Journal Homepage: http://ijcpe.uobaghdad.edu.iq
Iraqi Journal of Chemical and Petroleum Engineering
Vol. 27 No. 1 (March 2026) 85 – 92
EISSN: 2618-0707, PISSN: 1997-4884

Multi-output machine learning prediction of CO_2 enhanced oil recovery and geological storage performance

Abdullah A. Hashim [a, *], **Farqad A. Hadi** [a], **Dheiaa Alfarge** [b]

a Petroleum Engineering Department, College of Engineering, University of Baghdad, Baghdad, Iraq
b Oil and Gas Engineering Department, College of Engineering, University of Warith Al-Anbiyaa, Iraq

Abstract

Carbon dioxide-enhanced oil recovery (CO_2-EOR) is one of the best commercial carbon capture, utilization, and storage (CCUS) technologies, enhancing oil recovery and achieving safe geological storage. However, data density and computational costs hinder conventional reservoir simulations. This research presents a multi-output machine learning model to simultaneously predict both incremental oil production and CO_2 storage volume. A selected global dataset of 173 projects was compiled, preprocessed, and optimized. Among the evaluated algorithms (random forest, gradient boosting, and random forest chain), gradient boosting achieved the best predictive performance, with a weighted coefficient of determination (R^2) of 0.722. Model analysis showed that permeability, depth, and miscibility strongly influence oil recovery (R^2 for gradient boosting = 0.87), while the predictability of storage volume is low (R^2 for gradient e boosting = 0.5), which is significantly affected by reservoir depth and porosity. This study highlights the potential of machine learning as a rapid screening tool for evaluating carbon capture, utilization, and storage projects, identifying the performance of key factors, and emphasizing the need for standardized storage reporting to improve prediction accuracy.

Keywords: CO_2-EOR; CO_2-Storage; CCUS; Subsurface Storage; Geological Field Data; Machine Learning.

Received on 18/01/2026, Received in Revised Form on 14/03/2026, Accepted on 14/03/2026, Published on 30/03/2026

https://doi.org/10.31699/IJCPE.2026.1.7

1- Introduction

With global energy demand continuing to rise and concerns about human-induced climate change growing, technologies for utilizing and storing carbon dioxide have emerged as radical solutions. Among these technologies, carbon dioxide-enhanced oil recovery (CO_2-EOR) represents one of the most commercially viable applications [1]. Injecting carbon dioxide into oil reservoirs can achieve two objectives: first, enhance oil recovery through miscible or immiscible displacement mechanisms, and second, safely store carbon dioxide in geological formations, contributing to carbon capture, utilization, and storage strategies (CCUS) [2].

Conventional reservoir simulations require extensive data inputs (geological and fluid characteristics, etc.) and specialized expertise. These requirements often delay a rapid site investigation [3]. To address these challenges, many recent studies have turned to exploring artificial intelligence techniques, including hybrid neural networks and genetic algorithms, to improve injection strategies [4].

Predicting the behavior of oil reservoirs has become easier thanks to developments in artificial intelligence and machine learning technologies. Machine learning algorithms can handle the nonlinear relationships and interactions found in reservoir engineering [5]. Applications in petroleum engineering include production prediction [6] reservoir characterization [7], and predicting the performance of the carbon capture, utilization, and storage (CCUS) process [8].

A number of studies have investigated individual enhanced oil recovery projects, and some have developed artificial intelligence models for specific reservoirs. For instance, Zhou et al. [3] provided a comprehensive review of machine learning applications in reservoir engineering, noting that while single-reservoir models achieve high accuracy, their generalizability remains limited. Similarly, various methods for estimating geological storage capacity range from simple volumetric calculations to complex dynamic reservoir simulations [9]. Although researchers have developed empirical relationships to predict storage volume, these often lack generalizability across different geological environments [10].

Machine learning algorithms can capture complex, nonlinear relationships directly from data without the need for explicit physical equations. This capability makes them particularly well-suited for problems involving complex and rapid prediction requirements [11]. In petroleum engineering and CCUS applications, methods such as random forests, support vector machines, and artificial neural networks have been widely used to predict oil production in both conventional and

---

*Corresponding Author: Email: abd.abbas2208@coeng.uobaghdad.edu.iq

unconventional reservoirs and address CCUS challenges, including forecasting CO_2 plume migration, estimating storage capacity, optimizing well placement, and designing monitoring systems.

However, a critical gap remains in the literature: no research based on real-world, diverse field datasets has yet been conducted to create a general predictive model capable of simultaneously forecasting both oil recovery and CO_2 storage. Previous studies, such as those by Sun [8] and Aminu et al. [9], have either focused on simulated data or developed site-specific correlations that cannot be applied across different geological settings.

Multiple related objectives are predicted simultaneously using multi-output regression, which provides advantages over separate single-objective models when the outputs are physically connected [12]. Techniques such as multi-output regression models, regression chains, and clustering methods have demonstrated promising results in environmental and energy applications [13].

