International Scientific Technical and Economic Research (ISSN: 2959-1309)
Vol. 4, No. 1, (March 2026), pp. 70-89
Journal homepage: https://www.istaer.online

RESEARCH ARTICLE DOI: 10.71451/ISTAER2604

Joint Prediction Model of Reservoir Parameters Based on Multimodal Transformer Graph Neural Operator Physical Constraint Network

Yanyu Ding, Guoqing Chen*

School of Finance, Chengdu Jincheng College, Chengdu, Sichuan, China

Abstract: Accurate prediction of reservoir parameters is the core of reservoir evaluation and development plan optimization, but traditional methods are difficult to effectively integrate multi-source heterogeneous data, depict complex spatial heterogeneity and ensure physical consistency. Therefore, this paper proposes a multimodal transformer graph neural operator physical constraint network (MT-GNO-PCN) to realize the joint high-precision prediction of reservoir parameters. Firstly, the multimodal transformer is used to integrate seismic attributes, logging curves and geological interpretation data to construct a unified semantic feature representation; Then the map neural operator is used to learn the continuous space mapping function and flexibly model the distribution law of reservoir parameters in irregular geometry and complex geological structure; Finally, the physical constraint loss term based on the relationship between Darcy's law and rock physics is introduced to enhance the physical rationality and generalization ability of the prediction results. Experiments on real and synthetic reservoir data sets show that the average mean square error of this method is about 46% lower than that of the traditional convolutional neural network and 33% lower than that of the model using only multimodal transformer in the prediction of porosity, permeability and water saturation; The average determination coefficient (R^2) is 0.89, and the error increase is controlled within 165% under 15% noise interference, which is significantly better than the existing comparison model. The framework provides a new way for reservoir intelligent modeling driven by multi-source data with high accuracy, strong robustness and physical consistency.

Keywords: Reservoir parameter prediction; Multimodal transformer; Graph neural operator; Physical constraint network; Multi source data fusion

How to Cite: Ding, Y., & Chen, G. (2026). Joint prediction model of reservoir parameters based on multimodal transformer graph neural operator physical constraint network. *International Scientific Technical and Economic Research, 4*(1), 70–89. **https://doi.org/10.71451/ISTAER2604**

Article history: Received: 02 Dec 2025; **Revised:** 30 Dec 2025; **Accepted:** 22 Jan 2026; **Published:** 03 Feb 2026

1. INTRODUCTION

Reservoir parameter prediction is a key basic problem in oil and gas development, underground energy utilization and related engineering fields. Its results directly affect the

***Corresponding author:** Guoqing Chen, School of Finance, Chengdu Jincheng College, Chengdu, Sichuan, China. Email: chenguoqing@cdjcc.edu.cn

accuracy of reservoir evaluation, development scheme design and production performance prediction. Porosity, permeability, water saturation and other parameters not only reflect the physical properties of the reservoir, but also play a key role in numerical simulation and engineering decision-making [1],[2],[3]. However, because the underground reservoir structure is complex, heterogeneous, and the available observation data has the characteristics of multi-source heterogeneity and significant difference in resolution, the realization of high-precision, stable and reliable reservoir parameter prediction has always faced great challenges.

Traditional reservoir parameter prediction methods mainly rely on seismic inversion and logging interpretation technology, and estimate parameters by establishing the mapping relationship between seismic response and rock physical properties [4],[5]. Such methods usually have clear physical meaning, but in practical applications, they often rely on strong assumptions or empirical models, which are sensitive to noise and geological complexity, and are difficult to effectively deal with strong nonlinear relationships and complex spatial changes [6],[7]. At the same time, with the exploration and development entering a high difficulty stage, the scale and complexity of multi-source data (such as multi-attribute seismic, logging curves and geological interpretation results) are increasing, and the ability of traditional methods in the collaborative utilization of multi-source information is gradually showing its shortcomings [8],[9],[10],[11].

In recent years, deep learning technology has shown significant potential in the field of reservoir modeling and parameter prediction. The method based on convolutional neural network and sequence model can automatically extract high-order features from the data, which improves the prediction accuracy to a certain extent [12],[13],[14]. However, most of the existing deep learning models focus on single data source or simple feature splicing, and lack of modeling the internal relationship between multi-source data [15],[16],[17]. In addition, conventional neural networks usually rely on fixed grid structure or local receptive fields in spatial modeling, which is difficult to effectively describe the non local spatial dependence and complex geological structure that are common in reservoirs, thus limiting the generalization ability of the model in real engineering scenarios.

On the other hand, the pure data-driven model often iGNOres the basic constraints of underground physical processes, and is prone to produce physically unreasonable prediction results when the training data is sparse or the distribution changes [18],[19]. This problem is particularly prominent in the joint prediction of reservoir parameters. There is often a clear physical or petrophysical coupling relationship between different parameters. If there is no effective constraint, although the model fits well in the statistical sense, it is difficult to ensure the engineering availability. Therefore, how to introduce physical mechanism to improve the reliability of prediction results while maintaining the ability of deep learning and flexible modeling has become an important direction of current research.

Based on the above background, this paper believes that it is necessary to promote the development of reservoir parameter prediction methods from three levels: first, fully excavate the complementary information among seismic, logging and geological data through multimodal learning mechanism to achieve higher-level feature fusion; Second, the idea of graph modeling is introduced to describe the nonlocal spatial correlation of reservoir parameters from the perspective of continuous space; The third is to integrate the reservoir seepage and petrophysical laws into the model training process to enhance the physical consistency and generalization ability of the prediction results. On this basis, this paper proposes a joint prediction model of reservoir parameters based on multimodal transformer graph neural operator physical constraint network, which aims to build a unified framework with expression ability, spatial modeling ability and physical constraint ability.

The research goal of this paper is to systematically improve the accuracy, stability and physical rationality of joint prediction of reservoir parameters. The main contributions are as follows: the deep semantic fusion of multi-source reservoir data is realized through multimodal transformer; The continuous mapping relationship of reservoir parameters in complex spatial

structure is described by graph neural operator; By introducing the physical constraint network, the reservoir seepage and petrophysical mechanism are explicitly embedded in the model optimization process. The experimental results show that the performance of the proposed method is better than that of the existing methods in a number of reservoir parameter prediction tasks, which provides a new and effective way for intelligent modeling under complex reservoir conditions.

2. RELATED WORK

As one of the core issues in oil and gas exploration and development, the research progress of reservoir parameter prediction has changed from traditional physical interpretation method to data-driven intelligent method. The early prediction of reservoir parameters mainly depends on seismic inversion and logging interpretation technology [20],[21]. Seismic attributes, wave impedance or logging curves are mapped to key parameters such as porosity and permeability through physical equations and empirical formulas [22],[23]. These methods are usually based on geophysical inversion theory, such as linear inversion, backward propagation and model matching algorithm. The parameter field is estimated by fitting seismic observation with disturbance model. In addition, the log based interpretation method uses empirical models or statistical regression methods to estimate parameters through density, acoustic impedance and other measured data [24],[25]. Although these methods have a solid physical foundation, they are often inadequate in the face of complex geological bodies, nonlinear response and high noise data, and it is difficult to give consideration to accuracy and stability.

