Advances in Geo-Energy Research

Vol. 17, No. 3, p. 179-183, 2025

Perspective

Artificial intelligence applications and challenges in oil and gas exploration and development

Gang Hui[1,2,3], Yili Ren[4], Jianfei Bi[4], Muming Wang[3]*, Chengqian Liu[1]

[1]*State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum, Beijing 102249, P. R. China*

[2]*State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu 610059, P. R. China*

[3]*Department of Chemical and Petroleum Engineering, University of Calgary, Calgary, AB, T2N 1N4, Canada*

[4]*Research Institute of Petroleum Exploration and Development CNPC, Beijing 100083, P. R. China*

Keywords:
Artificial intelligence
data-driven
image analytics
physics-informed
multimodal

Cited as:
Hui G., Ren Y., Bi J., Wang M., Liu C. Artificial intelligence applications and challenges in oil and gas exploration and development. Advances in Geo-Energy Research, 2025, 17(3): 179-183.
https://doi.org/10.46690/ager.2025.09.01

Abstract:
The rapid integration of artificial intelligence into oil and gas exploration and development offers transformative opportunities within the context of the global energy transition. This article highlights the key advancements and challenges in artificial intelligence applications. Machine learning algorithms enable data-driven shale sweet spot prediction, overcoming the limitations of traditional methods by capturing complex controlling factors. Intelligent core image analysis, leveraging computer vision and foundation models, enables automatic mineral identification, pore analysis, and rock structure characterization, thereby providing a comprehensive framework for microscopic reservoir appraisal. Physics-informed neural networks address the limitations of purely data-driven reservoir simulation by embedding governing seepage equations into their loss functions, thereby ensuring physical consistency and improved generalization. Multimodal architectures significantly enhance unconventional shale gas production prediction by integrating geological heterogeneity with dynamic production behavior, leading to more accurate and stable forecasts. Collectively, these AI-driven approaches underscore the importance of combining domain expertise, multi-source data, and physics-aware modeling to achieve efficient and intelligent oil and gas development.

1. Introduction

Amid the ongoing global energy transition and intelligent development, artificial intelligence (AI) technology is integrating deeply into various aspects of oil and gas exploration and development at an unprecedented pace. This integration is bringing revolutionary opportunities to enhance geological prediction accuracy, development efficiency, and operational decision-making. AI-driven approaches are increasingly employed to address complex subsurface challenges and optimize hydrocarbon recovery. This article highlights key advances in AI applications within the sector, such as data-driven sweet spot prediction for shale gas resources, intelligent rock image analysis using advanced image processing techniques, and data-physics hybrid modeling for smart reservoir simulation. Furthermore, the application of multimodal AI for production forecasting and management in unconventional shale gas reservoirs is discussed. Each technological innovation not only enhances interpretation accuracy and reduces human bias but also promotes sustainable extraction practices. These developments signify a shift toward digitalized and intelligent oilfield operations, forming a new paradigm for future energy systems.

Yandy Scientific Press

*Corresponding author.
E-mail address: hui.gang@cup.edu.cn (G. Hui); renyili@petrochina.com.cn (Y. Ren); bjf2024@petrochina.com.cn (J. Bi); muming.wang@ucalgary.ca (M. Wang); liucq@student.cup.edu.cn (C. Liu).

Received July 13, 2025; revised July 29, 2025; accepted August 6, 2025; available online August 7, 2025.