This work aims to fill this gap by: (1) compiling a global dataset of enhanced oil recovery and carbon capture and storage projects from real-world field data; (2) developing a multi-output machine learning model, trained on this diverse dataset, for predicting both incremental oil production and CO_2 storage volume; and (3) providing an interpretable analysis of the significance of key reservoir characteristics on both objectives, thereby offering a rapid screening tool for CCUS project evaluation.

2- Data collection and methodology

The quality of any data-driven study fundamentally depends on a high-quality, well-curated dataset. This research relies on field data from published CO_2-EOR projects.

2.1. Data sources

In this study, the analysis was based on a large dataset (173 case studies) collected through numerous published research papers, global reports, and case studies. This multi-source approach was used to ensure an accurate representation of global CO_2 EOR and storage projects.

2.2. Types of data and variables

- Geological and petrophysical parameters: Reservoir depth, thickness, temperature, pressure, porosity, permeability.
- Fluid properties: Oil API gravity, oil viscosity.
- Operational parameters: minimum miscible pressure (MMP)
- Performance parameters (Target Variables) incremental oil production and estimated CO_2 storage volume.

2.3. Preprocessing of data

- Handling Missing Values: Parameters with a high proportion of missing data (>70% missing values) were excluded from quantitative statistical analyses (e.g., correlation studies, mean calculations). In contrast, parameters with a small percentage of missing data were not subjected to any procedure, this is to ensure that all of the presented statistics and conclusions were based only on the reported measurements and avoid the introduction of potential bias. Removes any row missing either target, 173 rows become 105 rows. It can't predict missing targets in supervised learning. For feature variables, replaces missing values with column average (median for numbers, mode for categories).
- Unit Standardization: All values have been converted to standard units to enable accurate comparisons and correlations. Depth is expressed in meters, pressure in psi, permeability in millidarcies (mD), viscosity in centipoise (cP), and temperature in degrees Celsius.
- Feature Engineering for a machine learning model development:

1. Reservoir quality index (RQI), in μm, quantifies the flow capacity of reservoir rock by combining permeability and porosity (Eq. 1), where k is permeability, ϕ is porosity, and 0.0314 is a conversion factor. RQI combines permeability and porosity into a single quality metric [14].

$$RQI = 0.0314 \times \sqrt{(k / \phi)} \tag{1}$$

2. Flow zone indicator (FZI) classifies the reservoir into hydraulic flow units with similar pore characteristics as presented in Eq. 2, where φz represents the normalized porosity, which is [ϕ / (1 - ϕ)]. FZI is Better for heterogeneous reservoirs [14].

$$FZI = RQI / (\phi z) \tag{2}$$

3. The miscibility ratio (MR) is defined as the ratio of the reservoir pressure to the minimum miscible pressure (Eq. 3).

$$MR = \frac{P_{reservoir}}{MMP} \tag{3}$$

4. Theoretical storage capacity (SC) represents the potential amount of stored CO_2 in a formation. It highlights the importance of pore volume. As presented in Eq. 4, it can be calculated using porosity (ϕ), thickness (h), areal extent (A), density of CO_2 (ρ_{CO2}), and the storage efficiency factor (E).

$$SC = \phi \times h \times A \times \rho_{CO2} \times E \tag{4}$$

5. Mobility ratio (MI) quantifies the ease with which fluid flows in the reservoir. It is defined as the ratio of permeability to fluid viscosity (μ).

$$MI = \frac{k}{\mu} \tag{5}$$

6. Porosity-permeability ratio: Captures the nonlinear relationship between porosity and permeability, Eq. 6 [15].

$$\phi - k\ ratio\ =\ (\phi/100)\ /\ ln(1+k) \qquad (6)$$

2.4. Machine learning model

The primary objective of building the model is to perform multi-objective forecasting of incremental oil production and CO_2 storage. A range of common machine learning algorithms will be implemented and compared to ensure robustness, including Random Forest, Gradient Boosting, and Random Forest Chain. The choice of these algorithms was justified by their proven success in solving similar problems in petroleum engineering, including production forecasting.

Model training and validation

1. *Data splitting*: The compiled dataset was randomly split into a training set (80%) and a hold-out test set (20%).
2. *Cross-validation*: The training set was further used in a 5-fold cross-validation scheme to tune the model hyperparameters. This process helps in reducing the risk of overfitting and provides a more reliable estimate of the model's performance [16].
3. *Performance metrics*: The developed models were evaluated and compared using two performance metrics: The determination coefficient (R^2), which represents the proportion of variance in the target variable explained by the model; the Root Mean Squared Error (RMSE), which represents the average magnitude of the predicted errors in the units of the target variable.

3- Results and discussion

3.1. Most Influential Parameters on Incremental Oil Production

Average Permeability: As can be seen from Fig. 1, the average permeability and incremental oil production are strongly and positively correlated. This can be explained by the correlation between injectivity and permeability, meaning that the higher the permeability, the higher the injection rate and, consequently, the higher the displacement efficiency, which increases the production rate. According to Chen et al. [15] the oil recovery increases as the average permeability increases.

