Advanced Energy Systems and Intelligent Geoscience (AESIG)
https://www.mospbs.com/journal/aesig

• *Review* •

A Review of Integrated CCUS-EOR and Storage Prediction and Optimization: From Accelerated Compositional Simulation to Data–Physics Coupled Intelligent Models

Yunfeng Xu[1, 2], Zhuyi Zhu1[1]

[1]School of Petroleum Engineering, Yangtze University, Wuhan, 430100, China

[2]Western Research Institute, Yangtze University, Karamay 834000, China

[3]State Key Laboratory of Low Carbon Catalysis and Carbon Dioxide Utilization (Yangtze University), Wuhan 430100, China.

[*]**Corresponding Author:** Yunfeng Xu **Email:** 201972114@yangtzeu.edu.cn

Received: 14 June 2025 **Accepted:** 17 October 2025

Abstract: Carbon dioxide enhanced oil recovery with geological storage has attracted increasing attention because it can simultaneously improve hydrocarbon recovery and reduce emissions. Accurate and efficient prediction of development performance, together with reliable support for injection–production design and optimization, has therefore become a central scientific and engineering challenge. Numerical simulation for carbon dioxide flooding has evolved from improved black-oil and pseudo-compositional formulations to K-value approaches and, more recently, equation-of-state-based compositional models. Although compositional simulation offers high mechanistic fidelity, it suffers from severe computational burdens in high-resolution three-dimensional models and in iterative workflows for history matching and optimization. To alleviate these constraints, three complementary acceleration routes have been developed, including multiscale methods that embed fine-scale heterogeneity into coarse-scale solutions, streamline-based methods that leverage convection-dominated flow characteristics, and reduced-order models that compress the state space for rapid iterative evaluation. In parallel, data-driven surrogate models have progressed rapidly with the growing availability of production, monitoring, and simulation data. These approaches enable fast forecasting, sensitivity analysis, and multi-objective decision support, yet their reliability remains limited under complex phase behavior and out-of-distribution operating conditions. Recent data–physics coupling paradigms, represented by simplified mechanistic models, network-based flow models, and physics-constrained deep-learning frameworks, provide promising pathways to reconcile physical consistency with computational efficiency. This review synthesizes the evolution, applicability boundaries, and engineering performance of these methods, and highlights future directions toward trustworthy, field-oriented intelligent simulation and closed-loop optimization in highly heterogeneous reservoirs.

Keywords: CCUS-EOR; Numerical simulation; Reduced-order models; Deep learning; Data–physics coupling; Physics-informed neural networks

1 Introduction

In the evolution of enhanced oil recovery (EOR) technologies, CCUS-EOR has attracted extensive attention because it can simultaneously increase oil recovery and contribute to carbon-emission reduction. Accurately predicting the development performance of CCUS-EOR and providing reliable support for the rational design of field injection–production schemes have become central scientific and engineering issues in integrated CO_2-EOR and storage research [1]. Since the 1970s, numerical simulation methods, owing to their strong capability in mechanism-based description and broad adaptability, have gradually become the most important technical approach in both CCUS-EOR research and practice [2]. Focusing on the phase behavior and multicomponent mass-transfer characteristics of CCUS-EOR, three major numerical simulation frameworks have been established: the improved black-oil (pseudo-compositional) model, the K-value model, and the equation-of-state (EOS)-based compositional model. These approaches differ in physical fidelity, computational complexity, and scope of application, and each has played a critical role at different stages of development[3].

In early studies, the pseudo-compositional model introduced the concept of a "solvent" into the traditional black-oil framework, enabling an engineering-level approximation of CCUS-EOR with relatively high computational efficiency [4]. However, its pseudo-component configuration is largely experience-dependent, making it difficult to faithfully represent CO_2 dissolution, extraction, and swelling effects under high-pressure and high-temperature conditions. Its capability to describe the continuous evolution of phase behavior under miscible and immiscible conditions is also limited. The K-value model, which employs empirical or semi-empirical gas–liquid equilibrium constant formulations, can capture

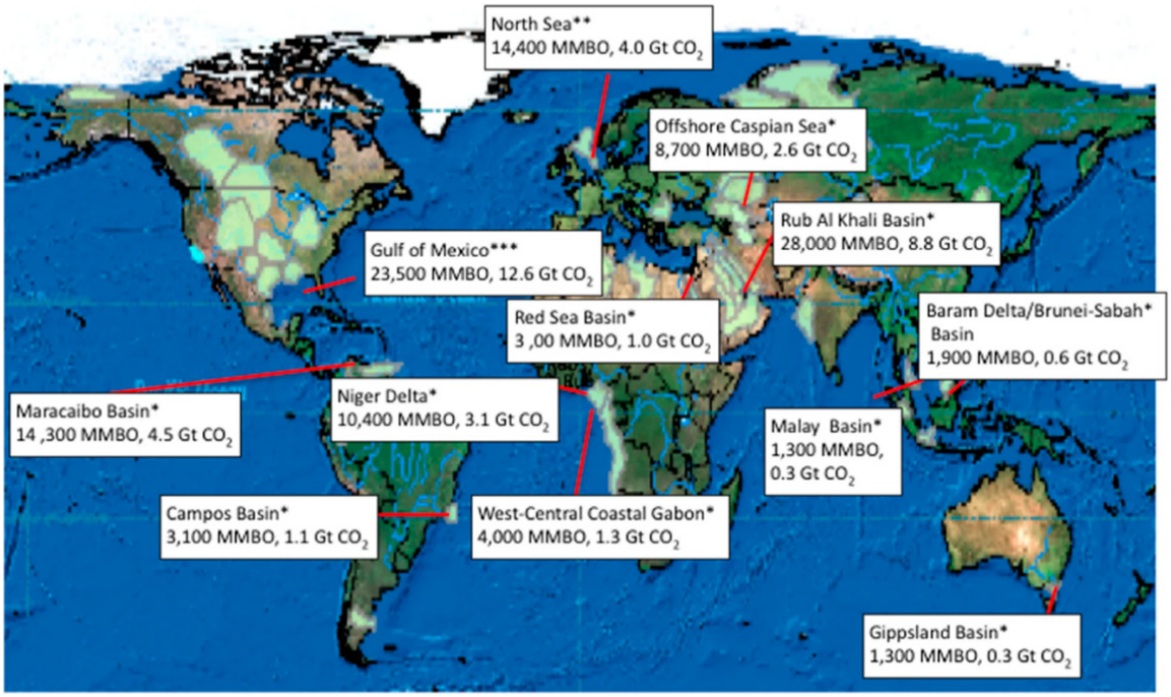


Fig. 1 Basins for which the potential for incremental oil production and CO2 storage have been assessed [5].

limited dissolution and mass-transfer processes under immiscible conditions and thus offers strong engineering adaptability. Yet, once the pressure exceeds the minimum miscibility pressure, the coupled effects of complex multicomponent mass transfer and phase behavior cause the accuracy of empirical K-value approximations to deteriorate significantly. To overcome these limitations, EOS-based compositional models have gradually become the mainstream for CCUS-EOR simulation, as they can more comprehensively characterize the phase behavior of the CO_2–crude oil system and the key associated effects such as viscosity reduction and interfacial tension lowering. Nevertheless, in large-scale three-dimensional high-resolution geological models, EOS iterations and strongly coupled solutions lead to substantial computational overhead. As a result, the runtime of a single simulation increases markedly, while the cost of history matching and optimization in outer-loop workflows can escalate almost exponentially, thereby constraining their application in rapid prediction and engineering decision-making.

To address the practical bottlenecks faced by compositional models, researchers have proposed various acceleration strategies without altering the underlying conservation laws and governing physical equations. These efforts can be broadly categorized into three logically distinct yet complementary technical routes: multiscale methods from the perspective of spatial discretization, streamline-based methods grounded in flow-dynamics characteristics, and reduced-order methods centered on state-space rank reduction. Each route offers unique advantages in preserving dominant heterogeneity effects, improving computational efficiency for convection-dominated systems, and enabling fast iterations for history matching and optimization, thereby providing an important foundation for “physically controllable acceleration” in integrated evaluation of CCUS-EOR and storage.

Meanwhile, with the accumulation of production and monitoring data and the rapid advancement of computational intelligence, purely data-driven machine learning and deep learning surrogate models have increasingly emerged as a major research direction for CCUS-EOR simulation and optimization. By learning the input–output mapping, these methods significantly reduce dependence on geological modeling and numerical discretization and deliver clear advantages in predictive efficiency. However, under complex phase-behavior conditions and when extrapolating to unseen operating scenarios, their reliability may be insufficient due to the lack of physical constraints. To balance “physical-mechanism fidelity” and “data-driven efficiency,” recently developed data–physics coupled paradigms—such as simplified-mechanism-based CRM/INSIM, networked/graph-structured surrogates, and physics-constrained deep models—have provided a more unified and engineering-promising framework and new pathways for efficient prediction, connectivity identification, and closed-loop optimization in integrated CO_2-EOR and storage.

Building on the above background, this paper addresses the key demands of integrated prediction and optimization for CO_2-EOR and storage. We systematically review the evolution and applicability boundaries of three major numerical simulation frameworks, summarize the core concepts and field performance of physics-based acceleration approaches including multiscale, streamline, and reduced-order methods, and examine the latest advances in machine learning/deep learning surrogates for joint CO_2-EOR and storage prediction, history matching, and multi-objective optimization. Particular emphasis is placed on data–physics coupled approaches,

with a focused discussion of their advantages and remaining challenges in enhancing model reliability, generalizability, and interpretability. Through comparative analyses of these technical pathways, this paper aims to provide a reusable technical roadmap and informed perspectives on future research directions for rapid evaluation and decision support of integrated CO_2-EOR and storage in complex heterogeneous reservoirs.

2 Computational Bottlenecks in CCUS-EOR Simulation and Dimension-Reduction Acceleration Pathways

2.1 Computational Bottlenecks in CCUS-EOR Simulation

In the evolution of enhanced oil recovery (EOR) technologies, CCUS-EOR has attracted extensive attention because it can achieve the dual objectives of improving oil recovery and reducing carbon emissions. Accurately predicting the development performance of CCUS-EOR, and thereby supporting the rational design of field injection–production schemes, remains a core research problem in this domain. Since the 1970s, numerical simulation methods—owing to their strong mechanism-based description capability and broad adaptability—have gradually become the primary technical tool in CCUS-EOR research and practice [6]. Over decades of development, three major numerical simulation frameworks have been established for CCUS-EOR: the improved black-oil (pseudo-compositional) model, the K-value model, and the fully compositional model. These approaches differ in physical fidelity, computational complexity, and application scope, and each has played an important role at different stages of technological evolution.

Early numerical simulations were largely built upon the conventional black-oil framework, which cannot directly represent the complex phase behavior between CO_2 and crude oil. In 1972, Todd and Longstaff proposed an improved black-oil model (also referred to as a pseudo-compositional model) [7]. The basic idea is to introduce “solvent” as a new pseudo-component into the traditional three-phase black-oil model, so that oil, water, dissolved gas, and injected gas are treated as four components, with separate mass conservation equations formulated and solved for each. In 2020, Sandve [8] proposed an extended black-oil model in which black-oil properties such as density and viscosity within each grid cell dynamically vary with the CO_2 fraction; black-oil functions are calibrated against experimental data, enabling predictions that are closer to those of fully compositional simulations [9]. This approach retains the advantages of the black-oil model, including high computational efficiency and a relatively simple structure. However, the definition of pseudo-components is still largely dependent on empirical approximations, making it difficult to faithfully capture CO_2-induced dissolution, extraction, and swelling effects under high-pressure and high-temperature conditions. Moreover, the model does not fully describe the continuous evolution of phase behavior across miscible and immiscible regimes, which can lead to substantial deviations when predicting displacement performance in complex reservoir settings.

As research progressed, scholars recognized that relying solely on pseudo-component approximations was no longer sufficient to accurately represent the complex phase behavior during displacement. Consequently, the K-value model was proposed and widely applied [10]. This approach determines gas–liquid equilibrium constants using empirical correlations or semi-empirical formulations, thereby describing the partitioning of multiple components between two phases.

