

A multi-objective reinforcement learning framework for real-time drilling optimization based on symbolic regression and perception

Zehua Song^a, Yu Song^{a,*}, Jin Yang^b, Baosheng Liu^c, Bingzhen Gao^b, Jizhou Tang^d

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

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

^c CNOOC China Limited, Tianjin Branch, Tianjin, 300459, China

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



ARTICLE INFO

Keywords:

Drilling parameter
Multi-objective optimization
Symbolic regression
Drilling sensing
Deep reinforcement learning

ABSTRACT

In the context of the global energy transition, optimizing deep-water oil and gas drilling parameters is crucial for ensuring safety while improving efficiency. Traditional methods face limitations in highly dynamic and nonlinear drilling environments, struggling to balance speed and cost-effectiveness. Furthermore, these methods rely on real-time logging-while-drilling (LWD) data for decision-making, but delays in data collection and processing hinder timely adjustments of drilling parameters, affecting decision accuracy and responsiveness. This paper proposes a multi-objective drilling parameter optimization framework, incorporating symbolic regression, time-series networks, and Markov decision processes to precisely predict ROP, formation conditions, and optimize drilling parameters in real time. Key innovations include a multi-population evolutionary symbolic regression algorithm for constructing empirical equations, the integration of variational mode decomposition (VMD) and sample entropy for data preprocessing, and multi-head self-attention time-series networks to enhance prediction accuracy. Quantile regression further estimates the range of drilling parameter adjustments. Additionally, a drilling parameter optimization deep deterministic policy gradient (DPODDPG) algorithm was developed to automate real-time parameter adjustments. Empirical analysis on the Ledong 10-1 block in the South China Sea demonstrated significant improvements: ROP increased from 54.18 m/hr to 122.17 m hr, mechanical specific energy (MSE) decreased from 100.82 MPa to 97.78 MPa, and cost per foot reduced from 121.16×10^2 CNY/m to 51.31×10^2 CNY/m. Compared to traditional methods, the proposed framework showed clear advantages in enhancing ROP, reducing MSE, and controlling costs, further validating its superiority in complex drilling environments. This method not only significantly improves drilling efficiency and economic benefits but also adapts to complex and changing drilling conditions, showing broad application potential, particularly in challenging deep-water oil and gas drilling operations, where it can provide more efficient and reliable optimization solutions.

1. Introduction

In the context of the global energy structure transition and low-carbon development strategies, deep-water oil and gas drilling technologies, which are key technologies for developing deep-sea oil and gas resources, are crucial for ensuring energy security and achieving sustainable development (Liu et al., 2023; Wang and Gao, 2022; Yang, 2022). However, deep-water drilling faces challenges such as extreme environmental conditions, complex geological structures, and high economic costs, making the precise optimization of drilling parameters a core task to ensure drilling safety, improve operational efficiency, and

reduce drilling risks (Chen et al., 2024). Therefore, exploring effective and reliable strategies for drilling parameter optimization is of significant theoretical and practical value for advancing deep-water drilling technology and optimizing drilling operations.

In the field of drilling parameter optimization, numerous researchers have attempted to explore various methods for enhancing drilling efficiency. Kendall and Goins Jr. Investigated the hydraulic parameters in drilling; established relationships between the maximum impact force, maximum hydraulic power, and maximum jet velocity; and developed a program to optimize these parameters (Kendall and Goins Jr., 1960). Eckel further explored the impact of drilling fluid properties and

* Corresponding author.

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

hydraulic parameters on ROP, and delineated the correlations between drilling fluid density, kinematic viscosity, and changes in hydraulic parameters (Eckel, 1967). Through multivariate regression analysis, Bourgoyne Jr. And Young Jr. Examined the effects of drilling pressure, rotational speed, bit wear, hydraulic parameters, and drillability coefficients on ROP and proposed a regression model for ROP (Bourgoyne Jr and Young Jr, 1974). Waughman et al. effectively identified instances of inefficient drilling by comparing the real-time mechanical specific energy (MSE) with standard values (Waughman et al., 2002). Dupriest and Koederitz used real-time MSE to assess the efficiency of drilling bits, whereas Iversen et al. developed methods for real-time optimization of drilling parameters (Dupriest and Koederitz, 2005; Iversen et al., 2006). Zhang et al. observed that reducing the nozzle diameter and increasing jet velocity significantly improved the drilling efficiency by enhancing fluid jet speed and penetration capability, corroborating earlier findings by ExxonMobil (Zhang et al., 2011). Sui et al. devised a multivariate extremum searching method for the real-time optimization of drilling parameters to minimize the MSE and enhance drilling efficiency and safety (Sui et al., 2013). Guria et al. applied a non-dominated sorting genetic algorithm with elitist strategies for the multi-objective optimization of drilling parameters, demonstrating the potential of multi-objective optimization (Guria et al., 2014). Zhang et al. proposed an improved multi-objective particle swarm optimization algorithm focused on ROP, bit life, and MSE to avoid local optima, showing optimum convergence and distribution (Zhang et al., 2017). Numerical investigations and combined percussive-rotary drilling techniques have significantly enhanced ROP and drill bit longevity (Xi et al., 2022, 2023; Zha et al., 2017). Moreover, precise control of pressure and flow rate in the wellbore can promote penetration and enhance drilling performance (Gavidia et al., 2024; Mohammadizadeh et al., 2021, 2023; Rangel Gavidia et al., 2023; Vásconez Garcia et al., 2024). To achieve such improvements, Zheng et al. and Khaleel et al. introduced a multi-objective cellular particle swarm optimization algorithm and applied genetic algorithms to estimate the coefficients of the Bourgoyne and Young model, significantly enhancing drilling efficiency, reducing operational costs, and ensuring more effective drilling operations (Khaleel et al., 2021; Zheng et al., 2019).

Although intelligent computational methods have made significant exploratory progress in drilling parameter optimization, traditional empirical equations, neural network black-box models, and multi-objective optimization algorithms still have limitations. Traditional empirical equations are based on simplified physical assumptions and lack comprehensive descriptive capabilities for complex drilling processes (Yimwadsana, 2023). Although neural network models can capture complex nonlinear relationships, their opaque nature limits their interpretability and generalization of the models (Zhang et al., 2021). In addition, relying solely on real-time logging-while-drilling (LWD) parameters for optimization has inherent limitations. By the time the data were acquired, the corresponding drilling activities had already occurred, often resulting in real-time parameter adjustments that did not achieve the expected effects. Such methods do not adequately consider latency effects in drilling operations or the need for drilling parameters to adapt dynamically to different drilling moments (Hegde et al., 2018; Matias et al., 2021). Moreover, multi-objective optimization during the drilling process faces numerous challenges, such as the complexity and uncertainty of the geological environment, which render the drilling parameter optimization problem highly dynamic and nonlinear. These factors make it difficult for traditional algorithms, such as heuristic algorithms, to adjust drilling parameters in real time to respond to changes in formations, often resulting in low computational efficiency and inadequate real-time response capabilities, which makes it challenging to simultaneously meet multiple objectives, such as increasing ROP, controlling bit wear, and optimizing energy consumption (Jahandideh and Jafarpour, 2020).

Symbolic regression, drilling perception, and reinforcement learning offer new solutions to the aforementioned challenges in drilling

parameter optimization. Symbolic regression, a regression analysis method based on genetic programming, discovers mathematical relationships in data and represents them as symbolic expressions, free from the constraints of traditional model forms (Kim and Han, 2020). This gives symbolic regression a unique advantage in revealing the complex relationships between the real-time drilling parameters and drilling performance. Drilling perception utilizes time-series prediction models to analyze historical and current drilling parameters and predict future parameters, thereby enabling the preemptive optimization of future drilling parameters (Magnusson et al., 2023; Osarogiagbon et al., 2020). Reinforcement learning, a machine learning method based on trial-and-error mechanisms, has also shown great potential for drilling parameter optimization (Sutton and Barto, 2018; Wu et al., 2018). Through interactions with the environment, reinforcement learning algorithms can learn decision-making strategies for optimal drilling operations across different drilling processes, ensuring safety and economic efficiency while effectively improving ROP. Thus, integrating symbolic regression, drilling perception, and reinforcement learning can create a more accurate, interpretable, and adaptive framework for drilling parameter optimization, effectively addressing the complex challenges in deep-water drilling.

This study proposes a comprehensive multi-objective framework for the real-time optimization of drilling parameters by integrating symbolic regression, time-series forecasting networks, and Markov decision processes to precisely predict the ROP and real-time optimization of drilling parameters. It includes the design of the multi-evolutionary symbolic regression for ROP (MESR-DS) algorithm to construct empirical equations for ROP and to establish equations that comprehensively map the complex relationships among drilling, mud logging, and real-time LWD parameters. This provides an accurate quantitative description of the dynamic characteristics of drilling and the interaction mechanisms among the parameters. Furthermore, the framework integrates a time-series prediction network equipped with a multi-head self-attention mechanism, enhancing the prediction accuracy of real-time drilling parameters by designing variational mode decomposition and sample entropy reconstruction data processing methods. In addition, by incorporating quantile regression, we estimated the adjustment ranges for the drilling parameters. A markov decision process (MDP) environment for drilling parameter optimization was established, and the drilling parameters optimization deep deterministic policy gradient (DPODDPG) algorithm was designed. This integrated framework achieved automated real-time adjustments of drilling parameters, thereby significantly enhancing the economic and mechanical efficiencies of drilling operations. The evaluation results from the Ledong 10-1 block in the South China Sea demonstrate the framework's significant potential for efficiency enhancement in drilling applications, providing technological reserves and support for the development and practice of future deep-water drilling operations. The following sections of this paper are organized as follows. Section 3 introduces the framework for drilling parameter optimization based on symbolic regression and time-series forecasting networks; Section 4 explains the data preprocessing methods and the design of the multi-head self-attention mechanism; Section 5 presents the drilling parameter optimization approach based on the DPODDPG algorithm and the establishment of a MDP environment; Section 6 provides an empirical analysis using the Ledong 10-1 block in the South China Sea to validate the effectiveness of the proposed framework; Section 7 concludes the paper and discusses future research directions.

2. Workflow

Construction of empirical equations for ROP based on symbolic regression: The MESR-DS algorithm was designed to construct empirical equations accurately during the drilling process. Using a multi-evolutionary strategy, this algorithm automatically refined equations that optimally described the complex relationships between drilling,

mud logging, and real-time LWD parameters. Using drilling data from operations in the South China Sea, geological and operational parameters were integrated, and data preprocessing steps such as missing value handling, outlier removal, and wavelet denoising were implemented to ensure high-quality input data. The training process incorporated parallel computing and simulated annealing strategies, which significantly enhanced the efficiency and accuracy of the empirical equation construction for the ROP. The developed equations supported real-time decision-making and optimization of drilling operations, ensuring high adaptability and reliability in practical drilling applications.

Drilling perception based on VMD-time series architecture: Focuses on a deep learning architecture that utilizes one-dimensional convolutional neural networks (1D-CNN), bidirectional long short-term memory (BiLSTM) networks, and multihead attention (MA) mechanisms for predicting key real-time LWD parameters. Variational mode decomposition (VMD) and sample entropy-based data reconstruction methods were used to preprocess the original real-time logging parameters, effectively extract key frequency components, and reduce noise to enhance the input data quality. The deep learning architecture precisely captured local features through the 1D-CNN layer and assesses forward and backward dependencies in time-series data via the BiLSTM layer. The MA mechanism enhanced the capacity of the model to understand and capture complex dependencies. Adaptive learning rate adjustments and early stopping strategies were implemented during model training to enhance precision and generalization across diverse geological formations. Additionally, quantile regression methods and a quantile loss function were introduced to predict the adjustment range of the drilling parameters, further enhancing the adaptability and responsiveness of real-time decision-making in drilling operations.

Design of a Markov decision environment for real-time drilling parameters: A MDP environment was constructed for the optimization of drilling parameters to achieve precise control and optimize economic

and mechanical benefits. This environment integrated empirical equations for ROP based on the MESR-DS algorithm and used these equations to predict the ROP in real time and evaluate the cost per foot and MSE. The reward function assessed the economic efficiency and cost-effectiveness of drilling parameter adjustments based on the predicted ROPs and actual drilling data, with the aim of optimizing the trade-off between cost and equipment efficiency. The design of the state and action spaces relied closely on the drilling perception model, which continuously sensed real-time LWD parameters through real-time data, providing dynamic inputs for the MDP environment to support real-time adaptability adjustments and the optimization of drilling parameters. The MDP environment enhanced the performance of the model in dynamic and complex drilling environments, providing data support for precise and accurate drilling operation decisions.

Implementation and deployment of the Deep Deterministic Policy Gradient algorithm for drilling parameters optimization: In conjunction with the MDP environment, the DPODDPG algorithm was designed to optimize the drilling parameters. Through finely tuned Q-networks, target networks, and policy networks, the algorithm optimized the economic benefits and mechanical efficiency of drilling operations, ensuring the automation and precision of the decision-making processes. The DPODDPG algorithm combined experience replay and priority-adjustment mechanisms to enhance learning efficiency and, through an exploration noise strategy, boosted the model's exploration capabilities, thus adapting in real time to geological conditions during actual drilling operations. Additionally, the process was validated in the Ledong 10-1 block in the South China Sea, achieving an optimal balance between ROP and cost through continuous parameter adjustments, thereby confirming its practical value in enhancing the economic and efficiency aspects of drilling operations. The workflow is illustrated in Fig. 1.

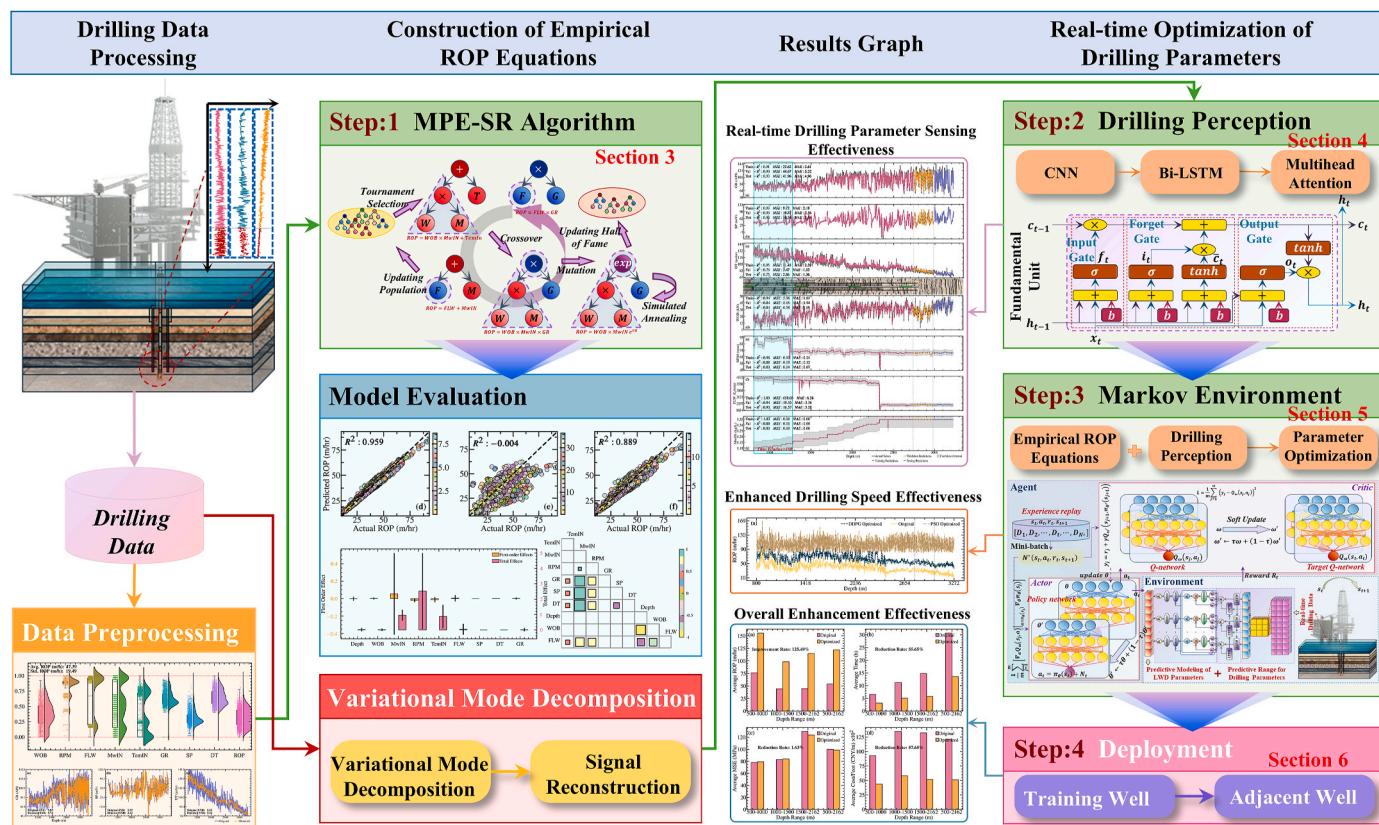


Fig. 1. Workflow process.

3. Construction of empirical ROP equation based on the MESR-DS algorithm

The construction of empirical ROP equations using the MESR-DS algorithm was conducted to predict ROP, involving the evaluation of various equation complexities. A quantitative global sensitivity analysis was performed using the Sobol variance method to assess the influence of key parameters on the predicted ROP.

3.1. MESR-DS algorithm

3.1.1. Application of symbolic regression in ROP prediction and limitations of traditional methods

ROP prediction, a core issue in real-time drilling engineering, traditionally relies on empirical formulas and data-driven machine learning models. Although widely used, empirical formulas such as the Bourgoyne and Young equations are limited by the selection of parameters and manually set coefficients, often resulting in insufficient accuracy (Bourgoyne Jr and Young Jr, 1974). Advances in machine learning have led to increasing research adopting algorithms, such as support vector machines and random forests, to model the relationship between drilling parameters and ROP (Barbosa et al., 2019; Mehrad et al., 2020). However, despite excellent performance on specific datasets, the "black box" nature of these algorithms limits their practical application, particularly due to poor model interpretability and low transferability. Models trained on specific blocks struggle to adapt to the drilling conditions in other blocks and require retraining, significantly limiting their widespread application in the drilling industry (Ahmed et al., 2019).

Symbolic regression, a flexible data-driven modeling approach that uses genetic programming to discover optimal mathematical expressions describing data, overcomes the limitations of traditional methods (Udrescu and Tegmark, 2020). These expressions not only provide high accuracy but also offer excellent interpretability, making them particularly suitable for predicting ROP. Symbolic regression can automatically construct mathematical expressions based on LWD, mud logging, and drilling parameters without the need to set specific forms or coefficients manually. Moreover, the high transparency and interpretability of its models provide deep insights into the drilling optimization and decision-making process, revealing the significant impact of geological parameters on ROP. These findings can be directly applied to optimize drilling operations to ensure precision and cost efficiency.