Average porosity: Fig. 2, shows average porosity versus incremental oil production. There is clear scatter in the results, indicating that oil production does not depend solely on porosity. This is because porosity is a measure of the reservoir's ability to store oil (hydrocarbon pore volume), not its ability to flow or produce oil. While porosity is necessary for oil in place, it is permeability and other flow-related properties that govern recovery efficiency.

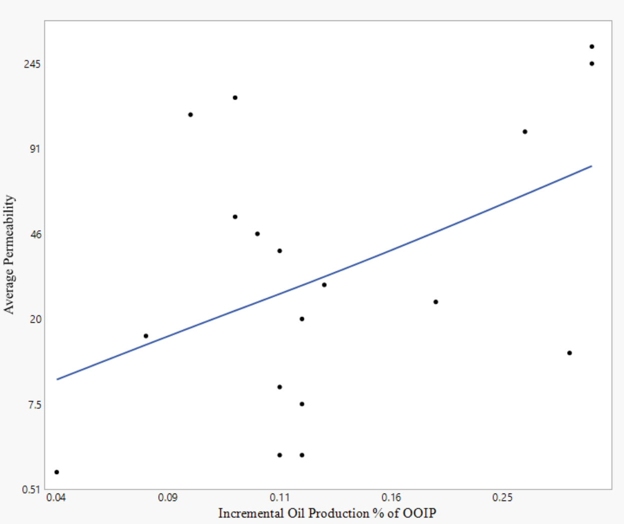

Fig. 1. Incremental oil production vs average permeability

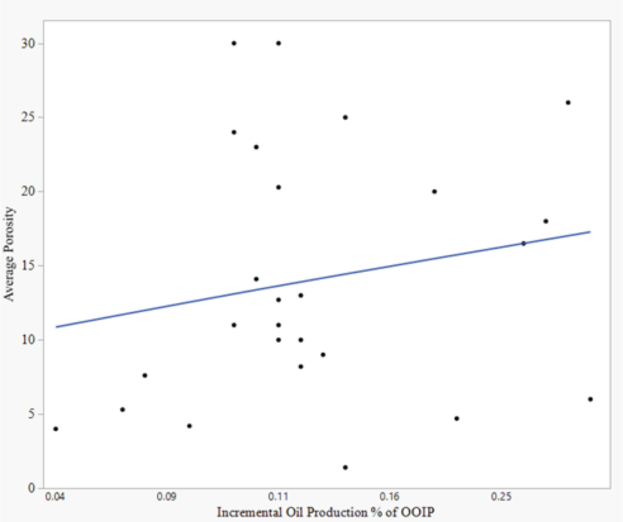

Fig. 2. Incremental oil production vs average porosity

Reservoir temperature: Fig. 3, shows the relationship between temperature and incremental oil production. The results indicate that temperature does not directly influence incremental oil production. Instead, temperature exerts an indirect effect through its relationship with oil viscosity and minimum miscibility pressure (MMP). Higher temperatures typically reduce oil viscosity, which can improve mobility, but they also increase MMP, making miscibility more difficult to achieve. This complex interaction explains the lack of a clear direct correlation.

Oil API: Fig. 4, shows the relationship between API gravity and incremental oil production. The positive trend indicates that oil production increases as API gravity increases. This is consistent with the physics of miscible displacement: lighter oils (higher API) achieve miscibility with CO_2 more readily and at lower pressures, leading to more efficient oil recovery, as documented in the classic work of Holm and Josendal [13] on the mechanisms of oil displacement by carbon dioxide.