Under immiscible conditions, the K-value model can reasonably capture the limited dissolution and mass-transfer processes between CO_2 and hydrocarbons. As a result, it was extensively used in the 1980s and 1990s to simulate immiscible CCUS-EOR [11][12]. Its key advantages include relatively low computational cost, ease of implementation, and strong engineering adaptability through integration with the traditional black-oil framework [13]. However, the applicability of the K-value model under miscible flooding conditions is clearly limited. When reservoir pressure exceeds the minimum miscibility pressure (MMP), complex multicomponent mass transfer and phase evolution occur between the oil and gas phases, which cannot be adequately reproduced by K-value approximations based purely on empirical expressions, resulting in a substantial loss of simulation accuracy. Therefore, although the K-value approach has played an important role in immiscible-flooding simulations, its accuracy bottleneck has gradually driven researchers toward fully compositional modeling.

To address the limitations of the aforementioned models, fully compositional models have increasingly become the mainstream approach for CCUS-EOR simulation. The EOS-based compositional model proposed by Coats in 1980 [14] established the theoretical foundation for subsequent commercial reservoir simulators such as ECLIPSE Compositional and CMG-GEM. These methods rely on equations of state (EOS), including the Peng–Robinson and Soave–Redlich–Kwong equations, to perform phase-behavior calculations for each component, enabling accurate determination of vapor–liquid equilibrium constants and simulation of the dynamic evolution of multicomponent phase behavior. The primary advantage of fully compositional models lies in their ability to comprehensively capture the complex interactions between CO_2 and crude oil [15], including dissolution, extraction, and swelling mechanisms, as well as the associated effects such as viscosity reduction and interfacial tension lowering [16][17][18]. Nevertheless, these models also present non-negligible limitations. First, their computational cost is high: in large-scale three-dimensional geological models, EOS iterations must be performed for every grid cell, and a single simulation often requires hours or even days. Second, during history matching, hundreds to thousands of simulation runs may be needed, causing the overall computational burden to increase dramatically and severely constraining their use in rapid prediction and optimization workflows [19]. In addition, constructing high-resolution 3D models itself requires substantial manpower and multi-source data, making the overall cost extremely high. More importantly, the large number of model degrees of freedom leads to pronounced non-uniqueness in parameter inversion, further increasing the uncertainty of predictive outcomes.

2.2 Progress in Dimension-Reduction Methods for CCUS-EOR

To address the practical bottlenecks of fully compositional models—particularly the heavy computational burden and the high iterative cost of history matching in high-resolution 3D geological models—researchers worldwide have proposed a series of acceleration strategies with clear engineering value, without altering the fundamental conservation laws and governing physical equations. In general, these strategies follow three logically distinct yet complementary technical routes: (i) multiscale methods from the perspective of spatial discretization; (ii) streamline-based methods leveraging flow-dynamics characteristics; and (iii) reduced-order methods centered on state-

space rank reduction.

2.2.1 Multiscale methods

The core idea of multiscale methods is to construct multiscale basis functions that can "carry" fine-scale heterogeneity information onto coarse-grid scales. In doing so, the global solution of the pressure–velocity field is shifted to the large scale, while the influence of fine-scale geological textures is embedded into the coarse-scale solution through basis-function corrections. This philosophy differs from simple homogenization/equivalent-parameter upscaling and also from blind coarsening. By using numerical operators to incorporate fine-scale information into the global solution in a structured manner, multiscale methods can significantly reduce the degrees of freedom and solution cost while preserving dominant heterogeneity effects. Representative approaches include the multiscale finite element method proposed by Hou and Wu, which constructs basis functions by solving local elliptic subproblems and demonstrates numerical stability and convergence [20]. The multiscale finite volume method developed by Jenny, Lee, and Tchelepi, by contrast, aligns more closely with industrial reservoir simulators in terms of control-volume partitioning and flux-conservation implementation [21].

Compared with conventional coarse-grid upscaling, a key advantage of multiscale methods is that they do not simply average fine-scale parameters. Instead, fine-scale impacts participate in the global solution through "correction–restriction" operators. As a result, these methods can still maintain reliable descriptions of pressure propagation and velocity distribution in settings characterized by strong interlayer contrasts and permeability spanning several orders of magnitude, especially near-well regions. In 2018, Torres et al.[23] constructed a cross-scale model in the Bakken tight oil reservoir to elucidate the interplay between CO_2-enhanced oil recovery and storage mechanisms. In 2024, Li et al. [24] developed a multiscale framework for oil–water and gas-flooding simulations, quantitatively analyzing the impact of multiscale heterogeneity on displacement efficiency. In 2025, Peng et al. [25] employed coupled pore-network and reservoir-scale multiscale simulation to reveal the mechanisms of CO_2 foam flooding in low-permeability reservoirs.

For CCUS-EOR scenarios involving strong phase-behavior coupling, pronounced capillary-pressure effects, or non-Darcy flow,

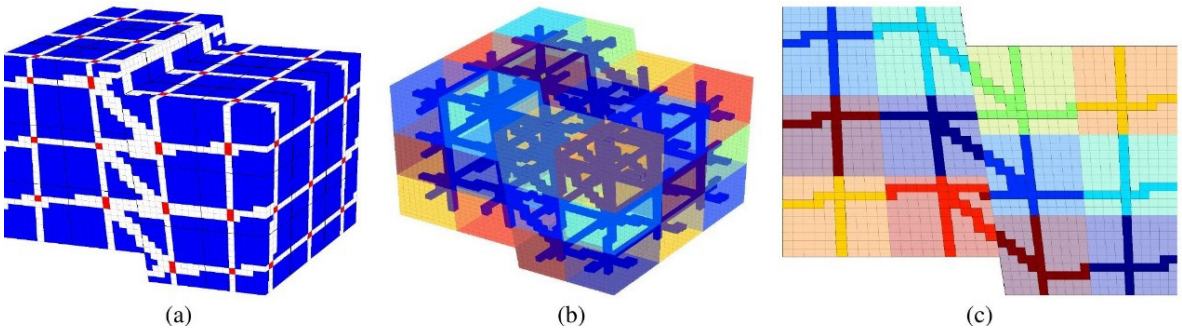


Fig. 2 Primal and dual coarse grids generated for a 40 x 30 x 31 realization of the single-fault model. The plot in (a) shows the 'inner' cells (blue), the 'face' cells (white), the 'edge' cells (red), and the primal coarse grid (thick line) for a subset of the whole model. Plot (b) shows the different blocks in the primal 4 x 3 x 3 coarse grid colored in different transparent colors to show the edge' cells inside, colored in blue. Plot (c) shows a side-view of the same plot, in which the 'edge' cells have been given a different color for each primal block [22].

retaining a traditional "pressure–saturation sequential" coupling framework may still introduce non-negligible errors in approximating phase behavior and mass-transfer terms within multiscale formulations. Therefore, in engineering practice, multiscale solvers are often combined with adaptive mesh refinement, local near-well fine grids, or hybrid grids (e.g., unstructured or nested meshes) [26]. This strategy maintains high fidelity in critical regions while allowing coarsening in less sensitive areas, achieving a practical balance between accuracy and efficiency. From the standpoint of time discretization, multiscale methods are also frequently paired with implicit-pressure/explicit-saturation schemes or fully implicit solvers to stably handle high compressibility and strongly coupled source terms. Across multiple comparative studies, such combined strategies can often deliver one to two orders of magnitude speedup relative to "full fine-grid—fully implicit" baselines, while keeping the influence on full-field pressure distribution and macroscopic sweep within acceptable engineering limits.

2.2.2 Streamline methods

Streamline methods are built on instantaneous velocity fields and transform 3D, convection-dominated multiphase flow problems into 1D transport problems along streamlines in a Lagrangian coordinate system. This mapping decomposes a high-dimensional convection–diffusion problem into a series of 1D transport equations that can be solved in parallel, markedly reducing numerical dissipation and dispersion during time marching. Thiele et al. [27] proposed a streamline-based history-matching strategy that adjusts parameters such as interwell connectivity and relative permeability, enabling simulated production to approach observed data under constraints of "streamline time of flight" and "volumetric allocation," thus achieving multi-well history matching with reduced full-field reassembly cost. Datta-Gupta and King [28] provided a systematic exposition of streamline theory and engineering implementation, highlighting that in waterflooding and gas-flooding problems with clear injection–production relationships and dominant macroscopic convection, streamline simulation offers clear advantages over full-field finite-difference or finite-volume methods in computational efficiency and parameter-sensitivity analysis[29].

In CCUS-EOR applications, the advantages of streamline methods can be particularly pronounced. In 2022, Islam and Woobaidullah performed integrated CO_2 storage and EOR assessments by combining streamline analysis with compositional simulation [30]. In 2025, Zhang et al. [31] proposed a three-phase streamline history-matching method that significantly improved field-data matching efficiency. Streamline approaches typically produce lower numerical dispersion when describing displacement fronts, viscous fingering, and gravity override—phenomena characterized by "strong convection—weak diffusion." This makes them well suited for rapid screening of multiple development scenarios, sensitivity analyses of injection–production strategies, and Monte Carlo evaluations across multiple geological realizations.

However, when phase behavior and source terms (e.g., viscosity and density variations induced by dissolution and extraction) are strongly coupled, the instantaneous velocity field and phase-state parameters require more frequent iterative updates, which can partially offset the computational advantages of streamline methods. To mitigate this issue, engineering practice often introduces relaxed iterations between pressure updates and streamline reconstruction, and increases the reconstruction frequency during critical time

periods to achieve a controlled trade-off between accuracy and efficiency[32].

2.2.3 Reduced-order methods

Reduced-order models (ROMs) focus on state-space compression. Methods such as principal component analysis and singular value decomposition are used to construct low-dimensional subspaces from "training snapshots." The original high-dimensional state is then projected onto this subspace for evolution, and physical fields are recovered through reconstruction mappings [33]. Chen and Durlofsky systematically discussed the construction, stability, and error sources of subsurface-flow ROMs, proposing techniques such as selective snapshot selection, blockwise reduction, and adaptive updating to alleviate "out-of-sample drift," thereby maintaining engineering usability under complex heterogeneity [34]. The linearized ROM proposed by Cardoso and Durlofsky adopts a methodological structure of "linear approximation—projected evolution—error correction," enabling effective dimensionality reduction for coupled pressure–saturation systems, and thus demonstrating clear computational advantages in parameter-sensitivity analyses and repeated simulation scenarios [35].

Compared with multiscale and streamline methods, ROM places greater emphasis on dimensionality reduction at the operator level. Theoretically, this enables ROM to interface with solvers using arbitrary grid structures and boundary conditions. From an engineering standpoint, ROM can be embedded as a "fast prediction kernel" into the outer loops of history matching and optimization, thereby substantially shortening the wall-clock time of the iterative "simulate–evaluate–re-simulate" cycle. It should be noted, however, that the effectiveness of ROM is highly dependent on the representativeness of training samples and the availability of robust online calibration mechanisms. When the displacement process involves phase-regime switching, strong nonlinear transitions, or significant changes in well patterns and control parameters, a static subspace may fail to cover the true evolution trajectory. In such cases, model accuracy must be maintained through incremental snapshots, local-subspace stitching, or error-aware retraining strategies [36].

In the context of CCUS-EOR, ROM is often combined in practice with tabulated treatments of EOS-based compositional calculations and surrogate approximations of key phase-behavior parameters. This hybrid strategy reduces the per-step cost of phase-equilibrium iterations without sacrificing critical phase-transition windows, leading to a notable reduction in overall runtime. In 2022, Zhao et al. [37] explored the feasibility of CCUS-EOR in low-pressure reservoirs by integrating EOS simulations with surrogate approximations. In 2025, Ma et al. compared multiple ROM approaches and systematically evaluated their predictive accuracy in CO_2 enhanced oil recovery (CO_2-EOR) scenarios [38].

3 Progress in Data-Driven Methods for CCUS-EOR

3.1 Progress in Machine-Learning-Based Approaches for CCUS-EOR

With the continuous accumulation of reservoir production-dynamics data and the rapid development of computational intelligence, purely data-driven approaches based on machine learning and deep learning have increasingly become an important research direction for numerical simulation and optimization of CCUS-EOR. Unlike traditional physics-based modeling methods, purely data-driven approaches typically

do not rely on explicit governing equations. Instead, they learn input–output mappings directly from large-scale historical data. Such methods can substantially reduce dependence on geological modeling and numerical discretization, and offer clear advantages in predictive efficiency [39].