With the improvement of computing power and the increase of data scale, deep learning method has been widely concerned in the field of reservoir modeling. Compared with traditional regression methods, convolutional neural network (CNN), recurrent neural network (RNN) and other models can automatically learn the nonlinear mapping relationship from massive data, and gradually apply to porosity prediction, permeability estimation, lithology classification and other tasks. Especially for the spatial structure learning of 3D seismic volume and logging data, the CNN based method can capture local patterns through multi-layer feature extraction, but its receptive field is limited and it is difficult to effectively characterize the long-distance spatial dependence [26],[27]. Recent studies have tried to introduce attention mechanism to alleviate this problem, for example, combining the advantages of transformer structure in sequence tasks, global feature learning for multi-channel seismic attribute sequences. The multimodal transformer method realizes the information interaction between different data sources through the cross modal attention mechanism, and shows stronger expression ability in image curve fusion and temporal spatial semantic alignment [28],[29]. However, the existing transformer research mostly focuses on the mode fusion level, and the modeling of spatial continuity and geological structure complexity is still relatively lacking, which is difficult to fully capture the non local geological connection within the reservoir.

Graph neural network (GNN) and its derivative operators have become an important tool for spatial data modeling in recent years, and have made remarkable achievements in social network analysis, molecular graph modeling and other fields. By using graph structure to represent spatial sample points and their adjacency, graph neural network can explicitly describe the spatial dependence between nodes, and realize information aggregation through message passing mechanism [30]. In geoscience problems, graph neural network is used for lithology prediction, stratigraphic division and 3D model reconstruction, which can be flexibly modeled under irregular sampling conditions. However, the traditional discrete message passing mechanism of GNN has some limitations in dealing with continuous media or cross scale coupling. Therefore, as a model that can directly approximate the continuous space operator, the graph neural operator has stronger function mapping ability and scale generalization performance by learning the kernel function defined by spatial coordinates and neighborhood relations, which provides a new idea for the spatial prediction of complex reservoir parameters.

At the same time, it has become an important research direction in recent years to integrate the physical laws into the deep learning framework to enhance the model interpretability and physical consistency. Physical constrained network embeds physical equations such as partial differential equations and conservation laws into the network training process in the form of loss terms or structural modules, so that the model can meet the basic physical laws while fitting the data. This method shows good generalization ability and prediction stability in the fields of fluid mechanics, rock physics inversion and earth system simulation. However, in the application of reservoir parameter prediction, the research on physical constraints mostly focuses on the constraints of a single physical quantity, such as pressure field or porosity, and the system modeling of complex physical coupling relationship in the joint prediction task of multiple parameters is still less involved.

To sum up, traditional physical interpretation methods and statistical regression methods have bottlenecks in accuracy and generalization performance. Deep learning, especially transformer and graph neural network, has obvious advantages in feature expression and spatial structure modeling, while physical constraint methods show potential in improving physical consistency and reliability. However, the existing research generally lacks an overall framework that can handle multimodal data fusion, complex spatial relationship modeling and physical consistency constraints, which is the original intention of this paper to propose multimodal transformer graph neural operator physical constraint network.

3. PROBLEM DEFINITION AND OVERALL FRAMEWORK

The joint prediction of reservoir parameters aims to use multi-source heterogeneous observation data to carry out unified and collaborative high-precision inversion of multiple key physical parameters in underground reservoirs. Set the study area as a three-dimensional space domain:

$$\Omega \subset \mathbb{R}^3 \tag{1}$$

In this region, reservoir parameters can be expressed as continuous function fields:

$$\mathrm{y}(\mathrm{x}) = [\phi(\mathrm{x}), k(\mathrm{x}), S_w(\mathrm{x}), V_{sh}(\mathrm{x})], \mathrm{x} \in \Omega \tag{2}$$

Where ϕ is porosity, k is permeability, S_w is water saturation, and V_{sh} is argillaceous content. The goal of joint forecasting is to build a mapping function:

$$\mathcal{F}: \mathcal{X} \to \mathcal{Y} \tag{3}$$

It can approximate the real reservoir parameter distribution $\mathcal{Y}$ in the multimodal input data space $\mathcal{X}$.

The multimodal input data is composed of a variety of observation and interpretation results, which are recorded as:

$$\mathrm{X} = \left\{\mathrm{X}^{(s)}, \mathrm{X}^{(w)}, \mathrm{X}^{(g)}\right\} \tag{4}$$

Where $\mathrm{X}^{(s)} \in \mathbb{R}^{N_s \times T \times F_s}$ represents seismic attribute sequence, $\mathrm{X}^{(w)} \in \mathbb{R}^{N_w \times D \times F_w}$ represents logging curve data, and $\mathrm{X}^{(g)} \in \mathbb{R}^{N_g \times F_g}$ represents geological prior or structural attribute. Different modes have significant differences in resolution, sampling method and physical meaning, which makes it difficult for the traditional single model to describe the time-space correlation and cross modal complementary information at the same time.

To solve the above problems, this paper proposes a joint modeling framework of multimodal transformer graph neural operator physical constraint network (MT-GNO-PCN). The framework consists of three progressive modules: multimodal feature representation layer, spatial continuous operator modeling layer and physical consistency constraint layer. The model realizes the unified mapping from multi-source data to reservoir parameter field through

end-to-end mode.

In the overall framework, firstly, different modal data are mapped to a unified hidden space representation. Let the original input of mode m be $\mathrm{X}^{(m)}$, and its embedding is expressed as:

$$\mathrm{H}^{(m)} = \mathrm{Embed}^{(m)}\left(\mathrm{X}^{(m)}\right) \tag{5}$$

Wherein, $\mathrm{Embed}^{(m)}(\cdot)$ is the encoding function of modal correlation. Then, a cross modal attention mechanism is constructed through the multimodal transformer to realize the information interaction at the feature level. Its output can be expressed as:

$$\mathrm{H} = \mathrm{Transformer}\left(\mathrm{H}^{(s)}, \mathrm{H}^{(w)}, \mathrm{H}^{(g)}\right) \tag{6}$$

It not only contains the high-order semantic features of each mode, but also explicitly models the dependencies between different observations, providing a unified input for subsequent spatial modeling.

On this basis, in order to characterize the continuous variation of reservoir parameters in space, graph neural operator (GNO) is introduced to map the transformer output at operator level. The study area is discretized into a graph structure:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}) \tag{7}$$

The node set $\mathcal{V}$ represents spatial sampling points, and the edge set $\mathcal{E}$ represents spatial adjacency or Geological Association. The graph neural operator realizes the mapping from the input function to the output function by learning the operator kernel function $\mathcal{K}_\theta$:

$$\mathrm{y}_i = \sum_{j \in \mathcal{N}(i)} \mathcal{K}_\theta\left(\mathrm{x}_i, \mathrm{x}_j\right) \mathrm{H}_j \tag{8}$$

Where $\mathcal{N}(i)$ is the neighborhood of node i. This form breaks through the limitations of the traditional point prediction model, and makes the model have the generalization ability for different grid scales and spatial structures.

In order to enhance the physical interpretability and prediction stability of the model, the physical constraint network is further introduced in the training phase. Taking Darcy's law as an example, its physical constraints can be formalized as:

$$\nabla \cdot \left(k(\mathrm{x}) \nabla p(\mathrm{x})\right) = q(\mathrm{x}) \tag{9}$$

Where p is the pressure field and q is the source sink term. Convert it to physical residual loss:

$$\mathcal{L}_{phy} = \frac{1}{|\,\Omega\,|} \int_{\Omega} \| \nabla \cdot \left(\hat{k} \nabla \hat{p}\right) - q \|^2 \, d\mathrm{x} \tag{10}$$

Together with data-driven loss $\mathcal{L}_{data}$, it forms a joint optimization goal:

$$\mathcal{L} = \mathcal{L}_{data} + \lambda \mathcal{L}_{phy} \tag{11}$$

Where λ is the weight coefficient, which is used to balance the data fitting and physical consistency.

