diversity in the search space and speeds up the convergence towards the global optimum. Similar results are also reported in prior studies (Abualigah et al., 2022). When evaluated against the base case, all models, such as MLP-RSA, MLP-PSO, and MLP-GA, demonstrated significant improvements in enhancing cumulative oil production and CO_2 storage. The overall performance of the models can be represented as MLP-RSA > MLP-PSO > MLP-GA.

Pressure plays an important role in enhancing oil production from depleted reservoirs because it directly affects the displacement and sweep efficiency in WAG processes (Sallam et al., 2015). The RSA algorithm results in an optimal solution with higher objective values by suggesting higher injection rates that substantially increase cumulative oil production and CO_2 storage. RSA-based solutions yield 5.74% and 8.10% more cumulative oil production and 9.33% and 19.10% more CO_2 storage in comparison to PSO and GA, respectively. In contrast to GA, PSO resulted in an increase of 2.92% and 10.17% in oil production and CO_2 storage, respectively. As the reservoir pressure is primarily controlled by water injections, the lower injection rates proposed by the GA led to comparatively decreased pressure, oil recovery, and CO_2 storage.

Effective optimization of injection rates can significantly improve the outcomes of CO_2-EOR projects, enhancing both economic viability and carbon storage efficiency. Additionally, the total oil recovery factor over 20 years was compared. The MLP-RSA, MLP-PSO, and MLP-GA methods yielded oil revenue of 146.84, 150.44, and 159.07 million USD, respectively. MLP-RSA resulted in an increase of 5.74% and 8.33% in the oil recovery factors in comparison to the MLP-PSO and MLP-GA models, respectively. These results indicate that our proposed optimization technique led to significant improvements across all key performance metrics examined in this study.

4.5.2. Model's performance and accuracy

The evaluation of the RSA, PSO, and GA revealed exceptional effectiveness in enhancing the specified simulation model across various operational factors, including water and gas injection rates as well as bottomhole pressures. Each model results in distinct optimal results owing to variations in their computational approaches and relevant adjustments to specific parameters that affect the efficiency. The RSA, inspired by the hunting habits of reptiles, utilizes adaptive movement strategies to achieve a balance between exploration and exploitation (Abualigah et al., 2022; Elgamal et al., 2022). PSO, which is rooted in swarm intelligence

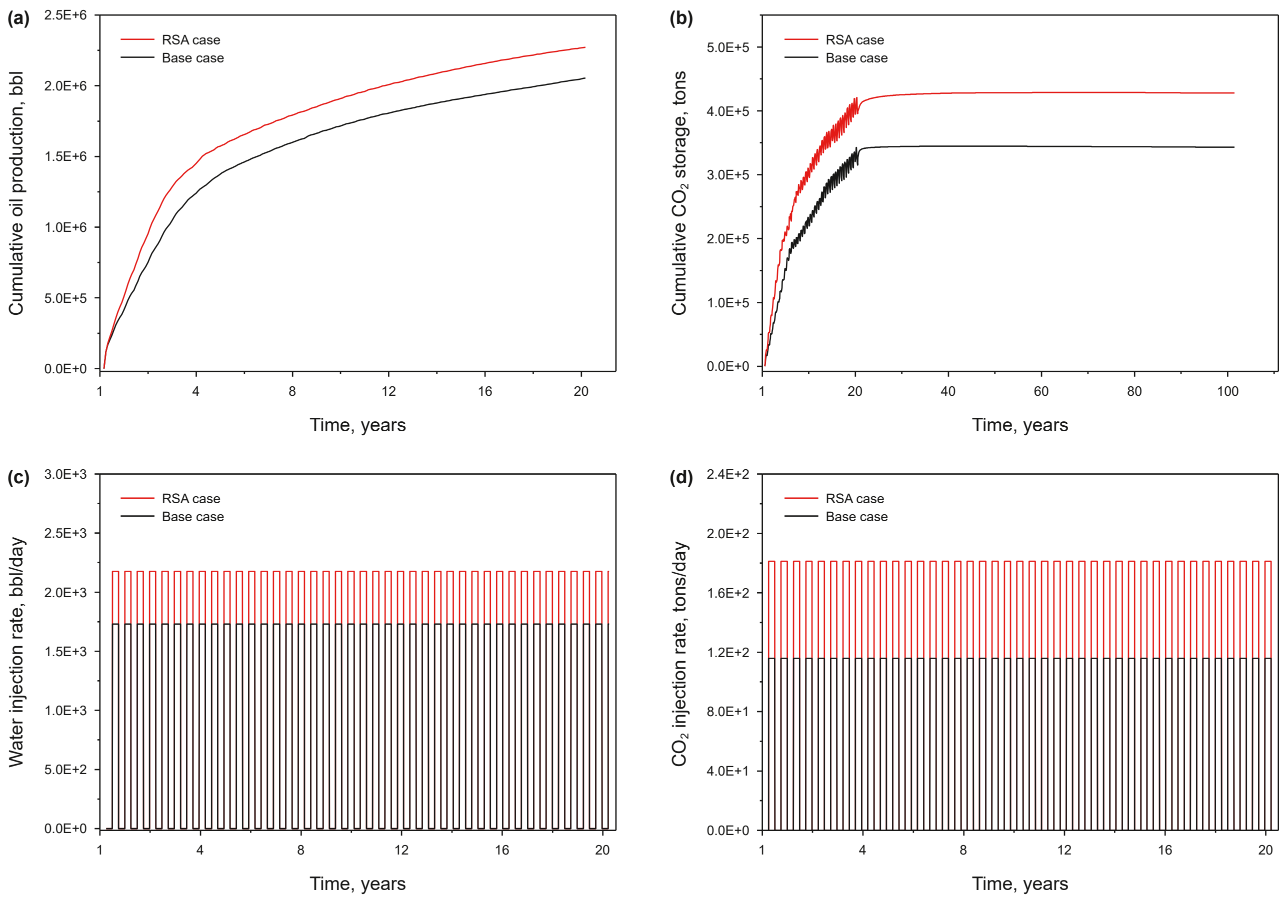

Fig. 17. RSA optimal model oil production, storage, and injection rates plots. **(a)** Cumulative oil production, **(b)** CO_2 storage, **(c)** water injection rate, and **(d)** CO_2 injection rate.