As one of the earliest classes of tools introduced into CCUS-EOR prediction and optimization, traditional machine-learning methods have played a significant role. These methods include support vector machines (Support Vector Machine, SVM), random forests (Random Forest, RF), and gradient-boosting models (Extreme Gradient Boosting, XGBoost). Owing to their relatively simple structures, stronger interpretability, and efficient training, they have been explored in various CO_2-EOR scenarios [40]. In 2016, Hosseinzadeh Helaleh and Alizadeh employed SVM combined with three optimization algorithms—ant colony optimization, particle swarm optimization, and genetic algorithms—to predict recovery factor [41].

In recent years, XGBoost has been frequently adopted in CCUS-EOR and CO_2-WAG scenarios, often in combination with surrogate and optimization algorithms to improve prediction and decision-making efficiency. In 2023, Gao et al. proposed an XGBoost-PSO workflow. By generating 10,000 samples with varying geological and operational parameters, they trained the model to predict CO_2-WAG production and optimize injection parameters [43]. Thanh et al. [44] compared multiple machine-learning models in CO_2-foam flooding and demonstrated the superiority of XGBoost in recovery prediction. Nevertheless, XGBoost typically entails relatively high training and tuning costs and exhibits strong sensitivity to hyperparameters.

Shen et al. developed a new CO_2-EOR potential-evaluation method based on a BO-LightGBM framework. Through feature fusion and importance ranking, they established a nonlinear mapping between reservoir parameters and development potential [45]. This method not only improved predictive accuracy but also provided an efficient tool for rapidly screening blocks suitable for CO_2-EOR implementation. Lv et al. [46] modeled the minimum miscibility pressure (Minimum Miscibility Pressure, MMP) of CO_2–crude oil systems using deep learning and tree-based models, and integrated thermodynamic models to achieve accurate MMP prediction. Esfand et al. [47] constructed surrogate models using various machine-learning algorithms to capture the relationship between well-placement design and reservoir heterogeneity in CO_2-EOR. Their results suggest that the proposed models can effectively identify sensitivity features of well-pattern optimization to production dynamics, thereby providing quantitative references for well deployment and development-scheme design.

3.2 Progress in Deep-Learning-Based Approaches for CCUS-EOR

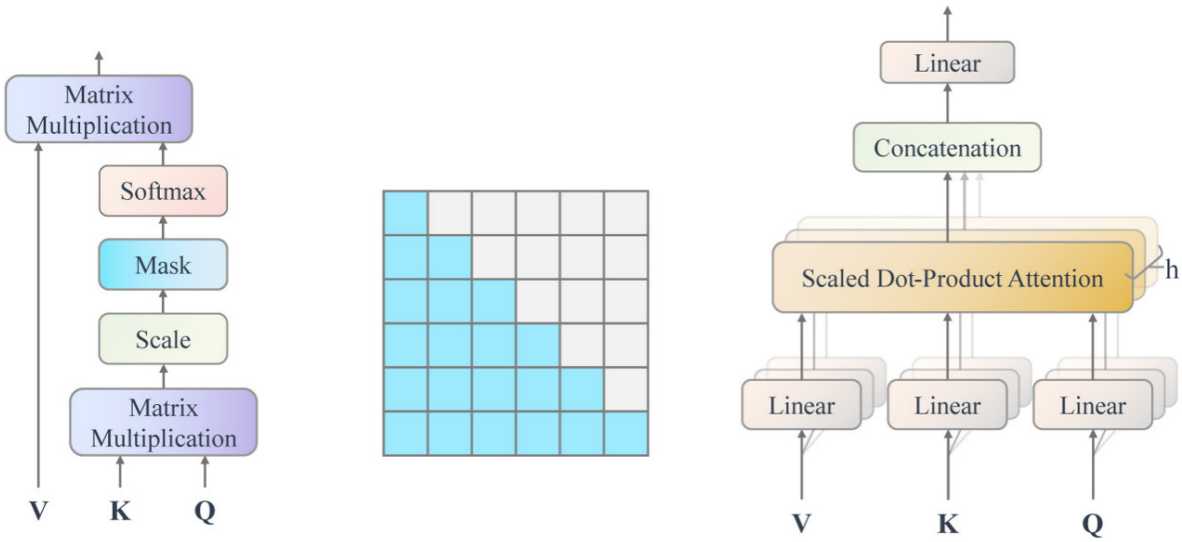
In practical CO_2 enhanced oil recovery (CO_2-EOR) applications, the rapid growth of monitoring data, numerical-simulation outputs, and experimental datasets has increasingly positioned deep learning (Deep Learning, DL) as an important tool in CO_2-EOR research. Compared with traditional machine-learning models, deep learning offers stronger nonlinear representation and automatic feature-extraction capabilities, enabling end-to-end modeling under spatiotemporal coupling, nonlinear boundary conditions, and high-dimensional complex inputs. As a result, DL has demonstrated distinctive advantages in rapid prediction, history matching, optimal control, and uncertainty analysis for CCUS-EOR problems [48][49][50][51]. At present, deep-learning-based surrogate models

exhibit diverse application scopes and strengths in CCUS-EOR, and a multi-path development landscape is gradually taking shape.

3.2.1 Convolutional neural networks

Convolutional neural networks (Convolutional Neural Networks, CNN) are well suited for processing images and structured grid data owing to their local receptive fields and weight-sharing mechanisms. They have therefore been introduced into reservoir modeling to extract spatial features associated with reservoir heterogeneity. In CCUS-EOR applications, CNNs are commonly used to predict oil and gas production under different injection–production settings, as well as the distribution of CO_2 plumes [52]. Kim et al. [53] constructed surrogate models using CNNs with average pooling and fully connected (Fully Connected, FC) layers, and demonstrated the effectiveness of deep CNNs in selecting well locations and well types, thereby broadening the application scope of neural networks in petroleum development. Meng et al. [54] developed a deep-learning surrogate based on a Res-U-Net architecture, organizing multi-source inputs—including porosity, permeability, well locations, and control conditions—into "image-like" tensors. Through multi-scale feature extraction and skip-connection fusion within an encoder–decoder structure, the model achieved end-to-end rapid prediction of state variables such as pressure and saturation fields. These results indicate that CNNs can substantially reduce dependence on high-fidelity full-physics simulations, making them suitable for rapid scenario screening and sensitivity analysis.

In 2024, Yan et al. [55] proposed a CNN-based surrogate for forecasting field production dynamics and performing history matching. By using Latin hypercube sampling and the MATLAB Reservoir Simulation Toolbox to generate diverse discrete-fracture model samples, they employed six input channels to comprehensively characterize



(a) (b) (c)

Fig. 3 (a) CNN-Transformer architecture. The static geological and time-varying engineering parameters are encoded to the latent space by the CNN encoder and MLP, respectively. Transformers are then applied to process the latent sequence, followed by the CNN decoder to generate predictions at all time steps simultaneously. Skip connections are used to enhance information flow between the encoder and decoder. The numbers at the corner of each convolutional block denote the channel. (b) Details of the convolutional blocks in the CNN encoder. (c) Details of the ResNet blocks in the CNN decoder[52].

reservoir and fracture attributes, and used oil and water production as the training and validation targets. The results showed that this CNN surrogate achieved up to 99% accuracy in predicting oil and water production. Moreover, when combined with a genetic-algorithm-based intelligent optimization strategy, the CNN surrogate exhibited higher efficiency and stability in automated history matching compared with conventional approaches.

Mo et al. [56] employed dense blocks to construct a model for simulating the evolution of subsurface CO_2 storage. Their approach encodes input permeability to generate a series of feature maps, concatenates time with these maps, and then decodes them to produce pressure and saturation distributions. Although a segmentation loss was introduced to address discontinuities at saturation fronts, the model is primarily effective for interpolation within the training-time window, while its extrapolation performance beyond that window remains limited. Han et al. [57] proposed a recurrent R-U-Net (Recurrent R-U-Net), which integrates temporal recurrent units into the U-Net backbone for spatial feature extraction, thereby capturing time dependence during injection–production evolution and enabling history matching within a hierarchical Markov chain Monte Carlo framework. Building on CNN and temporal-network architectures, Xu et al. introduced a multi-head attention module to more accurately represent complex interwell dynamics and long-term dependency features during CO_2-WAG development [58].

3.2.2 Recurrent neural networks

Recurrent neural networks (RNN) and their variants—such as long short-term memory (Long Short-Term Memory, LSTM)—have been widely used in recent years for production-dynamics prediction and history matching in CO_2-EOR, owing to their strengths in time-series modeling[59]. Through recursive state propagation, RNN architectures can capture temporal fluctuations in injection–production strategies and operating conditions, and are sensitive to long-term dependencies embedded in reservoir pressure, injection rates, and production profiles. Utomo Pratama Iskandar and Kurihara [60] utilized LSTM to predict time-varying oil, water, and gas production for existing wells, demonstrating improved generalizability and predictive stability over traditional machine-learning models across multiple well patterns. Ruijie Huang et al. [61] used up to 15 years of production data to develop an LSTM model for predicting oil production, gas–oil ratio, and water cut, and compared the results with numerical simulations, highlighting the efficiency of LSTM in handling real operational changes and interwell connectivity characteristics.

Davoodi et al. [64] applied LSTM models to a five-spot CO_2-EOR system to predict CO_2 storage capacity and oil recovery. Their results indicated that under complex spatial distributions and time-varying conditions, LSTM achieved markedly higher prediction accuracy than baseline models, while substantially reducing the time required for training and inference. Feng et al. [59] proposed an encoder–decoder-based ConvLSTM surrogate for rapidly predicting the evolution of pressure and saturation under dynamic CO_2 injection scenarios. Despite these successes, RNN/LSTM approaches still face limitations, such as vanishing or exploding gradients in ultra-long sequences. When injection–production strategies change frequently or data sampling intervals are inconsistent, additional architectural enhancements—such as attention mechanisms [66], segmented training strategies [60], or sliding-window formulations [62]—are often required to ensure training stability and predictive accuracy.

3.2.3 Graph neural networks

In recent years, graph neural networks (GNNs) have developed rapidly in the oil and gas engineering domain [63]. Their core advantage lies in their natural compatibility with unstructured grids, complex well-network connectivity, and irregular geological configurations. By directly leveraging graph-structured data to model reservoirs, GNNs can overcome the limitations of conventional CNNs that rely heavily on regular grids [70]. In CCUS-EOR, strong reservoir heterogeneity and complex inter-well connectivity often make it difficult for traditional methods to simultaneously achieve high efficiency and accuracy. The introduction of GNNs therefore provides a new and promising pathway to address this challenge.

Huang et al. [68] proposed an improved GNN that employs attention mechanisms to invert the dynamic layer-by-layer connectivity between injection and production wells. Du et al. [67] combined a graph convolutional network (Graph Convolutional Network, GCN) with gated recurrent units to compute daily oil production. In this framework, the GCN is used to quantify inter-well connectivity, while the gated recurrent unit extracts well features and outputs production information. As a result, when computing the response of a target well, wells that are not actually connected to the target may still be involved in the calculation. In 2025, Li et al. [69] developed a heterogeneous spatiotemporal fusion model integrating relational graph convolution, Transformer, and LSTM for dynamic prediction of water cut between injectors and producers in waterflooding. They further introduced a hybrid optimization framework that combines genetic algorithms with multi-objective particle swarm optimization, enabling closed-loop prediction–optimization control. Numerical experiments demonstrated that this approach can effectively improve net present value while maintaining predictive accuracy.

