

Optimization of Multi-Objective Real-Time Drilling Operations Considering Lag Effects and Formation Variability

Yu Song^a, Zehua Song^a, Jin Yang^b, Kejin Chen^b, Jizhou Tan^c

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

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

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

Abstract

In the context of global energy structure transformation, the accurate optimization of offshore oil and gas drilling parameters is crucial for enhancing operational efficiency and ensuring drilling safety. Traditional optimization methods rely primarily on historical data from drilled wells, resulting in models that often exhibit low predictive accuracy and limited generalizability in unexplored formations. Moreover, conventional methods depend on logging-while-drilling (LWD) parameters for decision-making; however, the delay in data acquisition and processing leads to significant latency effects in adjusting drilling parameters, thereby affecting the timeliness and precision of decisions. This study introduces a learning framework that focuses on a multi-objective drilling parameter approach using real-time optimization reinforcement that combines time-series forecasting networks, domain adversarial networks, and Markov decision processes to accurately predict drilling speed and formation pressure constraints while optimizing drilling parameters in real time. The framework utilizes multilayer convolutional networks as feature extractors in conjunction with domain discriminators to adapt to different geological environments, and a drilling parameter optimization actor–critic (DPOAC) algorithm is designed to automatically adjust drilling parameters in real time. Empirical analysis demonstrates that this framework significantly enhances the drilling speed and economic benefits in the Caofidian 6-4 block of the Bohai Sea in China. Drilling speed increased from an initial 58.76 m/hr to 210.81 m hr, mechanical specific energy (MSE) was reduced from 11.42 MPa to 10.01 MPa, and the unit cost per foot decreased from 184.39×10^2 CNY/m to 28.52×10^2 CNY/m, showcasing its wide application potential in modern drilling operations.

Keywords: Drilling Parameter Real-time Optimization; Drilling Sensing; Domain Adversarial Network; Deep Reinforcement Learning; Efficiency Improvement and Cost Reduction

1. Introduction

In the context of the global shift towards a transformed energy structure and low-carbon development strategy, offshore drilling technology is pivotal for developing marine oil and gas resources, which is crucial for ensuring energy security and achieving sustainable development [1-3]. However, offshore drilling operations face significant challenges owing to extreme environmental conditions, complex geological structures, and high economic costs. These factors make the precise optimization of drilling parameters a central task for ensuring safety, enhancing operational efficiency, and minimizing drilling risks [4]. Thus, exploring effective and reliable

1 strategies for optimizing drilling parameters has substantial theoretical and practical value for advancing marine
2 drilling technologies and optimizing operations.

3 In the field of drilling parameter optimization, numerous researchers have dedicated efforts to exploring
4 various methods for enhancing drilling efficiency. Kendall and Goins studied the drilling hydraulic parameters,
5 established relationships among the maximum impact force, maximum hydraulic power, and maximum jet
6 velocity, and developed a hydraulic parameter optimization program [5]. Eckel further explored the impact of
7 drilling fluid performance and hydraulic parameters on drilling speed by summarizing the correlations between
8 the drilling fluid density, kinematic viscosity, and changes in the hydraulic parameters [6]. Bourgoyn and
9 Young utilized multiple regression analysis to examine the effects of drill pressure, revolutions per minute, bit
10 wear, hydraulic parameters, and drillability coefficients on drilling speed and proposed a multiple regression
11 drilling speed equation [7]. More recently, Waughman et al. identified inefficient drilling scenarios by comparing
12 real-time calculated mechanical specific energies with standard values [8]. Dupriest and Koederitz employed
13 real-time mechanical specific energy to assess drill bit penetration efficiency, while Iversen et al. developed a
14 real-time drilling parameter optimization method [9, 10]. Zhang et al. noted that reducing the nozzle diameter
15 and increasing the jet velocity significantly enhanced the drilling efficiency by increasing the fluid jet speed and
16 penetration capabilities, aligning with the early findings of ExxonMobil [11]. Sui et al. developed a multivariable
17 extremum-seeking method for real-time optimization of drilling parameters to minimize the MSE and enhance
18 drilling efficiency and safety [12]. Guria et al. applied an elitist non-dominated sorting genetic algorithm for
19 multi-objective drilling parameter optimization, demonstrating the potential of multi-objective optimization [13].
20 Zhang et al. introduced an improved multi-objective particle swarm optimization algorithm targeting the drilling
21 rate, bit life, and mechanical specific energy to avoid local optima and display good convergence and distribution
22 [14]. Moreover, Zheng et al.'s multi-objective cellular particle swarm optimization algorithm and Khaleel et al.
23 's application of genetic algorithms to estimate coefficients for the Bourgoyn and Young model significantly
24 enhanced drilling efficiency, reduced operational costs, and ensured more effective drilling operations [15, 16].

25 Despite significant exploratory progress in drilling parameter optimization using intelligent computing
26 methods, neural network black-box models and multi-objective optimization algorithms still exhibit significant
27 limitations. Neural network models can capture complex nonlinear relationships; however, their opaque nature
28 restricts their interpretability and generalizability [17]. Particularly concerning is that most drilling speed
29 prediction models are trained on data from previously drilled wells, and when applied to unexplored formations
30 in target wells, they often demonstrate lower prediction accuracy and generalizability. This phenomenon usually
31 stems from the inadequate adaptation of the models to new geological conditions or changes in drilling
32 operational parameters [18]. Furthermore, relying solely on real-time logging data for optimization has inherent
33 limitations. Once data are acquired, corresponding drilling activities have already occurred, often causing real-
34 time drilling parameter adjustments to fall short of expectations. Such methods do not adequately consider the
35 latency effects in drilling operations and the need for dynamic adaptation of drilling parameters at different

1 drilling times [19, 20]. Additionally, multi-objective optimization during drilling faces numerous challenges
2 owing to the complexity and uncertainty of geological environments, which impart a high degree of dynamism
3 and nonlinearity to drilling-parameter optimization [21]. Traditional algorithms, such as heuristic optimization
4 algorithms, are therefore inefficient and inadequate for the real-time adjustment of drilling parameters to cope
5 with formation changes, and struggle to balance multiple objectives, such as enhancing drilling speed,
6 controlling bit wear, and optimizing energy consumption.

7 Drilling perception, transfer, online learning, and reinforcement learning offer new solutions to the
8 challenges of drilling parameter optimization. Drilling perception employs time-series prediction models that
9 analyze both historical and current logging-while-drilling (LWD) parameters to predict future drilling parameters,
10 allowing for the pre-optimization of future drilling conditions [22, 23]. This approach enhances the adaptability
11 of models by introducing domain-adversarial networks and incremental parameter tuning, enabling models to
12 quickly adjust by continually learning from new drilling data, thereby improving the accuracy of prediction
13 models in the unexplored strata of target wells [24, 25]. Reinforcement learning, a trial-and-error learning
14 mechanism-based machine learning method, demonstrates tremendous potential for drilling-parameter
15 optimization [26, 27]. Through interactions with the environment, reinforcement learning algorithms learn
16 decision-making strategies for optimal drilling operations in various drilling processes, ensuring the safety and
17 economic efficiency of drilling operations while effectively enhancing the drilling speed. Thus, combining
18 drilling perception, transfer, online learning, and reinforcement learning can create a more accurate,
19 generalizable, and adaptive drilling parameter optimization framework to effectively address the complex
20 challenges in offshore drilling.

21 This study introduces a comprehensive multi-objective framework for real-time optimization of drilling
22 parameters by integrating time-series forecasting networks, domain adversarial networks, and Markov decision
23 processes (MDPs) to precisely predict drilling speeds and optimize drilling parameters in real time. Initially, the
24 mechanical properties of the rock were assessed through triaxial compression tests, and the formation pressure
25 gradients were calculated using the Coulomb-Mohr criterion. Drilling rate prediction and formation property
26 awareness were addressed using a fully connected neural network and the iTransformer model, which detects
27 changes in geological parameters using data from fixed time windows. Additionally, a domain adversarial
28 mechanism was designed for strata not yet drilled into the target wells, employing a multilayer convolutional
29 network as a feature extractor, which works in collaboration with a domain discriminator to adapt to new
30 geological environments. Simultaneously, a Markov decision environment for drilling parameter optimization
31 was established, and the drilling parameter optimization actor-critic (DPOAC) algorithm was developed,
32 utilizing Q-, target, and policy networks to adjust drilling parameters and enhance economic and mechanical
33 benefits, while boosting the real-time adaptability of the model to geological changes. This integrated framework
34 facilitates automated and instantaneous adjustments of drilling parameters, significantly enhancing the economic
35 and mechanical efficiency of drilling operations. The evaluation results for the Caofidian 6-4 block of the Bohai

1 Sea demonstrated the framework's significant potential to enhance the efficiency of drilling applications,
2 providing technical reserves and support for the development and practice of future offshore drilling operations.

3 **2. Workflow**

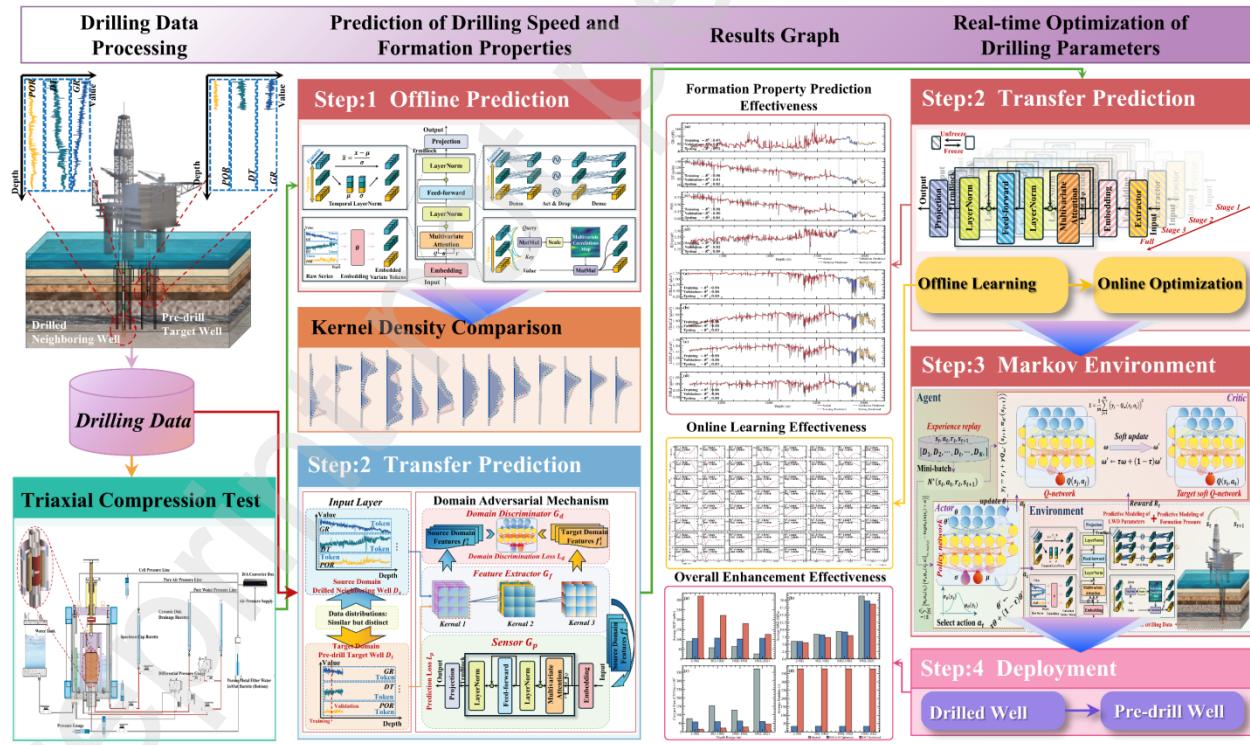
4 **Offline sensing of drilling speeds and formation properties in already-drilled neighboring wells:** Rock
5 mechanical properties at various depths were assessed through triaxial compression tests. Using the Coulomb-
6 Mohr criterion, the results were precisely modeled to calculate the formation pressure gradients, providing
7 essential data for evaluating wellbore stability. Two core models, a fully connected neural network (FCNN) and
8 an iTransformer, were designed for the perception of the drilling speed and formation properties. Based on
9 operational and geological parameters such as weight on bit (WOB), revolutions per minute (RPM), pump flow
10 rate (FLW), and mud input density (MwIN), gamma ray (GR), delta time (DT), porosity (POR), and rock density
11 (RD), the FCNN model predicted drilling speeds. Using fixed-time-window time-series data, the iTransformer
12 model detects short-term changes in geological parameters and expanded this capability to include complex
13 formation pressure perception. Adaptive learning rate adjustments and an early stopping strategy were utilized
14 in model training and evaluation to enhance the predictive performance and generalizability of the model in
15 diverse geological settings.

16 **Online Sensing and Transfer of Drilling Speeds and Formation Properties in Undrilled Target Wells:**
17 Models optimized for neighboring wells were employed as sensors for the transfer and online perception of
18 drilling speed and formation properties in the unexplored strata of target wells. A domain adversarial mechanism
19 incorporating multilayer convolutional networks as feature extractors to distinguish features from high-
20 dimensional data was designed to train a binary domain discriminator. To align the feature distributions between
21 the drilled neighboring wells and the unexplored strata of the target well, the feature extractor and domain
22 discriminator were cooperatively engaged in a min-max game. Model parameter freezing and fine-tuning
23 techniques were applied to adapt to the new geological environments, thereby enhancing the transferability of
24 the model. Using data from neighboring wells, the models were pre-trained, followed by incremental unfreezing
25 of the key layers to precisely adapt to the specific geological conditions of the unexplored strata of the target
26 wells. An online learning strategy, coupled with depth-triggered conditions and a forgetting mechanism, was
27 implemented to ensure real-time model updates, thereby optimizing the long-term learning efficiency and
28 prediction accuracy.

29 **Design of a Markov Decision Process Environment for Real-Time Drilling Parameters:** A Markov
30 decision process (MDP) environment was developed specifically for optimizing drilling parameters to precisely
31 control drilling operations and enhance both economic and mechanical efficiency. This environment
32 incorporates a transfer prediction model to estimate drilling speeds in the unexplored strata of target wells,
33 enabling real-time evaluation of the cost per foot and mechanical specific energy for various combinations of
34 drilling parameters. The reward function was designed to assess the economic and efficiency impacts of

1 adjustments in the drilling parameters accurately, thereby promoting an optimal balance between cost and
 2 equipment efficiency. The design of the state and action spaces integrates the models for the transfer of drilling
 3 parameter perception and formation pressure sensing, continuously updating the drilling parameters and
 4 formation pressures using real-time data. This setup provides essential dynamic inputs to the MDP environment
 5 while ensuring that the wellbore stability constraints are satisfied, thereby optimizing drilling operations under
 6 dynamic geological conditions and ensuring safety.

7 **Implementation and Deployment of Deep Reinforcement Learning Algorithms:** In conjunction with
 8 an MDP environment, the DPOAC algorithm was designed to optimize the drilling parameters using finely
 9 configured Q-, target, and policy networks. This algorithm enhances the economic benefits and mechanical
 10 efficiency of drilling operations while ensuring the automation and precision of the decision-making process.
 11 The DPOAC algorithm utilizes experience replay and priority-adjustment mechanisms to improve learning
 12 efficiency, and employs exploration noise strategies to enhance the ability of the model to adapt to changes in
 13 geological conditions in real time during actual drilling operations. Additionally, this process was validated in
 14 the Caofidian 6-4 block in the Bohai Sea, China, where continuous adjustments of drilling parameters achieved
 15 an optimized balance between drilling speed and cost, thereby demonstrating the practical potential of this
 16 method to enhance the economic and operational efficiency of drilling. The workflow is illustrated in Fig. 1.



17
 18 **Fig. 1. Workflow**

1 **3. Offline Perception of Drilling Speed and Formation Properties in Adjacent**
2 **Wells**

3 Conventional methods for optimizing drilling parameters have not sufficiently accounted for the latency
4 effects inherent in drilling operations, primarily relying on real-time LWD parameters for decision making.
5 However, owing to the time lag in data acquisition and processing, drilling activities often proceed before the
6 completion of data analysis, resulting in real-time drilling parameter adjustments that fail to achieve the intended
7 effects. To address this, a perception-based method has been proposed that involves predicting the LWD
8 parameters in advance to preemptively optimize the drilling parameters, thus more effectively adapting to
9 dynamic changes during the drilling process and enhancing the drilling efficiency and precision. This analysis,
10 based on data from a previously drilled well in the Caofeidian 6-4 block of the Bohai Sea, China, involved
11 predictions related to the drilling speed, LWD parameter perception, and formation pressure sensing. It primarily
12 utilized pre- and mid-drilling data, excluding post-drilling data. Pre-drilling and mid-drilling data provide direct
13 observations and measurements that are crucial for real-time decision-making and drilling-speed optimization.
14 In contrast, although post-drilling data offer comprehensive value for retrospective analysis, they are generally
15 processed and have lower timeliness, making them unsuitable for immediate response needs in the model.

16 For drilling speed prediction, the analysis focuses on drilling parameters, such as WOB, RPM, FLW, and
17 MwIN, which directly reflect the physical conditions of drilling operations and are critical for optimizing drilling
18 speed and process control. With respect to the LWD parameter perception, the dataset concentrates on GR, DT,
19 POR, and RD, providing essential data for the real-time assessment of the lithology, electrical, and acoustic
20 properties of the encountered formations, which are crucial for formation evaluation. Formation pressure
21 perception encompasses LWD data and rock mechanics test data, such as the Poisson's ratio, cohesion,
22 overburden pressure, and internal friction angle of rocks, including data on collapse pressure, fracture pressure,
23 loss pressure, and pore pressure. These parameters, derived from rock mechanics tests and calculations, offer an
24 in-depth understanding of the formation mechanical behavior and real-time acquisition of mechanical properties.

25 **3.1 Determination Methods for Formation Pressure Parameters**

26 **3.1.1 Measurement of Rock Mechanical Parameters**

27 Triaxial compression testing is a fundamental experimental method for assessing the mechanical properties
28 of rocks. By applying controlled axial and radial stresses, this test simulates the stress conditions that rocks
29 experience at various subterranean depths, enabling the determination of critical mechanical parameters such as
30 elastic modulus, Poisson's ratio, cohesion, and internal friction angle. These parameters are essential for
31 understanding and predicting the behavior of rocks in actual geological environments [28]. The experimental
32 setup is illustrated in Fig. 2.