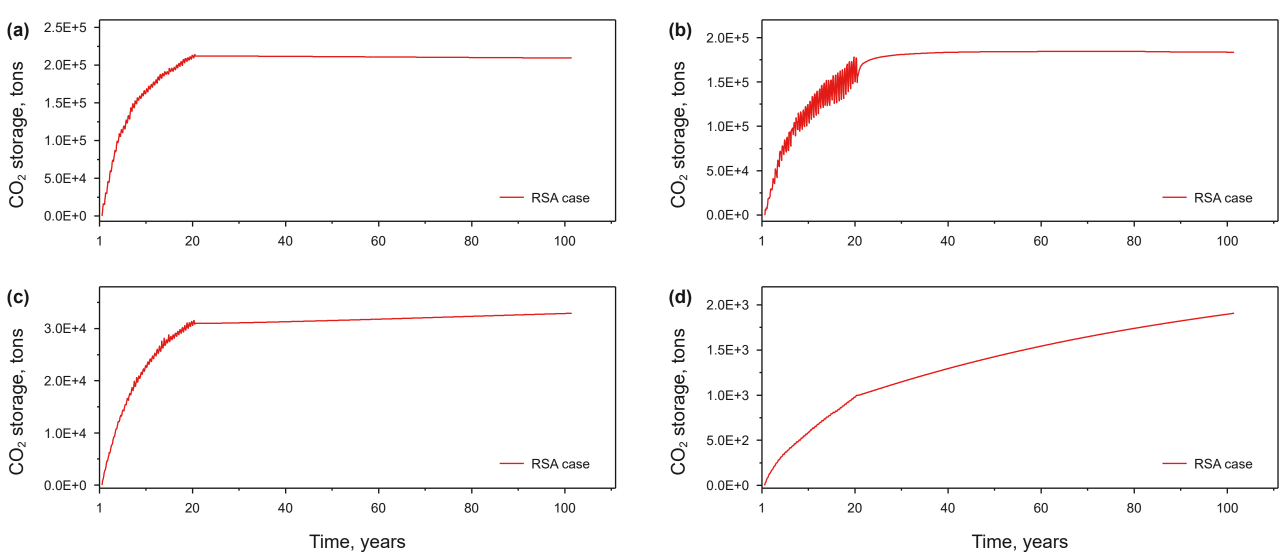

Fig. 18. Optimum model (RSA) based CO_2 trapping by all mechanisms. **(a)** Stratigraphic trapping, **(b)** residual trapping, **(c)** solubility trapping, and **(d)** mineralization trapping.

principles, depends on social and cognitive factors, along with inertial weights, to direct particles toward ideal solutions (Hassan et al., 2005; Huang et al., 2023). GA, an evolution-based algorithm, utilizes genetic mechanisms, such as selection, crossover, and mutation, to progressively improve solutions (Hassan et al., 2005). The selection of these parameters significantly influences the convergence patterns and the ultimate results of each algorithm (Naga and Prakash, 2020; Shanthi and Chethan, 2022). Despite their methodological distinctions, all three algorithms exhibited outstanding optimization capabilities.

After careful structural examination of the algorithms, the mean absolute relative error between the actual and predicted

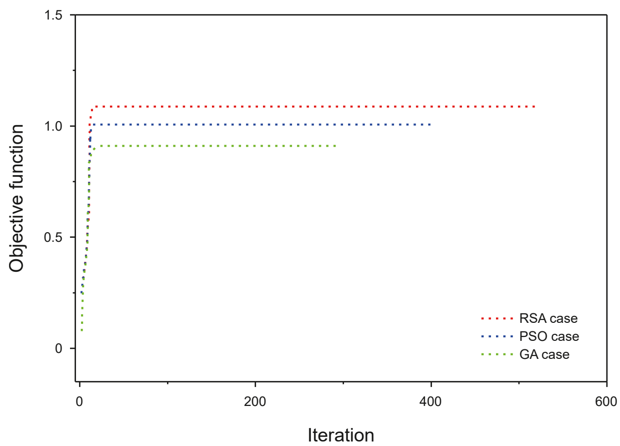

Fig. 19. Proposed workflow performance indicator, objective function vs. iteration plot.

Table 6
Base case and optimal case results and comparison.