Jiang proposed a GCN-based multiphase-flow simulation framework capable of efficiently solving pressure and saturation distributions on unstructured grids. The results indicate that the GCN model can substantially reduce computational cost while maintaining accuracy comparable to that of finite-volume methods, validating the feasibility and reliability of GNNs in reservoir numerical simulation [70]. Gudalad and Yan et al. [71] introduced a novel GNN-based deep-learning framework to accelerate simulations of fractured geothermal reservoirs. By encoding node and edge features, the GNN systematically represents the influence of fracture networks on fluid flow, heat transport, and geomechanical behavior, and predicts the dynamic response of complex geothermal systems under injection–

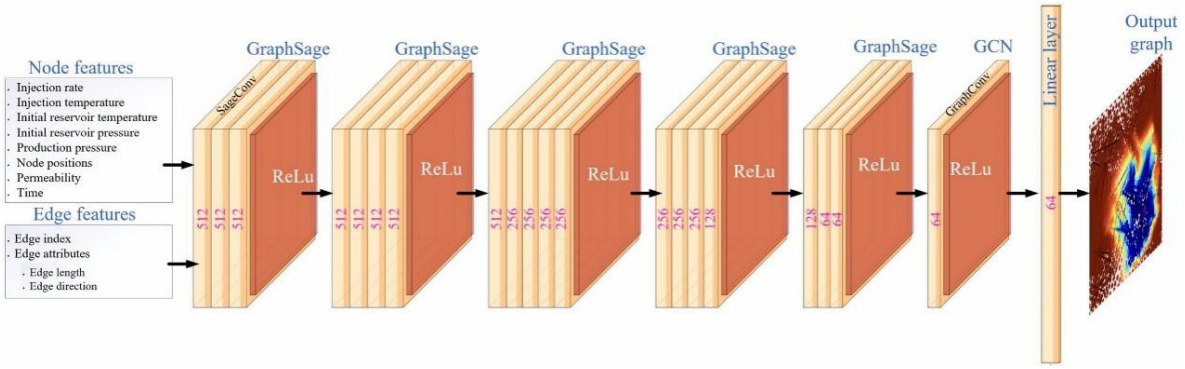


Fig. 4 Schematic of sequential sage (SeqSage) GNN model [71].

production operations. Although this work focuses on geothermal systems, its findings offer valuable insights for understanding multi-physics processes in fractured reservoirs subjected to CCUS-EOR.

In 2024, Ju et al. [72] combined graph convolution with LSTM to develop a graph convolutional long short-term memory network that effectively captures the spatiotemporal migration of CO_2 plumes in faulted reservoirs. The model maintained high accuracy under various fault geometries while substantially reducing computational cost, with single-run inference time reportedly around one percent of that required by conventional simulations. In 2025, Zhuang et al. [73] constructed a GNN–Transformer surrogate to predict CCUS-EOR dynamics and storage capacity, and integrated a multi-objective particle swarm optimization algorithm to adjust well placement and controls for maximizing both oil production and storage. Tariq et al. [74] proposed an enhanced GCN based on a U-Net backbone for rapid prediction of saturation and pressure evolution of CO_2 plumes in saline aquifers. Compared with traditional numerical simulation, this approach significantly reduced computational cost while achieving near-physics-level accuracy across multiple test scenarios.

Overall, GNNs have demonstrated clear advantages in reservoir simulation and CCUS-EOR studies. They can represent complex unstructured grids and well-network topologies, achieving high predictive accuracy with markedly reduced computational overhead. Current research has primarily focused on plume-migration prediction, connectivity analysis in faulted reservoirs, and fast surrogate computation of pressure–saturation fields. Nevertheless, it should be noted that much of the existing work remains at the numerical-experiment stage. The adaptability to field-scale dynamic data, model generalization capability, and deeper integration with physics-constrained approaches still require further investigation.

3.2.4 Fourier Neural Operator

The Fourier Neural Operator (FNO) is an efficient framework for solving partial differential equations (PDEs) popularized by Caltech [75]. The method first transforms input features into the frequency domain via a Fourier transform, evolves the representations in the spectral space, and then applies an inverse Fourier transform to obtain updated features. In 2022, Zhang et al. [76] applied FNO to solve two-dimensional oil–water two-phase flow PDEs. By introducing the fast Fourier transform to extract PDE information and incorporating physical constraints, their method efficiently predicted the evolution of saturation and pressure. It demonstrated excellent accuracy and generalization in both forward and inverse problems, highlighting its potential as an alternative to traditional numerical simulation.

Also in 2022, Chu et al. [77] addressed the high computational cost of CO_2 storage simulation in shale–sandstone composite reservoirs by proposing a deep-learning-based RU-FNO architecture to predict CO_2 plume migration and pressure buildup under complex reservoir conditions. The method reportedly achieved more than an 8000× speedup over conventional numerical simulation for saturation and pressure prediction, with improved accuracy, making it particularly suitable for investigating CO_2 migration governed by thin shale layers. In 2023, Liu et al. [78] developed a 4D (x, y, z, t) flow simulation framework that combines FNO with domain decomposition, successfully extending FNO from conventional 3D spatiotemporal settings to a four-dimensional space–time domain. This framework trains multiple 3D FNO networks in parallel and couples them in the z

direction, enabling efficient simulation of complex flows in fractured reservoirs within CCUS-EOR compositional models.

In 2024, Yang et al. [79] proposed an FNO-based surrogate model for CCUS-EOR in three-dimensional heterogeneous reservoirs to enable rapid prediction of reservoir-property distributions. The study showed that the model can accurately capture key displacement features in CCUS-EOR, including front propagation, gravitational effects, and crossflow. Compared with conventional compositional simulation, the approach achieved an approximately 360× improvement in computational efficiency, demonstrating strong potential for rapid well-placement optimization and parameter inversion in CCUS-EOR projects. In 2025, Liu et al. [80] proposed a DL-NRS framework that integrates FNO with low-resolution numerical simulation results to reconstruct high-resolution pressure and saturation evolution. By introducing physics-constrained loss functions and leveraging low-resolution simulation data for up sampling and training, the framework produced results close to high-resolution numerical solutions. The reported average relative errors for both pressure and saturation were below 1%, with a substantial reduction in computational cost, indicating strong applicability to rapid dynamic prediction for large-scale reservoirs.

3.2.5 Large Language Models

Large language models (LLMs) typically refer to models built upon Transformer backbones. Owing to their advantages in long-sequence modeling and adaptive feature learning, Transformer architectures have increasingly entered the field of reservoir engineering modeling and prediction. Compared with RNN/LSTM, Transformers employ multi-head self-attention (Multi-head self-attention, MSA) to effectively model long-range dependencies, mitigate vanishing-gradient issues, and deliver improved performance in multi-scale dynamic forecasting. In CCUS-EOR research, the introduction of Transformers has provided new methodological support for history matching, production forecasting, and identification of gas-channeling pathways.

In 2023, Zhang et al. [81] proposed a Transformer-based surrogate to address the limited parallelism and high training overhead of traditional RNN approaches for reservoir production prediction. Through multi-head self-attention, the model effectively captured complex dependencies between input features—such as water-injection rates, drilling decisions, well locations, porosity, and permeability—and production responses[82]. A sequence-to-sequence structure further enabled long-horizon extrapolation. The results showed that incorporating additional information such as drilling and well locations significantly improved prediction accuracy, yielding more reliable performance than models that rely solely on

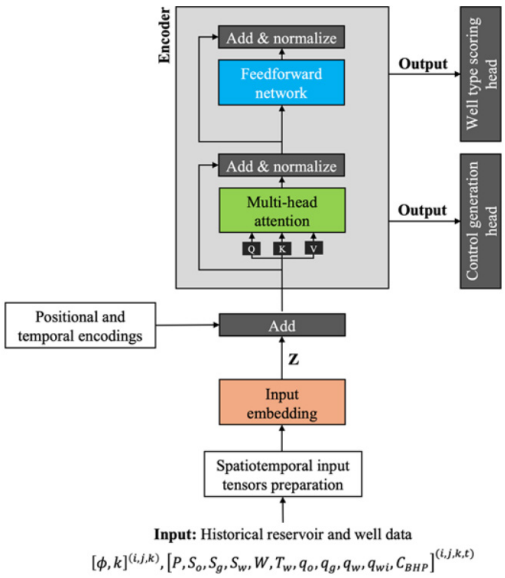


Fig. 5 Architecture of the proposed FDO-Transformer showing inputs, outputs, and encoder internals. The same forward pass is used across train and test splits.

injection-rate inputs. Meanwhile, compared with RNN baselines, the Transformer achieved roughly a fourfold speedup while maintaining similar predictive accuracy.

In 2024, Jia et al. [84] explored a new application of Transformers for production forecasting by proposing a Transformer-based deep neural network to predict productivity after refracturing operations. The method effectively captured nonlinearities and long-range dependencies in production curves, substantially improving prediction accuracy and offering new insights for performance evaluation in unconventional reservoirs. In 2025, Liu et al. [85] further proposed an automated history-matching framework that integrates GNNs, Transformers, and optimization algorithms. By enhancing inter-well connectivity representations with attention mechanisms and combining intelligent optimization to invert injection–production relationships, the framework achieved improved matching accuracy and stability in complex injection–production systems, providing technical support for closed-loop control.

Also in 2025, Feng et al. [52] developed a CNN–Transformer hybrid surrogate that combines the local feature-extraction capability of convolutional networks with the global temporal modeling strength of Transformers for multi-objective robust optimization in geological carbon storage. The results indicated stronger robustness and generalization under complex reservoir conditions. Xiao et al. [82] proposed a unified deep-learning framework that fuses static geological parameters with dynamic production data and enhances temporal-dependency modeling via Transformers, enabling joint prediction of CO_2-EOR performance and carbon storage capacity. The results showed clear accuracy improvements over conventional data-driven models, while providing high-confidence assessments of both static and dynamic performance. Li et al. [86] offered a representative example of a graph-attention plus self-attention spatiotemporal fusion model. In fractured–vuggy carbonate reservoirs characterized by complex geological connectivity, nonstationary data, and strong well-pattern interactions, the method demonstrated superior predictive performance compared with traditional time-series models.

4 Progress in Data–Physics Coupled Methods for CCUS

In mechanistic studies and engineering practice of reservoir development, improving computational efficiency while preserving physical consistency has long been a central challenge for CCUS-EOR simulation and history matching. Conventional full-physics compositional simulators can comprehensively capture complex mechanisms such as pressure propagation, phase-behavior evolution, and the formation of gas-channeling pathways. However, their computational cost is prohibitive. This limitation is particularly pronounced in high-resolution three-dimensional geological models, where such simulators often cannot support the hundreds or even thousands of runs required for history matching and optimization. In contrast, purely data-driven methods offer clear advantages in computational speed, yet their lack of essential physical constraints commonly leads to biased predictions when extrapolating to unseen operating conditions or complex phase-regime scenarios, falling short of the reliability demands of field applications.

Against this backdrop, a series of data–physics coupled approaches has been proposed in recent years to reconcile “physical-mechanism fidelity” with “data-driven efficiency.” The central

idea is twofold: on the one hand, to introduce physical priors or constraints to varying degrees so as to reduce the effective freedom of purely data-based models under complex reservoir conditions, thereby enhancing predictive accuracy and extrapolation capability; on the other hand, to leverage data-driven techniques to accelerate or compensate for traditional numerical models, alleviating their high computational burden.

Broadly, data–physics coupled methods can be categorized into the following types: (1) Analytical or semi-analytical models based on simplified governing equations and data fitting, with representative examples including the Capacitance Resistance Model (CRM) and the Inter-well Simulation Model (INSIM). By simplifying physical control equations and performing parameter inversion using production data, these models have been widely used for rapid history matching and quantitative characterization of inter-well connectivity. (2) Physics-constrained deep-learning methods, typified by Physics-Informed Neural Networks (PINNs), which directly embed PDE residuals into loss functions to enable end-to-end prediction of pressure fields, saturation evolution, and component transport.

4.1 Simplified Physics-Based Approaches

In the evolution of data–physics coupled methods, one of the earliest lines of research focused on introducing production-dynamics data into frameworks with simplified governing equations for parameter inversion. This idea was first proposed by Albertoni et al. [87] and Yousef et al. [88]. The Capacitance Resistance Model (CRM) conceptualizes a reservoir as a flow-network system composed of connectivity units between injectors and producers. Along each connectivity pathway, CRM solves simplified flow equations associated with pressure and saturation, while wellhead allocation is determined using empirical relations. With a relatively low computational burden, CRM can rapidly reflect inter-well connectivity and has therefore been widely applied to dynamic matching and connectivity analysis in multi-well systems.