3.1.2. Architecture of the MESR-DS algorithm

To accurately identify and describe ROPs, a novel symbolic regression algorithm called the MESR-DS was developed. This algorithm leverages the core principles of genetic algorithms tailored for complex drilling data environments to adapt to continuously changing geological conditions. MESR-DS employs a multi-population evolutionary strategy, allowing the asynchronous execution of multiple evolutionary

processes, with each population operating independently based on classical evolutionary algorithms (Wu et al., 2016). Employing tournament selection as the individual selection mechanism, MESR-DS not only maintains population diversity, but also efficiently filters individuals with high adaptability (Bäck and Schwefel, 1993). By evaluating multivariate parameter expressions related to ROP, this strategy optimizes the models that best describe the drilling efficiency. Additionally, MESR-DS implements mutation and crossover operations to innovate individuals and simulates natural variations and genetic recombination. Mutations may involve altering operators and adding or removing function nodes, whereas crossover involves the exchange of structural parts of expressions under different geological conditions, producing offspring with new traits that are adapted to diverse drilling environments. The architecture of the MESR-DS algorithm is illustrated in Fig. 2. The workflow of the MESR-DS algorithm can be found in Table B1 of Appendix B.

The MESR-DS algorithm incorporates the simulated annealing method to adjust the fitness function, controlling the probability of accepting solutions during the mutation process through a temperature parameter $T \in [0, 1]$. The acceptance or rejection of mutations is based on their contribution to the fitness, and the rejection probability, p calculated according to Eq. (1) (Delahaye et al., 2019):

$$p = \exp\left(\frac{LF - LE}{\alpha T}\right) \quad (1)$$

where LF and LE are the fitness of an individual after and before mutation, respectively, and α is a hyperparameter that adjusts the temperature scale. The alternating use of high- and low-temperature phases in this method increases individual diversity and introduces new innovative potential during the search process. Particularly, when dealing with complex geological parameters, this allows the algorithm to explore a broad range of unknown mathematical relationships, thereby enhancing the model's exploratory power and adaptability (Li et al., 2020a).

The MESR-DS algorithm was optimized to enhance the evolutionary efficiency of mathematical expressions using a tournament selection method. It incorporates a "simplify-optimize" cycle strategy (Coello and Montes, 2002), where the simplification phase aims to transform expressions related to mud logging, drilling parameters, and while drilling logging into more concise equivalent forms. The optimization phase focuses on the precise adjustment of constants within the expressions, such as rock physics constants. Simplification operations periodically apply rules of algebraic equivalence to eliminate redundancy in expressions, while maintaining their descriptive power. The optimization phase iteratively adjusts the constants using classical optimization algorithms, thereby enhancing the fit of the model to the experimental data. This strategy explores new possibilities through redundant but necessary intermediate expression states, effectively balancing the

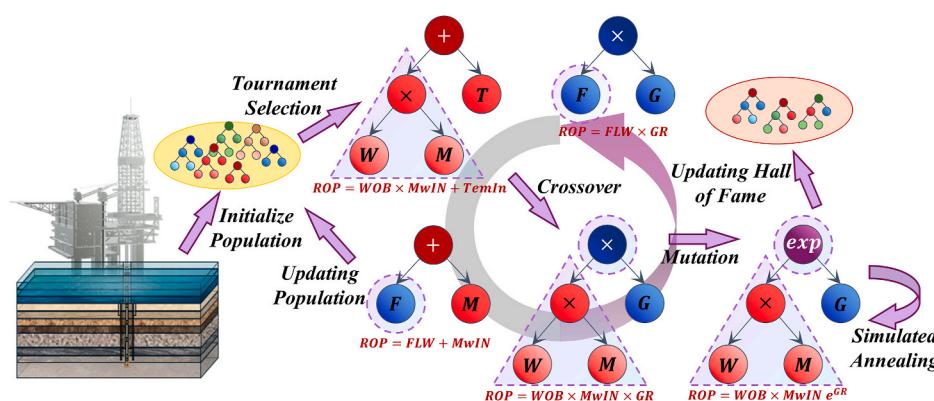


Fig. 2. Architecture of the MESR-DS algorithm.

search efficiency and exploring new expressions. In addition, the MESR-DS algorithm employs an adaptive simplicity index to dynamically adjust the complexity penalty of the expressions, optimizing the traditional loss function's penalty for complexity, which is usually adjusted according to Eq. (2) (Cranmer, 2023):

$$L(E) = L_{\text{pred}}(E) + \lambda \cdot C(E) \quad (2)$$

where $L_{\text{pred}}(E)$ represents the predictive loss measured by mean squared error, $C(E)$ is the complexity of the expression E , and λ is a hyper-parameter regulating the impact of complexity. Equation (2) effectively prevents overfitting by imposing direct penalties on complex expressions, although this may restrict the exploration of potentially effective yet slightly complex expressions using the model. To address this limitation, the MESR-DS algorithm implements an adaptive simplicity mechanism that dynamically adjusts the penalty intensity for complexity levels, thereby achieving a balance between the expressions of varying complexities within the population. The expression for this mechanism is given by Eq. (3) (Cranmer, 2023):

$$L(E) = L_{\text{pred}}(E) \cdot \exp(\text{"frecency"} [C(E)]) \quad (3)$$

where "frecency" is a metric that measures the frequency and recency of the complexity $C(E)$ within the population.

The empirical equation for ROP is constructed using the MESR-DS algorithm. The parameters utilized in this process include drilling parameters (WOB, RPM, and FLW), mud logging parameters (MwIN and TemIN), and LWD parameters (GR, SP, and DT).

3.2. Data preprocessing strategies

3.2.1. Data overview

Drilling data collected from specified well sites in the Ledong 10-1 block in the South China Sea focused on pre- and during-drilling data, excluding post-drilling data. Pre-drilling and during-drilling data provide direct observations and measurements crucial for real-time decision-making and optimization of ROP. In contrast, although post-drilling data are valuable for retrospective analysis, they are often processed in a less timely manner, rendering them unsuitable for immediate-response model applications. The data encompass multidimensional geological and operational parameters divided into three categories: drilling, mud logging, and LWD parameters. Drilling parameters, including the WOB, RPM, and FLW, directly reflect the physical operating conditions of the drilling process. Mud logging parameters, such as MwIN and TemIN, provide key information on fluid dynamics during drilling, which is essential for understanding downhole fluid behavior and controlling subterranean pressure. LWD parameters, including GR, SP, and DT, provide real-time data on formation lithology, electrical properties, and acoustic characteristics, offering a direct and scientific basis for the immediate evaluation of encountered formations.

3.2.2. Handling missing values and removing outliers

In the analysis of while-drilling data, ensuring data quality is a fundamental requirement that encompasses the effective handling of missing values and the precise identification and removal of outliers. For frequently occurring missing data within a dataset, a statistically based imputation strategy is utilized, particularly through imputation using the means of related attributes, to maintain the fundamental statistical properties of the data (García et al., 2015). For crucial drilling parameters such as rotational speed, linear interpolation is employed to fill gaps by inserting estimates between known data points, thus ensuring the continuity and accuracy of the time series data. The distributions of the crucial drilling, mud logging, and LWD parameters are illustrated in Fig. 3.

An interquartile range (IQR)-based strategy was used to identify and exclude outliers. This method calculates the first (Q1) and third quartiles (Q3) for each variable, defining the IQR as the difference between these

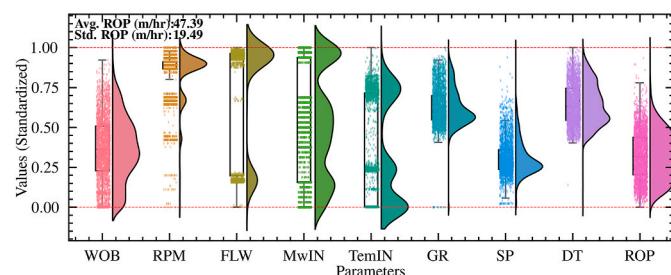


Fig. 3. Statistical distribution of key drilling, mud logging, and LWD parameters (Standardized).

two quartiles. Outliers are identified as data points lying below Q1 minus 1.5 times the IQR or above Q3 plus 1.5 times the IQR (Wan et al., 2014). This approach effectively distinguishes deviations caused by measurement errors, operational errors, or anomalous geological conditions, thereby cleaning the dataset and enhancing the accuracy of the integrated analyses of drilling dynamics and geological characteristics. This method ensures that the analysis models are not affected by extreme values, thereby improving the reliability and accuracy of ROP predictions.

3.2.3. Wavelet denoising

In the optimization of while-drilling data features, wavelet denoising was employed to minimize the impact of random noise on the accuracy of the empirical ROP equations (Ding et al., 2023). This technique involves decomposing the data signal into different frequency levels through wavelet transformation, allowing for independent analysis and processing of each frequency component. The before- and after-effects of wavelet denoising on the key LWD parameters (GR, SP, and DT) are shown in Fig. 4.

During this process, the Daubechies series wavelet DB6, which is widely applied because of its excellent performance in data analysis, was used for multilevel decomposition. The decomposed data were processed using a hard thresholding technique in which a fixed threshold was set to eliminate all wavelet coefficients below this threshold, thereby removing the noise components. This process not only retains key signals within the data, but also effectively reduces background noise. Subsequently, the wavelet reconstruction technique reassembles the threshold wavelet coefficients, restoring the original structure of the signal to a clearer and less noisy form. The hard-thresholding technique is particularly advantageous for handling sharp signal changes and preserving the abrupt characteristics of the signal, which is crucial for anomaly detection in while-drilling data (Ding et al., 2023). Wavelet denoising enhances the signal-to-noise ratio of the data features and improves the sensitivity and accuracy of the empirical ROP equations in response to complex drilling conditions.

3.3. Configuration and training of the MESR-DS model

In the construction of empirical ROP equations through symbolic regression analysis, a precise model configuration and optimization of the training process are crucial for achieving high accuracy and efficiency. The meticulous selection of model configuration parameters aims to enhance the predictive capabilities and operational efficiency. Core parameters, such as population size, expression complexity, and iteration count, were finely optimized through multiple trials to enhance the performance of the symbolic regression algorithm. Adjusting the population size directly boosts the parallel processing capability of the algorithm, allowing simultaneous exploration in multiple search spaces and effectively accelerating the optimization process. Moreover, implementing the maximum complexity and depth limits for expressions prevents model overfitting and ensures generalizability and predictive accuracy.

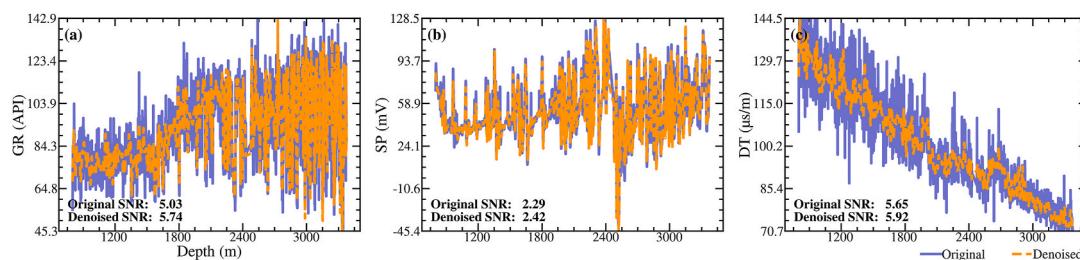


Fig. 4. Comparison of wavelet denoising effects on key LWD parameters: (a) Before and after denoising of GR parameter; (b) Before and after denoising of SP parameter; (c) Before and after denoising of DT parameter.

The training process incorporated a parallel computing framework and a simulated annealing strategy combined with crossover and mutation operations to systematically optimize the expression structure and parameters. The initialization phase involves randomly generating diverse expressions within each population, thereby providing a broad range of starting points for the model. Evolution cycles and control parameters guide the execution of the model in a parallel environment, where multithreading technology significantly enhances computational efficiency. The simulated annealing strategy dynamically adjusts the temperature parameter T to control the stringency of mutations, facilitating a gradual transition from broad exploration to a targeted local search. Comprehensive configuration and training strategies enabled this symbolic regression model to handle high-dimensional complex while-drilling data and accurately construct empirical ROP equations. The MESR-DS training hyperparameters are listed in Table 1.

Using the hyperparameters listed in Table 1, a score metric defined by Eq. (4) (Crammer, 2023):

$$\text{score} = -\log(\text{loss}_i / \text{loss}_{i-1}) / (\text{complexity}_i - \text{complexity}_{i-1}) \quad (4)$$

where loss_i represents the loss value under the current model configuration; loss_{i-1} indicates the loss value under the previous model configuration, with their ratio reflecting the evolution of model performance; $\text{complexity}_i - \text{complexity}_{i-1}$ denote the complexities of the current and previous model configurations, respectively, with their difference showing the magnitude of change in model structure. By calculating the negative logarithm of the ratio, the score metric aims to quantify the relative improvement in the model performance with a given increase in complexity.

Fig. 5 illustrates the iterative process of the loss values and evaluation scores for various complexities of the MESR-DS model.

Fig. 5 displays the training effects of the ROP empirical equations across a range of complexities, from simple to complex, with equations from complexities 1 to 20 showing a broad variation in loss values, clearly reflecting the trade-off between model complexity and prediction accuracy. Specifically, the equation at a complexity of 20 demonstrated optimal predictive performance with the lowest loss value (0.01).

3.4. Validation of the MESR-DS model

3.4.1. Performance evaluation of the empirical ROP equation

Following the training process described above, empirical ROP equations were developed for 20 complexities, with 200 sets trained for each level. In the case of the MESR-DS algorithm, hyper-parameters were determined by conducting 200 performance evaluations. Fig. 6 displays the calculated Pareto frontier during the non-dominated sorting process, which identifies the optimal trade-offs between model complexity and loss value. In Fig. 6, the loss values are plotted along the vertical axis, and the complexity along the horizontal axis shows the scatter of the 20 groups of empirical ROP equations. These points reveal the optimal balance points at which loss minimization and model complexity are achieved.

Specifically, the distribution of data points in Fig. 6 horizontally reflects an increase in complexity, and vertically reflects a gradual reduction in loss values. At lower complexities, higher loss values were observed; when the complexity reached a moderate level (within the range of 8–12), loss values significantly decreased, and data points quickly moved to the area of lower losses. However, once the complexity exceeded 15, the reduction in loss values slowed significantly, and the data points tended to stabilize at the lower part of the chart,

Table 1
Training strategies for the MESR-DS algorithm.

Hyperparameter	Description	Value
PopulationCount	Number of independent populations, each representing a parallel search path.	40
ExpressionsPerPopulation	Number of expressions per population to enhance search breadth.	1000
MutationCycles	Number of mutation cycles within the evolution period to improve convergence speed and solution quality.	300000
CrossoverProbability	Probability of crossover operation to introduce genetic diversity.	0.01
GlobalBestReplacementRate	Probability of replacing with global best solution to promote elite gene exchange among populations.	0.05
InterPopulationBestRate	Probability of replacing with the best solutions from other populations to enhance solution space exploration.	0.05
TournamentSize	Number of competitors in the tournament selection to balance selection pressure and diversity.	12
WinnerProbability	Probability of selecting the winner in the tournament to ensure the inheritance of high-quality solutions.	0.9
MutationIntensity	Constant mutation intensity to adjust the amplitude of value variations.	1
MinimalPerturbation	Minimum perturbation value to prevent excessive mutation that could destabilize the model.	0.1
AnnealingTemperature	Simulated annealing temperature coefficient to control the acceptance probability of solutions.	0.1
MaxComplexity	Maximum expression complexity to limit overfitting.	20
MaxExpressionDepth	Maximum depth of the expression tree to maintain model interpretability.	10
ConstantComplexityLimit	Complexity constraint for constant terms.	10
VariableComplexityWeight	Weight of variable complexity to optimize model interpretability.	0.0001
ParsimonyCoefficient	Parsimony coefficient to balance performance and simplicity of expressions.	0.001
BatchSize	Batch size to optimize memory usage and speed.	50
AdaptiveParsimonyScaling	Adaptive parsimony adjustment to dynamically adjust the parsimony ratio based on search dynamics.	10
ProcessingUnits	Number of computing units to enhance parallel computing capability.	16
CyclesPerUpdate	Number of evolution cycles per iteration update to optimize iterative effect.	500
Frecency	Metric that measures the frequency and recency of the complexity within the population.	0.5

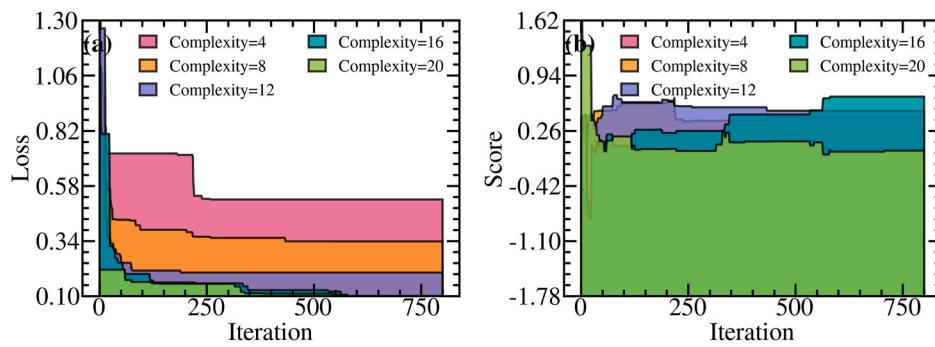


Fig. 5. Iterative dynamics under different complexities of the MESR-DS algorithm: (a) Loss values over iterations; (b) Evaluation scores over iterations.

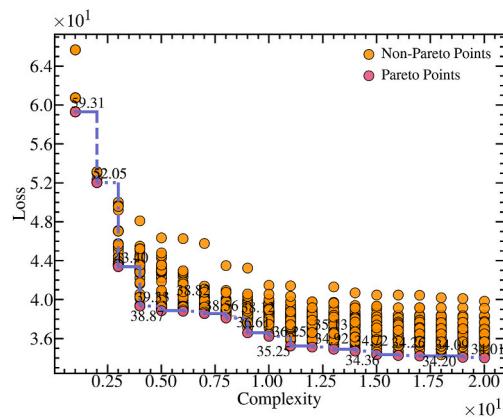


Fig. 6. Pareto frontier of the empirical ROP equation: optimal trade-off distribution between model complexity and loss value.

demonstrating the diminishing marginal returns of increasing complexity. Therefore, when selecting empirical ROP equations, emphasis should be placed on solutions along the Pareto frontier, which achieve an optimal balance between equation complexity and predictive accuracy. In particular, the ROP empirical equations with complexities between 15 and 20 maintained low loss values while avoiding excessive complexity, thus preserving computational efficiency. The selected 20 sets of empirical ROP equations are listed in Table B2 of Appendix B.

To comprehensively evaluate the performance of different complexities of the ROP empirical equations, validations were conducted at specific drilling locations in the Ledong 10-1 block of the South China Sea, focusing on complexities from 16 to 20. The results presented in Fig. 7 demonstrate the predictive performance of these empirical equations for varying complexities.

Fig. 7 illustrates the predictive performance of the empirical ROP

equations at specific drilling locations with complexities ranging from simple to complex. As the complexity of the ROP empirical equations increased, the MSE values gradually decreased from 116.51 to 29.85, clearly indicating that more complex empirical equations can predict ROP more accurately. This indicates that the predictive accuracy of the ROP empirical equations is directly proportional to their complexity. Equations with higher complexity are better at capturing subtle variations and complex relationships in drilling data, thereby providing more accurate predictions of the ROP.

