

Enhancing energy efficiency and sustainability in offshore drilling through real-time multi-objective optimization: Considering lag effects and formation variability

Yu Song ^{a,*}, Zehua Song ^a, Jin Yang ^b, Longgui Wei ^c, Jizhou Tang ^d

^a College of Artificial Intelligence, China University of Petroleum (Beijing), Beijing, 102249, China

^b College of Safety and Ocean Engineering, China University of Petroleum (Beijing), Beijing, 102249, China

^c Engineering Technology Branch, CNOOC Energy Development Co., Ltd, Tianjin, 300452, China

^d School of Ocean and Earth Science, Tongji University, Shanghai, 200092, China

ARTICLE INFO

Keywords:

AI-Empowered Optimization
Energy Efficiency in Industrial Drilling
Operational Reliability
Real-Time Energy Planning
Deep Reinforcement Learning

ABSTRACT

Advancements in offshore drilling demand AI-driven solutions to handle unpredictable geological conditions and complex subsurface environments, ensuring reliability and sustainability. Traditional methods relying on historical well data lack accuracy and adaptability, while logging-while-drilling (LWD) causes delays that hinder timely decision-making and reduce operational resilience. To address these challenges, this study presents an AI-empowered real-time multi-objective optimization framework designed to enhance the reliability and efficiency of offshore drilling systems. By integrating time-series forecasting networks, domain adversarial networks, and Markov decision processes, the framework accurately predicts the rate of penetration (ROP) and formation physical properties in real time despite limited data. Perceiving formation pressure gradients to constrain mud density adjustments and incorporating an online learning strategy enable adaptation to diverse geological environments, enhancing decision support. The double parameter optimization actor-critic (DPOAC) algorithm facilitates real-time adjustments, boosting operational efficiency and infrastructural reliability. Empirical analysis in the Caofeidian 6-4 block of the Bohai Sea, China, demonstrates significant improvements: ROP increased from 58.76 m/hr to 210.81 m hr, mechanical specific energy reduced from 11.42 MPa to 10.01 MPa, and unit cost per meter decreased from 18,439 CNY/m to 2,852 CNY/m. These results validate the framework's effectiveness in enhancing the sustainability and resilience of offshore drilling operations.

1. Introduction

In the global shift towards a transformed energy structure and low-carbon development strategies, offshore drilling technology is crucial for exploiting marine oil and gas resources, ensuring energy security, and achieving sustainable development [1–5]. However, offshore drilling operations encounter significant challenges due to unpredictable geological conditions, complex subsurface environments, and high economic costs [6–8]. These challenges necessitate the precise optimization of drilling parameters to enhance operational reliability, ensure safety, and minimize risks [9,10]. Integrating AI-driven solutions offers innovative and reliable strategies for optimizing drilling parameters, providing substantial theoretical and practical value for advancing marine drilling technologies and enhancing the resilience and

sustainability of offshore operations [11–13].

In the field of drilling parameter optimization, numerous studies have focused on methodologies to improve drilling efficiency while ensuring operational reliability and resilience. Kendall and Goins investigated drilling hydraulic parameters, establishing relationships among maximum impact force, hydraulic power, and jet velocity, thereby improving operational reliability [14]. Eckel examined the impact of drilling fluid density and viscosity on ROP, enhancing the sustainability of drilling operations [15]. Ramba et al. introduced a new metric, hydraulic drilling impact, to optimize drilling performance, enhancing reliability and enabling effective evaluation of operational parameters [16]. Zhang et al. demonstrated that reducing nozzle diameter and increasing jet velocity significantly enhanced drilling efficiency, supporting sustainable practices aligned with ExxonMobil's early findings

* Corresponding Author.

E-mail address: songyu_cup@163.com (Y. Song).

[17]. Bani Mustafa et al. employed response surface methodology for the optimization of controllable drilling parameters, enhancing drilling efficiency and operational performance [18]. Bajolvand et al. and Zang et al. applied an elitist non-dominated sorting genetic algorithm for multi-objective drilling parameter optimization, showcasing advancements in operational resilience [19,20]. Delavar et al. proposed a hybrid machine learning method combining multi-objective particle swarm optimization and random forest to optimize drilling parameters, significantly improving rate of penetration predictions and reducing potential hazards in complex subsurface environments [21]. Furthermore, Zheng et al. and Khaleel et al. implemented multi-objective cellular particle swarm optimization and genetic algorithms, respectively, to estimate coefficients for the Bourgoyn and Young model, significantly enhancing drilling efficiency, reducing operational costs, and ensuring more effective and resilient drilling operations [22,23].

Despite significant progress in drilling parameter optimization using intelligent computing methods, existing AI-driven models and multi-objective optimization algorithms present notable limitations in reliability and adaptability [24]. Neural network models, while capable of capturing complex nonlinear relationships, suffer from a black-box nature that hinders their interpretability and generalizability [25]. This lack of transparency is particularly problematic as most ROP prediction models are trained on data from previously drilled wells and perform poorly when applied to unexplored geological formations, limiting their reliability in dynamic environments [26]. Furthermore, reliance solely on logging-while-drilling (LWD) data introduces substantial delays, preventing timely decision-making and reducing operational resilience [27,28]. Simultaneously, past studies have often ignored the issues of borehole stability resulting from formation pressure gradient, leading to a somewhat arbitrary adjustment of the mud input density (MwIN). This lack of focus on formation pressure gradient and its impact on borehole stability may result in suboptimal MwIN under varying geological conditions. Such oversights can not only affect drilling efficiency but also increase the risks of borehole collapse and well kicks, ultimately impacting overall operational safety and economic viability. Importantly, multi-objective optimization in drilling is challenged by the inherent complexity and uncertainty of geological environments, resulting in dynamic and nonlinear parameter adjustments [29]. Traditional heuristic algorithms are often inefficient and inadequate for real-time parameter optimization, struggling to balance objectives such as enhancing ROP, controlling bit wear, and minimizing energy consumption, thereby compromising both reliability and sustainability.

Drilling perception, transfer, online learning, and reinforcement learning offer new solutions to the challenges of drilling parameter optimization. Drilling perception employs time-series prediction models that analyze both historical and current LWD parameters along with formation pressure gradient parameters to predict future drilling parameters, allowing for the pre-optimization of future drilling conditions [30,31]. This approach enhances the adaptability of models by introducing domain-adversarial networks and incremental parameter tuning, enabling models to quickly adjust by continually learning from new drilling data, thereby improving the accuracy of prediction models in the unexplored strata of target wells [32,33]. Reinforcement learning, a trial-and-error learning mechanism-based machine learning method, demonstrates tremendous potential for drilling-parameter optimization [34–36]. Through interactions with dynamic geological environments, reinforcement learning models are able to iteratively improve decision-making strategies by learning from real-time feedback, ensuring the safety and economic efficiency of drilling operations while effectively enhancing the ROP. Thus, combining drilling perception, transfer learning, online learning, and reinforcement learning establishes a more accurate, generalizable, and adaptive drilling parameter optimization framework, effectively addressing the complex challenges in offshore drilling with enhanced reliability, resilience, and sustainability. This approach uniquely resolves the trilemma of real-time adaptability, cross-well generalizability, and borehole stability

assurance-limitations persistently unaddressed in prior studies.

This study introduces a comprehensive multi-objective framework for real-time optimization of drilling parameters by integrating time-series forecasting networks, domain adversarial networks, and Markov decision processes (MDPs) to precisely predict ROP and optimize drilling parameters in real time. Initially, the mechanical properties of the rock were assessed through triaxial compression tests, and the formation pressure gradients were calculated using the Coulomb-Mohr criterion. These gradients serve as upper and lower bounds for the adjustment of MwIN, crucial for maintaining wellbore stability and operational safety. The MwIN range is defined to exceed the formation pore and collapse pressure gradients, thereby preventing the influx of formation fluids and ensuring wellbore stability against collapse, while remaining below the fracture and loss pressure gradients to mitigate fracture formation and mud loss. Drilling rate prediction and formation property awareness were addressed using a fully connected neural network and the iTransformer model, which detects changes in geological parameters using data from fixed time windows. Additionally, a domain adversarial mechanism was designed for strata not yet drilled into the target wells, which works in collaboration with an online learning strategy to adapt to new geological environments. Simultaneously, a Markov decision environment for drilling parameter optimization was established, and the double parameter optimization actor-critic (DPOAC) algorithm was developed, utilizing Q-, target, and policy networks to adjust drilling parameters and enhance economic and mechanical benefits, while boosting the real-time adaptability of the model to geological changes. This integrated framework facilitates automated and instantaneous adjustments of drilling parameters, significantly enhancing the economic and mechanical efficiency of drilling operations. The evaluation results from the Caofeidian 6-4 block in the Bohai Sea underscore the framework's substantial potential to improve drilling efficiency, resilience, and adaptability in complex geological environments, offering valuable technical insights and reliable strategies to support the sustainable development of future offshore drilling operations.

2. Workflow

Offline perception of ROP and formation properties in already-drilled neighboring wells: Rock mechanical properties at various depths were evaluated through triaxial compression tests. Applying the Coulomb-Mohr criterion, formation pressure gradients were accurately modeled to assess wellbore stability, ensuring operational reliability and safety. Two primary models, a fully connected neural network (FCNN) and an iTransformer, were developed for predicting ROP and formation properties. The FCNN utilized operational and geological parameters such as weight on bit (WOB), revolutions per minute (RPM), pump flow rate (FLW), mud input density (MwIN), gamma ray (GR), delta time (DT), porosity (POR), and rock density (RD) to forecast ROP, enhancing process control and reliability. The iTransformer leveraged fixed-time-window time-series data to detect short-term geological variations, extending its capability to accurately perceive complex formation pressure gradients. Incorporating adaptive learning rate adjustments and an early stopping strategy during model training and evaluation significantly improved the predictive performance and generalizability of the models across diverse geological settings. This approach enhances the reliability and safety of drilling operations by providing accurate real-time predictions and robust adaptation to varying subsurface conditions.

Online Perception and Transfer of ROP and Formation Properties in Undrilled Target Wells: Models optimized for neighboring wells were utilized to transfer and perceive ROP and formation properties in the unexplored strata of target wells. A domain adversarial mechanism, incorporating multilayer convolutional networks as feature extractors, was designed to train a binary domain discriminator, effectively distinguishing features from high-dimensional data. To align feature distributions between neighboring wells and target wells, the feature

extractor and domain discriminator engaged in a cooperative min-max training process. Model parameter freezing and fine-tuning techniques were employed to adapt to new geological environments, enhancing the model's transferability. Pre-trained on data from neighboring wells, the models underwent incremental unfreezing of key layers to precisely adapt to the specific geological conditions of target wells. An online learning strategy, combined with depth-triggered conditions and a forgetting mechanism, was implemented to ensure real-time model updates, thereby optimizing long-term learning efficiency and prediction accuracy. This strategy enhances the reliability and resilience of drilling operations by maintaining accurate and adaptable predictions in dynamic geological settings.

Design of a Markov Decision Process Environment for Real-Time Drilling Parameters: A MDP environment was developed to optimize drilling parameters, thereby enhancing the reliability and safety of drilling operations while improving economic and mechanical efficiency. This environment integrates a transfer prediction model to estimate the ROP in the unexplored strata of target wells, facilitating real-time assessment of cost per foot and mechanical specific energy across various drilling parameter combinations. The reward function is meticulously designed to evaluate the economic and efficiency impacts of drilling parameter adjustments, ensuring an optimal balance between cost management and equipment performance. The state and action spaces are structured to incorporate the transfer models for drilling parameter perception and formation pressure gradient perception, enabling continuous updates of drilling parameters and formation pressure gradients based on real-time data. This configuration provides critical dynamic inputs to the MDP environment, ensuring adherence to wellbore stability constraints. Consequently, the system optimizes drilling operations under dynamic geological conditions, reinforcing the safety and reliability of the drilling process.

Implementation and Deployment of Deep Reinforcement Learning Algorithms: In conjunction with an MDP environment, the DPOAC algorithm was designed to optimize the drilling parameters

using finely configured Q-, target, and policy networks. This algorithm enhances the economic benefits and mechanical efficiency of drilling operations while ensuring the automation and precision of the decision-making process. The DPOAC algorithm utilizes experience replay and priority-adjustment mechanisms to improve learning efficiency, and employs exploration noise strategies to enhance the ability of the model to adapt to changes in geological conditions in real time during actual drilling operations. These features collectively bolster the system's resilience against unexpected geological variations, thereby safeguarding the integrity of the drilling process. Additionally, this process was validated in the Caofidian 6-4 block in the Bohai Sea, China, where continuous adjustments of drilling parameters achieved an optimized balance between ROP and cost, thereby demonstrating the practical potential of this method to enhance the economic and operational efficiency of drilling. The workflow is illustrated in Fig. 1.

The following sections are organized as follows. Section 3 describes offline perception of ROP and formation properties using triaxial compression tests to determine rock mechanics, FCNN and iTransformer models to predict ROP and geological parameters, and formation pressure gradients to constrain mud density adjustments. Section 4 details transfer and online perception, where domain adversarial networks align feature distributions across wells, freezing-fine-tuning strategies adapt models to unexplored strata, and online learning enables incremental updates at 200 m intervals. Section 5 presents reinforcement learning optimization that employs a Markov decision process to define state and action spaces, utilizes the DPOAC algorithm for real-time drilling parameter optimization, and applies reward functions to balance mechanical specific energy and unit cost per foot.

3. Offline Perception of ROP and Formation Properties in Adjacent Wells

Conventional drilling parameter optimization methods inadequately address latency effects by relying primarily on real-time LWD

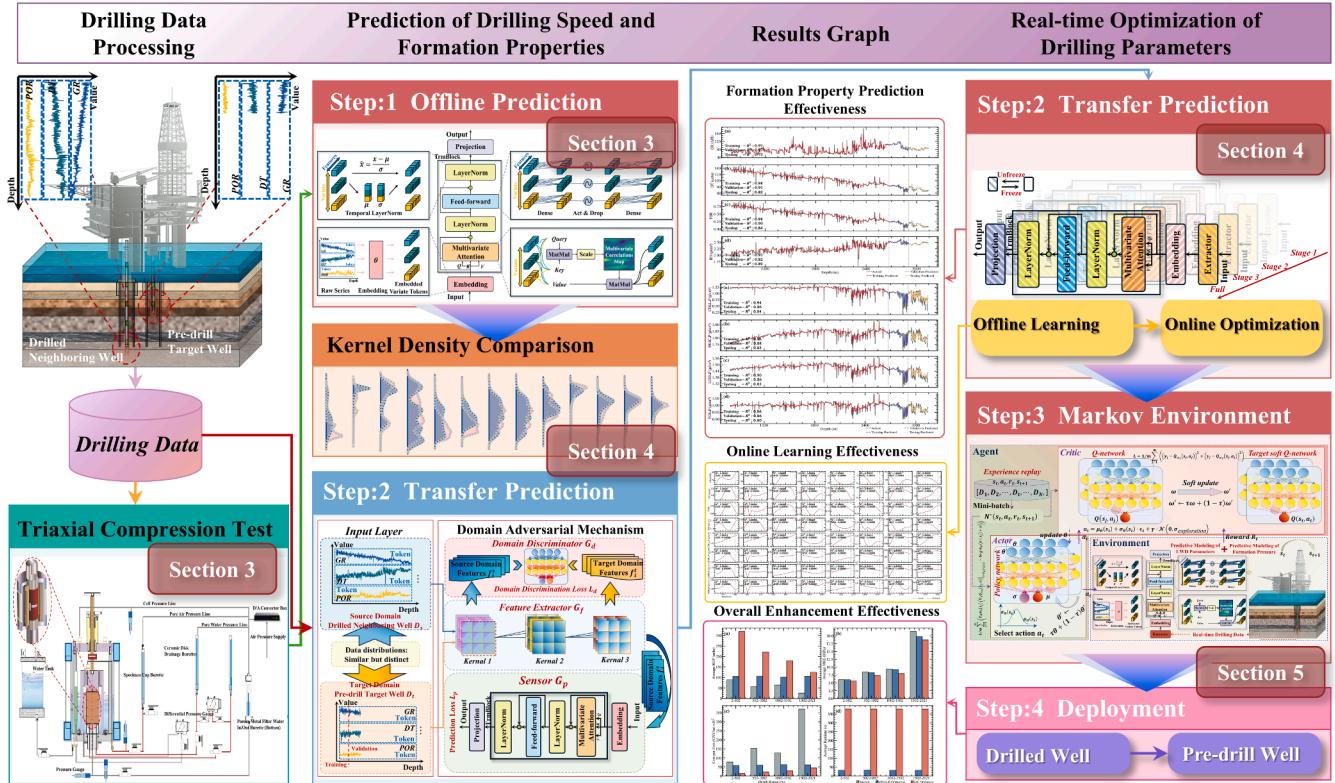


Fig. 1. Workflow.

parameters, causing delays and ineffective adjustments. To enhance operational reliability and safety, a perception-based method was developed to predict LWD parameters in advance, enabling proactive optimization. Utilizing pre- and mid-drilling data from the Caofidian 6-4 well in the Bohai Sea, China, this approach leverages direct observations essential for real-time decision-making and drilling-speed optimization. In contrast, post-drilling data, though valuable for retrospective analysis, are less timely and unsuitable for immediate model responses.

For ROP prediction, the analysis targets drilling parameters—WOB, RPM, FLW, and MwIN—that directly influence operational reliability and process control. LWD parameter perception focuses on GR, DT, POR, and RD, providing critical data for real-time assessment of lithological and geophysical properties essential for safe formation evaluation. Formation pressure gradient perception integrates LWD and rock mechanics data, including Poisson's ratio, cohesion, overburden pressure, internal friction angle, and pressure gradients such as collapse, fracture, loss, and pore pressure. These parameters enable a comprehensive understanding of formation mechanical behavior, ensuring reliable and safe drilling operations through accurate real-time mechanical property acquisition.

3.1. Determination Methods for Formation pressure gradient Parameters

3.1.1. Measurement of Rock Mechanical Parameters

Triaxial compression testing is a fundamental experimental method for assessing the mechanical properties of rocks. By applying controlled axial and radial stresses, this test simulates the stress conditions that rocks experience at various subterranean depths, enabling the determination of critical mechanical parameters such as elastic modulus, Poisson's ratio, cohesion, and internal friction angle. These parameters are essential for understanding and predicting the behavior of rocks in

actual geological environments [37]. The experimental setup is illustrated in Fig. 2.

In the preliminary phase of the experiment, core samples were prepared from irregularly shaped rocks that were obtained directly from the drilling site. In the laboratory, cylindrical specimens approximately 25 mm in diameter were extracted from the original core using a diamond-core drill bit. The ends of the samples were then machined and polished to achieve a smooth and even surface, ensuring a length-to-diameter ratio within the standard range of 1.8 to 2.0. All coring operations were performed at a controlled temperature of 22°C, using kerosene as a circulating coolant to prevent temperature fluctuations from affecting rock properties.

The triaxial compression test apparatus is crucial for simulating subsurface stress conditions, comprising a high-pressure triaxial chamber capable of withstanding confining pressures up to 200 MPa and temperatures up to 200°C. The chamber is equipped with systems for confining pressure, axial loading, temperature control, and data acquisition. It features an internal design that compensates for the confining pressure during loading, automatically neutralizing the effects of this pressure to ensure that the longitudinal stress applied to the rock samples is precisely equal to the differential stress. Control of the confining and axial pressures was achieved using an electrohydraulic servo system, ensuring precise and stable pressure adjustments.

During the experiments, a series of controlled steps ensured accuracy and repeatability. Finely prepared rock samples were placed in a high-pressure autoclave designed to withstand intense pressure. The HP3054A data acquisition system was then activated and adjusted to the appropriate settings to commence the tests. In the uniaxial tests, a hydraulic press directly applied an axial load to the rock samples, and the data acquisition system continuously recorded the stress and strain until fracture. Triaxial tests began with the application of predetermined

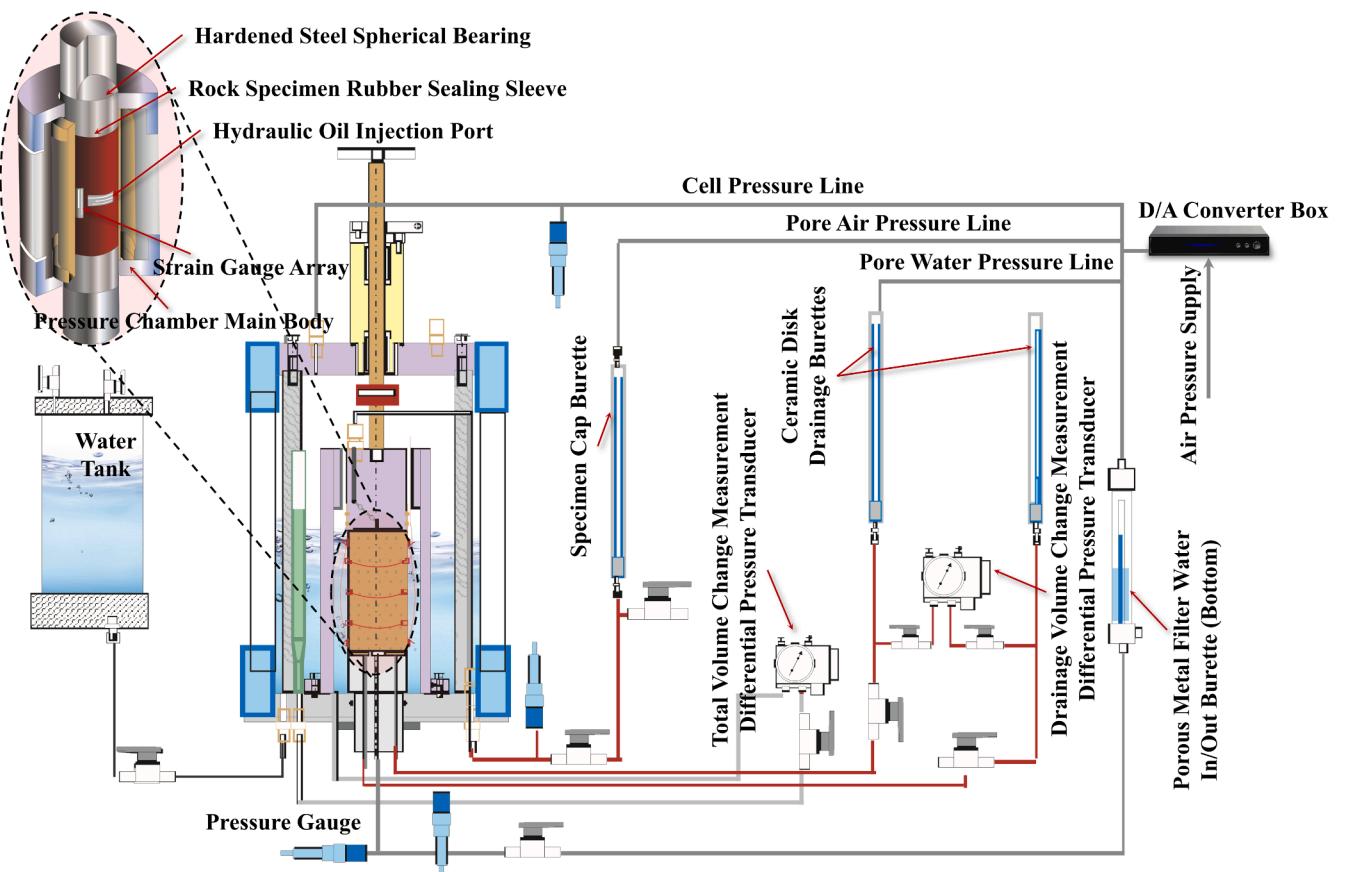


Fig. 2. Triaxial compression test apparatus.

confining pressures using a high-pressure pump to simulate actual subsurface stress conditions. Once the confining pressure stabilized, the axial pressure was increased, and the stress and strain data were recorded until the sample fractured. The analysis of the stress and strain data revealed the elastic modulus and Poisson's ratio of the rocks, whereas data near the yield point were used to calculate the ultimate strength, cohesion, and internal friction angle of the rocks. These precise mechanical parameters provide a scientific basis for formation pressure gradient calculations and the assessment of formation stability during drilling.