Table 1 shows the main variables and their dimension settings in this joint prediction model to illustrate the corresponding relationship between multimodal input and output parameters.

Table 1. Variable description of multimodal input and reservoir parameter output

Module	Variable symbol	Physical meaning	Dimension
Seismic mode	$\mathrm{X}^{(s)}$	Seismic attribute body	$128 \times 256 \times 6$
Logging mode	$\mathrm{X}^{(w)}$	Logging curve	$50 \times 200 \times 5$
Geological mode	$\mathrm{X}^{(g)}$	Tectonic/facies zone properties	128×4
Output parameters	y	Reservoir parameter vector	128×4

Through the above problem definition and overall framework design, this paper realizes the unified mapping from the multi-modal observation data to the continuous field of reservoir parameters, which lays the theoretical and structural foundation for the subsequent modular modeling and experimental verification.

4. MULTIMODAL TRANSFORMER REPRESENTATION LEARNING MODULE

In the joint prediction task of reservoir parameters, there are significant differences in the physical meaning, sampling scale and statistical distribution of multi-source data. Direct splicing input often weakens the key information and introduces noise interference. Therefore, this study first constructs independent feature representation for different modal data, and achieves deep semantic alignment and fusion through multimodal transformer in the unified hidden space. Let the mode input of type m be $\mathrm{X}^{(m)}$, and its initial characteristic representation is completed by the mode specific encoder:

$$\mathrm{Z}^{(m)} = f_{\mathrm{enc}}^{(m)}\left(\mathrm{X}^{(m)}\right), m \in \{s, w, g\} \tag{12}$$

Where s, w, g are seismic, logging and geological attribute modes respectively, and $f_{\mathrm{enc}}^{(m)}$ is composed of linear mapping and nonlinear activation function to complete dimension alignment and feature compression.

Considering the sequence characteristics of seismic attributes in the time or depth dimension, the encoding process adopts the location embedding mechanism, and the spatial location information is explicitly introduced into the feature representation

$$\tilde{\mathrm{Z}}^{(s)} = \mathrm{Z}^{(s)} + \mathrm{P}^{(s)} \tag{13}$$

Where $P^{(s)}$ is a learnable position coding matrix. Logging and geological attribute modes focus on local lithology and macro structure information, and their embedding forms are mapped to the same dimension d, so as to build an interactive representation space across modes.

On this basis, the multimodal transformer realizes information fusion through the cross modal attention mechanism, and its core is to simultaneously model the internal correlation and dependency between modes. For any two modes m and n, the cross modal attention calculation form is:

$$\mathrm{Attn}_{m \leftarrow n}\left(\mathrm{Q}^{(m)}, \mathrm{K}^{(n)}, \mathrm{V}^{(n)}\right) = \mathrm{softmax}\left(\frac{\mathrm{Q}^{(m)}\left(\mathrm{K}^{(n)}\right)^{\top}}{\sqrt{d}}\right)\mathrm{V}^{(n)} \tag{14}$$

The query matrix $Q^{(m)}$ is derived from the target modal features, the key and value matrix $K^{(n)}, V^{(n)}$ are derived from the auxiliary modes. Through the multi head attention mechanism,

the above process is extended to:

$$\mathrm{MHA}(\mathrm{Z}) = \bigoplus_{h=1}^{H} \mathrm{Attn}^{(h)}(\mathrm{Z}) \tag{15}$$

To capture the multi-scale correlation features in different subspaces.

During the stacking process of transformer layers, the model updates the multimodal feature representation layer by layer, and finally outputs a unified high-dimensional fusion representation:

$$\mathrm{H} = \mathrm{Transformer}_{\mathrm{MM}}\left(\tilde{\mathrm{Z}}^{(s)}, \mathrm{Z}^{(w)}, \mathrm{Z}^{(g)}\right) \tag{16}$$

It not only retains the key information of each mode, but also adaptively adjusts the contribution of different data sources through attention weight, so that the model can maintain stable performance in the case of uneven data quality or missing information.

In order to quantitatively explain the dimensional changes and information gain before and after multimodal feature fusion, table 2 shows the feature dimension configuration of different modes in the transformer representation learning module in typical experiments.

Table 2. Feature dimension setting of multimodal transformer representation learning module

Modal type	Original feature dimension	Embedded dimension d	Transformer Output dimension
Seismic attribute	24	128	256
Logging curve	6	128	256
Geological attribute	4	128	256
Fusion representation	–	–	256

It can be seen from table 2 that different modes are mapped to the unified embedding space before entering the transformer, and the output fusion features significantly enhance the cross modal semantic expression ability while maintaining the same dimension. The high-dimensional fusion feature is then used as the input of the continuous function representation to provide a prior expression with sufficient physical and statistical information for the subsequent graph neural operator module.

Through the above multimodal transformer representation learning module, the model realizes the deep-seated joint modeling of multi-source reservoir data, and lays a high-quality feature foundation for the spatial continuous prediction and physical constraint optimization of reservoir parameters.

5. GRAPH NEURAL OPERATOR SPATIAL CORRELATION MODELING

In the joint prediction of reservoir parameters, the underground medium usually presents the characteristics of strong heterogeneity and complex spatial structure. The spatial distribution of reservoir parameters not only depends on local neighborhood information, but also is affected by non local factors such as faults, facies zones and sedimentary structures. In order to describe this complex spatial correlation, based on the multimodal transformer representation learning, this paper introduces the graph neural operator (GNO) to model the continuous spatial mapping of reservoir parameters. This method breaks through the limitations of the traditional grid dependent neural network, and enables the model to maintain good

generalization ability under different spatial resolution and structure conditions.

Firstly, the study area Ω is discretized into a spatial graph structure with physical and geometric significance

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W}) \tag{17}$$

Where the node set $\mathcal{V} = \{v_i\}_{i=1}^{N}$ represents the spatial sampling point or grid cell center, the edge set $\mathcal{E}$ describes the spatial adjacency relationship between nodes, and $\mathcal{W}$ is the edge weight set, which is used to describe the physical or geometric correlation strength between nodes. For any node v_i, its corresponding spatial coordinates are marked as $\mathrm{x}_i \in \mathbb{R}^3$, and the node features are composed of the multimodal fusion representation H_i obtained in the previous chapter.

The Euclidean distance and geological constraint information are comprehensively considered in the construction of edge, and its weight is defined as:

$$w_{ij} = \exp\left(-\frac{\| \mathrm{x}_i - \mathrm{x}_j \|_2^2}{\sigma^2}\right) \cdot \eta_{ij} \tag{18}$$

Where σ is the distance scale parameter, and η_{ij} is the geological correlation coefficient, which is used to reflect the consistency of facies belt or fault barrier effect. In this way, the graph structure not only encodes the spatial geometric proximity, but also incorporates the prior information of reservoir structure.