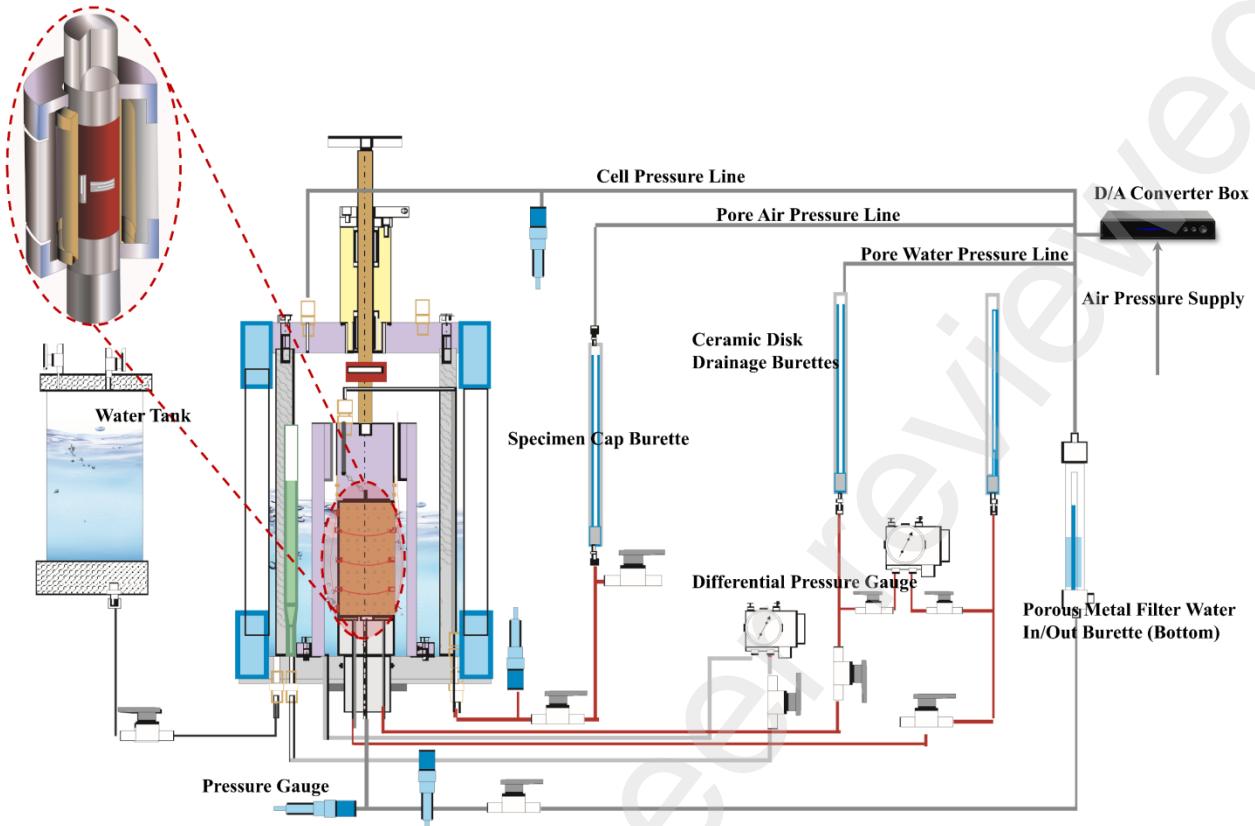


Fig. 2. Triaxial compression test apparatus

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

1 uniaxial tests, a hydraulic press directly applied an axial load to the rock samples, and the data acquisition system
2 continuously recorded the stress and strain until fracture. Triaxial tests began with the application of
3 predetermined confining pressures using a high-pressure pump to simulate actual subsurface stress conditions.
4 Once the confining pressure stabilized, the axial pressure was increased, and the stress and strain data were
5 recorded until the sample fractured. The analysis of the stress and strain data revealed the elastic modulus and
6 Poisson's ratio of the rocks, whereas data near the yield point were used to calculate the ultimate strength,
7 cohesion, and internal friction angle of the rocks. These precise mechanical parameters provide a scientific basis
8 for formation pressure calculations and the assessment of formation stability during drilling.

9 **3.1.2 Calculation of Formation Pressure Parameters**

10 Accurately calculating the formation collapse pressure, pore pressure, fracture pressure, and loss pressure
11 is crucial for maintaining wellbore stability and optimizing the drilling fluid parameters [29]. The core issue with
12 formation collapse pressure is that insufficient fluid column pressure may fail to support the stress state of the
13 surrounding rock, potentially leading to rock shear failure, which in turn reduces the drilling speed and may
14 cause severe drilling accidents [30]. The fracture pressure is related to the initiation of cracks or the propagation
15 of existing fractures in the exposed formation under drilling fluid pressure. This typically occurs because of
16 excessive mud density in the wellbore, which causes the circumferential stress to exceed the tensile strength of
17 the rock. Accurate calculation of fracture pressure is critical for selecting the maximum mud density that directly
18 affects wellbore stability and safety [31]. The accurate determination of the loss pressure is particularly crucial
19 in fault zones, where faults act as weak planes in the rock mass. The mechanical strength of faults is generally
20 lower than that of the surrounding rock, leading to significant discontinuities in the stress field and deformation
21 near the faults. These discontinuities can easily trigger cracks and fissures near faults, causing fluid loss and,
22 consequently, wellbore instability and reduced drilling efficiency [32].

23 The calculation of the formation collapse pressure was based on the Coulomb-Mohr strength criterion,
24 which provides the conditions for rock shear failure [33]. The specific expressions are given by Eq. 1.

$$25 \quad \begin{cases} \tau \geq C + \mu_f \sigma \\ \mu_f = \tan(\Phi) \end{cases} \quad (2)$$

26 where τ denotes the shear stress in MPa; C represents the cohesion of the rock in MPa; Φ is the internal
27 friction angle of the rock in degrees; and σ denotes the normal stress on the rock in MPa.

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

$$1 \quad \begin{cases} \sigma_H = \sigma_h \cot^2(45^\circ - \Phi/2) + 2C \cot(45^\circ - \Phi/2) \\ \sigma_c = \frac{2C \cos(\Phi)}{1 - \sin(\Phi)} \end{cases} \quad (3)$$

2 where σ_H denotes the maximum horizontal principal stress in MPa; σ_h the minimum horizontal principal
3 stress in MPa; and σ_c the uniaxial compressive strength in MPa.

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

$$6 \quad (\sigma_H - \alpha P_p) = (\sigma_h - \alpha P_p) \cot^2(45^\circ - \Phi/2) + 2C \cot(45^\circ - \Phi/2) \quad (4)$$

7 where P_p denotes the pore pressure in the rock in MPa; and α is the effective stress coefficient, $0 < \alpha$, 1
8 (dimensionless).

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

$$13 \quad \begin{cases} \sigma'_r = p_i - ap_p \\ \sigma'_\theta = \eta(3\sigma_H - \sigma_h - p_i) - ap_p \\ \tau_{r\theta} = 0 \end{cases} \quad (5)$$

14 where σ'_r and σ'_θ denote the minimum and maximum effective principal stresses at the collapse point in
15 MPa, respectively; η is the stress nonlinear correction coefficient (dimensionless); and $\tau_{r\theta} = 0$ indicates that
16 the tangential shear stress is zero under the considered stress state, applicable to axisymmetric conditions, in
17 MPa.

18 Using the Coulomb-Mohr criterion, the wellbore collapse pressure gradient is given by Eq. 5.

$$19 \quad \rho_m = \frac{\eta(3\sigma_H - \sigma_h) - 2CK + aP_p(K^2 - 1)}{(K^2 + \eta)H} \quad (6)$$

1 where ρ_m denotes the formation collapse pressure gradient in g/cm³; σ_H and σ_h are the maximum and
2 minimum horizontal principal stresses in MPa; and H is the well depth in m.

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

7
$$\sigma_\theta = 3\sigma_H - \sigma_h - p_i \quad (7)$$

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

11
$$\rho_f = \frac{\left(\frac{2\mu}{1-\mu} - K\right)(\sigma_V - \alpha P_p) + \alpha P_p + S_t}{H} \quad (8)$$

12 where ρ_f is the formation fracture pressure gradient in g/cm³; S_t is the formation tensile strength in MPa,
13 σ_V is the overburden pressure in MPa, μ is Poisson's ratio (dimensionless). Formation loss pressure
14 approximates the minimum horizontal principal stress with sufficient engineering accuracy [34]. The
15 calculations are given by Eq. 8.

16
$$\rho_l = \frac{\sigma_h}{H} \quad (9)$$

17 where ρ_l is the formation loss pressure gradient in g/cm³.

18 **3.2 Design of Offline Models for Drilling Speed and Formation Property Perception in
19 Neighboring Wells**

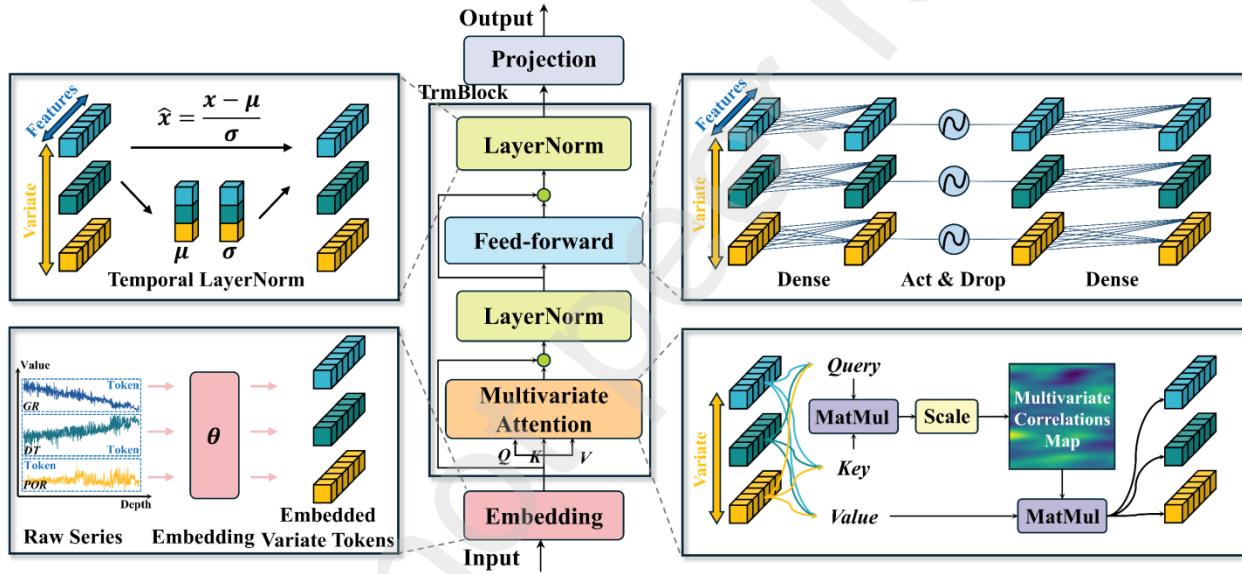
20 **3.2.1 Fully Connected Neural Network Model**

21 An FCNN comprises multiple layers of neurons, each connected to every neuron in the preceding layer,
22 hence the term "fully connected." This architecture is suitable for learning nonlinear and complex function
23 mappings from fixed-size input data [35]. It comprises an input layer, multiple hidden layers, and an output
24 layer. The input layer receives preprocessed drilling parameters: WOB, RPM, FLW, and MwIN. Each hidden
25 layer contains neurons that perform nonlinear transformations on the inputs via activation functions that
26 gradually extract higher-level features. The output layer produces the predicted drilling speed (ROP). In an

1 FCNN, forward propagation refers to the process in which data flow from the input layer to the output layer,
 2 with each layer's output serving as the input for the next layer. The model training relies on backpropagation,
 3 which is an efficient gradient descent method designed to optimize network weights and minimize prediction
 4 errors. During training, the loss function evaluates the deviation between the model predictions and actual data
 5 by adjusting the network parameters to minimize this deviation [36].

6 3.2.2 iTransformer Model

7 For formation property prediction, the iTransformer model provides an efficient architecture for predicting
 8 multivariable time-series data within specific time windows, such as GR, DT, POR, and RD [37]. This model
 9 effectively processes and analyzes these complex datasets to achieve accurate geological feature perception. The
 10 architecture of the model is illustrated in Fig. 3.



11
 12 **Fig. 3. Architecture of the iTransformer model**

13 **Model Architecture:** The iTransformer model employs an encoder-only architecture, omitting the
 14 traditional decoder, and focuses on using the encoder to directly process and predict time-series data [38]. This
 15 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
 16 $X = \{x_1, \dots, x_T\} \in \mathbb{R}^{T \times N}$ without an additional decoding process. In this model, $X_{(t,i)}$ represents the time points
 17 recorded simultaneously at time step t , and $X_{(:,n)}$ denotes the entire time series of each variable indexed by n .
 18 The iTransformer performs data embedding and feature projection tasks through encoder blocks composed of
 19 multilayer perceptrons, handles multivariable interactions in the time series via a self-attention mechanism, and
 20 effectively captures dynamic correlations between variables. The iTransformer can process time series with
 21 complex geological and physical properties to predict the formation pore pressure, fracture pressure, and collapse
 22 pressure during drilling. Sequential information is implicitly stored in the arrangement of the neurons in the

1 preceding network, which simplifies the complex positional encoding required by traditional transformers.

2 **Self-Attention Mechanism:** In formation property perception studies, the iTransformer model emphasizes
3 the utilization of the self-attention mechanism to process and predict time-series data. The model dynamically
4 adjusts its focus on various variables across different time steps, such as GR, DT, POR, and RD, using self-
5 attention layers. This mechanism allows the model to automatically identify and emphasize the most critical
6 variables for predicting subsequent formation properties such as the pore, fracture, collapse, and loss pressures.
7 The self-attention layers calculate the interdependencies between different variables to optimize the prediction
8 accuracy [39]. In this process, the multivariable data at each time series point are transformed into high-
9 dimensional vectors, each linearly transformed to obtain the corresponding queries (Q), keys (K), and values
10 (V), as shown in Eq. 9.

11
$$Q = HW^Q, K = HW^K, V = HW^V \quad (10)$$

12 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 ,
13 W^K , W^V are the weight matrices mapping the input to their respective spaces. The model calculates the
14 interactions between each pair of Q and K vectors, normalized using the softmax function, to obtain attention
15 weights, as shown in Eq. 10.

16
$$A = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) \quad (11)$$

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

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

25 **Embedding and Projection:** In iTransformer, the primary function of the embedding layer is to transform
26 the input multivariable time-series data into vector representations in a high-dimensional space. This process is
27 illustrated in Eq. 11.

1
$$h_n^0 = \text{Embedding}(X_{(:,n)}) \quad (12)$$

2 where h_n^0 denotes the vector processed by the embedding layer. These vectors interact through the self-
 3 attention mechanism and learn the complex relationships between the variables. Each vector is first mapped to
 4 a fixed-dimensional embedding space, aiding in standardizing the processing of the input data. The projection
 5 layer of the model converts the encoded high-dimensional features back to their original data dimensions, as
 6 expressed in Eq. 12.

7
$$Y_{(:,n)}^\wedge = \text{Projection}(h_n^L) \quad (13)$$

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

12 In the iTransformer architecture, the interactions between variables are achieved using a self-attention
 13 mechanism and independently processed and shared into feedforward networks (FFNs) in each TrmBlock. The
 14 sequence order is implicitly stored by the arrangement of neurons in the feedforward network, thereby
 15 eliminating the need for positional embedding. This process is illustrated in Eq. 13.

16
$$H^{(l+1)} = \text{TrmBlock}(H^l), l = 0, \dots, L-1 \quad (14)$$

17 where $H = \{h_1, \dots, h_N\} \in \mathbb{R}^{N \times D}$ and contains N embedding vectors with dimension D , the superscript
 18 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
 19 implemented through multilayer perceptrons.

20 **Feedforward Network and Layer Normalization:** In the iTransformer model, each encoding block
 21 integrates a FFN to process the output from the self-attention layer. The FFN enhances the ability of the model
 22 to represent time-series data through dense nonlinear transformations [40] and is given by Eq. 14.

23
$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (15)$$

24 where W_1 , W_2 , and b_1 , b_2 denote the weights and biases of the network layers, respectively. This network
 25 structure ensures that each data point is processed independently, strengthening the ability of the model to handle

1 the variability of complex geological data. In addition, the FFN effectively captures the geological features
2 derived from historical data windows and predicts subsequent geological changes.

3 Layer Normalization (LayerNorm) stabilizes the training process and accelerates model convergence. This
4 technique standardizes each feature within a layer and reduces the internal covariate shifts during training [41].
5 This process is illustrated in Eq. 15.

6

$$\text{LayerNorm}(x) = \frac{x - \mu}{\sigma} \quad (16)$$

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

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

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

20 **3.3 Model Training and Evaluation**

21 To enhance the accuracy of the drilling speed prediction, LWD parameter sensing, and formation pressure
22 perception, refined training strategies were implemented. Different dataset-partitioning strategies were
23 employed for various prediction tasks. Random partitioning was utilized for drilling speed prediction to ensure
24 the randomness of the data distribution and the generalizability of the model. For the LWD parameter sensing
25 and formation pressure perception, an 8:1:1 depth ratio was employed to better simulate the data distribution
26 during actual drilling processes [42]. An adaptive learning rate-adjustment strategy was employed to adjust the
27 learning rate dynamically through learning-rate annealing. This strategy enables rapid convergence with a higher
28 learning rate at the beginning of training, and gradual reduction based on changes in the loss function in later
29 stages [43]. In addition, the AdamW optimizer, which combines adaptive moment estimation and weight decay,
30 was selected to effectively improve the training efficiency and generalizability of complex datasets [44]. The

1 Mean Squared Error (MSE) and Smooth L1 Loss were used as loss functions to directly quantify the deviation
2 between the predicted and actual values, aiding in the precise optimization of the model outputs [45].

3 Dropout layers and regularization techniques were incorporated into the model to prevent overfitting and
4 validate its generalizability. These techniques effectively avoid overfitting the training data, while maintaining
5 training efficacy [46]. In addition, an Early Stopping strategy was employed to dynamically adjust the training
6 epochs based on the performance of the validation set. Training was halted when the performance of the model
7 on the validation set ceased to improve, thereby preventing overtraining and preserving the model's optimal
8 performance. These comprehensive training strategies and techniques ensured high accuracy and reliability in
9 predicting the LWD parameters. The training strategies for the prediction and perception of the offline models
10 are presented in Table 1.

11 **Table 1 Training strategies for prediction and perception of offline models**

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

12

13 **3.3.1 Offline Prediction of Drilling Speed in Adjacent Wells**

14 To optimize the hyperparameters of the FCNN model, grid search, Bayesian optimization, and particle
15 swarm optimization algorithms were employed [47-49]. These three distinct optimization strategies were utilized
16 to systematically adjust and optimize the hyperparameters of the model. Fig. 4 illustrates the changes in loss
17 values during the training and validation phases with iterations for the original model and the FCNN model
18 optimized by the three algorithms.

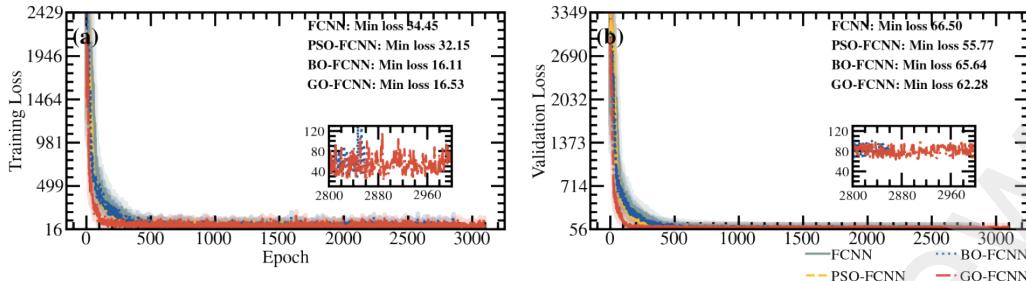


Fig. 4. Dynamics of training and validation losses for drilling speed models in adjacent wells under different optimization algorithms

Fig. 4(a) illustrates the changes in the loss values during the training phase for the original FCNN model and the FCNN models optimized by the grid search (GO-FCNN), Bayesian optimization (BO-FCNN), and particle swarm optimization (PSO-FCNN) models. As depicted in Fig. 4(a), the PSO-FCNN model demonstrates the fastest convergence speed and lowest loss values, indicating that the particle swarm optimization algorithm is highly efficient and accurate in parameter adjustment. The BO-FCNN model exhibits a significant drop in loss values during the initial iterations, but tends to stabilize in the later stages. In contrast, the GO-FCNN model has a slower convergence rate and higher final loss values, possibly because of its large search space and fixed search step size, which lead to inefficiency.