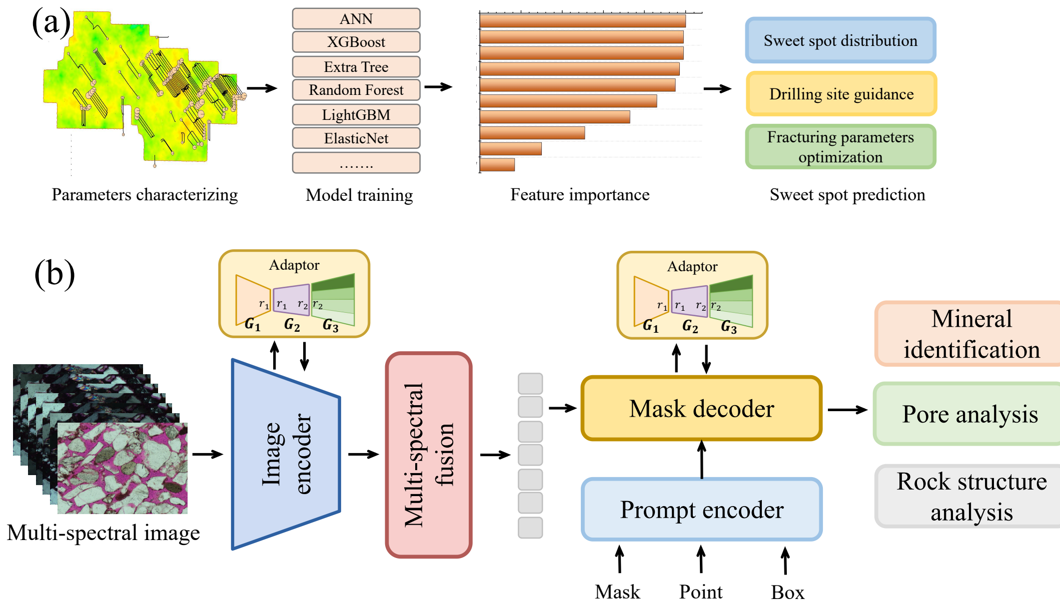

Fig. 1. (a) Schematic of multi-disciplinary geological-engineering data-driven shale sweet spot prediction (Hui et al., 2021) and (b) schematic of intelligent rock image analysis technology based on FalconCore intelligent image processing (Ren et al., 2025).

2. Applications

2.1 Sweet spot prediction

Multiple factors govern the productivity of shale reservoir horizontal wells, where large-scale datasets often display complex and non-linear patterns (Mustafa et al., 2022). Conventional analytical methods are prone to subjectivity and inherent limitations, frequently resulting in ambiguous identification of dominant controlling factors and unreliable productivity forecasts (Liu et al., 2021). In contrast, machine learning (ML) algorithms serve as powerful tools for extracting insights from high-dimensional data, enabling the detection of non-obvious relationships among variables and facilitating the identification of critical parameters that define shale gas sweet spots (Hui et al., 2023).

As an illustrative case, the intelligent productivity prediction study for the Duvernay Shale gas reservoir in the Western Canadian Sedimentary Basin employed an integrated workflow comprising: comprehensive characterization of geological and engineering parameters, application of multiple ML methods to train input-output relationships, comparative evaluation of predictive models under varying input combinations, and optimization of key parameters for production forecasting (Fig. 1(a)). The workflow begins with a systematic evaluation of key reservoir properties to establish a robust data foundation. Multiple ML methods, including artificial neural network (ANN), XGBoost, etc., are then employed to model input-output relationships, with comparative performance analysis guiding the selection of the optimal predictive algorithm. Subsequent parameter contribution analysis refines input variables to enhance model efficiency and interpretability. It culminates in three operational applications: a sweet spot prediction module leveraging critical geological parameters, a drilling location guidance module for optimized horizontal well placement, and a fracturing parameter optimization module that employs inverse modeling to maximize production (Hui et al., 2021). This structured approach offers a reliable data-driven foundation for intelligent and efficient shale gas development.

2.2 Intelligent reservoir evaluation

Core analysis is a fundamental technique for microscopic reservoir evaluation, typically involving visual inspection, image analysis, and laboratory testing, among which image analysis has found the widest application (Chen et al., 2025). Recent advances in AI, including computer vision, deep learning, and large foundation models, have significantly advanced the capabilities of intelligent core image interpretation (Liu et al., 2024; Qu et al., 2024). Building on this progress, the FalconCore intelligent analysis technology has been developed, utilizing multispectral thin-section images and supported by the CoreSAM segmentation foundation model (Fig. 1(b)), providing a versatile and high-precision framework for microscopic reservoir characterization.