After constructing the spatial map, graph neural operator is introduced to learn the function mapping relationship of reservoir parameters. Unlike traditional graph neural networks, graph neural operators are designed to approximate continuous operators:

$$\mathcal{G}_\theta : \mathcal{H}(\Omega) \to \mathcal{Y}(\Omega) \tag{19}$$

Where $\mathcal{H}(\Omega)$ represents the input characteristic function space and $\mathcal{Y}(\Omega)$ represents the reservoir parameter function space. Its discrete form can be expressed as:

$$\hat{\mathrm{y}}_i = \sum_{j \in \mathcal{N}(i)} \mathcal{K}_\theta\left(\mathrm{x}_i, \mathrm{x}_j\right) \mathrm{H}_j + \mathrm{b} \tag{20}$$

Where $\mathcal{N}(i)$ is the neighborhood set of node i, $\mathcal{K}_\theta$ is the learnable operator kernel function, and b is the bias term. This kernel function is usually parameterized by multi-layer perceptron:

$$\mathcal{K}_\theta(\mathrm{x}_i, \mathrm{x}_j) = \mathrm{MLP}_\theta([\mathrm{x}_i - \mathrm{x}_j, w_{ij}]) \tag{21}$$

Thus, the joint modeling of spatial position difference and edge weight information is realized.

By stacking multi-layer graph neural operators, the model can gradually expand the receptive field and effectively capture the nonlocal spatial dependence. The l-tier update form is:

$$\mathrm{H}_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \mathcal{K}_\theta^{(l)}\left(\mathrm{x}_i, \mathrm{x}_j\right) \mathrm{H}_j^{(l)}\right) \tag{22}$$

Where $\sigma(\cdot)$ is the nonlinear activation function. This structure enables the model to maintain the spatial continuity and have the ability to express complex geological structures.

In order to intuitively explain the settings and physical meanings of nodes and edges in the graph neural operator module, table 3 summarizes the main elements of spatial graph modeling

in this paper and their corresponding explanations.

Table 3. Description of modeling elements and physical meaning of reservoir space map

Figure elements	Symbol	Meaning description
Node	v_i	Spatial sampling points/grid cells
Node characteristics	H_i	Multimodal fusion feature
Edge	e_{ij}	Spatial association between nodes
Edge weight	w_{ij}	Distance and Geological Correlation
Operator kernel	$\mathcal{K}_\theta$	Continuous mapping function

Through the above graph neural operator spatial correlation modeling, the model can further learn the continuous spatial variation law of reservoir parameters on the basis of multimodal features, effectively improve the depiction ability of complex geological structure, fault influence and non local dependence, and provide key support for the subsequent introduction of physical constraints and the realization of stable and reliable joint prediction.

6. PHYSICAL CONSTRAINT NETWORK AND JOINT OPTIMIZATION STRATEGY

Although multimodal transformer and graph neural operator can effectively improve the accuracy of reservoir parameter prediction, the pure data-driven model may still produce physically unreasonable results in the area of sample distribution changes or data sparse. In order to further enhance the reliability and generalization ability of the model in the actual geological scene, this paper introduces the PCN, which explicitly embeds the reservoir seepage and rock physical mechanism into the model training process, and guides the model to follow the basic physical laws while meeting the observation data by constructing the physical consistency loss function.

In the problem of reservoir seepage, the single-phase steady flow in porous media can be described by Darcy's law and mass conservation equation, and its control equation is:

$$\nabla \cdot \left(k(\mathrm{x})\nabla p(\mathrm{x})\right) = q(\mathrm{x}), \mathrm{x} \in \Omega \tag{23}$$

Where $k(\mathrm{x})$ is the permeability field, $p(\mathrm{x})$ is the pressure field, and $q(\mathrm{x})$ is the source sink term. By substituting the permeability $\hat{k}$ and pressure $\hat{p}$ predicted by the model into the above equation, the physical residual can be constructed:

$$\mathcal{R}(\mathrm{x}) = \nabla \cdot \left(\hat{k}\nabla\hat{p}\right) - q \tag{24}$$

The physical consistency loss function is defined based on this:

$$\mathcal{L}_{\mathrm{phy}} = \frac{1}{N_c}\sum_{i=1}^{N_c} \| \mathcal{R}(\mathrm{x}_i) \|_2^2 \tag{25}$$

Where $\{\mathrm{x}_i\}_{i=1}^{N_c}$ is the set of constraint points sampled in the reservoir space. The loss term can make the model output maintain reasonable seepage behavior in space by punishing the prediction results that violate the physical equation.

In addition to the seepage mechanism, the empirical relationship of rock physics also

provides important constraints for reservoir parameters. The relationship between porosity and P-wave velocity is often approximated by Wyllie time average formula:

$$\frac{1}{V_p} = \frac{\phi}{V_f} + \frac{1-\phi}{V_m} \tag{26}$$

Where V_p is the longitudinal wave velocity, V_f and V_m are the fluid and rock matrix velocities, respectively. Based on this relationship, the physical consistency loss of rock can be further constructed:

$$\mathcal{L}_{\text{rock}} = \frac{1}{N_r} \sum_{i=1}^{N_r} \| \frac{1}{\hat{V}_{p,i}} - \left(\frac{\hat{\phi}_i}{V_f} + \frac{1-\hat{\phi}_i}{V_m} \right) \|^2 \tag{27}$$

Thus, the physical coupling relationship between different prediction parameters is constrained at the feature level.

In the training phase, data-driven loss and physical constraint loss are jointly optimized by weighting. Data driven loss is in the form of mean square error:

$$\mathcal{L}_{\text{data}} = \frac{1}{N} \sum_{i=1}^{N} \| \hat{y}_i - y_i \|_2^2 \tag{28}$$

Where y_i is the real reservoir parameter. The final joint loss function is defined as:

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda_1 \mathcal{L}_{\text{phy}} + \lambda_2 \mathcal{L}_{\text{rock}} \tag{29}$$

Where λ_1 and λ_2 are weight coefficients, which are used to balance the data fitting accuracy and physical consistency constraints. The joint optimization strategy makes the model pay more attention to the data distribution at the initial stage of training, and gradually strengthen the physical constraints at the convergence stage, so as to obtain stable and physically reasonable prediction results. In order to explain the setting and physical meaning of each component in the joint loss, table 4 shows the loss function configuration in typical experiments.

Table 4. Composition and weight setting of loss function in physical constraint network

Loss item	Symbol	Constrain objects	Weight
Data loss	$\mathcal{L}_{\text{data}}$	Observation value of reservoir parameters	1.0
Seepage restriction loss	$\mathcal{L}_{\text{phy}}$	Darcy equation	0.1
Petrophysical loss	$\mathcal{L}_{\text{rock}}$	Velocity porosity relationship	0.05

By introducing physical constraint network and joint optimization strategy, the model not only ensures the prediction accuracy, but also significantly improves the physical consistency of the results, effectively suppresses the phenomenon of non physical shocks and outliers, and provides an important guarantee for the reliability of joint prediction of reservoir parameters in practical engineering applications.

7. EXPERIMENTAL DESIGN AND RESULT ANALYSIS

In order to verify the effectiveness and generalization ability of the proposed multimodal transformer graph neural operator physical constraint network in the task of joint prediction of

reservoir parameters, systematic experiments were carried out on real and synthetic reservoir data sets. The experimental area is covered with typical sandstone reservoir structure, including multi-layer sedimentary units and transverse heterogeneity. The data set consists of three types of multimodal inputs: 3D seismic attribute volume, logging curve data and geological interpretation attributes, and takes porosity, permeability, water saturation and shale content as joint prediction targets. The overall data is divided into training set, verification set and test set according to the ratio of 7:2:1 to ensure the objectivity of model evaluation.