Results	Base case	Optimal case	Difference
Water injection rate, × 10^3 bbl/day	1.7	2.7	1
Gas injection rate, × 10^2 tons/day	1.1	1.8	0.7
Cumulative oil production, × 10^6 bbl	2.05	2.27	10.65%
Cumulative CO_2 storage, × 10^5 tons	3.43	4.28	24.72%
Oil revenue, million USD bbl × oil price	143.76	159.07	15.31
Objective function ($w = 0.5$)	0.88	1.088	0.208
Slug size, PV	0.98	1.58	0.6
Pressure profile (average), MPa	14.15	15.65	1.47

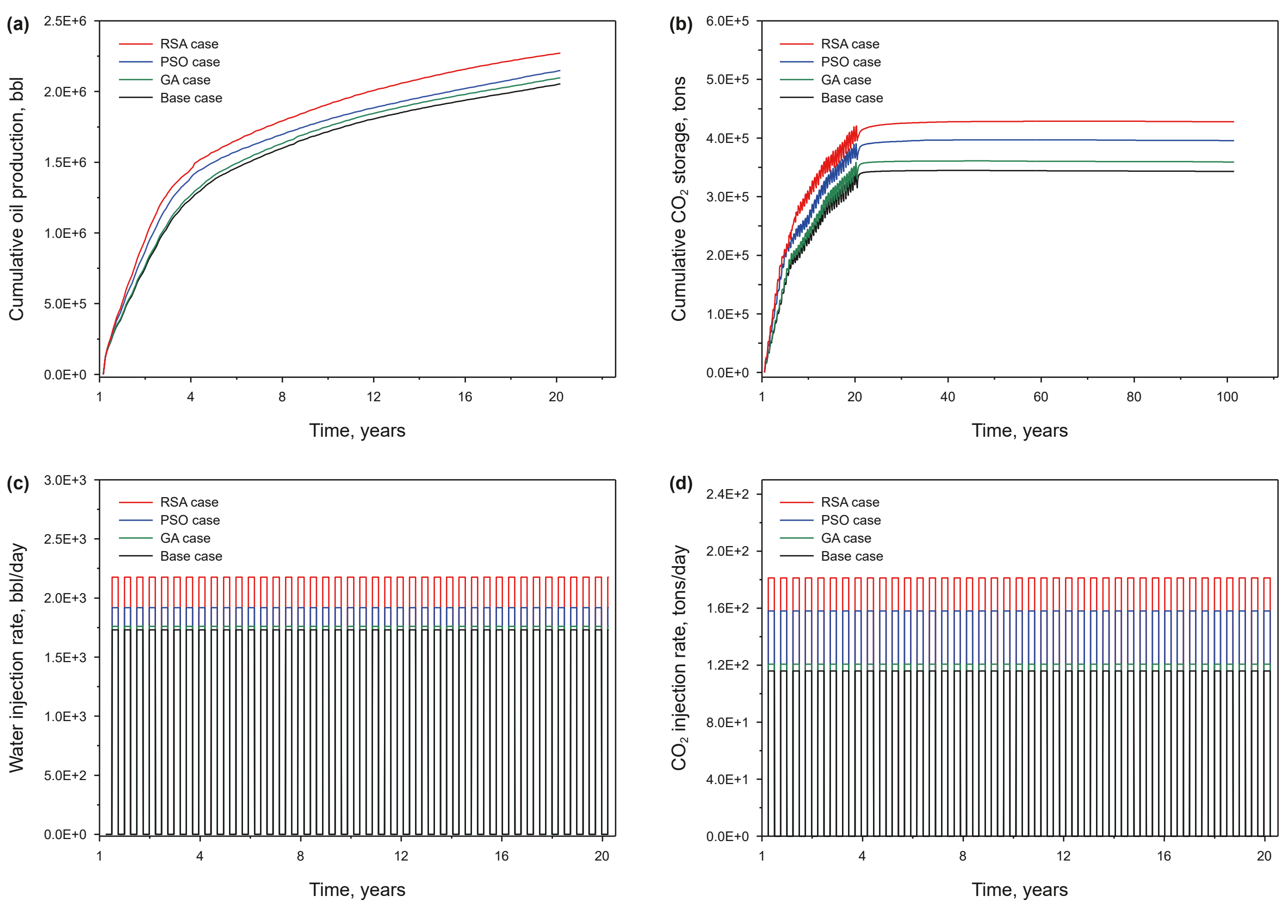

Fig. 20. Base GA, PSO, and RSA optimal models oil production, storage, and injection rates plots. **(a)** Cumulative oil production, **(b)** CO_2 storage, **(c)** water injection rate, and **(d)** CO_2 injection rate.

values falls between 0 and 1, 0–1.2, and 0–1.2 for RSA, PSO, and GA, respectively. The loss vs. epoch plots showed that the proxy models trained by these algorithms converged effectively without overfitting. Additionally, the R^2 values of all the models fall in the range of 0.99 and 0.97 for Proxy 1 and Proxy 2. Based on the above evaluation and optimization, the RSA algorithm outperformed both the PSO and GA algorithms by suggesting higher operational parameters, resulting in high cumulative oil production and CO_2 storage. Compared with the base case, all three algorithms outperformed the base case. The optimal values suggested by these algorithms indicate the robustness and effectiveness of the proposed model. Furthermore, these optimization techniques

Table 7
System configuration and environment.

System configuration	Description
Operating system	Windows 11 Pro, 64-bit
Processor	Intel (R) Core i7-8550U CPU 1.80 GHz 2.0 GHz (8 processors)
RAM	16 Gb

number of wells, and geological complexity, compared to real field conditions. The absence of field datasets meant that certain heterogeneities, operational constraints, and site-specific uncertainties could not be fully incorporated. The study focused on nine input parameters (such as oil saturation, water saturation, salinity, pH, average pressure, bottom hole pressure production, and injection) and two output parameters (cumulative oil production and CO_2 storage), sufficient for the scope of the study, but