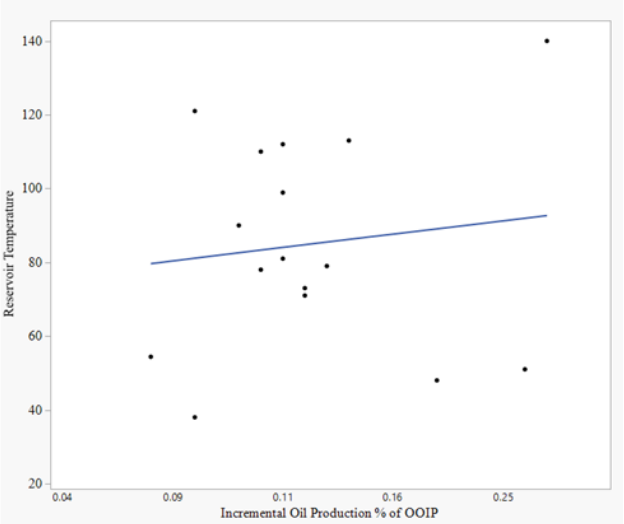

Fig. 3. Incremental oil production vs Reservoir Temperature

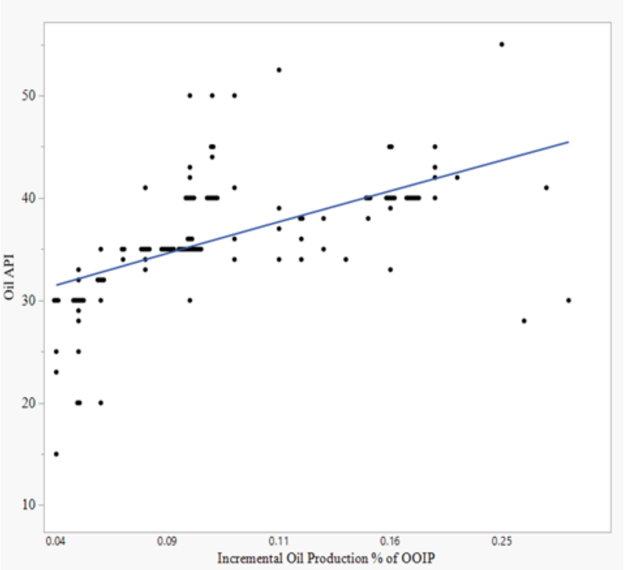

Fig. 4. Incremental oil production vs Oil API

Oil viscosity: Implementation of carbon dioxide would be of significance in unconventional reservoirs, as it can reduce the crude oil viscosity during the injection of CO_2 (Carbon dioxide) and increase the sweep efficiency in porous media [16]. As shown in Fig. 5, as viscosity decreases, the incremental oil production increases. This inverse relationship reflects the improved mobility ratio achieved when a less viscous CO_2 bank displaces oil.

3.2. Most influential parameters on CO_2 storage volume

Average Porosity: Fig. 6, shows a clear positive relationship between porosity and CO_2 storage volume. This is expected because porosity directly determines the pore volume available for fluid storage. As expressed in the theoretical storage capacity equation (Eq. 4), porosity is a primary control on the maximum potential CO_2 that can be stored in a formation. The scatter in the data reflects the additional influence of other factors such as reservoir thickness, areal extent, and CO_2 density, which vary with depth and temperature.

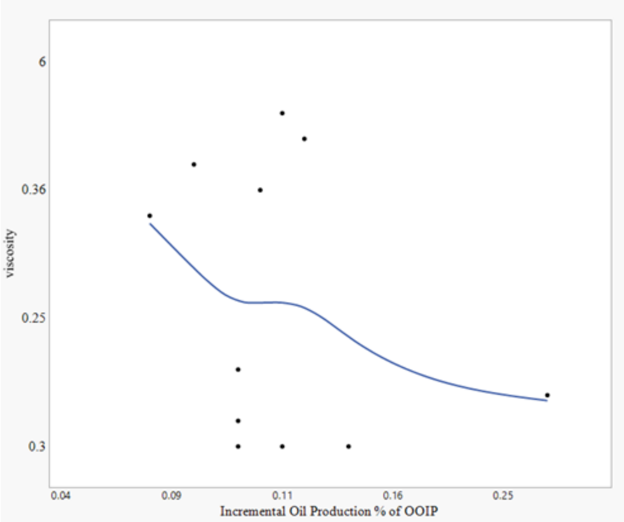

Fig. 5. Incremental oil production vs Oil viscosity

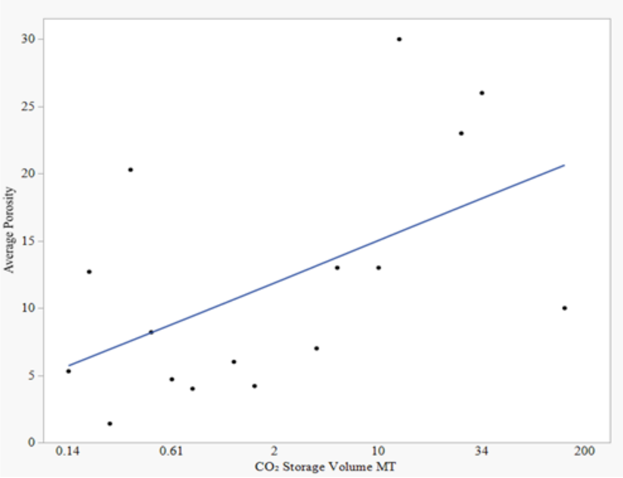

Fig. 6. CO_2 storage volume vs average porosity

3.3 Overall model performance

Table 1 summarizes the performance of the evaluated algorithms for predicting oil recovery and CO_2 storage. The comparative evaluation of the machine learning models reveals that the Gradient Boosting algorithm yields the most precise and resilient predictions for both oil recovery and CO_2 sequestration. The Gradient Boosting provides the highest weighted determination coefficient (R^2) of 0.722 (with an average of 0.685) rather than other machine learning models.

The high R^2 value for incremental oil production (0.87) indicates that reservoir and fluid properties captured in our dataset strongly influence the EOR process. As shown in Fig. 7, the predicted values closely align with actual field measurements. Feature importance analysis revealed that permeability, depth, miscibility ratio, and formation type were the most influential parameters. The strong performance demonstrates that machine learning can effectively learn the physical relationships governing miscible and immiscible CO_2 displacement from field data. Notably, the predictive power of our general model ($R^2 = 0.87$) is comparable to or exceeds that of many single-reservoir neural network models reported in the literature [3]. However, the key advantage of our approach lies in its generalizability across diverse geological settings, addressing a limitation of previous empirically-derived relationships [11].