3.1.2. Calculation of Formation pressure gradient Parameters

Accurately calculating the formation collapse pressure, pore pressure, fracture pressure, and loss pressure is crucial for maintaining wellbore stability and optimizing the drilling fluid parameters [38]. The core issue with formation collapse pressure is that insufficient fluid column pressure may fail to support the stress state of the surrounding rock, potentially leading to rock shear failure, which in turn reduces the ROP and may cause severe drilling accidents [39]. The fracture pressure is related to the initiation of cracks or the propagation of existing fractures in the exposed formation under drilling fluid pressure. This typically occurs because of excessive mud density in the wellbore, which causes the circumferential stress to exceed the tensile strength of the rock. Accurate calculation of fracture pressure is critical for selecting the maximum mud density that directly affects wellbore stability and safety [40]. The accurate determination of the loss pressure is particularly crucial in fault zones, where faults act as weak planes in the rock mass. The mechanical strength of faults is generally lower than that of the surrounding rock, leading to significant discontinuities in the stress field and deformation near the faults. These discontinuities can easily trigger cracks and fissures near faults, causing fluid loss and, consequently, wellbore instability and reduced drilling efficiency [41].

The calculation of the formation collapse pressure was based on the Coulomb-Mohr strength criterion, which provides the conditions for rock shear failure [42]. The specific expressions are given by Eq. 1.

$$\begin{cases} \tau \geq C + \mu_f \sigma \\ \mu_f = \tan(\Phi) \end{cases} \quad (1)$$

where τ denotes the shear stress, (MPa); C represents the cohesion of the rock, (MPa); Φ is the internal friction angle of the rock, ($^{\circ}$); and σ denotes the normal stress on the rock, (MPa).

The strength was determined via triaxial compression tests under varying confining pressures. The Coulomb-Mohr criterion is rewritten in terms of the principal stresses, as shown in Eq. 2.

$$\begin{cases} \sigma_H = \sigma_h \cot^2(45^{\circ} - \Phi/2) + 2C \cot(45^{\circ} - \Phi/2) \\ \sigma_c = \frac{2C \cos(\Phi)}{1 - \sin(\Phi)} \end{cases} \quad (2)$$

where σ_H denotes the maximum horizontal principal stress, (MPa); σ_h the minimum horizontal principal stress, (MPa); and σ_c the uniaxial compressive strength, (MPa).

In the analysis of wellbore collapse, considering the influence of pore pressure, the Coulomb-Mohr criterion was modified, as shown in Eq. 3.

$$(\sigma_h - \alpha P_{por}) = (\sigma_h - \alpha P_{por}) \cot^2(45^{\circ} - \Phi/2) + 2C \cot(45^{\circ} - \Phi/2) \quad (3)$$

where P_{por} denotes the pore pressure in the rock, (MPa); and α is the effective stress coefficient, $0 < \alpha \leq 1$ (dimensionless).

As shown in Eq. 3, rock shear failure is primarily controlled by the maximum and minimum principal stresses. The greater the difference between σ_h and σ_H , the more likely wellbore collapse is to occur. Wellbore collapse instability occurs at $\theta = 90^{\circ}$ and $\theta = 270^{\circ}$, where the effective differential stress $\sigma'_\theta - \sigma'_r$ is maximal. The effective stress at the collapse point is given by Eq. 4.

$$\begin{cases} \sigma'_r = p_i - \alpha p_p \\ \sigma'_\theta = \eta(3\sigma_H - \sigma_h - p_i) - \alpha p_p \\ \tau_{r\theta} = 0 \end{cases} \quad (4)$$

where σ'_r and σ'_θ denote the minimum and maximum effective principal stresses at the collapse point, (MPa), respectively; η is the stress nonlinear correction coefficient (dimensionless); and $\tau_{r\theta} = 0$ indicates that the tangential shear stress is zero under the considered stress state, applicable to axisymmetric conditions, (MPa). Using the Coulomb-Mohr criterion, the wellbore collapse pressure gradient is given by Eq. 5.

$$\rho_{coll} = \frac{\eta(3\sigma_H - \sigma_h) - 2CK + \alpha P_p(K^2 - 1)}{(K^2 + \eta)H} \quad (5)$$

where ρ_{coll} denotes the formation collapse pressure gradient, (g/cm^3); σ_H and σ_h are the maximum and minimum horizontal principal stresses, (MPa); and H is the well depth, (m).

An accurate calculation of the fracture pressure is critical, because if the mud density in the wellbore is excessive, the circumferential stress on the wellbore rock can surpass the rock's tensile strength, causing formation fractures. Formation fractures typically occur where circumferential stress is minimal, specifically at $\theta = 0^{\circ}$ or 180° . The circumferential stress at this point is expressed by Eq. 6.

$$\sigma'_\theta = 3\sigma_H - \sigma_h - p_i \quad (6)$$

Considering the wellbore in a plane strain state and based on rock mechanics theory, a calculation model for wellbore tensile failure under non-uniform in situ stress was constructed. The wellbore fracture pressure gradient is given by Eq. 7.

$$\rho_{frac} = \frac{\left(\frac{2\mu}{1-\mu} - K\right)(\sigma_V - \alpha P_p) + \alpha P_p + S_t}{H} \quad (7)$$

where ρ_{frac} is the formation fracture pressure gradient, (g/cm^3); S_t is the formation tensile strength, (MPa), σ_V is the overburden pressure, (MPa), μ is Poisson's ratio (dimensionless). Formation loss pressure approximates the minimum horizontal principal stress with sufficient engineering accuracy [43]. The calculations are given by Eq. 8.

$$\rho_{loss} = \frac{\sigma_h}{H} \quad (8)$$

where ρ_{loss} is the formation loss pressure gradient, (g/cm^3).

3.2. Design of Offline Models for ROP and Formation Property Perception in Neighboring Wells

3.2.1. Fully Connected Neural Network Model

An FCNN comprises multiple layers of neurons, each connected to every neuron in the preceding layer, hence the term "fully connected." This architecture is suitable for learning nonlinear and complex function mappings from fixed-size input data [44]. It comprises an input layer, multiple hidden layers, and an output layer. The input layer receives preprocessed drilling parameters: WOB, RPM, FLW, and MWIN. Each hidden layer contains neurons that perform nonlinear transformations on the inputs via activation functions that gradually extract higher-level features. The output layer produces the predicted ROP. In an FCNN, forward propagation refers to the process in which data flow from the input layer to the output layer, with each layer's output serving as the input for the next layer. The model training relies on backpropagation, which is an efficient gradient descent method designed to optimize network weights and minimize prediction errors. During training, the loss function evaluates the deviation between the model predictions and actual data by adjusting the network parameters to minimize this deviation [45].

3.2. 2 iTransformer Model

For formation property prediction, the iTransformer model provides an efficient architecture for predicting multivariable time-series data within specific time windows, such as GR, DT, POR, and RD [46]. This model effectively processes and analyzes these complex datasets to achieve accurate geological feature perception. The architecture of the model is illustrated in Fig. 3. Detailed descriptions of the model's architecture, including its self-attention mechanism and embedding and projection methods, are provided in Appendix D.

3.3. Offline Perception of ROP and Formation Properties in Adjacent Wells

To enhance the accuracy of the ROP prediction, LWD parameter perception, and formation pressure gradient perception, refined training strategies were implemented. Different dataset-partitioning strategies were employed for various prediction tasks. Random partitioning was utilized for ROP prediction to ensure the randomness of the data distribution and the generalizability of the model. For the LWD parameter perception and formation pressure gradient perception, an 8:1:1 depth ratio was employed to better simulate the data distribution during actual drilling processes [47]. An adaptive learning rate-adjustment strategy was employed to adjust the learning rate dynamically through learning-rate annealing. This strategy enables rapid convergence with a higher learning rate at the beginning of training, and gradual reduction based on changes in the loss function in later stages [48]. In addition, the AdamW optimizer, which combines adaptive moment estimation and weight decay, was selected to effectively improve the training efficiency and generalizability of complex dataset [49]. The Mean Squared Error (MSE) and Smooth L1 Loss were used as loss functions to directly quantify the deviation between the predicted and actual values, aiding in the precise optimization of the model outputs [50].

Dropout layers and regularization techniques were incorporated into the model to prevent overfitting and validate its generalizability. These techniques effectively avoid overfitting the training data, while maintaining training efficacy [51,52]. In addition, an Early Stopping strategy was employed to dynamically adjust the training epochs based on the performance of the validation set. Training was halted when the performance of the model on the validation set ceased to improve, thereby preventing overtraining and preserving the model's optimal performance. These comprehensive training strategies and techniques ensured high accuracy and reliability in predicting the LWD parameters.

The training strategies for the prediction and perception of the offline models are presented in Table 1.

To optimize the hyperparameters of the FCNN model, Particle swarm optimization (PSO) algorithm was employed [53]. Fig. 4 shows the R^2 fit of the PSO-FCNN model for the training, validation, and test datasets.

In Fig. 4(a), the PSO-FCNN model achieved an R^2 of 0.9378 on the training set, indicating high predictive accuracy with most points near the fit line. In Fig. 4(b), the model's validation R^2 was 0.9172, reflecting good generalizability, though points are more dispersed than in training. Fig. 4(c) shows a test set R^2 of 0.8658, illustrating that while predictive accuracy decreases slightly, acceptable performance is maintained on independent data.

Fig. 5 illustrates the performance of the iTransformer model in predicting the key LWD parameters across the training, validation, and test datasets.

Regarding GR, both the training and validation phases exhibited R^2 values of 0.97, indicating a high model fit and minimal error in these phases. However, the R^2 value decreased to 0.72 in the testing phase, showing a slight decline in the generalizability of the model for unseen data, although it maintained a relatively reasonable prediction accuracy. For the prediction of DT, the model achieved an R^2 value of 0.98 during the training phase, demonstrating excellent fitting capability. Despite a decrease in R^2 values to 0.91 and 0.82 during the validation and testing phases, respectively, the model continued to show good generalizability, particularly in the analysis of complex geological data. The prediction results for POR in the training phase reached an almost perfect R^2 value of 0.98, indicating a near-zero error. Although the R^2 values declined slightly in the validation and testing phases, the model maintained high prediction accuracy, demonstrating its efficiency and reliability in perceiving formation POR. For RD prediction, the model achieved an R^2 value of 0.93 in the training phase and R^2 values of 0.82 and 0.89 in the validation and testing phases, respectively. Despite having the highest fit in the training phase, the model showed consistent performance across different datasets, indicating its stability and ability to accurately capture variations in RD. Fig. 6 shows the perception effectiveness of the model for the formation pressure gradient across these datasets.

For the collapse pressure gradients, the model performed best during the training phase ($R^2 = 0.94$), with a slightly lower performance during the validation and testing phases, but still maintained high accuracy. This indicates the strong capability of the model to fit the data. The results for fracture pressure gradients were similar, with an R^2 of 0.96

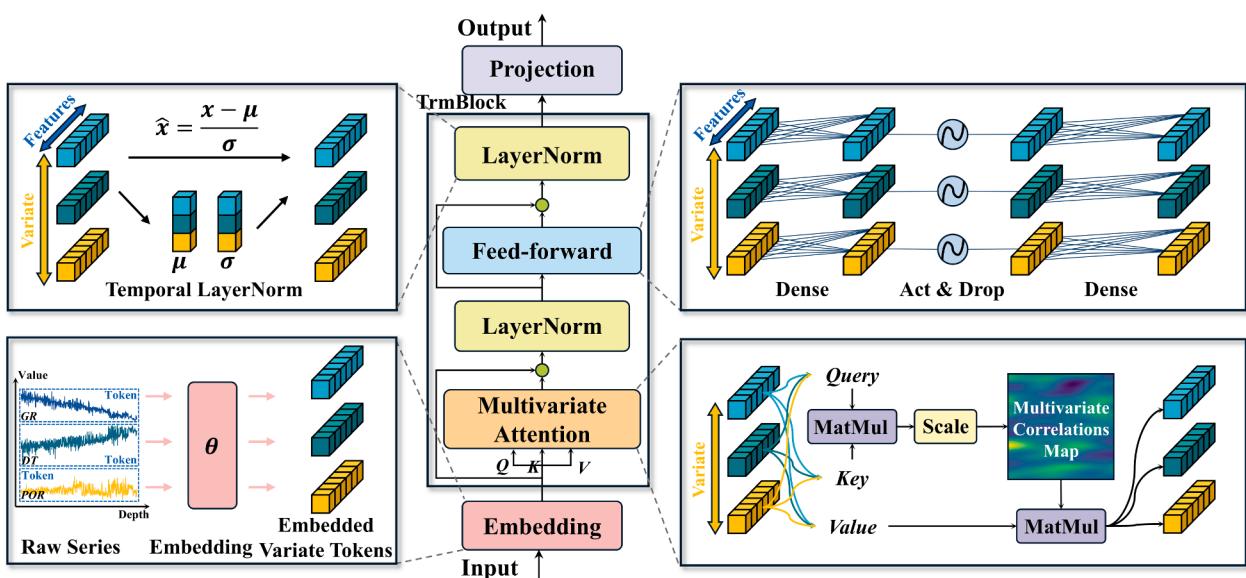


Fig. 3. Architecture of the iTransformer model.

Table 1

Training strategies for prediction and perception of offline models.

Hyperparameter	Description	ROP	Downhole Parameter	Formation pressure gradient
LearningRate	Learning rate for the model.	0.001	0.0002	0.0003
Optimizer	Optimization algorithm.	AdamW	AdamW	AdamW
BatchSize	Samples per training batch.	64	48	24
NumEpochs	Iterations over the dataset.	50	400	600
WeightDecay	Regularization to reduce overfitting.	0.02	0.015	0.02
DropoutRate	Dropout rate to prevent overfitting.	0.3	0.3	0.35
EarlyStoppingPatience	Patience for early stopping mechanism.	7	8	12
LearningRateScheduler	Scheduler for learning rate adjustment.	StepLR	StepLR	StepLR
StepSize	Step size for learning rate adjustment.	20	30	50
Gamma	Decay factor for learning rate adjustment.	0.2	0.2	0.2
Expand	Expansion factor for Mamba.	—	3	4
DConv	Convolution kernel size for Mamba.	—	5	3
TopK	Top-k pooling for TimesBlock.	—	8	10
NumKernels	Number of kernels for Inception.	—	7	8
DModel	Dimension of model.	—	512	512
NHeads	Number of heads in multi-head attention.	—	8	10
ELayers	Number of encoder layers.	—	3	4
DFf	Dimension of feed-forward network.	—	2048	2560
Dropout	Dropout probability.	—	0.25	0.3
Activation	Activation function used in the model.	—	gelu	gelu
UseNorm	Whether to use normalization.	—	1	1

during the training phase, with a slight decrease during the validation and testing phases. The high R^2 values for these two metrics indicate that the model can accurately predict the formation pressure gradients despite slight performance fluctuations across different datasets. For the loss and pore pressure gradients, although the R^2 values were high during the training phase, they decreased during the validation and testing phases. In particular, R^2 for the pore pressure in the testing set was only 0.80, indicating some challenges in the generalizability of the model to unseen data. Overall, the prediction accuracy of the model was reliable.

4. Transfer and Online Perception of ROP and Formation Properties for Untapped Target Wells

For the transfer and online perception of ROP and formation properties in the unopened formations of the target well, the data were sourced from target well locations within the same block of neighboring drilled wells in the Caofidian 6-4 block of the Bohai Sea, China. The data for the open formations of the target well included 100 records covering depths ranging from 1 to 100 m. The selection of 100 data points represents a trade-off between prediction accuracy and timely deployment during initial drilling stages. This dataset comprises pre-drilling forecast data and real-time acquisition data during drilling, focusing on direct observations and measurements during the real-time drilling process as well as predictive analysis for future drilling activities. Fig. 7 compares the kernel density of the drilling, LWD, and formation pressure gradient parameters between the drilled neighboring wells and the unopened formations of the target well.

Fig. 7 illustrates the kernel density comparison of the drilling and formation pressure gradient parameters between the drilled neighboring wells and the unopened formations of the target well in the Caofidian 6-4 block of the Bohai Sea, China. Because the wells were located within the same block, the kernel density comparison of the LWD and formation pressure gradient parameters showed minor differences, reflecting the high consistency of the formation properties and pressure conditions within the block. In contrast, the comparison of drilling parameters revealed significant differences, highlighting variations in drilling operational conditions, such as WOB, RPM, and FLW. These factors directly affect the drilling efficiency and safety. The significant differences in drilling parameters may stem from the localized heterogeneity in geological structures or variations in drilling techniques and operational strategies.

The target well formation dataset comprises 100 records spanning depths of 1–100 m, as shown in Table A2, whereas the neighboring drilled well covers a much broader depth range (depths from 696 m to 3142 m) and are detailed in Table A1 of Appendix A. Table A3 provides a statistical summary of key drilling and logging parameters (including WOB, RPM, FLW, MwIN, GR, DT, POR, RD, and ROP) by presenting the minimum, maximum, mean, standard deviation, and skewness. This summary visually reflects the distribution characteristics of the data. In addition, a wavelet denoising method was employed to preprocess the

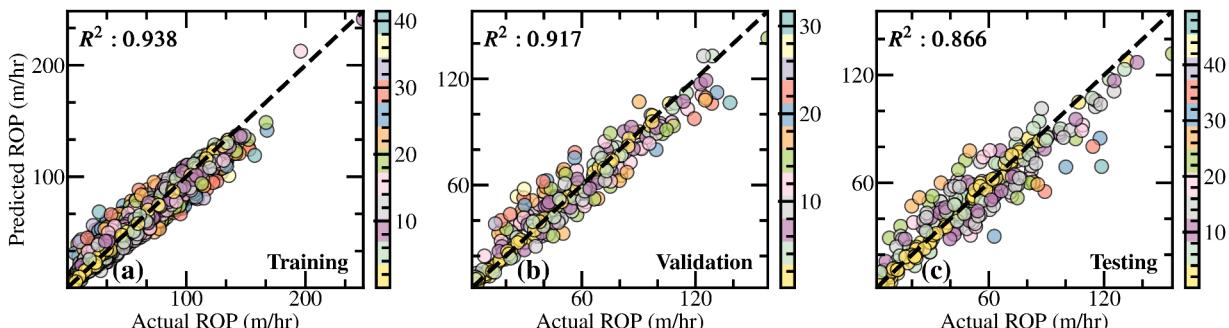


Fig. 4. R^2 fit of ROP prediction models for different data in adjacent wells.

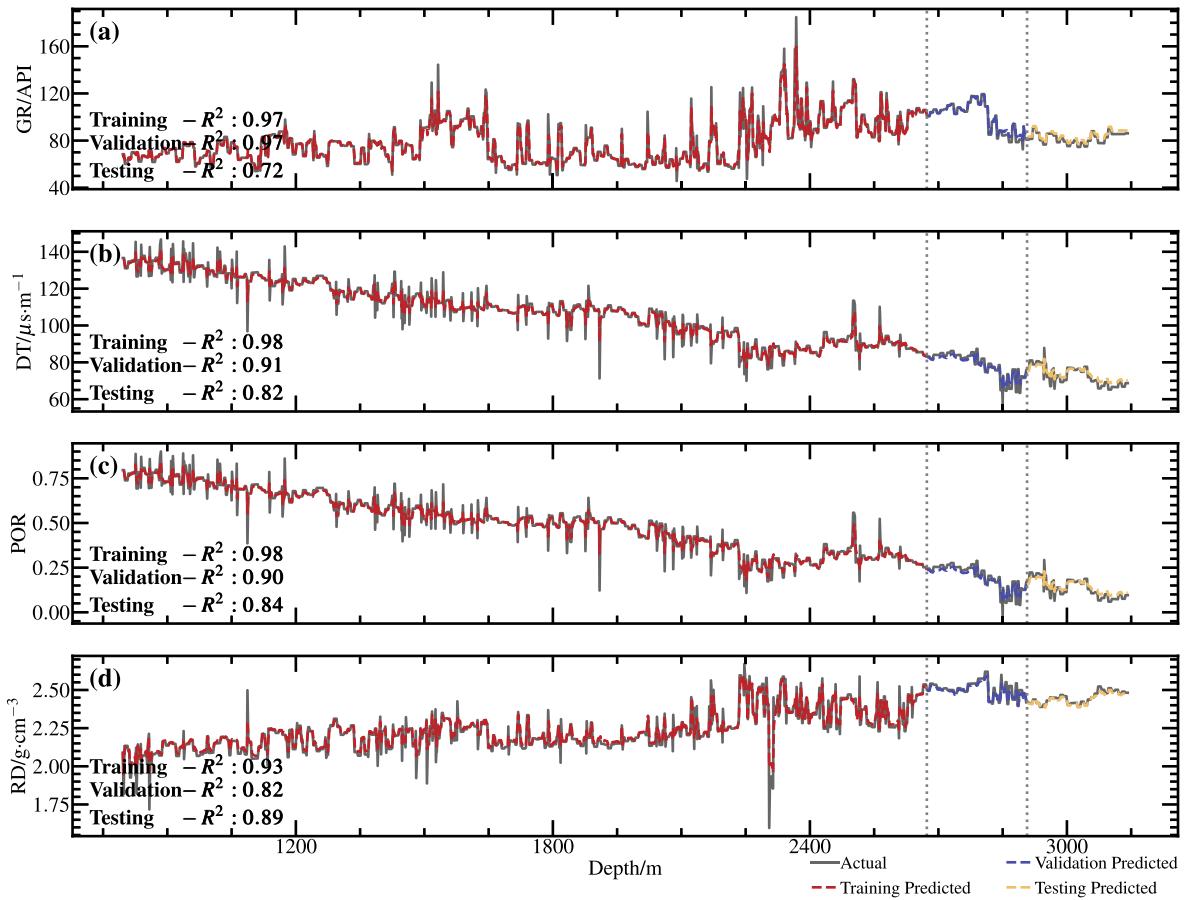


Fig. 5. Prediction performance of the offline perception iTransformer model for LWD parameters in adjacent wells.

raw data. The wavelet denoising effect for the key LWD parameters (GR, DT, POR, RD) is shown in Fig. A1 of Appendix A.

4.1. Transfer of the Perception Model Based on the Domain Adversarial Mechanism

4.1.1. Architecture of the Domain Adversarial Mechanism
To achieve effective transfer and online perception of the ROP and formation property parameters, a domain adversarial mechanism was designed to handle data from drilled neighboring wells and unopened formations of the target well in the Caofidian 6-4 block of the Bohai Sea, China. The architecture of the domain adversarial mechanism is shown in Fig. 8. The complete process for optimizing the transfer perception of ROP and formation property parameters using the domain adversarial mechanism is presented in Table B1 of Appendix B.

4.1.2. Freezing and Fine-Tuning Model Parameters

The strategy of parameter freezing and fine-tuning is crucial for model adaptability, particularly in cross-well application scenarios. The three models involved PSO-FCNN for ROP prediction, iTransformer for LWD parameter perception, and iTransformer for formation pressure gradient perception—underwent parameter freezing and fine-tuning to optimize their performance for undrilled formation data in the target well. The architecture of the parameter freezing and fine-tuning for the iTransformer model is illustrated in Fig. 9.

In the adaptation of the predictive and perceptive transfer models, parameter freezing primarily targets deep network parameters that exhibit stable performance during the source-domain training phase, particularly the primary and intermediate feature extraction layers

responsible for capturing basic patterns and features. These layers capture features that demonstrate generalizability across different geological conditions, and are crucial for the cross-scenario robustness of the model. Freezing these parameters ensures that the model relies on fundamental features to maintain its predictive performance when processing new data [54]. Conversely, fine-tuning focuses on the upper layers of the model, particularly the advanced, fully connected layers, whose weights and biases are directly associated with the output of the model. Fine-tuning aims to optimize the predictive accuracy and relevance according to the characteristics of the target domain, adapting to the unique properties of the target geological data, such as ROP variation patterns, POR distribution characteristics, and dynamic changes in formation pressure gradient [55]. This strategy is implemented using a progressive approach. Initially, most key layer parameters remain frozen, and the fine-tuned layers are adjusted with a low learning rate. As the performance of the model on the target domain data improved, more layers gradually became unfrozen, allowing for extensive parameter updates. This approach prevents the model from overfitting the source domain data while encouraging it to gradually adapt to the new geological environment, leveraging the retained knowledge. The specific progressive unfreezing strategies are listed in Table 2.