To further evaluate and compare the performance of the selected optimized empirical ROP equations, a fully connected neural network (FCNN) model for ROP prediction and the classical Bourgoyne–Young ROP empirical equation were introduced (Borozdin et al., 2020; McCulloch and Pitts, 1943). The FCNN model's specific parameters were determined using a Bayesian optimization algorithm. The configurations of specific parameters are listed in Table 2.

To comprehensively evaluate and compare the various ROP prediction methods, the models and equations were thoroughly tested at the same training well site and adjacent locations. The results of these tests are illustrated in Fig. 8, with the aim of comparing the applicability and accuracy of the different prediction methods under similar geological conditions.

Fig. 8(a) shows the performance of the FCNN model and the two types of empirical ROP equations for the training well. The FCNN model demonstrated the lowest MSE value of 5.64, outperforming both the high-complexity empirical equation proposed in this study (complexity = 20) and the classical Bourgoyne–Young ROP empirical equation, which recorded MSE values of 29.85 and 136.74, respectively. These results highlight the advantages of the FCNN model in terms of accuracy and data fitting. However, the high-complexity empirical equation showed significant improvements in reducing prediction errors, particularly compared to the Bourgoyne–Young equation, whose performance lagged significantly, reflecting the potential inflexibility of traditional models under modern drilling conditions owing to fixed parameter

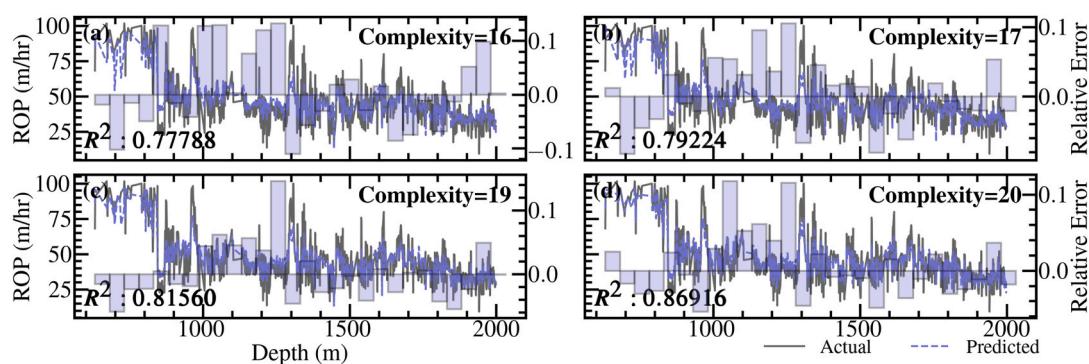


Fig. 7. Comparison of predictive performance of the empirical ROP equation under different complexities: (a) Complexity 16; (b) Complexity 17; (c) Complexity 19; (d) Complexity 20.

Table 2

Parameter configurations of the FCNN model and the Bourgoyne–Young empirical ROP equation.

Application	Hyperparameter	Description	Value
FCNN	<i>NetworkLayers</i>	Number of network layers.	4 layers
	<i>NeuronsPerLayer</i>	Number of neurons per layer.	[64, 128, 64, 1]
	<i>ActivationFunction</i>	Activation function used for neurons.	ReLU
	<i>LearningRate</i>	Learning rate for the optimizer.	0.001
	<i>BatchSize</i>	Number of samples per training batch.	32
	<i>TrainingEpochs</i>	Number of epochs for training on the entire dataset.	100
	<i>Optimizer</i>	Optimization algorithm used for training.	Adam
Bourgoyne–Young Empirical Equation	<i>LossFunction</i>	Loss function used for training.	MSE
	<i>a</i>	Coefficient for the relationship between ROP, WOB, and RPM.	0.1
	<i>b</i>	Coefficient for the relationship between ROP and well depth.	0.2
	<i>c</i>	Coefficient for the relationship between ROP and mud properties.	0.3
	<i>d</i>	Coefficient for the relationship between ROP and formation properties.	0.4

settings. The scatter plots of the R^2 fit in Fig. 8(b)–8(d) further confirm the superior fit of the FCNN model ($R^2 = 0.959$), while the high-complexity empirical equation, although less fitting ($R^2 = 0.781$), was significantly better than the Bourgoyne–Young ROP empirical equation (R^2 nearly zero). This indicates that, despite not reaching the level of the FCNN model, the high-complexity empirical equation has a clear advantage in capturing geological variability. By contrast, the traditional form and parameter settings of the Bourgoyne–Young equation limit its adaptability to new geological conditions. Fig. 8(e) compares the performance of the FCNN model and the high-complexity empirical equation in predicting ROPs at adjacent wells, showing close performance with MSE values of 106.48 and 116.94, respectively, far superior to the Bourgoyne–Young ROP empirical equation of 307.69. This demonstrates the high complexity of the empirical equation's similarity to modern machine-learning methods in terms of generalization. The R^2 fit scatter plots in Fig. 8(f)–8(h) show that, although the FCNN model has a slight advantage, the adaptability of the high-complexity empirical equation remains stronger than that of the Bourgoyne–Young ROP empirical equation, whose low R^2 value (0.022) further emphasizes its performance limitations in varying drilling environments.

3.4.2. Sensitivity analysis of the empirical ROP equation

To further deconstruct the empirical equations for high-complexity ROP, a quantitative global sensitivity analysis was conducted using the Sobol variance method. This method quantitatively assessed the impact of each input variable on the output of the ROP predictions, revealing the main and interaction effects of drilling, mud logging, and LWD parameters on the model. Through Sobol variance analysis, the

influence of individual parameters can be identified as well as the combined impact of multiple parameters on prediction outcomes is elucidated (Li et al., 2020b). Fig. 9(a) and (b) show the sensitivity analysis of the high-complexity ROP empirical equations using the quantitative Sobol variance method, assessing the overall sensitivity, first-order, and second-order sensitivities of different parameters on the ROP.

The quantitative Sobol variance analysis displayed in Fig. 9(a) reveals the overall sensitivity of the key drilling parameters, such as the RPM and TemIN, with values of 2.5242 and 0.8718, respectively. These figures underscore their critical role in predicting ROP. Adjustments in RPM directly affect the efficiency of the interaction between the drill bit and rock, whereas an appropriate mud temperature is crucial for ensuring effective cooling of the drill bit and transport of cuttings. Although the overall sensitivity index for depth is low (0.0163), its presence reflects the impact of increasing depth on the ROP, which is likely related to the increased hardness and complexity of the formations. The high overall sensitivity of the Mwin (0.9550) indicates its vital importance in controlling bottom-hole pressure and preventing wellbore collapse, which is directly related to the stability and safety of the ROP. The first-order sensitivity indices for the SP and DT were close to zero, indicating their minimal direct impact on the ROP; however, their interactions with RPM could lead to significant changes in the ROP. This suggests that in the drilling process, the impact of a single parameter might be obscured by variations in other geological or operational conditions; however, under multi-parameter interactions, even seemingly insignificant variables may significantly influence the ROP. Furthermore, the high sensitivity of TemIN to ROP is attributed to its

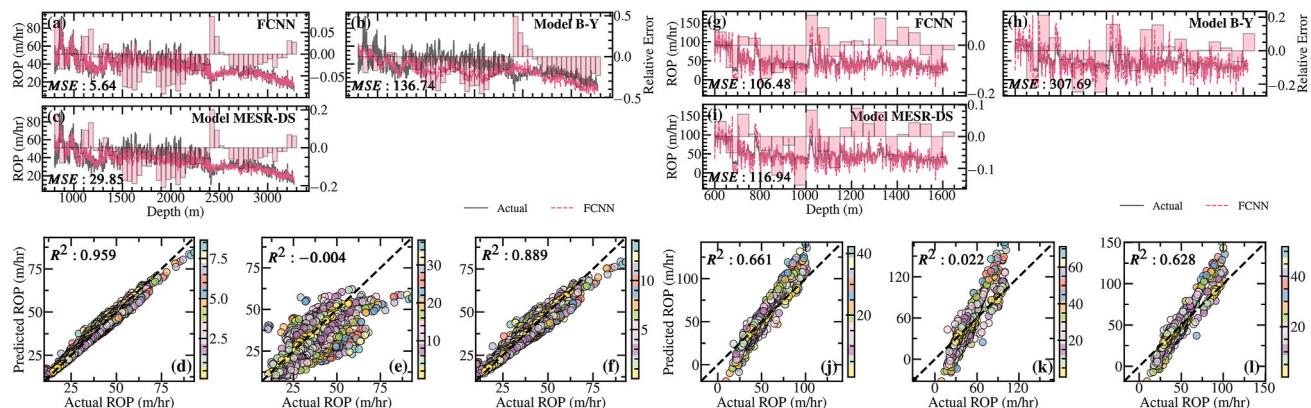


Fig. 8. Comparative analysis of the performance of different ROP prediction models on training wells and neighboring wells: (a) FCNN model's ROP prediction on training well (actual vs. predicted); (b) Bourgoyne–Young model's ROP prediction on training well; (c) MESR-DS model's ROP prediction on training well; (d) R^2 scatter plot of FCNN model on training well; (e) R^2 scatter plot of Bourgoyne–Young model on training well; (f) R^2 scatter plot of MESR-DS model on training well; (g) FCNN model's ROP prediction on neighboring well; (h) Bourgoyne–Young model's ROP prediction on neighboring well; (i) MESR-DS model's ROP prediction on neighboring well; (j) R^2 scatter plot of FCNN model on neighboring well; (k) R^2 scatter plot of Bourgoyne–Young model on neighboring well; (l) R^2 scatter plot of MESR-DS model on neighboring well.

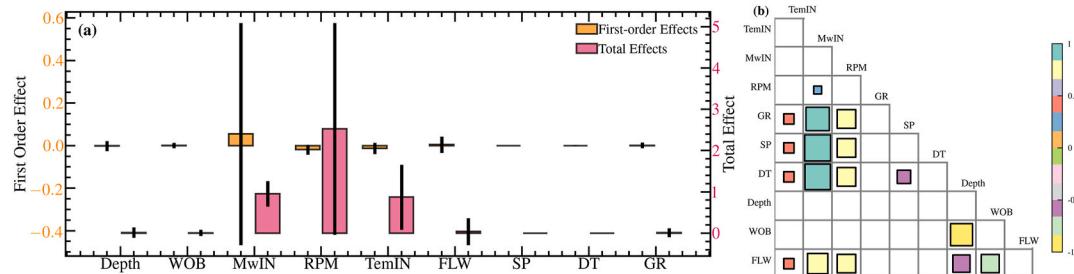


Fig. 9. Global sensitivity analysis and parameter interaction heatmap of the high-complexity empirical ROP equation: (a) Sobol variance-based global sensitivity indices for drilling parameters; (b) Heatmap of second-order parameter interactions among drilling parameters.

essential role in regulating drill bit cooling and cuttings transport. Elevated mud temperatures can diminish the drilling fluid's capacity to effectively cool the drill bit and transport cuttings, thereby reducing drilling efficiency and negatively impacting the ROP.

Fig. 9(b) shows the combinations of key parameters, especially those exhibiting significant second-order sensitivity indices. The combination of RPM and TemIN showed a notable positive second-order sensitivity (0.0180), indicating a significant joint impact on ROP predictions. Properly matching RPM with the mud temperature is crucial for maintaining the cooling efficiency and cutting carrying capacity. For instance, higher RPM should be complemented by lower mud temperatures to prevent overheating, thereby optimizing the drill bit performance and reducing wear. Additionally, the combination of the MwIN and FLW also showed a higher positive value (0.0190), indicating their important interactive role in regulating the drilling pressure and cleaning efficiency at the drill bottom. Proper relationships between mud density and flow are essential for drilling efficiency and safety, particularly for preventing wellbore collapse and loss of drilling fluid. Conversely, the combination of depth and WOB shows a negative second-order sensitivity (-0.0049), which might suggest that under certain conditions, variations in these parameters counteract each other, reducing their positive impact on the ROP. An appropriate configuration of the depth and WOB is crucial for optimizing the ROP, as excessive WOB in deep drilling may lead to premature wear of the drilling equipment or reduced drilling efficiency.

4. Drilling perception based on the VMD-CNN-BiLSTM-MA model

Conventional methods for optimizing drilling parameters have not adequately considered the hysteresis effects inherent in drilling operations and rely primarily on LWD parameters for decision-making. However, owing to the time lag in data acquisition and processing, drilling activities often proceed before the completion of data analysis, resulting in real-time drilling parameter adjustments that fail to achieve the intended effects. To address this issue, a method based on the drilling perception was proposed. This method involves predicting the LWD parameters in advance and optimizing the drilling parameters accordingly to adapt more effectively to dynamic changes during the drilling process, thereby enhancing the drilling efficiency and accuracy.

Table 3
Configuration parameters of the PSO algorithm optimized by variational mode decomposition.

Hyperparameter	Description	Value
PopulationSize	Balances search efficiency and computational resource consumption.	30
MaxIterations	Ensures thorough search by setting the maximum iteration count before stopping the algorithm.	50
ModeCount	Adapts to varying geological data characteristics by considering possible spectrum variations.	3–10
Regularization	Adjusts the regularization strength in variational mode decomposition to improve decomposition stability.	100–2000
InertiaWeight	Controls the inertia of the search.	0.5–0.9
PersonalBest	Degree to which particles adjust their position based on their historical best solution.	1.5
GlobalBest	Degree to which particles adjust their position based on the global best solution.	1.5
VelocityLimit	Prevents particles from escaping the optimal region by limiting their search speed.	Dynamic

4.1. Variational mode decomposition and data reconstruction

4.1.1. Variational mode decomposition of LWD parameters

In the processing of LWD parameters, the Variational Mode Decomposition (VMD) method was introduced to enhance the precision and efficiency of data analysis, which is particularly suitable for handling nonstationary data characteristics (Dragomiretskiy and Zosso, 2013; Yun and Jian, 2020). The VMD effectively isolates key frequency components from complex LWD data, thereby revealing potential information and features during drilling (Liu et al., 2016; Yan et al., 2020). The detailed mathematical formulation and implementation of the VMD method are provided in Appendix E.

To further optimize the performance of VMD, the Particle Swarm Optimization (PSO) algorithm was introduced, specifically for global optimization of the modal count K and the penalty parameter α (Kennedy and Eberhart, 1995). The PSO algorithm, using the minimization of local envelope entropy as a fitness function, effectively explores the solution space to find the optimal K and α , ensuring the best performance and minimal reconstruction error of the data decomposition (Xue et al., 2016).

The PSO algorithm settings for globally optimizing the modal count and penalty factor for the VMD of the LWD data are presented in Table 3.

The iterative dynamics of the envelope entropy, mode count, and penalty parameters for the GR, SP, and DT are shown in Fig. 10.

Taking GR as an example, the local envelope entropy reached its minimum value of 38.99 during the third iteration, indicating minimal data noise and the most significant features at this point. The penalty factor α stabilized after the fifth iteration at approximately 451.56, and the mode count K stabilized at 5 during the third iteration, confirming the optimal solution for K .

The LWD parameters for GR, SP, and DT were decomposed using the VMD method, each yielding five IMFs, with the decomposition results shown in Fig. 11(a)–11(e). Subsequently, these IMFs underwent Fast Fourier Transform (FFT) analysis, and the results are shown in Fig. 11(f)–11(j), confirming the uniform distribution of the frequency components in the time domain. This analysis validates the effectiveness of VMD decomposition and independence between modes, ensuring data integrity and analytical accuracy during processing (Duhamel and Vetterli, 1990).

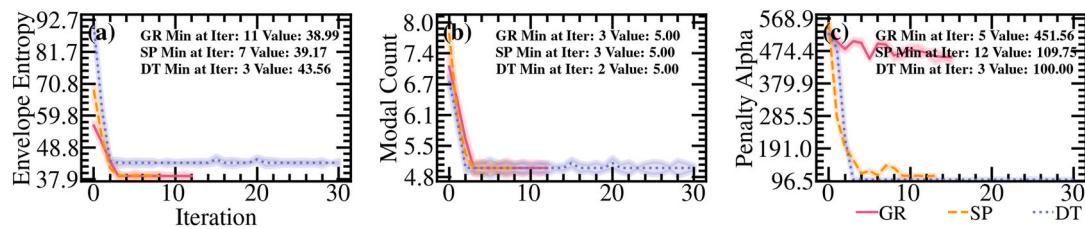


Fig. 10. Iterative dynamics of the envelope entropy, mode count, and penalty parameters for GR, SP, and DT data: (a) Envelope entropy over iterations; (b) Mode count over iterations; (c) Penalty parameters over iterations.

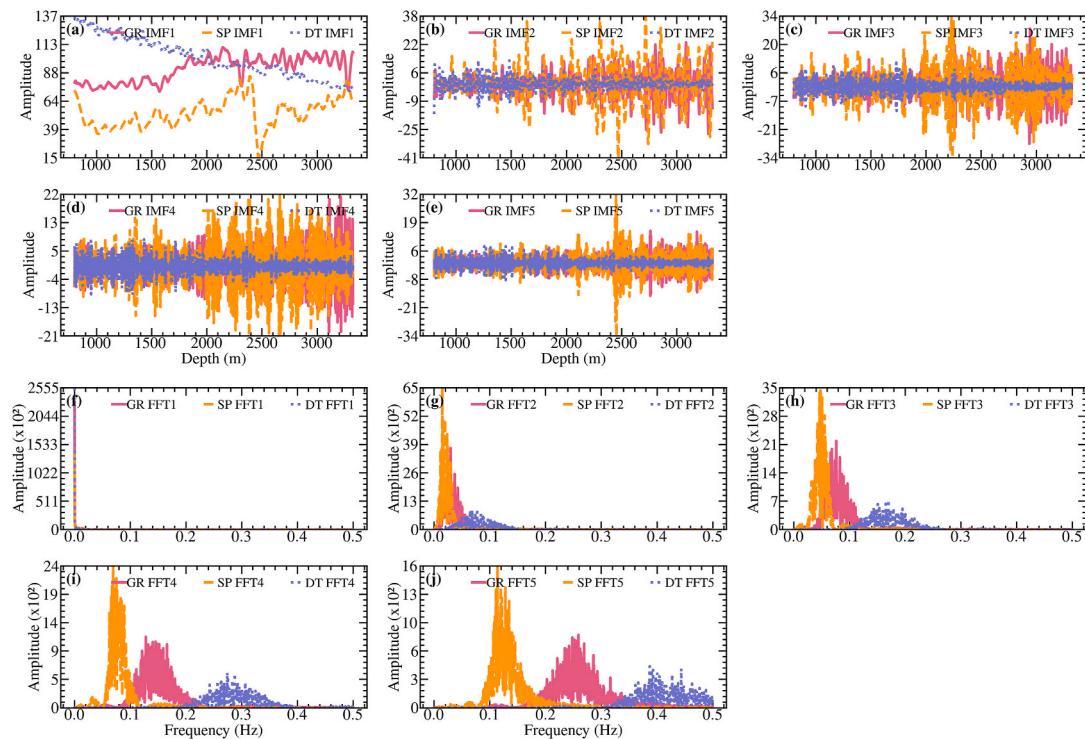


Fig. 11. IMF decomposition and frequency verification of GR, SP, and DT LWD parameters: (a) Intrinsic Mode Function (IMF) 1 obtained from VMD decomposition; (b) IMF 2; (c) IMF 3; (d) IMF 4; (e) IMF 5; (f) FFT analysis of IMF 1 confirming frequency distribution; (g) FFT analysis of IMF 2; (h) FFT analysis of IMF 3; (i) FFT analysis of IMF 4; (j) FFT analysis of IMF 5.