The system performs three primary functions: it first enables automated mineral identification and quantitative analysis across diverse lithologies and regions, leveraging the generalization capability of CoreSAM. Next, it facilitates precise extraction of pore geometric parameters (e.g., such as aspect ratio and tortuosity) as well as characterization of their spatial distribution. Finally, through the integration of multispectral thin-section images, it reconstructs spatial relationships among grains, matrix, and pores to reveal key microstructural features

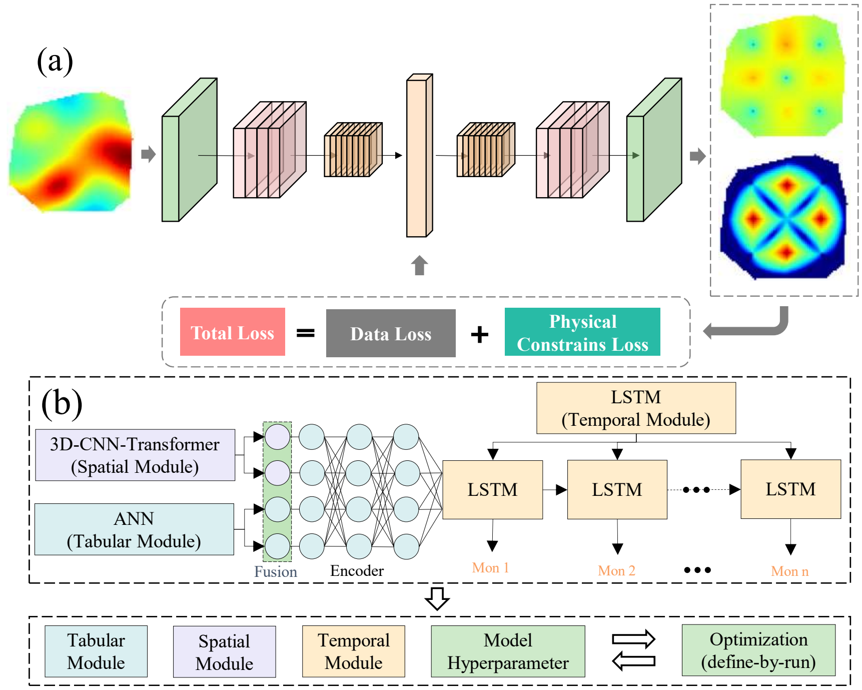

Fig. 2. (a) Schematic of PINNs architecture and loss function composition (Bi et al., 2024) and (b) schematic of multimodal-based shale gas production prediction (Wang et al., 2025a).

(Ren et al., 2025).

This integrated approach establishes a triaxial evaluation framework covering mineral composition, pore system properties, and microstructure configuration. The methodology has been extended to applications such as automated cuttings classification and AI-assisted fossil identification, thereby providing a robust microscopic data foundation for enhanced hydrocarbon recovery and development precision.

2.3 Intelligent reservoir simulation

In recent years, deep learning-based surrogate models for reservoir simulation have gained increasing research attention and practical adoption (Kim et al., 2025). However, purely data-driven surrogate models often fail to incorporate oil and gas seepage theory, making them highly data-dependent and unable to consistently ensure physical fidelity in their predictions (Nande and Patwardhan , 2024; Zhang et al., 2025). This limitation stems largely from the inherent low-frequency bias of neural networks, which hinders accurate representation of fine-scale phenomena such as pressure funnel effects and waterflood front propagation.

Physics-informed neural networks (PINNs) offer a dual-driven approach that integrates governing differential equations of reservoir seepage into the neural network's loss function (Raissi et al., 2019). This allows the model to conform simultaneously to observed data and underlying physical principles during training. As shown in (Fig. 2(a)), the network architecture adopts a multi-level feature extraction mechanism, sequentially performing local feature capture through convolutional layers, feature dimensionality reduction via pooling layers, and global information integration with fully connected layers, ultimately outputting prediction results that comply with physical laws. Notably, by constructing a multi-objective loss function containing both data fitting terms and physical constraint terms, the model simultaneously achieves observational data feature learning and strict adherence to physical conservation laws during parameter optimization, this embodies the principle of collaborative optimization between deep learning and mechanistic models. Consequently, PINNs not only enhance the physical plausibility of predictions but also improve generalization performance beyond the training data (Bi et al., 2024), addressing key weaknesses of purely data-driven surrogates.