4.2. Model Training and Evaluation

In the process of transfer learning for predicting the ROP, perceiving the LWD parameters, and perception the formation pressure gradient, optimizing the feature extraction and domain discrimination strategies for the source and target domains is crucial. The source domain data were used directly for comprehensive training without splitting,

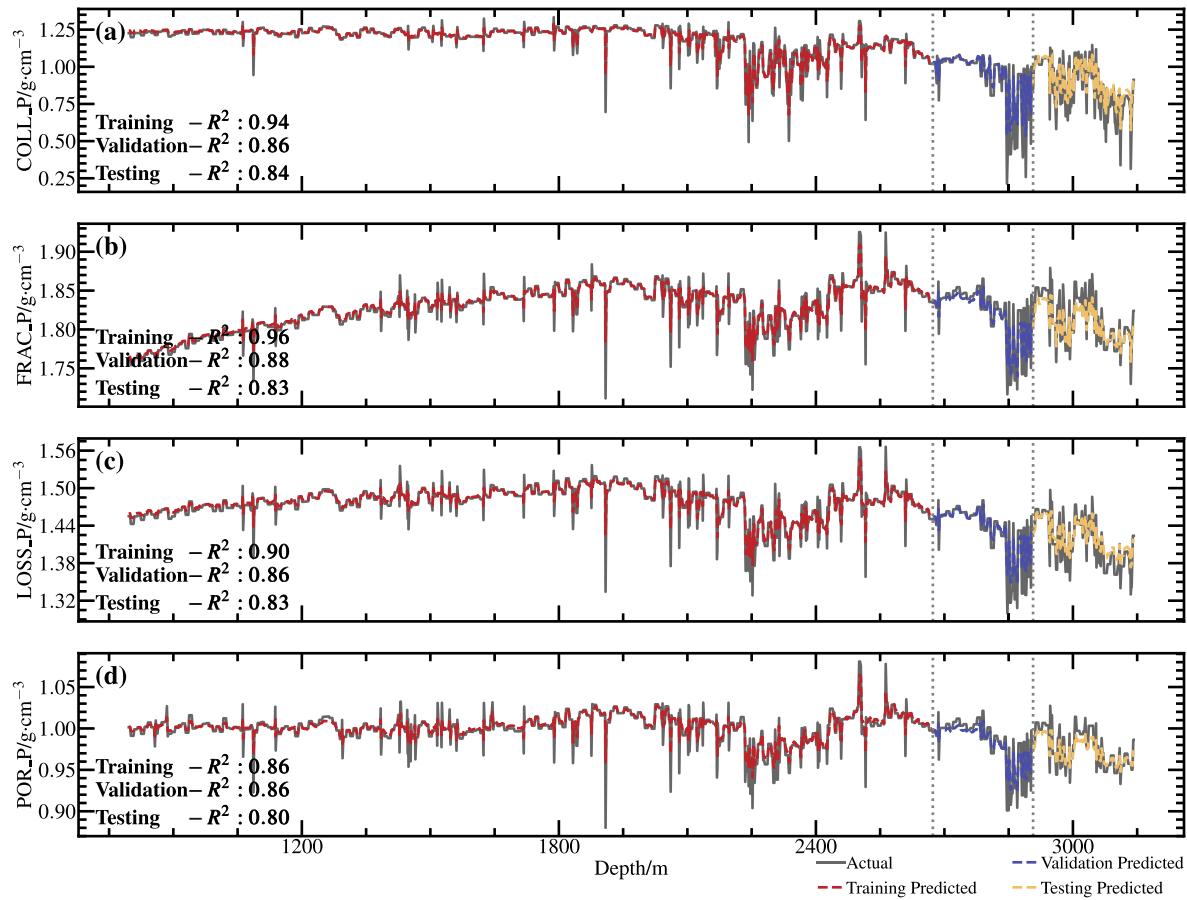


Fig. 6. Prediction performance of the offline perception iTransformer model for formation pressure gradients in adjacent wells.

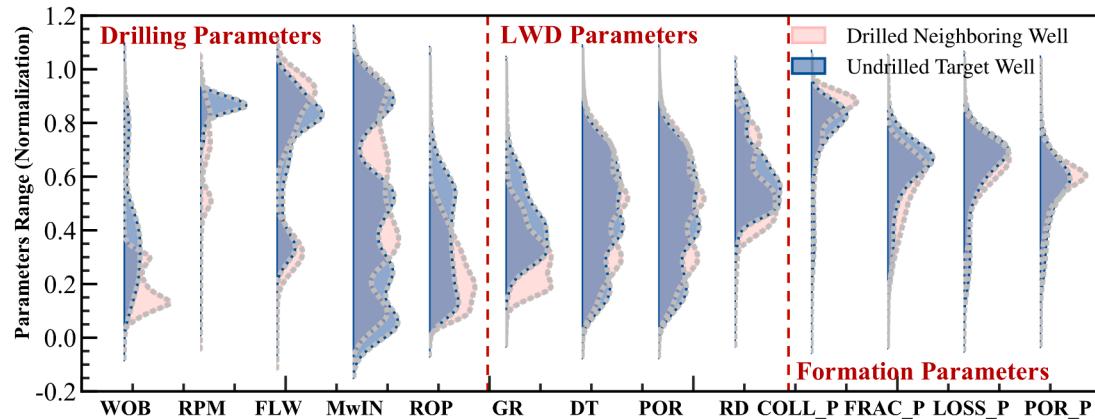


Fig. 7. Comparison of kernel density for key parameters between adjacent wells and untapped target wells.

whereas the target domain data were divided into training and validation sets in a 5:5 ratio. This division strategy aims to assess the adaptability of the model to new geological environments.

In the implementation of the transfer model training, the strategies remained consistent with those used in the offline perception of ROP and formation properties. A high initial learning rate was employed, and the AdamW optimizer was used to rapidly reduce the loss. As the training progressed, the learning rate was gradually adjusted using an annealing strategy to finely tune the model parameters. In addition, regularization and dropout methods were employed to enhance the generalization of

the transfer model. An early stopping method was used to monitor the performance of the validation set and halt training appropriately to prevent overfitting, ensuring that the transfer model maintained a high accuracy and stability in the new data environment. These measures collectively constitute the core strategies for transfer model training aimed at optimizing the analysis and prediction accuracy of drilling and geological data. The training strategies for the predictive and perceptive transfer models are listed in Table 3.

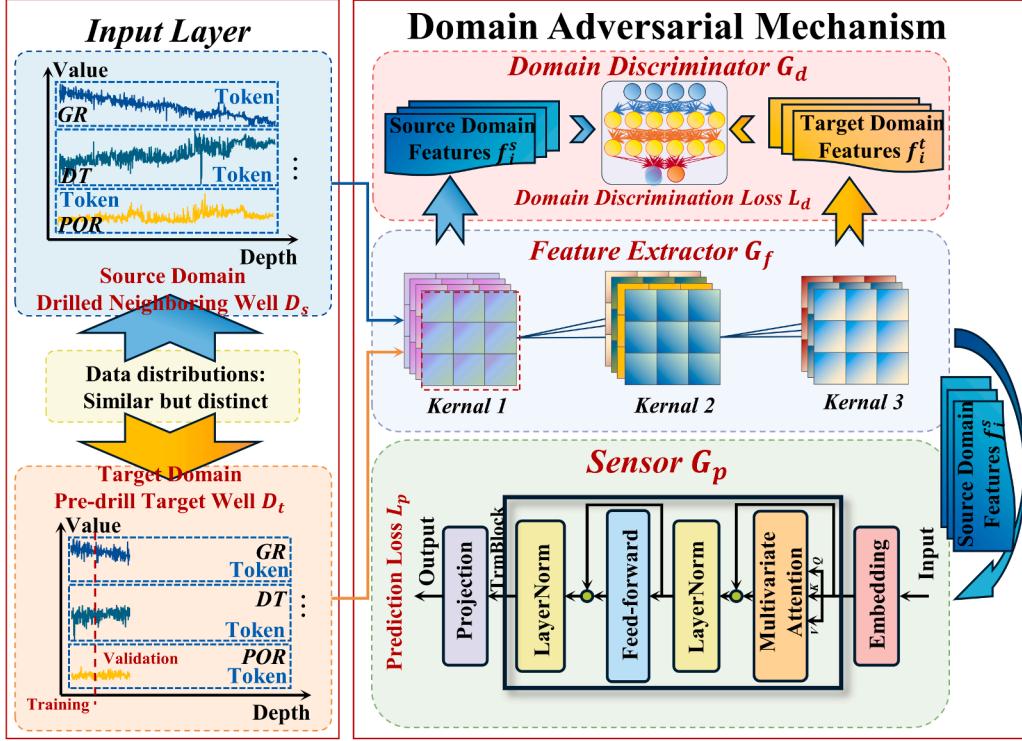


Fig. 8. Architecture of domain adversarial mechanism for adjacent and untapped target wells.

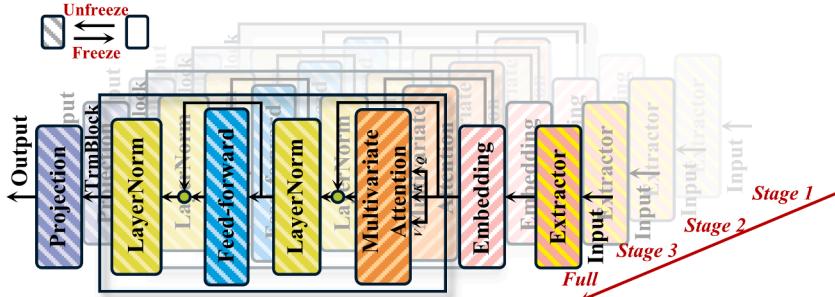


Fig. 9. Architecture of parameter freezing and fine-tuning for the iTransformer model.

4.2.1. Transfer Prediction of ROP for Untapped Target Wells

To investigate the effect of different learning rates on the transfer prediction model for the ROP from drilled neighboring wells to untapped target formations, three learning rate levels were set: 0.00005, 0.0001, and 0.00025. These were applied to the feature extractor, domain discriminator, and predictor to evaluate their specific effects on the performance of the transfer model. Fig. 10 illustrates the changes in the prediction loss values and domain discriminator loss values during the training and validation stages under three different learning rate settings for the ROP prediction transfer model.

In Fig. 10(a), a higher learning rate (0.00025) leads to the fastest loss reduction, facilitating rapid adaptation and convergence of the transfer model, which is crucial for iterative optimization in environments with frequently updated drilling data. Fig. 10(b) demonstrates that higher learning rates enable the transfer model to swiftly adjust to new validation data, enhancing prediction efficiency. Conversely, lower learning rates (0.00005 and 0.0001) result in a more gradual and stable loss decrease, improving model generalizability and reducing the risk of overfitting. Fig. 10(c) shows that the domain discriminator achieves the fastest loss reduction at a learning rate of 0.00025, effectively distinguishing feature differences between source and target domains and

thereby increasing the transfer model's prediction accuracy under new geological and operational conditions. In contrast, lower learning rates slow the loss decrease, requiring more iterations to reach similar performance levels.

After determining that 0.00025 is the optimal learning rate for the transfer model to predict the ROP of the undrilled strata of the target well, the transfer model underwent comprehensive training. Subsequently, to enhance its adaptability to the undrilled strata of the target well, all transfer model parameters were frozen. The specific progressive unfreezing strategy is detailed in Table 2. Fig. 11 shows the change in the prediction loss of the predictor during training and validation across the four unfreezing stages.

In Fig. 11(a), the initial stage with all layers frozen results in high training loss, indicating that the ROP prediction transfer model has not yet adapted to the specific geological characteristics of the target well. As the model enters the first stage by unfreezing the output layer, the training loss begins to decrease, demonstrating initial adaptation to the target domain's geological features. In the second stage, unfreezing the penultimate layer further reduces the training loss, enhancing the model's ability to accurately map the relationship between ROP and geological parameters. Fig. 11(b) shows that validation loss decreases

Table 2
Progressive unfreezing strategy for prediction and perception transfer models.

Task Category	Stage	Unfrozen Components	Description
ROP	Initial Training	None	All layers frozen to stabilize initial outputs.
	Stage 1	Output Layer	Unfreezes final layer for initial domain-specific adaptation.
	Stage 2	Output + Penultimate Layers	Unfreezes deeper layers to adjust intermediate features.
	Stage 3	Full Predictor Network	Extends unfreezing to entire predictor for enhanced flexibility.
	Full Adaptation	Entire Network	All layers unfrozen for comprehensive model fine-tuning.
LWD Parameter Formation pressure gradient	Initial Training	None	All layers frozen to ensure stability and prevent overfitting early on.
	Stage 1	Output Layer	Unfreezes final output layer to begin domain-specific adjustments.
	Stage 2	Output + Attention Layers	Unfreezes attention and output layers for refining key temporal features.
	Stage 3	Attention + Intermediate Layers	Broadens unfreezing to include more of the intermediate layers for deeper adaptation to domain-specific features.
	Full Adaptation	Entire Network	All layers unfrozen to maximize adaptability and fine-tune across all aspects of the network for optimal performance.

through the first three stages, with the lowest validation loss achieved in the third stage, indicating optimal generalizability and effective adaptation to complex geological conditions. However, in the final stage, fully unfreezing the network leads to a significant increase in validation loss, signaling overfitting.

Fig. 12(a-d) shows that without transfer learning, the baseline R^2 is 0.2279. Adjusting the learning rates to 0.00005, 0.0001, and 0.00025 significantly improves the fit, with the highest R^2 of 0.8169 achieved at a learning rate of 0.00025, highlighting the critical role of learning rate in enhancing model performance. Fig. 12(e-h) demonstrates the impact of progressively unfreezing model layers. In the first stage, unfreezing only the output layer results in a slight R^2 decrease to 0.7732. The second

stage, with the penultimate layer unfrozen, maintains a stable R^2 of 0.7711. Fully unfreezing the prediction network in the third stage boosts R^2 to 0.8783, indicating optimal adaptation. However, in the final stage, R^2 drops to 0.7575, signaling overfitting and reduced generalizability. These results emphasize the importance of balancing model adaptation with generalization to ensure reliable and resilient ROP predictions in dynamic drilling environments.

4.2.2. Transfer Perception of Formation Properties for Untapped Target Wells

Fig. 13(a-c), a learning rate of 0.0005 resulted in training and validation losses of 0.0146 and 0.0011, respectively, indicating effective model adaptation and generalization. Increasing the learning rate to 0.001 further reduced the training loss to 0.0105 and validation loss to 0.0009, while a learning rate of 0.002 maintained a training loss of 0.0123 and achieved the lowest validation loss of 0.0009, demonstrating superior perception capability. In Fig. 13(d-f), the formation pressure gradient perception transfer model showed stable but limited adaptation with a learning rate of 0.0005 (losses of 0.1976 and 0.1898). Higher learning rates of 0.001 and 0.002 decreased the training losses to 0.1473 and 0.1453, respectively, while validation losses remained low at 0.1903 and 0.1899. Additionally, the domain discriminator's loss significantly decreased from 0.0752 to 0.0068 with higher learning rates, enhancing the model's ability to distinguish between source and target domains.

Fig. 14(a-b) shows the training and validation loss dynamics of the formation properties perception transfer model across four unfreezing stages for key geological parameters (GR, DT, POR, RD). Initially, all layers are frozen, resulting in high losses. In the first stage, unfreezing the output layer reduces training loss to 0.0011 and validation loss to 0.0073, enhancing model responsiveness to new drilling data. The second stage, unfreezing output and attention layers, further decreases training loss to 0.0009 while maintaining validation loss, improving sensitivity to POR and RD changes. In the third stage, unfreezing intermediate layers lowers training loss to 0.0008 and validation loss to 0.0068, demonstrating high adaptability. Fully unfreezing all layers in the final stage reduces training loss to 0.0007 but slightly increases validation loss to 0.0069, indicating potential overfitting.

Fig. 14(c-d) depicts the training and validation loss dynamics of the formation pressure gradient perception transfer model across four unfreezing stages. Initially, all layers were frozen, resulting in baseline perception accuracy. In the first stage, unfreezing the output layer reduced the training loss to 0.1317 and validation loss to 0.1888, enhancing adaptation to formation pressure gradients. In the second stage, unfreezing both output and attention layers further decreased

Table 3
Training strategies for prediction and perception transfer models.

Parameter Class	Hyperparameter	ROP	Downhole Parameter	Formation pressure gradient
General Configurations	<i>LearningRate</i>	-	-	-
	<i>Optimizer</i>	AdamW	AdamW	AdamW
	<i>BatchSize</i>	32	32	32
	<i>WeightDecay</i>	0.02	0.015	0.02
	<i>DropoutRate</i>	0.3	0.3	0.35
	<i>EarlyStoppingPatience</i>	100	100	100
Learning Rate Scheduler	<i>Scheduler</i>	StepLR	StepLR	StepLR
	<i>StepSize</i>	20	30	50
	<i>Gamma</i>	0.2	0.2	0.2
Feature Extractor	<i>NumberOfLayers</i>	2	3	3
	<i>ConvolutionKernels</i>	[32,64]	[32, 64, 128]	[32, 64, 128]
	<i>KernelSizes</i>	[3, 3]	[3, 5, 3]	[5, 5, 3]
	<i>PoolingLayers</i>	[Max, Max, Avg]	[Max, Max, Avg]	[Max, Max, Avg]
	<i>ActivationFunction</i>	ReLU	ReLU	ReLU
Domain Discriminator	<i>Normalization</i>	BatchNorm	BatchNorm	BatchNorm
	<i>Layers</i>	2	3	4
	<i>ActivationFunction</i>	ReLU	ReLU	ReLU
Predictor	<i>OutputSize</i>	2	2	2
	<i>TaskRelated</i>	Same as offline task	Same as offline task	Same as offline task

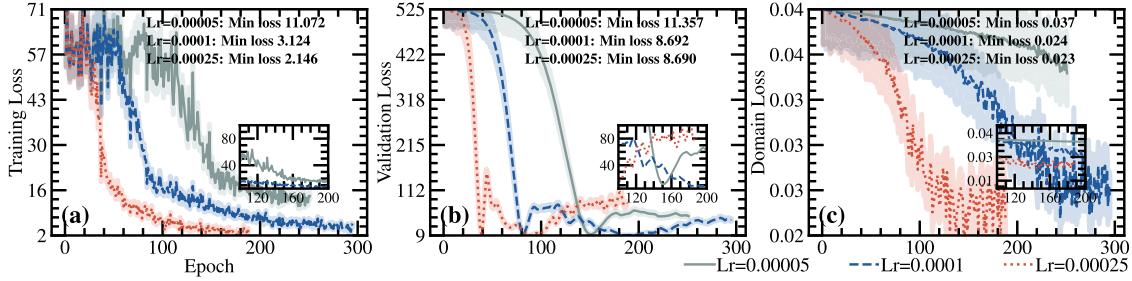


Fig. 10. Dynamics of losses for the ROP prediction transfer model and domain discriminator under different learning rate settings.

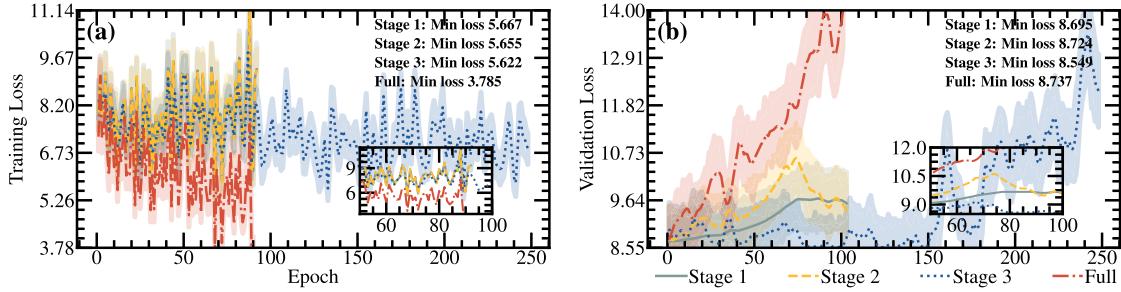


Fig. 11. Dynamics of training and validation losses for the predictor in the ROP prediction transfer model across four unfreezing stages.

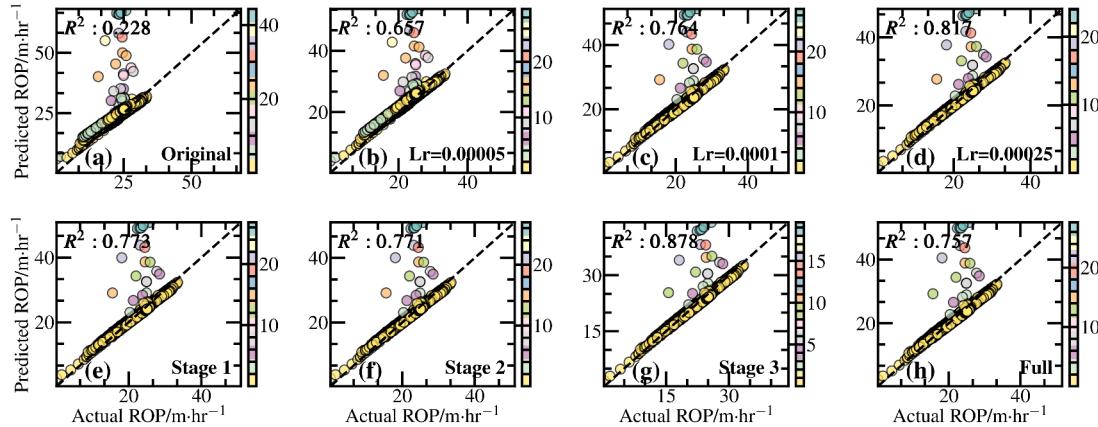


Fig. 12. R^2 fit of the ROP prediction transfer model under different configurations (learning rates and unfreezing stages).

training loss to 0.1285 while slightly increasing validation loss to 0.1901, highlighting challenges in capturing temporal features. The third stage involved unfreezing intermediate layers, which lowered training loss to 0.1151 and validation loss to 0.1925, demonstrating improved adaptability. However, in the final stage, fully unfreezing all layers reduced training loss to 0.1086 but increased validation loss to 0.1955, indicating potential overfitting. These findings underscore the necessity of balancing model adaptation and generalizability to ensure reliable and resilient pressure gradient predictions in dynamic drilling environments.

Fig. 15(a-d) illustrates the MAE progression of the LWD parameter-perception transfer model for GR, DT, POR, and RD. Training MAE values indicate strong initial learning, while validation MAE increased by over 200% for GR and DT. Testing MAE escalated to nine to fifteen times training values, highlighting severe generalization challenges in heterogeneous geological environments despite domain-adversarial adaptation.

Fig. 15(e-h) analyzes MAE trends for formation pressure gradients. Collapse pressure gradients exhibited testing MAE exceeding twelve times training values under overpressure conditions, reflecting limited cross-well transferability. Fracture and loss pressure gradients showed testing errors seven to nine times higher than training, suggesting insufficient freezing-fine-tuning adjustments for localized stress regimes. Pore pressure predictions demonstrated relative stability with testing MAE four times training values, attributed to domain-invariant data distributions.

4.3. Online Perception of Formation Properties

Performance declines of LWD parameter and formation pressure gradient transfer models during validation reveal significant limitations in adaptability and generalization under complex geological conditions. These shortcomings highlight the inadequacies of traditional static training methods in dynamic environments, emphasizing the necessity

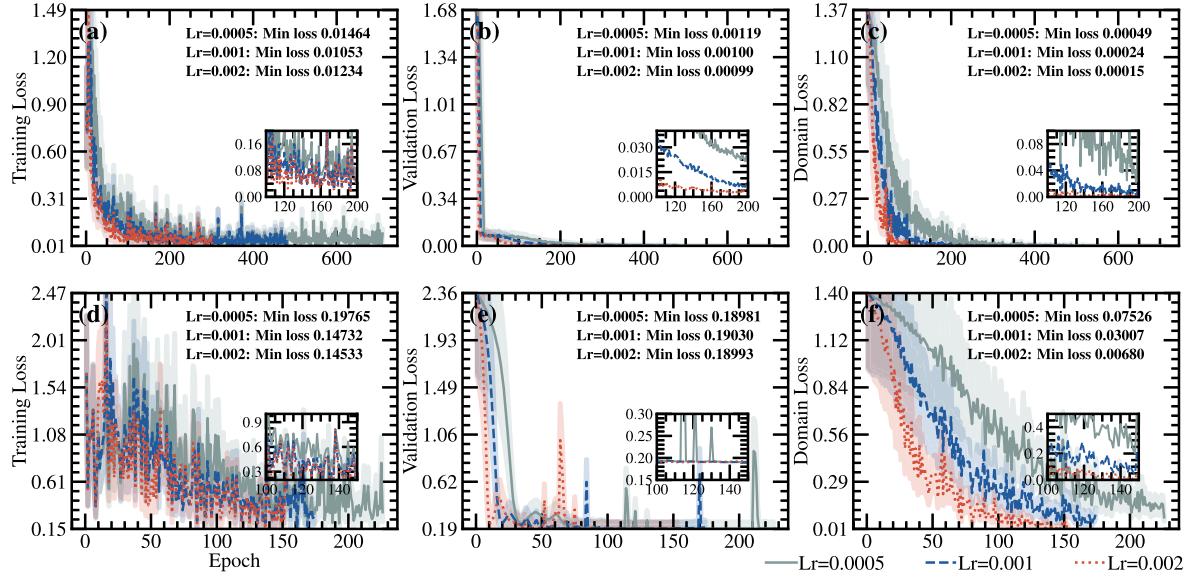


Fig. 13. Dynamics of losses for the formation properties perception transfer model and domain discriminator under different learning rate settings.