4.1.2. LWD data reconstruction based on sample entropy

Sample entropy (SampEn), developed by Richman from Pincus' approximate entropy (ApEn), is a method for assessing the complexity of time-series data (Richman et al., 2004). Compared to its predecessor, SampEn exhibits superior performance, particularly in the selection of dimensions (m) and thresholds (r), effectively measuring the frequency of repetitive patterns in a time series to assess their complexity (Delgado-Bonal and Marshak, 2019). An increase in sample entropy reflects greater dispersion and uncertainty in time-series information (Ebrahimi et al., 1994). In the analysis of the LWD parameters, sample entropy helps distinguish meaningful information from noise components. For data reconstruction, sample entropy evaluates each data component and sets thresholds to effectively exclude components with high randomness and low information content while retaining those with rich information and low irregularity. This strategy enhances the accuracy and efficiency of data processing. The mathematical expression for sample entropy is given by Eq. (5) (Govindan et al., 2007):

$$\text{SampEn}(m, r, N) = -\ln[B^{m+1}(r) / B^m(r)] \quad (5)$$

where N is the length of the time series, $B^m(r)$ represents the probability of matching two time series at m points under threshold r , and after

reconstructing the sequence into an $m + 1$ dimensional vector, $B^{m+1}(r)$ represents the matching probability at $m + 1$ points.

The calculation of SampEn depends on two critical parameters: the embedding dimension m and similarity threshold r . In practice, m is typically set to 1 or 2 and r is chosen to be between 10% and 25% of the standard deviation of the data. Using embedding dimensions m of 1 and 2 and similarity thresholds r of 0.1 and 0.2 (Zurek et al., 2012), the SampEn values of the IMFs for GR, SP, and DT were obtained, as illustrated in Fig. 12.

It is evident from Fig. 12 that the largest variations in the SampEn values of the IMFs occurred at $m = 2$ and $r = 0.1$; hence, these parameters were set. Following the SampEn analysis, the individual IMFs of the GR, SP, and DT were further processed to identify and reconstruct them into trend, periodic, and residual components. This classification, based on the relative SampEn values of each IMF, enabled the precise capture of key dynamic features in the drilling data.

Specifically, the trend component is typically composed of a weighted average of the IMFs with the lowest SampEn values, reflecting the data's long-term trends or smooth changes. The periodic component is represented by IMFs with evident periodicity but not by the lowest SampEn, which captures periodic fluctuations. The residual component consists of the remaining high-entropy IMFs, which typically represent

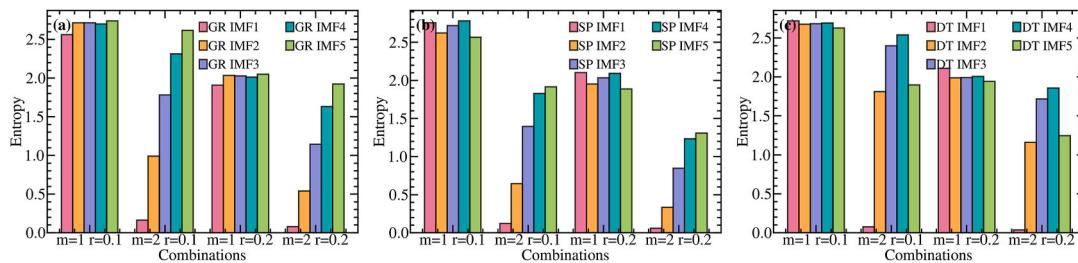


Fig. 12. Analysis results of IMF sample entropy under different embedding dimensions and similarity thresholds: (a) SampEn values of IMFs for GR parameter; (b) SampEn values of IMFs for SP parameter; (c) SampEn values of IMFs for DT parameter.

noise or irregular fluctuations (Harvey and Koopman, 2000). The reconstruction results for GR, SP, and DT are illustrated in Fig. 13.

By integrating the VMD method with SampEn analysis, it was possible to efficiently extract key and information-rich data components from the LWD data, thereby facilitating further data-driven decision-making.

4.2. Construction of the CNN-BiLSTM-MA model

The CNN-BiLSTM-MA model integrates neural network architectures to enhance predictive performance for LWD parameters. Specifically, it combines a One-Dimensional Convolutional Neural Network (1D-CNN) for feature extraction, a Bidirectional Long Short-Term Memory (BiLSTM) network for capturing temporal dependencies, and a Multi-head Attention (MA) mechanism for focusing on relevant information. Detailed descriptions of these components are provided in Appendix E.

In the predictive model for LWD parameters, accurately identifying parameters, such as GR, SP, and DT, is crucial. Fig. 15 demonstrates that the MA mechanism, through its independent operation of multiple attention heads in parallel, precisely captured the physical and chemical property changes in the formations. Each attention head specializes in processing specific types of data and identifying key information from vast and complex geological data, thereby providing decisive value for the interpretation and analysis of geological layers. In the multi-head self-attention mechanism, each head is trained on specific geological data variations, allowing the model to instantly reflect geological characteristics and predict their trends. This functionality greatly enhanced the responsiveness and accuracy of the model to changes in formation characteristics, providing important data support for drilling decisions.

Similarly, Fig. 16 demonstrates that in the predictive model for adjusting drilling parameters such as WOB, RPM, FLW, MWIN, and TemIN, the MA mechanism also plays a crucial role. Through its independently operating attention heads, this mechanism tracked the key dynamics of drilling operations in real time, such as changes in WOB or adjustments in RPM, ensuring that the model could predict changes in operational conditions instantly and accurately. The focused approach of each attention head improves the responsiveness and accuracy of key events during the drilling process, thus optimizing the efficiency of parameter adjustments and enhancing decision-making support for drilling operations. The network architecture of the CNN-BiLSTM-MA model is illustrated in Fig. 14.

4.3. Model establishment and training

4.3.1. Numerical prediction of LWD parameters

The primary objective of the model was to enhance the prediction accuracy of LWD parameters, including GR, SP, and DT. These parameters are crucial for the real-time geological assessment and optimization of drilling decisions. Prior to training, the model underwent a series of pre-processing steps on the LWD data, including VMD and data reconstruction based on SampEn, as previously described. These methods helped purify the data, reduce interference from irrelevant noise, and ensure that the dataset used for training met high standards of quality and relevance.

The LWD data were divided by depth into training, validation, and test sets at a ratio of 8:1:1. For the training strategy, an adaptive learning rate adjustment was implemented using a learning rate annealing method that dynamically adjusts the learning rate based on loss changes, enabling the model to converge quickly in the early stages of training and approach a global optimum through fine-tuning in later stages (Zeiler, 2012). Additionally, the AdamW optimizer, which combines adaptive moment estimation with weight decay adjustments, has been employed to enhance the training efficiency and generalization capability of the model on complex datasets (Loshchilov and Hutter, 2017). The mean squared error (MSE) loss function was selected to provide a direct quantification of the deviation between the predicted and actual values, which is crucial for optimizing the model's output accuracy (Willmott and Matsura, 2005). To reduce overfitting and verify the generalization ability of the model, dropout layers and regularization strategies were added to prevent overfitting while maintaining effective training (Srivastava et al., 2014). Moreover, training utilizes the early stopping method, dynamically adjusting training epochs based on performance on the validation set, effectively preventing overtraining and ensuring that the model state is saved promptly upon reaching optimal performance (Prechelt, 2002). These comprehensive training strategies and methods ensured the high accuracy and reliability of the model in predicting LWD parameters. The training strategy parameters are presented in Table 4.

The training and validation loss values for the models, as shown in Fig. 15(a), vary over the iterations. The training and validation loss values for the models, as shown in Fig. 15(a), vary over the iterations.

Fig. 15(a) shows the loss values of the four models (BiLSTM, CNN-BiLSTM, CNN-BiLSTM-MA, and VMD-CNN-BiLSTM-MA) across the iterations on the training set. Among these, the VMD-CNN-BiLSTM-MA model exhibited the lowest loss values and converged rapidly,

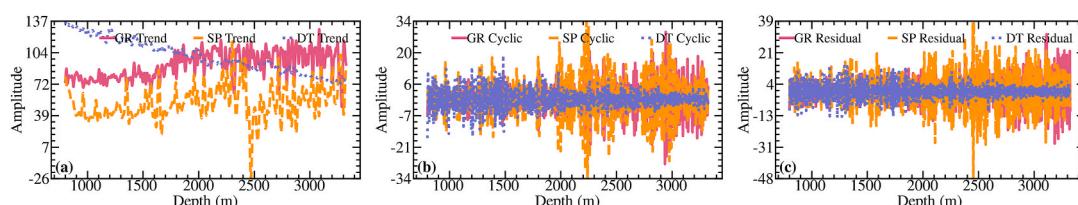


Fig. 13. Reconstruction results of trends, cycles, and residual components: (a) GR parameter; (b) SP parameter; (c) DT parameter.

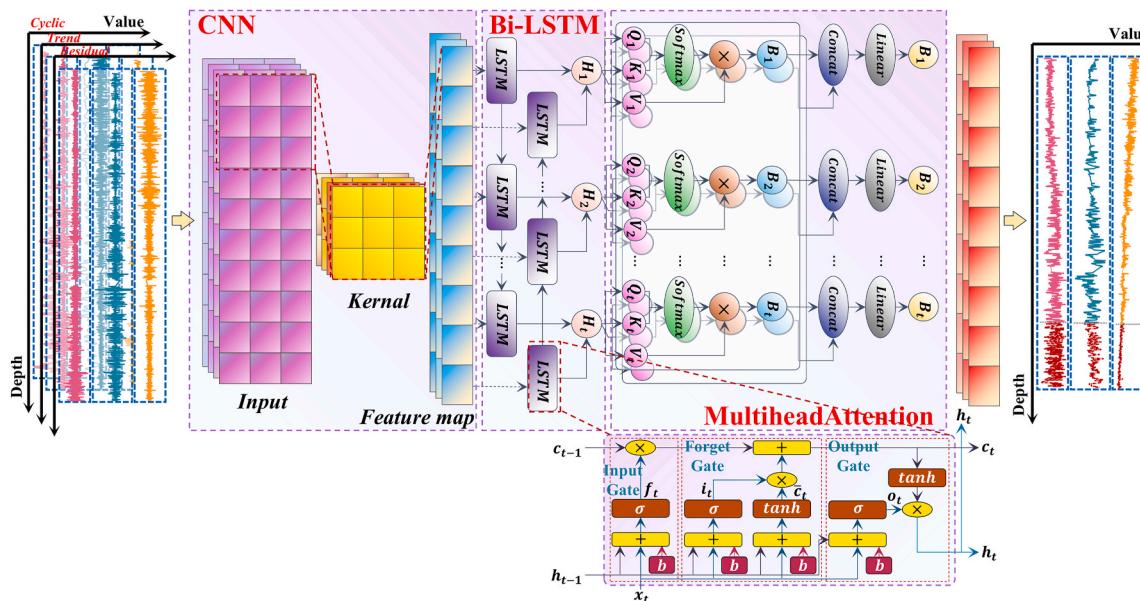


Fig. 14. Architecture of the CNN-BiLSTM-MA algorithm.

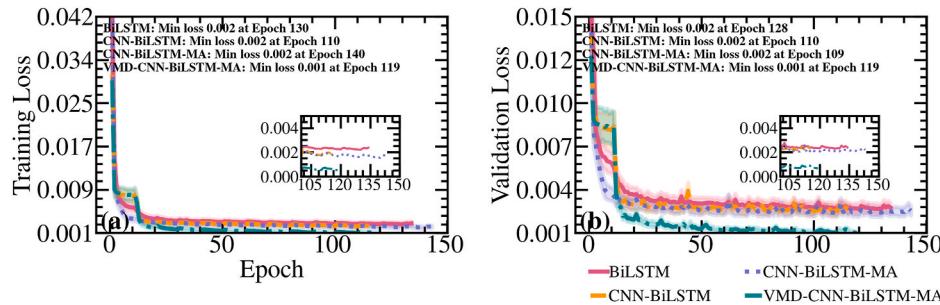


Fig. 15. Training and validation loss dynamics of the LWD parameter prediction model: (a) Training loss values over iterations; (b) Validation loss values over iterations.

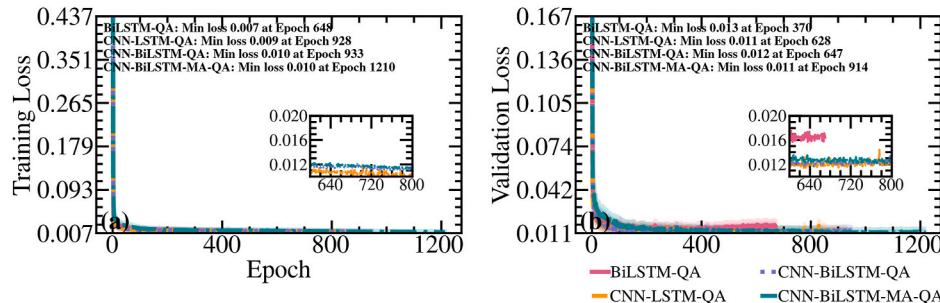


Fig. 16. Training and validation loss dynamics of the drilling parameter adjustment range prediction model: (a) Training loss values over iterations; (b) Validation loss values over iterations.

reflecting its superior data processing and feature extraction capabilities, which effectively reduced the training errors. Conversely, despite higher loss values, the BiLSTM model demonstrated rapid convergence, indicating its effectiveness in processing time-series data. The CNN-BiLSTM and CNN-BiLSTM-MA models exhibited closely aligned loss values in the final stages, indicating that these architectures tended to stabilize after extended training periods.

Fig. 15(b) shows the changes in the loss values for the four models in the validation set. Similarly, the VMD-CNN-BiLSTM-MA model exhibited the smallest loss values, thereby validating its performance on the training set and reflecting its strong generalization capability. Although

the BiLSTM model recorded higher loss values on the training set, it exhibited greater stability on the validation set than CNN-BiLSTM and CNN-BiLSTM-MA, likely because of its simpler structure, which reduced the overfitting risk. In contrast, while the performances of CNN-BiLSTM and CNN-BiLSTM-MA on the validation set were stable, they showed larger fluctuations, suggesting that the adaptability of the models to new data at different iterations could be improved.

4.3.2. Prediction of drilling parameter adjustment range

To precisely adjust the key drilling parameters, including the WOB, RPM, FLW, MwIN, and TemIN, quantile regression was employed to

Table 4

Configuration parameters for training the LWD parameter prediction model.

Hyperparameter	Description	Value
LearningRate	Learning rate for CNN-BiLSTM-MultiheadAttention model.	0.0001
OptimizationAlgorithm	Algorithm used for model optimization.	AdamW
BatchSize	Number of samples in each training batch.	32
TrainingEpochs	Number of iterations over the entire dataset for training.	100
WeightDecay	Regularization term to reduce model overfitting.	0.01
DropoutRate	Dropout rate applied in CNN-BiLSTM-MultiheadAttention to reduce overfitting.	0.5
EarlyStoppingPatience	Patience parameter for early stopping mechanism to prevent overfitting.	10
LearningRateScheduler	Scheduler for adjusting the learning rate.	StepLR
SchedulerStepSize	Step size for learning rate adjustment.	10
LearningRateDecayFactor	Decay factor for learning rate adjustment.	0.1
TimeWindows	Number of time steps considered in the input sequence.	100

predict the potential adjustment ranges of these parameters. Quantile regression differs from traditional mean regression in that it does not model the conditional mean, but instead models specific quantiles of the

conditional distribution, offering a comprehensive view of the potential outcome distribution. Quantile regression introduces a Quantile Loss function that enables the model to adapt to various quantiles of data

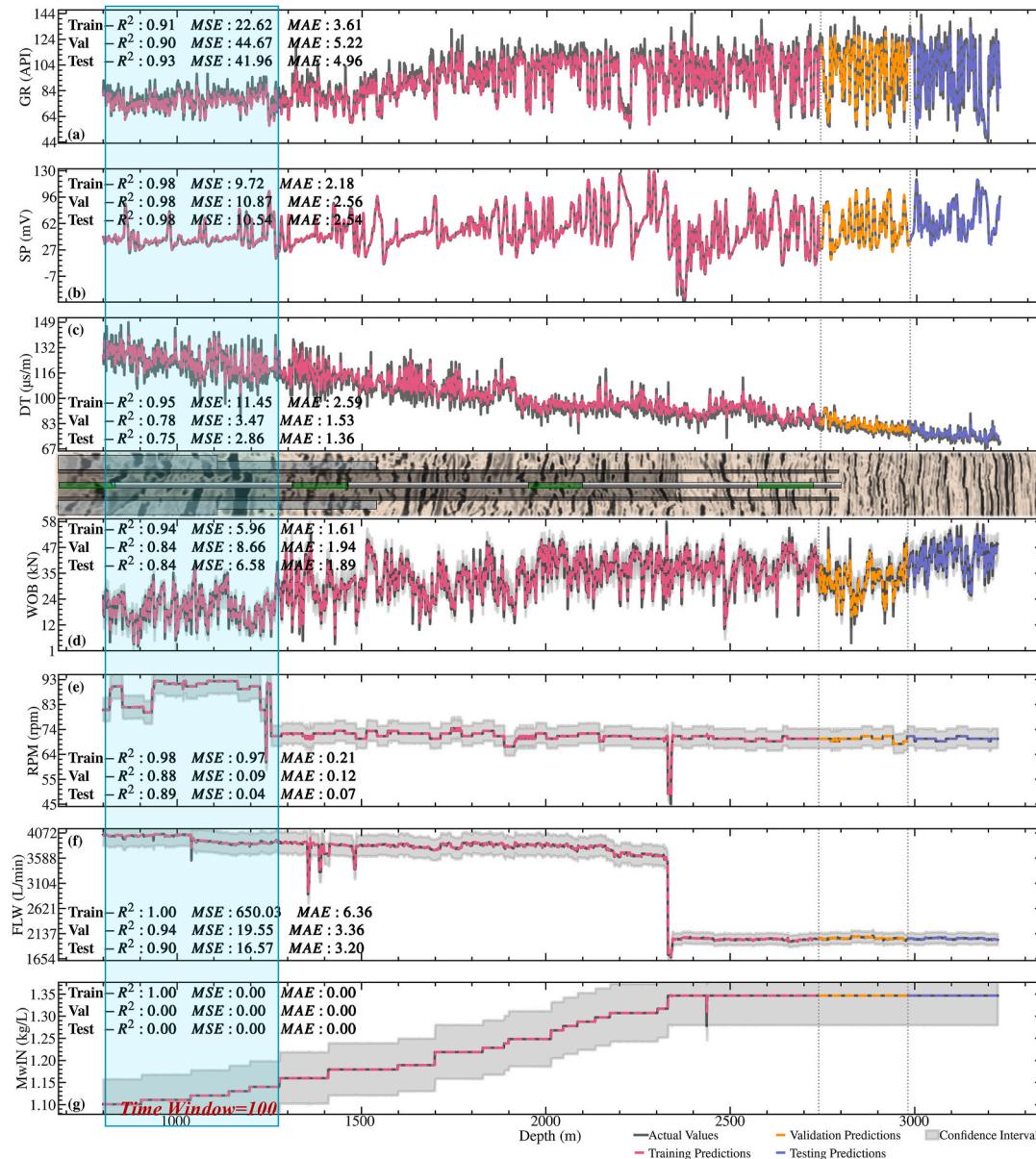


Fig. 17. Performance of the VMD-CNN-BiLSTM-MA and CNN-BiLSTM-MA-QA models in predicting LWD and drilling parameters: (a) Prediction results for GR parameter; (b) Prediction results for SP parameter; (c) Prediction results for DT parameter; (d) Prediction results for WOB parameter; (e) Prediction results for RPM parameter; (f) Prediction results for FLW parameter; (g) Prediction results for MwIN parameter.

distribution. The mathematical expressions for the Quantile Loss function is given by Eq. (6) (Buchinsky, 1998):

$$L(y, \hat{y}) = q \cdot \max(y - \hat{y}, 0) + (1 - q) \cdot \max(\hat{y} - y, 0) \quad (6)$$

where y represents the actual observations, \hat{y} represents the predictions, and q is a quantile parameter controlling the quantile position on which the model focuses. When $q = 0.5$, the Quantile Loss is equivalent to the median absolute error (MAE), causing the model to focus on fitting the data median. This allows the model to provide robust predictions in the presence of outliers or skewed distributions (Regression, 2017).