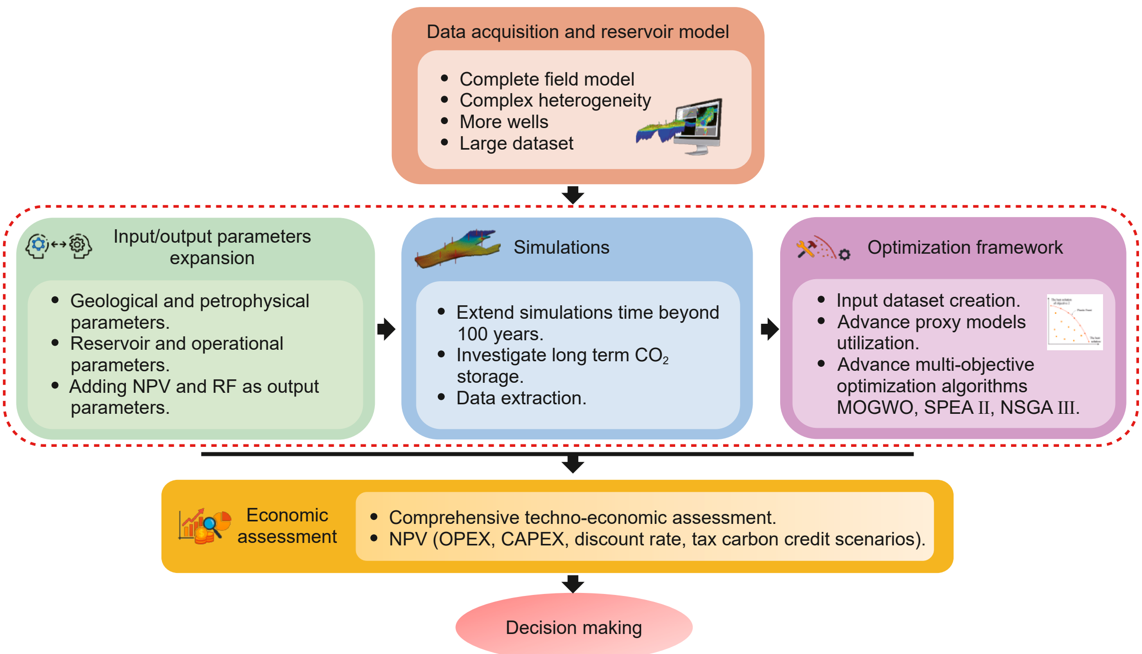

Fig. 21. Future workflow to extend the optimization study.

eliminate the need for time-consuming numerical simulations, which are typically employed to determine the optimal input parameters for the reservoir. Traditional commercial simulators require hours to complete the optimization task, while machine learning techniques can complete the task in mere seconds. A normal conventional simulator requires 7–8 min for each case, hence needs hours to complete an optimization study, and requires a large storage space (in GBs). On the other hand, the designed models used for the optimization cases required 359, 470, and 556 s for the RSA, PSO, and GA, respectively, and significantly reduced storage space (mere MBs). To compare the convergence of the three algorithms, RSA requires 15 iterations to converge, whereas PSO and GA require 20 and 24 iterations to converge, respectively, as shown in Fig. 19. Furthermore, details of the PC environment used in this study are listed in Table 7. The computational effort required for optimization is substantially decreased by this methodology, making it a practical alternative for conventional simulations, and can help engineers to conduct optimization studies with ease.

4.6. Limitations and future work

This study provides a robust optimization framework for CO_2-WAG process, considering a representative model based on Bell Creek Oilfield, providing a platform for testing and demonstrating the proposed optimization workflow. However, a representative model simplifies several aspects of a complete reservoir model, which can be considered as limitations, such as heterogeneity, needs to be extended to capture the reservoir intricate details. To incorporate mineralization trapping, minerals like anorthite, calcite, and kaolinite were taken into account because other minerals like quartz have slow kinetic reactions and require long simulation time (>100 years) and computational power. Furthermore, for profitability, the study only considered oil revenues to reflect the incremental oil recovery without considering other important parameters such as operational expenditure (OPEX), capital expenditure (CAPEX), discount rate, and carbon credits.

In future, this optimization framework can be extended based on complex models, having complex heterogeneity, a large number of wells, and more operational constraints as given in Fig. 21. The input parameters will be updated to diverse geological, petrophysical, and operational parameters, which will enable the proposed framework to capture more complex, uncertain, and nonlinear interactions. For the economic side, a full techno-economic analysis will be adopted to incorporate all the important parameters for NPV calculations (OPEX, CAPEX, discount rate, and carbon credit) that affect decision-making. In terms of CO_2 emission reduction and complete analysis, it is necessary to conduct a full life cycle analysis (LCA) to see the overall impact of CO_2-EOR operations. Furthermore, to strengthen the methodology, future work will integrate RSA algorithms with other algorithms and can also utilize multi-objective optimization algorithms (Wang et al., 2023) (such algorithms perform better when having conflicting objective functions), as given in Fig. 21. Considering all these elements, future studies will further enhance the optimization workflow in complex reservoir models.

5. Conclusions

This study introduced a machine-learning-driven workflow for the coupled optimization of cumulative oil production and CO_2 storage for CO_2-WAG EOR project. A dataset of 700 samples was generated by utilizing CMG-GEM, consisting of 9 input and two output parameters, that served as an input for training MLP proxy models for predictions. The proxy models were then coupled with evolutionary and metaheuristic optimization algorithms such as GA, PSO, and RSA to predict optimal operating strategies. The proposed workflow gives an alternate approach to conventional optimization techniques by reducing computational time while maintaining accuracy.