Fig. 4(b) illustrates the changes in the loss values during the validation phase for the GO-FCNN, BO-FCNN, PSO-FCNN, and original models. The PSO-FCNN model also exhibited the lowest loss values in the validation set, validating its generalizability. Although the BO-FCNN model performed well during the training phase, its performance on the validation set was slightly inferior, suggesting that this method may overfit the training data at certain stages. The GO-FCNN model exhibited large fluctuations in the loss values in the validation set, indicating its potential inadequacy in adapting to different datasets. The original model, while having higher loss values in both the training and validation sets, exhibited stable performance, reflecting its potential as a baseline model for handling complex data.

Fig. 5 shows the R^2 fit of the PSO-FCNN model for the training, validation, and test datasets.

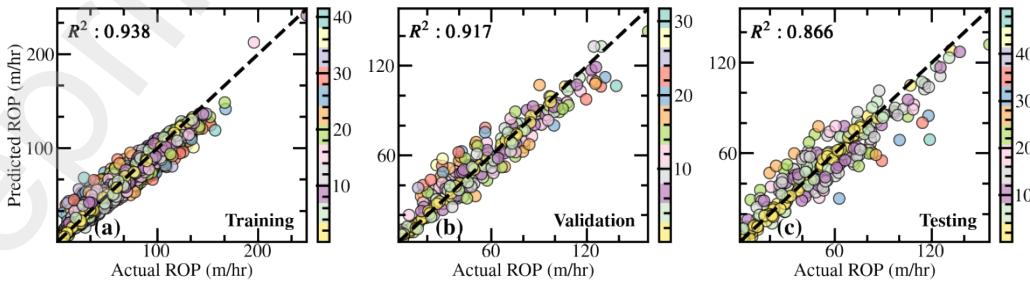


Fig. 5. R^2 fit of drilling speed prediction models for different data in adjacent wells

Fig. 5(a) shows the R^2 -fit scatter plot of the PSO-FCNN model on the training set, where the R^2 score was 0.9378, indicating a high degree of fit with the training data. Most of the scatter points are closely clustered around the fit line, suggesting that the model can accurately predict the drilling speed for most training samples, thereby demonstrating the effectiveness of the hyperparameter optimization. Fig. 5(b) presents the R^2 fit scatter plot of the PSO-FCNN model on the validation set. In this figure, the R^2 score is 0.9172, indicating the good generalizability of the model. Although the fit was slightly lower than that of the training set, high predictive accuracy was maintained. The scatter points are slightly more dispersed than those in the training phase, suggesting that the model faces certain challenges in handling unseen data; however, the overall performance remains stable. Fig. 5(c) presents the R^2 -fit scatter plot of the PSO-FCNN model on the test set, where the R^2 score is 0.8658, indicating the performance of the model on the independent test set. Although the R^2 value was lower than that of the training and validation sets, it still reflected an acceptable predictive capability.

3.3.2 Offline Perception of LWD Parameters in Adjacent Wells

The Long Short-Term Memory (LSTM), transformer, and Informer models were used as baseline comparisons to evaluate the performance enhancements of the iTransformer in the perception of the LWD parameters [38, 39, 50]. Fig. 6 illustrates the dynamics of the loss values during the training and validation phases for LSTM, transformer, Informer, and iTransformer models.

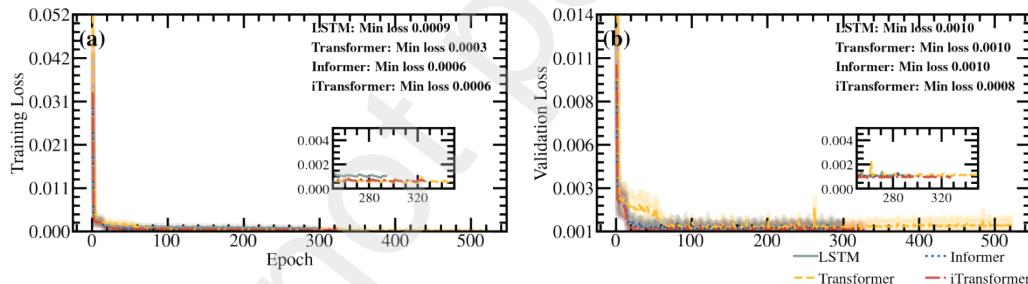


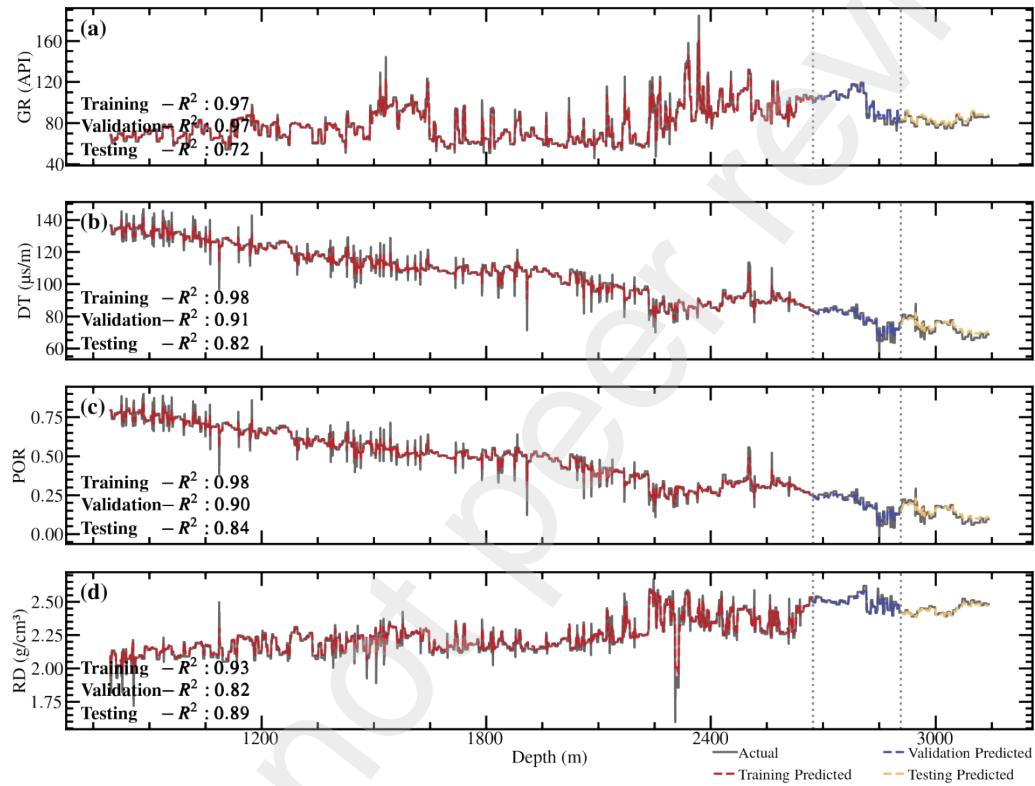
Fig. 6. Dynamics of training and validation losses for offline perception models of LWD parameters in adjacent wells

Fig. 6(a) illustrates the dynamics of the loss values during the training phase for the LSTM, Transformer, Informer, and iTransformer models over the iterations. The iTransformer model exhibited a faster convergence rate and continuously decreased the loss values, demonstrating its efficiency in handling multivariate time-series data. In contrast, while the LSTM and Transformer models showed a rapid initial loss reduction, they experienced fluctuations in loss values during the iterations, suggesting potential performance bottlenecks in long-sequence learning. The Informer model demonstrated an impressive loss reduction speed, but lacked the stability of the iTransformer, reflecting improvements in the iTransformer's ability to capture time-series features.

Fig. 6(b) shows the changes in loss values during the validation phase for these models. The iTransformer model also exhibited the lowest loss values in the validation set, proving its excellent generalizability and

1 adaptability to unseen data. The Transformer and Informer models exhibited increased loss values in the
2 validation phase compared with the training phase, indicating possible overfitting issues. Additionally, although
3 the LSTM model showed higher loss values on the training set, it demonstrated relatively good stability on the
4 validation set, indicating consistency in handling different datasets.

5 Fig. 7 illustrates the performance of the iTransformer model in predicting the key LWD parameters across
6 the training, validation, and test datasets.



7
8 **Fig. 7. Prediction performance of the offline perception iTransformer model for LWD parameters in**
9 **adjacent wells**

10 Regarding GR, both the training and validation phases exhibited R^2 values of 0.97, indicating a high model
11 fit and minimal error in these phases. However, the R^2 value decreased to 0.72 in the testing phase, showing a
12 slight decline in the generalizability of the model for unseen data, although it maintained a relatively reasonable
13 prediction accuracy. For the prediction of DT, the model achieved an R^2 value of 0.98 during the training phase,
14 demonstrating excellent fitting capability. Despite a decrease in R^2 values to 0.91 and 0.82 during the validation
15 and testing phases, respectively, the model continued to show good generalizability, particularly in the analysis
16 of complex geological data. The prediction results for POR in the training phase reached an almost perfect R^2
17 value of 0.98, indicating a near-zero error. Although the R^2 values declined slightly in the validation and testing
18 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.

3.3.3 Offline Perception of Formation Pressure in Adjacent Wells

To evaluate the performance improvement of the iTransformer in formation pressure perception, the LSTM, Transformer, and Informer models were used as baseline comparison groups. Fig. 8 shows the variation in the loss values during the training and validation phases for the LSTM, Transformer, Informer, and iTransformer models.

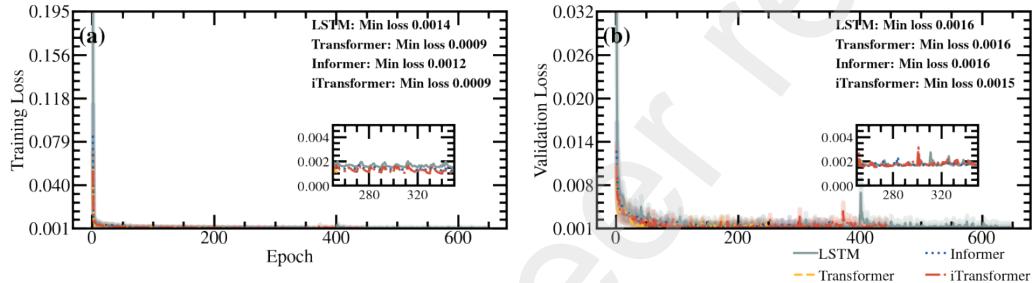
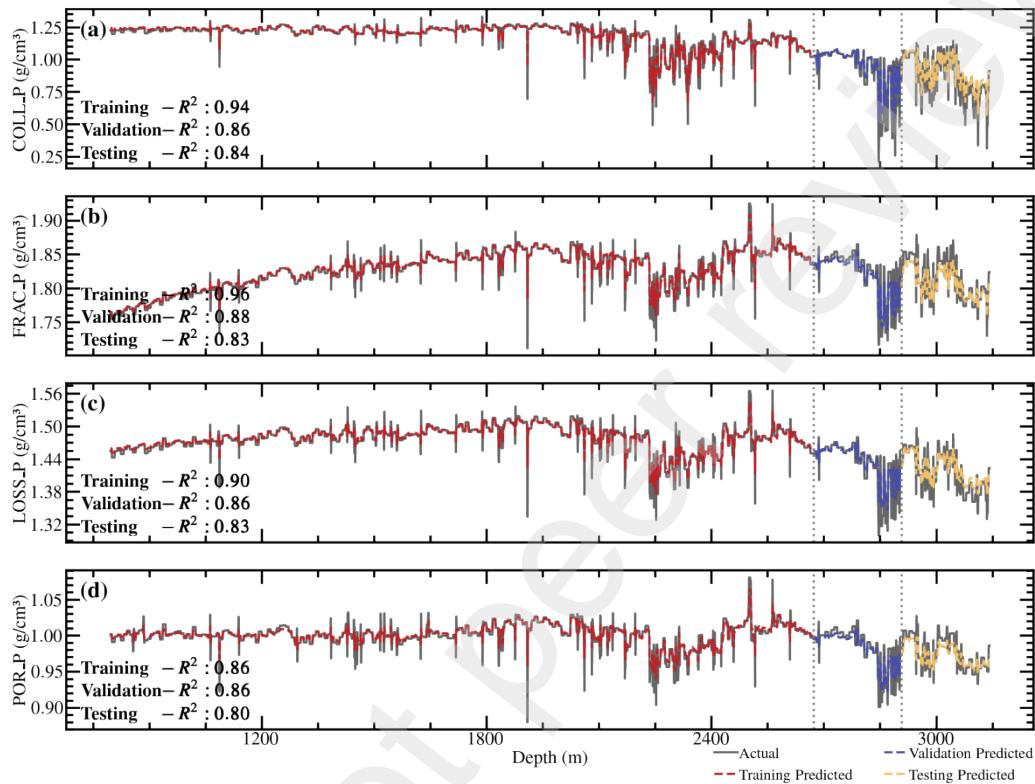


Fig. 8. Dynamics of training and validation losses for offline perception models of formation pressure in adjacent wells

Fig. 8(a) illustrates the changes in the loss values for the LSTM, Transformer, Informer, and iTransformer models during the training phase. Because of its ability to handle more complex input features, the iTransformer model demonstrated significantly lower loss values and faster convergence speeds. This reflects the advantage of iTransformer in capturing nonlinear relationships and multiparameter dependencies in the formation pressure data. In contrast, although the LSTM and Transformer models initially exhibited rapid loss reduction, they later exhibited significant fluctuations, indicating potential limitations in handling high-dimensional features and complex data relationships. Despite having a higher overall loss, the Informer model exhibited a consistent downward trend, indicating its stability in processing long-sequence data.

Fig. 8(b) shows the loss value changes in these models during the validation phase. In this phase, the iTransformer model exhibits the lowest loss values, highlighting its superior generalizability and adaptability to unseen data. This further validates the effectiveness of the iTransformer in formation-pressure prediction, particularly when integrating multiple geomechanical parameters for prediction with high accuracy. Although the Informer model had smaller fluctuations, its overall loss values were higher, indicating the need for further optimization to enhance its performance in processing complex geological data.

1 In the evaluation of the iTransformer model, the performance of the model in predicting the collapse,
 2 fracture, loss, and pore pressures was presented using R^2 values across the three datasets (training, validation,
 3 and testing). Fig. 9 shows the perception effectiveness of the model for the formation pressure across these
 4 datasets.

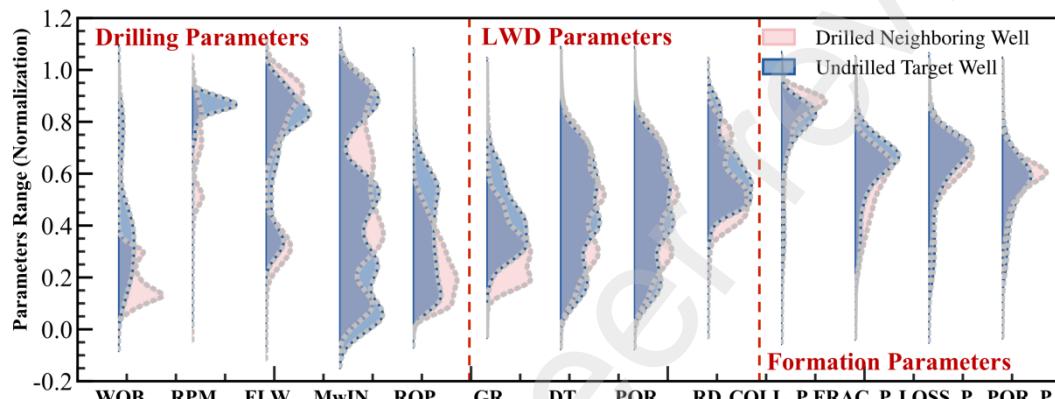


5
 6 **Fig. 9. Prediction performance of the offline perception iTransformer model for formation pressure in**
 7 **adjacent wells**

8 For the collapse pressure, the model performed best during the training phase ($R^2 = 0.94$), with a slightly
 9 lower performance during the validation and testing phases, but still maintained high accuracy. This indicates
 10 the strong capability of the model to fit the data. The results for fracture pressure were similar, with an R^2 of
 11 0.96 during the training phase, with a slight decrease during the validation and testing phases. The high R^2 values
 12 for these two metrics indicate that the model can accurately predict the formation pressure despite slight
 13 performance fluctuations across different datasets. For the loss and pore pressures, although the R^2 values were
 14 high during the training phase, they decreased during the validation and testing phases. In particular, R^2 for the
 15 pore pressure in the testing set was only 0.80, indicating some challenges in the generalizability of the model to
 16 unseen data. Overall, the prediction accuracy of the model was reliable.

17 **4. Transfer and Online Perception of Drilling Speed and Formation Properties**
 18 **for Untapped Target Wells**

1 For the migration and online perception of drilling speed and formation properties in the unopened
 2 formations of the target well, the data were sourced from target well locations within the same block of
 3 neighboring drilled wells in the Caofeidian 6-4 block of the Bohai Sea, China. The data for the open formations
 4 of the target well included 100 records covering depths ranging from 1 to 100 m. This dataset comprises pre-
 5 drilling forecast data and real-time acquisition data during drilling, focusing on direct observations and
 6 measurements during the real-time drilling process as well as predictive analysis for future drilling activities.
 7 Fig. 10 compares the kernel density of the drilling, LWD, and formation pressure parameters between the drilled
 8 neighboring wells and the unopened formations of the target well.



9
 10 **Fig. 10. Comparison of kernel density for key parameters between adjacent wells and untapped target**
 11 **wells**

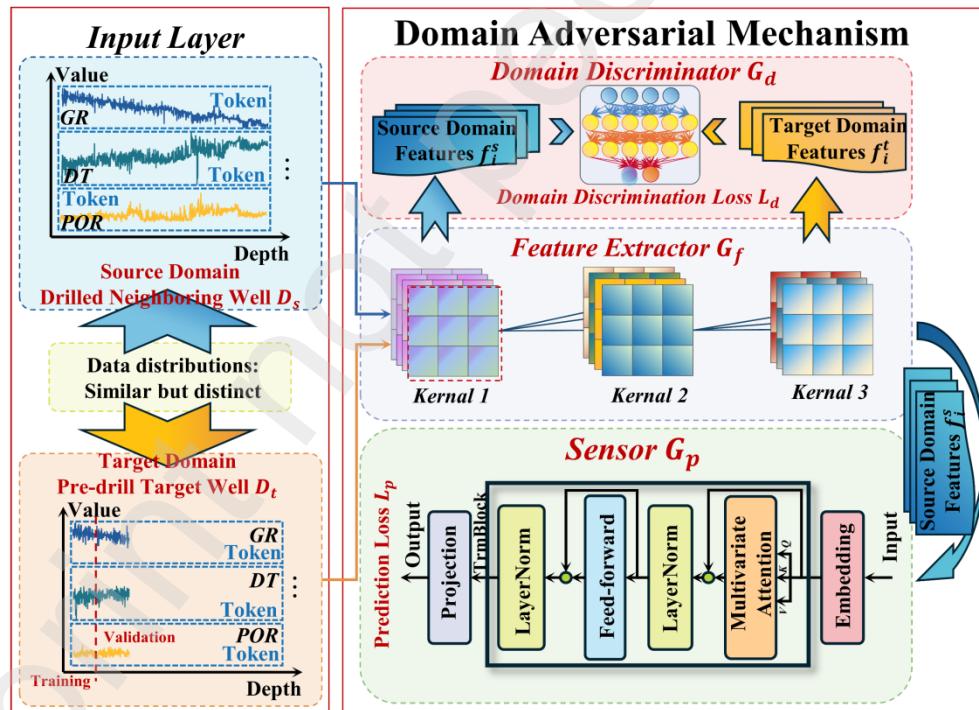
12 Fig. 10 illustrates the kernel density comparison of the drilling and formation pressure parameters between
 13 the drilled neighboring wells and the unopened formations of the target well in the Caofeidian 6-4 block of the
 14 Bohai Sea, China. Because the wells were located within the same block, the kernel density comparison of the
 15 LWD and formation pressure parameters showed minor differences, reflecting the high consistency of the
 16 formation properties and pressure conditions within the block. In contrast, the comparison of drilling parameters
 17 revealed significant differences, highlighting variations in drilling operational conditions, such WOB, RPM, and
 18 FLW. These factors directly affect the drilling efficiency and safety. The significant differences in drilling
 19 parameters may stem from the localized heterogeneity in geological structures or variations in drilling techniques
 20 and operational strategies. Identifying these differences is crucial for adjusting the migration and online learning
 21 models, particularly for addressing the drilling challenges of unopened formations of the target well.