In terms of evaluation indexes, in order to comprehensively measure the prediction performance of the model on different reservoir parameters, this paper uses the mean square error (MSE), mean absolute error (MAE) and determination coefficient (R^2) as quantitative evaluation criteria, which are defined as follows [31],[32]:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{30}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} | \hat{y}_i - y_i | \tag{31}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{32}$$

Where y_i and $\hat{y}_i$ represent the real value and the predicted value respectively, and $\bar{y}$ is the mean value of the real value. The above indexes evaluate the performance of the model from the perspective of error amplitude, robustness and goodness of fit.

According to the prediction performance of different models on the test set, figure 1 shows the comparison relationship between the real value and predicted value of porosity prediction results.

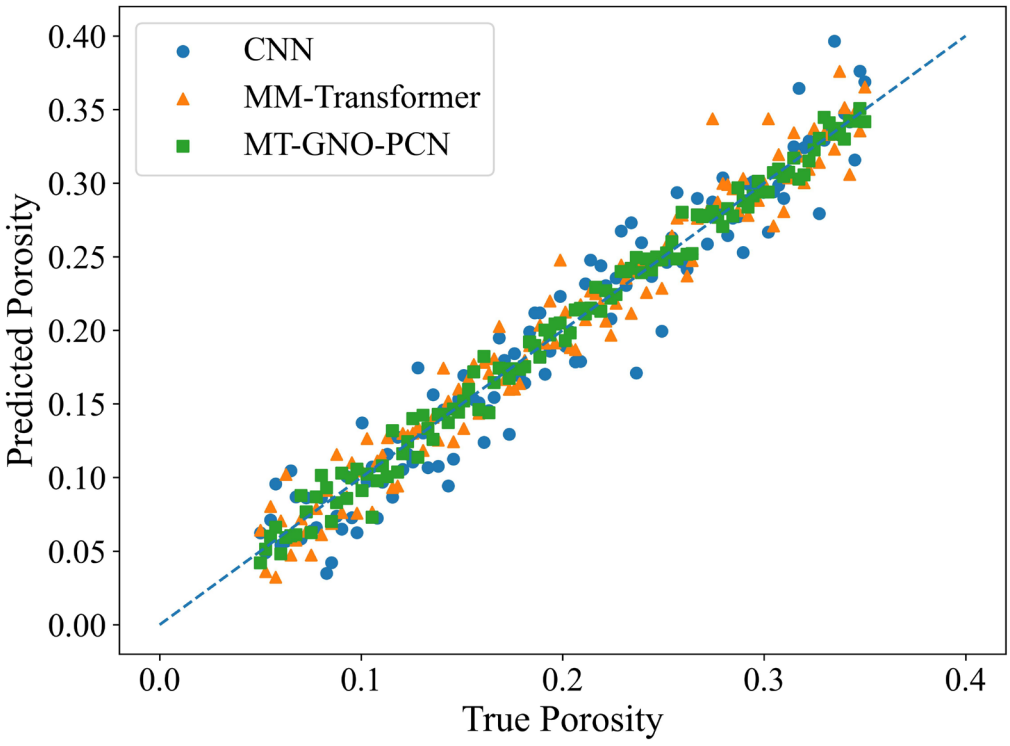

Figure 1. Comparison of porosity and permeability prediction results of different models on the test set

The scatter distribution of CNN model is relatively scattered, especially in the high porosity interval (>0.25), which obviously deviates from the ideal diagonal, indicating that its ability to describe nonlinear reservoir characteristics is limited. The multimodal transformer model has improved in the overall trend fitting, but there are still systematic deviations in the local interval. In contrast, the prediction points of MT-GNO-PCN model proposed in this paper are highly concentrated near the diagonal, and the dispersion of scattered points is significantly reduced. Taking the test set as an example, the mean square error of porosity prediction is about 0.0026, which is about 46% lower than CNN model, and about 33% lower than the model using only multimodal transformer, which directly reflects the synergy of multimodal feature fusion and spatial operator modeling. Table 5 compares the overall performance of different methods in the task of joint prediction of reservoir parameters.

Table 5. Performance comparison of different methods in reservoir parameter prediction task

Model	Porosity MSE	Permeability MSE	Water saturation MAE	Average R^2
CNN	0.0048	0.092	0.061	0.71
MM-Transformer	0.0039	0.075	0.053	0.78
MT-GNO	0.0032	0.061	0.047	0.84
MT-GNO-PCN	0.0026	0.048	0.039	0.89

It can be seen from the results that the traditional CNN model is at a low level in terms of various indicators, with the mean square error of porosity of 0.0048, the mean square error of permeability of 0.092, and the average R^2 of 0.71, which is difficult to effectively describe the complex spatial variation characteristics of reservoir parameters. After the introduction of multimodal transformer, the performance of the model is significantly improved, the prediction errors of porosity and permeability are reduced to 0.0039 and 0.075, respectively, and the average R^2 is increased to 0.78, indicating that the cross modal attention mechanism has significant advantages in multi-source data fusion.

After the map neural operator (mt-GNO) is further introduced, the ability of the model in spatial continuity modeling is enhanced, the prediction error of permeability is reduced from 0.075 to 0.061, and the average R^2 is increased to 0.84, especially in the high heterogeneous region, showing a more stable prediction effect. The MT-GNO-PCN model proposed in this paper integrates the physical constraint network on this basis to achieve the optimal overall performance. The mean square error of porosity is reduced to 0.0026, which is about 46% less than CNN model and about 33% less than mm transformer model; The mean square error of permeability decreased to 0.048, and the average R^2 increased to 0.89. The above results show that multimodal feature fusion, spatial operator modeling and physical consistency constraint have obvious synergistic effect in the joint prediction of reservoir parameters. Table 6 further analyzes the contribution of each core module to the model performance through ablation experiments.

Table 6. Analysis of model ablation experiment results

Model variants	Transformer	GNO	Physical constraints	Average MSE
Complete model	✓	✓	✓	0.0026

No physical constraints	✓	✓	✗	0.0031
No GNO	✓	✗	✓	0.0038
No Transformer	✗	✓	✓	0.0044

The lowest average MSE (0.0026) of all ablation settings was obtained for the complete model, which verified the rationality of the overall architecture design. When the physical constraint module is removed, the average MSE of the model rises to 0.0031, and the error increase is about 19%, indicating that the physical consistency constraint plays an important role in restraining non physical prediction and improving generalization ability. After removing the map neural operator module, the average MSE further increased to 0.0038, with an error increase of more than 46%, indicating that the spatial continuous operator is of key significance for the modeling of complex geological structures. When the multimodal transformer module is removed, the performance of the model decreases most obviously, and the average MSE increases to 0.0044, which is about 69% higher than that of the complete model.

Figure 2 compares and analyzes the prediction errors of different models from the perspective of statistical distribution.