The conclusions drawn are as follows:

1) Machine learning-driven MLP proxy models developed for CO_2-WAG projects exhibited precise and fast predictions of cumulative oil production and CO_2 storage, providing engineers with a better understanding of fluid dynamics, carbon-trapping mechanisms, and reservoir behavior.
2) The newly developed MLP-GA, MLP-PSO, and MLP-RSA workflows demonstrated its capability to optimize cumulative oil production and CO_2 storage for CO_2-WAG projects in various operational settings. This process exhibits excellent efficiency, stability, and speed. Using these models allows engineers to make wise operational decisions without resorting to extensive numerical simulation.
3) In terms of performance, all three optimization models outperformed the base model, particularly MLP-RSA. Optimizing both cumulative oil production and CO_2 storage by 10.65% and 24.72% respectively, and increasing the oil revenue, followed by MLP-PSO and MLP-GA.
4) By comparing the three algorithms, RSA optimization outperformed both PSO and GA algorithms, resulting in 5.74% and 8.10% higher cumulative oil production and 9.33% and 19.10% greater CO_2 storage. MLP-RSA, MLP-PSO, and MLP-GA demonstrated exceptional computing efficiency compared with traditional reservoir simulators, requiring minutes for optimization, thereby accelerating the optimization process.

Furthermore, the proposed methodology can also be employed in other EOR techniques (chemical flooding), and fields encounter similar optimization problems due to its flexibility. Therefore, the development of this optimization workflow will provide additional insights into the resolution of multi-objective optimization problems in various fields.

CRediT authorship contribution statement

Shahab Ud Din: Writing – original draft, Validation, Software, Conceptualization. **Liang Xue:** Supervision, Methodology, Conceptualization. **Fu-Long Ning:** Writing – review & editing, Conceptualization. **Dong-Dong Guo:** Writing – review & editing. **Qin-Zhuo Liao:** Writing – review & editing. **Hussan Zeb:** Writing – review & editing.

Data availability

Data will be made available upon request.

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.

Acknowledgment

This work was funded by the National Natural Science Foundation of China (Grant No. 52274048), and National Science and Technology Major Project (Grant No. 2025ZD1405202).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.petsci.2026.01.026.

References

Abualigah, L., Elaziz, M.A., Sumari, P., Geem, Z.W., Gandomi, A.H., 2022. Reptile search algorithm (RSA): a nature-inspired meta-heuristic optimizer. Expert Syst. Appl. 191. https://doi.org/10.1016/j.eswa.2021.116158.
Al-Bayati, D., Saeedi, A., Myers, M., White, C., Xie, Q., Clennell, B., 2018. Insight investigation of miscible $SCCO_2$ water alternating gas (WAG) injection performance in heterogeneous sandstone reservoirs. J. CO_2 Util. 28, 255–263. https://doi.org/10.1016/j.jcou.2018.10.010.
Aliyuda, K., Howell, J., Humphrey, E., 2020. Impact of geological variables in controlling oil-reservoir performance: an insight from a machine-learning technique. SPE Reserv. Eval. Eng. 23 (4), 1314–1327. https://doi.org/10.2118/201196-PA.
Bachu, S., 2001. Screening and ranking of hydrocarbon reservoirs for CO_2 storage in the Alberta Basin, Canada. In: First National Conference on Carbon Sequestration.
Chen, R.C., Dewi, C., Huang, S.W., Caraka, R.E., 2020. Selecting critical features for data classification based on machine learning methods. J. Big Data 7 (1), 52. https://doi.org/10.1186/s40537-020-00327-4.
Christensen, J.N., Watkins, J.M., Devriendt, L.S., DePaolo, D.J., Conrad, M.E., Voltolini, M., Yang, W., Dong, W., 2021. Isotopic fractionation accompanying CO_2 hydroxylation and carbonate precipitation from high pH waters at the Cedars, California, USA. Geochem. Cosmochim. Acta 301, 91–115. https://doi.org/10.1016/j.gca.2021.01.003.
Christensen, J.R., Stenby, E.H., Skauge, A., 2001. Review of WAG field experience. SPE Reservoir Eval. Eng. 4 (2), 97–106. https://doi.org/10.2118/71203-PA.
Czernichowski, L.I., Berenblyum, R., Bigi, S., Car, M., Gastine, M., Persoglia, S., Poulsen, N., Schmidt-Hattenberger, C., Stead, R., Vincent, C.J., Wildenborg, T., 2018. CO_2 GeoNet actions in Europe for advancing CCUS through global cooperation. Energy Proc. 154, 73–79. https://doi.org/10.1016/j.egypro.2018.11.013.
De Weck, O., Kim, I.Y., 2004. Adaptive weighted sum method for Bi-objective optimization. In: 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference. American Institute of Aeronautics and Astronautics, Reston, Virigina. http://arc.aiaa.org/doi/abs/10.2514/6.2004-1680.
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. CO_2 accounting and risk analysis for CO_2 sequestration at enhanced oil recovery sites. Environ. Sci. Technol. 50 (14), 7546–7554. https://doi.org/10.1021/acs.est.6b01744.
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.
Delerce, S., Bénézeth, P., Schott, J., Oelkers, E.H., 2023. The dissolution rates of naturally altered basalts at pH 3 and 120 °C: implications for the in-situ mineralization of CO_2 injected into the subsurface. Chem. Geol. 621, 121353. https://doi.org/10.1016/j.chemgeo.2023.121353.
Dziejarski, B., Krzyżyńska, R., Andersson, K., 2023. Current status of carbon capture, utilization, and storage technologies in the global economy: a survey of technical assessment. Fuel 342, 127776. https://doi.org/10.1016/j.fuel.2023.127776.
Elgamal, Z., Sabri, A.Q.M., Tubishat, M., Tbaishat, D., Makhadmeh, S.N., Alomari, O.A., 2022. Improved reptile search optimization algorithm using chaotic map and simulated annealing for feature selection in medical field. IEEE Access 10, 51428–51446. https://doi.org/10.1109/ACCESS.2022.3174854.
Emam, M.M., Houssein, E.H., Ghoniem, R.M., 2023. A modified reptile search algorithm for global optimization and image segmentation: case study brain MRI images. Comput. Biol. Med. 152, 106404. https://doi.org/10.1016/j.compbiomed.2022.106404.
Forray, V., Király, C., Demény, A., Cseresznyés, D., Szabó, C., Falus, G., 2021. Mineralogical and geochemical changes in conglomerate reservoir rocks induced by CO_2 influx at Mihályi-Répcelak natural analogue, NW-Hungary. Environ. Earth Sci. 80 (23), 765. https://doi.org/10.1007/s12665-021-10050-9.
Ghaderi, S.M., Clarkson, C.R., Chen, S., 2012. Optimization of WAG process for coupled CO_2 EOR-storage in tight oil formations: an experimental design approach. In: SPE Canadian Unconventional Resources Conference. https://doi.org/10.2118/161884-MS.

