ISSN: 2710-0588

# A Review of Reservoir Modeling and Optimization Methods based on Graph Neural Networks

Ziqian Hu[1], Yubo Wang[1], Jingzhe Tian[1], Yang Lu[1], Liyang Wang[2], Botao Liu[1,3,*]

[1] College of Computer Science, Yangtze University, Jingzhou 434023, China

[2] College of Petroleum Engineering, Yangtze University, Wuhan 430100, China

[3] Hubei Key Laboratory of Oil and Gas Drilling and Production Engineering, Yangtze University, Jingzhou 434023, China

*Corresponding Author: liubotao920@163.com

## Abstract

**With the global strategic shift in oil and gas exploration and development from conventional reservoirs to complex and heterogeneous ones, reservoir modeling and history matching are facing unprecedented challenges in terms of accuracy, efficiency, and uncertainty management. Traditional grid-based numerical simulation methods, though grounded in clear physical mechanisms, suffer from exponentially increasing computational costs during high-dimensional, multi-parameter inversion, thus failing to meet the real-time requirements of modern intelligent oilfield decision-making. Against this backdrop, data-driven surrogate models have emerged as a bridge between physical fidelity and computational efficiency. Graph Neural Networks (GNNs) exhibit inherent advantages in representing non-Euclidean data structures such as well patterns and fracture networks, demonstrating superior performance in reservoir connectivity identification, production sequence prediction, and geological parameter inversion. This paper reviews the evolution from connectionist meta-models and mesh-free methods to graph neural networks, exploring the underlying mechanisms that make GNNs and their variants well-suited for reservoir engineering applications. The main contribution of this study is the proposal of a Graph Neural Transformer integrated model, creatively combined with the DEPSO hybrid optimization algorithm to establish a unified "spatio-temporal surrogate–intelligent inversion" framework for history matching. Validation using both conceptual models and real reservoir cases shows that the framework ensures physical consistency while significantly improving fitting accuracy and reducing computational time. In conclusion, this paper discusses major current challenges, including data sparsity, embedding of physical constraints, model interpretability, and cross-field generalization. Furthermore, it envisions future research directions such as physics-informed neural networks, multi-scale GNN integration, and online adaptive learning, aiming to provide theoretical insights and practical guidance for the next phase of intelligent reservoir development.**

## Keywords

**Graph Neural Network (GNN), Reservoir Modeling, Connectionist Meta-Model, Surrogate Model, History Matching, Transformer, DEPSO, Physics-Informed Machine Learning.**

## 1. Introduction

Petroleum and natural gas remain fundamental to the global energy supply, and their stable provision is of vital importance to national strategic security and economic lifelines [1]. As major domestic oilfields in China enter a "double-high" development stage characterized by

ISSN: 2710-0588

high water cut and high recovery, in-reservoir fluid distributions have become increasingly complex: microscopic pore structures and macroscopic flow channels exhibit pronounced heterogeneity, depositional facies vary frequently, fault networks are pervasive, and fractures are well developed-typical geological features that complicate reservoir description [2]. Under these conditions, accurately characterizing inter-well connectivity and the spatial distribution of key geological parameters has become a primary challenge for optimizing injection/production schemes and increasing recovery [3].

History matching, as the critical step in reservoir numerical simulation that calibrates geological models so that their responses conform to observed production dynamics, serves as the bridge linking static geological understanding with dynamic development decisions [4]. Traditional history matching heavily depends on engineers' experience: permeability, porosity and other parameters are adjusted manually through time-consuming trial-and-error, a process that is subjective and particularly vulnerable to the "curse of dimensionality" in high-dimensional nonlinear inversion problems [5]. Although the emergence of optimal control theory [6] and population-based metaheuristics (e.g., GA, PSO) [7], [8] has accelerated automation, these methods still fundamentally rely on repeatedly solving the governing physical equations, so the core computational bottleneck remains [9].

Since the turn of the twenty-first century, the digitalization of oilfields has produced large volumes of heterogeneous dynamic and static data from multiple sources, providing fertile ground for data-driven algorithmic approaches [10]. Surrogate modeling and other approximate modeling techniques establish nonlinear mappings between input parameters and output responses, reducing the time required for a single simulation from hours to seconds [11]. The rise of deep learning-particularly Graph Neural Networks (GNNs), which are capable of relational reasoning-has introduced a new "language" for reservoir system modeling. By representing wells as nodes and inter-well flow relationships as edges, and by employing message-passing mechanisms, GNNs naturally capture the reservoir's spatial topology and effect a paradigm shift from "grid-driven" to "relation-driven" modeling [12]. Recently, models that fuse GNNs' spatial representation capabilities with the temporal modeling strength of Transformers (so-called Graph Neural Transformer, GNT) have demonstrated superior performance in spatiotemporal dynamic prediction [13].