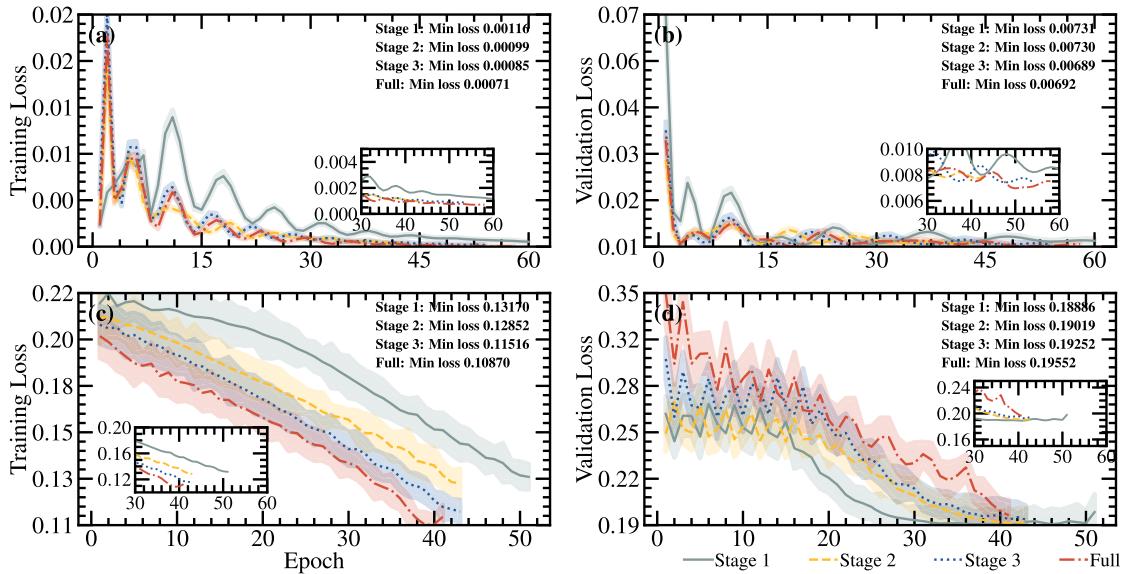


Fig. 14. Dynamics of training and validation losses for the perceiver in the formation properties perception transfer model across four unfreezing stages.

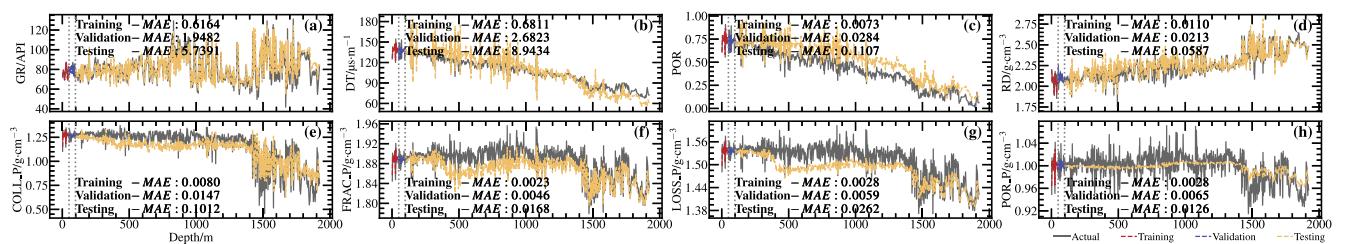


Fig. 15. Dynamics of perception performance for the LWD parameters perception transfer model in stage 3 and the formation pressure gradient perception transfer model in stage 1.

for an online learning perception strategy to enhance model reliability and resilience.

The online learning strategy facilitates continuous adaptive optimization of formation properties transfer models, including LWD parameters and pressure gradient models, through incremental learning. This enables immediate network updates with new drilling data via a rolling window, reducing training resource demands [56]. Real-time data undergo strict preprocessing, such as feature standardization, interpolation, and denoising. The iTransformer model, combined with feature extractors, processes drilling parameters and pressure gradient data in real time, delivering key perception results. Model adaptability is periodically assessed, with learning strategies and parameters adjusted to ensure precision and robustness in complex geological conditions. The specific parameters of the incremental learning strategy are listed in **Table 4**.

Implementing the online learning strategy in offshore drilling operations requires balancing computational efficiency and real-time responsiveness. Incremental updates limit GPU memory usage to below 2 GB per cycle by processing data in 100-point windows, enabling real-time deployment via Huawei Cloud edge computing units. Integration with existing LWD systems utilizes WITSML 2.0 interfaces, avoiding extensive hardware modifications. These design choices meet computational constraints and ensure compatibility with industry-standard data protocols.

Fig. 16 shows the perception performance of the formation-property transfer model in an online learning environment. The upper part of **Fig. 16** illustrates the performance of the transfer model for the LWD parameters, particularly its ability to handle GR, DT, POR, and RD. The performance evaluation of this transfer model was focused on a narrow range of 11 m after each model update, encompassing the ranges before and after the transfer model training point. In this local range, the transfer model demonstrated extremely high prediction accuracy, achieving a perfect R^2 value of 1.0000 for all the parameters at each test point. This high accuracy reflects the high sensitivity of the transfer model to minor formation changes and its rapid adaptability to variations in geological data, thereby underscoring its importance in practical applications.

The lower part of **Fig. 16** shows the ability of the transfer model to perceive formation pressure gradients, including the collapse, fracture, loss, and pore pressures, in an online learning environment. Predictions for the collapse pressure gradient and fracture pressure gradient showed nearly perfect R^2 values at multiple depth intervals, such as between 608 and 617 m, where both pressure gradients achieved an R^2 of 1.0000.

Table 4
Incremental learning strategy for prediction and perception transfer models.

Hyperparameter	Description	LWD Parameter	Formation pressure gradient
<i>UpdateInterval</i>	Frequency of model updates based on new data.	Every 200m	Every 200m
<i>TriggerDepth</i>	Depth interval for triggering model retraining.	Every 1m	Every 1m
<i>WindowSize</i>	Number of recent data points used for model updates.	100	100
<i>LearningRate</i>	Learning rate for incremental updates.	0.0002	0.0003
<i>BatchSize</i>	Batch size for incremental learning.	48	24
<i>WeightDecay</i>	Regularization to reduce overfitting.	0.015	0.02
<i>DropoutRate</i>	Dropout rate to prevent overfitting.	0.3	0.35
<i>EarlyStoppingThreshold</i>	Patience for early stopping mechanism.	40	40

Similarly, the predictions for the loss and pore pressure gradients were outstanding, especially at depth intervals of 1261–1269 meters and 1694–1704 meters, where the pore pressure gradient achieved an R^2 of 1.0000. This validates the high efficiency and adaptability of the transfer model after the integration of new formation data.

Fig. 17 demonstrates the forgetting mechanism's impact on GR perception in an online learning environment for undrilled wells. Initially, significant discrepancies between predictions and actual values indicate effective forgetting of outdated data. In the 1800–2000 meter interval, the model's predictions closely match the actual geological data, showcasing enhanced adaptability and accuracy. This highlights the mechanism's role in maintaining real-time responsiveness and reliability in dynamic drilling conditions, ensuring accurate formation property predictions and supporting safe and efficient drilling operations.

Fig. 18(a-c) shows the exceptional online learning capabilities of the LWD parameter-and formation pressure gradient perception transfer models in dynamic geological environments. The performances of these models were verified through continuous evaluations following each training update, with average R^2 values exceeding 0.99, demonstrating their high accuracy in capturing and responding to formation changes. All parameters exhibited extremely low errors in the MSE metric, indicating high prediction precision. Although some formation pressure gradient parameters had higher Mean Absolute Percentage Errors (MAPE), such as an MAPE of 0.0187 for the collapse pressure gradient, this still emphasizes the strong applicability and reliability of the models in real drilling environments.

Continuous perception and real-time updates of the formation pressure gradient and LWD parameter perception transfer models significantly improved the adaptability of the models to unknown or complex geological conditions. Regular model evaluations and updates ensured accurate real-time predictions of formation properties during drilling operations, greatly reducing the risks posed by geological uncertainties. The effectiveness of the entire process, especially in continuous learning through the gradual integration of new formation information, ensures the accuracy and robustness of the models under complex formation conditions.

To assess broader applicability, the framework was validated across five exploration wells with varying geological complexities. The domain adversarial mechanism aligned feature distributions between the Caofidian 6-4-2-trained model and the target wells, while freezing-finetuning preserved essential geological knowledge. Online learning enabled continuous adaptation through updates at 200 m intervals. **Fig. E1** of Appendix E presents performance patterns, showing R^2 decreasing from 0.9952 at Caofidian 6-4-4 to 0.6749 at Ledong 10-1-1, inversely correlated with stratigraphic divergence in key formations. Detailed stratigraphic depth and thickness for these wells are provided in **Table E1** of Appendix E. These results quantitatively confirm the framework's scalability to heterogeneous geological environments.

5. Real-Time Multi-Objective Deep Reinforcement Learning Optimization for Drilling Parameters in Untapped Target Wells

To efficiently optimize the drilling parameters for undrilled target formations, a FCNN algorithm adjusted via PSO was designed to construct a ROP prediction model. Simultaneously, an online-trained iTransformer model based on adversarial domain adaptation, was used to sense the LWD parameters and formation pressure gradients. The ROP estimator provided accurate predictions through the PSO-FCNN model, whereas the LWD sensor provided operators with real-time critical geological information, such as lithological characteristics and fluid dynamics, thereby enhancing the precision of the drilling process control and decision-making. Additionally, real-time perception of the formation pressure gradient acted as a constraint on the mud density, ensuring wellbore stability and preventing formation fluid invasion. By providing real-time feedback on key pressure parameters, the mud

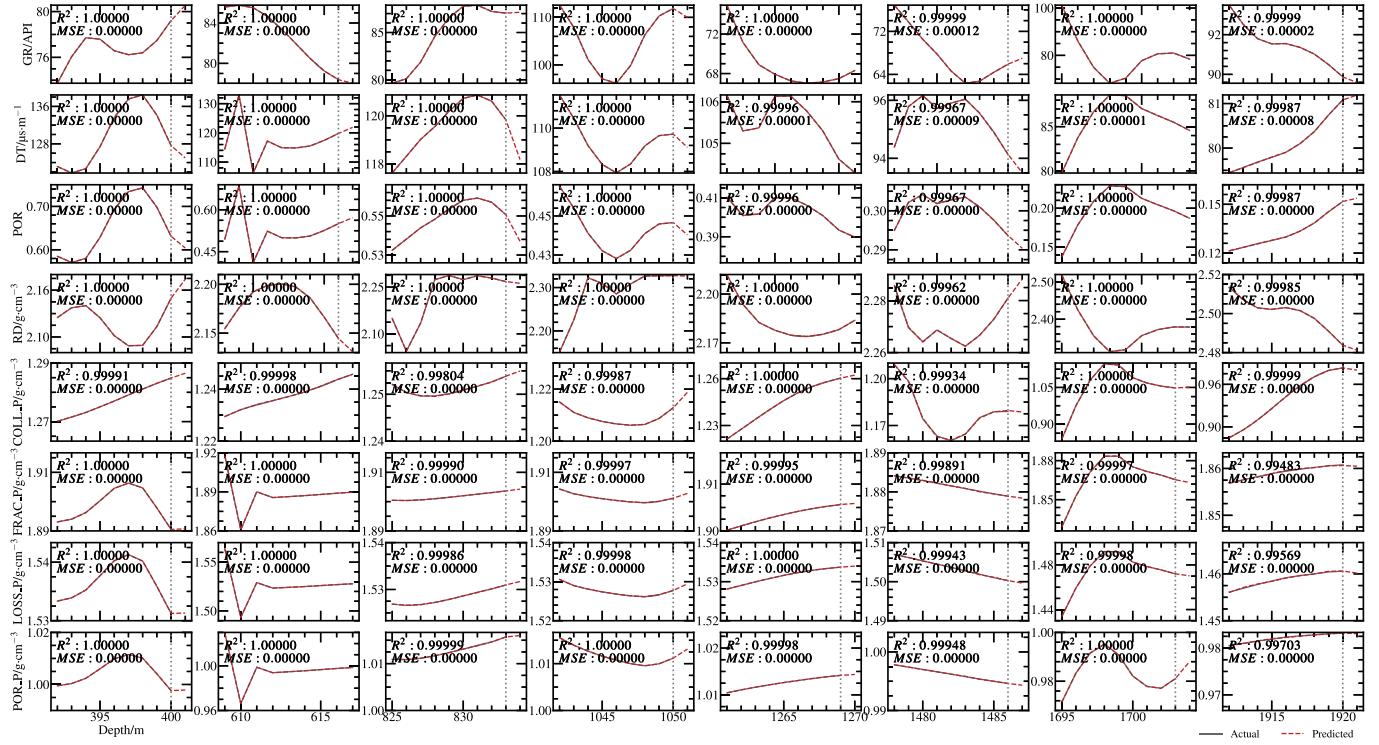


Fig. 16. Online learning performance of the formation property perception transfer model. Upper Panels: LWD parameters (GR, DT, POR, RD) over 11-meter intervals post-update. Lower Panels: Formation pressure gradients (collapse, fracture, loss, pore) over 11-meter intervals post-update.

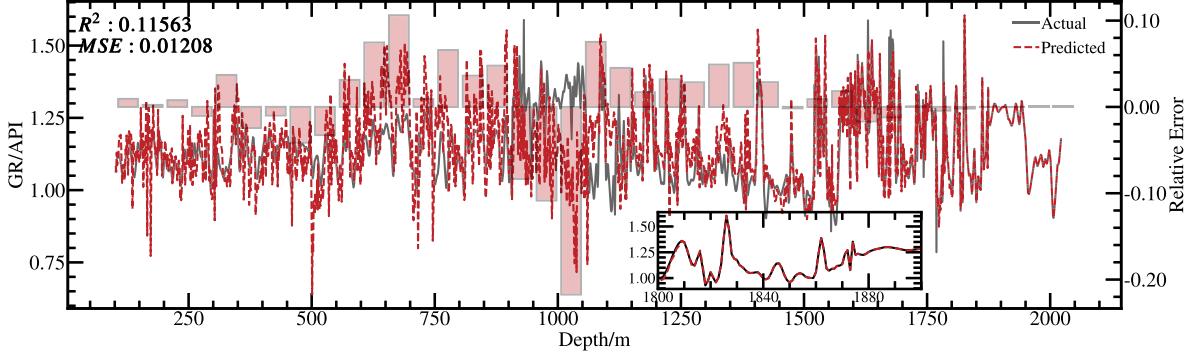


Fig. 17. Impact analysis of the forgetting mechanism on the perception of GR in untapped target wells in an online learning environment.

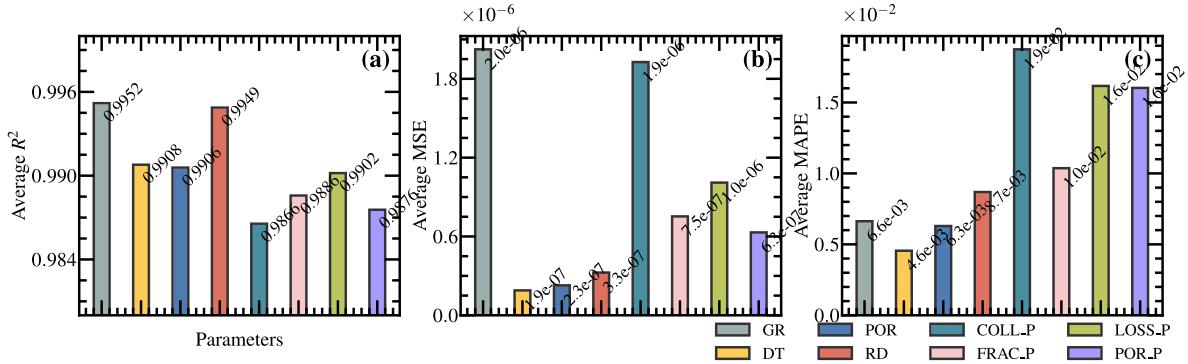


Fig. 18. Performance evaluation of the formation property perception transfer model in a dynamic geological environment.

density settings were optimized to avoid drilling issues caused by inappropriate mud density. The integration of these components formed a MDP for drilling parameter adjustments, allowing for the dynamic optimization of drilling parameters based on continuous data streams. This enabled real-time adjustments to the drilling operation to adapt to changing geological conditions and optimize drilling efficiency.

5.1. Design of Multi-Objective Optimization Functions

5.1.1. Unit Cost per Foot Function

In the drilling process, the unit cost per foot function is a core indicator for evaluating the drilling economic efficiency, with the key being the accurate measurement of costs and efficiency during drilling[57]. The expression for the unit cost per foot is given by Eq. 9.

$$C_{pm} = \frac{C_r \left[\left(\frac{C_b}{C_r} + t_f \right) \cdot A_f (a_1 n + a_2 n^3) + \left(h_f + \frac{C_1}{2} h_f^2 \right) \right]}{v_{pc} \cdot \left(h_f + \frac{C_1}{2} h_f^2 \right)} \quad (9)$$

where C_{pm} is the cost per foot, (CNY/m); C_b is the cost of the drill bit, (CNY); C_r is the rig operation fee, (CNY/hr); t_f is the time required to make trips and connections, (hr); A_f is the abrasiveness coefficient of the formation; n is the RPM, (r/min); a_1, a_2 are the speed influence coefficients determined by the type of drill bit; h_f denotes the amount of wear corresponding to the drill-bit lifespan; v_{pc} is the ROP prediction transfer model. The detailed derivation of Eq. 9 is provided in Appendix D.

This cost per foot function illustrates the combined impact of drill bit cost and rig operating fees on the unit cost of drilling. This provides a precise method for calculating drilling costs, thus supporting scientific and accurate decision-making in drilling operations.

5.1.2. Mechanical Specific Energy Function

The mechanical specific energy (MSE) model correlates the energy required to break a unit volume of rock with the efficiency of the bit in breaking the rock, and serves as a critical method for quantifying drilling efficiency. The MSE is calculated using parameters such as the ROP, WOB, RPM, torque, and bit diameter. Higher MSE values indicate lower drilling efficiency and poorer adaptability of the bit to the formation, suggesting that the drilling parameters need to be optimized [58]. The ideal-state MSE is calculated using Eq. 10.

$$MSE = E_f \left(\frac{4W}{\pi D_b^2} + 0.16 \frac{n\mu W}{D_b v_{pc}} \right) \quad (10)$$

where MSE is the mechanical specific energy, (MPa); W is the WOB, (kN); D_b is the drill bit diameter, (m); n is the RPM, (r/min); μ is the sliding friction coefficient of the drill bit; v_{pc} is the ROP prediction transfer model. The detailed derivation of Eq. 10 is provided in Appendix D.

5.1.3. Objective Function

Building upon single-objective optimization, a multi-objective optimization model is constructed that considers both drilling costs and rock-breaking energy consumption based on an empirical equation for ROP. This model aims to optimize the WOB, RPM, FLW, and MWIN to achieve the lowest unit cost per foot and the minimum MSE.

The objective function, also known as the evaluation index, serves as the criterion for assessing the quality of the drilling design schemes. In practical drilling operations, the goal is to maximize the drilling efficiency while minimizing costs. Therefore, the lowest unit cost per foot and the minimum MSE were selected as the objective functions for the multi-objective optimization. The objective function model F is expressed in Eq. 11.

$$\begin{cases} F = \min \{ \min(C_{pm}), \min(MSE_m) \} \\ \min(C_{pm}) = \frac{C_r \left[\left(\frac{C_b}{C_r} + t_f \right) \cdot A_f (a_1 n + a_2 n^3) + \left(h_f + \frac{C_1}{2} h_f^2 \right) \right]}{v_{pc} \cdot \left(h_f + \frac{C_1}{2} h_f^2 \right)} \\ \min(MSE_m) = E_f \left(\frac{4W}{\pi D_b^2} + 0.16 \frac{n\mu W}{D_b v_{pc}} \right) \end{cases} \quad (11)$$

5.2. Markov Decision Process Environment

5.2.1. Reward Function Design

In the deep reinforcement learning framework, the reward function plays a crucial role by providing immediate feedback to guide decision optimization [59]. It was specifically designed to evaluate the effects of adjustments to drilling parameters, primarily the unit cost per foot and MSE rewards. The unit cost per foot reward reflects the economic efficiency of the drilling costs relative to progress, whereas the MSE reward focuses on the energy efficiency and mechanical load during the drilling process. These rewards enable the agent to find the optimal balance in adjusting the drilling parameters, thereby maximizing the cost-effectiveness and equipment efficiency.

The unit cost per foot reward evaluates the economic benefit of adjusting drilling parameters based on changes in the unit cost per foot. This measure reflects the relationship between the bit costs, rig operation expenses, and drilling progress. If the operation results in a reduced unit cost per foot, the reward is positive; conversely, if the costs increase, the reward is negative. This design allows the reward function to directly reflect changes in economic efficiency and promotes cost-effectiveness.

The MSE reward evaluates the energy efficiency and mechanical load of the drilling operation through mechanical specific energy. A reduction in MSE indicated improved drilling efficiency and reduced risk of equipment wear; thus, operations that reduced MSE received positive rewards, whereas those that increased MSE received negative rewards. This reward design emphasizes the importance of energy efficiency, promoting the rational use of resources and the long-term stability of equipment. The reward function model is expressed in Eq. 12.

$$\begin{cases} R_1 = -(C_{pm} - C_{pm,prev}) \\ R_2 = -(MSE_m - MSE_{m,prev}) \end{cases} \quad (12)$$

where R_1 represents the improvement reward for unit cost per foot, reflecting the cost change from the previous to the current iteration; C_{pm} is the unit cost per foot for the current iteration; $C_{pm,prev}$ is the unit cost per foot for the previous iteration. R_2 represents the improvement reward for mechanical specific energy, assessing the change in MSE from the previous to the current iteration; MSE_m is the MSE for the current iteration; $MSE_{m,prev}$ is the MSE for the previous iteration.

The combination of these two rewards forms a multi-objective optimization framework, allowing the deep reinforcement learning algorithm to determine the optimal balance between cost and efficiency, thereby optimizing the drilling parameters.

5.2.2. State Space and Action Space

In the framework of deep reinforcement learning applied to optimize drilling operations, the definition of the state space depends on the key control parameters of the drilling process, including WOB, RPM, FLW, and MWIN. Real-time measurements of these parameters constitute the current state during the drilling process, providing the necessary data foundation for monitoring and adjusting drilling operations. The action space defines the range of the adjustable drilling parameters and serves as the operational domain in which the intelligent agent adjusts the drilling performance. By implementing specific actions, that is fine-tuning the drilling parameters, the agent directly influences the drilling state. The results of these actions, along with the changes in the

drilling state, are fed back to the agent through a well-defined reward function. The state and action spaces are expressed using Eqs. 13 and 14, respectively.

$$S = \{w, r, f, m\}$$

$$\left\{ \begin{array}{l} w \in [w_{min}, w_{max}] \\ r \in [r_{min}, r_{max}] \\ f \in [f_{min}, f_{max}] \\ m \in [m_{min}, m_{max}] \end{array} \right. \quad (13)$$

$$A = \{\Delta w, \Delta r, \Delta f, \Delta m\}$$

$$\left\{ \begin{array}{l} \Delta w \in [-\delta w, \delta w] \\ \Delta r \in [-\delta r, \delta r] \\ \Delta f \in [-\delta f, \delta f] \\ \Delta m \in [-\delta m, \delta m] \end{array} \right. \quad (14)$$

where w, r, f, m represent the current values of the four drilling parameters, each with its own permissible limits; $\Delta w, \Delta r, \Delta f, \Delta m$ denote the corresponding changes in these parameters; $\delta w, \delta r, \delta f, \delta m$ are the maximum allowable adjustments in the drilling operation.