However, the physical simplifications adopted by CRM also impose clear limitations. First, the productivity index and allocation (splitting) coefficients are typically assumed to be constant throughout the simulation, making it difficult to represent dynamic adjustments caused by changes in field injection–production strategies or wellbore operating conditions. Second, the fractional-flow formulation used to compute phase rates is derived mainly from empirical correlations rather than rigorous physics-based descriptions. Third, constrained by its analytical formulation, CRM cannot directly impose bottom-hole pressure (BHP) control in prediction and optimization workflows [89].

In subsequent developments, researchers have attempted to extend CRM to CO_2-EOR and carbon-storage scenarios. In 2008, Sayarpour integrated CRM with an allocation model to establish a history-matching and performance-prediction tool applicable to both waterflooding and CCUS-EOR, and validated the approach using multiple field cases[90]. The results indicated that, as a lightweight surrogate model, CRM can substantially reduce computational cost in history matching and optimization, while also supporting reservoir uncertainty quantification and performance evaluation. In 2012, Nguyen et al. [91] proposed the Integrated Capacitance Resistance Model and validated it in an Omani oilfield as well as in multiple synthetic CCUS-EOR cases. In 2016, Eshraghi et al. [92] introduced the Gentil fractional-flow formulation into the CRM framework and, combined with heuristic intelligent optimization algorithms, developed an optimization approach

suitable for miscible CCUS-EOR.

In 2014, Zhao Hui et al. [93] proposed a simplified representation of injection–production systems and developed an inter-well connectivity model based on dynamic injection–production data. Through dynamic matching, this method enables rapid inversion of inter-well formation parameters while capturing connectivity relationships. Building on this work, Zhao et al. (2015) [94] proposed the INSIM framework that couples physics-based modeling with data-driven strategies, enabling efficient prediction of two-phase oil–water flow [95]. In 2016, Zhao et al. [96] incorporated aquifer effects into the model and explored its potential for production optimization. In the same year, they proposed a layered connectivity analysis approach and further developed an inter-well connectivity inversion model for multilayer waterflooding reservoirs [97].

Regarding broader applications, Guo et al. (2018) [98] developed the INSIM-FT model, in which a front-tracking approach was used to improve saturation computations based on the Buckley–Leverett equation, thereby enhancing water-cut prediction accuracy and enabling production-optimization studies [99]. In 2019, Zhao et al. proposed a rapid evaluation method for injector efficiency and fast optimization of injection–production strategies within the INSIM framework [100]. In the same year, Guo incorporated gravity effects and developed a corresponding three-dimensional extension [101]. Further advances include the INSIM-FPT model proposed by Zhao et al. (2020) [102], which integrates node densification and path-tracking algorithms to compute key parameters such as inter-well flow paths and allocation coefficients. In 2021, Liu et al. [103] proposed a rapid injection–production rate-optimization method for oil and water wells based on oil-saturation evaluation. Zhao et al. further expanded model capabilities by developing the INSIM-FPT-3D model with gravity effects and introducing a new well-index formulation to improve BHP calculation accuracy [104]. Meanwhile, Liu et al. (2021) also employed these models for dynamic prediction of profile-control and plugging treatments, demonstrating that the approach can effectively forecast post-treatment reservoir behavior and guide optimization design [105].

Despite the demonstrated efficiency and practical value of INSIM and its variants in rapid dynamic prediction, history matching, and production optimization, several limitations remain. First, the basic INSIM framework was originally designed for two-phase oil–water displacement. Although later extensions have been made to polymer flooding and three-dimensional settings, the model still struggles to fully capture key mechanisms in CCUS-EOR scenarios with pronounced phase-behavior complexity, such as dissolution, extraction, swelling, and gas channeling. Second, while improvements such as front tracking and path tracing have enhanced performance, the underlying physical approximation remains largely rooted in Buckley–Leverett theory. This can lead to non-negligible deviations in representing capillarity, gravity, and strongly nonlinear displacement conditions.

The limitations of CRM and INSIM have motivated the development of numerical flow-network models [106]. In such models [107], the one-dimensional inter-well connectivity pathways are discretized, and the governing equations are solved numerically. Compared with CRM or INSIM, numerical flow-network models offer greater flexibility. For example, owing to the discretized formulation, each well can be assigned an independent well grid, which allows the direct use of conventional well models for bottom-hole pressure (BHP) calculations. This feature enables

the model to incorporate BHP data during history matching and to impose BHP control during forecasting. In addition, the mathematical framework of this approach allows the introduction of arbitrary governing equations, making it possible to extend flow-network modeling to more complex physical processes, such as thermal recovery or compositional flow. This flexibility, however, comes at the cost of increased computation. In most cases, nevertheless, the number of grids required by flow-network models remains far smaller than that of full 3D geological models.

In 2018, Lutidze proposed the StellNet model [107], whose key distinction from traditional flow-network models lies in the introduction of the concept of "well cells." Well cells refer to reservoir grid blocks perforated by a wellbore. Different connectivity pathways can interact through these well cells without requiring actual fluid entry into or exit from the wellbore. This enhancement significantly increases the number of potential flow paths and improves the model's capability to represent complex connectivity. In this sense, the well-cell concept plays a role similar to that of "virtual wells".

In 2019, Ren et al. [108] developed GPSNet at Chevron as a network model built on a general-purpose numerical simulator. GPSNet inherits the essential ideas of StellNet, while using a commercial simulator to evolve the network model. Based on this platform, layered workflows were proposed and successfully applied to reservoir history matching and waterflooding optimization [109][110]. GPSNet was subsequently extended to GPSNet-2D, where two-dimensional connectivity pathways were used for history matching and optimization in steamflood reservoirs [111]. Wang et al. [112] further expanded GPSNet to address rapid decline analysis and interwell interference in unconventional reservoirs, and introduced a dual-grid system for coupled solutions.

Along a similar line, Kiærr et al. [113] proposed and open-sourced FlowNet for production prediction and optimization. Rao et al. [114] further explored how to integrate physical constraints with data-driven methods within a more generalized framework. A representative outcome is the recently proposed general physics-driven data-driven paradigm, which takes governing equations as the core and combines data assimilation with fast surrogate modeling to achieve unified modeling and solution for reservoir simulation and history matching. Like FlowNet, this framework emphasizes interwell connectivity and network-based representation, but goes further by coupling physical equations with data-driven algorithms. It thus reduces computational cost while preserving physical consistency, and offers strong generalizability and flexibility.

In 2022, Zhao et al. [115] proposed a FlowNet-based method tailored to the development characteristics of hydraulically fractured wells in shale gas and tight-oil reservoirs. Without relying on complex geological modeling, the approach constructs an inter-well connectivity network and incorporates fracture-treatment parameters to enable rapid history matching and production forecasting for fractured wells. This method effectively reduces the modeling cost of tight-reservoir simulation and achieved high predictive accuracy, demonstrating the applicability of FlowNet to unconventional-reservoir development. In 2024, Xu et al. [116] applied FlowNet to CO_2 water-alternating-gas (WAG) scenarios and established a rapid history-matching and production-optimization framework. By using network-based representations to simplify injector–producer connectivity calculations and coupling optimization algorithms to identify rational injection–production strategies, the method was able to reproduce production his-

tory in typical CO_2-WAG cases while significantly reducing computational time and supporting operational optimization.

Despite the promising potential of GPSNet and FlowNet for rapid history matching, production forecasting, and optimization control, several limitations remain. Under complex network structures or long-term iterative workflows, convergence issues may arise, manifested as reduced numerical stability or slower convergence during parameter inversion. These issues can be particularly pronounced for large-scale well networks and long-horizon forecasts.

In 2023, Zhao Hui et al. [117] further developed the theoretical framework of the Connecting Element Method (CEM) based on INSIM, positioning it not only as a computational approach but also as a physics–data-driven modeling tool. Typical test cases verified its feasibility and extensibility for multiscale two-phase flow simulations. Liu et al. [118] combined CEM with data-space inversion and proposed an efficient modeling scheme oriented toward closed-loop optimization, offering a new pathway for history matching and production optimization in complex reservoirs. Zhao et al. [119] applied CEM to two-phase flow simulation in multiscale fractured reservoirs, significantly reducing dependence on fine-grid discretization and enabling efficient characterization of fracture–matrix coupled flow. Subsequent work by Xu Yunfeng et al. [120] extended this line of research to CCUS-EOR, proposing CEM-based CO_2 simulation and channeling-pathway prediction methods. This provides a new physics–data-driven tool for gas-channeling identification and control.

4.2 Physics-Informed-Neural-Network-Based Surrogate Models

As a representative approach that couples deep learning with physical constraints, physics-informed neural networks (Physics-Informed Neural Networks, PINNs) have been increasingly introduced into modeling for oil and gas reservoir development in recent years. The core idea of PINNs is to embed the governing partial differential equations (PDEs) of reservoir flow—such as mass conservation, Darcy's law, and phase-equilibrium relations—directly into the loss function. In this way, neural networks are trained not only to fit observational data but also to satisfy fundamental physical laws [121]. Compared with purely data-driven models, PINNs generally exhibit stronger physical consistency when training samples are limited or when predictions must be extrapolated to unseen operating conditions. This makes them particularly suitable for complex multiphase flow problems such as CCUS-EOR [122].

In 2023, Han et al. [123] proposed a domain-decomposition PINN framework (PINN-DD). By partitioning the reservoir into two subdomains—regions with wells and regions without wells—and imposing governing-equation and boundary-condition constraints separately, the method improved modeling accuracy under sparse production-data conditions. The results indicated that PINN-DD can maintain reliable predictive performance with limited monitoring data and shows stronger stability and generalization than conventional PINNs. In 2024, Han et al. [124] introduced a Criss-Cross Physics-Informed Convolutional Neural Network (CC-PINN) to address a key shortcoming of traditional PINNs in strongly heterogeneous reservoirs, namely the difficulty in ensuring inter-cell flux continuity. By combining automatic differentiation with 2D CNNs and introducing criss-cross physical constraints within the convolutional structure, the network learns parameterized PDE solutions while enforcing flux continuity between neighboring grid cells. Benchmark tests on multiple strongly heterogeneous reservoir problems

showed that CC-PINN substantially outperforms standard PINNs, improving cross-boundary flux accuracy and demonstrating greater robustness under complex geological conditions.

Also in 2024, Liu et al. [125] proposed a hard–soft constrained PINN framework (HS-PINN). This method introduced a Lorentz-based soft constraint and hard boundary-condition enforcement for predicting near-well pressure fluctuations, thereby maintaining high accuracy and stability under complex geology. Their results reported prediction errors below 1% in both single-well and multi-well cases, while the computational cost was only about 8% of that of conventional simulations, indicating a promising tool for rapid uncertainty quantification[126]. In 2025, Liu et al. [127] applied PINNs to the convection equations of polymer-flooding reservoirs and developed a high-accuracy model for predicting water saturation and polymer concentration distributions. They observed that using a single network to handle the two variables outperformed a dual-network structure in both accuracy and convergence speed. Numerical stability was further enhanced by introducing an artificial viscosity coefficient.

Nevertheless, a single PINN framework still faces challenges when dealing with high-dimensional complex equations, irregular grids, and strongly nonlinear physical processes. These challenges include convergence difficulties, high training cost, and sensitivity to boundary conditions. To further enhance the adaptability and generalization of deep-learning models in reservoir-development studies, researchers have begun exploring more diverse physics–data fusion strategies. A shared characteristic of these methods is the explicit or implicit incorporation of physical knowledge in network architectures, loss design, or training workflows. This helps overcome the “fast but unstable” limitation of purely data-driven models, while also alleviating optimization bottlenecks associated with the single-route PINN paradigm.

In 2020, Wang et al. [128] proposed theory-guided neural networks (Theory-guided Neural Networks, TgNN) for groundwater-flow and solute-transport modeling, demonstrating that the method can retain high accuracy even under limited-data conditions. In 2021, Xu et al. [129] further developed a weak-form theory-guided network, replacing strong-form constraints with weak-form formulations and thereby significantly improving numerical stability under complex boundary conditions. In the same year, Wang et al. [130] proposed a physics-constrained training strategy that combines the finite difference method (Finite Difference Method, FDM) with CNNs. In this approach, governing equations are first discretized using FDM. The residuals of the discretized PDEs are then computed based on CNN predictions and minimized during training to introduce physical constraints. Unlike Sobel filtering, FDM can simultaneously compute spatial and temporal gradients.