Gao, M., Liu, Z., Qian, S., Liu, W., Li, W., Yin, H., Cao, J., 2023. Machine-learning-based approach to optimize CO_2-WAG flooding in low permeability oil reservoirs. Energies 16 (17), 6149. https://doi.org/10.3390/en16176149.

Goodfellow, I.,B.Y., C. A, 2016. Deep Learning. MIT Press.

Hassan, R., Cohanim, B., de Weck, O., Venter, G., 2005. A comparison of particle swarm optimization and the genetic algorithm. In: 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference. https://doi.org/10.2514/6.2005-1897.

He, R., MA, W., Ma, X., Liu, Y., 2021. Modeling and optimizing for operation of CO_2-EOR project based on machine learning methods and greedy algorithm. Energy Rep. 7, 3664–3677. https://doi.org/10.1016/j.egyr.2021.05.067.

Huang, K.W., Wu, Z.X., Jiang, C.L., Huang, Z.H., Lee, S.H., 2023. WPO: a whale particle optimization algorithm. Int. J. Comput. Intell. Syst. 16 (1), 115. https://doi.org/10.1007/s44196-023-00295-6.

Huang, X., Gu, L., Li, S., Du, Y., Liu, Y., 2022. Absolute adsorption of light hydrocarbons on organic-rich shale: an efficient determination method. Fuel 308, 121998. https://doi.org/10.1016/j.fuel.2021.121998.

Iskandar, U.P., Kurihara, M., 2022. Time-series forecasting of a CO_2-EOR and CO_2 storage project using a data-driven approach. Energies 15 (13), 4768. https://doi.org/10.3390/en15134768.

Jaber, A.K., 2022. Genetic algorithm to optimize miscible water alternate CO_2 flooding in heterogeneous clastic reservoir. Arabian J. Geosci. 15 (8), 714. https://doi.org/10.1007/s12517-022-09958-8.

Jaiswal, J.K., Samikannu, R., 2017. Application of random forest algorithm on feature subset selection and classification and regression. In: 2017 World Congress on Computing and Communication Technologies (WCCCT), pp. 65–68. https://doi.org/10.1109/WCCCT.2016.25.

Jiang, J., Rui, Z., Hazlett, R., Lu, J., 2019. An integrated technical-economic model for evaluating CO_2 enhanced oil recovery development. Appl. Energy 247, 190–211. https://doi.org/10.1016/j.apenergy.2019.04.025.

Jin, L., Pekot, L.J., Smith, S.A., Salako, O., Peterson, K.J., Bosshart, N.W., Hamling, J.A., Mibeck, B.A.F., Hurley, J.P., Beddoe, C.J., Gorecki, C.D., 2018. Effects of gas relative permeability hysteresis and solubility on associated CO_2 storage performance. Int. J. Greenh. Gas Control 75, 140–150. https://doi.org/10.1016/j.ijggc.2018.06.002.

Jarrell, P.M., Fox, C.E., Stein, M.H., Webb, S.L., 2002. Practical Aspects of CO_2 Flooding. Society of Petroleum Engineers. https://doi.org/10.2118/9781555630966.

Karimaie, H., Nazarian, B., Aurdal, T., Nøkleby, P.H., Hansen, O., 2017. Simulation study of CO_2 EOR and storage potential in a north sea reservoir. Energy Proc. 114, 7018–7032. https://doi.org/10.1016/j.egypro.2017.03.1843.

Kashkooli, S.B., Gandomkar, A., Riazi, M., Tavallali, M.S., 2022. Coupled optimization of carbon dioxide sequestration and CO_2 enhanced oil recovery. J. Petrol. Sci. Eng. 208, 109257. https://doi.org/10.1016/j.petrol.2021.109257.

Khather, M., Yekeen, N., Al-Yaseri, A., Al-Mukainah, H., Giwelli, A., Saeedi, A., 2022. The impact of wormhole generation in carbonate reservoirs on CO_2-WAG oil recovery. J. Petrol. Sci. Eng. 212. https://doi.org/10.1016/j.petrol.2022.110354.

Kolawole, O., Ispas, I., Kumar, M., Weber, J., Zhao, B., Zanoni, G., 2021. How can biogeomechanical alterations in shales impact caprock integrity and CO_2 storage? Fuel 291, 120149. https://doi.org/10.1016/j.fuel.2021.120149.

Kong, S., Huang, X., Li, K., Song, X., 2019. Adsorption/desorption isotherms of CH_4 and C_2H_6 on typical shale samples. Fuel 255, 115632. https://doi.org/10.1016/j.fuel.2019.115632.