Precise control of the MwIN is crucial for maintaining drilling safety and efficiency. Therefore, they must be accurately regulated to match the mechanical and hydrodynamic properties of the formation. The key to setting the MwIN is to ensure that it exceeds both the formation pore and collapse pressures, thereby preventing the backflow of formation fluids and maintaining wellbore stability to avoid collapse. Simultaneously, the MwIN must be lower than the fracture and loss pressures to prevent the formation of fractures and mud loss when these pressure values are exceeded. Thus, m_{min} and m_{max} are defined by Eq. 15.

$$\left\{ \begin{array}{l} m_{min} = \max(\rho_{por}, \rho_{coll}) \\ m_{max} = \min(\rho_{frac}, \rho_{loss}) \end{array} \right. \quad (15)$$

where m_{min} and m_{max} represent the minimum and maximum allowable MwIN under the current operational conditions, respectively, set ac-

cording to the dynamic formation pressure gradient conditions. m_{min} is defined as the greater formation pore pressure gradient ρ_{por} and the collapse pressure gradient ρ_{coll} to ensure that the MwIN is sufficient to prevent the backflow of formation fluids and maintain wellbore stability. m_{max} is set as the lower fracture pressure gradient ρ_{frac} and loss pressure gradient ρ_{loss} . The values of ρ_{por} , ρ_{coll} , ρ_{frac} , and ρ_{loss} are determined using the iTransformer model.

5.2.3. Learning Mechanism

To enhance the decision quality and efficiency of drilling operations, a multi-objective deep reinforcement learning algorithm, DPOAC, was specifically designed for drilling parameter optimization. This algorithm is based on the soft actor (SAC) framework and is customized to meet the optimization needs of continuous action spaces, adapting to the complex decision environment and dynamic operational requirements of drilling [60]. Fig. 19 shows the architecture of the DPOAC algorithm.

Online Q-Network and Target Q-Network: The online Q-network and target Q-network are responsible for evaluating the reward values in the drilling state, including the WOB, RPM, FLW, and MwIN. By analyzing the drilling parameters in real time, these networks output the expected rewards, focusing on metrics such as the unit drilling cost and mechanical specific energy. DPOAC uses a double Q-network structure, which significantly reduces reward estimation bias and improves decision quality and policy efficiency. The online Q-network is continuously updated to optimize the relationship between the drilling cost and performance, whereas the target Q-network periodically synchronizes the parameters from the online Q-network to maintain training stability. This dual-network synchronization strategy, combined with intermittent parameter updates, not only solidifies the learning objectives, but also mitigates fluctuations during the training process, thereby enhancing the overall performance and predictive reliability of the model in a dynamically complex drilling environment.

Policy Network: The policy network significantly enhances the exploration efficiency through entropy regularization, preventing premature convergence to suboptimal solutions. This strategy receives drilling parameters, such as WOB, RPM, FLW, and MwIN, and outputs

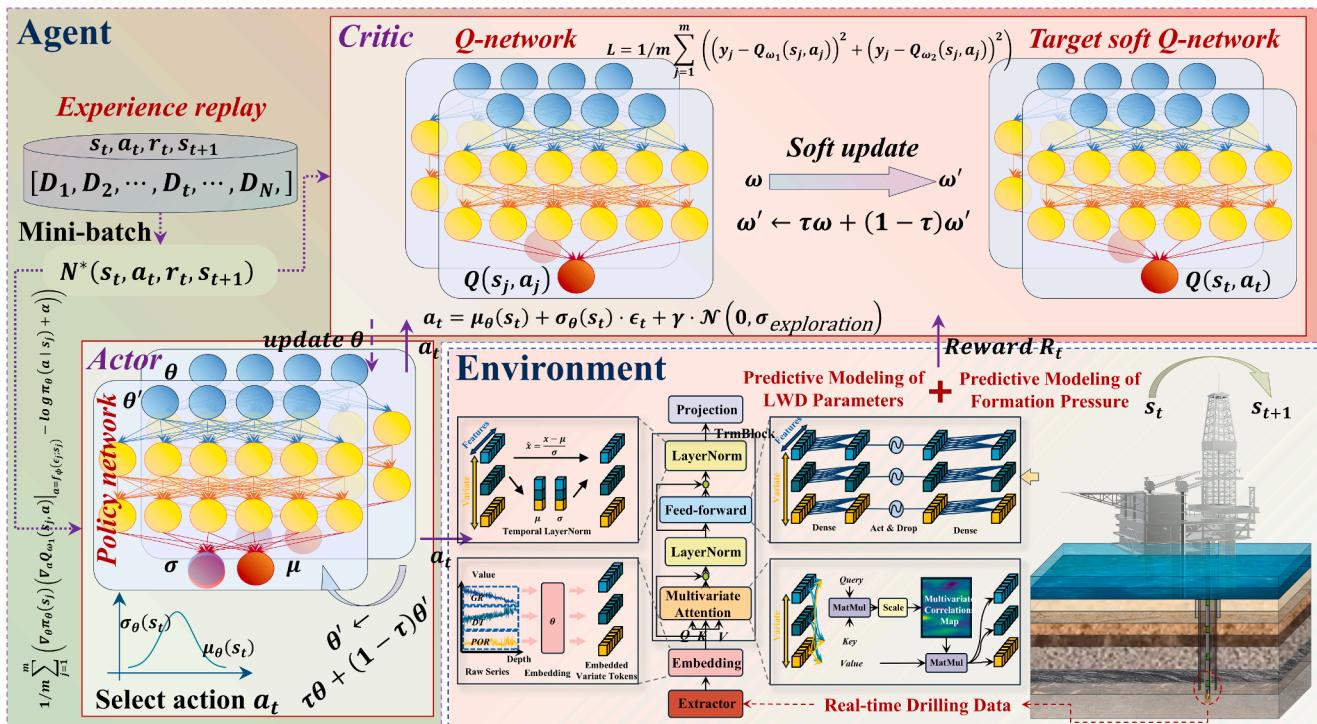


Fig. 19. Architecture of the DPOAC algorithm for drilling parameter optimization.

the probability distribution of actions. Combining expected reward maximization and behavior entropy maximization increases the behavior diversity and expands the exploration state space. The DPOAC algorithm adopts a double Q-network structure, reducing the bias from overestimation and improving the learning efficiency. Each update uses the smaller of the two Q-values for optimization, reducing the reward estimation fluctuations and bias. Regular synchronization between the online Q-network and the target Q-network enhances learning goal stability and optimizes long-term reward accumulation by balancing exploration and exploitation. The policy network updates strategies through interactive feedback, adjusts action selection using accumulated experience, ensures precise and effective decisions in complex drilling environments, and significantly improves the cost efficiency and safety of drilling operations.

Prioritized Experience Replay: Prioritized experience reinforces the learning of critical experiences through a sampling mechanism based on high temporal difference (TD) errors, accelerating error correction, and performance optimization. This mechanism was particularly suitable for continuous action spaces in drilling operations, ensuring that the model learned from significant transitions to improve the decision quality. The DPOAC algorithm increases behavior diversity through entropy regularization and minimizes bias by updating with the smaller of the two independent Q-values, thereby enhancing learning stability and efficiency. This combination of entropy regularization and prioritized experience replay not only optimizes the balance between exploration and exploitation, but also improves model adaptability and decision accuracy in complex drilling scenarios by dynamically adjusting the sampling strategy of the training data. The prioritized experience replay ensures that the model could effectively learn and adapt to new operating conditions in changing drilling environments, thereby significantly enhancing the cost efficiency and safety of drilling operations.

Exploration Noise Strategy: The exploration noise strategy, integrated into the policy network through entropy regularization, significantly enhances the exploration efficiency of the drilling parameter optimization process. This strategy dynamically adjusts the exploration behavior based on the current operating environment and decision stage. A unique aspect of DPOAC is its ability to maximize behavioral entropy to encourage diversity, effectively avoiding premature convergence to suboptimal solutions. Additionally, DPOAC utilizes a double Q-network structure, updating the policy using the smaller of two independent Q-value estimators, significantly reducing the risk of overestimation and improving learning accuracy and efficiency. The exploration noise added to the policy network was not merely random, but structurally integrated information closely related to the current policy performance, thereby more accurately guiding the exploration process and expanding the exploration of the state space. This finely tuned balance of exploration and exploitation, along with the periodic synchronization of the online Q-network and target Q-network parameters, ensures the stability and efficiency of the algorithm in long-term reward accumulation.

To enhance the DPOAC algorithm's performance in drilling parameter optimization and minimize the objective function, a reparameterization method is integrated and defined by Eq. 16.

$$a_t = \mu_\theta(s_t) + \sigma_\theta(s_t) \cdot e_t + \gamma \cdot \mathcal{N}(0, \sigma_{\text{exploration}}) \quad (16)$$

where $\mu_\theta(s_t)$ is the mean output of the policy network; $\sigma_\theta(s_t)$ is the standard deviation output of the policy network; e_t is a noise term sampled from the standard normal distribution $\mathcal{N}(0, 1)$; γ is the discount factor for future reward weighting; $\mathcal{N}(0, \sigma_{\text{exploration}}^2)$ is an additive Gaussian exploration noise with zero mean and adjustable standard deviation $\sigma_{\text{exploration}}$.

This formulation ensures differentiable policy sampling, enabling gradient-based optimization in complex drilling environments. The

differentiable action generation guarantees adaptive policy adjustments during real-time drilling operations, critical for maintaining stability under dynamic geological conditions. The complete process for optimizing the drilling parameters using the DPOAC algorithm is presented in Table C1 of Appendix C.

5.3. DPOAC Algorithm

5.3.1. Training of Q-, Target, and Policy Networks

For the drilling operations in the unpenetrated formations of the Caofidian 6-4 block in the Bohai Sea, China, Table 5 lists the coefficient settings required for the unit cost per foot function and mechanical specific energy function. The precise configuration of these coefficients is crucial for calculating and optimizing drilling cost efficiency.

For the optimization of the drilling parameters using the DPOAC algorithm, the training hyperparameters were set as listed in Table 6. Each parameter's specific function and selected value significantly affects the algorithm's performance and convergence speed.

In the training of the drilling parameter optimization, the Q-network was optimized by minimizing the loss function, which quantified the deviation between the actual and predicted rewards based on the current drilling state and action. The policy network derived the optimal action strategy from the current state, adjusting key drilling parameters, such as the WOB and RPM, to minimize the unit cost per foot and mechanical specific energy. The expected reward from these actions is used to estimate the values in each Q network, forming the loss of the policy network. This was optimized using gradient descent to adjust the network parameters and ensure policy effectiveness and adaptability. The target network adopts a soft update strategy, inheriting parameters from the Q-network and using the tau factor to control the proportion of parameter transfers from the main network. This is aimed at providing stable target values for calculating the loss function, performing gradient updates, reducing training fluctuations, and enhancing convergence. This strategy helps smooth the training process, stabilizes learning, and optimizes the system performance and reliability.

The role of the Q-network in drilling operations is extended to evaluate the immediate benefits and long-term cost-efficiency trade-offs. Through periodic network parameter adjustments and real-time reward feedback, the deep reinforcement learning model refines the decision-making process over continuous training cycles. This method allows the model to dynamically adjust drilling strategies, adapt to changes in the actual drilling environment, and optimize the overall drilling efficiency and cost management. Fig. 20 shows the training loss values of the Q1-, Q2-, and policy networks under the four different learning rate settings. The loss curves reflect the performance and convergence

Table 5

Coefficients for drilling cost and mechanical specific energy functions in the Caofidian 6-4 block, Bohai, China.

Application	Parameter	Description	Value
Unit Cost per Foot	C_b	Purchase cost of the drill bit.	900
	C_r	Operating cost per hour of the rig.	250
	h_f	Tooth wear of the bit at the end of its life.	0.6
	t_f	Time for tripping and connecting single joints.	102
	A_f	Effect of formation on bit wear.	0.00228
	a_1	Linear coefficient for the impact of RPM on drilling efficiency.	1.5
	a_2	Cubic coefficient for the impact of RPM on drilling efficiency.	6.53E-05
	C_1	Coefficient for reducing bit tooth wear rate.	5
	E_f	Energy conversion efficiency of the bit during rock breaking.	0.35
	μ	Sliding friction coefficient when the bit contacts the rock.	0.5
Mechanical Specific Energy			

Table 6

Learning strategies for the DPOAC algorithm.

Hyperparameter	Description	Value
<i>ActorLearningRate</i>	Learning rate for updating the Actor Network.	0.000010
<i>Q1NetLearningRate</i>	Learning rate for updating Q1-network parameters.	-
<i>Q2NetLearningRate</i>	Learning rate for updating Q2-network parameters.	-
<i>DiscountFactor</i>	Discount factor for weighting current vs future rewards.	0.7
<i>TauSoftUpdate</i>	Tau parameter for soft updates in target networks.	0.7
<i>LossWeights</i>	Weights for balancing importance among optimization targets in the loss function.	0.5,0.5
<i>BatchSize</i>	Number of samples drawn from the experience replay buffer per batch.	2048
<i>BufferCapacity</i>	Maximum capacity of the experience replay buffer.	16000
<i>MaxEpochSteps</i>	Maximum number of steps per epoch.	1000
<i>AlphaPriority</i>	Exponent parameter for calculating priority in experience replay.	0.6
<i>BetaStart</i>	Initial importance sampling parameter in experience replay.	0.4
<i>BetaEnd</i>	Final importance sampling parameter in experience replay.	1
<i>BetaDecaySteps</i>	Number of steps over which the importance sampling parameter decays in experience replay.	500
<i>NoiseStd</i>	Standard deviation of exploration noise	0.1
<i>EntropyTemp</i>	Temperature parameter for entropy regularization	0.01
<i>TargetUpdateFreq</i>	Frequency for soft updates of target networks	100
<i>RewardDecayFactor</i>	Decay factor for future reward weighting	0.99

characteristics of each network during the training phase.

Fig. 20 illustrates the impact of different learning rate settings on the Q1-, Q2-, and policy networks. For the Q1-network, all learning rate combinations resulted in a loss value that gradually converged from a high initial value to a value near zero. Notably, the setting of $a = 1.0 \times 10^{-5} c = 2.0 \times 10^{-5}$ achieved the smallest convergence loss, demonstrating the best training effect. This highlights the critical role of an appropriate learning rate in optimizing Q1-network performance. Similarly, for the Q2-network, the same learning rate settings exhibited an excellent performance, with the training loss value converging to the lowest point. This indicates that higher learning rates contribute to the rapid achievement of ideal loss levels. The loss values of the policy network showed a trend of transitioning from positive to negative values. The setting of $a = 5.0 \times 10^{-6} c = 1.0 \times 10^{-5}$ achieved the smallest negative value, suggesting a potential risk of overfitting. In contrast, the combination of $a = 2.5 \times 10^{-6} c = 5.0 \times 10^{-6}$ showed a slower convergence rate but exhibited more stable training loss values, potentially better suited for generalization needs. These results emphasize the decisive role of learning rate settings in the performance of the DPOAC algorithm, particularly in adapting to complex drilling environments. A reasonable adjustment of the learning rate significantly affects the convergence speed, performance, and stability of the model.

5.3.2. Drilling Parameter Optimization Process

In the process of drilling parameter optimization, Fig. 21 illustrates the variation in the unit cost per foot and mechanical specific energy as the drilling depth ranges from 102 to 2021 m. The results for each depth were recorded after 3000 iterations. Notably, the Pareto points are identified in the figure, indicating the positions where the optimal balance between cost and efficiency was achieved, distinguishing these points from the non-Pareto points.

Fig. 21 shows the trends of the unit cost per foot and mechanical specific energy as the drilling depth increased from 102 to 2021 m. Initially, both metrics fluctuated significantly, reflecting the exploratory nature of the parameter adjustments during the early drilling stages. During this phase, the Pareto boundary changed frequently, indicating that the DPOAC algorithm attempts to find an optimal balance between cost and efficiency. As drilling progressed to intermediate depths (600–1200 m), fluctuations in both metrics decreased, and the Pareto boundary stabilized, signifying a shift from broad exploration to focused exploitation of operational parameters. In later stages (beyond 1200–2021 m), despite the stabilization of unit costs and mechanical efficiency, the Pareto boundary continued to show changes, likely reflecting severe formation pressure gradient fluctuations, particularly in areas with potential abnormal pressure zones. These adjustments in drilling parameters, necessary to adapt to changing pressures, highlight the ongoing need to dynamically fine-tune operations and balance the economic and efficiency considerations in drilling operations. This iterative process demonstrates the capability of the DPOAC algorithm to dynamically adjust and optimize the drilling parameters, driving towards optimal cost-effectiveness and mechanical efficiency through sustained exploration and exploitation.

Fig. 22 illustrates the fluctuations in the WOB across different drilling depths, where the post-optimization values of the WOB exhibit notable variability. This variability underscores the response of the algorithm to the uncertainties encountered in real drilling scenarios while searching for the optimal WOB settings. In particular, during the initial phases of drilling, the optimized WOB values adapt to diverse geological conditions, adjusting to variations in the formation hardness and fracturability, reflecting the dynamic adaptation of the algorithm to changing drilling layers.

Fig. 23 and 24 illustrate the variations in the RPM and FLW across different drilling depths. Compared to the initial settings, revolutions per minute increased by an average of 17.5% throughout the drilling process, whereas the pump flow rate initially increased by an average of 13.5% but subsequently exhibited a declining trend in the mid-phases, occasionally dropping below the initial settings before rising again in the later stages. These fluctuations demonstrate the adaptation of the optimization strategy to the formation hardness and other drilling conditions. Particularly in hard formations, increasing the revolutions per minute and pump flow rate significantly improves the ROP and wellbore cleanliness; increased power and cleaning capability are crucial to prevent bit wear and mitigate issues related to cuttings accumulation.

Fig. 25 shows the adjustments in MwIN across various drilling depths, with an average increase of approximately 22%. This emphasizes the criticality of meticulous mud density management throughout

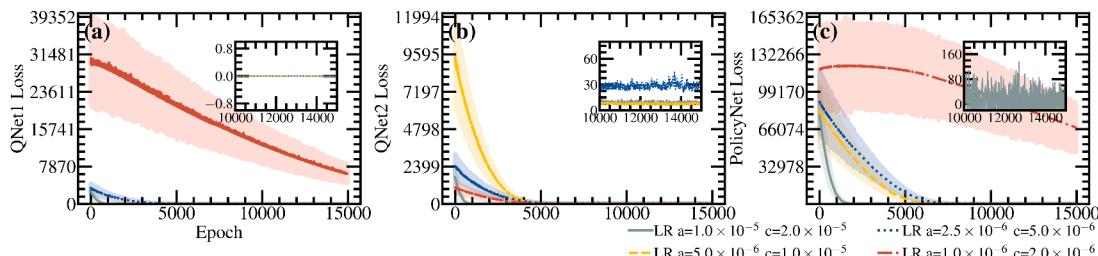


Fig. 20. Dynamics of training losses for the Q1-, Q2-, and policy networks under different learning rate settings.

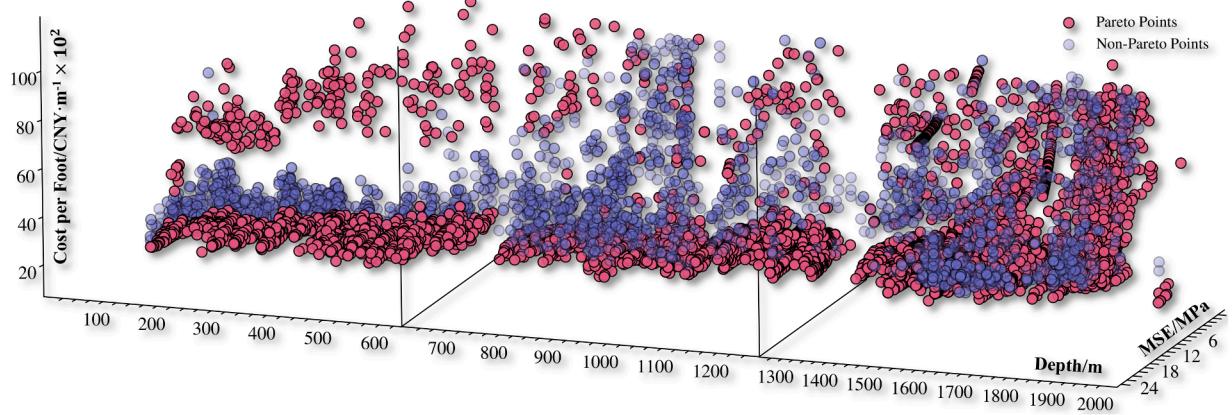


Fig. 21. Dynamics of unit cost per foot and mechanical specific energy with drilling depth and Pareto frontier.

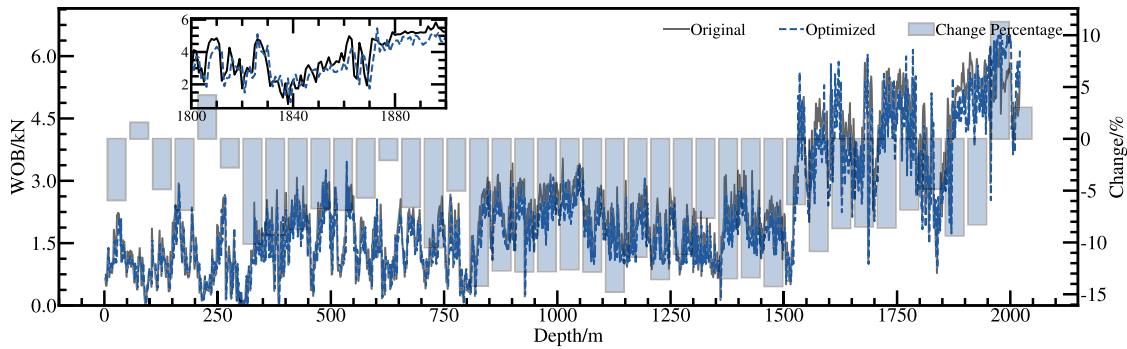


Fig. 22. Optimization changes in weight on bit (WOB) with drilling depth.

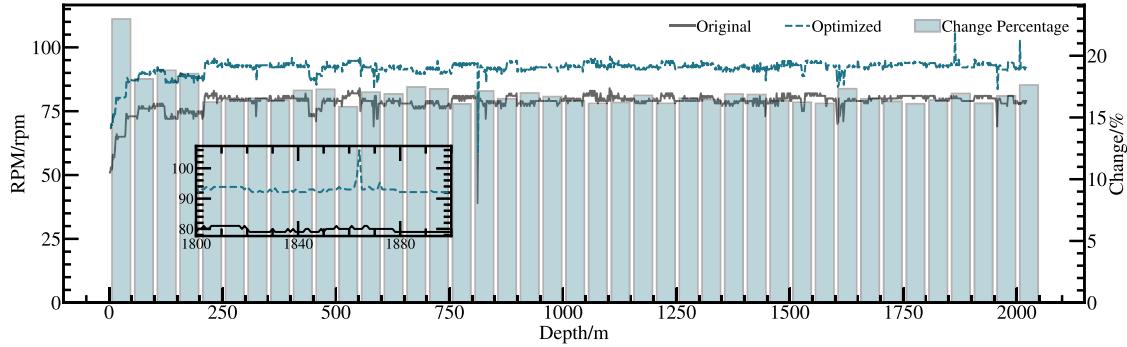


Fig. 23. Optimization changes in revolutions per minute (RPM) with drilling depth.

the drilling process. This systematic increase primarily aims to accommodate dynamic formation conditions and fulfill wellbore stability requirements. Given the stringent constraints imposed by the formation pore, collapse, fracture, and loss pressures, the mud density adjustments were precisely regulated within a safe operational window. This is pivotal for maintaining wellbore stability and optimizing the transport efficiency of the cuttings. By maintaining the mud density within this safe threshold, drilling operations can effectively prevent issues, such as wellbore collapse triggered by abnormal formation pressure gradient. This approach safeguards the wellbore and ensures continuity and efficiency of drilling operations. Fine-tuning the mud density facilitates a balance in wellbore pressures, enhances the interaction between the drill bit and formation, minimizes bit wear, and boosts ROP.