In 2025, Chen et al. [131] proposed physics-informed graph neural networks (PIGNN), integrating physical constraints with graph structures to address the insufficient flux accuracy of traditional PINNs in irregular grids and strongly heterogeneous reservoirs. PIGNN substantially reduced errors and improved computational efficiency in pressure prediction for spatially heterogeneous reservoirs, with an average R^2 of 0.998, demonstrating superior performance over standard PINNs. Zhang et al. [132] introduced a physics-based deep CNN framework for simulating and predicting reservoir pressure fields. In this work, a finite-volume formulation was used to compute governing-equation residuals, thereby incorporating physics into the learning process. To handle transient problems, a

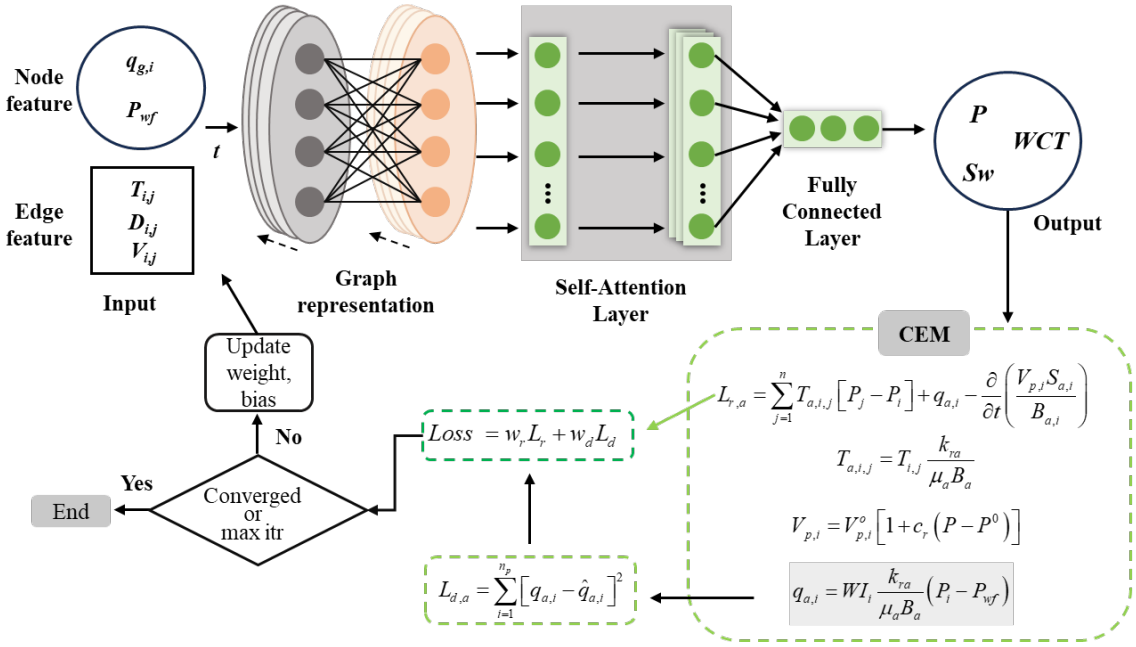


Fig. 6 The overall workflow of the PINN-GCEM model[135]

sequence of CNNs was trained, with each network responsible for predicting pressure at a single time step. The output of one CNN served as the input to the next, enabling stepwise temporal forecasting.

Wang et al. [133] developed a Hybrid Physics-Informed Data-Driven Neural Network (HPDNN) for simulating CO_2 storage dynamics in depleted shale reservoirs. Built on a fully connected neural network (FCNN), HPDNN captures relationships between static and spatiotemporal reservoir attributes (e.g., thickness, porosity, wellbore storage, injection rate) and injection responses (e.g., CO_2 injection pressure differential and stored CO_2 volume). The model adopts a hybrid training strategy that integrates data-driven learning with physics-based constraints. Training datasets are generated by numerical simulators, while multiscale CO_2 transport mechanisms—including diffusion, adsorption, dissolution, slip flow, and Darcy flow—are incorporated as physical information to constrain network training. After training, HPDNN can accurately predict CO_2 injection dynamics and storage capacity. It also serves as a valuable tool for estimating reservoir parameters from dynamic injection profiles, which are critical for inverse modeling in geological storage. By leveraging HPDNN, engineers can bypass conventional, labor-intensive, and computationally expensive numerical simulations required for complex CO_2 storage systems, thereby significantly improving inversion efficiency.

In 2024, Nagao and Datta-Gupta [134] proposed a physics-constrained hybrid neural framework that combines simplified physical models with PINNs for production prediction and interwell-connectivity identification in CO_2-EOR scenarios. The framework first generates approximate solutions using simplified physics to reduce model complexity, and then introduces PDE constraints and physical regularization in the loss function to ensure physically consistent predictions. Tests on benchmark cases and field data showed that this hybrid approach outperforms purely machine-learning methods in predicting

multiphase production rates and identifying inter-well connectivity, offering stronger physical interpretability and prediction stability. The results suggest that the method can balance computational efficiency and physical realism, providing a feasible pathway toward real-time optimization and closed-loop control for CO_2-EOR.

In 2025, Xu et al. [135] proposed the PINN-GCEM framework, integrating the connecting element method, GNNs, and PINNs to achieve inter-well connectivity identification and dynamic production prediction. This approach reportedly outperforms conventional models in both accuracy and efficiency, accurately capturing complex inter-well connectivity and demonstrating potential for real-time applications in CO_2-EOR optimization.

5 Conclusions and Perspectives

5.1 Conclusions

(1) Numerical simulation of CCUS-EOR has evolved from improved black-oil/pseudo-compositional models, to K-value models, and ultimately to EOS-based fully compositional models. Fully compositional simulation can more systematically capture key mechanisms such as dissolution, extraction, swelling, and miscibility effects. However, its high computational cost in large-scale three-dimensional models and in iterative history matching/optimization renders the conflict between "high fidelity" and "high efficiency" a primary bottleneck that constrains rapid field decision-making.

(2) To address this bottleneck, three major acceleration routes—multiscale, streamline, and reduced-order modeling—have each demonstrated distinct strengths. Multiscale methods embed fine-scale heterogeneity into coarse-scale solutions from the perspective of spatial discretization; streamline methods enable efficient scenario screening in convection-dominated settings; and reduced-order models are well suited to serve as fast prediction kernels in the outer loops of history matching and optimization. Together, these approaches have laid the groundwork for engineering-grade CCUS-EOR simulation frameworks with "controllable accuracy" and "affordable runtimes."

(3) Purely data-driven machine-learning and deep-learning methods have shown clear efficiency advantages in production forecasting, rapid MMP estimation, and well-pattern/well-control sensitivity analysis. In particular, emerging architectures such as GNNs, FNOs, and Transformers offer stronger representational power for complex well-network topologies and spatiotemporal coupling. Nevertheless, their reliability remains constrained by the representativeness of training samples, physical consistency, and generalization to out-of-distribution operating conditions.

(4) Data–physics coupled approaches have become one of the most engineering-promising directions. Simplified-mechanism models such as CRM/INSIM, network-based frameworks such as FlowNet/GPSNet, and physics-constrained deep-learning paradigms including the connecting element method and PINN-based models provide scalable and unified frameworks for inter-well-connectivity identification, dynamic prediction, and rapid optimization in complex heterogeneous reservoirs.

5.2 Perspectives

(1) Enhancing low-dimensional representations of complex phase behavior. For near-miscible or miscible CCUS-EOR involving intricate dissolution, extraction, and asphaltene-precipitation phenomena, future research should

develop more efficient phase-behavior surrogates. Alternatively, more accurate compositional flash algorithms can be embedded into simplified physics frameworks (e.g., connecting-element and network models) to overcome current limitations in representing complex thermodynamics.

(2) Building high-trust physics–data deep-fusion architectures. Future efforts should further explore how physical constraints are imposed, shifting from soft constraints (loss-function penalties) toward hard constraints (architecture design and operator learning). For example, physics-encoded Transformers or GNNs could be advanced to satisfy conservation laws while retaining strong capability in handling unstructured grids and long-range spatiotemporal dependencies.

(3) Improving robustness under small-sample and noisy data regimes. Field data are often sparse and noisy. Future models should integrate uncertainty quantification with transfer learning or few-shot learning to enhance inversion accuracy and practical value under limited monitoring conditions.

(4) Toward real-time optimization for closed-loop control. The ultimate goal is to embed efficient predictive models into real-time production-optimization and closed-loop management systems, enabling a transition from "offline evaluation" to "online regulation," thereby maximizing both the economic performance and storage efficiency of CCUS-EOR.

Acknowledgement

None.

Funding Statement

None.

Author Contributions

The authors confirm their responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Availability of Data and Materials

None.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Miao L, Feng L, Ma Y. Comprehensive evaluation of CCUS technology: A case study of China's first million-tonne CCUS-EOR project[J]. Environmental Impact Assessment Review, 2025, 110: 107684.

[2] Tan F, Jiang R, Ma C, et al. CO2 oil displacement and geological storage status and prospects[J]. Energy Science & Engineering, 2025, 13(2): 475-511.

[3] Watkins J W, Mardock E S. Use of Radioactive Iodine as a Tracer in Water-Flooding Operations[J]. Journal of Petroleum Technology, 1954, 6(09): 117-124.

[4] Lake L W. Enhanced Oil Recovery[M]. Englewood Cliffs: Prentice Hall, 1989.

[5] Eide L I, Batum M, Dixon T, et al. Enabling large-scale carbon capture, utilisation, and storage (CCUS) using offshore carbon dioxide (CO2) infrastructure developments—a review[J]. Energies, 2019, 12(10): 1945.

[6] Green D W, Willhite G P. Enhanced Oil Recovery[M]. Richardson, TX: SPE Textbook Series, 1998.

[7] Todd M R, Longstaff W J. The development, testing, and application of a numerical simulator for predicting miscible flood performance[J]. Journal of Petroleum Technology, 1972, 24(07): 874-882.

[8] Sandve T H, Vareid O S, Aavatsmark I. Improved Extended Blackoil Formulation for CO2-EOR Simulations[J], ECMOR XVII, Sep 2020, Vol 2020, 1-22.

[9] Sandve T H, Sævareid O, Aavatsmark I. Dynamic PVT model for CO2-EOR black-oil simulations[J]. Computational Geosciences, 2022, 26(4): 1029-1043.

[10] Bolling J D. Development and application of a limited-compositional, miscible flood reservoir simulator[C]//SPE Reservoir Simulation Conference. SPE, 1987: SPE-15998-MS.

[11] Van-Quy N, Simandoux P, Corteville J. A numerical study of diphasic multicomponent flow[J]. Society of Petroleum Engineers Journal, 1972, 12(02): 171-184.

[12] Young L C, Stephenson R E. A generalized compositional approach for reservoir simulation[J]. Society of Petroleum Engineers Journal, 1983, 23(05): 727-742.

[13] Rannou G, Voskov D, Tchelepi H. Tie-line-based k-value method for compositional simulation[J]. SPE Journal, 2013, 18(06): 1112-1122.

[14] Coats K H. An equation of state compositional model[J]. Society of Petroleum Engineers Journal, 1980, 20(5): 363–376.

[15] Laochamroonvorapongse R, Srisuriyachai F, Chewaroungroaj J. A parametric study of polymer-enhanced CO2 injection in a light-oil reservoir using full-physics simulation[J]. Journal of Petroleum Exploration and Production Technology, 2025, 15(2): 26.

[16] Acs G, Doleschall S, Farkas E. General purpose compositional model[J]. Society of Petroleum Engineers Journal, 1985, 25(04): 543-553.

[17] Chien M C H, Lee S T, Chen W H. A new fully implicit compositional simulator[C]. SPE Reservoir Simulation Conference. SPE, 1985: SPE-13385-MS.

[18] Orr Jr F M, Dindoruk B, Johns R T. Theory of multicomponent gas/oil displacements[J]. Industrial and engineering chemistry research, 1995, 34(8): 2661-2669.

[19] Christie M A, Blunt M J. Tenth SPE comparative solution project: a comparison of upscaling techniques [J]. SPE Reservoir Evaluation and Engineering, 2001, 4(4): 308-317.

[20] Hou T Y, Wu X H. A multiscale finite element method for elliptic problems in composite materials and porous media[J]. Journal of computational physics, 1997, 134(1): 169-189.