22 **4.1 Transfer of the Perception Model Based on the Domain Adversarial Mechanism**

23 **4.1.1 Architecture of the Domain Adversarial Mechanism**

24 To achieve effective migration and online perception of the drilling speed and formation property
 25 parameters, a domain adversarial mechanism was designed to handle data from drilled neighboring wells and
 26 unopened formations of the target well in the Caofeidian 6-4 block of the Bohai Sea, China. These datasets are

1 considered as the source domain $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{n_s}$ and the target domain $D_t = \{(x_i^t, y_i^t)\}_{i=1}^{n_t}$, where x_i represents
 2 feature data, including drilling and LWD parameters, and y_i indicates the related prediction targets, such as the
 3 drilling speed and formation property parameters. The domain adversarial mechanism encompasses a feature
 4 extractor G_f , a perceiver G_p , and a domain discriminator G_d . The feature extractor G_f uses a multilayer
 5 convolutional network structure to extract abstract and discriminative features from high-dimensional data,
 6 which are crucial for subsequent drilling speed and formation property perception. The feature extractor
 7 processes the data from both the source and target domains, extracting features denoted as f_i^s and f_i^t , which
 8 are then input into the domain discriminator and perceiver. The domain discriminator G_d is designed as a binary
 9 classification network tasked with distinguishing whether the input features are from the source or the target
 10 domain. Thus, the domain discriminator and feature extractor form an adversarial relationship that aims to
 11 accurately distinguish the feature sources, whereas the feature extractor attempts to generate features that are
 12 difficult for the domain discriminator to differentiate, thus aligning the feature distributions of the source and
 13 target domains. The architecture of the domain adversarial mechanism is shown in Fig. 11.



14
 15 **Fig. 11. Architecture of domain adversarial mechanism for adjacent and untapped target wells**

16 The feature extractor G_f played a critical role in processing highly collinear multidimensional drilling data,
 17 ensuring the accurate extraction of key information for tasks, such as drilling speed prediction, LWD parameter
 18 and formation pressure perceptions. To handle these complex data, the feature extractor employs a multilayer
 19 convolutional neural network (CNN) architecture that is specifically optimized for the characteristics of various

1 geological and operational parameters to ensure efficient and accurate feature extraction. Drilling parameters,
2 such as WOB, RPM, and FLW, typically exhibit strong operational dependency. To process these data
3 effectively, feature extractor G_{f1} utilized a deep learning framework with convolutional layers to analyze the
4 dynamic changes and physical properties during drilling operations. By applying multi-sized convolution kernels
5 across multiple convolutional layers, both local and global dependencies of the drilling parameters were captured.
6 Alternating convolution and pooling operations increased the processing depth of the network, reduced feature
7 dimensions, and enhanced computational efficiency and noise resistance [51]. For LWD parameter perception,
8 such as GR, DT, POR, and RD, which contain critical geological information essential for formation evaluation,
9 the feature extractor G_{f2} employed refined convolutional layers to capture subtle variations in the logging
10 parameters, extracting geological features from time-series data to identify formation continuity and
11 heterogeneity. Each convolutional layer was designed to identify geological signals of different frequencies,
12 ensuring the extraction of representative features from the drilling data to support precise formation analysis. In
13 the context of formation pressure perception, feature extractor G_{f3} accurately captured trends in collapse,
14 fracture, lost circulation, and pore pressures, identifying abnormal patterns that could lead to formation
15 instability. Through deep feature learning, high-level features were extracted from complex input data to predict
16 formation behavior.

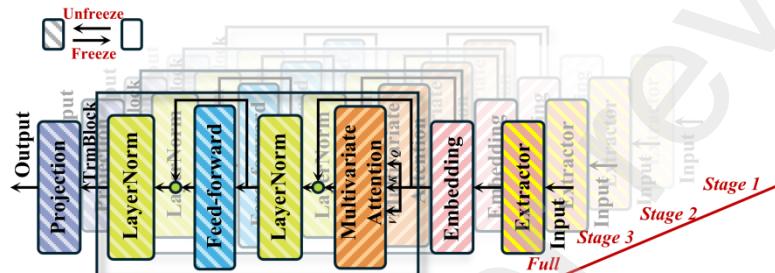
17 The domain discriminator G_d was designed as a binary classifier to enhance the generalizability of the
18 model when handling the target domain data, particularly when the target domain label data are scarce. Its
19 primary responsibility was to differentiate features from the source domain (drilled neighboring well data) and
20 the target domain (undrilled formation data). The domain discriminator G_d and feature extractor G_f engaged
21 in a min-max game. The objective of G_f was to extract features that made it difficult for G_d to distinguish their
22 domains of origin, whereas G_d aimed to classify these features as accurately as possible. If G_f generated
23 features that G_d could not effectively distinguish, these features were considered domain invariant, indicating
24 cross-domain generalizability. Adversarial training iteratively optimized this process by continuously adjusting
25 the parameters of G_f and G_d . In this game, G_f strived to make features from the source and target domains
26 appear similar to G_d , whereas G_d sought to enhance its discrimination ability. This training strategy led to G_f
27 learning feature representations, which gradually became difficult for G_d to accurately classify, promoting
28 feature confusion and improving the adaptability of the model to new domains.

29 The predictor/perceiver G_p employed three models trained and optimized on drilled neighboring well data:
30 the PSO-FCNN model, G_{p1} , for drilling speed perception, iTransformer model, G_{p2} , for LWD parameter

1 perception, and iTransformer model, G_{p3} , for formation pressure perception.

2 **4.1.2 Freezing and Fine-Tuning Model Parameters**

3 The strategy of parameter freezing and fine-tuning is crucial for model adaptability, particularly in cross-
4 well application scenarios. The three models involvedPSO-FCNN for drilling speed prediction, iTransformer for
5 LWD parameter perception, and iTransformer for formation pressure perception—underwent parameter freezing
6 and fine-tuning to optimize their performance for undrilled formation data in the target well. The architecture of
7 the parameter freezing and fine-tuning for the iTransformer model is illustrated in Fig. 12.



8 **Fig. 12. Architecture of parameter freezing and fine-tuning for the iTransformer model**

9 In the adaptation of the predictive and perceptive transfer models, parameter freezing primarily targets deep
10 network parameters that exhibit stable performance during the source-domain training phase, particularly the
11 primary and intermediate feature extraction layers responsible for capturing basic patterns and features. These
12 layers capture features that demonstrate generalizability across different geological conditions, and are crucial
13 for the cross-scenario robustness of the model. Freezing these parameters ensures that the model relies on
14 fundamental features to maintain its predictive performance when processing new data [52]. Conversely, fine-
15 tuning focuses on the upper layers of the model, particularly the advanced, fully connected layers, whose weights
16 and biases are directly associated with the output of the model. Fine-tuning aims to optimize the predictive
17 accuracy and relevance according to the characteristics of the target domain, adapting to the unique properties
18 of the target geological data, such as drilling speed variation patterns, POR distribution characteristics, and
19 dynamic changes in formation pressure [53]. This strategy is implemented using a progressive approach. Initially,
20 most key layer parameters remain frozen, and the fine-tuned layers are adjusted with a low learning rate. As the
21 performance of the model on the target domain data improved, more layers gradually became unfrozen, allowing
22 for extensive parameter updates. This approach prevents the model from overfitting the source domain data while
23 encouraging it to gradually adapt to the new geological environment, leveraging the retained knowledge. The
24 specific progressive unfreezing strategies are listed in Table 2.
25

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

Task Category	Stage	Unfrozen Components	Description
Drilling Speed	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.
	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.
Downhole Parameter / Formation Pressure	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.

1 4.2 Model Training and Evaluation

2 In the process of transfer learning for predicting the drilling speed, perceiving the LWD parameters, and
 3 sensing the formation pressure, optimizing the feature extraction and domain discrimination strategies for the
 4 source and target domains is crucial. The source domain data were used directly for comprehensive training
 5 without splitting, whereas the target domain data were divided into training and validation sets in a 5:5 ratio.
 6 This division strategy aims to assess the adaptability of the model to new geological environments. The loss
 7 functions involve an adversarial game between feature extractor G_f and domain discriminator G_d . During training,
 8 feature extractor G_f was adjusted to minimize the prediction error while maximizing the domain discrimination
 9 error. This adversarial game was realized through joint optimization of the loss functions \mathcal{L}_p and \mathcal{L}_d , where \mathcal{L}_p
 10 represents the prediction loss and \mathcal{L}_d represents the domain discrimination loss. Specifically, \mathcal{L}_p was designed as
 11 the mean squared error (MSE), and \mathcal{L}_d was designed as the cross-entropy loss to differentiate the source and
 12 target domain data.

13 In the implementation of the transfer model training, the strategies remained consistent with those used in
 14 the offline perception of drilling speed and formation properties. A high initial learning rate was employed, and
 15 the AdamW optimizer was used to rapidly reduce the loss. As the training progressed, the learning rate was
 16 gradually adjusted using an annealing strategy to finely tune the model parameters. In addition, regularization
 17 and dropout methods were employed to enhance the generalization of the transfer model. An early stopping
 18 method was used to monitor the performance of the validation set and halt training appropriately to prevent
 19 overfitting, ensuring that the transfer model maintained a high accuracy and stability in the new data environment.
 20 These measures collectively constitute the core strategies for transfer model training aimed at optimizing the
 21 analysis and prediction accuracy of drilling and geological data. The training strategies for the predictive and
 22 perceptive transfer models are listed in Table 3.

23 **Table 3. Training strategies for prediction and perception transfer models**

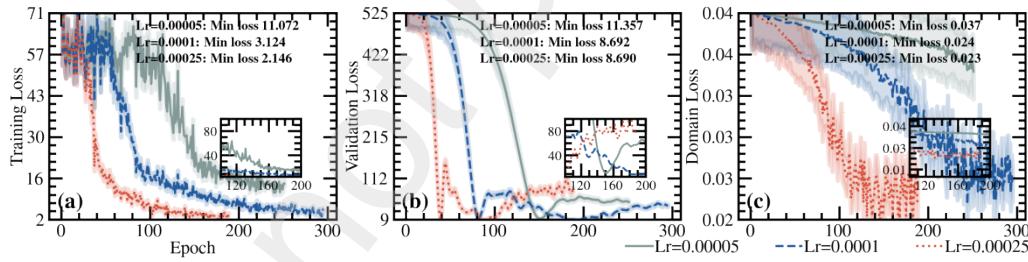
Parameter Class	Hyperparameter	Drilling Speed	Downhole Parameter	Formation Pressure
General Configurations	LearningRate	-	-	-

	<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]
Domain Discriminator	<i>ActivationFunction</i>	ReLU	ReLU	ReLU
	<i>Normalization</i>	BatchNorm	BatchNorm	BatchNorm
	<i>Layers</i>	2	3	4
Predictor	<i>ActivationFunction</i>	ReLU	ReLU	ReLU
	<i>OutputSize</i>	2	2	2
Predictor	<i>TaskRelated</i>	Same as offline task	Same as offline task	Same as offline task

1

2 **4.2.1 Transfer Prediction of Drilling Speed for Untapped Target Wells**

3 To investigate the effect of different learning rates on the transfer prediction model for the drilling speed
4 from drilled neighboring wells to untapped target formations, three learning rate levels were set: 0.00005, 0.0001,
5 and 0.00025. These were applied to the feature extractor, domain discriminator, and predictor to evaluate their
6 specific effects on the performance of the transfer model. Fig. 13 illustrates the changes in the prediction loss
7 values and domain discriminator loss values during the training and validation stages under three different
8 learning rate settings for the drilling speed prediction transfer model.

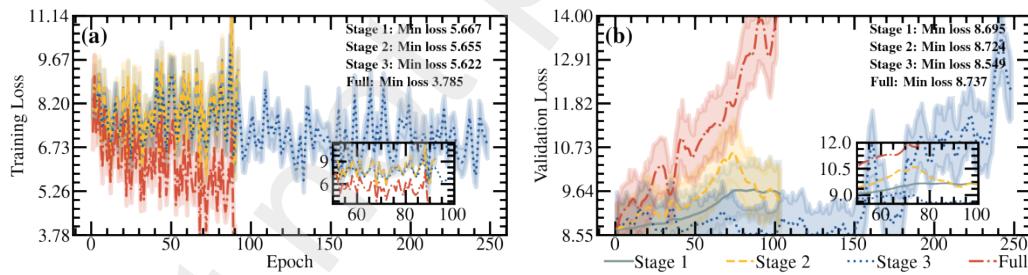


9
10 **Fig. 13. Dynamics of losses for the drilling speed prediction transfer model and domain discriminator**
11 **under different learning rate settings**

12 Fig. 13(a) illustrates that, with a higher learning rate (0.00025), the loss value decreases the fastest. This
13 phenomenon indicates that a high learning rate can facilitate the quick adaptation of the transfer model to new
14 data environments in the initial stages of training, thereby accelerating convergence. This is crucial for the rapid
15 iteration and optimization of the transfer model, especially in scenarios where drilling data are frequently updated.
16 Fig. 13(b) further emphasizes the impact of the learning rate on the transfer model during the validation phase.
17 As the learning rate increases, the rapid decrease in loss values indicates that a higher learning rate helps the
18 transfer model quickly adjust and adapt when encountering new validation data, which is beneficial for achieving
19 efficient transfer predictions. However, lower learning rates (0.00005 and 0.0001), although slower in reducing
20 the loss, provide a more stable decrease in the loss values in later iterations. This is advantageous for improving

1 the generalizability of the transfer model when facing unseen data and for reducing the risk of overfitting. Fig.
 2 13(c) shows the performance differences of the domain discriminator under different learning rates. The loss
 3 value decreases the fastest with a learning rate of 0.00025, indicating that under this setting, the domain
 4 discriminator could more effectively learn and distinguish the feature differences between the source and target
 5 domains. This is crucial for enhancing the prediction accuracy of the transfer model under new geological and
 6 operational conditions. By contrast, the setting with a learning rate of 0.00005 exhibits the slowest decrease in
 7 loss, which may lead to a transfer model requiring more iterations to achieve similar performance levels.

8 After determining that 0.00025 is the optimal learning rate for the transfer model to predict the drilling
 9 speed of the undrilled strata of the target well, the transfer model underwent comprehensive training.
 10 Subsequently, to enhance its adaptability to the undrilled strata of the target well, all transfer model parameters
 11 were frozen. The specific progressive unfreezing strategy is detailed in Table 3, covering the four key unfreezing
 12 stages, from initial training to complete adaptation. In the initial stage, all layers remained frozen, and the transfer
 13 model gradually unfroze more layers in subsequent stages, allowing it to adapt to the specific characteristics of
 14 the target domain and adjust the deep features. In the full adaptation stage, the entire network was unfrozen and
 15 comprehensive fine-tuning of the transfer model was performed to maximize the prediction accuracy and
 16 generalizability of the undrilled strata of the target well. This strategy ensured that the transfer model was
 17 effectively adapted to the new data environment at each stage, thereby optimizing its performance. Fig. 14 shows
 18 the change in the prediction loss of the predictor during training and validation across the four unfreezing stages.



19
 20 **Fig. 14. Dynamics of training and validation losses for the predictor in the drilling speed prediction**
 21 **transfer model across four unfreezing stages**

22 Fig. 14(a) illustrates the change in the training loss of the drilling-speed prediction transfer model across
 23 the four unfreezing stages with respect to the number of iterations. In the initial stage, when all layers remained
 24 frozen, the training loss was relatively high, indicating that the transfer model had not yet been optimized for the
 25 specific geological characteristics and drilling parameters of the undrilled strata of the target well. As the model
 26 entered the first stage and the output layer was unfrozen, it began to adapt to the geological features of the target
 27 domain, leading to a gradual decrease in the training loss. In the second stage, with the unfreezing of the
 28 penultimate layer, the training loss continued to decrease, demonstrating the effectiveness of adjusting the
 29 intermediate layer feature representations to model the relationship between the drilling speed and geological

parameters more accurately. In the third stage, where the entire prediction network was unfrozen, the training loss decreased significantly, indicating that the transfer model had substantially adapted to the new drilling data environment, optimizing its response to bottom-hole conditions and drilling operations.

Fig. 14(b) shows the changes in the validation loss across the stages. The lowest validation loss was observed in the third stage, indicating that the transfer model had the best generalizability at this point and could effectively adapt to the complex geological conditions of the undrilled areas. However, in the full adaptation stage, although the training loss further decreased to the lowest point, the validation loss increased significantly, indicating overfitting. At this stage, the transfer model was overly optimized for the training data, resulting in reduced generalizability and difficulty in adapting to new drilling environments.