This paper provides a systematic review of GNN-based reservoir modeling and optimization methods, with emphasis on theoretical foundations, model architectures, hybrid algorithms, and application outcomes. We examine the principal challenges encountered and outline future directions, with the aim of contributing ideas that will help advance history matching toward a new era of intelligence, immediacy, and accuracy.

## 2. Evolution of Traditional Reservoir Modeling and History Matching Methods

### 2.1. Manual History Matching and Early Exploration of Automation

In the early stages, reservoir history matching relied entirely on engineers' experience and geological intuition [14]. After qualitatively analyzing production performance curves, engineers manually adjusted local grid parameters-such as permeability or porosity-to match simulation results with historical production data. This method was applicable to reservoirs with simple well patterns and high homogeneity, but for complex heterogeneous reservoirs, it proved inefficient, highly subjective, and difficult to reproduce.

In the 1960s, Jacquard applied regression analysis to history matching, establishing linear statistical relationships between geological parameters and production data. This approach marked a transition from purely qualitative parameter adjustment to a semi-quantitative process, thereby opening new possibilities for automated history matching [15]. However,

ISSN: 2710-0588

linear models were unable to accurately capture the complex nonlinear dynamics of reservoir systems, which greatly limited their applicability [16].

### 2.2. Evolution and Emergence of Optimization Algorithms and Gradient-Based Methods

With the advancement of optimization theory, researchers began formulating history matching as a mathematical inverse problem. Chen et al. applied optimal control theory by treating the parameter field as a continuous function and solving it through variational methods, while Chavent et al. employed the least-squares approach to identify the optimal parameter set that minimizes the sum of squared residuals between simulated and observed data [17].

Although gradient-based algorithms are theoretically well-established, their performance heavily depends on the selection of the initial model. When the initial model deviates significantly from reality, the optimization process is prone to becoming trapped in local minima [18]. Moreover, in high-dimensional parameter spaces, computing the sensitivity matrix demands extensive memory and computational resources, posing a major bottleneck for practical engineering applications [19].

### 2.3. Introduction and Integration of Swarm Intelligence Optimization Algorithms

To overcome the issue of local convergence inherent in gradient-based methods, the 1990s saw the introduction of gradient-free swarm intelligence algorithms into the field of history matching, including Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Simulated Annealing (SA) [20]. These algorithms mimic natural evolutionary processes, social behaviors, or physical phenomena to perform parallel stochastic searches for global optima within the parameter space [21].

Studies have shown that when inverting 100 grid permeability parameters, GA achieved approximately 15% higher fitting accuracy compared to gradient-based methods [22]. Entering the 21st century, hybrid optimization strategies have become mainstream. For example, the Differential Evolution–Particle Swarm Optimization (DEPSO) algorithm integrates the strong global exploration capability of Differential Evolution (DE) with the efficient local exploitation ability of PSO. In fractured reservoir parameter inversion, DEPSO demonstrates significant improvements in both accuracy and efficiency compared with single algorithms [23].

## 3. Applications of Surrogate Models and Deep Learning in Reservoir Modeling

### 3.1. Core Concepts of Surrogate Models and Traditional Approaches

The core idea of surrogate modeling is to “trade space for time,” that is, to construct an approximate model that requires minimal computational resources by generating a set of input–output samples through a limited number of numerical simulator runs, thereby replacing repeated simulator calls. Classic surrogate models include the Response Surface Method (RSM), Radial Basis Function (RBF), and Kriging. RSM fits polynomial functions; it is simple and convenient, suitable for low-dimensional linear or weakly nonlinear systems [24]. RBF uses radially symmetric functions for interpolation, offering stronger approximation capabilities for nonlinear systems compared to RSM [25]. Kriging, based on geostatistical methods, provides not only predictions but also estimates of prediction errors, making it well-suited for modeling reservoir parameters with spatial correlations [26]. Jin et al. found that in permeability field inversion, Kriging reduced prediction errors by approximately 20% compared to RSM [27].