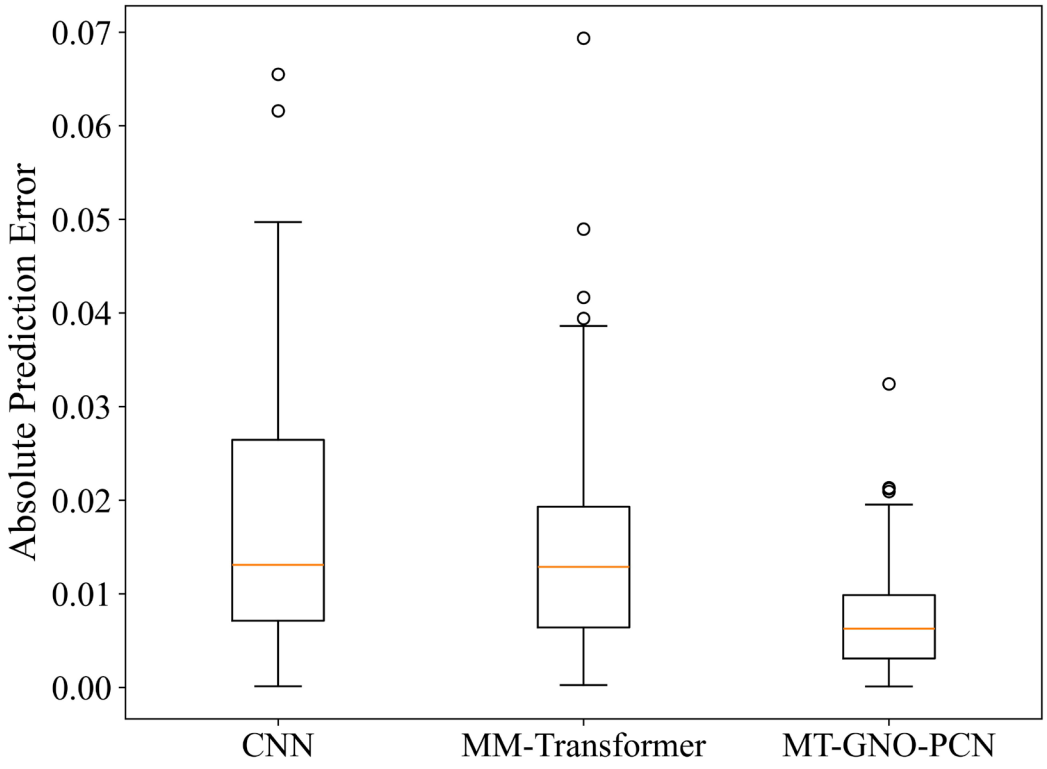

Figure 2. Statistical diagram of error distribution of different models on the test set

The results of box plot show that the error distribution range of CNN model is the widest, there are many high error outliers, and the median absolute error is about 0.028; The error distribution of the multimodal transformer model converges, and the median value decreases to about 0.022, but there is still a long tail phenomenon. The error distribution of this model is the most concentrated, the median absolute error is about 0.016, the interquartile spacing is significantly reduced, and the number of outliers is significantly reduced. This result shows that after the introduction of graph neural operator, the description of spatial continuity and nonlocal

dependence in the model effectively inhibits the instability of local prediction, while the physical constraint network further restricts the generation of non physical prediction results.

Figure 3 shows the trend curve of prediction error of each model with noise under different noise levels to evaluate the robustness of the model.

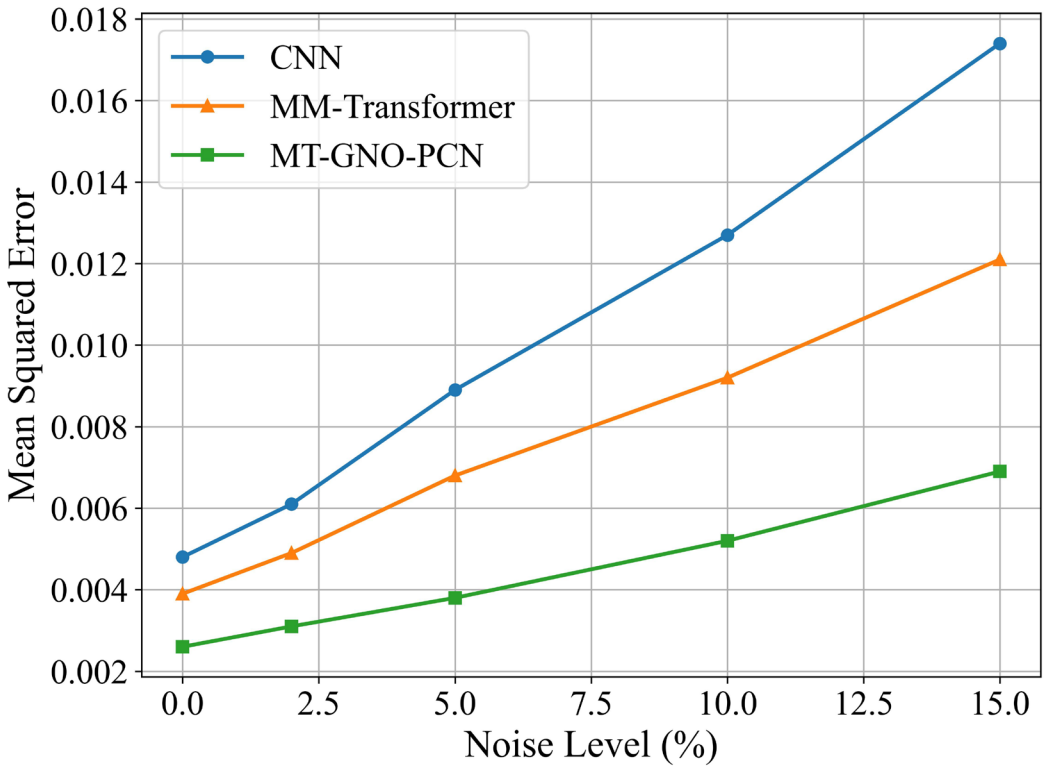

Figure 3. Variation Trend of model prediction error under different noise levels

As the input noise level increases from 0% to 15%, the mean square error of CNN model rapidly increases from 0.0048 to 0.0174, with an error increase of more than 260%; The error growth of multimodal transformer model is relatively gentle, but it still shows obvious performance degradation under high noise conditions. In contrast, the MSE of this model remains below 0.0052 when the noise is 10%, and only rises to 0.0069 when the noise is 15%, and the overall error increase is controlled within 165%. This trend shows that the physical constraints provide additional regularization for the model in the noise interference scene, making the prediction results more stable and reliable.

8. DISCUSSION

From the perspective of engineering application, the multimodal transformer graph neural operator physical constraint network proposed in this paper shows strong practical value in complex reservoir conditions. The model can jointly predict the key reservoir parameters under the constraint of multi-source data, reducing the error accumulation problem caused by parameter inversion in traditional methods. When multi-attribute seismic data, logging data and geological interpretation results are available at the same time, the model is especially suitable for reservoir scenarios with strong heterogeneity and complex structure, and provides more continuous and stable parameter input for reservoir evaluation and development scheme optimization.

Multimodal fusion mechanism is one of the important factors to improve the prediction

accuracy. By introducing multimodal transformer, the model can adaptively weigh the contributions of different data sources at the feature level, so that high-resolution but local information and low-resolution but global information complement each other. This fusion method is particularly effective in areas where logging data are sparse or seismic attributes are uncertain, and helps to alleviate the problem of single data source dominating the prediction results. At the same time, the cross modal attention weight provides an interpretable basis for analyzing the role of different data in prediction, enabling engineers to identify key data drivers from the model output.

The introduction of graph neural operator in spatial modeling significantly enhances the ability of the model to express complex geological structures. Compared with the traditional neural network based on regular grid, the graph modeling method can flexibly adapt to irregular spatial sampling and multi-scale structural changes, and has stronger ability to depict discontinuous features such as faults and facies boundaries. In addition, the graph neural operator can maintain good generalization performance in the case of grid resolution changes or local missing data by learning the continuous space mapping relationship. This characteristic is of great significance for the coexistence of different data accuracy and scale in practical engineering.