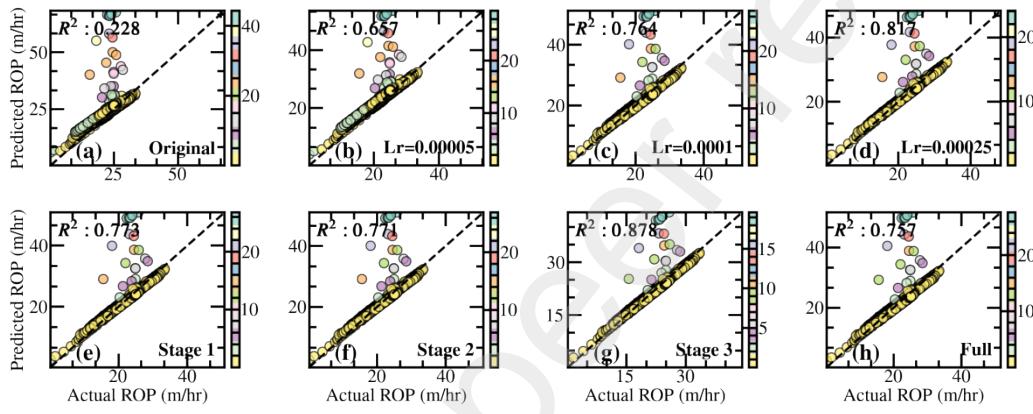
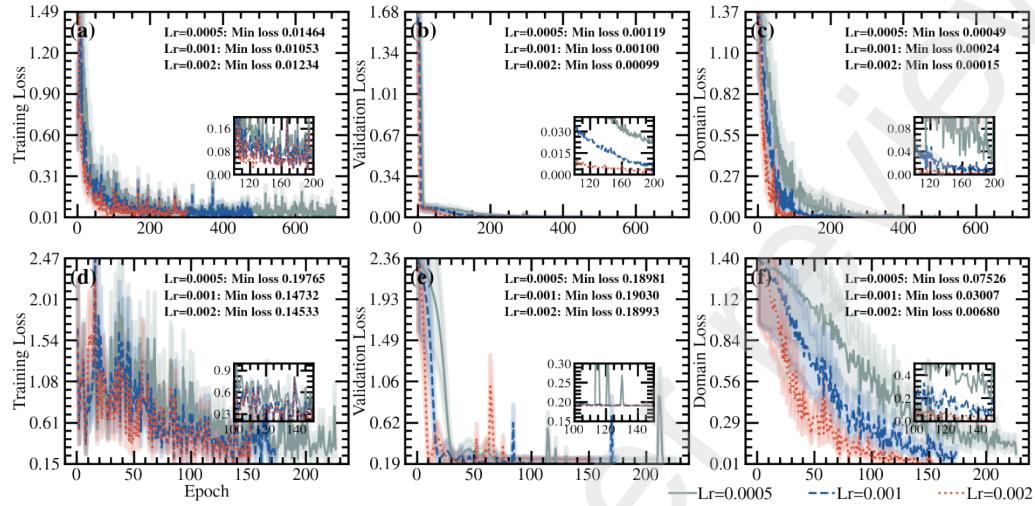


Fig. 15. R^2 fit of the drilling speed prediction transfer model under different configurations (learning rates and unfreezing stages)

Fig. 15(a–h) show the R^2 performance of the drilling speed prediction transfer model under different configurations, reflecting the fit of the model to the undrilled strata data of the target well. Fig. 15(a–d) shows that in the original state of the transfer model, that is, without transfer learning, the R^2 value was 0.2279. This indicates a low baseline fit level before specific training. Subsequently, by adjusting the learning rates to 0.00005, 0.0001, and 0.00025, the fitting accuracy of the transfer model improved significantly, particularly when the learning rate was set to 0.00025, where R^2 reached its highest value of 0.8169. This emphasizes the importance of increasing the learning rate to enhance the model performance. Fig. 15(e–h) illustrates the performance after applying the strategy of freezing all the parameters, followed by four progressive unfreezing stages. In the first stage, only the output layer was unfrozen, resulting in a slight decrease in the R^2 value to 0.7732, reflecting the initial impact of unfreezing on model adaptability. In the second stage, the penultimate layer was unfrozen, and the R^2 value remained stable at 0.7711, indicating the stable performance of the model. Upon reaching the third stage, with the entire prediction network unfrozen, the R^2 value significantly increased to 0.8783, demonstrating the excellent adaptation and optimization performance of the model for the new data environment. However, in the fully adapted stage, although the training loss further decreased to its lowest point, the R^2 value dropped to

1 0.7575, indicating a reduced generalizability due to overfitting to the training data. This stage highlights the
 2 importance of preventing overfitting in the drilling speed prediction to ensure the reliability and accuracy of the
 3 transfer model in practical applications.

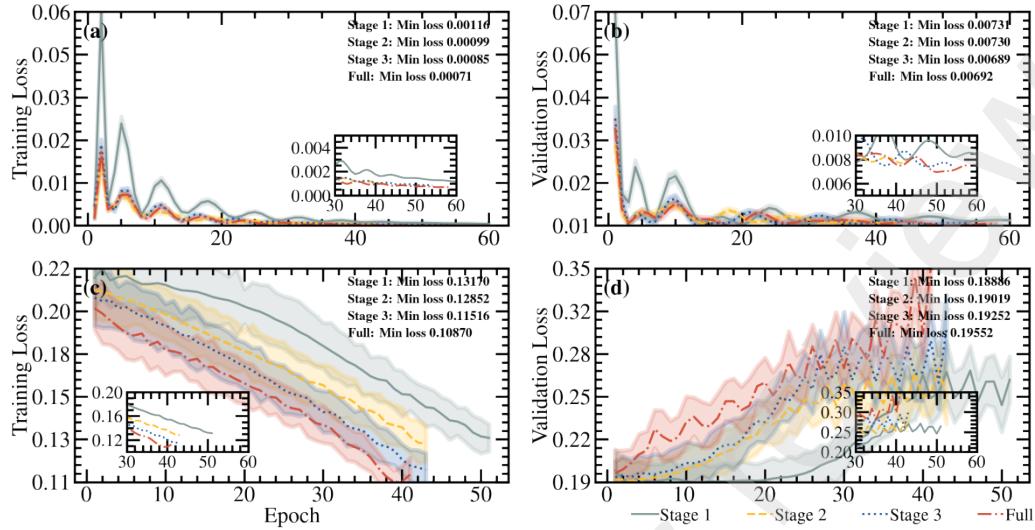
4 4.2.2 Transfer Perception of Formation Properties for Untapped Target Wells



5 **Fig. 16. Dynamics of losses for the formation properties perception transfer model and domain
 6 discriminator under different learning rate settings**

8 Fig. 16(a–c) demonstrates the impact of different learning rates on the training and validation losses in the
 9 formation property sensing transfer model. The training and validation losses of the transfer model exhibited
 10 specific trends at various learning rates. Specifically, with a learning rate of 0.0005, the transfer model exhibited
 11 a baseline training loss of 0.0146 and a validation loss of 0.0011, indicating good adaptation to the training set
 12 and effective generalization to new data. As the learning rate increased to 0.001 and 0.002, the training loss
 13 decreased to 0.0105 and increased slightly to 0.0123, respectively, whereas the minimum validation loss
 14 decreased to 0.0009. This suggests that increasing the learning rate helps the transfer model adapt more quickly
 15 to unseen data, especially at the high learning rate of 0.002, where the model exhibited the best sensing capability
 16 for the validation data. Fig. 16(d–f) focuses on the formation pressure sensing transfer model. With a learning
 17 rate of 0.0005, the training and validation losses of the transfer model were 0.1976 and 0.1898, respectively,
 18 indicating a stable processing ability but limited adaptation to new environments. Increasing the learning rate to
 19 0.001 reduced the training loss to 0.1473, whereas the validation loss increased slightly to 0.1903, emphasizing
 20 the need to balance the generalization when increasing the learning rate. Further increasing the learning rate to
 21 0.002 maintained the training loss at 0.1453 and the validation loss decreased slightly to 0.1899, indicating
 22 improved adaptability under extreme learning rates. Additionally, the loss value of the domain discriminator
 23 decreased significantly from 0.0752 to 0.0068, indicating that higher learning rates enhanced the sensitivity and
 24 accuracy of the transfer model in distinguishing changes in drilling environments, which is crucial for accurately

1 predicting the formation pressure under dynamic geological conditions.



2
3 **Fig. 17. Dynamics of training and validation losses for the perceiver in the formation properties perception**
4 **transfer model across four unfreezing stages**

5 Fig. 17(a–b) shows the phased training and validation performance of the formation property sensing
6 transfer model concerning key geological parameters, such as GR, DT, POR, and RD. In the initial stage, all
7 layers of the transfer model were frozen to stabilize the outputs and prevent early overfitting. During the first
8 stage, the output layer was unfrozen, resulting in an initial training loss of 0.0011 and a validation loss of 0.0073,
9 which significantly enhanced the responsiveness of the model to new drilling data. In the second stage, further
10 unfreezing of the output and attention layers reduced the training loss to 0.0009 and slightly decreased the
11 validation loss to 0.0073, thereby improving the ability of the model to capture changes in POR and RD. In the
12 third stage, unfreezing is extended to the intermediate layers, deepening the adaptation of the model to complex
13 geological features. The training loss decreased to 0.0008, and the validation loss decreased to 0.0068,
14 demonstrating high adaptability in complex formations. In the full adaptation stage, all network layers were
15 unfrozen, resulting in the lowest training loss of 0.0007; however, the validation loss slightly increased to 0.0069,
16 indicating a potential risk of overfitting that must be monitored to maintain generalizability.

17 Fig. 17(c–d) focuses on the formation pressure-sensing transfer model, showing the phased training and
18 validation performance across the collapse, fracture, loss, and pore pressures. In the initial stage, all the layers
19 were frozen to ensure baseline sensing accuracy. After unfreezing the output layer in the first stage, the initial
20 training loss was 0.1317, and the validation loss was 0.1888, marking the preliminary adaptation of the model
21 to the new formation pressure data. In the second stage, unfreezing the output and attention layers further refined
22 the model, reducing the training loss to 0.1285. However, the validation loss increased slightly to 0.1901,
23 reflecting the complex dynamics of extracting the key temporal features. In the third stage, unfreezing was

extended to the intermediate layers, reducing the training loss to 0.1151; however, the validation loss increased to 0.1925, highlighting the challenges in deeper adaptation processes. In the final full adaptation stage, with all network layers unfrozen, the training loss further decreased to 0.1086, but the validation loss increased to 0.1955, indicating that variations in the actual geological conditions necessitate careful handling of overfitting risks.

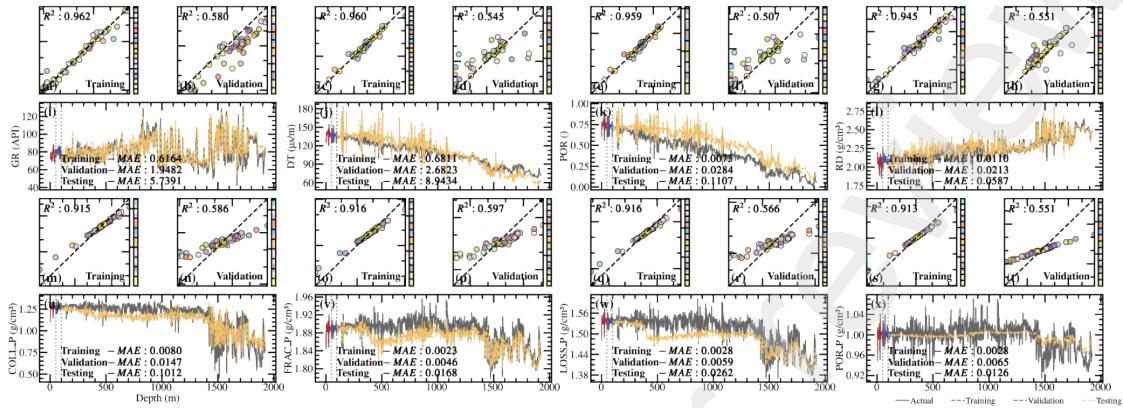


Fig. 18. Dynamics of perception performance for the LWD parameters perception transfer model in stage 3 and the formation pressure perception transfer model in stage 1

Fig. 18(a–l) illustrates the performance of the LWD parameter-sensing transfer model in the third stage for sensing the GR, DT, POR, and RD of the undrilled formation of the target well. The model demonstrated an excellent fit ($R^2=0.9620$) to the training set for GR and DT, indicating high efficiency in learning and fitting the known training data. However, the performance declined significantly for both the validation and test sets. The validation set R^2 for GR decreased to 0.5797, with the mean absolute error (MAE) increasing to 1.9482, and the test set MAE further increased to 5.7391. For DT, the validation set R^2 was 0.5454, with an MAE of 2.6823. The test set MAE surged to 8.9434. This performance revealed the generalization challenges of the model when faced with complex or variable geological data. Despite the high R^2 values (0.9589 and 0.9448, respectively) and exceptionally low MAE (both below 0.0200) during training, this performance was primarily driven by the small scale of the parameters rather than by the absolute predictive ability of the model. This was confirmed in the validation and test sets, where the MAE remained low but did not fully reflect the predictive accuracy of the model in real geological environments.

Fig. 18(m–x) shows the performance of the formation pressure-sensing transfer model in sensing the collapse, fracture, loss, and pore pressures of the undrilled formation of the target well. The collapse pressure had an R^2 of 0.9152 and an MAE of 0.0080 on the training set, demonstrating excellent data-fitting ability. However, the R^2 dropped to 0.5864, and MAE increased to 0.0147 on the validation set, with further increases in the test set to 0.1012, indicating the limited adaptability of the transfer model to new data. The fracture and loss pressures exhibited high R^2 and low MAE during training, but their performance fluctuated in subsequent test stages, emphasizing the need for possible adjustments to prevent overfitting. Although the transfer model

1 showed stable performance in sensing pore pressure during training, its generalization performance in later
2 stages was inadequate.

3 **4.3 Online Perception of Drilling Speed and Formation Properties**

4 Owing to the significant decline in the performance of the LWD parameter and formation pressure sensing
5 transfer models during the validation and testing phases, it is evident that these models exhibit notable limitations
6 in terms of adaptability and generalization when confronted with unknown or complex geological conditions.
7 This issue not only highlights the inadequacies of traditional static training methods in dynamic geological
8 environments but also underscores the need to enhance the responsiveness of the transfer model to real-time
9 changes. Therefore, designing an online learning perception strategy has become a crucial approach for
10 addressing the aforementioned issues.

11 **4.3.1 Incremental Learning Strategy**

12 The core of the online learning strategy is to achieve continuous adaptive optimization of the transfer model
13 for the drilling speed and formation properties. This strategy relies on incremental learning methods that allow
14 the transfer model to retrain the entire network for immediate updates when new drilling data are received.
15 Specifically, the incremental learning strategy enables the transfer model to gradually integrate new sample
16 information by setting a rolling window, thereby effectively reducing the training resource consumption caused
17 by data expansion [54].

18 During this process, the efficient automatic processing and quality control of real-time data streams are
19 crucial. All drilling data undergo strict preprocessing, including the standardization of features with drilled
20 neighboring well data, linear interpolation of missing values, and wavelet transform denoising, ensuring the
21 consistency and high quality of the input data. The completed iTransformer model, combined with feature
22 extractors, parses drilling parameters and formation properties in real time, and subsequently outputs key sensing
23 results, such as logging parameters and formation pressures. The adaptability of the transfer model is evaluated
24 using periodic performance reviews. If the transfer model performance declines or new formation conditions are
25 encountered, learning strategies and network parameters are immediately adjusted to ensure precision and
26 robustness in complex environments. Additionally, an annealing strategy was gradually adopted to adjust the
27 learning rate and optimize the step management within the learning cycle, effectively avoiding local optima. The
28 transfer model regularly saves and evaluates optimal parameters, ensuring flexible adaptation to different stages
29 of drilling operations and optimizing the actual drilling conditions. The specific parameters of the incremental
30 learning strategy are listed in Table 4.

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

Hyperparameter	Description	Downhole Parameter	Formation Pressure
<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

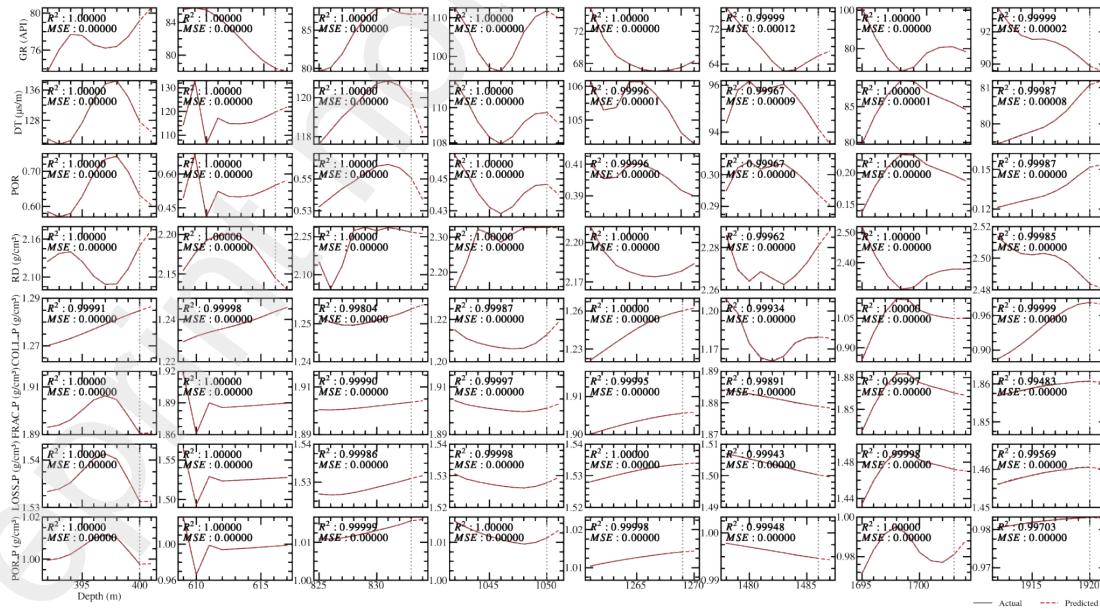
1

2 4.3.2 Trigger Conditions and Forgetting Mechanism

3 In drilling operations, the update frequency and trigger conditions are crucial for the adaptability of the
4 transfer model. The update mechanism was set based on the real-time acquisition of new data, specifically
5 triggering a transfer model update every 10 m of drilling depth. At this point, the transfer model was retrained
6 using the most recent 200 m of drilling data. This strategy ensures that the transfer model can continuously
7 integrate the latest formation data, adapting in real time to changing drilling conditions and geological
8 characteristics, thereby optimizing the prediction accuracy and operational efficiency.

9 Additionally, the forgetting mechanism dynamically adjusts the data weights to reduce the influence of old
10 data on the transfer model, thereby emphasizing the importance of new data. This approach is particularly
11 suitable for drilling environments in which geological conditions change rapidly, thereby significantly enhancing
12 the adaptability and flexibility of the transfer model. Effective data management strategies reduce resource
13 consumption during continuous learning while maintaining high efficiency in transfer model performance,
14 thereby enhancing the safety and economic viability of drilling operations. This time-window-based update
15 mechanism not only improves the timeliness of predictions, but also ensures the responsiveness and adaptability
16 of the transfer model when encountering different formation conditions.

17 4.3.3 Online Learning Process



18

19 **Fig. 19. Perception performance of the formation property perception transfer model in an online learning**
20 **environment**

Fig. 19 shows the perception performance of the formation–property transfer model in an online learning environment. The upper part of Fig. 19 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. The forgetting mechanism in online learning prioritizes the latest data directly related to the current drilling environment, gradually eliminating old and irrelevant data. This effectively optimizes the response speed of the model to newly emerging formation characteristics and reduces the computational burden. 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. 19 shows the ability of the transfer model to perceive formation pressures, including the collapse, fracture, leak-off, and pore pressures, in an online learning environment. Predictions for the collapse pressure and fracture pressure showed nearly perfect R^2 values at multiple depth intervals, such as between 608 and 617 m, where both pressures achieved an R^2 of 1.0000. Similarly, the predictions for the leak-off and pore pressures were outstanding, especially at depth intervals of 1261–1269 meters and 1694–1704 meters, where the pore pressure 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.