Furthermore, the hybrid PINN framework supports broader engineering applications. When combined with Monte Carlo algorithms, it enables quantification of uncertainty in production forecasts; when integrated with data assimilation techniques, it facilitates automated history matching and reservoir parameter inversion; and when coupled with heuristic optimization methods, it enhances the efficiency of development strategy optimization.

2.4 Multimodal production prediction

Multimodal AI architectures have shown superior accuracy and stability in predicting production from unconventional shale gas reservoirs compared to traditional single-source models, which often fail to adequately capture geological heterogeneity and dynamic production behavior (Mohammadi et al., 2025; Wang et al., 2025a). By integrating diverse data modalities, such as structured engineering parameters, 2D/3D geological property maps, and production time-series data, the proposed model achieves a more comprehensive feature representation and significantly improves the identification of underlying production mechanisms (Wang et al., 2025b).

Throughout the model development process, domain-specific engineering knowledge was incorporated to guide the representation of different modalities and to select appropriate processing algorithms. This integration enhances the physical consistency and interpretability of the model, increasing its reliability and practical value in real-world reservoir development applications (Wang et al., 2025a) (Fig. 2(b)). The architecture comprises three core modules: 3D-convolutional neural network (CNN)-transformer, as spatial module, processes geological spatial features, ANN (tabular module) handles structured parameters, with their outputs merging at the "Fusion" node before entering the central "Encoder" for feature extraction. Long short-term memory as temporal module, captures dynamic temporal patterns across sequential time steps. During development, domain-specific engineering knowledge was incorporated to guide modality representation and module selection, reinforcing the model's physical consistency and interpretability, thereby boosting its credibility and practical utility in reservoir development scenarios.

These findings underscore that future advances in AI applications within petroleum engineering will rely on the deep integration of engineering expertise with multi-source data. Such an approach is essential for developing intelligent systems that support operational decision-making and enable efficient interpretation and precise development of complex subsurface resources.

3. Challenges

Currently, data-driven ML methods are increasingly integrating multi-source data, such as seismic, logging, core, and dynamic production data, to enhance the accuracy of reservoir characterization and prediction. In parallel, these methods are evolving from purely data-driven approaches toward multi-modal and physics-constrained ML frameworks. A key future direction lies the comprehensive incorporation of multi-scale core images and experimental core analysis data to achieve holistic, precise, and quantitative reservoir evaluation in terms of mineral composition, pore systems, and microstructural attributes. The adoption of multimodal large models is expected to enable integrated interpretation of core, log, and seismic data, thereby extending the application of core-based insights to reservoir assessment, flow simulation, hydraulic fracturing design, and related engineering tasks. More importantly, multimodal large-model technology based on transformer architectures is breaking scale barriers between core, logging, and seismic data through cross-modal attention mechanisms. In this way, microscopic core features can directly constrain macroscopic physical inversion processes, allowing pore–permeability parameters obtained at the laboratory scale to inform reservoir numerical simulations and hydraulic fracturing optimization. While current challenges are often broadly acknowledged, a deeper analysis of their root causes is essential. The computational difficulty of embedding complex multiphase flow and seepage equations into neural networks arises not only from high dimensionality but also from the inherently stiff and nonlinear nature of the governing partial differential equations. These equations demand highly stable numerical schemes that often conflict with the gradient-based optimization used in ML. Similarly, the challenge in balancing data-fit and physics-based losses arises from a deeper methodological gap: the absence of a principled framework to reconcile statistical deviations with mechanistic constraints, especially under sparse or noisy data conditions. This frequently leads to adversarial learning dynamics where improving physical consistency may compromise predictive accuracy, and vice versa. Moreover, in highly heterogeneous reservoirs, the intrinsic multiscale nature of flow and transport phenomena further complicates the faithful encoding of physical laws into scalable loss functions.

Looking ahead, multimodal ML offers promising pathways for modeling complex subsurface systems by fusing diverse data types. A key challenge, however, lies in the effective and systematic incorporation of domain knowledge into both model architecture and validation processes to ensure physical plausibility and practical relevance across diverse field applications.