Table 1. Model performance comparison

Model	Weighted R^2	Avg. R^2	Oil Recovery R^2	CO_2 Storage R^2	Oil Recovery RMSE	CO_2 Storage RMSE
Random Forest	0.600	0.553	0.799	0.306	0.038	1.523
Gradient Boosting	0.722	0.685	0.870	0.500	0.028	1.206
Random Forest Chain	0.558	0.500	0.793	0.206	0.039	1.694

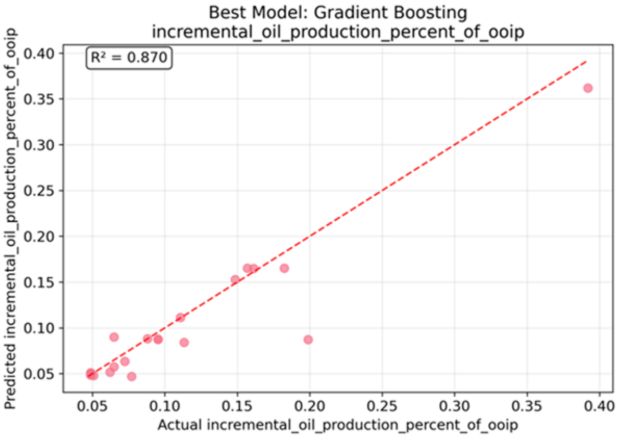

Fig. 7. incremental oil production, predicted vs actual

In contrast, all models exhibited significantly lower R^2 values for carbon dioxide storage prediction (Table 1, Fig. 8). The Gradient Boosting model achieved an R^2 of 0.50 with an RMSE of 1.206, which while the best among the tested algorithms, indicates a high level of uncertainty. This disparity in predictability between oil recovery and CO_2 storage is a key finding of this study.

This reduced predictive ability was interpreted as a direct result of data quality and reporting standards, rather than a failure of the modeling approach. While oil production is an economically vital metric that has been measured and reported with relative consistency for decades, CO_2 storage volume is a newer metric with no standardized reporting protocols. As discussed in Section 3.4, the lack of consistent definitions, measurement techniques, and reporting frameworks for "stored" CO_2 introduces significant noise and uncertainty into the target variable. This makes it fundamentally more difficult for any model to learn the underlying patterns. This result underscores the critical need for improved CO_2 storage measurement and reporting standards, as noted by previous researchers [12, 14]. Despite this limitation, our achieved R^2 of 0.5 provides a valuable real-world benchmark for future studies, as most prior machine learning work on CO_2 storage has relied on simulated rather than field data [8].

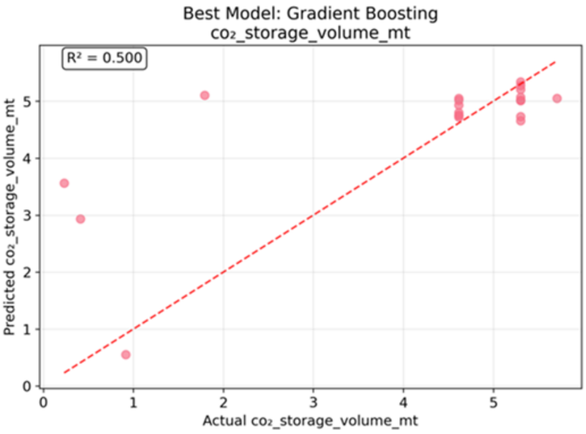

Fig. 8. CO_2 storage volume, actual vs predicted

According to the results of this study, the significance analysis indicates that formation permeability, depth, and miscibility are the most important factors affecting oil recovery. In contrast, carbon dioxide storage volume is primarily affected by reservoir thickness, porosity, and depth. For both objectives (oil recovery and carbon dioxide storage volume), engineered features, such as reservoir quality index (RQI) and storage efficiency, proved more important than many of the raw input variables.

3.4. Dataset limitations

- *Data completeness*: one of the challenges was the existence of missing data for key criteria in many of the 173 cases. This missing data can introduce uncertainty into statistical analysis and the identification of trends.
- *Reporting bias*: The dataset potentially suffers from a positive reporting bias, as successful and ongoing projects are more frequently documented and published than failed or suspended projects. This may lead to an overly optimistic representation of the average performance and success rates of CO_2 EOR projects.
- *Data variability*: The data were collected from multiple public sources, which may have used different measurement criteria, reporting protocols, and definitions (such as "successful"), which may affect the consistency of the analysis.
- *Storage mechanisms*: that will change the result of the storage model.