Kulkarni, M.M., Rao, D.N., 2005. Experimental investigation of miscible and immiscible water-alternating-gas (WAG) process performance. J. Petrol. Sci. Eng. 48 (1–2), 1–20. https://doi.org/10.1016/j.petrol.2005.05.001.

Kumar, R., Campbell, S., Sonnenthal, E., Cunningham, J., 2020. Effect of brine salinity on the geological sequestration of CO_2 in a deep saline carbonate formation. Greenh. Gases Sci. 10 (2), 296–312. https://doi.org/10.1002/ghg.1960.

Kurz, B.A., Heebink, L.V., Eylands, K.E., Smith, S.A., et al., 2013. Bell Creek Test site-preinjection Geochemical Report Plains CO_2 Reduction (PCOR) Partnership Phase III Task 4-deliverable D33 Task 4-milestone M12.

Lake, L., Johns, R.T., Rossen, W.R., Pope, G.A., 2014. Fundamentals of Enhanced Oil Recovery. Society of Petroleum Engineers. https://doi.org/10.2118/9781613993286.

Lazreg, B., Syed, M.M., Akmal, A., 2020. Random forest algorithm for CO_2 water alternating gas incremental recovery factor prediction. Int. J. Adv. Sci. Technol. 29 (1), 168–188.

Land, C.S., 1968. Calculation of imbibition relative permeability for Two- and three-phase flow from rock properties. Soc. Petrol. Eng. J. 1 (8), 149–156. https://doi.org/10.2118/1942-pa, 02.

Li, H., Gong, C., Liu, S., Xu, J., Imani, G., 2022. Machine learning-assisted prediction of oil production and CO_2 storage effect in CO_2-Water-Alternating-Gas injection (CO_2-WAG). Appl. Sci. 12 (21). https://doi.org/10.3390/app122110958.

Mac Dowell, N., Fennell, P.S., Shah, N., Maitland, G.C., 2017. The role of CO_2 capture and utilization in mitigating climate change. Nat. Clim. Change 7 (4), 243–249. https://doi.org/10.1038/nclimate3231.

Mohagheghian, E., James, L.A., Haynes, R.D., 2018. Optimization of hydrocarbon water alternating gas in the norne field: application of evolutionary algorithms. Fuel 223, 86–98. https://doi.org/10.1016/j.fuel.2018.01.138.

Mohanasundaram, R., Malhotra, A.S., Arun, R., Periasamy, P.S., 2019. Deep learning and semi-supervised and transfer learning algorithms for medical imaging. Deep Learn. Parall. Comp. Environ. Bioeng. Syst. 139–151. https://doi.org/10.1016/B978-0-12-816718-2.00015-4.

Matthew, D.A.M., Jahanbani Ghahfarokhi, A., Ng, C.S.W., Nait Amar, M., 2023. Proxy model development for the optimization of water alternating CO_2 gas for enhanced oil recovery. Energies 16 (8), 3337. https://doi.org/10.3390/en16083337.

Naga, P.Y.V.R., Prakash, K.B., 2020. Impact of Inertia Weight and Cognitive and Social Constants in Obtaining Best Mean Fitness Value for PSO, pp. 197–206. https://doi.org/10.1007/978-981-15-0184-5_18.

Ng, C.S.W., Jahanbani Ghahfarokhi, A., Nait Amar, M., 2021. Application of nature-inspired algorithms and artificial neural network in waterflooding well control optimization. J. Pet. Explor. Prod. Technol. 11 (7), 3103–3127. https://doi.org/10.1007/s13202-021-01199-x.

Ng, C.S.W., Nait Amar, M., Jahanbani Ghahfarokhi, A., Imsland, L.S., 2023. A survey on the application of machine learning and metaheuristic algorithms for intelligent proxy modeling in reservoir simulation. Comput. Chem. Eng. 170, 108107. https://doi.org/10.1016/j.compchemeng.2022.108107.

Nait Amar, M., Jahanbani Ghahfarokhi, A., Ng, C.S.W., Zeraibi, N., 2021. Optimization of WAG in real geological field using rigorous soft computing techniques and nature-inspired algorithms. J. Petrol. Sci. Eng. 206, 109038. https://doi.org/10.1016/j.petrol.2021.109038.

Nikbakht, S., Anitescu, C., Rabczuk, T., 2021. Optimizing the neural network hyperparameters utilizing genetic algorithm. J. Zhejiang Univ.-Sci. 22 (6), 407–426. https://doi.org/10.1631/jzus.A2000384.

Oskouei, S.J.P., Maini, B.B., Moore, R.G., Mehta, S.A., 2012. Effect of initial water saturation on the thermal efficiency of the steam-assisted gravity-drainage process. J. Can. Pet. Technol. 51 (5), 351–361. https://doi.org/10.2118/138846-PA.

Pancholi, S., Negi, G.S., Agarwal, J.R., Bera, A., Shah, M., 2020. Experimental and simulation studies for optimization of water–alternating-gas (CO_2) flooding for enhanced oil recovery. Pet. Res. 5 (3), 227–234. https://doi.org/10.1016/j.ptlrs.2020.04.004.

Qolomany, B., Maabreh, M., Al-Fuqaha, A., Gupta, A.K., 2017. Parameters optimization of deep learning models using particle swarm optimization. In: The 13th International Wireless Communications and Mobile Computing Conference (IWCMC 2017). https://doi.org/10.48550/arXiv.1711.10354.

Ren, B., Duncan, I.J., 2021. Maximizing oil production from water alternating gas (CO_2) injection into residual oil zones: the impact of oil saturation and heterogeneity. Energy 222. https://doi.org/10.1016/j.energy.2021.119915.