4. Conclusions

The integration of AI technologies into oil and gas exploration and development is reshaping the industry by addressing key challenges and unlocking new opportunities. Advancements ranging from data-driven sweet-spot prediction, intelligent reservoir evaluation, and intelligent reservoir simulation to multimodal production forecasting demonstrate AI's ability to handle complex, multi-source data and improve decision-making efficiency. These developments collectively highlight that the future of AI in petroleum engineering lies in the deep integration of domain expertise and diverse data sources, which ultimately supports more efficient and sustainable oil and gas resource development within the context of the global energy transition.

Acknowledgements

This research was supported by State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation (No. CDUT-PLC2024006) and China National Petroleum Corporation Technology Project (No. 2023DJ84).

Conflict of interest

The authors declare no competing interest.

References

Bi, J., Li, J., Wu, K., et al. A physics-informed spatial-temporal neural network for reservoir simulation and uncertainty quantification. SPE Journal, 2024, 29(4): 2026-2043.

Chen, G., Jiao, J., Wang, Z., et al. Multi-fidelity machine learning with knowledge transfer enhances geothermal energy system design and optimization. Advances in Geo-Energy Research, 2025, 16(3): 244-259.

Hui, G., Chen, S., He, Y., et al. Machine learning-based production forecast for shale gas in unconventional reservoirs via integration of geological and operational factors. Journal of Natural Gas Science and Engineering, 2021, 94: 104045.

Hui, G., Chen, Z., Wang, Y., et al. An integrated machine learning-based approach to identifying controlling factors of unconventional shale productivity. Energy, 2023, 266: 126512.

Kim, S., Kim, T., Jo, S. Artificial intelligence in geoenergy: Bridging petroleum engineering and future-oriented applications. Journal of Petroleum Exploration and Production Technology, 2025, 15(2): 35.

Liu, H., Ren, Y., Li, X., et al. Research status and application of artificial intelligence large models in the oil and gas industry. Petroleum Exploration and Development, 2024, 51(4): 1049-1065.

Liu, Y. , Ma, X., Zhang, X., et al. A deep-learning-based prediction method of the estimated ultimate recovery (EUR) of shale gas wells. Petroleum Science, 2021,18(5): 1450-1464.

Mohammadi, M., Tajik, E., Martinez-Maldonado, R., et al. Artificial intelligence in multimodal learning analytics: A systematic literature review. Computers and Education: Artificial Intelligence, 2025, 8: 100426.

Mustafa, A., Tariq, Z., Mahmoud, M., et al. Data-driven machine learning approach to predict mineralogy of organic-rich shales: An example from Qusaiba Shale, Rub' al Khali Basin, Saudi Arabia. Marine and Petroleum Geology, 2022, 137: 105495.

Nande, S. B., Patwardhan, S. D. Automated reservoir characterization of carbonate rocks using deep learning image segmentation approach. SPE Journal, 2024, 29(8): 4356-4375.

Qu, F., Liao, H., Lu, M., et al. Recognition of drill string vibration state based on WGAN-div and CNN-IWPSO-SVM. Geoenergy Science and Engineering, 2024, 243: 213342.

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. Journal of Computational Physics, 2019, 378: 686-707.

Ren, Y., Zeng, C., Li, X., et al. Intelligent evaluation of sandstone rock structure based on a visual large model. Petroleum Exploration and Development, 2025, 52(2): 548-558.

Wang, M., Wang, H., Hui, G., et al. A hybrid tabular-spatial-temporal model with 3D geomodel for production prediction in shale gas formations. SPE Journal, 2025a, 30(6): 3281-3293.

Wang, M., Zhang, X., Wang, H., et al. Integrating geological model via A multimodal machine learning approach in shale gas production forecast. Gas Science and Engineering, 2025b, 139: 205617.

Zhang, D., Li, H. Efficient prediction of subsurface CO_2 plume migration in heterogeneous reservoirs using the TransUNet deep learning network. Geoenergy Science and Engineering, 2025, 255: 214106.