4- Conclusion

This study presents a global, field-scale machine learning framework for forecasting the incremental oil recovery and CO_2 storage volume through the application of multi-output regression using physics-informed engineered features with a variety of real-world field data. The findings of this study can be summarized in the following key points:

1. The results show that the influencing parameters on the incremental oil production and CO_2 storage volume are different. The incremental oil production is influenced by average permeability, depth, and miscibility ratio, while the CO_2 storage volume is influenced by reservoir thickness, porosity, and depth.
2. Considering the composite physical features can support better capture of the complex interactions between reservoir properties. For example, the Reservoir Quality Index (RQI) and the miscibility ratio support better results for both incremental oil production and CO2 storage volume, rather than several individual inputs.
3. Gradient Boosting successfully predicted both incremental oil production and CO_2 storage volume with a weighted R^2 of 0.722, demonstrating ML's potential for coupled CCUS performance prediction.
4. incremental oil production ($R^2 = 0.87$) proved more predictable than CO_2 storage volume ($R^2 = 0.5$), Demonstrating the need for improved CO_2 storage measurement and reporting standards.
5. Using a diverse, real-world field data of CO_2-EOR and storage projects improves the models' generality, robustness, and applicability across different geological environments.

References

[1] V. A. Kuuskraa, M. L. Godec, and P. Dipietro, "CO_2 utilization from 'next generation' CO_2 enhanced oil recovery technology," in *Energy Procedia*, Elsevier Ltd, 2013, pp. 6854–6866. Volume: 37 https://doi.org/10.1016/j.egypro.2013.06.618

[2] M. M. M. Alam, A. Hassan, M. Mahmoud, N. Sibaweihi, and S. Patil, "Dual Benefits of Enhanced Oil Recovery and CO2 Sequestration: The Impact of CO2 Injection Approach on Oil Recovery," *Front Energy Res*, vol. 10, Apr. 2022, https://doi.org/10.3389/fenrg.2022.877212

[3] W. Zhou, C. Liu, Y. Liu, Z. Zhang, P. Chen, and L. Jiang, "Machine Learning in Reservoir Engineering: A Review," Jun. 01, 2024, *Multidisciplinary Digital Publishing Institute,* Volume: 12, Issue: 6, Article Number: 1219 (MDPI). https://doi.org/10.3390/pr12061219

[4] S. B. Anieto, E. G. Chukwuka, A. Adekomi, O. O. Abdullahi, E. O. Great, C. I. Collins, A. Fifo, and C. A. Chinonyerem, "machine learning assisted optimization of Carbon Capture and storage in oil and gas reservoirs through Advanced Geomodelling," *International Journal of Engineering Processing and Safety Research*, Nov. 2025 Volume: 9, Issue: 5, pp. 14-24, https://doi.org/10.70382/caijepsr.v9i5.016

[5] T. Kyriazos and M. Poga, "Application of Machine Learning Models in Social Sciences: Managing Nonlinear Relationships," *Encyclopedia*, vol. 4, no. 4, pp. 1790–1805, Dec. 2024, https://doi.org/10.3390/encyclopedia4040118

[6] N. Qi, X. Li, Z. Wu, Y. Wan, N. Wang, G. Duan, L. Wang, J. Xiang, Y. Zhao, and H. Zhan,, "Machine Learning-Based Research for Predicting Shale Gas Well Production," *Symmetry (Basel)*, vol. 16, no. 5, May 2024, https://doi.org/10.3390/sym16050600

[7] B. I. Odoh, C. V. Ahaneku, C. D. Muogbo, C. N. Mmaduka, D. S. Aderokaye, H. C. Anozie, C. P. Arukwe-Moses, I. E. Eze, V. O. Oguadinma, and R. O. Arinze, "The Impact of Artificial Intelligence and Machine Learning on Reservoir Characterization: A Review of Recent Advances," 2024, *International Journal of Research and Innovation in Applied Science*. Volume: 9, Issue: 7, pp. 625-637 https://doi.org/10.51584/IJRIAS.2024.907052

[8] Y. Sun, "Machine Learning Models for CO Storage Estimation in Natural Gas Reservoirs," *Applied and Computational Engineering*, vol. 179, no. 1, pp. 90–94, Aug. 2025, https://doi.org/10.54254/2755-2721/2025.LD26467

[9] M. D. Aminu, S. A. Nabavi, C. A. Rochelle, and V. Manovic, "A review of developments in carbon dioxide storage," *Applied Energy*, 2017 Volume: 208, pp. 1389-1419. https://doi.org/10.1016/j.apenergy.2017.09.015

[10] Y. Sun, Q. Sun, and K. Qin, "Physics-based deep learning for flow problems," *Energies (Basel)*, vol. 14, no. 22, Nov. 2021 Volume: 14, Issue: 22, Article Number: 7760, https://doi.org/10.3390/en14227760

[11] J. O. Amaefule, M. Altunbay, D. G. Kersey, and D. K. Keelan, "Enhanced Reservoir Description: Using Core and Log Data to Identify Hydraulic (Flow) Units and Predict Permeability in Uncored Intervals/Wells," 1993. *SPE Annual Technical Conference and Exhibition*, Houston, Texas, https://doi.org/10.2118/26436-MS

[12] X. Shi, H. Bo, Q. Zheng, F. Yi, Y. Yin, Y. Chen, M. Dong, D. Zhang, J. Li, and X. Zhou. "A novel model for oil recovery estimate in heterogeneous low-permeability and tight reservoirs with pseudo threshold pressure gradient," *Energy Reports*, vol. 7, pp. 1416–1423, Nov. 2021, https://doi.org/10.1016/j.egyr.2021.01.017