Ren, D., Wang, X., Kou, Z., Wang, S., Wang, H., Wang, X., Tang, Y., Jiao, Z., Zhou, D., Zhang, R., 2023. Feasibility evaluation of CO_2 EOR and storage in tight oil reservoirs: a demonstration project in the Ordos basin. Fuel 331, 125652. https://doi.org/10.1016/j.fuel.2022.125652.

Rodrigues, H., Mackay, E., Arnold, D., Silva, D., 2019. Optimization of CO_2-WAG and calcite scale management in pre-salt carbonate reservoirs. In: Paper Presented at the Offshore Technology Conference Brasil, Rio De Janeiro, Brazil. https://doi.org/10.4043/29823-MS.

Sallam, S., Ahmad, M.M., Nasr, M., 2015. The Effect of Water Injection on Oil Well Productivity the Effect of Water Injection on the Oil Well Productivity the Effect of Water Injection on Oil Well Productivity. Flexible Automation and Intelligent Manufacturing, FAIM2015. https://doi.org/10.13140/RG.2.1.2892.8486.

Sanchez, N.L., 1999. Management of water alternating gas (WAG) injection projects. In: Latin American and Caribbean Petroleum Engineering Conference. https://doi.org/10.2118/53714-MS.

Sen, D., Chen, H., Datta, G.A., 2022. Inter-well connectivity detection in CO_2 WAG projects using statistical recurrent unit models. Fuel 311, 122600. https://doi.org/10.1016/j.fuel.2021.122600.

Shanthi, D.L., Chethan, N., 2022. Genetic algorithm based hyper-parameter tuning to improve the performance of machine learning models. SN Comp. Sci. 4 (2), 119. https://doi.org/10.1007/s42979-022-01537-8.

Shen, B., Yang, S., Gao, X., Li, S., Ren, S., Chen, H., 2023. A novel CO_2-EOR potential evaluation method based on BO-LightGBM algorithms using hybrid feature mining. Geoenergy Sci. Eng. 222, 211427. https://doi.org/10.1016/j.geoen.2023.211427.

Subasi, A., 2020. Machine Learning Techniques. Practical Machine Learning for Data Analysis Using Python, pp. 91–202. https://doi.org/10.1016/B978-0-12-821379-7.00003-5.

Sun, X., Liu, J., Dai, X., Wang, X., Yapanto, L.M., Zekiy, A.O., 2021. On the application of surfactant and water alternating gas (SAG/WAG) injection to improve oil recovery in tight reservoirs. Energy Rep. 7, 2452–2459. https://doi.org/10.1016/j.egyr.2021.04.034.

Tang, Y., Hou, C., He, Y., Wang, Y., Chen, Y., Rui, Z., 2021. Review on pore structure characterization and microscopic flow mechanism of CO_2 flooding in porous media. Energy Technol 9 (1), 2000787. https://doi.org/10.1002/ente.202000787.

Vo Thanh, H., Sugai, Y., Nguele, R., Sasaki, K., 2020. Robust optimization of CO_2 sequestration through a water alternating gas process under geological uncertainties in Cuu Long Basin, Vietnam. J. Nat. Gas Sci. Eng. 76, 103208. https://doi.org/10.1016/j.jngse.2020.103208.

Wan, Y., Du, S., Lin, G., Zhang, F., Xu, T., 2017. Dissolution sequestration mechanism of CO_2 at the Shiqianfeng saline aquifer in the Ordos Basin, northwestern China. Arabian J. Geosci. 10 (3), 71. https://doi.org/10.1007/s12517-017-2858-7.

Wang, L., Yao, Y., Luo, X., Daniel A, C., Zhao, G., Lai, F., 2023. A critical review on intelligent optimization algorithms and surrogate models for conventional and unconventional reservoir production optimization. Fuel 350, 128826. https://doi.org/10.1016/j.fuel.2023.128826.

Xue, L., Li, D., Dou, H., 2023a. Artificial intelligence methods for oil and gas reservoir development: current progresses and perspectives. Adv. Geo-Energy Res. 10 (1), 65–70. https://doi.org/10.46690/ager.2023.10.07.

Xue, L., Wang, J., Han, J., Yang, M., Mwasmwasa, M., Nanguka, F., 2023b. Gas well performance prediction using deep learning jointly driven by decline curve analysis model and production data. Adv. Geo-Energy Res. 8 (3), 159–169. https://doi.org/10.46690/ager.2023.06.03.

Xue, L., Xu, S., Nie, J., Qin, J., Han, J.X., Liu, Y.T., Liao, Q.Z., 2024. An efficient data-driven global sensitivity analysis method of shale gas production through convolutional neural network. Pet. Sci. 21 (4), 2475–2484. https://doi.org/10.1016/j.petsci.2024.02.010.

Xue, L., Zhu, Y., Ren, J., Liao, H., Dai, Q., Tu, B., 2025. Coupled optimization method for CO_2-EOR and storage based on machine learning. J. Porous Media 28 (7), 37–53. https://doi.org/10.1615/JPorMedia.2024052865.

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 CO_2-EOR project. J. Clean. Prod. 260. https://doi.org/10.1016/j.jclepro.2020.120866.

Zubarev, D.I., 2009. Pros and cons of applying proxy-models as a substitute for full reservoir simulations. In: SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/124815-MS.

Zhong, Z., Liu, S., Carr, T.R., Takbiri-Borujeni, A., Kazemi, M., Fu, Q., 2019. Numerical simulation of water-alternating-gas process for optimizing EOR and carbon storage. Energy Proc. 158, 6079–6086. https://doi.org/10.1016/j.egypro.2019.01.507.