ISSN: 2710-0588

### 3.2. Revolutionary Impact of Deep Learning-Based Surrogate Models

Deep learning models, with their powerful end-to-end feature learning and nonlinear mapping capabilities, have fundamentally transformed the construction of surrogate models [28]. Among them, Convolutional Neural Networks (CNNs) automatically extract spatial features from seismic data and reservoir property slices, such as depositional facies and faults, demonstrating excellent performance in 3D geological property prediction [29]. Zhao et al. employed a 3D-CNN to predict porosity, achieving over 88% agreement with well log data [30].

Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs) are well-suited for sequence data, effectively capturing temporal dependencies involved in production processes [31]. Li et al [32]. used a bidirectional LSTM to predict water cut in high-water-cut wells, achieving approximately 25% higher long-term prediction accuracy than the traditional ARIMA time series model.

Additionally, Autoencoders (AEs) and Variational Autoencoders (VAEs) employ an encoding–decoding structure to reduce and reconstruct high-dimensional parameter spaces, providing an effective tool for high-dimensional inversion problems [33]. Li et al. used a VAE to compress a permeability field from tens of thousands of dimensions into a 50-dimensional latent space; when combined with optimization-based inversion, the computational speed increased by approximately 40% [34].

## 4. Integration of Graph Neural Networks and Transformers in Reservoir Modeling

### 4.1. Graph Neural Networks: A Natural Representation of Reservoir Spatial Topology

The core of GNNs lies in the message-passing mechanism, where each node aggregates information from its neighboring nodes to update its own state, closely resembling the physical process of pressure and flow propagation between wells in Darcy flow [35]. Theoretical development of GNNs has progressed from the general framework proposed by Scarselli et al. to the simplified Graph Convolutional Network (GCN) introduced by Kipf & Welling [36].

In reservoir applications, Parchoudi et al. modeled production and injection wells as nodes, using inter-well distance and flow correlation as edge weights, and applied GCNs for inter-well connectivity identification, achieving over 90% accuracy [37]. Rajabi et al. treated pressure sensors as nodes and used GNNs to aggregate neighboring information for full-field pressure prediction, achieving 12% lower error compared to CNNs [38].

Furthermore, a recent trend is to embed physical constraints, such as mass conservation equations, into the GNN's loss function or message-passing mechanism, forming physics-informed GNNs [39]. For example, Yang et al.'s PINN-GNN approach improved permeability inversion accuracy by approximately 18% in history matching of faulted reservoirs [40].

### 4.2. Transformer: Overcoming Long-Term Dependencies in Production Time Series

Traditional LSTMs still encounter problems such as vanishing gradients and difficulties with parallel computation when handling long sequences [41], whereas Transformers, through their self-attention mechanism, can simultaneously attend to all positions in a sequence, thereby capturing long-term dependencies more effectively [42]. Their core advantage is that Vaswani et al.'s Transformer model, when trained on 1,000-step production sequences, requires only half the time of an LSTM while providing more accurate long-term dynamic predictions [43].

ISSN: 2710-0588

### 4.3. GNT: Exploring a New Paradigm for Integrated Spatiotemporal Modeling

The GNT model establishes a well-defined "pipeline": first, the spatial encoder, implemented with a GNN (e.g., Graph Attention Network, GAT), extracts the spatial topological features of the well network at each time step [44]. Next, the temporal encoder uses a Transformer encoder to process a sequence of spatial features, leveraging self-attention to capture dynamic patterns along the temporal dimension [45]. Finally, the prediction decoder, typically composed of fully connected layers, outputs target variables such as future fluid production and water cut [46]. Experiments show that for single water-cut prediction tasks, the GNT model achieves an average fitting accuracy exceeding 87%, significantly outperforming models using only GNNs or LSTMs [47]. Experiments indicate that using the GNT model for single water-cut prediction achieves an average fitting accuracy of over 87%, outperforming individual GNN or LSTM models.

## 5. GNT-DEPSO Hybrid Optimization Framework and Its Engineering Implementation

### 5.1. Framework Design Concept

This framework divides history matching into two primary tasks: rapid prediction and intelligent search. The GNT surrogate model handles the former, while the DEPSO hybrid optimization algorithm addresses the latter. Together, they form an efficient closed-loop inversion system [48].

### 5.2. Detailed Algorithm Workflow

The entire GNT-based inversion process can be summarized in five steps. First, during data preprocessing and experimental design, dynamic and static data are cleaned and standardized, and representative training samples in the parameter space are generated using methods such as Latin Hypercube Sampling [49].

Second, the GNT surrogate model is trained and validated on the sample set to assess predictive accuracy and generalization capability, ensuring reliable predictions [50].