[13] L. W. Holm and V. A. Josendal, "Mechanisms of Oil Displacement By Carbon Dioxide," *Journal of Petroleum Technology*, 1974 Volume: 26, Issue: 12, pp. 1427-1438. https://doi.org/10.2118/4736-PA

[14] S. Bachu, "Review of CO2 storage efficiency in deep saline aquifers," *International Journal of Greenhouse Gas Control*, 2015 Volume: 40, pp. 188-202. https://doi.org/10.1016/j.ijggc.2015.01.007

[15] P. Chen, S. Bose, A. Selveindran, and G. Thakur, "Application of CCUS in India: Designing a CO2 EOR and storage pilot in a mature field," *International Journal of Greenhouse Gas Control*, vol. 124, Mar. 2023, https://doi.org/10.1016/j.ijggc.2023.103858

[16] Y. Wang, Z. Jiao, R. Surdam, L. Zhou, R. Gao, Y. Chen, T. Luo, and H. Wang. "A feasibility study of the integration of enhanced oil recovery (CO2 flooding) with CO2 storage in the mature oil fields of the Ordos Basin, China," *Geological CO2 Storage Characterization*, 2013, vol. 37, pp. 6846–6853. https://doi.org/10.1007/978-1-4614-5788-6_13

[17] O. Hedriana, Sugihardjo, and Usman, "Assessment of CO2 - EOR and Storage Capacity in South Sumatra and West Java Basins," in *Energy Procedia*, Elsevier Ltd, 2017, vol:114 pp. 4666–4678. https://doi.org/10.1016/j.egypro.2017.03.1598

[18] S. Hannisa, J. Lub, A. Chadwicka, S. Hovorkab, K. Kirka, K. Romanakb, and J. Pearce, "CO2 Storage in Depleted or Depleting Oil and Gas Fields: What can We Learn from Existing Projects?," in *Energy Procedia*, Elsevier Ltd, 2017, vol:114 pp. 5680–5690. https://doi.org/10.1016/j.egypro.2017.03.1707

[19] V. Honari, J. Underschultz, X. Wang, A. Garnett, X. Wang, R. Gao, and Q. Liang, "CO2 sequestration-EOR in Yanchang oilfield, China: Estimation of CO2 storage capacity using a two-stage well test," in *Energy Procedia*, Elsevier Ltd, 2018, Volume: 154 pp. 9–14. https://doi.org/10.1016/j.egypro.2018.11.003

[20] S. S. Ganguli, N. Vedanti, I. Akervoll, and V. P. Dimri, "Assessing the Feasibility of CO2-Enhanced Oil Recovery and Storage in Mature Oil Field: A Case Study from Cambay Basin," *Journal of the Geological Society of India*, 2016 Volume: 88, Issue: 3, pp. 273-280. https://doi.org/10.1007/s12594-016-0490-x

[21] F. Gozalpour, S. R. Ren, and B. Tohidi, "CO2 EOR and Storage in Oil Reservoirs," *Oil & Gas Science and Technology*, 2005 Volume: 60, Issue: 3, pp. 537-546. https://doi.org/doi:10.2516/ogst:2005036

[22] G. WANG, "Carbon dioxide capture, enhanced-oil recovery and storage technology and engineering practice in Jilin Oilfield, NE China," *Petroleum Exploration and Development*, vol. 50, no. 1, pp. 245–254, Feb. 2023, https://doi.org/10.1016/S1876-3804(22)60384-7

[23] Y. Zhang, L. Zhang, B. Niu, and S. Ren, "Integrated assessment of CO2-enhanced oil recovery and storage capacity," *in Society of Petroleum Engineers - Canadian Unconventional Resources and International Petroleum Conference 2010*, Society of Petroleum Engineers, 2010, pp. 1780–1788. https://doi.org/10.2118/137615-MS

[24] S. Mishra, A. Haagsma, M. Valluri, and N. Gupta, "Assessment of CO2-enhanced oil recovery and associated geologic storage potential in the Michigan Northern Pinnacle Reef Trend," *Greenhouse Gases: Science and Technology*, vol. 10, no. 1, pp. 32–49, Feb. 2020, https://doi.org/10.1002/ghg.1944

[25] L. Zhang, B. Ren, H. Huang, Y. Li, S. Ren, G. Chen, and H. Zhang"CO2 EOR and storage in Jilin oilfield, China: Monitoring program and preliminary results," *J Pet Sci Eng*, vol. 125, pp. 1–12, Jan. 2015, https://doi.org/10.1016/j.petrol.2014.11.005

[26] Y. Tan, Q. Li, L. Xu, A. Ghaffar, X. Zhou, and P. Li, "A critical review of carbon dioxide enhanced oil recovery in carbonate reservoirs," *Fuel*, vol. 328, Nov. 2022, https://doi.org/10.1016/j.fuel.2022.125256

[27] A. O. Gbadamosi, J. Kiwalabye, R. Junin, and A. Augustine, "A review of gas enhanced oil recovery schemes used in the North Sea," Dec. 01, 2018, *Springer Verlag. Journal of Petroleum Exploration and Production Technology*. Volume: 8, Issue: 4, pp. 1373-1387. https://doi.org/10.1007/s13202-018-0451-6