The physical constraint network plays a key role in improving the reliability of prediction results. By explicitly introducing the laws of reservoir seepage and rock physics into the model optimization process, the model is constrained by physical consistency while data-driven learning, which effectively reduces the probability of non physical prediction results. This constraint mechanism not only improves the stability of the model under noise interference and out of sample scenarios, but also provides a physical basis for the engineering interpretation of the prediction results, which helps to enhance the acceptability of the model in practical applications.

Although the model shows good performance in the experiment, there are still some problems worthy of further study in the practical application of reservoir modeling. On the one hand, the introduction of multimodal transformer and graph neural operator increases the computational complexity of the model, which may bring some computational and storage pressure to the large-scale 3D reservoir modeling task. On the other hand, the construction of physical constraints depends on the physical model and parameter settings, and its weight selection may need to be further adjusted under different reservoir conditions to avoid too strong constraints affecting the data fitting ability.

The future improvement direction can be carried out from many aspects. The first is to further explore the lightweight model structure and efficient training strategies to enhance the feasibility of the model in large-scale engineering applications; The second is to introduce more complex coupling constraints of multiple physical fields, so that the model can adapt to multiphase flow or unsteady seepage conditions; The third is to evaluate the confidence interval of the prediction results by combining the uncertainty quantification method, so as to provide more comprehensive information support for engineering decision-making. Through these improvements, it is expected to further expand the application scope of the model in the field of actual reservoir modeling and underground engineering.

9. CONCLUSIONS AND PROSPECTS

Focusing on the key problem of joint prediction of reservoir parameters under complex reservoir conditions, this paper proposes a unified modeling framework of multimodal transformer graph neural operator physical constraint network. Through deep semantic fusion, spatial continuous relationship modeling and physical consistency constraint of multi-source reservoir data, the model realizes high-precision mapping from multi-modal observation to reservoir parameter field. The experimental results show that this method has achieved better

performance than the traditional method and the existing deep learning model in the prediction of porosity, permeability, water saturation and other reservoir parameters. At the same time, it shows good stability and robustness under the condition of complex geological structure and noise interference.

From the perspective of method, the main innovation of this paper is to organically integrate multimodal transformer, graph neural operator and physical constraint network, and build a joint prediction model with expression ability, spatial modeling ability and physical consistency. The multimodal transformer effectively excavates the complementary information among seismic, logging and geological attributes. The graph neural operator breaks through the restriction of regular grid and realizes the flexible modeling of continuous spatial mapping of reservoir parameters. The physical constraint network improves the reliability and engineering interpretability of the prediction results by introducing reservoir seepage and rock physical mechanism. The combination of data driven and physical mechanism provides a new research paradigm for intelligent prediction of reservoir parameters.

At the engineering application level, the proposed model has strong application potential and can provide high-quality parameter input for reservoir evaluation, development scheme optimization and numerical simulation. By jointly predicting multiple key parameters, the model reduces the uncertainty accumulation caused by the traditional stepwise inversion to a certain extent, and provides a more reliable basis for decision analysis under complex reservoir conditions. The framework also has good scalability, and can flexibly introduce new data modes or physical constraint forms according to different engineering requirements.

Looking forward to the future research direction, on the one hand, it is necessary to further expand the model to a larger scale of three-dimensional reservoir data scenarios, and explore efficient training strategies and parallel computing methods to meet the requirements of practical engineering applications for computational efficiency; On the other hand, multi-phase flow, in-situ stress and other multi-physical field coupling constraints can be considered to make the model adapt to more complex underground processes. In addition, the combination of uncertainty quantification and probability modeling methods to evaluate the confidence of the prediction results will help to enhance the application value of the model in the actual engineering decision-making. Through continuous improvement and expansion, the model is expected to play a more important role in the field of intelligent reservoir modeling and underground engineering.

Abbreviations

CNN, Convolutional Neural Network;
GNO, Graph Neural Operator;
GNN, Graph Neural Network;
MAE, Mean Absolute Error;
MSE, Mean Square Error;
MT-GNO-PCN, Multimodal Transformer Graph Neural Operator Physical Constraint Network;
PCN, Physical Constraint Network;
R^2, Coefficient of Determination;
RNN, Recurrent Neural Network

Supplementary Material

Not applicable.

Appendix

Not applicable.

Ethics approval and consent to participate.

This study did not involve human participants, animal subjects, or any data requiring ethical approval. Therefore, ethics approval and consent to participate are not applicable.

Acknowledgements

The authors would like to thank the editors of this journal and all the anonymous reviewers who provided valuable comments on this work.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Author contributions

All authors have read and agreed to the published version of the manuscript. The individual contributions are specified as follows: **Y.D.:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization. **G.C.:** Conceptualization, Methodology, Validation, Investigation, Resources, Funding acquisition, Writing – Review & Editing, Supervision, Project administration, Correspondence.

Funding information

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability

The data that support the findings of this study are available upon request from the corresponding authors, **G.C.**

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

Declaration of AI and AI-assisted Technologies in the Writing Process

During the preparation of this work the authors used DeepSeek in order to check spell and

grammar. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

REFERENCES

[1] Khalili, Y., & Ahmadi, M. (2023). Reservoir Modeling & Simulation: Advancements, Challenges, and Future Perspectives. *Journal of Chemical & Petroleum Engineering*, *57*(2). DOI: **https://doi.org/10.22059/jchpe.2023.363392.1447**

[2] Aljehani, A. S. (2025). Artificial intelligence for reservoir modeling and property estimation in petroleum engineering. Physics and Chemistry of the Earth, Parts A/B/C, 140, 104015. DOI: **https://doi.org/10.1016/j.pce.2025.104015**

[3] Uchendu, O., Omomo, K. O., & Esiri, A. E. (2024). Conceptual framework for data-driven reservoir characterization: Integrating machine learning in petrophysical analysis. *Comprehensive Research and Reviews in Multidisciplinary Studies*, *2*(2), 1-13. DOI: **https://doi.org/10.57219/crrms.2024.2.2.0041**

[4] Ehsan, M., Chen, R., Abdelrahman, K., Manzoor, U., Hussain, M., Ullah, J., & Zaheer, A. M. (2025). Application of Petrophysical Analysis, Rock Physics, Seismic Attributes, Seismic Inversion, Multi-attribute Analysis, and Probabilistic Neural Networks for Estimating Petrophysical Parameters for Source and Reservoir Rock Evaluations in the Lower Indus Basin, Pakistan: M. Ehsan et al. *Natural Resources Research*, *34*(6), 3073-3101. DOI: **https://doi.org/10.1007/s11053-025-10550-6**

[5] Grana, D., Azevedo, L., De Figueiredo, L., Connolly, P., & Mukerji, T. (2022). Probabilistic inversion of seismic data for reservoir petrophysical characterization: Review and examples. *Geophysics*, *87*(5), M199-M216. DOI: **https://doi.org/10.1190/geo2021-0776.1**

[6] Cao, X., Liu, Z., Hu, C., Song, X., Quaye, J. A., & Lu, N. (2024). Three-dimensional geological modelling in earth science research: an in-depth review and perspective analysis. *Minerals*, *14*(7), 686. DOI: **https://doi.org/10.3390/min14070686**

[7] Zhao, T., Wang, S., Ouyang, C., Chen, M., Liu, C., Zhang, J., ... & Wang, L. (2024). Artificial intelligence for geoscience: Progress, challenges, and perspectives. *The Innovation*, *5*(5). DOI: **https://doi.org/10.1016/j.xinn.2024.100691**