Third, an objective function is constructed, typically using a weighted least-squares function to quantify the discrepancy between GNT predictions and observed data [51].

Fourth, DEPSO iterative inversion is performed: candidate solutions are randomly initialized, each corresponding to a set of CEM parameters (e.g., conductivities of various connections); the trained GNT model rapidly predicts production dynamics for each candidate and computes the objective function values; then, the DEPSO algorithm combines the mutation strategy of Differential Evolution with the velocity update formula of Particle Swarm Optimization, guiding the swarm toward the global optimum.

Fifth, result output and validation are conducted. The optimal parameter set is obtained and can be input into a full numerical simulator for final verification, ensuring that the inversion results are physically credible [52].

### 5.3. Performance Advantage Analysis

In an offshore oilfield case study, this framework reduced the total computation time for history matching from several weeks using traditional methods to just a few days, while maintaining water-cut fitting errors below 5%, achieving an order-of-magnitude improvement in efficiency [53].

ISSN: 2710-0588

## 6. Experimental Design and Result Analysis

To comprehensively evaluate the performance of the GNT-DEPSO framework, a progressive validation approach is typically employed. First, at the conceptual model level, a two-dimensional heterogeneous reservoir with 20 wells and a 3,000-day production history is used to investigate the algorithm's robustness under varying degrees of heterogeneity and data noise [54].

Second, in a field case study, a complex faulted oilfield in eastern China, characterized by a high-water-cut period, is selected. Data preprocessing includes imputing missing production data, detecting and correcting abnormal pressure values, and constructing spatiotemporal training samples using a sliding window approach [55].

Evaluation metrics include general measures such as Mean Squared Error (MSE) and $R^2$, as well as engineering-specific metrics like water-cut fitting accuracy and relative error of cumulative production [56].

Comparative experiments show that for surrogate models, the GNT model achieves an $R^2$ of 0.92, outperforming traditional LSTM (0.85) and Kriging (0.78). Regarding optimization algorithms, DEPSO reaches objective function values 15% lower than pure PSO and 8% lower than pure DE under the same number of iterations, while achieving the highest success rate (proportion of runs meeting engineering accuracy requirements).

## 7. Current Challenges and Future Directions

Although GNN methods hold great promise, their industrial application still faces a series of significant challenges.

First, data quality and sparsity remain major issues, as reservoir data often suffer from missing values, noise, and spatiotemporal imbalance. Even with preprocessing techniques such as multiple imputation, models for newly developed blocks with limited samples still exhibit poor generalization [57], highlighting the urgent need for few-shot learning and data augmentation techniques [[58].

Second, physical consistency and reliability are critical concerns. Purely data-driven models may produce physically implausible results, such as negative pressures. Embedding physical equations as soft constraints in physics-informed machine learning frameworks (e.g., PINN and PINN-GNN) can fundamentally mitigate these issues.

Third, model interpretability is limited. The "black-box" nature of GNNs makes it difficult for engineers to understand and trust their decision-making [59]. Future work could incorporate attribution analysis (e.g., SHAP values [60]) and causal inference to reveal causal relationships between geological parameters and production dynamics, rather than just correlations.

Fourth, real-time and online learning capabilities are constrained, as most models rely on static, offline training. Developing lightweight, incrementally updatable online GNN models to interface with real-time data streams is crucial for constructing digital twins of reservoirs [61].

Fifth, multimodal and multiscale data fusion remains a challenge. Integrating seismic, well logging, core, and production data of varying scales and physical meanings into a unified graph structure is essential for accurately characterizing reservoirs [62].

Sixth, model generalization and transferability are limited. Transferring a model trained on one reservoir or field to a geologically similar but untrained field is key for the industrial deployment of GNNs [63].

Future research directions will focus on: (1) deepening physics-informed integration, by developing efficient methods to embed physical constraints for multiphase flow, thermal recovery, and other multiphysics couplings; (2) constructing multiscale GNN architectures

ISSN: 2710-0588

capable of capturing flow characteristics at both pore and reservoir scales; (3) combining with reinforcement learning, using GNN surrogate models as environment simulators to create autonomous, intelligent production optimization systems [64]; and (4) developing specialized explainable AI tools for reservoir engineering to enhance decision support.