[21] Jenny P, Lee S H, Tchelepi H A. Multi-scale finite-volume method for elliptic problems in subsurface flow simulation[J]. Journal of computational physics, 2003, 187(1): 47-67.

[22] Moyner O, Lie K A. The multiscale finite volume method on unstructured grids[C]//2013 SPE Reservoir Simulation Symposium.2013.

[23] Torres J A, Jin L, Bosshart N W, Pekot L J, Sorensen J A, Peterson K J, et al. Multiscale modeling to evaluate the mechanisms controlling CO_2-based enhanced oil recovery and CO_2 storage in the Bakken Formation[C]. URTeC 2902837, 2018.

[24] Li S, Liang Y, Jiang F, et al. Multiscale simulation of water/oil displacement with dissolved CO2: Implications for geological carbon storage and CO2-enhanced oil recovery[J]. Chemical Engineering Journal, 2024, 499: 155936.

[25] Peng M, Zhou Y, Yang J, Wang J, Lin Z, Zhao J, et al. Experimental investigation and multi-scale simulations of CO_2 foam flooding for enhanced oil recovery[J]. Capillarity, 2025, 16(2): 39-50.

[26] Tchelepi H A, Lee S H, Jenny P. Multiscale finite-volume method for multiphase flow simulations in porous media[J]. Journal of Computational Physics, 2006, 211(2): 665-680.

[27] Thiele M R, Batycky R P, Blunt M J. A streamline-based 3d field-scale compositional reservoir simulator[C]. SPE Annual Technical Conference and Exhibition. SPE, 1997: SPE-38889-MS.

[28] Datta-Gupta A, King M J. Streamline simulation: theory and practice[M]. Society of Petroleum Engineers, 2007.

[29] Tanaka S, Datta-Gupta A, King M J. Compositional streamline simulation of CO2 injection account-

ing for gravity and capillary effects using orthogonal projection[C]. SPE Improved Oil Recovery Conference. SPE, 2014: SPE-169066-MS.

[30] Islam M A, Woobaidullah A S M. Integrated study of CO_2 sequestration and CO_2-EOR in oil reservoir by compositional simulation[J]. Dhaka University Journal of Science, 2022, 70(2): 23-28.

[31] Zhang R, Park J, Nunna K, et al. Streamline-Based Three-Phase History Matching for Field-Scale Compositional Modeling Under CO2 Enhanced-Oil-Recovery in the Permian Basin, Texas[C]. SPE Reservoir Simulation Conference. SPE, 2025.

[32] Chen H, Li A. CO2 Plume Monitoring and Visualization Using Pathlines, Source Cloud and Time Cloud[J]. Geoenergy Science and Engineering, 2025: 214093.

[33] Jin, Z. L., Durlofsky, L. J. Reduced-order modeling of CO2 storage operations[J]. International Journal of Greenhouse Gas Control, 2018, 68: 49-67.

[34] Chen Y, Durlofsky L J, Gerritsen M, et al. A coupled local–global upscaling approach for simulating flow in highly heterogeneous formations[J]. Advances in Water Resources, 2003, 26(10):1041-1060.

[35] Cardoso M A, Durlofsky L J. Linearized reduced-order models for subsurface flow simulation[J]. Journal of Computational Physics, 2011, 229(3):681-700.

[36] Zhao J, Jin L, Bosshart N, et al. Extending CO2 EOR and associated storage to low-pressure oil reservoirs by adding rich gas components to the injection stream[C]. Proceedings of the 16th Greenhouse Gas Control Technologies Conference (GHGT-16). 2022: 23-24.

[37] Ma H, McCoy S T, Chen Z. Development and comparison of reduced-order models for CO2-enhanced oil recovery predictions[J]. Energy, 2025, 320: 135313.

[38] Panja P, Velasco R, Pathak M, et al. Application of artificial intelligence to forecast hydrocarbon production from shales[J]. Petroleum, 2018, 4(1): 75-89.

[39] Wu H, Lubbers N, Viswanathan H S, et al. A multi-dimensional parametric study of variability in multi-phase flow dynamics during geologic CO2 sequestration accelerated with machine learning[J]. Applied Energy, 2021, 287: 116580.

[40] Tariq Z, Abu-Al-Saud M O, Mahmoud M, et al. Fast Tracking of Safe CO2 Trapping Indices using Machine Learning for Smarter Reservoir Management[J]. Petroleum, 2025.

[41] Helaleh A H, Alizadeh M. Performance prediction model of miscible surfactant-CO2 displacement in porous media using support vector machine regression with parameters selected by ant colony optimization[J]. Journal of Natural Gas Science and Engineering, 2016, 30: 388-404.

[42] Yue M, Dai Q, Liao H, et al. Prediction of ORF for Optimized CO2 Flooding in Fractured Tight Oil Reservoirs via Machine Learning[J]. Energies, 2024, 17(6): 1303.

[43] Gao M, Liu Z, Qian S, et al. Machine-learning-based approach to optimize CO2-WAG flooding in low permeability oil reservoirs[J]. Energies, 2023, 16(17): 6149.

[44] Thanh H V, Dashtgoli D S, Zhang H, et al. Machine-learning-based prediction of oil recovery factor for experimental CO2-Foam chemical EOR: Implications for carbon utilization projects[J]. Energy, 2023, 278: 127860.

[45] Shen B, Yang S, Gao X, et al. A novel CO2-EOR potential evaluation method based on BO-LightGBM algorithms using hybrid feature mining[J]. Geoenergy Science and Engineering, 2023, 222: 211427.

[46] Lv Q, Zheng R, Guo X, et al. Modelling minimum miscibility pressure of CO2-crude oil systems using deep learning, tree-based, and thermodynamic models: Application to CO2 sequestration and enhanced oil recovery[J]. Separation and Purification Technology, 2023, 310: 123086.

[47] Esfandi T, Sadeghnejad S, Jafari A. Effect of reservoir heterogeneity on well placement prediction in CO2-EOR projects using machine learning surrogate models: Benchmarking of boosting-based al-

gorithms[J]. Geoenergy Science and Engineering, 2024, 233: 212564.

[48] Xiao C, Lin H X, Leeuwenburgh O, et al. Surrogate-assisted inversion for large-scale history matching: Comparative study between projection-based reduced-order modeling and deep neural network[J]. Journal of Petroleum Science and Engineering, 2022, 208: 109287.

[49] Han Y, Hamon F P, Durlofsky L J. Accelerated training of deep learning surrogate models for surface displacement and flow, with application to MCMC-based history matching of CO2 storage operations[J]. Geoenergy Science and Engineering, 2025, 246: 213589.

[50] Ju X, Hamon F P, Wen G, et al. Learning CO2 plume migration in faulted reservoirs with Graph Neural Networks[J]. Computers and Geosciences, 2024, 193: 105711.

[51] Tang H, Durlofsky L J. Graph network surrogate model for optimizing the placement of horizontal injection wells for CO2 storage[J]. International Journal of Greenhouse Gas Control, 2025, 145: 104404.

[52] Feng Z, Yan B, Shen X, et al. A hybrid CNN-transformer surrogate model for the multi-objective robust optimization of geological carbon sequestration[J]. Advances in Water Resources, 2025, 196: 104897.

[53] Kim J, Yang H, Choe J. Robust optimization of the locations and types of multiple wells using CNN based proxy models[J]. Journal of Petroleum Science and Engineering, 2020, 193: 107424.

[54] Meng S, Fu Q, Tao J, et al. Predicting CO2-EOR and storage in low-permeability reservoirs with deep learning-based surrogate flow models[J]. Geoenergy Science and Engineering, 2024, 233: 212467.

[55] Yan B, Zhong Z, Bai B. A convolutional neural network-based proxy model for field production prediction and history matching[J]. Gas Science and Engineering, 2024, 122: 205219.

[56] Mo S, Zhu Y, Zabaras N, et al. Deep convolutional encoder-decoder networks for uncertainty quantification of dynamic multiphase flow in heterogeneous media[J]. Water Resources Research, 2019, 55(1): 703-728.

[57] Han Y, Hamon F P, Jiang S, et al. Surrogate model for geological CO2 storage and its use in hierarchical MCMC history matching[J]. Advances in Water Resources, 2024, 187: 104678.

[58] Xu Y, Zhao H, Pathe Gama Gamage R, et al. Advanced attention-based spatial-temporal neural networks for enhanced CO2 water-alternating-gas performance prediction and history matching[J]. Physics of Fluids, 2024, 36(9): 097116.

[59] Feng Z, Tariq Z, Shen X, et al. An encoder-decoder ConvLSTM surrogate model for simulating geological CO2 sequestration with dynamic well controls[J]. Gas Science and Engineering, 2024, 125: 205314.

[60] Iskandar U P, Kurihara M. Time-series forecasting of a CO2-EOR and CO2 storage project using a data-driven approach[J]. Energies, 2022, 15(13): 4768.

[61] Huang R, Wei C, Li B, et al. Prediction and optimization of WAG flooding by using LSTM neural network model in Middle East carbonate reservoir[C]. Abu Dhabi International Petroleum Exhibition and Conference. SPE, 2021: D011S012R003.

[62] Dong S, Wang L, Zeng L, et al. Fracture identification in reservoirs using well log data by window sliding recurrent neural network[J]. Geoenergy Science and Engineering, 2023, 230: 212165.

[63] Wang H, Han J, Zhang K, et al. An interpretable interflow simulated graph neural network for reservoir connectivity analysis[J]. SPE Journal, 2021, 26(04): 1636-1651.

[64] Davoodi S, Thanh H V, Wood D A, et al. Machine learning insights to CO2-EOR and storage simulations through a five-spot pattern–a theoretical study[J]. Expert Systems with Applications, 2024, 250: 123944.

[65] Wu K, Zhang J, Feng Q, et al. Co-optimization of well schedule and conformance control parameters assisted with Transformer-LSTM for CO2-EOR

and storage in oil reservoirs[J]. Energy, 2025, 331: 136891.

[66] Huang H, Gong B, Sun W. A deep-learning-based graph neural network-long-short-term memory model for reservoir simulation and optimization with varying well controls[J]. SPE Journal, 2023, 28(06): 2898-2916.

[67] Du L, Liu Y, Xue L, et al. A deep learning framework using graph convolutional networks for adaptive correction of interwell connectivity and gated recurrent unit for performance prediction[J]. SPE Reservoir Evaluation and Engineering, 2022, 25(04): 815-831.

[68] Huang Z Q, Wang Z X, Hu H F, et al. Dynamic interwell connectivity analysis of multi-layer waterflooding reservoirs based on an improved graph neural network[J]. Petroleum Science, 2024, 21(2): 1062-1080.

[69] Li B, Zhao H, Liu B, et al. Graph neural networks and hybrid optimization for water-flooding regulation[J]. Physics of Fluids, 2025, 37(8).

[70] Jiang J. Simulating multiphase flow in fractured media with graph neural networks[J]. Physics of Fluids, 2024, 36(2).

[71] Gudala M, Yan B. Graph neural network for multi-physics geothermal simulation with discrete fracture network[J]. Advances in Water Resources, 2025, 205: 105062.

[72] Ju X, Hamon F P, Wen G, et al. Learning CO2 plume migration in faulted reservoirs with Graph Neural Networks[J]. Computers and Geosciences, 2024, 193: 105711.

[73] Zhuang X Y, Wang W D, Su Y L, et al. Deep learning-assisted optimization for enhanced oil recovery and CO2 sequestration considering gas channeling constraints[J]. Petroleum Science, 2025, 22(8): 3397-3417.

[74] Tariq Z, Abualsaud M, He X, et al. A U-Net Enhanced Graph Neural Network to Simulate Geological Carbon Sequestration[J]. SPE Journal, 2025, 30(7): 1-19.

[75] Li Z, Kovachki N, Azizzadenesheli K, et al. Fourier neural operator for parametric partial differential equations[J]. arXiv preprint arXiv:2010.08895, 2020.

[76] Zhang K, Zuo Y, Zhao H, et al. Fourier neural operator for solving subsurface oil/water two-phase flow partial differential equation[J]. SPE Journal, 2022, 27(03): 1815-1830.

[77] Chu A K, Benson S M, Wen G. Deep-learning-based flow prediction for CO2 storage in shale–sandstone formations[J]. Energies, 2022, 16(1): 246.