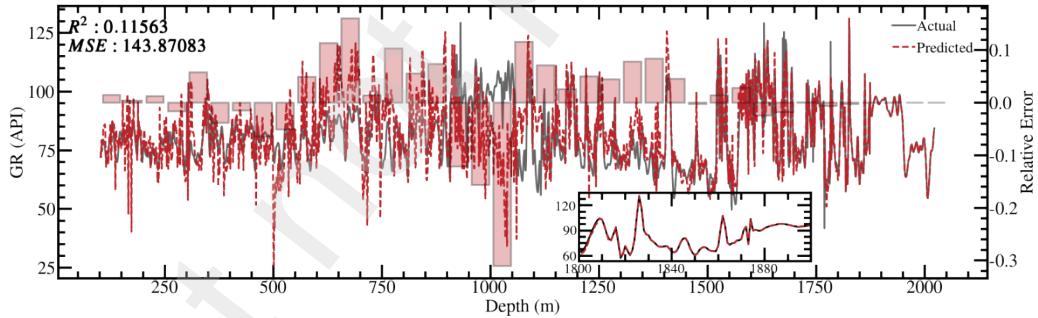
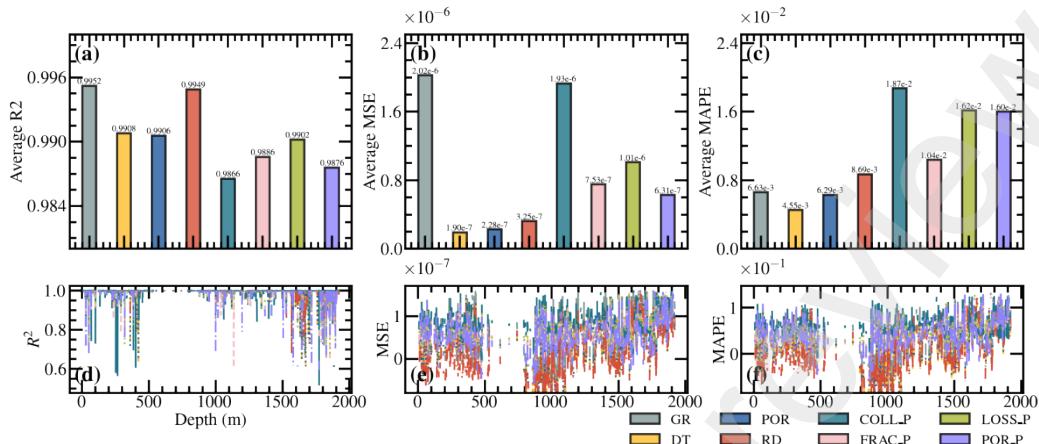


Fig. 20. Impact analysis of the forgetting mechanism on the perception of GR in untapped target wells in an online learning environment

Fig. 20 illustrates the application of the forgetting mechanism in an online learning environment, showing the results of using the latest updated transfer model to sense GR across the full depth of the undrilled formation of the target well. Although there were significant differences between the predictions and actual values in the early depth intervals, indicating that the model effectively forgot old data, the sensing performance significantly improved in the 1800–2000 meter depth interval, closely matching the actual geological data. This phenomenon highlights the importance of the forgetting mechanism in drilling applications and demonstrates that this

1 mechanism ensures that the model maintains real-time adaptability and accuracy. This also reflects the
2 effectiveness of online learning strategies in addressing changing geological environments.



3
4 **Fig. 21. Performance evaluation of the formation property perception transfer model in a dynamic**
5 **geological environment**

6 Fig. 21(a-f) shows the exceptional online learning capabilities of the LWD parameter-and formation
7 pressure-sensing transfer models in dynamic geological environments. The performances of these models were
8 verified through continuous evaluations following each training update, with average R^2 values exceeding 0.99,
9 demonstrating their high accuracy in capturing and responding to formation changes. All parameters exhibited
10 extremely low errors in the MSE metric, indicating high prediction precision. Although some formation pressure
11 parameters had higher Mean Absolute Percentage Errors (MAPE), such as an MAPE of 0.0187 for the collapse
12 pressure, this still emphasizes the strong applicability and reliability of the models in real drilling environments.

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

19 **5. Real-Time Multi-Objective Deep Reinforcement Learning Optimization for** 20 **Drilling Parameters in Untapped Target Wells**

21 To efficiently optimize the drilling parameters for undrilled target formations, a FCNN algorithm adjusted
22 via PSO was designed to construct a drilling speed prediction model. Simultaneously, an online-trained
23 iTransformer model based on adversarial domain adaptation, was used to sense the LWD parameters and
24 formation pressures. The drilling speed estimator provided accurate predictions through the PSO-FCNN model,

1 whereas the LWD sensor provided operators with real-time critical geological information, such as lithological
 2 characteristics and fluid dynamics, thereby enhancing the precision of the drilling process control and decision-
 3 making. Additionally, real-time sensing of the formation pressure acted as a constraint on the mud density,
 4 ensuring wellbore stability and preventing formation fluid invasion. By providing real-time feedback on key
 5 pressure parameters, the mud density settings were optimized to avoid drilling issues caused by inappropriate
 6 mud density. The integration of these components formed a MDP for drilling parameter adjustments, allowing
 7 for the dynamic optimization of drilling parameters based on continuous data streams. This enabled real-time
 8 adjustments to the drilling operation to adapt to changing geological conditions and optimize drilling efficiency.

9 ***5.1 Design of Multi-Objective Optimization Functions***

10 **5.1.1 Unit Cost per Foot Function**

11 In the drilling process, the unit cost per foot function is a core indicator for evaluating the drilling economic
 12 efficiency, with the key being the accurate measurement of costs and efficiency during drilling [55]. The
 13 expression for the unit cost per foot is given by Eq. 16.

$$14 \quad C_{pm} = \frac{C_b + C_r(t_t + t)}{H} \quad (17)$$

15 where C_{pm} is the cost per meter in currency units per meter; C_b is the bit cost in currency units; C_r is the
 16 rig operation cost in currency units per hour; t is the drilling time in hours; t_t is the trip and connection time in
 17 hours; and H is the total drilled depth in meters.

18 The bit wear rate is given by Eq. 17, reflecting the dynamics of the bit tooth wear.

$$19 \quad \frac{dh}{dt} = \frac{A_f(a_1n + a_2n^3)}{Z_2 - Z_1W(1 + C_1h)} \quad (18)$$

20 Eq. 17 can be transformed to relate time dt to the wear amount dh , as shown in Eq. 18.

$$21 \quad dt = \frac{(Z_2 - Z_1W)(1 + C_1h)}{A_f(a_1n + a_2n^3)} dh \quad (19)$$

22 where A_f is the formation abrasiveness coefficient; n is the revolutions per minute (RPM); a_1 and a_2 are
 23 the revolutions per minute impact coefficients, determined by the bit type; Z_1 and Z_2 are the WOB impact
 24 coefficients, related to the bit diameter; and C_1 is the bit tooth wear deceleration coefficient.

1 By incorporating the drilling speed prediction transfer model $v_{pc}(w, n, q, k_d)$ into the aforementioned
 2 bit wear rate equation, a mathematical model for bit footage H and working time t was established, as shown
 3 in Eq. 19.

$$4 \quad \begin{cases} H = \int_0^{h_f} v_{pc}(w, n, q, k_d) \cdot \frac{(Z_2 - Z_1 W)(1 + C_1 h)}{A_f(a_1 n + a_2 n^3)} dh \\ t = \int_0^{h_f} \frac{(Z_2 - Z_1 W)}{A_f(a_1 n + a_2 n^3)} \left(h + \frac{C_1}{2} h^2 \right) dh \end{cases} \quad (20)$$

5 where h_f is the amount of bit tooth wear corresponding to the bit life. Accordingly, the unit cost per foot
 6 function was transformed into a function of the WOB and RPM, resulting in the expression shown in Eq. 20.

$$7 \quad C_{pm} = \frac{C_r \left[t_E \cdot A_f (a_1 n + a_2 n^3) + \left(h_f + \frac{C_1}{2} h_f^2 \right) \right]}{v_{pc}(w, n, q, k_d) \cdot \left(h_f + \frac{C_1}{2} h_f^2 \right)} \quad (21)$$

8 where $t_E = \frac{C_b}{C_r} + t_t$ is the equivalent unit cost-time term. The unit cost per foot function demonstrates the
 9 combined impact of the bit and rig operation costs on the overall drilling cost. This approach provides a precise
 10 method for calculating drilling costs and supporting scientific and accurate decision-making in drilling
 11 operations.

12 5.1.2 Mechanical Specific Energy Function

13 The mechanical specific energy (MSE) model correlates the energy required to break a unit volume of rock
 14 with the efficiency of the bit in breaking the rock, and serves as a critical method for quantifying drilling
 15 efficiency. The MSE is calculated using parameters such as the drilling speed (ROP), weight on bit (WOB),
 16 revolutions per minute (RPM), torque, and bit diameter. Higher MSE values indicate lower drilling efficiency
 17 and poorer adaptability of the bit to the formation, suggesting that the drilling parameters need to be optimized
 18 [56]. The ideal-state MSE is calculated using Eq. 21.

$$19 \quad MSE = \frac{4W}{\pi D_b^2} + \frac{480nT_b}{D_b^2 P} \quad (22)$$

20 where MSE is the mechanical specific energy in MPa; W is the weight on bit in kN; D_b is the bit diameter
 21 in m; n is the revolutions per minute in r/min; T_b is the bit torque in kN·m; and P is the penetration rate in
 22 m/hr.

1 In actual drilling operations, the actual torque at the bit is often unavailable and must be calculated using
 2 the bit sliding friction coefficient and weight-on-bit (WOB). Based on the double-integral theorem, the torque
 3 expression during the drilling process is given by Eq. 22.

$$4 T_b = \frac{1}{1000} \int_0^{\frac{D_b}{2}} \int_0^{2\pi} r^2 \frac{4\mu W}{\pi D_b^2} dr d\theta = \frac{\mu WD_b}{3000} \quad (23)$$

5 where r represents the infinitesimal length of the bit radius in mm, and μ is the bit sliding friction
 6 coefficient.

7 Owing to factors such as friction and vibration, the energy utilization efficiency in actual drilling is typically
 8 between 30% and 40%, and the required MSE is approximately three times the rock strength. The effective
 9 energy utilization rate of the bit is defined as E_f . By introducing the drilling speed prediction transfer model
 10 $v_{pc}(w, n, q, k_d)$, the modified mechanical specific energy model MSE_m can be expressed as Eq. 23.

$$11 T_b = \frac{1}{1000} \int_0^{\frac{D_b}{2}} \int_0^{2\pi} r^2 \frac{4\mu W}{\pi D_b^2} dr d\theta = \frac{\mu WD_b}{3000} \quad (24)$$

12 5.1.3 Objective Function

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

17 The objective function, also known as the evaluation index, serves as the criterion for assessing the quality
 18 of the drilling design schemes. In practical drilling operations, the goal is to maximize the drilling efficiency
 19 while minimizing costs. Therefore, the lowest unit cost per foot and the minimum MSE were selected as the
 20 objective functions for the multi-objective optimization. The objective function model FFF is expressed in Eq.
 21 24.

$$22 \begin{cases} F = \min \left\{ \min(C_{pm}), \min(MSE_m) \right\} \\ \min(C_{pm}) = \frac{C_r \left[t_E \cdot A_f \left(a_1 n + a_2 n^3 \right) + \left(h_f + \frac{C_1}{2} h_f^2 \right) \right]}{v_{pc}(w, n, q, k_d) \cdot \left(h_f + \frac{C_1}{2} h_f^2 \right)} \\ \min(MSE_m) = E_f \left(\frac{4W}{\pi D_b^2} + 0.16 \frac{n \mu W}{D_b v_{pc}(w, n, q, k_d)} \right) \end{cases} \quad (25)$$

1 **5.2 Markov Decision Process Environment**

2 **5.2.1 Reward Function Design**

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

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

15 The MSE reward evaluates the energy efficiency and mechanical load of the drilling operation through
16 mechanical specific energy (i.e., the energy required to drill a unit volume of rock). A reduction in MSE indicated
17 improved drilling efficiency and reduced risk of equipment wear; thus, operations that reduced MSE received
18 positive rewards, whereas those that increased MSE received negative rewards. This reward design emphasizes
19 the importance of energy efficiency, promoting the rational use of resources and the long-term stability of
20 equipment. The reward function model is expressed in Eq. 25.

$$21 \quad \begin{cases} R_1 = -(C_{pm} - C_{pm,prev}) \\ R_2 = -(MSE_m - MSE_{m,prev}) \end{cases} \quad (26)$$

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

27 The combination of these two rewards forms a multi-objective optimization framework, allowing the deep
28 reinforcement learning algorithm to determine the optimal balance between cost and efficiency, thereby
29 optimizing the drilling parameters. Depending on the specific drilling requirements and strategic objectives,
30 these rewards can be appropriately weighted to reflect the priority of different drilling goals and ensure the

1 optimization of drilling operations across multiple dimensions.

2 **5.2.2 State Space and Action Space**

3 In the framework of deep reinforcement learning applied to optimize drilling operations, the definition of
4 the state space depends on the key control parameters of the drilling process, including WOB, RPM, FLW, and
5 MwIN. Real-time measurements of these parameters constitute the current state during the drilling process,
6 providing the necessary data foundation for monitoring and adjusting drilling operations. The action space
7 defines the range of the adjustable drilling parameters and serves as the operational domain in which the
8 intelligent agent adjusts the drilling performance. By implementing specific actions, that is fine-tuning the
9 drilling parameters, the agent directly influences the drilling state. The results of these actions, along with the
10 changes in the drilling state, are fed back to the agent through a well-defined reward function. The state and
11 action spaces are expressed using Eqs. 26 and 27, respectively.

$$S = \{w, r, f, m\}$$
$$\begin{cases} w \in [w_{min}, w_{max}] \\ r \in [r_{min}, r_{max}] \\ f \in [f_{min}, f_{max}] \\ m \in [m_{min}, m_{max}] \end{cases} \quad (27)$$

$$A = \{\Delta w, \Delta r, \Delta f, \Delta m\}$$
$$\begin{cases} \Delta w \in [-\delta w, \delta w] \\ \Delta r \in [-\delta r, \delta r] \\ \Delta f \in [-\delta f, \delta f] \\ \Delta m \in [-\delta m, \delta m] \end{cases} \quad (28)$$

12 where w , r , f , m represent the current values of the four drilling parameters, each with its own
13 permissible limits; Δw , Δr , Δf , Δm denote the corresponding changes in these parameters; δw , δr , δf ,
14 δm are the maximum allowable adjustments in the drilling operation.

15 Precise control of the mud input density is crucial for maintaining drilling safety and efficiency. Therefore,
16 they must be accurately regulated to match the mechanical and hydrodynamic properties of the formation. The
17 key to setting the mud density is to ensure that it exceeds both the formation pore and collapse pressures, thereby
18 preventing the backflow of formation fluids and maintaining wellbore stability to avoid collapse. Simultaneously,
19 the mud density must be lower than the fracture and loss pressures to prevent the formation of fractures and mud
20 loss when these pressure values are exceeded. Thus, m_{min} and m_{max} are defined by Eq. 28.

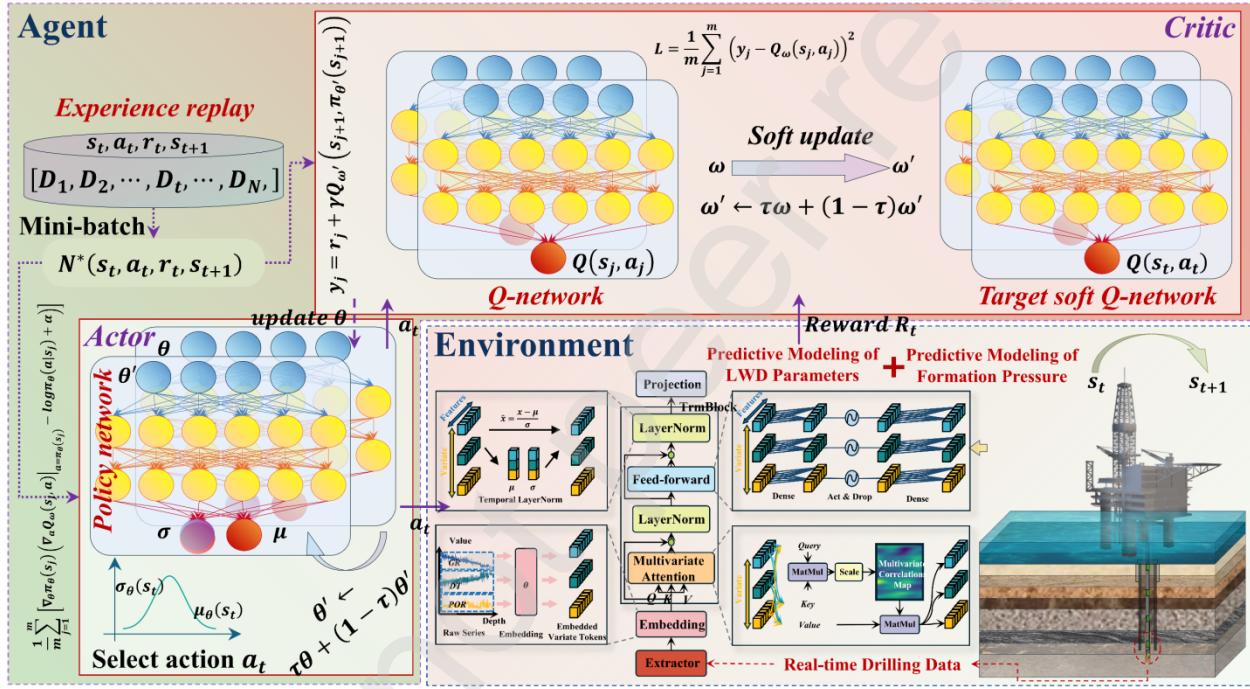
$$\begin{cases} m_{min} = \max(p_{por}, p_{coll}) \\ m_{max} = \min(p_{frac}, p_{loss}) \end{cases} \quad (29)$$

22 where m_{min} and m_{max} represent the minimum and maximum allowable mud input densities under the
23 current operational conditions, respectively, set according to the dynamic formation pressure conditions. m_{min}
24 is defined as the greater formation pore pressure p_{por} and the collapse pressure p_{coll} to ensure that the mud
25 density is sufficient to prevent the backflow of formation fluids and maintain wellbore stability. m_{max} is set as

1 the lower fracture pressure p_{frac} and loss pressure p_{loss} . The values of p_{por} , p_{coll} , p_{frac} , and p_{loss} are
2 determined using the iTransformer model.

3 5.2.3 Learning Mechanism

4 To enhance the decision quality and efficiency of drilling operations, a multi-objective deep reinforcement
5 learning algorithm, DPOAC, was specifically designed for drilling parameter optimization. This algorithm is
6 based on the soft actor (SAC) framework and is customized to meet the optimization needs of continuous action
7 spaces, adapting to the complex decision environment and dynamic operational requirements of drilling [57].
8 Fig. 22 shows the architecture of the DPOAC algorithm.



9
10 **Fig. 22. Architecture of the DPOAC algorithm for drilling parameter optimization**

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

1 **Policy Network:** The policy network significantly enhances the exploration efficiency through entropy
2 regularization, preventing premature convergence to suboptimal solutions. This strategy receives drilling
3 parameters, such as WOB, RPM, FLW, and MwIN, and outputs the probability distribution of actions.
4 Combining expected reward maximization and behavior entropy maximization increases the behavior diversity
5 and expands the exploration state space. The DPOAC algorithm adopts a double Q-network structure, reducing
6 the bias from overestimation and improving the learning efficiency. Each update uses the smaller of the two Q-
7 values for optimization, reducing the reward estimation fluctuations and bias. Regular synchronization between
8 the online Q-network and the target Q-network enhances learning goal stability and optimizes long-term reward
9 accumulation by balancing exploration and exploitation. The policy network updates strategies through
10 interactive feedback, adjusts action selection using accumulated experience, ensures precise and effective
11 decisions in complex drilling environments, and significantly improves the cost efficiency and safety of drilling
12 operations.

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

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

1 drilling parameters using the DPOAC algorithm is presented in Table B1 of Appendix B.