Fig. 26 demonstrates the effects of the DPOAC algorithm on the drilling responses to hard rock formations at depths between 1510 and 1530 m, highlighting the changes in key logging parameters, drilling parameters, and performance indicators. **Fig. 26(a-d)** show significant increases in GR, notable decreases in DT, reductions in POR, and increases in RD, which are indicators typically associated with encountering hard rock layers that demand precise and adaptive drilling strategies. As drilling reached a depth of 1521 m, **Fig. 26(e-g)** show an increased WOB and RPM, along with a reduced FLW. These algorithm-driven adjustments optimized the drilling parameters to efficiently address the unique challenges of hard rock layers, enhance rock-breaking efficiency, and reduce the risk of operational complications from excessive mud flow velocities. Economically, **Fig. 26(h-j)** indicated

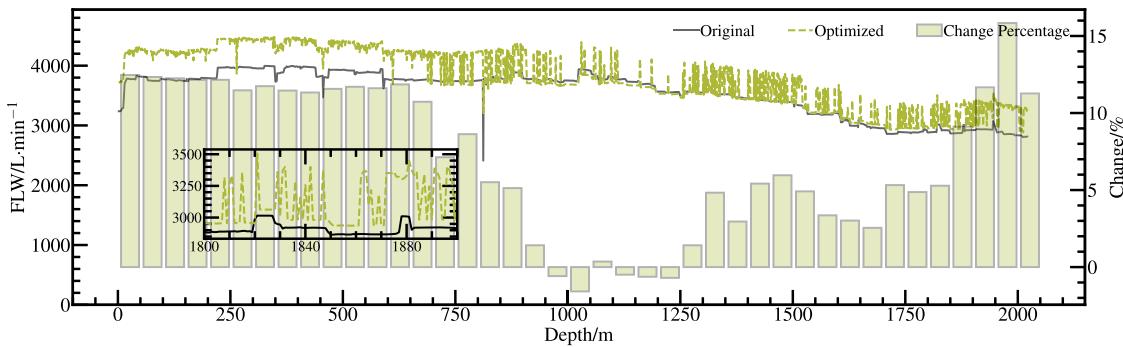


Fig. 24. Optimization changes in pump flow rate (FLW) with drilling depth.

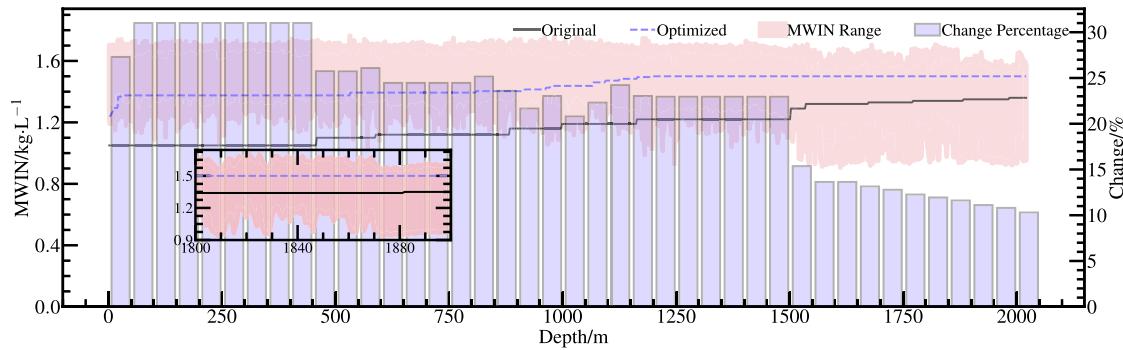


Fig. 25. Optimization changes in mud input density (MwIN) and wellbore stability constraints with drilling depth.

increases in the unit cost per foot and mechanical specific energy, along with a decrease in ROP, illustrating the direct impact of hard rock layers on drilling operations. Despite the initial cost increase when encountering hard rock, the adjustments made by the DPOAC algorithm resulted in substantial cost reductions and enhanced ROP compared with the pre-optimization settings, underscoring the efficacy of the algorithm in optimizing drilling operations under complex geological conditions. The continuous learning and adaptive strategy adjustments of the DPOAC algorithm dynamically respond to geological variations, optimizing the drilling efficiency and cost-effectiveness. Under challenging conditions, such as hard rock layers, the algorithm not only refines the operational parameters but also ensures the safety and economic viability of the drilling process, affirming its pivotal role in modern drilling operations.

Fig. 27 demonstrates the effectiveness of the DPOAC algorithm in optimizing key performance metrics, such as the unit cost per foot, mechanical specific energy, and ROP, within a multi-parameter drilling environment. The application of the DPOAC algorithm resulted in a substantial reduction of approximately 75% in the unit cost per foot, a 10% decrease in mechanical specific energy, and an astonishing 350% increase in ROP. These results not only confirm the ability of the DPOAC algorithm to adjust drilling parameters dynamically to suit variable geological conditions but also underscore its proficiency in slashing drilling costs and boosting operational efficiency.

Fig. 28 illustrates the effectiveness of the DPOAC algorithm in reducing CO₂ emissions per meter drilled throughout the drilling interval. Compared to the initial drilling conditions, the DPOAC algorithm achieved an average CO₂ emission reduction ranging from 5.57% to 16.97% across various depth intervals, with notable reductions exceeding 10% frequently observed beyond 800 m. Although minimal improvements occurred at shallow depths (2-250 m), substantial emission reductions emerged consistently from mid to deeper intervals (850-

2021 m), peaking at approximately 17%. These results highlight the capability of the DPOAC algorithm to dynamically adjust drilling parameters, significantly enhancing environmental sustainability while maintaining operational efficiency. The detailed derivation of the CO₂ reduction model is provided in Appendix D.

Fig. 29 shows how bit wear changes as drilling depth increases, comparing results before and after applying the DPOAC optimization algorithm. Before optimization, improper control of the RPM during drilling caused high bit wear per meter, averaging about 0.002 mm/m. Consequently, cumulative wear quickly reached the safety threshold of 0.6 mm, requiring frequent bit replacements—a total of seven times. After applying the DPOAC algorithm, the system dynamically adjusted drilling parameters such as RPM in real-time according to formation conditions. As a result, bit wear per meter dropped significantly to about 0.0003 mm/m on average. Consequently, cumulative wear decreased dramatically, requiring only one bit replacement during the entire drilling process. Specifically, cumulative wear at 340 m was just 0.084 mm, and at 600 m it increased only to 0.147 mm—far below the values observed before optimization. This substantial improvement demonstrates that the DPOAC algorithm effectively reduces the bit wear rate through precise control of bit rotational speed, significantly extending bit life, reducing replacement frequency, and greatly enhancing drilling efficiency and cost-effectiveness. The wear rate model of the drill bit is provided in Appendix D.

In addition to direct unit cost reductions, the DPOAC algorithm provided notable long-term economic benefits through reduced maintenance expenditures. Decreasing drill bit replacements from seven to one significantly lowered equipment downtime and related labor, logistics, and procurement costs compared to historical methods. Additionally, the cumulative reduction of 55.846 kg in CO₂ emissions further enhanced economic efficiency by reducing environmental compliance and carbon footprint-related expenses.

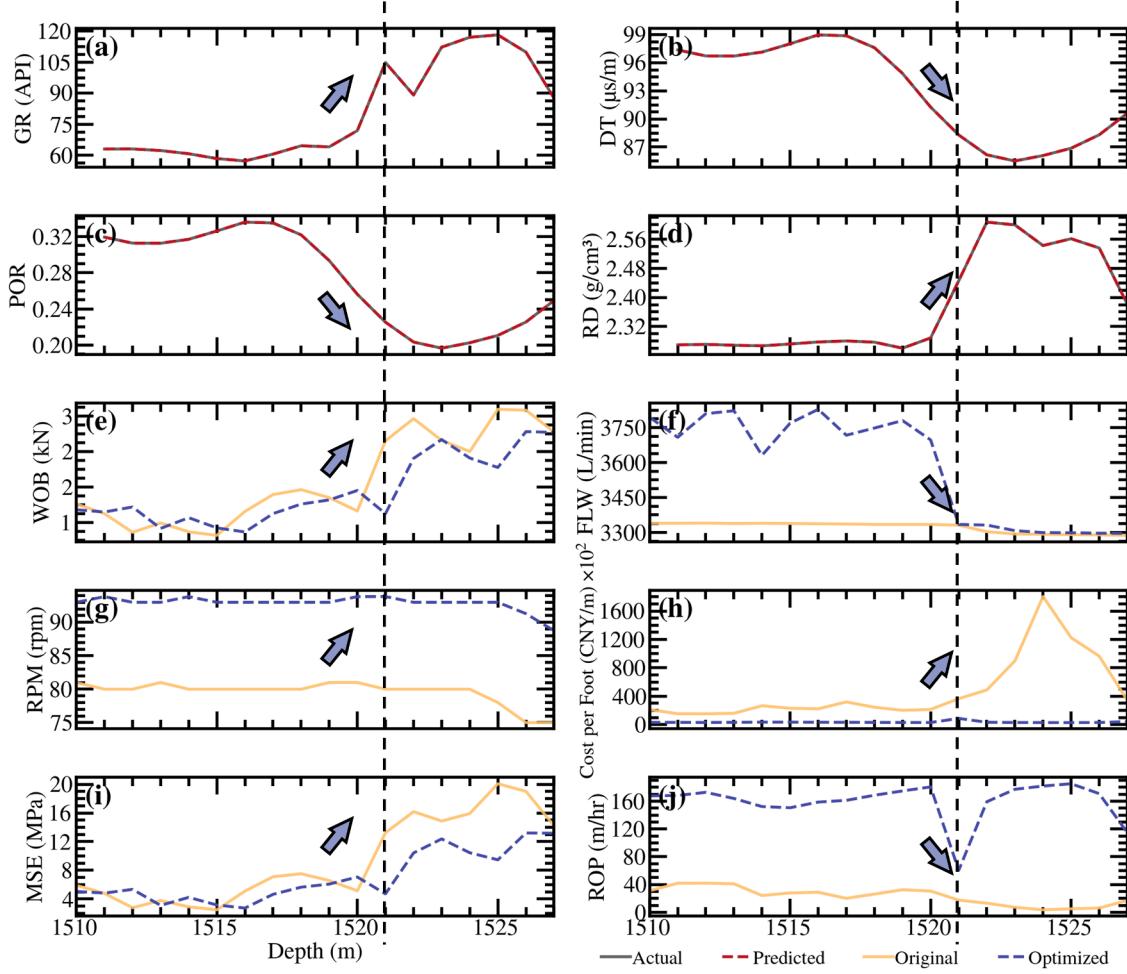


Fig. 26. Analysis of drilling response and DPOAC algorithm adjustments in the hard rock section from 1510 to 1530 m.

5.3.3. Drilling Parameter Optimization Performance Validation

Fig. 30 shows a comprehensive comparison of different algorithms in drilling optimization, focusing on the performance of the DPOAC algorithm versus the Non-dominated Sorting Genetic Algorithm II (NSGA-II) across four crucial drilling stages. In Fig. 30(a), a comparison of the ROP reveals a significant enhancement by the DPOAC algorithm in the early depth stages, indicating its effective broad exploration of the initial drilling parameters and rapid adaptation to geological conditions. As the drilling depth increased, the DPOAC continuously optimized the ROP and achieved stable and efficient performance, markedly surpassing that of the NSGA-II algorithm. This optimization of ROP demonstrates the capability of the DPOAC algorithm to ensure drilling efficiency while effectively managing drilling risks and reducing mechanical wear or operational errors caused by improper parameter selection.

Fig. 30(b) presents a comparison of the mechanical specific energies. Although the reduction in the mechanical specific energy by the DPOAC algorithm was not pronounced in the early depth stages, its optimization effects became increasingly apparent with depth. This trend underscores the adaptability and precision of the DPOAC algorithm under dynamic geological conditions, which is capable of fine-tuning drilling parameters to reduce energy consumption and enhance drilling machinery efficiency. Regarding the unit cost per foot, Fig. 30(c) shows that both the DPOAC algorithm and NSGA-II achieved significant cost reductions, with the performance of the DPOAC algorithm proving to be superior. This not only highlights the efficiency of the DPOAC algorithm in cost control but also reflects its advantages in maintaining the economic viability of drilling operations.

Fig. 30(d) shows the significant differences in efficiency between the

NSGA-II and DPOAC algorithms. As a heuristic algorithm, the NSGA-II algorithm relies on complex iterative processes, often requiring thousands of iterations to identify optimal drilling parameters, resulting in operational times that are more than ten times those of the DPOAC algorithm. Although this iterative approach explores a broad parameter space, its high time cost significantly reduces drilling efficiency in modern operations that demand rapid responses. In contrast, the DPOAC algorithm employs a model-based strategy for optimizing continuous action spaces, allowing swift adjustments to the drilling parameters in response to real-time geological changes. The optimization process of the DPOAC algorithm is not only responsive, but also continuously updates its decision-making strategy, ensuring real-time feedback from logging data while achieving an optimal balance between ROP and cost.

Fig. 31 presents a performance comparison of the DPOAC algorithm with the aforementioned methods. The DQN algorithm exhibits weaker optimization of mechanical specific energy and unit cost per foot due to its conservative exploration strategy and insufficient parameter adjustment. Consequently, it achieves lower ROP, longer training time, and lower efficiency. The DDQN algorithm alleviates Q-value overestimation with a double Q-network and shows slight improvements in mechanical specific energy and unit cost per foot, but its exploration ability and optimization efficiency remain limited compared with DPOAC. The DDPG algorithm handles continuous action spaces effectively and increases ROP, but reward delay and limited exploration result in weaker optimization of mechanical specific energy and unit cost per foot relative to DPOAC. The PPO algorithm enhances training stability by restricting policy update ranges; however, its complex parameter adjustment yields slightly inferior optimization of

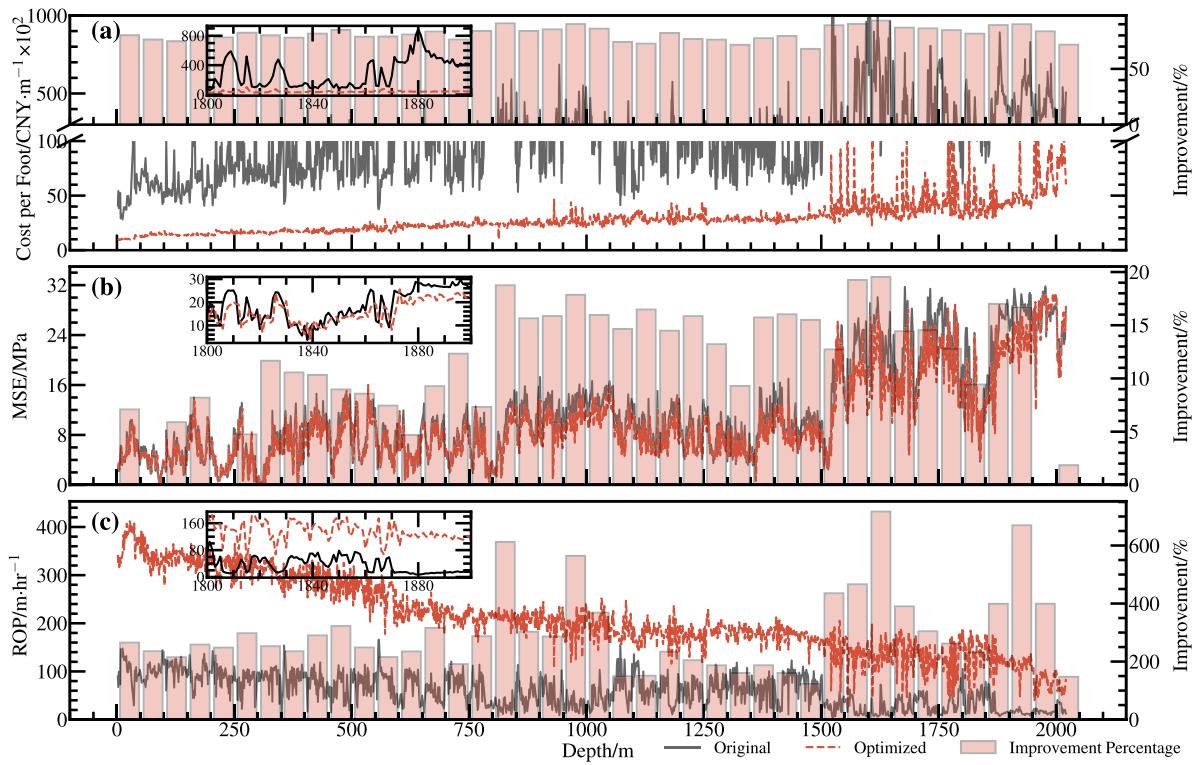


Fig. 27. Optimization changes in unit cost per foot, mechanical specific energy, and ROP with drilling depth.

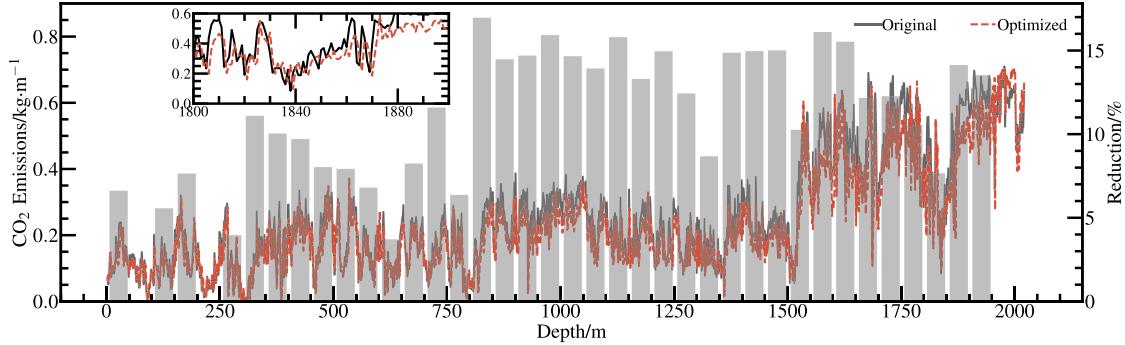


Fig. 28. Optimization impact on CO₂ emissions per meter drilled with drilling depth.

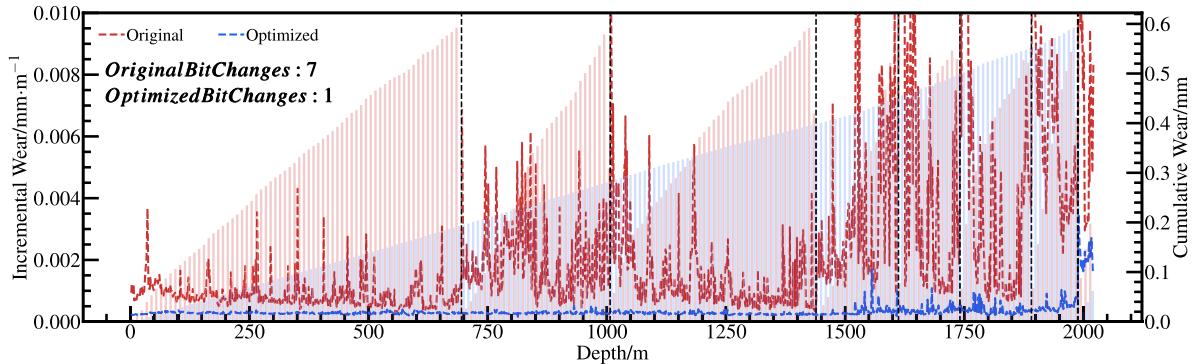


Fig. 29. Comparison of incremental and cumulative bit wear before and after optimization.

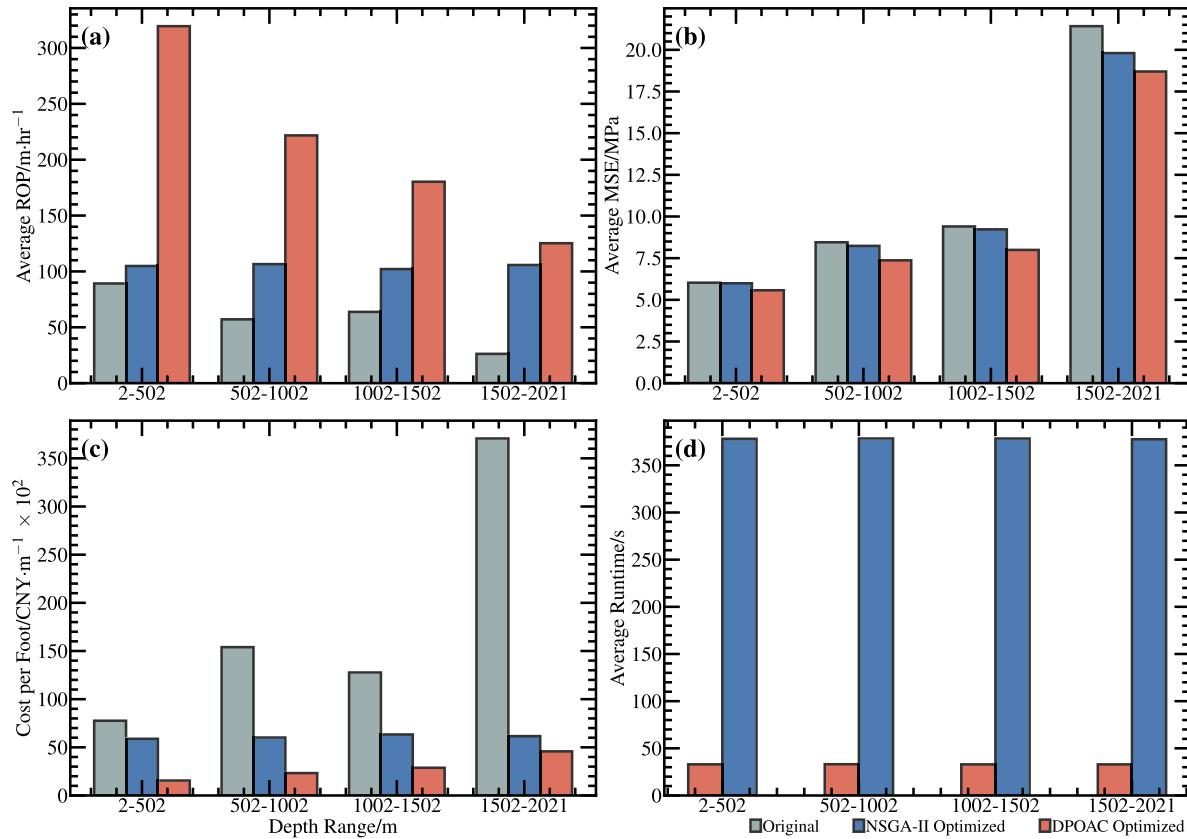


Fig. 30. Performance comparison between the DPOAC algorithm and NSGA-II in drilling optimization.

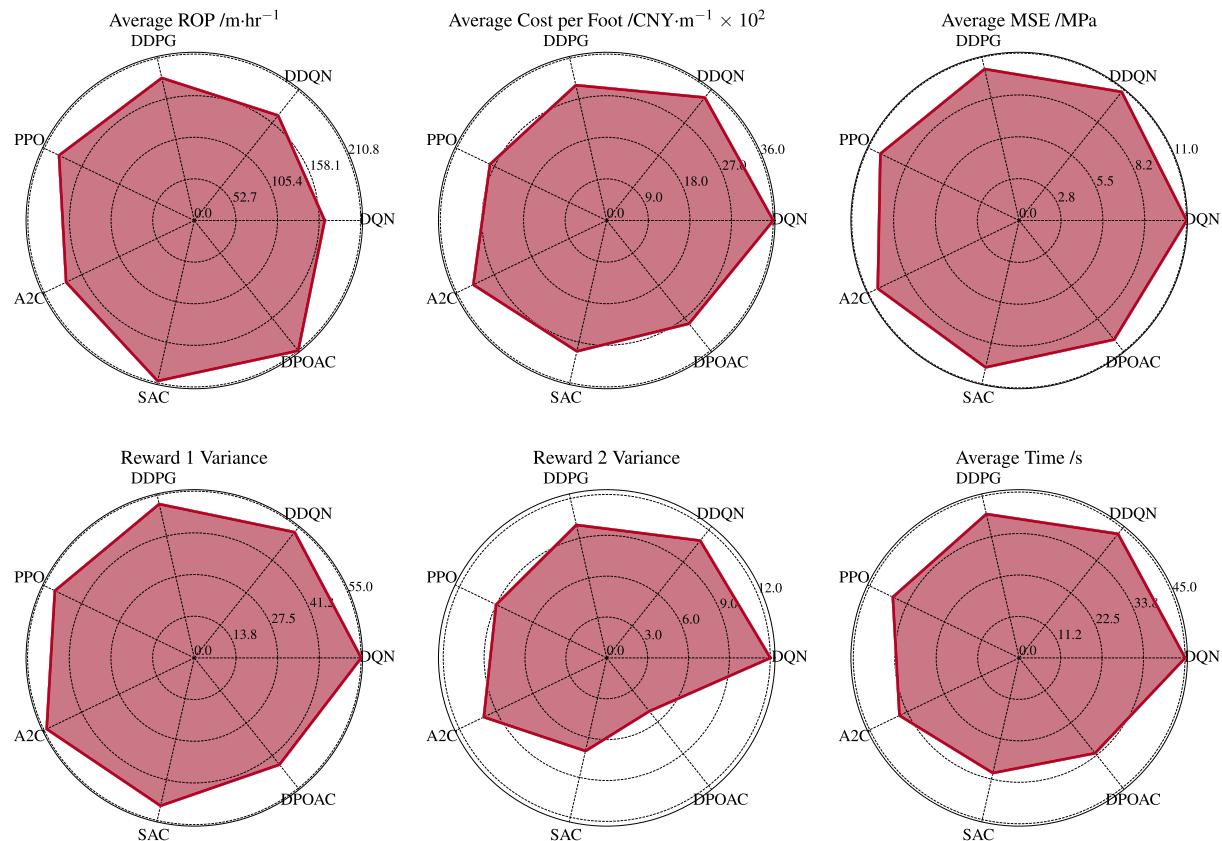


Fig. 31. Performance comparison of DPOAC and various reinforcement learning methods.

mechanical specific energy and unit cost per foot. The A2C algorithm provides stable training by combining policy optimization with value estimation, yet its optimization of mechanical specific energy and unit cost per foot is below that of DPOAC. The SAC algorithm balances exploration and exploitation to optimize ROP and performs well in mechanical specific energy and unit cost per foot, but the absence of reparameterization and a double Q-network leads to slightly lower training stability and exploration ability than DPOAC. Overall, the DPOAC algorithm demonstrates superior training stability, exploration efficiency, and optimization precision, especially in reducing mechanical specific energy and unit cost per foot, while achieving ROP improvements comparable to those of the SAC algorithm.