To evaluate the potential adjustment space for drilling parameters, quantile values q were set at 0.05 and 0.95, allowing predictions under varied risk adaptabilities. By mapping the operational conditions to risk preferences, this strategy is well suited to dynamic and complex drilling environments. Fig. 17 confirms that this method covers most potential operational situations in the model's output prediction intervals, providing crucial decision support and thereby optimizing the responsiveness and adaptability of drilling parameter adjustments during drilling operations. The architecture and training strategy used for the prediction of the drilling parameter adjustment ranges were consistent with those employed for the numerical prediction of the LWD parameters. The training and validation process loss values over the iterations are shown in Fig. 16.

Fig. 16 illustrates the dynamics of the training and validation losses for the drilling parameter adjustment range prediction model. Fig. 16(a) shows the loss changes in the four models (BiLSTM-QA, CNN-LSTM-QA, CNN-BiLSTM-QA, and CNN-BiLSTM-MA-QA) during training. In the training set, although the CNN-BiLSTM-MA-QA model initially converged slowly, it ultimately displayed the lowest loss values, demonstrating its superior learning capabilities and performance for complex datasets. Although converging faster, the other three models did not perform as well as the CNN-BiLSTM-MA-QA model during prolonged training. This highlighted the significance of the model structure and training strategy for handling highly complex drilling data. Fig. 16(b) shows similar trends in the validation set, where the CNN-BiLSTM-

MA-QA model, despite being the slowest to converge, exhibited the best generalization ability with the lowest and most stable loss values. Conversely, the BiLSTM-QA model performed the worst on the validation set, with the highest and most volatile loss values, likely owing to overfitting issues arising from its simpler structure. The CNN-LSTM-QA and CNN-BiLSTM-QA models showed similar capabilities in generalizing to new datasets, demonstrating that more complex neural network structures can provide more precise and stable performance for high-variability tasks, such as drilling parameter prediction.

4.4. Model performance evaluation

At a drilling site in the Ledong 10-1 block of the South China Sea, the VMD-CNN-BiLSTM-MA and CNN-BiLSTM-MA-QA models were evaluated for their effectiveness in predicting key LWD parameters such as GR, SP, and DT, along with drilling parameters such as WOB, RPM, FLW, MwIN, and TemIN. The prediction results, illustrated in Fig. 17, offer an intuitive comparison of the performances of the two models, highlighting their capabilities in a real-world drilling context.

Fig. 17(a)–17(c) show the VMD-CNN-BiLSTM-MA model in predicting three key LWD parameters: GR, SP, and DT. Each chart explicitly demonstrates the model performance across the training, validation, and test datasets evaluated using R^2 , mean squared error (MSE), and mean absolute error (MAE). For GR prediction, the model achieved an R^2 value of approximately 0.90 across all datasets, indicating high accuracy and consistency. Although the MSE and MAE were slightly higher in the validation and test sets, indicating a minor shortfall in the generalization of the model to unseen data, the overall performance remained robust. In predicting the SP, the model's performance was equally impressive across all datasets, with R^2 values reaching or approaching 0.98, reflecting its excellent fit and high sensitivity to changes in the SP. The high R^2 values combined with the relatively low error metrics highlight the superior predictive precision of the model for this logging parameter. However, the prediction of DT presented some challenges, particularly with a significant drop in R^2 values to approximately 0.75 in the validation and test sets, indicating a decrease in performance when handling

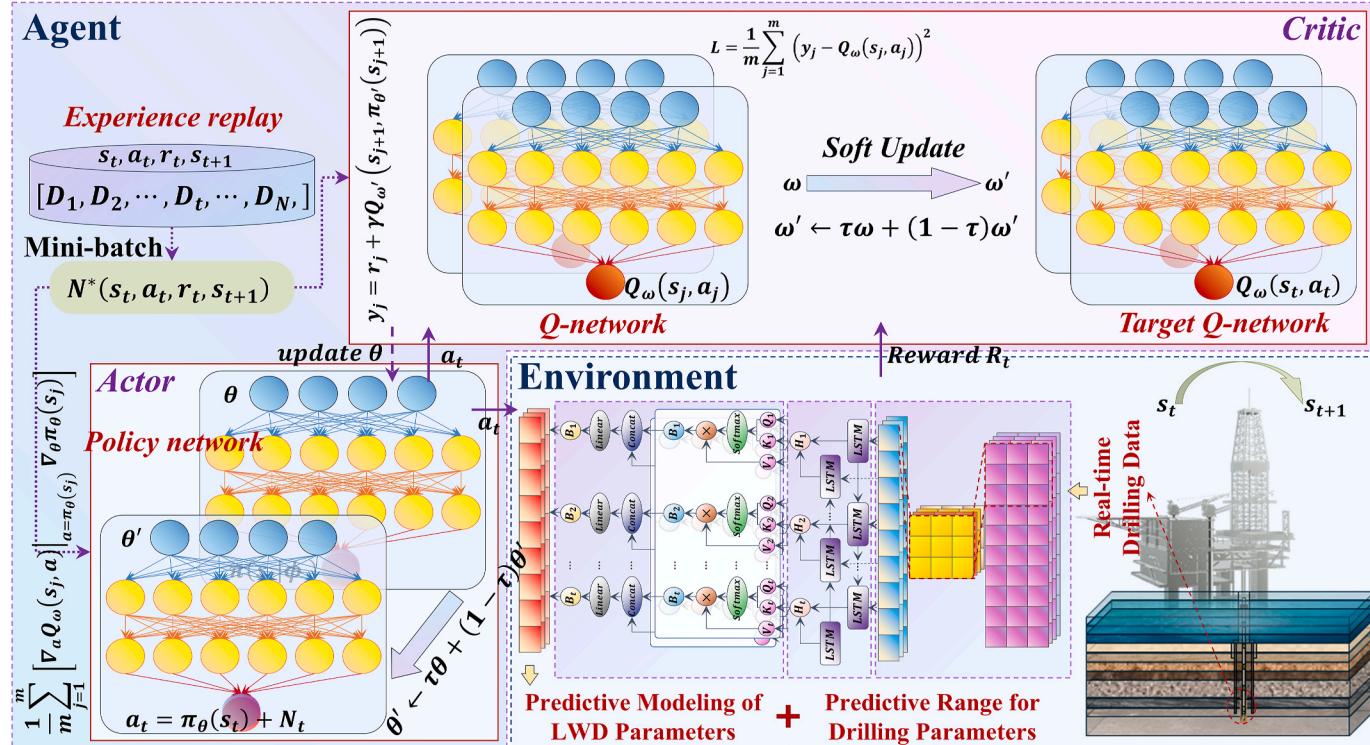


Fig. 18. Architecture of the DPODDPG algorithm.

the complexity and potential nonlinear relationships of DT data. This decline in performance may be attributed to noise or highly variable geological conditions within the data, necessitating further analysis and optimization of the model to enhance its robustness.

Fig. 17(d)–17(g) detailed the CNN-BiLSTM-MA-QA model's process in predicting drilling parameters, including WOB, RPM, FLW, MwIN, and TemIN. For WOB, the model achieved an R^2 value of 0.95 in the training set and 0.85 in the validation and test sets, indicating high accuracy in predicting changes in WOB. The prediction intervals generally encompassed the actual measurements, though they did not fully cover some high-variance outliers, likely due to the complexity of geological conditions and the inherent instability of WOB. The model's prediction, especially its handling of abrupt changes near 2000 m in the training set, demonstrating adaptability and responsiveness. Although the R^2 slightly decreased in the validation and test sets, it still showed high prediction accuracy. For MwIN and TemIN, the model exhibited perfect performance in the training set, although the R^2 values plummeted to 0 in the validation and test sets. This likely resulted from minimal variation or no change in the mud input parameters within these datasets, rendering the model ineffective at learning and predicting these static values. Considering that the calculation of R^2 depends on the variance ratio of predicted to actual values, the invariance of the real data (i.e., zero variance) directly limited the model's predictive capability under these conditions, resulting in a significant drop in R^2 values.

5. Real-time optimization of drilling parameters using multi-objective deep reinforcement learning

To optimize the drilling parameters effectively, an empirical ROP equation based on the MESR-DS algorithm and a drilling perception model based on VMD-CNN-BiLSTM-MA were proposed. The empirical ROP equation serves as a speed estimator and the drilling perception model serves as an environmental estimator, establishing a Markov Decision Process (MDP).

Table 5
Optimization process of drilling parameters using the DPODDPG algorithm.

Algorithm: DPODDPG for Optimizing Drilling Parameters

Input: T = Total number of episodes
 Parameters:
 θ = Initial actor network parameters
 ω = Initial critic network parameters
 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 θ and ω for the actor and critic networks
 Function DPODDPG
 1 .Initialize actor network π_θ and critic network Q_ω .
 2 .Initialize target networks: $\pi_{\theta'} \leftarrow \pi_\theta$, $Q_{\omega'} \leftarrow Q_\omega$.
 3 .for $i = 1$ to T do:
 4 .Observe initial state s_i .
 5 .while s_i is not terminal do:
 6 .Select action $a_t = \pi_\theta(s_t) + N_t$ using current policy and exploration noise.
 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 $y_j = r_j + \gamma Q_{\omega'}(s_{j+1}, \pi_\theta(s_{j+1}))$ if s_{j+1} is not terminal, else $y_j = r_j$.
 11 .Update critic by minimizing loss: $L = 1/m \sum_{j=1}^m (y_j - Q_\omega(s_j, a_j))^2$.
 12 .Update actor using policy gradient: $\nabla_\theta J \approx 1/m \sum_{j=1}^m \nabla_a Q_\omega(s_j, a) \Big|_{a=\pi_\theta(s_j)} \nabla_\theta \pi_\theta(s_j)$.
 13 .Soft update target networks: .
 14 . $\theta' \leftarrow \tau \theta + (1 - \tau) \theta'$.
 15 . $\omega' \leftarrow \tau \omega + (1 - \tau) \omega'$.
 16 . $s_t \leftarrow s_{t+1}$.
 17 .end while.
 18 .end for.
 19 .Return optimized network parameters θ , ω .

Table 6

Coefficients of drilling cost and mechanical specific energy functions in the Ledong 10-1 block, South China Sea.

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_t	Time for tripping and connecting single joints.	102
	A_f	Effect of formation on bit wear.	0.00228
	a_1	Linear coefficient for the impact of RPM on drilling efficiency.	1.5
	a_2	Cubic coefficient for the impact of RPM on drilling efficiency.	6.53E-05
	C_1	Coefficient for reducing bit tooth wear rate.	5
	E_f	Energy conversion efficiency of the bit during rock breaking.	0.35
	μ	Sliding friction coefficient when the bit contacts the rock.	0.5
Mechanical Specific Energy			

5.1. Design of multi-objective optimization functions

5.1.1. Unit cost per foot function

In drilling operations, the cost per foot function is a core indicator for assessing economic efficiency, focusing on accurately measuring costs and efficiency during drilling. The expression for cost per foot is given in Eq. (7) (Wang et al., 2015):

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

where C_{pm} is the cost per foot, (CNY/m); C_b is the cost of the drill bit, (CNY); C_r is the rig operation fee, (CNY/hr); t_t is the time required to make trips and connections, (hr); A_f is the abrasiveness coefficient of the formation; n is the RPM, (r/min); a_1, a_2 are the speed influence coefficients determined by the type of drill bit; h_f denotes the amount of wear corresponding to the drill-bit lifespan; $v_{pe}(w, n, q, k_d)$ is the ROP empirical equations developed through the MESR-DS algorithm. The detailed derivation of Eq. (7) is provided in Appendix E.

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

5.1.2. MSE function

The MSE model correlates the energy required to fracture a unit volume of rock with the rock-breaking efficiency of the drill bit and serves as a critical metric for quantifying drilling efficiency (Teale, 1965). The MSE is calculated using parameters such as the ROP, WOB, PRM, torque, and drill bit diameter. A higher MSE value indicates a lower drilling efficiency and poorer adaptability of the drill bit to the formation, suggesting that the drilling parameters require optimization. The ideal MSE calculation is given by Eq. (8) (Dupriest and Koederitz, 2005):

$$MSE = E_f \left(\frac{4W}{\pi D_b^2} + 0.16 \frac{n\mu W}{D_b v_{pc}(w, n, q, k_d)} \right) \quad (8)$$

where MSE is the mechanical specific energy, (MPa); W is the WOB, (kN); D_b is the drill bit diameter, (m); n is the RPM, (r/min); μ is the sliding friction coefficient of the drill bit; $v_{pc}(w, n, q, k_d)$ is the ROP empirical equations developed through the MESR-DS algorithm. The

detailed derivation of Eq. (8) is provided in Appendix E.

5.1.3. Objective functions

Building on single-objective optimization, a multi-objective optimization model was developed that considered both drilling costs and rock-breaking energy consumption. This model utilizes the empirical ROP equation to optimally select drilling parameters such as WOB, RPM, FLW, MwIN, and TemIN to minimize the unit cost per foot and MSE. The objective function, designed to maximize the drilling efficiency while minimizing costs, incorporates the lowest unit cost per foot and minimum MSE as the targets of multi-objective optimization. The model for the objective function, denoted by F , is given by Eq. (9):

$$\left\{ \begin{array}{l} F = \min(\min(C_{pm}), \min(MSE_m)) \\ \min(C_{pm}) = \frac{C_r \left[t_E \cdot A_f (a_1 n + a_2 n^3) + \left(h_f + \frac{C_1}{2} h_f^2 \right) \right]}{v_{pc}(w, n, q, k_d) \cdot \left(h_f + \frac{C_1}{2} h_f^2 \right)} \\ \min(MSE_m) = E_f \left(\frac{4W}{\pi D_b^2} + 0.16 \frac{n\mu W}{D_b v_{pc}(w, n, q, k_d)} \right) \end{array} \right. \quad (9)$$

5.2. MDP environment

A specific MDP environment was constructed to accurately control the operational parameters during the drilling process, a specific MDP environment was constructed. Given the deterministic physical characteristics of the drilling process, the state transition probabilities in the environment were set as constants.

5.2.1. Design of reward functions

In a deep reinforcement learning framework, the reward function plays a crucial role in providing immediate feedback to guide decision optimization (Arulkumaran et al., 2017). It was specifically designed to evaluate the effects of drilling parameter adjustments, including rewards for costs per foot and MSE. The cost per foot reward reflects the economic efficiency of drilling costs relative to the drilling process, whereas the MSE reward focuses on the energy efficiency and mechanical load during the drilling process. These rewards enable the agent to determine the optimal balance in adjusting the drilling parameters, thus maximizing the cost and equipment efficiency.

Cost per foot reward: This reward evaluates the economic benefits of drilling parameter adjustments based on changes in the cost per foot. This metric reflects the relationship between drill bit costs, drilling operation expenses, and drilling progress. A positive reward is given if the operation leads to a reduction in the cost per foot; conversely, if costs increase, the reward is negative. This design allows the reward function to directly reflect changes in economic efficiency and maximize cost effectiveness.

MSE reward: This reward evaluates the energy efficiency and mechanical load of drilling operations through the MSE, that is, the energy required to break a unit volume of rock. Reducing the MSE indicates improved drilling efficiency and reduced equipment wear risk; hence, operations that reduce the MSE receive positive rewards, whereas those that increase the MSE receive negative rewards. This reward design emphasizes the importance of energy efficiency, promoting the rational use of resources and the long-term stability of equipment. The reward function model is expressed by Eq. (10):

$$\left\{ \begin{array}{l} R_1 = -(C_{pm} - C_{pm,prev}) \\ R_2 = -(MSE_m - MSE_{m,prev}) \end{array} \right. \quad (10)$$

where R_1 represents the reward for improving the cost per foot, illustrating cost changes between iterations; C_{pm} denotes the cost per foot for the current iteration; $C_{pm,prev}$ for the previous iteration; R_2 signifies the reward for enhancing MSE, evaluating MSE modifications from one iteration to the next; MSE_m is the MSE for the current iteration; $MSE_{m,prev}$ for the previous iteration.

These two rewards combine to form a multi-objective optimization framework, allowing deep reinforcement learning algorithms to find the optimal balance between cost and efficiency and optimize the drilling parameters. Depending on the actual drilling needs and strategic goals, these rewards are appropriately weighted to reflect the priorities of different drilling objectives, ensuring the optimization of drilling operations across multiple dimensions.

5.2.2. State space and action space

In the deep reinforcement learning framework tailored for optimizing drilling operations, the state space is defined by critical drilling parameters such as WOB, RPM, FLW, MwIN, and TemIN. Real-time measurements of these parameters establish the current state of the drilling process and provide essential data for monitoring and adjustment purposes. The action space delineates the permissible range for adjusting the drilling parameters, constituting the domain in which the intelligent agent manipulates the drilling performance (Mnih et al., 2015). The agent directly influences the drilling state by executing specific actions. The outcomes resulting from these actions and the consequent state changes are communicated back to the agent via a meticulously defined reward function. The state and action spaces are expressed by Eq. (11) and Eq. (12):

$$\left\{ \begin{array}{l} S = \{w, r, f, m, t\} \\ w \in [w_{min}, w_{max}] \\ r \in [r_{min}, r_{max}] \\ f \in [f_{min}, f_{max}] \\ m \in [m_{min}, m_{max}] \\ t \in [t_{min}, t_{max}] \end{array} \right. \quad (11)$$

Table 7
Learning strategies of the DPODDPG algorithm.