[78] Liu J, Jing H, Pan H. New fast simulation of 4d (x, y, z, t) CO2 EOR by Fourier neural operator based deep learning method[C]. SPE Reservoir Simulation Conference. SPE, 2023: D010S001R006.

[79] Yang Z, Liao X, Dong P, et al. Rapid Prediction of CO2 Flooding Attribute Distribution in 3D Heterogeneous Reservoirs Using Fourier Neural Operator[J]. environment, 1: 5.

[80] Liu J, Pan H, Jiang S, et al. DL-NRS Net: A Physics-Informed Fourier Neural Operator Framework for High-Resolution Reconstruction Without High-Resolution Labels[J]. Mathematical Geosciences, 2025: 1-27.

[81] Zhang F, Nghiem L, Chen Z. Evaluating reservoir performance using a transformer based proxy model[J]. Geoenergy Science and Engineering, 2023, 226: 211644.

[82] Xiao Z, Shen B, Yang J, et al. Deep Learning Framework for Accurate Static and Dynamic Prediction of CO2 Enhanced Oil Recovery and Storage Capacity[J]. Processes, 2024, 12(8): 1693.

[83] Al-Ismael, Menhal A., and Abeeb A. Awotunde. Transformer-Based Generative Model for Automated Field Development Planning and Optimization in Hydrocarbon Reservoirs[J]. SPE Journal, 2025, SPE-231418-PA.

[84] Jia J, Li D, Wang L, et al. Novel Transformer-based deep neural network for the prediction of post-refracturing production from oil wells[J]. Advances in Geo-Energy Research, 2024, 13(2): 119-131.

[85] Liu B, Xu T, Xu Y, et al. Automated Reservoir History Matching Framework: Integrating Graph Neural Networks, Transformer, and Optimization for Enhanced Interwell Connectivity Inversion[J]. Processes, 2025, 13(5): 1386.

[86] Li J, Liu W, Yu M, et al. Reservoir production prediction based on improved graph attention network[J]. IEEE Access, 2023, 12: 50044-50056.

[87] Albertoni A, Lake L W. Inferring interwell connectivity only from well-rate fluctuations in waterfloods[J]. SPE reservoir evaluation and engineering, 2003, 6(01): 6-16.

[88] Yousef A A, Gentil P, Jensen J L, et al. A capacitance model to infer interwell connectivity from production-and injection-rate fluctuations[J]. SPE Reservoir Evaluation and Engineering, 2006, 9(06): 630-646.

[89] Lee K H, Ortega A, Nejad A M, et al. An active method for characterization of flow units between injection/production wells by injection-rate design[J]. SPE Reservoir Evaluation and Engineering, 2011, 14(04): 433-445.

[90] Sayarpour, Morteza. Development and application of capacitance-resistive models to water/carbon dioxide floods[D]. The University of Texas at Austin, 2008.

[91] Nguyen, A.P. Capacitance resistance modeling for primary recovery, waterflood and water-CO_2 flood[D]. University of Texas at Austin, 2012.

[92] Eshraghi S E, Rasaei M R, Zendehboudi S. Optimization of miscible CO2 EOR and storage using heuristic methods combined with capacitance/resistance and Gentil fractional flow models[J]. Journal of Natural Gas Science and Engineering, 2016, 32: 304-318.

[93] Zhao H, Kang Z, Zhang X, et al. INSIM: A data-driven model for history matching and prediction for waterflooding monitoring and management with a field application[C] SPE reservoir simulation symposium, 2015, SPE-173213-MS.

[94] Zhao H, Kang Z, Zhang X, et al. A physics-based data-driven numerical model for reservoir history matching and prediction with a field application[J]. SPE Journal, 2016, 21(06): 2175-2194.

[95] Zhao H, Li Y, Cui S, et al. History matching and production optimization of water flooding based on a data-driven interwell numerical simulation model[J]. Journal of Natural Gas Science and Engineering, 2016, 31: 48-66.

[96] ZHAO H, KANG Z, SUN H, et al. An interwell connectivity inversion model for waterflooded multilayer reservoirs[J]. Petroleum Exploration and Development, 2016, 43(1): 106-114.

[97] XIE X, ZHAO H, KANG X, et al. Prediction method of produced polymer concentration based on interwell connectivity[J]. Petroleum Exploration and Development, 2017, 44(2): 286-293.

[98] Guo Z, Reynolds A C, Zhao H. A physics-based data-driven model for history matching, prediction, and characterization of waterflooding performance[J]. SPE Journal, 2018, 23(02): 367-395.

[99] Guo Z, Reynolds A C, Zhao H. Waterflooding optimization with the INSIM-FT data-driven model[J]. Computational Geosciences, 2018, 22: 745-761.

[100] Zhao H, Xu L, Guo Z, et al. A new and fast waterflooding optimization workflow based on INSIM-derived injection efficiency with a field application[J]. Journal of Petroleum Science and Engineering, 2019, 179: 1186-1200.

[101] Guo Z, Reynolds A C. INSIM-FT in three-dimensions with gravity[J]. Journal of Computational Physics, 2019, 380: 143-169.

[102] Zhao H, Xu L, Guo Z, et al. Flow-path tracking strategy in a data-driven interwell numerical simulation model for waterflooding history matching and performance prediction with infill wells[J]. SPE Journal, 2020, 25(02): 1007-1025.

[103] Liu W, Zhao H, Sheng G, et al. A rapid waterflooding optimization method based on INSIM-FPT data-driven model and its application to three-dimensional reservoirs[J]. Fuel, 2021, 292: 120219.

[104] Zhao H, Liu W, Rao X, et al. INSIM-FPT-3D: A Data-driven Model for History Matching, Water-breakthrough Prediction and Well-connectivity

Characterization in Three-dimensional Reservoirs[C]. SPE Reservoir Simulation Conference, Galveston, Texas, USA, 4–6 October 2021, SPE-203931-MS.

[105] Liu W, Zhao H, Zhong X, et al. A Novel Data-Driven Model for Dynamic Prediction and Optimization of Profile Control in Multilayer Reservoirs[J]. Geofluids, 2021, 3272860.

[106] Lerlertpakdee P, Jafarpour B, Gildin E. Efficient production optimization with flow-network models[J]. SPE Journal, 2014, 19(06): 1083-1095.

[107] Lutidze G. StellNet: Physics-Based Data-Driven General Model for Closed-Loop Reservoir Management[D]. The University of Tulsa, 2018.

[108] Ren G, He J, Wang Z, et al. Implementation of physics-based data-driven models with a commercial simulator[C]. SPE Reservoir Simulation Conference. SPE, 2019: D010S017R010.

[109] Wang Z, He J, Milliken W J, et al. Fast history matching and optimization using a novel physics-based data-driven model: an application to a diatomite reservoir[J]. SPE Journal, 2021, 26(06): 4089-4108.

[110] Guan X, Wang Z, Kostakis F, et al. Fast history matching and optimization using a novel physics-based data-driven model: An application to a diatomite reservoir with hundreds of wells[J]. Geoenergy Science and Engineering, 2023, 228: 211919.

[111] Wang Z, Guan X, Milliken W, et al. Fast history matching and robust optimization using a novel physics-based data-driven flow network model: An application to a Steamflood sector model[J]. SPE Journal, 2022, 27(04): 2033-2051.

[112] Wang Z, Li C, King M J. Applications of Asymptotic Solutions of the Diffusivity Equation to Infinite Acting Pressure Transient Analysis[J]. SPE Journal, 2024, 29(08): 4069-4093.

[113] Kiærr A, Lødøen O P, De Bruin W, et al. Evaluation of a data-driven flow network model (FlowNet) for reservoir prediction and optimization[C]. ECMOR XVII. European Association of Geoscientists and Engineers, 2020, 2020(1): 1-18.

[114] Rao X, Xu Y, Liu D, et al. A general physics-based data-driven framework for numerical simulation and history matching of reservoirs[J]. Advances in Geo-Energy Research, 2021, 5(4): 422-436.

[115] Zhao H, Rao X, Liu D, et al. A flownet-based method for history matching and production prediction of shale or tight reservoirs with fracturing treatment[J]. SPE Journal, 2022, 27(05): 2793-2819.

[116] Xu Y, Liu W, Zhao H, et al. Application of the FlowNet model for history matching and production optimization in water alternating gas enhanced oil recovery[J]. Physics of Fluids, 2024, 36(8): 083323.

[117] Zhao H, Zhan W, Yuhui Z, et al. A connection element method: Both a new computational method and a physical data-driven framework—-Take subsurface two-phase flow as an example[J]. Engineering Analysis with Boundary Elements, 2023, 151: 473-489.

[118] Liu W, Xu Y, Rao X, et al. Reservoir closed-loop optimization method based on connection elements and data space inversion with variable controls[J]. Physics of Fluids, 2023, 35(11).

[119] Zhao H, Zhan W, Chen Z, et al. A novel connection element method for multiscale numerical simulation of two-phase flow in fractured reservoirs[J]. SPE Journal, 2024, 29(09): 4950-4973.

[120] Yunfeng Xu, Wei Liu, Hui Zhao, et al. Simulation of CO2 flooding and prediction method for conformance channels based on connection elements[J]. Journal of China University of Petroleum, 2025, 49(3): 158-168.

[121] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational physics, 2019, 378: 686-707.

[122] Wang N, Chen Y, Zhang D. A comprehensive review of physics-informed deep learning and its applications in geoenergy development[J]. The In-

novation Energy, 2025, 2(2): 100087-1-100087-15.

[123] Han J X, Xue L, Wei Y S, et al. Physics-informed neural network-based petroleum reservoir simulation with sparse data using domain decomposition[J]. Petroleum Science, 2023, 20(6): 3450-3460.

[124] Han J, Xue L, Jia Y, et al. Prediction of Porous Media Fluid Flow with Spatial Heterogeneity Using Criss-Cross Physics-Informed Convolutional Neural Networks[J]. Computer Modeling in Engineering and Sciences, 2024, 138(2), 1323-1340.

[125] Liu A, Li J, Bi J, et al. A novel reservoir simulation model based on physics informed neural networks[J]. Physics of Fluids, 2024, 36(11): 116617.

[126] Han J, Xue L, Jia Y, et al. Prediction of Porous Media Fluid Flow with Spatial Heterogeneity Using Criss-Cross Physics-Informed Convolutional Neural Networks[J]. CMES-Computer Modeling in Engineering and Sciences, 2024, 140(2): 1323-1340.

[127] Liu B, Wei J, Kang L, et al. Physics-informed neural network (PINNs) for convection equations in polymer flooding reservoirs[J]. Physics of Fluids, 2025, 37(3): 036622.

[128] Wang, N, Zhang D, Chang H, et al. Deep learning of subsurface flow via theory-guided neural network[J]. Journal of Hydrology, 2020, 584: 124700.

[129] Xu R, Zhang D, Rong M, et al. Weak form theory-guided neural network (TgNN-wf) for deep learning of subsurface single-and two-phase flow[J]. Journal of Computational Physics. 2021,436:110318.

[130] Wang N, Chang H, Zhang D. Theory-guided auto-encoder for surrogate construction and inverse modeling[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 385: 114037.

[131] Chen H Y, Xue L, Liu L, et al. Physics-informed graph neural network for predicting fluid flow in porous media[J]. Petroleum Science, 2025.

[132] Zhang Z. A physics-informed deep convolutional neural network for simulating and predicting transient Darcy flows in heterogeneous reservoirs without labeled data[J]. Journal of Petroleum Science and Engineering, 2022, 211: 110179.

[133] Wang Y W, Dai Z X, Wang G S, et al. A hybrid physics-informed data-driven neural network for CO2 storage in depleted shale reservoirs[J]. Petroleum Science, 2024, 21(1): 286-301.

[134] Nagao M, Datta-Gupta A. Physics Informed Machine Learning for Reservoir Connectivity Identification and Production Forecasting for CO2-EOR[C]. Paper presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA, September 2024.

[135] Xu, Y, Zhao H, Jia D, et al. Physics-Informed Graph Connected Element Model for Reservoir Connectivity Identification and Production Forecasting in CO2-EOR[C]. Paper presented at the GOTECH, Dubai City, UAE, April 2025.