6. Conclusions

A multi-objective real-time optimization framework for offshore drilling was developed using advanced AI techniques, including temporal prediction networks, domain adversarial networks, and Markov decision processes. Feature extraction and online learning enable robust prediction of formation properties and seamless model transfer to un-drilled formations. Real-time processing of LWD data and formation pressure gradient perception improve decision-making and optimize mud density adjustments, ensuring wellbore stability and operational safety. The DPOAC algorithm within a Markov decision framework provides precise control of drilling parameters and economic optimization.

Empirical analysis in the Caofidian 6-4 block of the Bohai Sea demonstrated the framework's effectiveness, with the ROP increasing from 58.76 m/hr to 210.81 m hr, mechanical specific energy decreasing from 11.42 MPa to 10.01 MPa, and unit cost per meter dropping from 18,439 CNY/m to 2,852 CNY/m. The framework demonstrates optimal performance in the Bohai Sea's Paleogene-Neogene successions, specifically within the Dongying Formation (Members 2 and 3) and Min-guazhen Formation, where siliciclastic lithologies dominate. These outcomes underscore the framework's potential to enhance the reliability, resilience, and sustainability of offshore drilling operations, contributing significantly to sustainable practices in geotechnical and underground engineering.

The scope of application of the framework is currently limited to offshore drilling operations with abundant historical data from neighboring wells. Key limitations include dependency on high-quality historical data. Immediate improvement directions include the development of adaptive sampling strategies to reduce training data requirements in data-scarce environments, and the integration of additional environmental metrics, such as real-time energy consumption and CO₂ emission metrics, into the reward function. Future research should also explore human-AI collaborative drilling by developing hybrid decision systems that combine reinforcement learning with expert

heuristic rules for critical well control scenarios.

Nomenclature

Symbol/Abbreviation	Definition	Unit
WOB	Weight on Bit	kN
RPM	Revolutions Per Minute	rpm
FLW	Pump Flow Rate	L/min
MwIN	Mud Input Density	kg/L
GR	Gamma Ray	API
POR	Porosity	-
DT	Delta Time	μs/m
RD	Rock Density	g/cm ³
ROP	Rate of Penetration	m/hr
MSE	Mechanical Specific Energy	MPa
NSGA-II	Non-dominated Sorting Genetic Algorithm II	-
POR_P	Pore Pressure Gradient	g/cm ³
COLL_P	Collapse Pressure Gradient	g/cm ³
FRAC_P	Fracture Pressure Gradient	g/cm ³
LOSS_P	Loss Pressure Gradient	g/cm ³
LWD	Logging-While-Drilling	-
DPOAC	Double Parameter Optimization Actor-Critic	-
FCNN	Fully Connected Neural Network	-
LSTM	Long Short-Term Memory	-
R ²	Coefficient of Determination	-

CRediT authorship contribution statement

Yu Song: Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Zehua Song:** Writing – original draft, Visualization, Investigation, Conceptualization. **Jin Yang:** Resources, Funding acquisition, Conceptualization. **Longgui Wei:** Data curation. **Jizhou Tang:** Resources, Formal analysis.

Declaration of competing interest

Jin Yang reports financial support was provided by the National Key Research and Development Program of China. Yu Song reports financial support was provided by the National Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the financial support from the National Key Research and Development Program of China (Grant No. 2022YFC28061004) and the National Natural Science Foundation of China (Grant No. 52204017).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ress.2025.111138.

Appendix A—Overview of Drilling Parameters for the Caofidian 6-4 Block of the Bohai Sea in China

Appendix A provides a overview of the drilling parameters and LWD parameters collected from wells in the Caofidian 6-4 block of the Bohai Sea in China.

Table A1

Overview of drilling parameters and LWD parameters for the neighboring well.

Depth (m)	WOB (kN)	RPM (rpm)	FLW (L/min)	MwIN (kg/L)	GR (API)	DT (μs/m)	POR	RD (g/cm³)	ROP (m/hr)
...
700	0.9	90	3998	1.04	42.923	135.788	0.787732	1.713	62.83
701	1.19	90	3998	1.04	45.002	128.208	0.709528	1.78	38.98
702	1	91	3998	1.04	62.065	133.604	0.765199	2.118	50.03
703	0.86	90	3999	1.04	66.293	131.685	0.745401	2.142	61.52
704	0.42	91	3998	1.04	52.668	156.295	0.999305	1.729	72.25
705	1.14	90	3998	1.04	67.734	134.867	0.77823	2.016	39.07
706	0.54	90	4001	1.08	73.222	112.446	0.54691	2.269	27
707	1.87	54	3127	1.08	62.679	139.042	0.821304	1.914	72.59
708	1.97	54	3124	1.08	67.809	134.854	0.778096	2.07	84.99
709	2.21	54	3125	1.08	79.69	137.019	0.800432	2.022	102.71
710	2.23	54	3123	1.08	74.663	140.132	0.832549	2.074	101.84
711	2.21	54	3124	1.08	72.33	133.939	0.768655	2.083	88.54
712	0.94	58	3176	1.08	66.645	117.128	0.595215	2.254	42.13
713	1.29	59	3185	1.08	67.562	135.642	0.786225	2.06	47.78
714	1.31	59	3184	1.08	64.899	140.957	0.841061	2.028	55.93
715	1.61	59	3183	1.08	68.364	133.528	0.764415	2.22	49.72
...

Table A2

Overview of drilling parameters and LWD parameters for the target well.

Depth (m)	WOB (kN)	RPM (rpm)	FLW (L/min)	MwIN (kg/L)	GR (API)	DT (μs/m)	POR	RD (g/cm³)	ROP (m/hr)
...
11	0.68	58	3294	1.05	67.031	147.788	0.83921	2.048	141.89
12	0.81	57	3292	1.05	71.778	147.602	0.837291	2.033	147.84
13	0.87	60	3483	1.05	72.684	145.479	0.815388	2.027	135.69
14	0.8	65	3626	1.05	71.208	143.236	0.792246	2.055	135.44
15	0.8	65	3703	1.05	67.641	126.11	0.615556	2.24	123.8
16	0.89	65	3723	1.05	68.411	140.977	0.76894	2.068	129.33
17	0.78	66	3761	1.05	70.05	143.982	0.799943	2.07	130.63
18	0.96	65	3762	1.05	75.352	141.548	0.774831	2.04	130.63
19	1	64	3785	1.05	73.43	137.749	0.735637	2.081	121.32
20	1.19	65	3785	1.05	72.904	130.754	0.663468	2.032	131.84
21	1.82	65	3779	1.05	70	162.273	0.988653	1.671	115.36
22	1.68	65	3779	1.05	71.511	131.35	0.669617	1.995	98.42
23	1.52	65	3780	1.05	76.398	139.114	0.749719	2.095	98.59
24	1.35	65	3780	1.05	77.812	141.96	0.779082	2.1	97.98
25	1.21	65	3778	1.05	70.008	144.598	0.806298	2.049	111.86
...

Table A3

Statistical summary of drilling and LWD parameters for neighboring and target wells.

	Statistic	Depth (m)	WOB (kN)	RPM (rpm)	FLW (L/min)	MwIN (kg/L)	GR (API)	DT (μs/m)	POR	RD (g/cm³)	ROP (m/hr)
Drilled Neighboring Well	Min	696	0.01	29	891	1.04	39.201	59.618	0.002	1.672	0.4
	Max	3142	17.44	91	4017	1.42	184.746	156.295	0.999	2.727	248.5
	Mean	-	2.644	72.557	3063.867	1.231	80.032	103.501	0.455	2.267	53.693
	Std	-	1.544	10.949	942.904	0.116	18.515	20.571	0.212	0.163	34.566
	Skewness	-	0.73	-0.571	-0.559	0.012	0.74	-0.015	-0.015	0.172	0.581
Pre-drill Target Well	Min	1	0.01	39	2412	1.05	43.478	66.503	0.001	1.565	3.55
	Max	101	6.31	84	3998	1.36	135.073	163.082	0.997	2.676	166.04
	Mean	-	2.336	78.596	3551.582	1.183	81.161	109.224	0.441	2.238	58.74
	Std	-	1.448	3.296	371.804	0.106	13.278	20.993	0.217	0.159	33.719
	Skewness	-	0.809	-4.216	-0.798	0.272	0.569	-0.095	-0.095	0.461	0.236

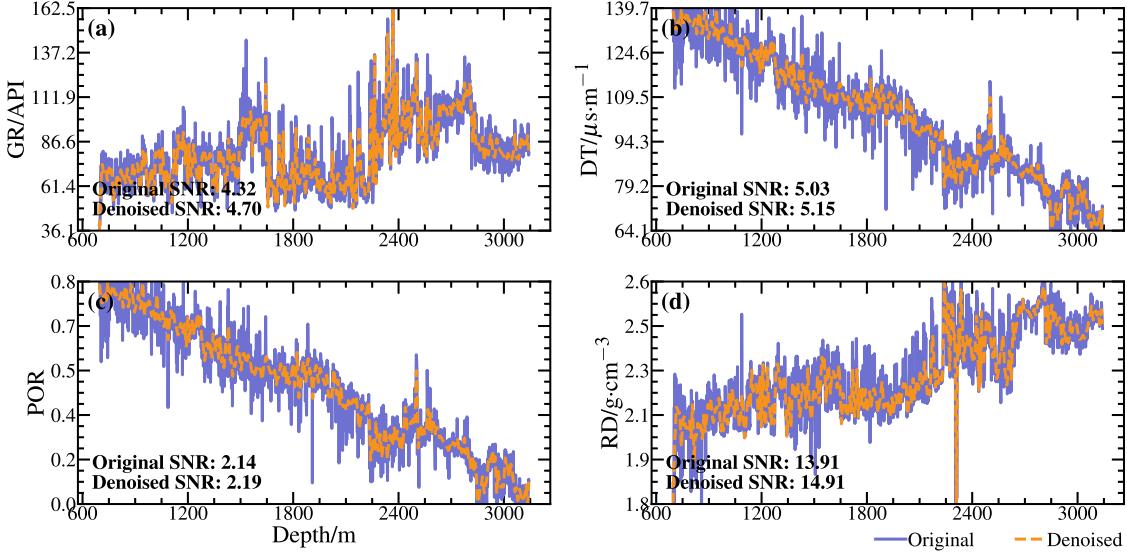


Fig. A1. Wavelet denoising effect on key LWD parameters (GR, DT, POR, and RD) in neighboring well.

Appendix B—Optimization Process for ROP and Formation Property Prediction Using the Domain Adversarial Mechanism

Appendix B outlines the detailed steps of the Domain Adversarial Mechanism, which is designed to optimize the perception and prediction of ROP and formation property parameters through adversarial domain adaptation.

Table B1

Optimization process for ROP and formation property parameters using the Domain Adversarial Mechanism.

Algorithm: Domain Adversarial Mechanism for ROP and Formation Property Prediction

Input:

$D_s = \{(x_i^s, y_i^s)\}_{i=1}^{n_s}$: Source domain dataset (drilled neighboring wells)

$D_t = \{(x_i^t, y_i^t)\}_{i=1}^{n_t}$: Target domain dataset (undrilled formations)

G_f : Initial feature extractor parameters

G_d : Initial domain discriminator parameters

$G_p = \{G_{p1}, G_{p2}, G_{p3}\}$: Initial perceiver models for ROP, LWD parameters, and formation pressure gradient

Parameters:

Learning rates η_f, η_d, η_p for G_f, G_d, G_p respectively

Number of training epochs E

Batch size B

Adversarial loss weight λ

Optimization algorithms

Output:

Trained feature extractor G_f^*

Trained domain discriminator G_d^*

Trained perceivers $G_p^* = \{G_{p1}^*, G_{p2}^*, G_{p3}^*\}$

Function: DomainAdversarialMechanism

1. **Initialize** feature extractor G_f , domain discriminator G_d , and perceivers G_{p1}, G_{p2}, G_{p3} with parameters $\theta_f, \theta_d, \theta_{p1}, \theta_{p2}, \theta_{p3}$

2. Define loss functions:

3. Domain classification loss L_d (Binary Cross-Entropy)

4. Perception losses L_{p1}, L_{p2}, L_{p3} (Mean Squared Error)

5. Gradient reversal layer parameter λ for adversarial training

6. **For** epoch = 1 to E **do**:

7. Shuffle and partition D_s and D_t into mini-batches of size B

8. **For** each mini-batch b from 1 to B **do**:

9. Sample batch from source domain: $\{x_i^s, y_i^s\}$

10. Sample batch from target domain: $\{x_i^t, y_i^t\}$

11. Forward pass through G_f :

13. Extract features $f_i^s = G_f(x_i^s)$

14. Extract features $f_i^t = G_f(x_i^t)$

15. Domain discriminator training:

16. Combine f_i^s and f_i^t with domain labels d_i (0 for source, 1 for target)

17. Compute domain classification loss $L_d = \text{BCE}(G_d(f_i), d_i)$

18. Backpropagate and update G_d parameters to minimize L_d

19. Feature extractor adversarial training:

(continued on next page)

Table B1 (continued)

Algorithm: Domain Adversarial Mechanism for ROP and Formation Property Prediction	
20.	Pass f_i^s and f_i^t through G_d with reversed gradients using gradient reversal layer
21.	Compute adversarial loss $L_{adv} = BCE(G_d(f_i), 1 - d_i)$
22.	Combine with perception losses $L_p = L_{p1} + L_{p2} + L_{p3}$
23.	Total loss $L_f = L_p + \lambda L_{adv}$
24.	Backpropagate and update G_f parameters to minimize L_f
25.	Perceiver training:
26.	For each perceiver G_{p1}, G_{p2}, G_{p3} :
27.	Compute prediction $\hat{y} = G_p(f_i^s)$
28.	Compute perception loss $L_{pj} = MSE(\hat{y}, y_i^s)$
29.	Backpropagate and update G_{pj} parameters to minimize L_{pj}
30.	End for
31.	End for
32.	Return trained models G_f^*, G_d^*, G_p^*

Appendix C—Optimization Process for Drilling Parameters Using the DPOAC Algorithm

Appendix C outlines the detailed steps of the DPOAC algorithm, which is designed to optimize drilling parameters through a reinforcement learning approach.

Table C1

Optimization process for drilling parameters using the DPOAC algorithm.

Algorithm: DPOAC for Optimizing Drilling Parameters	
Input: T = Total number of episodes	
Parameters:	
θ = Initial actor network parameters	
ω_1, ω_2 = Initial twin critic network parameters	
θ' = Initial target actor network parameters	
ω'_1, ω'_2 = Initial target critic network parameters	
α = Temperature parameter for entropy regularization	
τ = Soft update coefficient for target networks	
γ = Discount factor for future rewards	
D = Empty experience replay buffer	
N = Exploration noise process	
ϵ_t = Reparameterized noise term	
$\sigma_{\text{exploration}}$ = Standard deviation of exploration noise	
Output: Optimized parameters $\theta, \omega_1, \omega_2$ for the actor and critic networks	
Function DPOAC	
1.	Initialize actor network π_θ and twin critic networks Q_{ω_1} and Q_{ω_2} , and reparameterization function $f_\phi(\epsilon_j; s_j)$
2.	Initialize target critic networks: $\theta' \leftarrow \theta, \omega'_1 \leftarrow \omega_1, \omega'_2 \leftarrow \omega_2$
3.	For $i = 1$ to T do:
4.	Observe initial state s_1
5.	While s_i is not terminal do:
6.	Sample noise ϵ_t from N .
7.	Select action $a_t = \mu_\theta(s_t) + \sigma_\theta(s_t) \cdot \epsilon_t + \gamma \cdot \mathcal{N}(0, \sigma_{\text{exploration}}^2)$, where π_θ outputs parameters $\mu_\theta(s_t)$ and $\sigma_\theta(s_t)$ of a Gaussian distribution
8.	Execute action a_t , observe reward r_t , new state s_{t+1} , and terminal status
9.	Store transition (s_t, a_t, r_t, s_{t+1}) in D
10.	Sample a minibatch of m transitions (s_j, a_j, r_j, s_{j+1}) from D
11.	Set $y_j = r_j + \gamma \cdot \min(Q_{\omega'_1}(s_{j+1}, f_\phi(\epsilon_{j+1}; s_{j+1})), Q_{\omega'_2}(s_{j+1}, f_\phi(\epsilon_{j+1}; s_{j+1})))$ if s_{j+1} is not terminal, else $y_j = r_j$
12.	Update critic $Q_{\omega_1}, Q_{\omega_2}$ by minimizing loss: $\mathcal{L} = \frac{1}{m} \sum_{j=1}^m \left[(y_j - Q_{\omega_1}(s_j, a_j))^2 + (y_j - Q_{\omega_2}(s_j, a_j))^2 \right]$.
13.	Update actor network π_θ via policy gradient: $\nabla_\theta \mathcal{L} \approx \frac{1}{m} \sum_{j=1}^m \left[\nabla_\theta \pi_\theta(s_j) \left(\nabla_a Q_{\omega_1}(s_j, a) _{a=f_\phi(s_j)} - \log \pi_\theta(a s_j) + \alpha \right) \right]$
14.	Adjust temperature parameter α for entropy regularization
15.	Soft update target networks:
16.	$\omega'_1 \leftarrow \tau \omega_1 + (1 - \tau) \omega'_1$
17.	$\omega'_2 \leftarrow \tau \omega_2 + (1 - \tau) \omega'_2$
18.	$\theta' \leftarrow \tau \theta + (1 - \tau) \theta'$
19.	Update state: $s_t \leftarrow s_{t+1}$
20.	End while
21.	End for
22.	Return optimized network parameters $\theta, \omega_1, \omega_2$

Appendix D—Detailed Mathematical Formulations and Derivations***iTransformer model***

For formation property prediction, the iTransformer model provides an efficient architecture for predicting multivariable time-series data within specific time windows, such as GR, DT, POR, and RD [46]. This model effectively processes and analyzes these complex datasets to achieve accurate geological feature perception. To elucidate the mechanisms behind its performance, an examination of the model architecture, self-attention mechanism, and embedding and projection methods is presented below.

Model Architecture: The iTransformer model employs an encoder-only architecture, omitting the traditional decoder, and focuses on using the encoder to directly process and predict time-series data [61]. This structure allows learning feature representations of future data $Y = \{x_{T+1}, \dots, x_{T+S}\} \in \mathbb{R}^{S \times N}$ from historical data $X = \{x_1, \dots, x_T\} \in \mathbb{R}^{T \times N}$ without an additional decoding process. In this model, $X_{(t,i)}$ represents the time points recorded simultaneously at time step t , and $X_{(:,n)}$ denotes the entire time series of each variable indexed by n . The iTransformer performs data embedding and feature projection tasks through encoder blocks composed of multilayer perceptrons, handles multivariable interactions in the time series via a self-attention mechanism, and effectively captures dynamic correlations between variables. The iTransformer can process time series with complex geological and physical properties to predict the formation pore pressure, fracture pressure, and collapse pressure during drilling. Sequential information is implicitly stored in the arrangement of the neurons in the preceding network, which simplifies the complex positional encoding required by traditional transformers.

Self-Attention Mechanism: In formation property perception studies, the iTransformer model emphasizes the utilization of the self-attention mechanism to process and predict time-series data. The model dynamically adjusts its focus on various variables across different time steps, such as GR, DT, POR, and RD, using self-attention layers. This mechanism allows the model to automatically identify and emphasize the most critical variables for predicting subsequent formation properties such as the pore, fracture, collapse, and loss pressures. The self-attention layers calculate the interdependencies between different variables to optimize the prediction accuracy [62]. In this process, the multivariable data at each time series point are transformed into high-dimensional vectors, each linearly transformed to obtain the corresponding queries (Q), keys (K), and values (V), as shown in Eq. D1.

$$Q = HW^Q, K = HW^K, V = HW^V \quad (\text{D1})$$

where $H = \{h_0, \dots, h_N\} \in \mathbb{R}^{N \times D}$ and represents the embedded representation of the input sequence, and W^Q, W^K, W^V are the weight matrices mapping the input to their respective spaces. The model calculates the interactions between each pair of Q and K vectors, normalized using the softmax function, to obtain attention weights, as shown in Eq. D2.

$$A = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \quad (\text{D2})$$

where $A \in \mathbb{R}^{N \times N}$ represents the attention score matrix, determining the strength of interactions between variables. Each V vector is combined according to these weights to generate an output representation for predicting future attributes.

The self-attention mechanism enhances the sensitivity of the model to dynamic changes in the time series and significantly improves the accuracy of formation property predictions by precisely modeling the interdependencies between multiple variables. Fig. 4 demonstrates that the mechanism enables iTransformer to effectively capture both long- and short-term dependencies when processing drilling data with complex formation properties, thus providing more accurate decision support in predicting the formation properties.

Embedding and Projection: In iTransformer, the primary function of the embedding layer is to transform the input multivariable time-series data into vector representations in a high-dimensional space. This process is illustrated in Eq. D3.

$$h_n^0 = \text{Embedding}(X_{(:,n)}) \quad (\text{D3})$$

where h_n^0 denotes the vector processed by the embedding layer. These vectors interact through the self-attention mechanism and learn the complex relationships between the variables. Each vector is first mapped to a fixed-dimensional embedding space, aiding in standardizing the processing of the input data. The projection layer of the model converts the encoded high-dimensional features back to their original data dimensions, as expressed in Eq. D4.

$$Y_{(:,n)}^L = \text{Projection}(h_n^L) \quad (\text{D4})$$

where h_n^L is the output of the final transformer layer. This embedding and projection process ensures a continuous and consistent flow of information from the input to the output while preserving complex interdata relationships. In addition, the model effectively processes and predicts multivariate time-series data, providing accurate forecasts for future time points.

In the iTransformer architecture, the interactions between variables are achieved using a self-attention mechanism and independently processed and shared into feedforward networks (FFNs) in each TrmBlock. The sequence order is implicitly stored by the arrangement of neurons in the feedforward network, thereby eliminating the need for positional embedding. This process is illustrated in Eq. D5.

$$H^{(l+1)} = \text{TrmBlock}(H^l), l = 0, \dots, L - 1 \quad (\text{D5})$$

where $H = \{h_1, \dots, h_N\} \in \mathbb{R}^{N \times D}$ and contains N embedding vectors with dimension D , the superscript denotes the layer index. The embedding function $\mathbb{R}^T \rightarrow \mathbb{R}^D$ and the projection function $\mathbb{R}^D \rightarrow \mathbb{R}^S$ are implemented through multilayer perceptrons.

Feedforward Network and Layer Normalization: In the iTransformer model, each encoding block integrates a FFN to process the output from the self-attention layer. The FFN enhances the ability of the model to represent time-series data through dense nonlinear transformations [63] and is given by Eq. D6.

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (\text{D6})$$

where W_1 , W_2 , and b_1 , b_2 denote the weights and biases of the network layers, respectively. This network structure ensures that each data point is processed independently, strengthening the ability of the model to handle the variability of complex geological data. In addition, the FFN effectively captures the geological features derived from historical data windows and predicts subsequent geological changes.

Layer Normalization (LayerNorm) stabilizes the training process and accelerates model convergence. This technique standardizes each feature within a layer and reduces the internal covariate shifts during training [64]. This process is illustrated in Eq. D7.

$$\text{LayerNorm}(x) = \frac{x - \mu}{\sigma} \quad (\text{D7})$$

where μ and σ denote the mean and standard deviation of the vector x , respectively. This method addresses the impacts of different scales and dimensions, enabling data from various measurement standards to be compared and processed on the same scale.