Hyperparameter	Description	Value
<i>PolicyLearningRate</i>	Learning rate for the actor network, used to update the parameters of the policy network.	0.000010
<i>Q1LearningRate</i>	Learning rate for the Q1-network, used to update the parameters of each Q-network.	–
<i>Q2LearningRate</i>	Learning rate for the Q2-network, used to update the parameters of each Q-network.	–
<i>DiscountFactor</i>	Discount factor, used to balance the weight of current rewards and future rewards.	0.7
<i>ObjectiveWeights</i>	Weights of different optimization objectives in the loss function, used to balance the importance among objectives.	0.5,0.5
<i>BatchSize</i>	Number of batch samples drawn from the experience replay buffer.	2048
<i>BufferCapacity</i>	Maximum capacity of the experience replay buffer for storing previous experience samples.	16000
<i>MaxStepsPerEpoch</i>	Maximum number of steps executed in each epoch.	1000
<i>PriorityExponentAlpha</i>	Exponent parameter in the priority calculation of the experience replay buffer.	0.6
<i>ImportanceSamplingBetaStart</i>	Initial importance sampling parameter in the priority calculation of the experience replay buffer.	0.4
<i>ImportanceSamplingBetaEnd</i>	Final importance sampling parameter in the priority calculation of the experience replay buffer.	1
<i>ImportanceSamplingBetaDecay</i>	Decay steps for the importance sampling parameter in the priority calculation of the experience replay buffer.	500

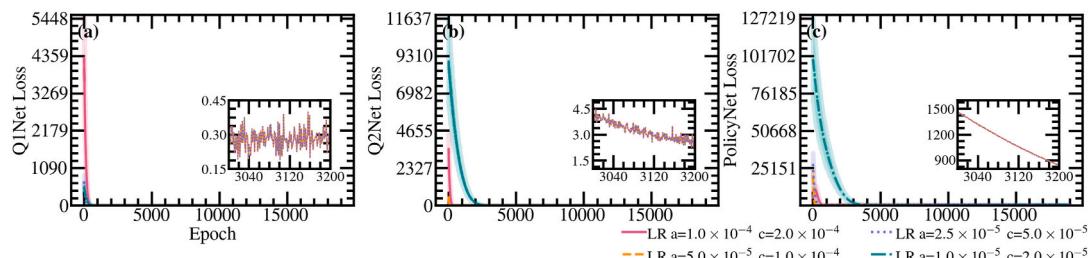


Fig. 19. Training loss dynamics of Q1-network, Q2-network, and policy network under different learning rates: (a) Training loss of Q1-network over iterations; (b) Training loss of Q2-network over iterations; (c) Training loss of policy network over iterations.

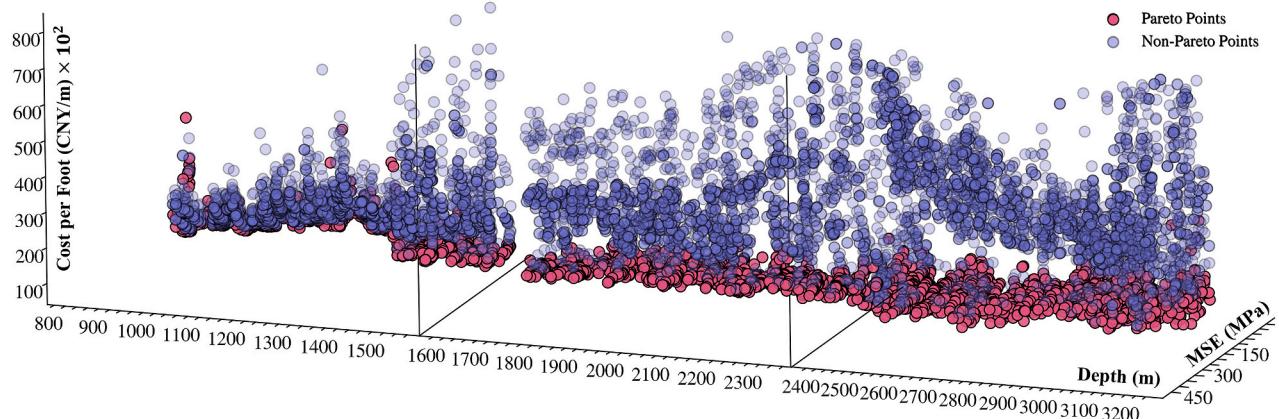


Fig. 20. Dynamic changes in unit cost per foot and MSE with drilling depth and Pareto frontier.

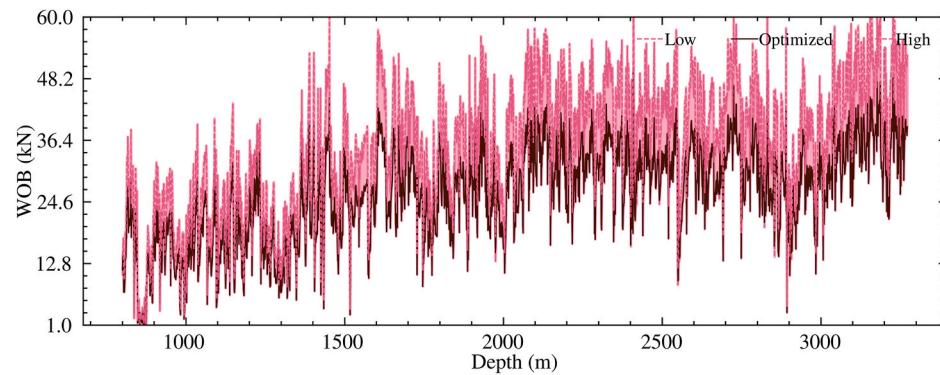


Fig. 21. Adjustment range and optimization dynamics of WOB with drilling depth.

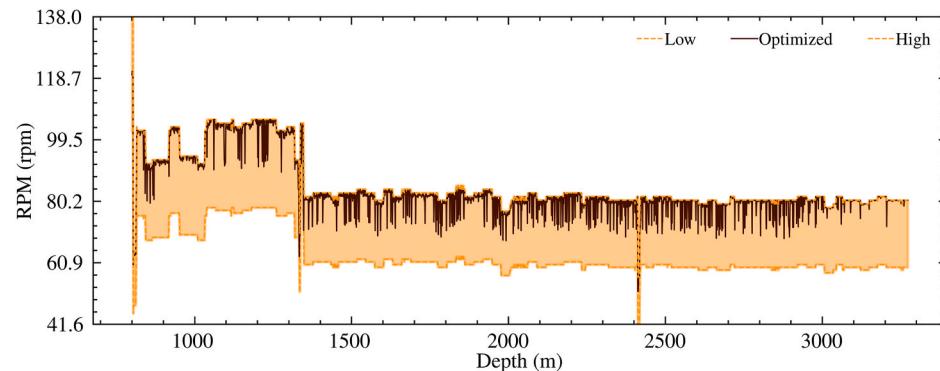


Fig. 22. Adjustment range and optimization dynamics of RPM with drilling depth.

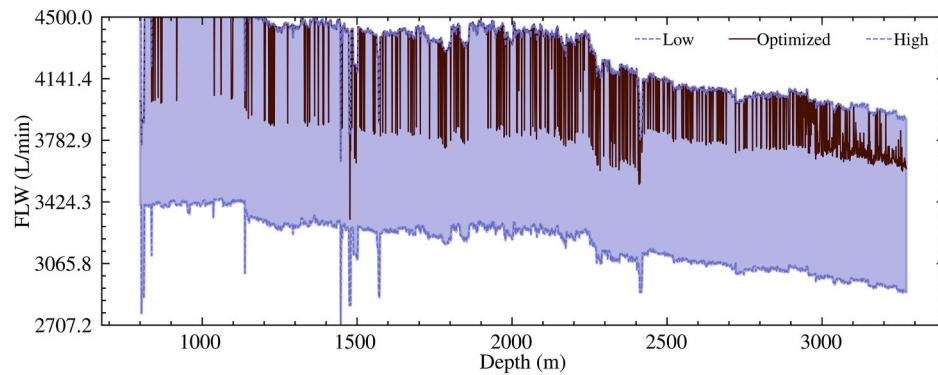


Fig. 23. Adjustment range and optimization dynamics of FLW with drilling depth.

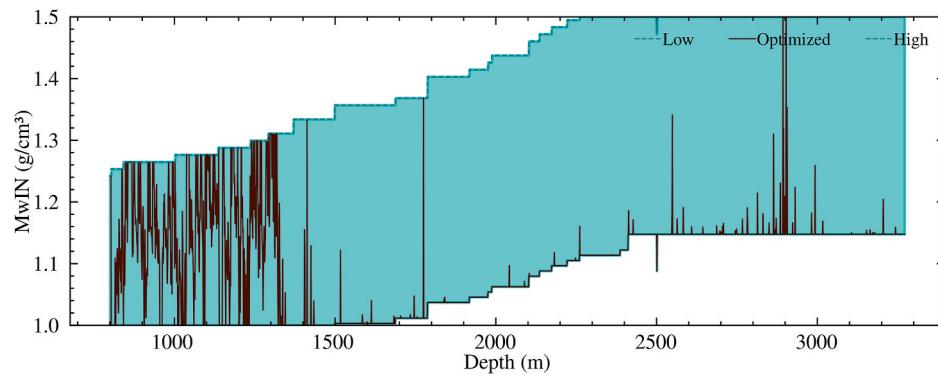


Fig. 24. Adjustment range and optimization dynamics of MwIN with drilling depth.

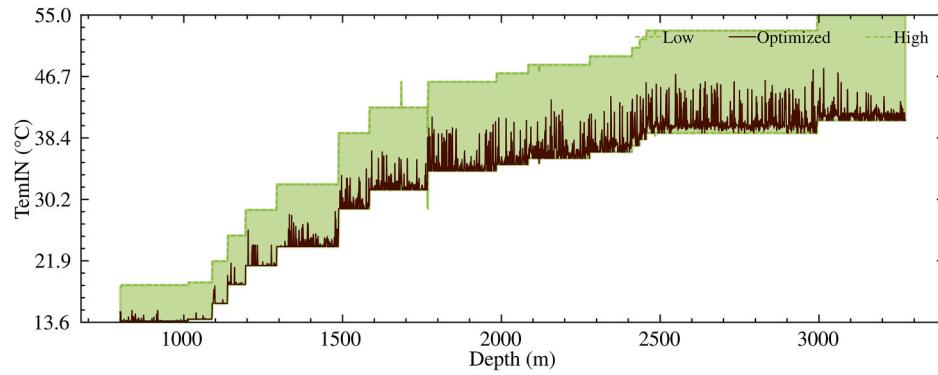


Fig. 25. Adjustment range and optimization dynamics of TemIN with drilling depth.

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

$$\left\{ \begin{array}{l} \Delta w \in [-\delta w, \delta w] \\ \Delta r \in [-\delta r, \delta r] \\ \Delta f \in [-\delta f, \delta f] \\ \Delta m \in [-\delta m, \delta m] \\ \Delta t \in [-\delta t, \delta t] \end{array} \right. \quad (12)$$

where w, r, f, m, t denote the current values of the five drilling parameters, each within specified limits; $\Delta w, \Delta r, \Delta f, \Delta m, \Delta t$ indicate the allowable changes in these parameters; $\delta w, \delta r, \delta f, \delta m, \delta t$ represent the maximum adjustments permitted during drilling operations.

To enhance the adaptability and efficiency of drilling operations, the

boundaries of the state space were dynamically modified based on the adjustment ranges of the drilling parameters forecasted by the quantile regression model. The predictive intervals of this model indicate potential variations in the drilling parameters at various risk levels, thereby enabling the state space to accommodate real-time changes in drilling operations and facilitate optimized decision-making. Furthermore, quantile regression allows an intelligent agent to anticipate and adjust to extreme conditions that may significantly influence drilling operations. Concurrently, the action space is updated flexibly to ensure precise control over the adjustment of the drilling parameters. It defines the scope of the parameter adjustments, with boundaries recalibrated according to the current limits of the state space. This adjustment creates a new range of actions, enabling the agent to make precise modifications at each decision-making step. The expressions for updating the allow-

able limits and variations in the drilling parameters are detailed in Eq. (13):

$$\left\{ \begin{array}{l} p_{\min} \rightarrow q_{0.05}^p, p_{\max} \rightarrow q_{0.95}^p \\ \Delta p = \frac{p_{\max} - p_{\min}}{20} \\ p \in \{w, r, f, m, t\} \end{array} \right. \quad (13)$$

where $q_{0.05}^p$ and $q_{0.95}^p$ represent the predicted lower and upper limits of the QR for the drilling parameters, respectively, indicating the potential adjustment spectrum at different risk thresholds.

By directly mapping changes in state boundaries to adjustments in the action range, deep reinforcement learning systems maintain sensitivity to the environment, while flexibly responding to the dynamic parameter changes required during drilling operations (François-Lavet et al., 2018). This approach ensures that the action space remains aligned with the latest state space, optimizes the agent's operational strategies, and enhances the drilling efficiency and adaptability. The adaptive adjustment mechanism of the action space, achieved through the continuous monitoring of state boundaries and the proportional adjustment of action limits, enables the agent to fine-tune the drilling parameters based on real-time data and predictive outcomes.

This synchronous update mechanism underscores the advantages of deep reinforcement learning in managing dynamic and complex environments, allowing the agent to adapt to immediate changes in drilling operations while flexibly addressing variations in geological conditions and equipment performance under operational efficiency and cost-effectiveness constraints. This flexibility is achieved through the ongoing monitoring of real-time data and the integration of model predictions, ensuring that the agent can adjust the drilling parameters to accommodate geological changes and equipment performance demands. This method endows the agent with the capability to continually optimize its decision-making strategies within dynamic and complex drilling environments, thereby enhancing the overall adaptability and efficiency of drilling operations.

5.2.3. Learning mechanisms

To achieve the learning objectives efficiently, a DPODDPG algorithm was designed. This algorithm is tailored for optimization in continuous action spaces and is particularly suitable for complex decision-making environments and dynamic operational requirements in drilling parameter optimization (Tan, 2021). Fig. 18 illustrates the architecture of DPODDPG.

Online Q-network and Target Q-network: The primary role of the online Q-network is to assess the reward value under the current drilling parameter conditions, including the WOB, RPM, FLW, MwIN, and TemIN. The network takes the current state of the drilling parameters as inputs and outputs the expected rewards for various actions taken in that state, including the unit cost per foot and MSE. The online Q-network continuously learns and updates, aiming to minimize the discrepancies between the predicted and actual drilling costs and performance (Henderson et al., 2018). The target Q-network is structurally identical to the online Q-network but is initialized with the weights of the online Q-network. This setup provided a stable learning target for mitigating volatility during the training process. Throughout the learning phase, the parameters of the target Q-network are updated at fixed intervals by periodically copying the parameters from the online Q-network. This update mechanism helps to reduce instability during the learning process, thereby enhancing the overall performance and reliability of the model (Tan, 2021).

Policy Network: The policy network was designed to identify the optimal action choices under the current drilling parameter states, receiving inputs such as the WOB, RPM, FLW, MwIN, and TemIN. Based on these parameters, the network outputs the probability distribution of actions. Utilizing this distribution, the model determines the optimal action and estimates the loss value using predictions from the online Q-network and target Q-network. This loss value was subsequently used to further refine the parameters of the policy and online Q-networks, thereby facilitating more precise adjustments in the drilling operations. To ensure the stability and continuity of the training process, the policy network design includes a mechanism for regularly synchronizing the

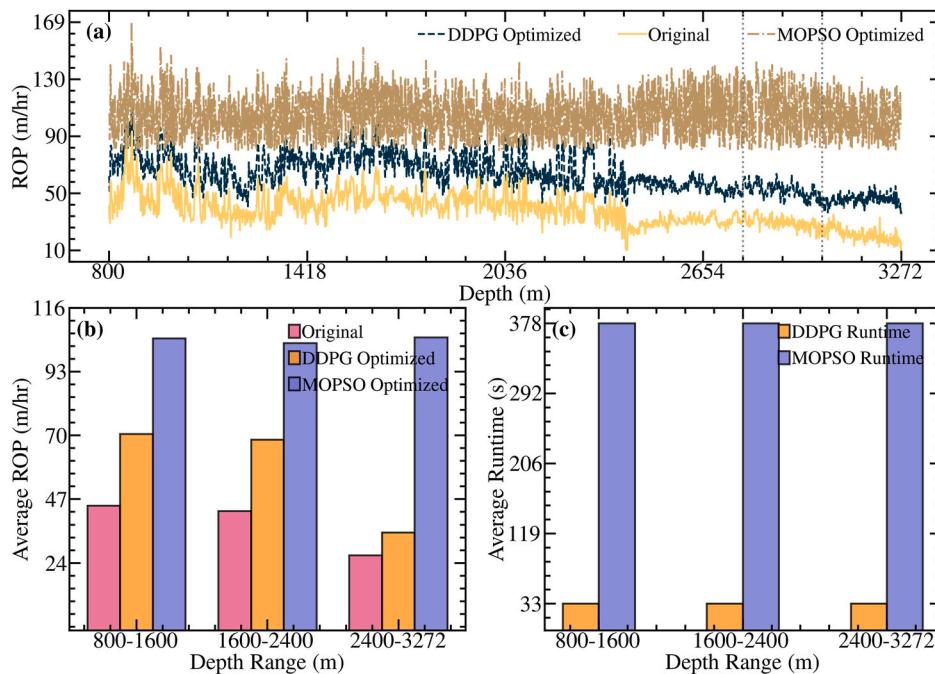


Fig. 26. Efficiency and time efficiency comparison of DPODDPG and MOPSO algorithms in ROP optimization: (a) Variation of ROP with well depth before and after optimization using DPODDPG; (b) Comparison of ROP enhancements between DPODDPG and MOPSO algorithms; (c) Comparison of time spent on each run between DPODDPG and MOPSO algorithms.