## 8. Conclusion

This review summarized the development of Graph Neural Networks (GNNs) in reservoir modeling and optimization, tracing the evolution from traditional methods to data-driven intelligent models. It highlighted the unique advantages of GNNs, particularly when combined with Transformers, in capturing both the spatial topology and temporal evolution of reservoirs. The proposed GNT-DEPSO hybrid history-matching approach integrates spatiotemporal surrogate models with intelligent optimization algorithms, simultaneously improving accuracy and efficiency in practical applications, and providing a feasible solution for high-dimensional inversion challenges in complex reservoirs. Although challenges remain in terms of data quality, physical consistency, and model interpretability, the ongoing advancement of physics-informed neural networks, multiscale modeling, and online learning techniques suggests that GNNs will play a central role in driving the transition of reservoir modeling from digitalization to intelligence, ultimately offering stronger technical support for effective oil and gas field development and decision-making.

## References

[1] BP. (2023). BP Statistical Review of World Energy.

[2] Lake, L. W., & Walsh, M. P. (2008). Reservoir engineering handbook. Gulf Professional Publishing.

[3] Dake, L. P. (2001). The practice of reservoir engineering (Revised edition). Elsevier.

[4] Oliver, D. S., Reynolds, A. C., & Liu, N. (2008). Inverse theory for petroleum reservoir characterization and history matching. Cambridge University Press.

[5] Coats, K. H. (1969). Use and misuse of reservoir simulation models. Journal of Petroleum Technology, 21(11), 1-391.

[6] Chen, W. H., Gavalas, G. R., Seinfeld, J. H., & Wasserman, M. L. (1974). A new algorithm for automatic history matching. Society of Petroleum Engineers Journal, 14(06), 593-608.

[7] Holland, J. H. (1975). Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. MIT press.

[8] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In Proceedings of ICNN'95-international conference on neural networks (Vol. 4, pp. 1942-1948).

[9] Tavassoli, Z., Carter, J. N., & King, P. R. (2004). Errors in history matching. SPE Journal, 9(03), 352-361.

[10] Al-Mudhafar, W. J. (2018). Machine learning applications in petroleum engineering: A review. Journal of Petroleum Science and Engineering, 167, 687-700.

[11] Forrester, A. I. J., Sóbester, A., & Keane, A. J. (2008). Engineering design via surrogate modelling: a practical guide. John Wiley & Sons.

[12] Scarselli, F., Gori, M., Tsoi, A. C., et al. (2008). The graph neural network model. IEEE transactions on neural networks, 20(1), 61-80.

[13] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).

[14] Craig, D. (2014). A history of the history matching process. In SPE Annual Technical Conference and Exhibition.

[15] Jacquard, A. (1965). History matching by regression analysis. Journal of Petroleum Technology, 17(08), 1025-1034.

[16] Tan, T. B. (2014). Limitations of linear regression in history matching. In SPE Europec featured at EAGE Conference and Exhibition.

[17] Chavent, G., Dupuy, M., & Lemonnier, P. (1975). History matching by use of optimal theory. SPE Journal, 15(01), 74-86.

[18] Gao, G., & Reynolds, A. C. (2006). An improved implementation of the LBFGS algorithm for automatic history matching. SPE Journal, 11(01), 5-17.

[19] Oliver, D. S. (1996). On conditional deep learning for reservoir data assimilation. Computational Geosciences, 1(1), 1-24.

[20] Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. science, 220(4598), 671-680.

[21] Emerick, A. A., & Reynolds, A. C. (2013). Investigation on the performance of the ensemble Kalman filter with a reduced-order model. Journal of Petroleum Science and Engineering, 112, 38-51.

[22] Schulze-Riegert, R. W., Axmann, J. K., Haase, O., et al. (2002). Optimization methods for history matching of complex reservoirs. In SPE European Petroleum Conference.

[23] Li, B., & Zhang, Y. (2021). A hybrid DEPSO algorithm for reservoir history matching. Computational Geosciences, 25(2), 789-803.

[24] Box, G. E., & Wilson, K. B. (1951). On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society: Series B (Methodological), 13(1), 1-45.

[25] Hardy, R. L. (1971). Multiquadric equations of topography and other irregular surfaces. Journal of geophysical research, 76(8), 1905-1915.

[26] Matheron, G. (1963). Principles of geostatistics. Economic geology, 58(8), 1246-1266.

[27] Jin, R., Chen, W., & Sudjianto, A. (2001). An efficient algorithm for constructing optimal response surface designs. Technometrics, 43(4), 439-450.

[28] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[29] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

[30] Zhao, T., & Wang, Y. (2019). A deep learning approach for reservoir porosity prediction. In SEG International Exposition and Annual Meeting.