2 **5.3 Drilling Parameters Optimization Actor-Critic (DPOAC) Algorithm**

3 **5.3.1 Training of Q-, Target, and Policy Networks**

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

8 **Table 5. Coefficients for drilling cost and mechanical specific energy functions in the Caofidian 6-4 block,**
9 **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_l	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
Mechanical Specific Energy	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

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

14 **Table 6. Learning strategies for the DPOAC algorithm**

Hyperparameter	Description	Value
$ActorLearningRate$	Learning rate for updating the Actor Network.	0.000010
$Q1NetLearningRate$	Learning rate for updating Q1-network parameters.	-
$Q2NetLearningRate$	Learning rate for updating Q2-network parameters.	-
$DiscountFactor$	Discount factor for weighting current vs future rewards.	0.7
$TauSoftUpdate$	Tau parameter for soft updates in target networks.	0.7
$LossWeights$	Weights for balancing importance among optimization targets in the loss function.	0.5,0.5
$BatchSize$	Number of samples drawn from the experience replay buffer per batch.	2048
$BufferCapacity$	Maximum capacity of the experience replay buffer.	16000
$MaxEpochSteps$	Maximum number of steps per epoch.	1000
$AlphaPriority$	Exponent parameter for calculating priority in experience replay.	0.6
$BetaStart$	Initial importance sampling parameter in experience replay.	0.4
$BetaEnd$	Final importance sampling parameter in experience replay.	1
$BetaDecaySteps$	Number of steps over which the importance sampling parameter decays in experience replay.	500

15 In the training of the drilling parameter optimization, the Q-network was optimized by minimizing the loss
16 function, which quantified the deviation between the actual and predicted rewards based on the current drilling
17 state and action. The policy network derived the optimal action strategy from the current state, adjusting key
18

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. 23 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 characteristics of each network during the training phase.

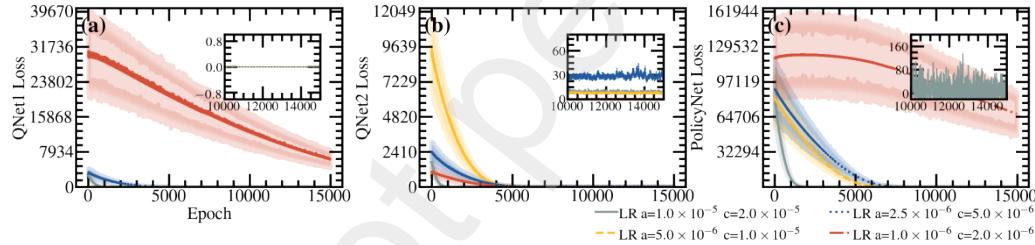


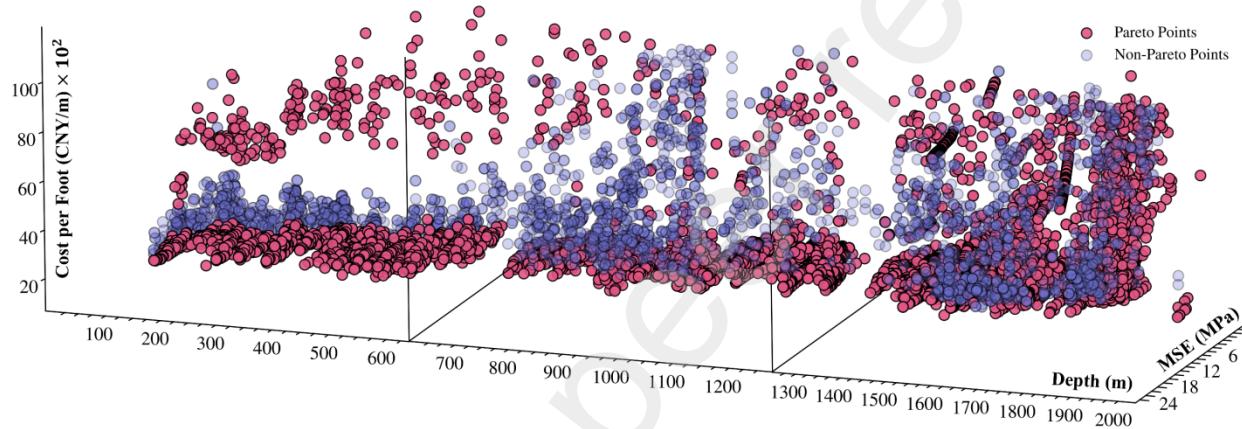
Fig. 23. Dynamics of training losses for the Q1-, Q2-, and policy networks under different learning rate settings

Fig. 23 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

1 role of learning rate settings in the performance of the DPOAC algorithm, particularly in adapting to complex
2 drilling environments. A reasonable adjustment of the learning rate significantly affects the convergence speed,
3 performance, and stability of the model.

4 **5.3.2 Drilling Parameter Optimization Process**

5 In the process of drilling parameter optimization, Fig. 24 illustrates the variation in the unit cost per foot
6 and mechanical specific energy as the drilling depth ranges from 102 to 2021 m. The results for each depth were
7 recorded after 3000 iterations. Notably, the Pareto points are identified in the figure, indicating the positions
8 where the optimal balance between cost and efficiency was achieved, distinguishing these points from the non-
9 Pareto points.



10 **Fig. 24. Dynamics of unit cost per foot and mechanical specific energy with drilling depth and Pareto**
11 **frontier**

12 Fig. 24 shows the trends of the unit cost per foot and mechanical specific energy as the drilling depth
13 increased from 102 to 2021 m. Initially, both metrics fluctuated significantly, reflecting the exploratory nature
14 of the parameter adjustments during the early drilling stages. During this phase, the Pareto boundary changed
15 frequently, indicating that the DPOAC algorithm attempts to find an optimal balance between cost and efficiency.
16 As drilling progressed to intermediate depths (600–1200 m), fluctuations in both metrics decreased, and the
17 Pareto boundary stabilized, signifying a shift from broad exploration to focused exploitation of operational
18 parameters. In later stages (beyond 1200–2021 m), despite the stabilization of unit costs and mechanical
19 efficiency, the Pareto boundary continued to show changes, likely reflecting severe formation pressure
20 fluctuations, particularly in areas with potential abnormal pressure zones. These adjustments in drilling
21 parameters, necessary to adapt to changing pressures, highlight the ongoing need to dynamically fine-tune
22 operations and balance the economic and efficiency considerations in drilling operations. This iterative process
23 demonstrates the capability of the DPOAC algorithm to dynamically adjust and optimize the drilling parameters,
24 driving towards optimal cost-effectiveness and mechanical efficiency through sustained exploration and
25 exploitation.

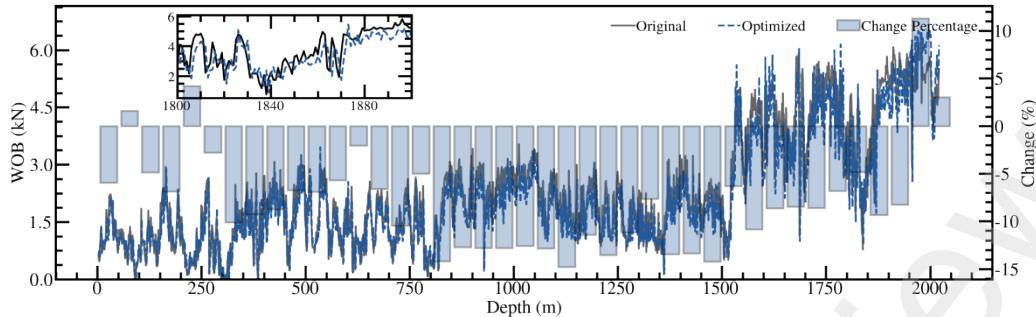


Fig. 25. Optimization changes in weight on bit (WOB) with drilling depth

Fig. 25 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.

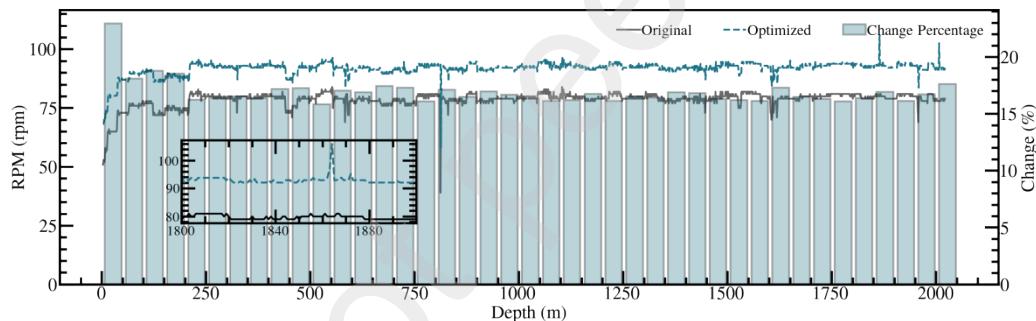


Fig. 26. Optimization changes in revolutions per minute (RPM) with drilling depth

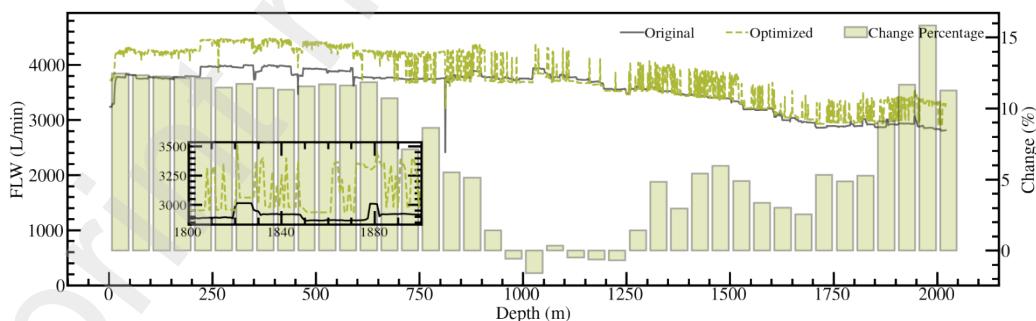
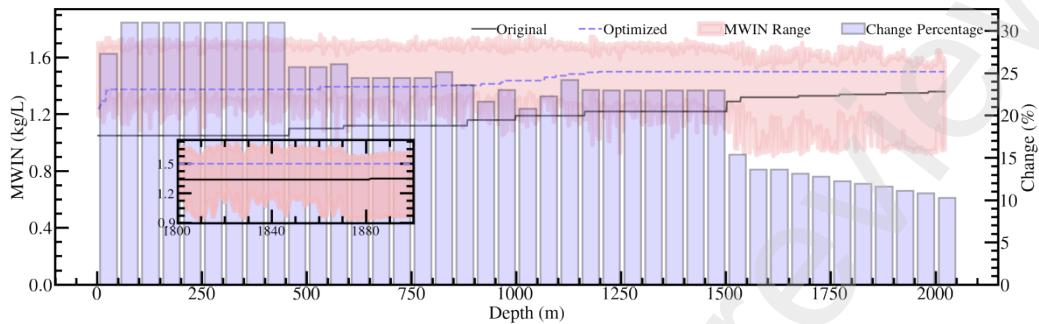


Fig. 27. Optimization changes in pump flow rate (FLW) with drilling depth

Fig. 26 and 27 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.

1 These fluctuations demonstrate the adaptation of the optimization strategy to the formation hardness and other
2 drilling conditions. Particularly in hard formations, increasing the revolutions per minute and pump flow rate
3 significantly improves the drilling speed and wellbore cleanliness; increased power and cleaning capability are
4 crucial to prevent bit wear and mitigate issues related to cuttings accumulation.



5
6 **Fig. 28. Optimization changes in mud input density (MwIN) and wellbore stability constraints with**
7 **drilling depth**

8 Fig. 28 shows the adjustments in MwIN across various drilling depths, with an average increase of
9 approximately 22%. This emphasizes the criticality of meticulous mud density management throughout the
10 drilling process. This systematic increase primarily aims to accommodate dynamic formation conditions and
11 fulfill wellbore stability requirements. Given the stringent constraints imposed by the formation pore, collapse,
12 fracture, and loss pressures, the mud density adjustments were precisely regulated within a safe operational
13 window. This is pivotal for maintaining wellbore stability and optimizing the transport efficiency of the cuttings.
14 By maintaining the mud density within this safe threshold, drilling operations can effectively prevent issues,
15 such as wellbore collapse triggered by abnormal formation pressure. This approach safeguards the wellbore and
16 ensures continuity and efficiency of drilling operations. Fine-tuning the mud density facilitates a balance in
17 wellbore pressures, enhances the interaction between the drill bit and formation, minimizes bit wear, and boosts
18 drilling speed.

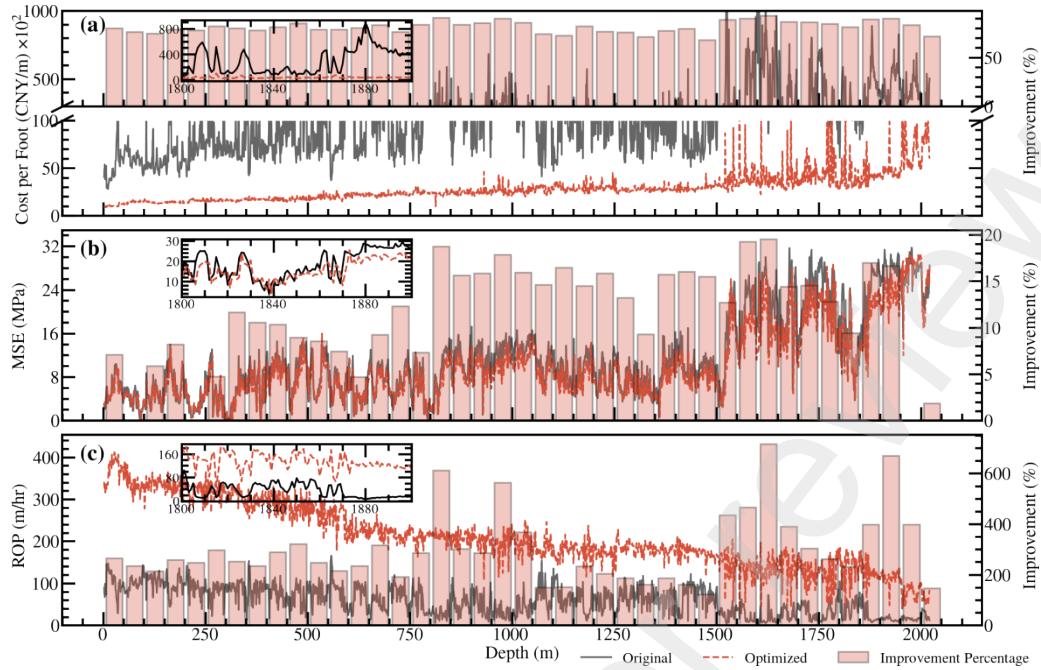


Fig. 29. Optimization changes in unit cost per foot, mechanical specific energy, and drilling speed with drilling depth

Fig. 29 demonstrates the effectiveness of the DPOAC algorithm in optimizing key performance metrics, such as the unit cost per foot, mechanical specific energy, and drilling speed, 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 drilling speed. 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.

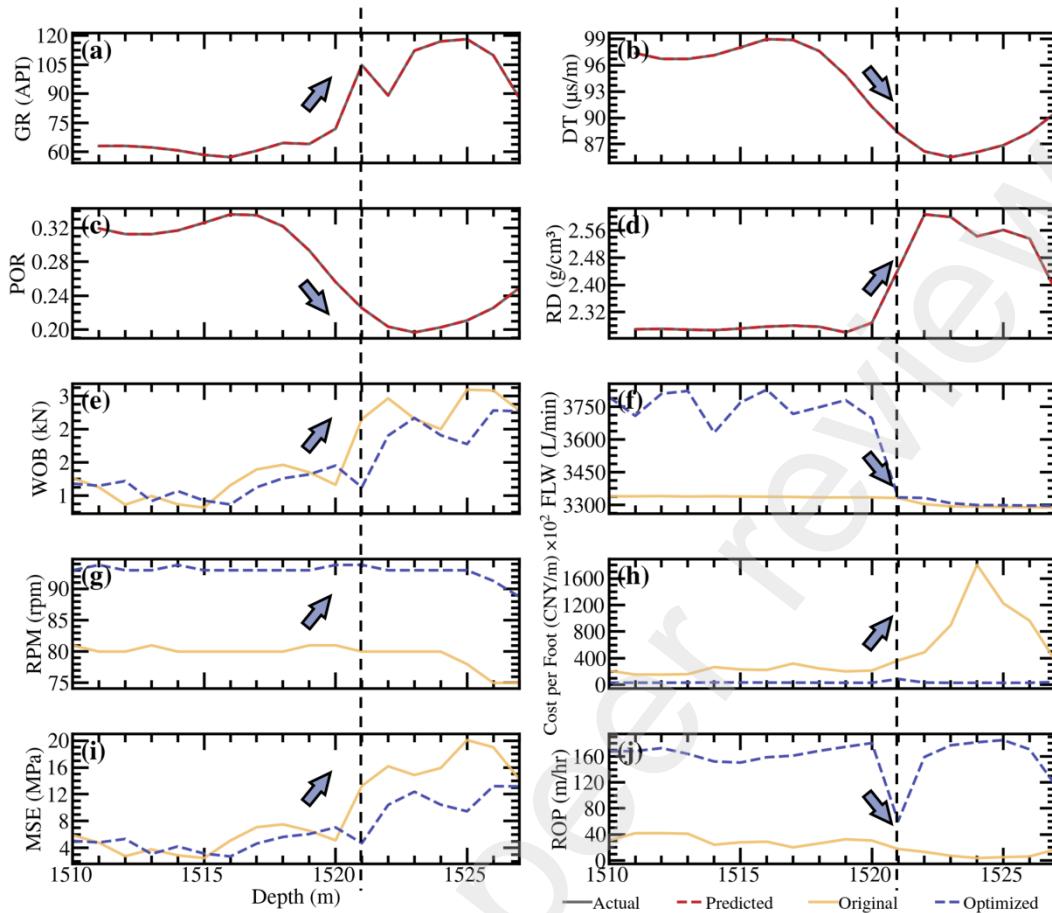
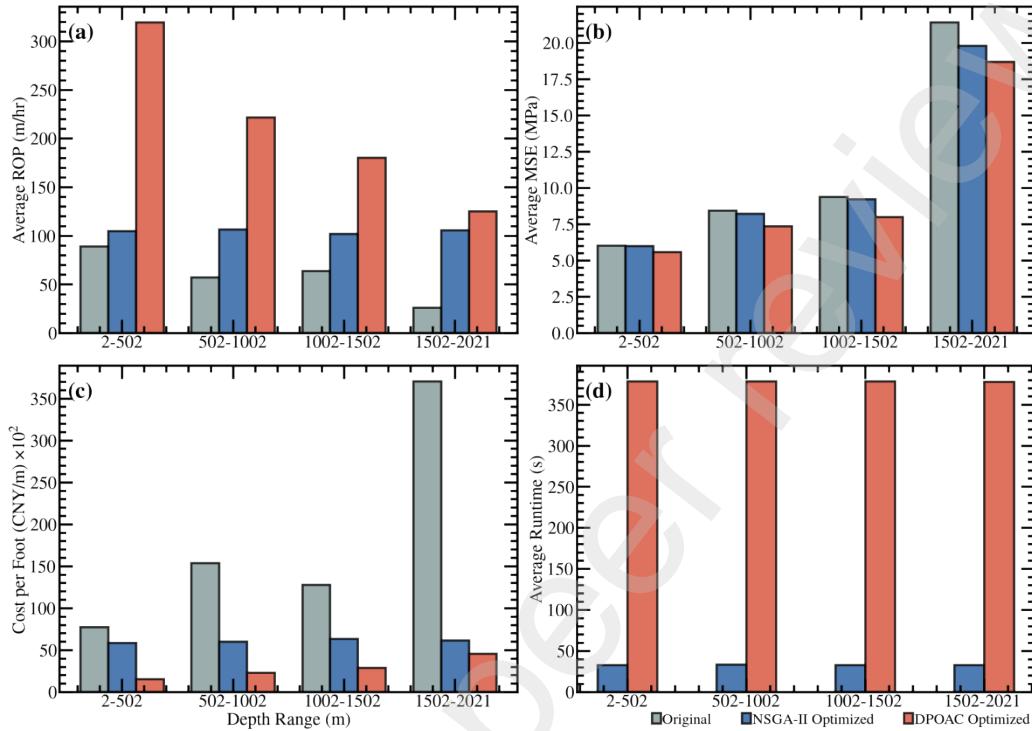


Fig. 30. Analysis of drilling response and DPOAC algorithm adjustments in the hard rock section from 1510 to 1530 m

Fig. 30 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. 30(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. 30(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. 30(h–j) indicated increases in the unit cost per foot and mechanical specific energy, along with a decrease in drilling speed, 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 drilling speed 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

1 variations, optimizing the drilling efficiency and cost-effectiveness. Under challenging conditions, such as hard
 2 rock layers, the algorithm not only refines the operational parameters but also ensures the safety and economic
 3 viability of the drilling process, affirming its pivotal role in modern drilling operations.