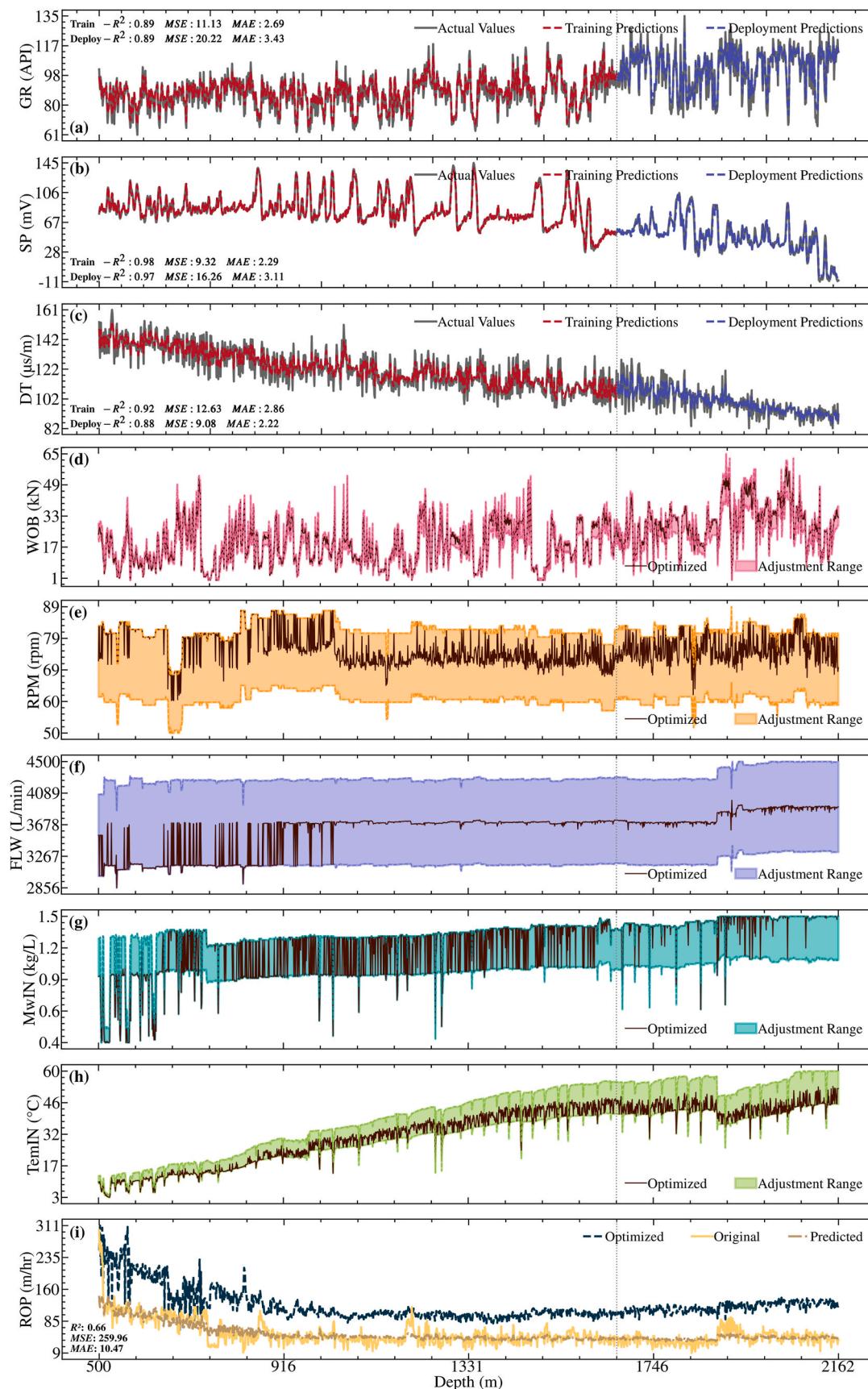


Fig. 27. Prediction of LWD parameters and optimization of drilling parameters: (a) Prediction of GR parameter; (b) Prediction of SP parameter; (c) Prediction of DT parameter; (d) Prediction of WOB parameter; (e) Prediction of RPM parameter; (f) Prediction of FLW parameter; (g) Prediction of MwIN parameter; (h) Prediction of TemIN parameter; (i) Forecasting and optimization of ROP using the DPODDPG algorithm.

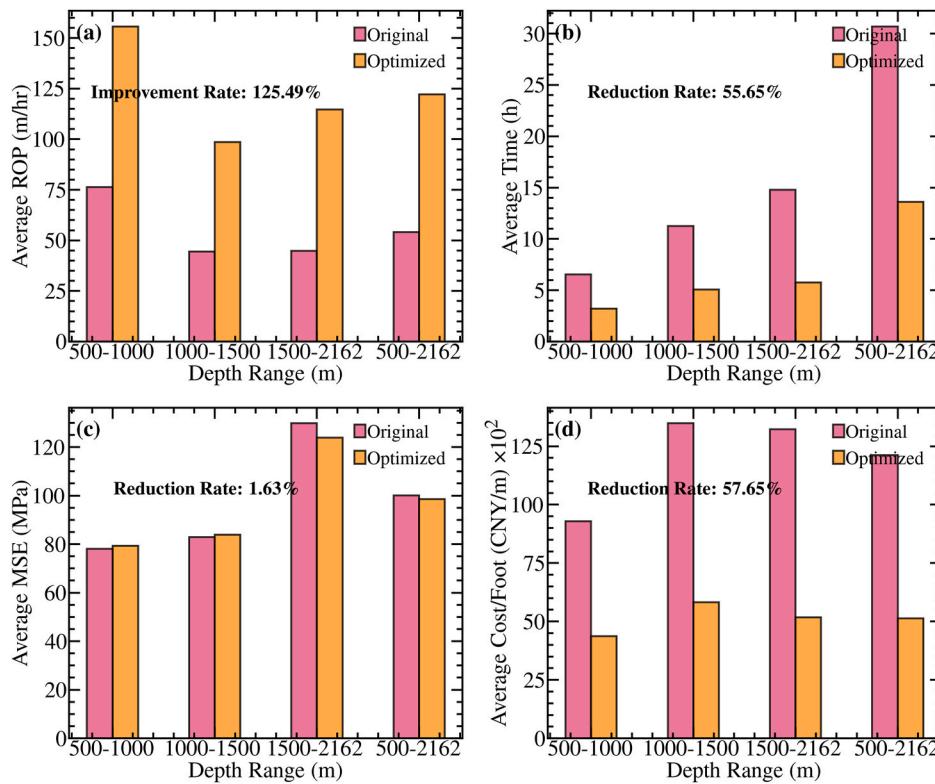


Fig. 28. Comparative analysis of drilling parameters before and after optimization across the full depth range: (a) ROP comparison; (b) Time consumption comparison; (c) Mean Squared Error (MSE) comparison; (d) Cost per foot comparison.

parameters between the online Q-network and the target Q-network. During specific phases, parameters from the online Q-network are copied to the target Q-network to maintain consistency in the learning objectives and minimize oscillations during the exploration phase (Tan, 2021). By integrating current operational states and historical data, this strategy enables the model to dynamically adjust its actions in a fluctuating drilling environment and optimize the decision-making process (François-Lavet et al., 2018).

Priority Experience Replay: In Priority experience replay plays a crucial role in the process of deep reinforcement learning for drilling parameter optimization. The experience pool, which is a centralized data structure that stores historical drilling states and their associated rewards, provides a continuous data input for subsequent learning. Employing a random sampling strategy ensures uniform and unbiased learning across various drilling parameter states, thereby promoting the model's generalizability and stability while disrupting data temporal correlations (Tan, 2021). More importantly, the prioritized sampling strategy accurately identifies experiences that are significantly valuable for model learning (Arulkumaran et al., 2017). In the process of adjusting the drilling parameters, some specific states may be more challenging to optimize owing to their different operational effects. By assigning higher learning priorities to these complex or anomalous drilling states, the model can more effectively optimize the parameters for these critical states, thereby enhancing the overall optimization outcomes and stability.

Exploration Noise Strategy: Exploration noise strategies play a critical role in the early stages of drilling parameter optimization. In the initial phase of the optimization process, these strategies encouraged a multi-objective deep reinforcement learning algorithm to extensively explore various possible states of the drilling parameters. As the understanding of the optimization process deepened, the focus of the strategy gradually shifted towards drilling states that could optimize the

drilling time predicted by the drilling parameter prediction model. By incorporating random noise in the action selection process, the strategy increases the randomness of the action policy, encouraging the agent not only to repeat known effective strategies but also to explore new possibilities, thereby discovering potentially higher-return action strategies (Tan, 2021). This approach is particularly crucial for drilling operations in complex geological environments, as these environments often conceal behavioral patterns that are difficult to predict using traditional methods. The complete process for optimizing the drilling parameters using the DPODDPG algorithm is presented in Table 5.

5.3. DPODDPG algorithm

5.3.1. Training of Q-network, target network, and policy network

For the drilling operations at a specific well location in the Ledong 10-1 Block of the South China Sea, Table 6 lists the coefficient settings required for the unit cost per foot function and the MSE function. The precise setting of these coefficients is crucial, as they directly influence the calculation and optimization of drilling cost efficiency.

When optimizing the drilling parameters using the DPODDPG algorithm, the training hyperparameters were set as listed in Table 7. The specific functions and selected values of each parameter significantly affected the performance and convergence speed of the algorithm.

Precise updates of Q-networks and policy networks are crucial during the training process for drilling parameter optimization. The Q-networks are optimized by minimizing a loss function that quantifies the discrepancy between the actual rewards obtained based on the current drilling states and actions taken, and the predicted rewards. The policy networks derive optimal action strategies from the current drilling states, adjusting key drilling parameters such as WOB and RPM to minimize the cost per foot and MSE. The generated actions are used to estimate the expected rewards in the Q-networks, forming a composite

loss for the policy networks, which is optimized through a gradient descent to adjust the network parameters and ensure policy effectiveness and adaptability. Target Q-networks utilize a parameter update strategy that does not involve learning through a gradient descent; instead, they periodically replicate the parameters of the main Q-network. This method aims to provide stable target values for loss function calculations and gradient updates, reduce volatility during training, and enhance algorithm convergence. This strategy for target networks contributes to the smoothness of the training process and enhances the stability of learning, thereby optimizing the overall system performance and reliability.

Q-networks evaluate not only the immediate benefits of drilling operations but also the long-term trade-offs between cost and efficiency. Through periodic network parameter adjustments and real-time reward feedback, the deep reinforcement learning model gradually refines its decision-making process over continuous training cycles. Fig. 19 illustrates the training loss changes in the Q1-network, Q2-network, and policy network under four different learning rate settings, revealing the performance and convergence characteristics of each network during the training phase.

Fig. 19 indicates that the Q1-network progressively converged towards nearly zero loss across all learning rate configurations, achieving the lowest convergence loss under a specific learning rate setting, which significantly optimized the training outcome. This highlighted the critical role of an appropriate learning rate in enhancing Q1-network's performance. For Q2-network, setting the learning rate to a specific value exhibited the best performance, with the training loss rapidly decreasing to the lowest point, indicating that a higher learning rate facilitates the rapid optimization of the network. The policy network loss showed a trend from positive to negative values, reaching the smallest negative value under a certain setting, suggesting a risk of overfitting. However, a configuration with a different set of parameters, although converging more slowly, exhibited a more robust training loss trend, which might better satisfy the generalization requirements in real-world applications.

These results demonstrate that appropriate learning rate settings decisively affect the performance of the DPODDPG algorithm, especially when adapted to complex drilling environments. Properly adjusting the learning rate affects not only the convergence speed of the model, but also its ultimate performance and stability.

5.3.2. Optimization process of drilling parameters

Fig. 20 illustrates the variations in cost per foot and MSE across drilling depths from 800 to 3375 m, with results recorded after 3000 iterations at each depth. Notably, the Pareto frontier is prominently displayed as a curve, marking an optimal tradeoff between cost and efficiency.

Fig. 20 reveals that as the drilling depth increases, both the cost per foot and MSE exhibit specific trends. In the initial phase (800–1500 m), the fluctuations in both metrics were substantial, reflecting the exploratory nature of the parameter adjustments during this drilling segment. The frequent changes in the Pareto frontier during this stage indicate that the optimization algorithm tested various operational combinations to find a balance between cost and efficiency. In the middle phase (approximately 1500–2500 m), fluctuations in cost and energy began to decrease, and the Pareto frontier gradually stabilized. This suggests that the algorithm shifts from broad exploration to more focused exploitation, thereby honing more effective parameter combinations. During this phase, the trade-off between cost and efficiency becomes more apparent, and the Pareto frontier indicates the optimal operational points more clearly. In the later stage, from depths of 2500–3375 m, changes in the Pareto frontier were minimal, and both the cost per foot and MSE tended to stabilize. At this point, the data points in the chart are

dense and converge towards a consistent range, indicating that the optimization process is near its optimal solution, and that the algorithm makes fine adjustments to the operational parameters to fine-tune the final balance between cost and efficiency.

The iterative results from the DPODDPG algorithm showed its dynamic capability in the gradual optimization of the drilling parameters. By leveraging the continuous exploration and exploitation phases, the algorithm effectively directs drilling operations towards optimal cost effectiveness and mechanical efficiency. Additionally, the algorithm's timely adjustments across different stages significantly enhanced the adaptability to complex geological conditions as the drilling depth increased, ultimately optimizing both the cost and efficiency of the drilling operations.

Detailed adjustments and optimized values for parameters such as drilling pressure, RPM, FLW, MwIN, and TemIN were documented with increasing drilling depths and are presented in Figs. 21–25, respectively.

Fig. 21 shows the fluctuations in WOB as the drilling depth increased, with the adjustments and optimized values exhibiting significant variability. This variability reflects the uncertainty of the algorithm in responding to actual drilling conditions as it seeks the optimal WOB setting. Particularly in the early stages of drilling, as geological conditions vary, the optimization of the WOB shows adaptive adjustments to different drilling strata to accommodate changes in formation hardness and fracturability.

Figs. 22 and 23 show that both the RPM and FLW tended to remain at the upper limit of their adjustment ranges throughout the drilling process. This phenomenon indicates that as the drilling depth increases, the algorithm tends to select a higher RPM and FLW to enhance drilling efficiency and wellbore cleaning effectiveness. In particular, when encountering harder formations, a higher RPM and FLW help accelerate the ROP and effectively remove cuttings, thereby reducing the wear on the drill bit and preventing complex downhole situations.

For the MwIN, Fig. 24 shows significant initial adjustments reflecting the initial uncertainty as the algorithm explores the optimal mud density settings. As the drilling depth increased, the mud density gradually decreased, stabilized, and trended toward the lower bound of the adjustment range. This indicates that the algorithm adapted to actual geological conditions, optimizing wellbore stability and cutting-carrying capacity, and thus enhancing drilling safety and efficiency.

Regarding the TemIN, the data in Fig. 25 show minimal adjustments during the initial stage. However, as the drilling depth increases, the optimized temperature values begin to fluctuate. This variability likely stems from the algorithm's response to changing subsurface temperatures and chemical properties, particularly during deep drilling. Drilling fluid temperature within the wellbore is influenced by multiple factors, including formation temperature and circulating heat transfer. However, Fig. 25 focuses on TemIN, which is primarily controlled at the surface. By adjusting TemIN, the mud temperature within the wellbore can be regulated to some extent, thereby maintaining the physicochemical properties of the mud and optimizing drill bit performance despite the complexities introduced by subsurface conditions. Appropriate temperature adjustments are crucial for maintaining mud properties and optimizing the working conditions of the drill bit.

The optimization results demonstrate the effectiveness of the DPODDPG algorithm in a multiparameter drilling environment, particularly in dynamically adjusting and precisely controlling drilling parameters to adapt to complex geological conditions and optimize drilling costs and efficiency. Additionally, based on predictions from the quantile regression model, dynamic adjustments of the drilling parameter ranges provide precise control over the MwIN and temperature. The dynamic lower-bound adjustments of the quantile regression model reflect progressively stringent formation requirements. Under this mechanism, although the optimized values of the mud parameters

closely follow the lower bounds of the adjustment range, these bounds gradually increase with depth, ensuring that wellbore stability and mud-carrying capacity are maintained while adapting to the specific requirements of deeper geological layers. This adjustment strategy not only improves operational adaptability, but also enhances the economic efficiency and effectiveness of the entire drilling process, allowing the algorithm to effectively adjust strategies at various drilling stages and achieve an optimal balance between cost and efficiency.

Fig. 26(a) shows the change in ROP with well depth, detailing the optimization results at each drilling depth, and clearly displays the specific changes in ROP before and after optimization using the DPODDPG algorithm within the dynamic adjustment constraints of the quantile regression model. Improvements in the ROP are significant yet reasonably constrained. **Fig. 26(b)** and **26(c)** compare the enhancements in the ROP and time spent on each run between the DPODDPG and Multiple Objective Particle Swarm Optimization (MOPSO) algorithms.

In the initial stages, a notable increase in the ROP was observed, highlighting the extensive exploration and adjustment of the drilling parameters of the algorithm. As the drilling depth increased, the rate of speed increase began to stabilize, reflecting the ability of the DPODDPG algorithm to effectively identify and apply the most effective operational settings. By integrating the unit cost per foot and MSE as optimization objectives, the algorithm not only optimized the ROP, but also ensured the stability and safety of the drilling process, avoiding mechanical failures and operational risks that could arise from excessively rapid ROPs. Additionally, **Fig. 26(a)** shows a significant reduction in ROP fluctuations with increasing depth, particularly in the mid to late stages (depths exceeding the median). This indicates that the DPODDPG algorithm successfully optimizes the drilling parameters, reducing ineffective and excessive operations, thereby minimizing resource wastage and enhancing drilling efficiency. The stability of the ROP was closely linked to the dynamic adjustment range provided by quantile regression, ensuring that the ROP did not pose risks owing to extreme operations. In contrast, although the MOPSO algorithm also achieved significant improvements in the ROP throughout the drilling process, its lack of dynamic adjustment constraints, especially in the later stages of drilling, increased the operational risks owing to unreasonable speed increases. This behavior of the MOPSO algorithm can lead to extreme operations, particularly at greater drilling depths, failing to adapt to changing geological conditions, and potentially causing safety incidents or equipment damage.

Fig. 26(b) and **26(c)** compare the performances of the DPODDPG and MOPSO algorithms in optimizing the ROP and duration of each run. The analysis indicated that the MOPSO algorithm required significantly more time for each optimization than the DPODDPG algorithm, which is often more than ten times longer and does not meet the requirements for real-time optimization. This time difference is primarily due to the iterative nature of MOPSO as a heuristic algorithm. For example, after acquiring real-time drilling data, MOPSO requires numerous iterations (such as 1000) to determine the optimal drilling parameters, which must be repeated at subsequent moments. This method is inefficient in dynamic drilling environments because each optimization takes too long to respond swiftly to real-time changes in the drilling parameters. In contrast, the DPODDPG algorithm adapts quickly to the current drilling conditions through continuous action space optimization. Not only does DPODDPG demonstrate extensive exploratory capabilities in the initial stages, but as drilling progresses, its optimization of the ROP stabilizes, effectively balancing cost and efficiency while ensuring the stability and safety of drilling operations. Additionally, DPODDPG's dynamic adjustment mechanisms of DPODDPG, such as range settings based on quantile regression, ensure that even in the face of geological changes, ROP adjustments do not fluctuate excessively, thus avoiding potential risks. This enhanced adaptability and real-time responsiveness make

DPODDPG superior in handling the dynamic nature of drilling parameters compared to traditional methods like MOPSO. Overall, the introduced DPODDPG algorithm ensures effective enhancement of the ROP while maintaining the safety and economic efficiency of drilling operations through sensible dynamic adjustments and strategy optimization.

6. Case study

To validate the universality and effectiveness of the multi-objective deep reinforcement learning real-time optimization process for drilling parameters, a neighboring well in the Ledong 10-1 block of the South China Sea (covering depths from 500 to 2162 m) was selected as a test case for empirical analysis. During the initial phase of the drilling operations (from 500 m to 1500 m depth), the VMD-CNN-BiLSTM-MA model combined with the CNN-BiLSTM-MA-QA model was trained to predict the numerical values of the LWD parameters and the adjustment range of the drilling parameters. These models are trained using the strategy parameters listed in **Table 4**. Based on the coefficients required for the cost per foot function and the MSE function, as detailed in **Table 6**, and constructed using the ROP empirical equation, these optimization targets reflect key considerations of cost-effectiveness and mechanical efficiency in drilling operations. In addition, the Q-network, target network, and policy network within the DPODDPG algorithm were trained using the hyperparameters specified in **Table 7**.

In the later phase (from 1500 to 2162 m depth), the models trained during the initial stage were deployed for the real-time optimization of drilling parameters. The effectiveness of the ROP empirical equation for predicting speeds in later stages, the performance of the VMD-CNN-BiLSTM-MA model for predicting LWD parameters, and the effectiveness of the DPODDPG algorithm in optimizing drilling parameters for later stages are illustrated in **Fig. 27**.