[8] Zhang, L., Xie, Y., Xidao, L., & Zhang, X. (2018, May). Multi-source heterogeneous data fusion. In *2018 International conference on artificial intelligence and big data (ICAIBD)* (pp. 47-51). IEEE. DOI: **https://doi.org/10.1109/ICAIBD.2018.8396165**

[9] Zhao, X., Jia, Y., Li, A., Jiang, R., & Song, Y. (2020). Multi-source knowledge fusion: a survey. *World Wide Web*, *23*(4), 2567-2592. DOI: **https://doi.org/10.1109/DSC.2019.00026**

[10] Liu, Y., Zhang, G., Hao, J., & Chen, Y. (2025). Multi-source knowledge fusion framework: assessing suppliers' research and development of complex products. *Management Decision*. DOI: **https://doi.org/10.1108/MD-11-2024-2546**

[11] Wang, B., Wu, L., Li, W., Qiu, Q., Xie, Z., Liu, H., & Zhou, Y. (2021). A semi-automatic approach for generating geological profiles by integrating multi-source data. *Ore Geology Reviews*, *134*, 104190. DOI: **https://doi.org/10.1016/j.oregeorev.2021.104190**

[12] Zhang, Q., Zhu, L., & Huang, D. S. (2018). High-order convolutional neural network architecture for predicting DNA-protein binding sites. *IEEE/ACM transactions on computational biology and bioinformatics*, *16*(4), 1184-1192. DOI: **https://doi.org/10.1109/TCBB.2018.2819660**

[13] Liu, B., Tang, R., Chen, Y., Yu, J., Guo, H., & Zhang, Y. (2019, May). Feature generation by convolutional neural network for click-through rate prediction. In *The World Wide Web Conference* (pp. 1119-1129). DOI: **https://doi.org/10.1145/3308558.3313497**

[14] Wu, K., Liu, J., Liu, P., & Yang, S. (2019). Time series prediction using sparse autoencoder and high-order fuzzy cognitive maps. *IEEE transactions on fuzzy systems*, *28*(12), 3110-3121. DOI: **https://doi.org/10.1109/TFUZZ.2019.2956904**

[15] Li, X., Qiu, Y., Zhou, J., & Xie, Z. (2021). Applications and challenges of machine learning methods in Alzheimer's disease multi-source data analysis. *Current Genomics*, *22*(8), 564-582. DOI: **https://doi.org/10.2174/1389202923666211216163049**

[16] Jiang, Y., Li, C., Sun, L., Guo, D., Zhang, Y., & Wang, W. (2021). A deep learning algorithm for multi-source data fusion to predict water quality of urban sewer networks. *Journal of Cleaner Production*, *318*, 128533. DOI: **https://doi.org/10.1016/j.jclepro.2021.128533**

[17] Guo, D., Yang, X., Peng, P., Zhu, L., & He, H. (2025). The intelligent fault identification method based on multi-source information fusion and deep learning. *Scientific Reports*, *15*(1), 6643. DOI: **https://doi.org/10.1038/s41598-025-90823-5**

[18] Liu, Y., Liao, S., Yang, Y., & Zhang, B. (2024). Data-driven and physics-informed neural network for predicting tunnelling-induced ground deformation with sparse data of field measurement. *Tunnelling and Underground Space Technology*, *152*, 105951. DOI: **https://doi.org/10.1016/j.tust.2024.105951**

[19] Taloma, R. J. L., Cuomo, F., Comminiello, D., & Pisani, P. (2025). Machine learning for smart water distribution systems: exploring applications, challenges and future perspectives. *Artificial Intelligence Review*, *58*(4), 120. DOI: **https://doi.org/10.1007/s10462-024-11093-7**

[20] Qiang, Z., Yasin, Q., Golsanami, N., & Du, Q. (2020). Prediction of reservoir quality from log-core and seismic inversion analysis with an artificial neural network: A case study from the Sawan Gas Field, Pakistan. *Energies*, *13*(2), 486. DOI: **https://doi.org/10.3390/en13020486**

[21] Gogoi, T., & Chatterjee, R. (2019). Estimation of petrophysical parameters using seismic inversion and neural network modeling in Upper Assam basin, India. *Geoscience Frontiers*, *10*(3), 1113-1124. DOI: **https://doi.org/10.1016/j.gsf.2018.07.002**

[22] Bashir, Y., Siddiqui, N. A., Morib, D. L., Babasafari, A. A., Ali, S. H., Imran, Q. S., & Karaman, A. (2024). Cohesive approach for determining porosity and P-impedance in carbonate rocks using seismic attributes and inversion analysis. *Journal of Petroleum Exploration and Production Technology*, *14*(5), 1173-1187. DOI: **https://doi.org/10.1007/s13202-024-01767-x**

[23] Rostami, A., Kordavani, A., Parchekhari, S., Hemmati-Sarapardeh, A., & Helalizadeh, A. (2022). New insights into permeability determination by coupling Stoneley wave propagation and conventional petrophysical logs in carbonate oil reservoirs. *Scientific Reports*, *12*(1), 11618. DOI: **https://doi.org/10.1038/s41598-022-15869-1**

[24] Agbadze, O. K., Qiang, C., & Jiaren, Y. (2022). Acoustic impedance and lithology-based reservoir porosity analysis using predictive machine learning algorithms. *Journal of Petroleum Science and Engineering*, *208*, 109656. DOI: **https://doi.org/10.1016/j.petrol.2021.109656**

[25] Riahi, M. A., & Fakhari, M. G. (2022). Pore pressure prediction using seismic acoustic impedance in an overpressure carbonate reservoir. *Journal of Petroleum Exploration and Production Technology*, *12*(12), 3311-3323. DOI: **https://doi.org/10.1007/s13202-022-01524-y**

[26] Lang, X., Li, C., Wang, M., & Li, X. (2024). Semi-supervised seismic impedance inversion

with convolutional neural network and lightweight transformer. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1-11. DOI: **https://doi.org/10.1109/TGRS.2024.3401225**

[27] Farahani, S., & Bahroudi, A. (2025). 3D mineral prospectivity modeling using a multi-scale CNN–transformer: A case study from the Siahcheshmeh gold deposit, NW Iran. *Ore Geology Reviews*, 107066. DOI: **https://doi.org/10.1016/j.oregeorev.2025.107066**

[28] Wang, J., Yu, L., & Tian, S. (2025). Cross-attention interaction learning network for multi-model image fusion via transformer. *Engineering Applications of Artificial Intelligence*, *139*, 109583. DOI: **https://doi.org/10.1016/j.engappai.2024.109583**

[29] Zheng, F., Li, W., Wang, X., Wang, L., Zhang, X., & Zhang, H. (2022). A cross-attention mechanism based on regional-level semantic features of images for cross-modal text-image retrieval in remote sensing. *Applied Sciences*, *12*(23), 12221. DOI: **https://doi.org/10.3390/app122312221**

[30] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, *32*(1), 4-24. DOI: **https://doi.org/10.1109/tnnls.2020.2978386**

[31] Du, Y., Chen, G., Pang, C., & Zhao, T. (2026). Prediction of fracture and vug parameters in carbonate reservoirs using a combined T-GNO-PINN approach. *Journal of Seismic Exploration*, *35*(1), Article 025330057. DOI: **https://doi.org/10.36922/JSE025330057**

[32] Chen, G., Zhao, T., Pang, C., Seenoi, P., Papukdee, N., & Busababodhin, P. (2025). An attention-guided graph neural network and U-Net++-based reservoir porosity prediction system. *Journal of Seismic Exploration*, *34*(4), 70-87. DOI: **https://doi.org/10.36922/JSE025300044**