4
 5 **Fig. 31. Performance comparison between the DPOAC algorithm and NSGA-II in drilling optimization**

6 Fig. 31 shows a comprehensive comparison of different algorithms in drilling optimization, focusing on the
 7 performance of the DPOAC algorithm versus the Non-dominated Sorting Genetic Algorithm II (NSGA-II)
 8 across four crucial drilling stages. In Fig. 31(a), a comparison of the drilling speeds reveals a significant
 9 enhancement by the DPOAC algorithm in the early depth stages, indicating its effective broad exploration of the
 10 initial drilling parameters and rapid adaptation to geological conditions. As the drilling depth increased, the
 11 DPOAC continuously optimized the drilling speed and achieved stable and efficient performance, markedly
 12 surpassing that of the NSGA-II algorithm. This optimization of drilling speed demonstrates the capability of the
 13 DPOAC algorithm to ensure drilling efficiency while effectively managing drilling risks and reducing
 14 mechanical wear or operational errors caused by improper parameter selection.

15 Fig. 31(b) presents a comparison of the mechanical specific energies. Although the reduction in the
 16 mechanical specific energy by the DPOAC algorithm was not pronounced in the early depth stages, its
 17 optimization effects became increasingly apparent with depth. This trend underscores the adaptability and
 18 precision of the DPOAC algorithm under dynamic geological conditions, which is capable of fine-tuning drilling
 19 parameters to reduce energy consumption and enhance drilling machinery efficiency. Regarding the unit cost
 20 per foot, Fig. 31(c) shows that both the DPOAC algorithm and NSGA-II achieved significant cost reductions,

1 with the performance of the DPOAC algorithm proving to be superior. This not only highlights the efficiency of
2 the DPOAC algorithm in cost control but also reflects its advantages in maintaining the economic viability of
3 drilling operations.

4 Fig. 31(d) shows the significant differences in efficiency between the NSGA-II and DPOAC algorithms.
5 As a heuristic algorithm, the NSGA-II algorithm relies on complex iterative processes, often requiring thousands
6 of iterations to identify optimal drilling parameters, resulting in operational times that are more than ten times
7 those of the DPOAC algorithm. Although this iterative approach explores a broad parameter space, its high time
8 cost significantly reduces drilling efficiency in modern operations that demand rapid responses. In contrast, the
9 DPOAC algorithm employs a model-based strategy for optimizing continuous action spaces, allowing swift
10 adjustments to the drilling parameters in response to real-time geological changes. The optimization process of
11 the DPOAC algorithm is not only responsive, but also continuously updates its decision-making strategy,
12 ensuring real-time feedback from logging data while achieving an optimal balance between drilling speed and
13 cost.

14 **6. Conclusions**

15 A multi-objective real-time optimization reinforcement learning framework integrating temporal prediction
16 networks, domain adversarial networks, and Markov decision processes was successfully developed and
17 deployed. This framework significantly enhanced the ability to process and predict data from undrilled
18 formations by integrating feature extractors and domain discriminators, and adapting to variable geological
19 environments. The domain discriminator ensured high generalizability across diverse geological conditions,
20 whereas the feature extractor accurately extracted key information from complex drilling data, supporting model
21 migration and real-time optimization in new formations. The online learning strategy reinforced model
22 adaptability through incremental learning methods, allowing immediate updates when new drilling data were
23 received, maintaining data processing continuity and high quality, and ensuring long-term model stability and
24 optimization.

25 Formation property perception utilized LWD to analyze geological parameter changes in real time,
26 effectively addressing the lag effect in the real-time data processing of traditional drilling optimization methods.
27 This strategy significantly improved decision-making timeliness and accuracy, enabling a rapid response to
28 changes during the drilling process, optimizing drilling operations, and reducing risks. Furthermore, the real-
29 time application of formation pressure perception, considering wellbore stability, optimized the adjustment of
30 mud density by imposing constraints such as formation collapse, pore, fracture, and lost circulation pressures,
31 ensuring the safety and economic efficiency of drilling operations.

32 The construction of a Markov decision environment for real-time drilling parameter optimization enabled
33 the precise control and economic optimization of drilling operations. The DPOAC algorithm optimized the
34 economic and mechanical efficiencies of the drilling operations. By incorporating experience replay and priority-

1 adjustment mechanisms, the learning efficiency was improved, and the exploration noise strategy enhanced the
2 model adaptability, ensuring real-time responses to changes in geological conditions. This precise control of
3 drilling parameters not only reduced mechanical wear but also effectively decreased drilling costs. An empirical
4 analysis of the Caofidian 6-4 block of the Bohai Sea, China, demonstrated that the framework significantly
5 increased the drilling speed and optimized the mechanical specific energy and unit cost per foot, proving its
6 broad applicability and practical potential in modern drilling operations.

7 **CRediT authorship contribution statement**

8 **Yu Song:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding
9 acquisition. **Zehua Song:** Conceptualization, Investigation, Writing – original draft, Visualization. **Jin Yang:**
10 Conceptualization, Resources, Funding acquisition. **Kejin Chen:** Data curation, **Jizhou Tang:** Formal analysis,
11 Resources.

12 **Declaration of competing interest**

13 The authors declare that they have no known competing financial interests or personal relationships that
14 could have appeared to influence the work reported in this paper.

15 **Data availability**

16 The data that has been used is confidential.

17 **Acknowledgements**

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

21 **References**

- 22 [1] Yang J. Jetting Installation Method of Oil-Gas Well Conductor. *Installation Methods of Offshore Oil-Gas Well*
23 *Conductor*: Springer; 2022. p. 107-62.
- 24 [2] Liu Z, Liu K, Chen X, Ma Z, Lv R, Wei C, et al. Deep-sea rock mechanics and mining technology: State of the
25 art and perspectives. *International Journal of Mining Science and Technology*. 2023;33:1083-115.
- 26 [3] Wang Y, Gao D. Study on the marine environment limiting conditions of deepwater drilling for natural gas
27 hydrate. *Applied Energy*. 2022;312:118802.
- 28 [4] Chen X, Du X, Weng C, Yang J, Gao D, Su D, et al. A real-time drilling parameters optimization method for
29 offshore large-scale cluster extended reach drilling based on intelligent optimization algorithm and machine
30 learning. *Ocean Engineering*. 2024;291:116375.
- 31 [5] Kendall H, Goins Jr W. Design and operation of jet-bit programs for maximum hydraulic horsepower, impact
32 force or jet velocity. *Transactions of the AIME*. 1960;219:238-50.
- 33 [6] Eckel JR. Microbit studies of the effect of fluid properties and hydraulics on drilling rate. *Journal of Petroleum*
34 *Technology*. 1967;19:541-6.
- 35 [7] Bourgoyne Jr AT, Young Jr F. A multiple regression approach to optimal drilling and abnormal pressure
36 detection. *Society of Petroleum Engineers Journal*. 1974;14:371-84.

- 1 [8] Waughman RJ, Kenner JV, Moore RA. Real-time specific energy monitoring reveals drilling inefficiency and
2 enhances the understanding of when to pull worn PDC bits. SPE/IADC Drilling Conference and Exhibition:
3 SPE; 2002. p. SPE-74520-MS.
- 4 [9] Dupriest FE, Koederitz WL. Maximizing drill rates with real-time surveillance of mechanical specific energy.
5 SPE/IADC Drilling Conference and Exhibition: SPE; 2005. p. SPE-92194-MS.
- 6 [10] Iversen F, Cayeux E, Dvergsnes E, Gravdal J, Vefring E, Mykletun B, et al. Monitoring and control of drilling
7 utilizing continuously updated process models. SPE/IADC Drilling Conference and Exhibition: SPE; 2006. p.
8 SPE-99207-MS.
- 9 [11] Zhang YX, Wang LF, Yao F, Wang H, Men XS, Ye BB. Experimental study on the influence of nozzle diameter
10 on abrasive jet cutting performance. Advanced Materials Research. 2011;337:466-9.
- 11 [12] Sui D, Nybø R, Azizi V. Real-time optimization of rate of penetration during drilling operation. 2013 10th IEEE
12 International Conference on Control and Automation (ICCA): IEEE; 2013. p. 357-62.
- 13 [13] Guria C, Goli KK, Pathak AK. Multi-objective optimization of oil well drilling using elitist non-dominated
14 sorting genetic algorithm. Petroleum Science. 2014;11:97-110.
- 15 [14] Zhang Q-Z, Wei-Xiao L, Sha L-X. Drilling parameters optimization based on chaotic multi-objective particle
16 swarm optimization algorithm. 2017 International Conference on Electrical Engineering and Automation
17 Control, ICEEAC2017.
- 18 [15] Zheng J, Li ZL, Dou B, Lu C. Multi-objective cellular particle swarm optimization and RBF for drilling
19 parameters optimization. Mathematical Biosciences and Engineering: MBE. 2019;16:1258-79.
- 20 [16] Khaleel AA, Adnan MS, Alhamd SJ. Estimation of Bourgoyne and young model coefficients to predict optimum
21 drilling rates and bit weights using genetic algorithms—a case study of the Faihaa oil field in Iraq. IOP Conference
22 Series: Materials Science and Engineering: IOP Publishing; 2021. p. 012154.
- 23 [17] Zhang Y, Tiño P, Leonardis A, Tang K. A survey on neural network interpretability. IEEE Transactions on
24 Emerging Topics in Computational Intelligence. 2021;5:726-42.
- 25 [18] Sabah M, Talebkeikhah M, Wood DA, Khosravianian R, Anemangely M, Younesi A. A machine learning
26 approach to predict drilling rate using petrophysical and mud logging data. Earth Science Informatics.
27 2019;12:319-39.
- 28 [19] Hegde C, Daigle H, Gray KE. Performance comparison of algorithms for real-time rate-of-penetration
29 optimization in drilling using data-driven models. Spe Journal. 2018;23:1706-22.
- 30 [20] Matias J, de Castro Oliveira JP, Le Roux GA, Jäschke J. Real-time optimization with persistent parameter
31 adaptation applied to experimental rig. Ifac-papersonline. 2021;54:475-80.
- 32 [21] Jahandideh A, Jafarpour B. Closed-loop stochastic oilfield optimization for hedging against geologic,
33 development, and operation uncertainty. Computational Geosciences. 2020;24:129-48.
- 34 [22] Magnusson LV, Olsson JR, Tran CTT. Recurrent neural networks for oil well event prediction. IEEE Intelligent
35 Systems. 2023;38:73-80.
- 36 [23] Osarogiagbon AU, Oloruntobi O, Khan F, Venkatesan R, Butt S. Gamma ray log generation from drilling
37 parameters using deep learning. Journal of Petroleum Science and Engineering. 2020;195:107906.
- 38 [24] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A comprehensive survey on transfer learning. Proceedings
39 of the IEEE. 2020;109:43-76.
- 40 [25] Fan C, Sun Y, Xiao F, Ma J, Lee D, Wang J, et al. Statistical investigations of transfer learning-based
41 methodology for short-term building energy predictions. Applied Energy. 2020;262:114499.
- 42 [26] Sutton RS, Barto AG. Reinforcement learning: An introduction2018.
- 43 [27] Wu J, He H, Peng J, Li Y, Li Z. Continuous reinforcement learning of energy management with deep Q network
44 for a power split hybrid electric bus. Applied energy. 2018;222:799-811.
- 45 [28] Jiang Q, Zhong S, Cui J, Feng X-T, Song L. Statistical characterization of the mechanical parameters of intact
46 rock under triaxial compression: An experimental proof of the Jinping marble. Rock Mechanics and Rock
47 Engineering. 2016;49:4631-46.
- 48 [29] Sen S, Kundan A, Kumar M. Modeling pore pressure, fracture pressure and collapse pressure gradients in
49 offshore panna, western India: implications for drilling and wellbore stability. Natural Resources Research.
50 2020;29:2717-34.
- 51 [30] Zhang L, Yan X, Yang X, Tian Z, Yang H. An elastoplastic model of collapse pressure for deep coal seam drilling
52 based on Hoek–Brown criterion related to drilling fluid loss to reservoir. Journal of Petroleum Science and
53 Engineering. 2015;134:205-13.
- 54 [31] Ma T, Zhang Q, Chen P, Yang C, Zhao J. Fracture pressure model for inclined wells in layered formations with
55 anisotropic rock strengths. Journal of Petroleum Science and Engineering. 2017;149:393-408.

- 1 [32] Lahann RW, Swarbrick RE. Fracture pressure, leak-off tests and Poisson's ratio. *Petroleum Geoscience*.
2 2022;28:petgeo2021-103.
- 3 [33] Handin J. On the Coulomb - Mohr failure criterion. *Journal of Geophysical Research*. 1969;74:5343-8.
- 4 [34] Breckels I, Van Eekelen H. Relationship between horizontal stress and depth in sedimentary basins. *Journal of*
5 *Petroleum Technology*. 1982;34:2191-9.
- 6 [35] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain.
7 *Psychological review*. 1958;65:386.
- 8 [36] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *nature*.
9 1986;323:533-6.
- 10 [37] Liu Y, Hu T, Zhang H, Wu H, Wang S, Ma L, et al. itransformer: Inverted transformers are effective for time
11 series forecasting. *arXiv preprint arXiv:231006625*. 2023.
- 12 [38] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. *Advances*
13 *in neural information processing systems*. 2017;30.
- 14 [39] Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H, et al. Informer: Beyond efficient transformer for long sequence
15 time-series forecasting. *Proceedings of the AAAI conference on artificial intelligence*2021. p. 11106-15.
- 16 [40] Hornik K. Approximation capabilities of multilayer feedforward networks. *Neural networks*. 1991;4:251-7.
- 17 [41] Xu J, Sun X, Zhang Z, Zhao G, Lin J. Understanding and improving layer normalization. *Advances in neural*
18 *information processing systems*. 2019;32.
- 19 [42] Joseph VR. Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*.
20 2022;15:531-8.
- 21 [43] Yao X. A new simulated annealing algorithm. *International Journal of Computer Mathematics*. 1995;56:161-8.
- 22 [44] Loshchilov I, Hutter F. Decoupled weight decay regularization. *arXiv preprint arXiv:171105101*. 2017.
- 23 [45] Girshick R. Fast r-cnn. *Proceedings of the IEEE international conference on computer vision*2015. p. 1440-8.
- 24 [46] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural
25 networks from overfitting. *The journal of machine learning research*. 2014;15:1929-58.
- 26 [47] Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *Journal of machine learning research*.
27 2012;13.
- 28 [48] Mockus J. Application of Bayesian approach to numerical methods of global and stochastic optimization. *Journal*
29 *of Global Optimization*. 1994;4:347-65.
- 30 [49] Kennedy J, Eberhart R. Particle swarm optimization. *Proceedings of ICNN'95-international conference on neural*
31 *networks: ieee*; 1995. p. 1942-8.
- 32 [50] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural computation*. 1997;9:1735-80.
- 33 [51] Huang W, Cheng J, Yang Y, Guo G. An improved deep convolutional neural network with multi-scale
34 information for bearing fault diagnosis. *Neurocomputing*. 2019;359:77-92.
- 35 [52] Kim H, Youn BD. A new parameter repurposing method for parameter transfer with small dataset and its
36 application in fault diagnosis of rolling element bearings. *Ieee Access*. 2019;7:46917-30.
- 37 [53] Siahkoohi A, Louboutin M, Herrmann FJ. The importance of transfer learning in seismic modeling and imaging.
38 *Geophysics*. 2019;84:A47-A52.
- 39 [54] Luo Y, Yin L, Bai W, Mao K. An appraisal of incremental learning methods. *Entropy*. 2020;22:1190.
- 40 [55] Hossain ME. Drilling costs estimation for hydrocarbon wells. *Journal of Sustainable Energy Engineering*.
41 2015;3:3-32.
- 42 [56] Chen X, Gao D, Guo B, Feng Y. Real-time optimization of drilling parameters based on mechanical specific
43 energy for rotating drilling with positive displacement motor in the hard formation. *Journal of Natural Gas*
44 *Science and Engineering*. 2016;35:686-94.
- 45 [57] Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: Off-policy maximum entropy deep reinforcement
46 learning with a stochastic actor. *International conference on machine learning: PMLR*; 2018. p. 1861-70.
- 47

48

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

2 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
 3 China.

4 **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
...

5 **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
...

1 Appendix B—Optimization Process for Drilling Parameters Using the DPOAC Algorithm

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

4 **Table B1. 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

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 $Q_{\omega_1}, Q_{\omega_2}$

2. **Initialize** target critic networks: $\pi_{\theta'} = \pi_\theta, Q_{\omega'_1} = Q_{\omega_1}, Q_{\omega'_2} = Q_{\omega_2}$

3. **for** $i=1$ **to** T **do:**

4. Observe initial state s_1

5. **while** s_t is not terminal **do:**

6. Select action $a_t \sim \pi_\theta(s_t) + N_t$, where π_θ outputs parameters $\mu_\theta(s_t)$ and $\sigma_\theta(s_t)$ of a Gaussian distribution

7. Execute action a_t , observe reward r_t , new state s_{t+1} , and terminal status

8. Store transition (s_t, a_t, r_t, s_{t+1}) in D

9. Sample a minibatch of m transitions (s_j, a_j, r_j, s_{j+1}) from D

10. Set $y_j = r_j + \gamma \min(Q_{\omega'_1}(s_{j+1}, \pi_{\theta'}(s_{j+1})), Q_{\omega'_2}(s_{j+1}, \pi_{\theta'}(s_{j+1})))$ **if** s_{j+1} is not terminal, else $y_j = r_j$

11. Update critic $Q_{\omega_1}, Q_{\omega_2}$ by minimizing loss: $L = \frac{1}{m} \sum_{j=1}^m \left((y_j - Q_{\omega_1}(s_j, a_j))^2 + (y_j - Q_{\omega_2}(s_j, a_j))^2 \right)$

12. Update actor using policy gradient: $\nabla_\theta J \approx \frac{1}{m} \sum_{j=1}^m \left(\nabla_\theta \pi_\theta(s_j) \left(\nabla_a Q_{\omega_1}(s_j, a) \Big|_{a=\pi_\theta(s_j)} - \log \pi_\theta(a | s_j) + \alpha \right) \right)$

13. Adjust temperature parameter α for entropy regularization

-
13. Soft update target networks:
 14. $\omega'_1 \leftarrow \tau\omega_1 + (1-\tau)\omega'_1$
 15. $\omega'_2 \leftarrow \tau\omega_2 + (1-\tau)\omega'_2$
 16. $\theta' \leftarrow \tau\theta + (1-\tau)\theta'$
 17. $s_t = s_{t+1}$
 18. **end while**
 19. **end for**
 20. **Return** optimized network parameters $\theta, \omega_1, \omega_2$
-

1