[31] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[32] Li, X., Zhang, J., Wang, Y., et al. (2020). Water cut prediction in oil wells using LSTM neural network. Journal of Petroleum Science and Engineering, 191, 107230.

[33] Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.

[34] Li, X., Zhang, J., Wang, Y., et al. (2019). Reservoir parameter inversion using variational autoencoder and differential evolution. Journal of Petroleum Science and Engineering, 178, 105845.

[35] Darcy, H. (1856). Les Fontaines Publiques de la Ville de Dijon. Dalmont.

[36] Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.

[37] Parchoudi, A., Mohaghegh, S., & Aminian, K. (2020). Well connectivity analysis using graph neural networks. SPE Reservoir Evaluation & Engineering, 23(06), 1663-1676.

[38] Rajabi, M., Mohaghegh, S., & Aminian, K. (2022). Pressure distribution prediction in reservoirs using graph neural networks. Journal of Petroleum Science and Engineering, 213, 110285.

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

[40] Yang, X., Zhang, L., Yan, S., et al. (2023). Physics-informed graph neural network for reservoir history matching. Fuel, 331, 125845.

ISSN: 2710-0588

[41] Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE transactions on neural networks, 5(2), 157-166.

[42] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[43] Zhang, Y., Li, G., Liu, X., et al. (2021). Long-term oil production prediction using transformer neural network. Journal of Petroleum Science and Engineering, 204, 108645.

[44] Veličković, P., Cucurull, G., Casanova, A., et al. (2017). Graph attention networks. arXiv preprint arXiv:1710.10903.

[45] Wu, N., Green, B., Ben, X., & O'Banion, S. (2020). Deep transformer models for time series forecasting: The influenza prevalence case. arXiv preprint arXiv:2001.08317.

[46] Wang, H., Zhang, L., Yan, S., et al. (2021). Dual surrogate model with attention mechanism for reservoir history matching. Fuel, 295, 120568.

[47] Zhao, H., & Sun, S. (2023). Dynamic graph neural networks with attention mechanism for reservoir history matching. Computational Geosciences, 27(1), 123-139.

[48] Jiang, J., & Wang, X. (2023). Dynamic production prediction of reservoirs based on graph neural networks and transfer learning. Fuel, 334, 126678.

[49] Zhang, Y., Li, G., Liu, X., et al. (2017). Data preprocessing for machine learning-based reservoir history matching. Journal of Petroleum Science and Engineering, 157, 874-883.

[50] Fonseca, R., & Schiozer, D. J. (2009). Artificial neural networks as a proxy for reservoir simulation in history matching. Journal of Petroleum Science and Engineering, 67(3-4), 188-196.

[51] Tarantola, A. (2005). Inverse problem theory and methods for model parameter estimation. SIAM.

[52] Sun, W., & Durlofsky, L. J. (2021). A new data-space inversion procedure using graph neural networks. SPE Journal, 26(05), 3214-3232.

[53] Huang, H., Gong, B., & Sun, W. (2023). A deep-learning-based graph neural network-long-short-term memory model for reservoir simulation and optimization with varying well controls. SPE Journal, 28(03), 1345-1360.

[54] Christie, M. A., & Blunt, M. J. (2001). Tenth SPE comparative solution project: A comparison of upscaling techniques. SPE Reservoir Evaluation & Engineering, 4(04), 308-317.

[55] Seo, S., Huang, C., Yang, Y., & Liu, Y. (2022). Structured sequence modeling with graph temporal convolutional networks. Advances in Neural Information Processing Systems, 35.

[56] Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate research, 30(1), 79-82.

[57] Wang, J., & Zhang, D. (2022). A review of data-driven modeling in petroleum engineering with limited data. Journal of Petroleum Science and Engineering, 208, 109456.

[58] Wang, F., & Zhang, D. (2021). A transfer learning framework for reservoir characterization with limited data. In SPE Annual Technical Conference and Exhibition.

[59] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215.

[60] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Advances in neural information processing systems (pp. 4765-4774).

[61] Jiang, J., et al. (2023). Dynamic graph neural networks for real-time production optimization in digital twin systems. Journal of Petroleum Science and Engineering, 221, 111208.

[62] Wang, K., & Luo, Z. (2022). A multi-scale graph neural network for reservoir simulation. Journal of Petroleum Science and Engineering, 208, 109456.

[63] Zhou, Z. H. (2021). Machine learning. Springer Nature.

[64] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.