Layer normalization was applied after the self-attention and FFNs, normalizing the output of each operation. Unlike batch normalization used in traditional transformer models, layer normalization does not depend on the batch size, making it more stable when processing data batches of varying sizes. Additionally, layer normalization helps eliminate the interaction noise caused by unsynchronized time-point collection or measurement errors, thereby enhancing the performance of the model in processing non-stationary time-series data.

Through these carefully designed structures, the iTransformer model effectively processes and predicts multivariable time-series data. In drilling parameter optimization applications, this model can predict future formation property changes based on historical data, providing scientific support for decision-making, and thereby enhancing drilling efficiency and safety.

Unit Cost per Foot Function

In drilling operations, the cost per foot function is a core indicator for assessing economic efficiency, focusing on accurately measuring costs and efficiency during drilling. The expression for cost per foot is given in Eq. D8 [65].

$$C_{pm} = (C_b + C_r(t_t + t)) / H \quad (\text{D8})$$

where C_{pm} is the cost per foot, (CNY/m); C_b is the cost of the drill bit, (CNY); C_r is the rig operation fee (CNY/h); t is the drilling time, (hr); t_t is the time required to make trips and connections, (hr) and H is the total footage drilled, (m); The wear rate of the drill bit is given by Eq. D9, illustrating the wear dynamics of the drill bit teeth.

$$\frac{dh}{dt} = \frac{A_f(a_1n + a_2n^3)}{Z_2 - Z_1W(1 + C_1h)} \quad (\text{D9})$$

Eq. D9 is transformed into a relationship between time dt and wear dh , as shown in Eq. D10.

$$dt = \frac{(Z_2 - Z_1W)(1 + C_1h)}{A_f(a_1n + a_2n^3)} dh \quad (\text{D10})$$

where A_f is the abrasiveness coefficient of the formation; n is the RPM, (r/min); a_1, a_2 are the speed influence coefficients determined by the type of drill bit; Z_1, Z_2 are the WOB influence coefficients related to the diameter of the drill bit; C_1 is the tooth wear deceleration coefficient.

By incorporating the ROP prediction transfer model v_{pc} into the aforementioned bit wear rate equation, a mathematical model for bit footage H and working time t was established, as shown in Eq. D11.

$$\begin{cases} H = \int_0^{h_f} v_{pc} \cdot \frac{(Z_2 - Z_1W)(1 + C_1h)}{A_f(a_1n + a_2n^3)} dh \\ t = \int_0^{h_f} \frac{(Z_2 - Z_1W)}{A_f(a_1n + a_2n^3)} \left(h + \frac{C_1}{2}h^2 \right) dh \end{cases} \quad (\text{D11})$$

where h_f denotes the amount of wear corresponding to the drill-bit lifespan. Consequently, the cost per foot function is reformulated as a function of the WOB and the RPM, as expressed in Eq. D12.

$$C_{pm} = \frac{C_r \left[t_E \cdot A_f(a_1n + a_2n^3) + \left(h_f + \frac{C_1}{2}h_f^2 \right) \right]}{v_{pc} \cdot \left(h_f + \frac{C_1}{2}h_f^2 \right)} \quad (\text{D12})$$

where $t_E = \frac{C_r}{C_p} + t_t$ is the equivalent time factor for the unit cost. This cost per foot function illustrates the combined impact of drill bit cost and rig operating fees on the unit cost of drilling. This provides a precise method for calculating drilling costs, thus supporting scientific and accurate decision-making in drilling operations.

MSE Function

The MSE model correlates the energy required to fracture a unit volume of rock with the rock-breaking efficiency of the drill bit and serves as a critical metric for quantifying drilling efficiency [66]. The MSE is calculated using parameters such as the ROP, WOB, PRM, torque, and drill bit diameter. A higher MSE value indicates a lower drilling efficiency and poorer adaptability of the drill bit to the formation, suggesting that the drilling parameters require optimization. The ideal MSE calculation is given by Eq. D13 [67].

$$MSE = \frac{4W}{\pi D_b^2} + \frac{480nT_b}{D_b^2P} \quad (\text{D13})$$

where MSE is the mechanical specific energy, (MPa); W is the WOB, (kN); D_b is the drill bit diameter, (m); n is the RPM, (r/min); T_b is the torque of the drill bit in, (kN·m) and P is the mechanical ROP, (m/hr).

In practice, real torque values at the drill bit bottom are often unavailable, necessitating calculations using the drill bit's sliding friction coefficient and WOB. The torque during drilling can be expressed using a double integral, as depicted in Eq. D14 [68].

$$T_b = \frac{1}{1000} \int_0^{D_b/2} \int_0^{2\pi} r^2 \frac{4\mu W}{\pi D_b^2} dr d\theta = \frac{\mu WD_b}{3000} \quad (\text{D14})$$

where r represents the radial increment of the drill bit, (mm), and μ is the sliding friction coefficient of the drill bit.

Considering factors such as friction and vibrations, the actual energy utilization rate in drilling is typically between 30% and 40% [67]. The required MSE is approximately three times the rock strength. Defining the effective energy utilization rate as E_f . By introducing the ROP prediction transfer model v_{pc} , the modified mechanical specific energy model MSE_m is given in Eq. D15.

$$MSE_m = E_f \left(\frac{4W}{\pi D_b^2} + 0.16 \frac{n\mu W}{D_b v_{pc}} \right) \quad (\text{D15})$$

CO₂ Reduction Model

The CO₂ reduction model quantifies emissions savings by linking fuel consumption to drilling energy efficiency, providing a critical sustainability metric. The CO₂ emissions calculation per meter drilled is given by Eq. D16.

$$CO_{2,\text{emission}} = FC \times C_f \quad (\text{D16})$$

where $CO_{2,\text{emission}}$ is CO₂ emissions per meter drilled, (kg/m); FC is fuel consumption per meter, (L/m); C_f is the diesel carbon intensity factor, (2.63 kg/L).

Fuel consumption is driven by the power demand of the drilling process. Power demand derives directly from the modified mechanical specific energy, which quantifies energy losses during rock fracturing. The power demand calculation based on the modified mechanical specific energy is expressed by Eq. D17.

$$P = \frac{MSE_m \times v_{pc} \times \pi(D_b/2)^2}{3.6} \quad (\text{D17})$$

where P is the power demand, (kW); MSE_m is the modified mechanical specific energy model, (MPa); v_{pc} is the ROP prediction transfer model, (m/hr); D_b is the drill bit diameter, (m); the factor 3.6 converts energy units from MJ/hr to kW.

The drivetrain efficiency and engine fuel efficiency are derived from field data and the official Caterpillar 3512B Industrial Engine Specifications for the Caterpillar 3512B diesel engines deployed in the Caofeidian 6-4 block. The theoretical fuel efficiency calculation based on the engine's brake specific fuel consumption (BSFC) and diesel density is presented in Eq. D18.

$$\epsilon_{f,\text{theory}} = \frac{BSFC}{\rho} \quad (\text{D18})$$

where $\epsilon_{f,\text{theory}}$ is the theoretical fuel efficiency; ρ is the diesel density, (0.835kg/L); E_d is the diesel energy content, (42.5MJ/kg); BSFC is the brake specific fuel consumption, (g/kWh); ϵ_f is the adjusted engine fuel efficiency, (0.28L/kWh); The adjusted ϵ_f accounts for drivetrain losses and auxiliary power demand.

The fuel consumption per meter drilled calculation, incorporating power demand and efficiencies, is shown in Eq. D19.

$$FC = \frac{P}{\eta} \times \epsilon_f \times \frac{1}{v_{pc}} \quad (\text{D19})$$

where η is the drivetrain efficiency, (67%).

Substituting Eq. D17 and ϵ_f into Eq. D16, the optimized CO₂ emissions calculation is presented in Eq. D20.

$$CO_{2,\text{opt}} = \frac{MSE_m \times \pi(D_b/2)^2}{3.6\eta} \times \epsilon_f \times C_f \quad (\text{D20})$$

Appendix E—Cross-Block Stratigraphic Constraints and Transfer Learning Performance

This appendix provides stratigraphic depth/thickness distributions of key formations across validation wells and quantifies R^2 performance of the Caofeidian 6-4-2-trained perception transfer model under online learning.

Table E1

Stratigraphic depth and thickness of key formations across the 5 blocks.

Formation	Caofeidian 6-4-4		Kenli 16-1-2		Penglai 7-6-1		Dongfang 13-2-1		Ledong 10-1-1	
	Depth (m)	Thickness (m)	Depth (m)	Thickness (m)	Depth (m)	Thickness (m)	Depth (m)	Thickness (m)	Depth (m)	Thickness (m)
Pingyuan Fm	500	500	515	435	678	656.5	-	-	-	-
Minghuazhen Fm	855	355	950	436	1414.2	736.2	-	-	-	-
Lower Minghua	1627	772	-	-	2275	860.8	-	-	-	-
Guantao Fm	2233	606	1074	124	3250	975	-	-	-	-
Dongying Fm	2483	250	1191	117	-	-	-	-	-	-
Shahjie Fm	-	-	1452	260	-	-	-	-	-	-
Mesozoic	-	-	1558	107	-	-	-	-	-	-
Ledong Fm	-	-	-	-	-	-	895	895	1465	490
Ledong Fm	-	-	-	-	-	-	1353	458	2336	871
Yinggehai Fm (Member 1)	-	-	-	-	-	-	2727	1374	3562	1226
Yinggehai Fm (Member 2)	-	-	-	-	-	-	-	-	-	-
Huangliu Fm (Member 1)	-	-	-	-	-	-	-	-	3951	389
Huangliu Fm (Member 2)	-	-	-	-	-	-	-	-	4098	147

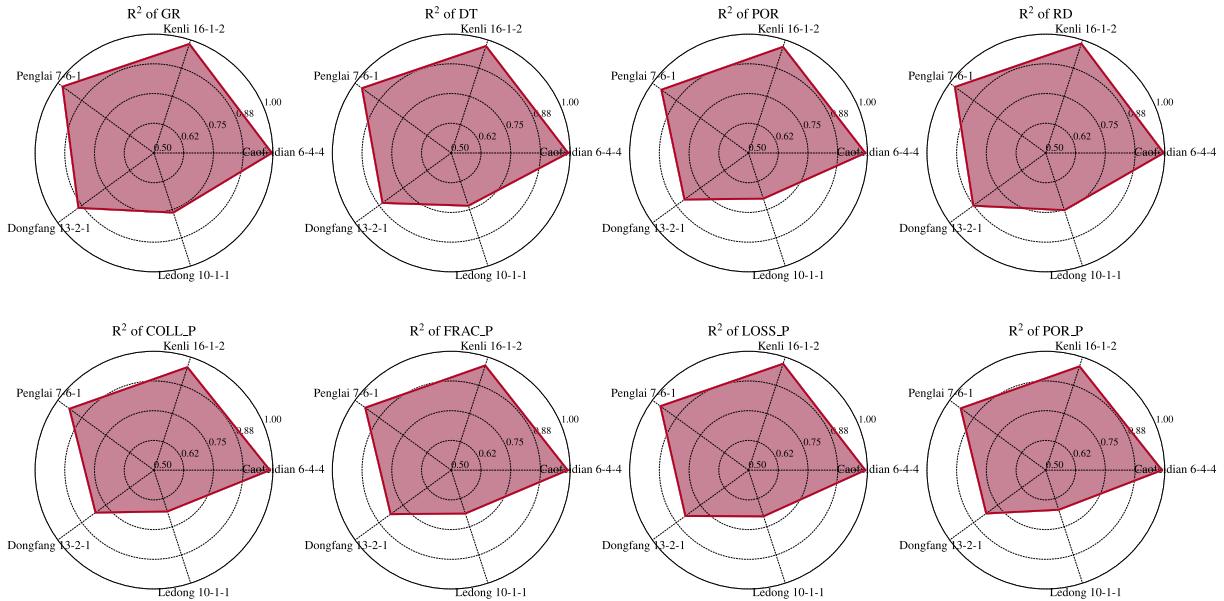


Fig. E1. R^2 performance evaluation of the formation property perception transfer model with online learning across five wells trained on Caofeidian 6-4-2.

Data availability

Data will be made available on request.

References

- [1] Yang J. Jetting Installation Method of Oil-Gas Well Conductor. Installation Methods of Offshore Oil-Gas Well Conductor. Springer; 2022. p. 107–62.
- [2] Wang Y, Gao D. Study on the marine environment limiting conditions of deepwater drilling for natural gas hydrate. Appl Energy 2022;312:118802.
- [3] Wu S, Zhang Q, Li B, Zhang L, Zheng W, Li Z, et al. Reliability analysis of subsea wellhead system subject to fatigue and degradation during service life. Reliab Eng Syst Saf 2023;239:109494.
- [4] Liu Z, Liu K, Chen X, Ma Z, Lv R, Wei C, et al. Deep-sea rock mechanics and mining technology: State of the art and perspectives. Int J Min Sci Technol 2023;33: 1083–115.
- [5] Cao B, Yin Q, Guo Y, Yang J, Zhang L, Wang Z, et al. Field data analysis and risk assessment of shallow gas hazards based on neural networks during industrial deep-water drilling. Reliab Eng Syst Saf 2023;232:109079.
- [6] Liu W, Li A, Fang W, Love PED, Hartmann T, Luo H. A hybrid data-driven model for geotechnical reliability assessment. Reliab Eng Syst Saf 2023;231:108985.
- [7] Skogdalen JE, Vinnem JE. Quantitative risk analysis of oil and gas drilling, using Deepwater Horizon as case study. Reliab Eng Syst Saf 2012;100:58–66.
- [8] Parhizkar T, Hogendoorn S, Vinnem JE, Utne IB. Data driven approach to risk management and decision support for dynamic positioning systems. Reliab Eng Syst Saf 2020;201:106964.
- [9] Zhang L, Lin P. Multi-objective optimization for limiting tunnel-induced damages considering uncertainties. Reliab Eng Syst Saf 2021;216:107945.
- [10] Chen X, Du X, Weng C, Yang J, Gao D, Su D, et al. A real-time drilling parameters optimization method for offshore large-scale cluster extended reach drilling based on intelligent optimization algorithm and machine learning. Ocean Eng 2024;291: 116375.
- [11] Silva LMR, Guedes, Soares C. Robust optimization model of an offshore oil production system for cost and pipeline risk of failure. Reliab Eng Syst Saf 2023; 232:109052.

- [12] Pan Y, Qin J, Hou Y, Chen J-J. Two-stage support vector machine-enabled deep excavation settlement prediction considering class imbalance and multi-source uncertainties. *Reliab Eng Syst Saf* 2024;241:109578.
- [13] Yu J, Zhou Z, Liu Y, Zhang H, Song Y, Dong X. Economic perspectives on renewable energy in offshore oil and gas development: Integrating multiple functions combination degradation maintenance strategies. *Reliab Eng Syst Saf* 2024;250:110271.
- [14] Kendall H, Goins Jr W. Design and operation of jet-bit programs for maximum hydraulic horsepower, impact force or jet velocity. *Transactions of the AIME* 1960; 219:238–50.
- [15] Eckel JR. Microbit studies of the effect of fluid properties and hydraulics on drilling rate. *J Pet Technol* 1967;19:541–6.
- [16] Ramba V, Selvaraju S, Subbiah S, Palanisamy M, Srivastava A. Optimization of drilling parameters using improved play-back methodology. *Journal of Petroleum Science and Engineering* 2021;206:108991.
- [17] Zhang YX, Wang LF, Yao F, Wang H, Men XS, Ye BB. Experimental study on the influence of nozzle diameter on abrasive jet cutting performance. *Advanced Materials Research* 2011;337:466–9.
- [18] Bani Mustafa A, Abbas AK, Alsaba M, Alameen M. Improving drilling performance through optimizing controllable drilling parameters. *Journal of Petroleum Exploration and Production* 2021;11:1223–32.
- [19] Bajolvand M, Ramezanzadeh A, Mehrad M, Roohi A. Optimization of controllable drilling parameters using a novel geomechanics-based workflow. *Journal of Petroleum Science and Engineering* 2022;218:111004.
- [20] Zang C, Lu Z, Ye S, Xu X, Xi C, Song X, et al. Drilling parameters optimization for horizontal wells based on a multiobjective genetic algorithm to improve the rate of penetration and reduce drill string drag. *Applied Sciences* 2022;12:11704.
- [21] Delavar MR, Ramezanzadeh A, Gholami R, Sanei M. Optimization of drilling parameters using combined multi-objective method and presenting a practical factor. *Comput Geosci* 2023;175:105359.
- [22] Zheng J, Li ZL, Dou B, Lu C. Multi-objective cellular particle swarm optimization and RBF for drilling parameters optimization. *Mathematical Biosciences and Engineering: MBE* 2019;16:1258–79.
- [23] Khaled AA, Adnan MS, Alhamd SJ. Estimation of Bourgoine and young model coefficients to predict optimum drilling rates and bit weights using genetic algorithms-a case study of the Faithaa oil field in Iraq. In: IOP Conf Ser: Mater Sci Eng. IOP Publishing; 2021, 012154.
- [24] Lin P, Zhang L, Tiong RLK. Multi-objective robust optimization for enhanced safety in large-diameter tunnel construction with interactive and explainable AI. *Reliab Eng Syst Saf* 2023;234:109172.
- [25] Zhang Y, Tião P, Leonards A, Tang K. A survey on neural network interpretability. *IEEE Trans Emerging Top Comput Intell* 2021;5:726–42.
- [26] Sabah M, Talebkeikhah M, Wood DA, Khosravani R, Anemangely M, Younesi A. A machine learning approach to predict drilling rate using petrophysical and mud logging data. *Earth Sci Inf* 2019;12:319–39.
- [27] Hegde C, Daigle H, Gray KE. Performance comparison of algorithms for real-time rate-of-penetration optimization in drilling using data-driven models. *SPE J* 2018; 23:1706–22.
- [28] Matias J, de Castro Oliveira JP, Le Roux GA, Jäschke J. Real-time optimization with persistent parameter adaptation applied to experimental rig. *Ifac-papersonline* 2021;54:475–80.
- [29] Jahandideh A, Jafarpour B. Closed-loop stochastic oilfield optimization for hedging against geologic, development, and operation uncertainty. *Comput Geosci* 2020; 24:129–48.
- [30] Osarogiobon AU, Oloruntobi O, Khan F, Venkatesan R, Butt S. Gamma ray log generation from drilling parameters using deep learning. *Journal of Petroleum Science and Engineering* 2020;195:107906.
- [31] Magnusson LV, Olsson JR, Tran CIT. Recurrent neural networks for oil well event prediction. *IEEE Intell Syst* 2023;38:73–80.
- [32] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A comprehensive survey on transfer learning. *Proc IEEE* 2020;109:43–76.
- [33] Fan C, Sun Y, Xiao F, Ma J, Lee D, Wang J, et al. Statistical investigations of transfer learning-based methodology for short-term building energy predictions. *Appl Energy* 2020;262:114499.
- [34] Wu J, He H, Peng J, Li Y, Li Z. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. *Appl Energy* 2018;222:799–811.
- [35] Zhang Q, Liu Y, Xiang Y, Xiahou T. Reinforcement learning in reliability and maintenance optimization: A tutorial. *Reliab Eng Syst Saf* 2024;251:110401.
- [36] Pinciroli L, Baraldi P, Zio E. Maintenance optimization in industry 4.0. *Reliab Eng Syst Saf* 2023;234:109204.
- [37] Jiang Q, Zhong S, Cui J, Feng X-T, Song L. Statistical characterization of the mechanical parameters of intact rock under triaxial compression: An experimental proof of the Jinping marble. *Rock Mech Rock Eng* 2016;49:4631–46.
- [38] Sen S, Kundan A, Kumar M. Modeling pore pressure, fracture pressure and collapse pressure gradients in offshore panna, western India: implications for drilling and wellbore stability. *Nat Resour Res* 2020;29:2717–34.
- [39] Zhang L, Yan X, Yang X, Tian Z, Yang H. An elastoplastic model of collapse pressure for deep coal seam drilling based on Hoek-Brown criterion related to drilling fluid loss to reservoir. *Journal of Petroleum Science and Engineering* 2015;134:205–13.
- [40] Ma T, Zhang Q, Chen P, Yang C, Zhao J. Fracture pressure model for inclined wells in layered formations with anisotropic rock strengths. *Journal of Petroleum Science and Engineering* 2017;149:393–408.
- [41] Lahann RW, Swarbrick RE. Fracture pressure, leak-off tests and Poisson's ratio. *Pet Geosci* 2022;28. petgeo2021-103.
- [42] Handin J. On the Coulomb-Mohr failure criterion. *J Geophys Res* 1969;74:5343–8.
- [43] Breckels I, Van Eekelen H. Relationship between horizontal stress and depth in sedimentary basins. *J Pet Technol* 1982;34:2191–9.
- [44] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* 1958;65:386.
- [45] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *nature* 1986;323:533–6.
- [46] Liu Y, Hu T, Zhang H, Wu H, Wang S, Ma L, et al. arXiv preprint arXiv: 231006625. 2023.
- [47] Joseph VR. Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal* 2022;15:531–8.
- [48] Yao X. A new simulated annealing algorithm. *Int J Comput Math* 1995;56:161–8.
- [49] Song Z, Song Y, Yang J, Liu B, Gao B, Tang J. A multi-objective reinforcement learning framework for real-time drilling optimization based on symbolic regression and perception. *Geoenergy Sci Eng* 2025;244:213392.
- [50] Chen C, Luo Y, Liu J, Yi Y, Zeng W, Wang S, et al. Joint sound denoising with EEMD and improved wavelet threshold for real-time drilling lithology identification. *Measurement* 2024;238:115363.
- [51] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 2014;15:1929–58.
- [52] Chen Z, Li X, Wang W, Li Y, Shi L, Li Y. Residual strength prediction of corroded pipelines using multilayer perceptron and modified feedforward neural network. *Reliab Eng Syst Saf* 2023;231:108980.
- [53] Lima JPS, Evangelista F, Guedes, Soares C. Hyperparameter-optimized multi-fidelity deep neural network model associated with subset simulation for structural reliability analysis. *Reliab Eng Syst Saf* 2023;239:109492.
- [54] Kim H, Youn BD. A new parameter repurposing method for parameter transfer with small dataset and its application in fault diagnosis of rolling element bearings. *IEEE Access* 2019;7:46917–30.
- [55] Siahkoohi A, Louboutin M, Herrmann FJ. The importance of transfer learning in seismic modeling and imaging. *Geophysics* 2019;84:A47–52.
- [56] Luo Y, Yin L, Bai W, Mao K. An appraisal of incremental learning methods. *Entropy* 2020;22:1190.
- [57] Hossain ME. Drilling costs estimation for hydrocarbon wells. *J Sustainable Energy Eng* 2015;3:3–32.
- [58] Chen X, Gao D, Guo B, Feng Y. Real-time optimization of drilling parameters based on mechanical specific energy for rotating drilling with positive displacement motor in the hard formation. *J Nat Gas Sci Eng* 2016;35:686–94.
- [59] Hao Z, Di Maio F, Zio E. A sequential decision problem formulation and deep reinforcement learning solution of the optimization of O&M of cyber-physical energy systems (CPESS) for reliable and safe power production and supply. *Reliab Eng Syst Saf* 2023;235:109231.
- [60] Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International conference on machine learning. PMLR; 2018. p. 1861–70.
- [61] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. *Advances in neural information processing systems* 2017;30.
- [62] Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting. In: Proceedings of the AAAI conference on artificial intelligence; 2021. p. 11106–15.
- [63] Hornik K. Approximation capabilities of multilayer feedforward networks. *Neural Networks* 1991;4:251–7.
- [64] Xu J, Sun X, Zhang Z, Zhao G, Lin J. Understanding and improving layer normalization. *Advances in neural information processing systems* 2019;32.
- [65] Wang D, Zhang W, Zhang X, Zhao G, Zuo R, Ni J, et al. Technical Economical Analysis. The China Continental Scientific Drilling Project: CCS-1 Well Drilling Engineering and Construction 2015:349–59.
- [66] Teale R. The concept of specific energy in rock drilling. *International journal of rock mechanics and mining sciences & geomechanics abstracts* 1965:57–73.
- [67] Dupriest FE, Koederitz WL. Maximizing drill rates with real-time surveillance of mechanical specific energy. In: SPE/IADC Drilling Conference and Exhibition: SPE; 2005. SPE-92194-MS.
- [68] Guo Hj, Luo Hd, Zhan Gd, Wang Bd, Zhu S. A Real-Time Friction Prediction Model for in Service Drill String Based on Machine Learning Methods Coupling with Mechanical Mechanism Analysis. In: SPE Middle East Oil and Gas Show and Conference. SPE; 2021. D031S3R02.