[28] P. O. Longe, J.-S. Tsau, S. Musgrove, J. Villalobos, J. D'Erasmo, M. M. Alhajeri, and R. Barati "An Overview of Stewart Field Unit Project: A Field Case Study of CO2 Capture, Utilization, and Storage," *in Society of Petroleum Engineers - SPE Energy Transition Symposium*, ETS 2024, Society of Petroleum Engineers, 2024. https://doi.org/10.2118/221422-MS

[29] Z. Daia, H. Viswanathana, T. Xiaob, R. Middletona, F. Panb, W. Ampomahc, C. Yangd, Y. Zhoud, W. Jiab, S.-Y. Leee, M. Catherc, R. Balchc, and B. McPherson"CO2 Sequestration and Enhanced Oil Recovery at Depleted Oil/Gas Reservoirs," in *Energy Procedia*, Elsevier Ltd, 2017, Volume: 114 pp. 6957–6967. https://doi.org/10.1016/j.egypro.2017.08.034

[30] A.R. Awan, R. Teigland, and J. Kleppe, "EOR Survey in the North Sea," *SPE Symposium on Improved Oil Recovery*, 2006. https://doi.org/10.2118/99546-MS

[31] E. P. Robertson, "Review of Ongoing CO2 Enhanced Oil Recovery Operations in Wyoming, USA," in *Proceedings - SPE Symposium on Improved Oil Recovery*, Society of Petroleum Engineers (SPE), 2024. https://doi.org/10.2118/218237-MS

[32] D. S. Mahdi, D. Al-bayati, A. Saeedi, M. Myers, and C. White, "Mechanisms of Miscible and Immiscible scCO2 Displacement Efficiency: Analytical Evaluation of Experimental conditions," *Iraqi Journal of Oil and Gas Research (IJOGR)*, vol. 2, no. 1, pp. 98–107, Apr. 2022, http://doi.org/10.55699/ijogr.2022.0201.1019

[33] A. Dudun, F. Boukadi, J. Osumanu, "Comparative Study of CO2 Storage Capacity Estimation in Depleted Oil & Gas Reservoir: A Case Study in Vermillion Basin Gulf of Mexico," *Petroleum & Petrochemical Engineering Journal*, vol. 8, no. 1, pp. 1–10, Jan. 2024, https://doi.org/10.23880/ppej-16000379

التنبؤ بالتعلم الآلي متعدد المخرجات لتعزيز استرجاع النفط وأداء التخزين الجيولوجي

عبدالله عباس هاشم [١، *]، فرقد علي هادي [١]، ضياء الفرج [٢]

[١] قسم هندسة النفط، كلية الهندسة، جامعة بغداد، بغداد، العراق

[٢] قسم هندسة النفط والغاز، كلية الهندسة، جامعة وارث الأنبياء، العراق

الخلاصة

يُعد استخلاص النفط المعزز بثاني أكسيد الكربون (CO_2-EOR) من أفضل تقنيات احتجاز الكربون واستخدامه وتخزينه (CCUS) التجارية، حيث يُحسّن استخلاص النفط ويحقق تخزينًا جيولوجيًا آمنًا. ومع ذلك، فإن كثافة البيانات والتكاليف الحسابية تُعيق عمليات محاكاة المكامن التقليدية. يقدم هذا البحث نموذجًا للتعلم الآلي متعدد المخرجات للتنبؤ في آنٍ واحد بكلٍ من زيادة إنتاج النفط وحجم تخزين ثاني أكسيد الكربون. تم تجميع مجموعة بيانات عالمية مختارة تضم ١٧٣ مشروعًا، ومعالجتها مسبقًا، وتحسينها. من بين الخوارزميات التي تم تقييمها (الغابة العشوائية، وتعزيز التدرج، وسلسلة الغابات العشوائية)، حقق تعزيز التدرج أفضل أداء تنبؤي، بمعامل تحديد مرجح (R^2) قدره ٠.٧٢٢. أظهر تحليل النموذج أن النفاذية والعمق وقابلية الامتزاج تؤثر بشكل كبير على استخلاص النفط (R^2 لتعزيز التدرج = ٠.٨٧)، بينما كانت إمكانية التنبؤ بحجم التخزين منخفضة (R^2 لتعزيز التدرج = ٠.٥)، والتي تتأثر بشكل كبير بعمق المكمن ومساميته. تسلط هذه الدراسة الضوء على إمكانات التعلم الآلي كأداة فحص سريعة لتقييم مشاريع احتجاز الكربون واستخدامه وتخزينه، وتحديد أداء العوامل الرئيسية، والتأكيد على الحاجة إلى تقارير تخزين موحدة لتحسين دقة التنبؤ.

الكلمات الدالة: CO_2-EOR، تخزين ثاني أكسيد الكربون، احتجاز الكربون واستخدامه وتخزينه، التخزين تحت السطحي، بيانات الحقول الجيولوجية، التعلم الآلي.