Fig. 27(a)–27(c) illustrate the precision with which the VMD-CNN-BiLSTM-MA model predicts the key LWD parameters. The predicted curves were closely aligned with the actual measurements, verifying the exceptional capability of the model for deciphering complex drilling data. This model effectively captures variations in the logging parameters during the drilling process, offering precise data support for drilling parameter optimization decisions, which significantly enhances both the efficiency and safety of drilling operations. Additionally, **Fig. 27(i)** shows the forecasting and optimization of the ROP, revealing the robust performance of the model in generalizing to neighboring wells with an R^2 value of 66%, thereby affirming its applicability and stability. A detailed segment is illustrated in **Appendix C** for clearer visualization.

Fig. 28 compares the key drilling parameters and performance indicators over the entire depth range (500–2162 m) before and after optimization. These indicators include ROP, time consumption, MSE, and cost per foot, illustrating the improvements in drilling operations facilitated by the optimization process.

Fig. 28 illustrates the substantial increase in ROP following optimization, with speeds escalating from an initial 54.18 m/h to 122.17 m/h, nearly doubling the rate of drilling. Concurrently, the time required for drilling was reduced from 30.68 h to 13.60 h, more than halving the duration, thus substantially enhancing operational efficiency. Despite only a marginal decrease in MSE from 100.82 MPa to 97.78 MPa, this reduction signifies efficiency improvements maintained alongside equipment stability and rock-breaking efficiency. Most significantly, the cost per foot drilled was drastically reduced from 121.16×10^2 CNY/m to 51.31×10^2 CNY/m, evidencing not only significant cost-effectiveness but also the economic success of the optimization strategy.

Intervention with the DPODDPG algorithm markedly improved the optimization of the ROP. Within this framework, the cost per foot and MSE were designated as the optimization targets. Through continual adjustments to the drilling parameters, the DPODDPG algorithm

ensured that increases in ROP did not compromise equipment safety or durability, thereby striking an optimal balance between economic efficiency and operational effectiveness. These outcomes not only validate the efficacy of the DPODDPG algorithm in tailoring drilling parameters to diverse geological conditions but also highlight its potential to boost operational efficiency in real drilling scenarios.

This case study thoroughly verified the efficacy and wide applicability of the proposed multi-objective deep reinforcement learning instant optimization process in actual drilling operations. Following validation in the Ledong 10-1 block in the South China Sea, this methodology successfully fine-tuned drilling parameters and exhibited remarkable optimization results across various geological conditions and drilling depths.

7. Conclusions

This study advances the field of drilling optimization by introducing a novel multi-objective reinforcement learning framework that enables instant optimization of drilling parameters, significantly enhancing operational efficiency and decision-making in complex drilling environments. Utilizing the ROP empirical equation built with the MESR-DS algorithm, the framework created empirical equations that comprehensively mapped the intricate relationships among drilling, mud logging, and LWD parameters, offering precise quantification of the dynamic characteristics of drilling and interactions among various parameters.

By enhancing the early detection of formation parameters and resolving latency issues inherent in traditional methods, the framework significantly improves decision-making timeliness. This advancement facilitates rapid responses to changes during drilling operations, optimizes drilling processes, and minimizes operational risks.

By establishing a Markov decision environment for real-time drilling parameters, the framework ensures precise control over drilling operations and optimization of economic benefits. The DPODDPG algorithm, through its innovative integration of experience replay, priority-adjustment mechanisms, and exploration noise strategies, enhances learning efficiency and model adaptability. This not only ensures real-time adaptability to changes in geological conditions but also contributes to broader goals of operational safety, cost reduction, and sustainability by optimizing both economic and mechanical efficiencies in drilling operations. The framework's application of quantile regression to estimate adjustments in the drilling parameters further boosts its dynamic adaptability and responsiveness, promoting more efficient and safer drilling practices.

The empirical analysis conducted in the Ledong 10-1 block in the South China Sea confirms that the framework significantly improves ROP and optimizes both the mechanical specific energy (MSE) and cost per foot. These results not only demonstrate the framework's effectiveness in this specific context but also highlight its potential for broader applicability across different drilling environments and geological conditions, indicating its value as a versatile tool for modern drilling operations worldwide.

Future research will focus on extending the framework to optimize drilling parameters in horizontal wells and developing adaptive decision-making strategies to further enhance operational efficiency and safety in increasingly complex drilling scenarios.

Appendix A. Overview of Drilling Parameters for the Ledong 10-1 Block in the South China Sea

This appendix A provides a overview of the drilling parameters, mud logging parameters, and LWD parameters collected from a well in the Ledong 10-1 block in the South China Sea.

Nomenclature

Symbol/ Abbreviation	Definition	Unit
WOB	Weight on Bit	kN
RPM	Revolutions Per Minute	rpm
FLW	Flow Rate	L/ min
MwIN	Mud Input Density	g/ cm ³
TemIN	Mud Input Temperature	°C
GR	Gamma Ray	API
SP	Spontaneous Potential	mV
DT	Delta Time	μs/m
ROP	Drilling Rate	m/hr
MSE	Mechanical Specific Energy	—
LWD	Logging-While-Drilling	—
IQR	Interquartile Range	—
VMD	Variational Mode Decomposition	—
SampEn	Sample Entropy	—
FFT	Fast Fourier Transform	—
1D-CNN	One-Dimensional Convolutional Neural Network	—
BiLSTM	Bidirectional Long Short-Term Memory	—
MA	Multithead Attention	—
DPODDPG	Drilling Parameter Optimization Deep Deterministic Policy Gradient	—
MOPSO	Multiple Objective Particle Swarm Optimization	—
FCNN	Fully Connected Neural Network	—
MESR-DS	Multi-Evolutionary Symbolic Regression for ROP	—
QA	Quantile Adjustment	—
QR	Quantile Regression	—
R ²	Coefficient of Determination	—

CRediT authorship contribution statement

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

Declaration of competing interest

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

Acknowledgements

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

Table A1

Overview of drilling parameters for the Ledong 10-1 block in the South China Sea.

Depth (m)	WOB (kN)	RPM (rpm)	FLW (L/min)	MwIN (g/cm ³)	TemIN (°C)	GR (API)	SP (mV)	DT (μs/m)	ROP (m/hr)
...
1500	26.46	71	3703	1.21	42.6	94.1068	73.1179	106.2119	21.89781
1507	25.48	71	3704	1.21	42.7	87.4601	75.8272	110.9728	34.28571
1509	17.64	72	3720	1.21	43.1	89.7698	73.5790	109.4237	31.08808
1510	18.62	72	3719	1.21	43.1	86.2728	72.8765	107.2750	29.70297
1511	19.60	72	3719	1.21	43.2	87.4200	72.8240	108.1329	33.70787
1513	38.22	72	3706	1.21	43.2	86.3932	73.8503	104.7009	33.70787
1514	23.52	72	3707	1.21	43.3	84.9335	73.7579	102.4631	44.77612
1515	20.58	72	3708	1.21	43.3	82.5137	73.0971	107.4914	51.28205
1516	19.60	72	3709	1.21	43.4	80.9934	73.0691	106.2352	48
1517	21.56	71	3710	1.21	43.4	81.4703	73.3718	105.0631	44.44444
1531	27.44	71	3710	1.21	42.6	77.9440	71.9530	102.2403	46.15385
1532	27.44	71	3710	1.21	43.2	85.4855	73.0598	117.9029	38.46154
1533	27.44	71	3711	1.21	43.4	80.4496	73.8538	112.1620	40
1534	27.44	71	3710	1.21	43.5	81.4296	75.4930	105.0582	33.70787
1535	29.40	71	3711	1.21	43.5	84.6899	76.2330	112.8852	37.26708
1536	25.48	71	3707	1.21	43.6	92.4549	73.5452	118.2079	34.09091
1541	27.44	71	3710	1.21	43.9	91.3245	71.4553	116.0237	27.90698
1542	18.62	72	3712	1.21	43.9	82.7569	72.7150	111.1800	41.95804
1543	19.60	72	3711	1.21	44.0	95.8315	73.3238	118.5333	37.5
1544	23.52	72	3712	1.21	44.0	98.0077	72.8364	123.1381	40.54054
1545	24.50	72	3711	1.21	44.1	101.8051	72.2139	122.9742	37.73585
1546	24.50	72	3712	1.21	44.1	100.2130	72.6016	129.9390	36.58537
1547	37.24	71	3714	1.21	44.1	92.3046	72.2680	113.8804	26.66667
1548	16.66	71	3712	1.21	44.3	79.9639	72.0313	105.6524	47.61905
1549	20.58	71	3712	1.21	44.3	109.4998	72.4242	131.9004	30.15075
1550	26.46	72	3712	1.21	44.3	97.7938	72.3836	122.0387	45.11278
1551	26.46	72	3714	1.21	44.3	90.5034	71.3847	105.7041	41.66667
1552	27.44	72	3715	1.21	44.3	88.0758	72.2953	120.3434	38.46154
1553	37.24	72	3713	1.21	44.2	95.6571	72.7400	121.4985	25.53191
1554	25.48	71	3715	1.21	44.2	82.5265	73.1272	104.3144	58.82353
1555	26.46	71	3716	1.21	44.2	81.8358	71.7444	102.8355	57.14286
1560	28.42	71	3718	1.20	42.3	85.5492	74.1663	109.6695	52.63158
1561	30.38	71	3718	1.20	43.0	90.7262	74.7131	112.1871	37.5
1563	22.54	72	3707	1.20	43.7	101.3362	73.9503	118.5695	35.08772
1564	31.36	71	3706	1.21	43.8	91.6869	74.2014	112.5692	40.26846
1565	23.52	71	3702	1.20	43.8	91.9497	74.4384	113.7755	37.26708
...

Appendix B. MESR-DS Algorithm for Empirical ROP Equation Construction: Process and Application

This appendix presents [Table B1](#), which details the process of using the MESR-DS algorithm to construct empirical ROP equations.

Table B1

Process of constructing empirical ROP equation using the MESR-DS algorithm.

Algorithm: Integrated MESR-DS Process

Input: X (dataset with features: [Depth, WOB, RPM, MwIN, TemIN, FLW, GR, SP, DT]), y (target variable ROP)

Parameters:

- np = number of populations,
- L = expressions per population,
- nc = mutations per evolution cycle,
- $pcross$ = probability of performing crossover,
- αH = fraction of replacements from global bests,
- αM = fraction of replacements from other populations' bests,
- ns = size of tournament,
- $ptournament$ = probability of selecting the fittest in tournament,
- f = scale of constant perturbation,
- ϵ = minimum perturbation,
- α = annealing temperature scale

Output: Optimal expressions for predicting ROP at each complexity level

Function MESR-DS (X)

- 1: **Initialize** population sets P_i **for** $i = 1$ **to** np , each containing L random expressions of initial complexity
- 2: **Initialize** M_i to store the best expressions from each P_i , create global set H for overall best expressions
- 3: Repeat for a set number of iterations or until convergence:
- 4: **for** each population P_i in parallel:
- 5: **for** each evolution cycle k **from** 1 **to** nc :
- 6: **if** random() < p_{cross} :
- 7: **for** each pair selection:
- 8: Sample a subset Q of size ns from P_i
- 9: **while** length(Q) > 1:
- 10: Determine E as the expression in Q with highest fitness by evaluating against X , y :
- 11: Adjust fitness score by expression complexity using a parsimony parameter
- 12: **if** random() < $ptournament$, **break** and use E
- 13: **else**, remove E from Q
- 14: $E_1, E_2 \leftarrow$ the remaining or selected expressions from Q
- 15: Identify crossover points in E_1 and E_2
- 16: Swap sub-trees at these points to generate new expressions E_1^*, E_2^*
- 17: Validate E_1^*, E_2^* ; if better, replace two oldest expressions in P_i
- 18: **else:**
- 19: Sample a subset Q of size ns from P_i for selection
- 20: **while** length(Q) > 1:
- 21: Determine E with the highest fitness by evaluating against X , y :
- 22: Fitness calculation includes adjustments for parsimony and accuracy metrics
- 23: **if** random() < $ptournament$, select E and break
- 24: **else**, remove E from Q
- 25: Compute $T = 1 - k/n$ as the annealing temperature for this cycle
- 26: $E^* \leftarrow$ apply mutation to E based on a weighted probability of mutation types:
- 27: For constant mutations, perturb a constant in E^* using T , f , ϵ
- 28: For operator mutations, randomly replace an operator in E^*
- 29: Other mutation types involve adding, removing, or modifying nodes
- 30: **if** E^* meets operational constraints, replace the least fit expression in P_i
- 31: Update M_i with the most accurate expression from P_i based on defined metrics
- 32: Aggregate the best expressions from all M_i to update H
- 33: For each population P_i , migrate expressions between populations:
- 34: **for** each expression E in P_i :
- 35: With probability αH , replace E with a randomly selected expression from H
- 36: With probability αM , replace E with an expression from a different M_i
- 37: Evaluate and update the global hall of fame H with the most accurate expressions across all complexities
- 38: **Return** H containing optimized expressions for predicting ROP

This appendix presents **Table B2**, which provides a detailed summary of the results from the application of the MESR-DS algorithm in constructing empirical ROP equations. **Table B2** includes information on the complexity of each model, the corresponding loss values, scores, and the parameters used.

Table B2

Results of empirical ROP equation construction using the MESR-DS algorithm.

Complexity	Loss	Score	Parameters	Equation
1	61.3293	0.0049	GR, FLW	$\text{ROP} = \frac{\text{FLW}}{\text{GR}}$
2	53.0174	0.0631	GR	$\text{ROP} = \frac{3453.92}{\text{GR}}$
3	49.6315	0.1150	GR, TemIN	$\text{ROP} = \frac{1723.78}{\text{GR} - \text{TemIN}}$
4	39.5033	0.0219	Depth, GR	$\text{ROP} = \frac{\text{Depth}}{\log(\text{GR})} - 304.17$
5	39.3367	0.0866	Depth, GR	$\text{ROP} = \left \frac{\text{Depth}}{\log(\text{GR})} - 303.76 \right $
6	37.5654	0.2052	Depth, GR, DT	$\text{ROP} = \frac{\text{Depth}}{\log(\text{GR})} - 303.31 - \cos(\text{DT})$
7	36.3383	0.2053	Depth, GR, TemIN	$\text{ROP} = \frac{\text{Depth}}{\log(\text{GR})} - 303.90 - \sin(\text{Depth} + \text{TemIN})$
8	30.0062	0.0328	Depth, GR, WOB, TemIN, FLW	$\text{ROP} = \frac{\text{GR}}{\text{WOB}} - \frac{\sin(\text{Depth})}{\cos(\text{FLW})} + \frac{\text{Depth}}{\text{GR} - \text{TemIN}}$
9	29.4904	0.0270	DT, Depth, GR, WOB, TemIN, FLW	$\text{ROP} = \frac{\text{DT}}{\text{WOB}} - \frac{\sin(\text{Depth})}{\cos(\text{FLW})} + \frac{\text{Depth} - \text{WOB}}{\text{GR} - \text{TemIN}}$
10	28.2886	0.0416	Depth, GR, WOB, TemIN, FLW	$\text{ROP} = \frac{\text{GR}}{\text{WOB}} + \left(\frac{\sin(\text{TemIN}) - \sin(\text{Depth})}{\cos(\text{FLW})} \right) - \frac{\text{Depth}}{\text{TemIN} - \text{GR}}$
11	27.5359	0.0173	Depth, GR, WOB, TemIN, FLW	$\text{ROP} = \frac{\text{GR}}{\text{WOB}} + \left(\frac{\cos(\text{Depth} - \text{FLW}) - \sin(\text{Depth})}{\cos(\text{FLW})} \right) - \frac{\text{Depth}}{\text{TemIN} - \text{GR}}$
12	26.6486	0.1915	Depth, GR, WOB, TemIN, MwIN	$\text{ROP} = \frac{\cos\left(\exp(\text{WOB}) + \frac{\text{GR}}{\text{GR}} - \text{TemIN}\right)}{0.1496} - \frac{\text{Depth}}{\text{TemIN} + \text{MwIN}} + \text{MwIN}$
13	21.7022	0.0332	Depth, GR, WOB, TemIN, MwIN, FLW	$\text{ROP} = \frac{\cos\left(\exp(\text{WOB}) - \frac{\text{TemIN}}{\text{MwIN}}\right)}{0.1679} - \frac{\sin(\text{Depth})}{\cos(\text{FLW})} + \frac{\text{Depth}}{\text{GR} - \text{TemIN}}$
14	17.6754	0.0461	Depth, GR, WOB, TemIN, MwIN, FLW	$\text{ROP} = \frac{\cos\left(\exp(\text{WOB}) - \frac{\text{TemIN}}{\text{MwIN}}\right)}{0.1679} - \frac{\sin(\text{Depth})}{\cos(\text{FLW})} + \frac{\text{Depth} + \text{GR}}{\text{GR} - \text{TemIN}}$
15	16.2096	0.0042	Depth, GR, WOB, RPM, TemIN, FLW	$\text{ROP} = \frac{\cos(\exp(\text{WOB}) + \log(\text{Depth}) - \text{TemIN})}{0.1391} - \frac{\sin(\text{Depth}) + \text{RPM} + \text{Depth}}{\cos(\text{FLW}) + \text{GR} - \text{TemIN}}$
16	15.8590	0.2282	Depth, GR, WOB, RPM, TemIN, FLW	$\text{ROP} = \frac{\sin(\cos(\exp(\text{WOB}) + \log(\text{Depth}) - \text{TemIN}))}{0.1391} - \frac{\sin(\text{Depth}) + \text{RPM} + \text{Depth}}{\cos(\text{FLW}) + \text{GR} - \text{TemIN}}$
17	14.1362	0.0660	Depth, SP, GR, WOB, RPM, TemIN, FLW	$\text{ROP} = \frac{\cos(\exp(\text{WOB}) + \sin(0.3699 \cdot \text{SP}) - \text{TemIN})}{0.1391} - \frac{\sin(\text{Depth}) + \text{RPM} + \text{Depth}}{\cos(\text{FLW}) + \text{GR} - \text{TemIN}}$
19	12.4605	0.1456	Depth, SP, GR, WOB, TemIN, FLW	$\text{ROP} = \frac{\cos(\exp(\text{WOB}) + \cos(\log(\sin(0.3699 \cdot \text{SP})))) - \text{TemIN}}{0.1391} - \frac{\sin(\text{Depth}) + \text{Depth} + \text{GR}}{\cos(\text{FLW}) + \text{GR} - \text{TemIN}}$
20	12.4002	0.3604	Depth, SP, GR, WOB, TemIN, MwIN, FLW, DT	$\text{ROP} = \frac{\cos(\exp(\text{WOB}) + \cos(\log(\sin(0.3699 \cdot \text{SP})))) - \text{TemIN}}{0.1391} - \frac{\sin(\text{Depth})}{\cos(\text{FLW}) \cdot \text{MwIN}} + \frac{\text{Depth} + \text{DT}}{\text{GR} - \text{TemIN}}$

The evaluation of ROP predictions across various complexities is illustrated, where each subplot compares actual and predicted ROP values for different model complexities. The relative error bars highlight the prediction accuracy across depth intervals, as demonstrated in **Figure B1**.

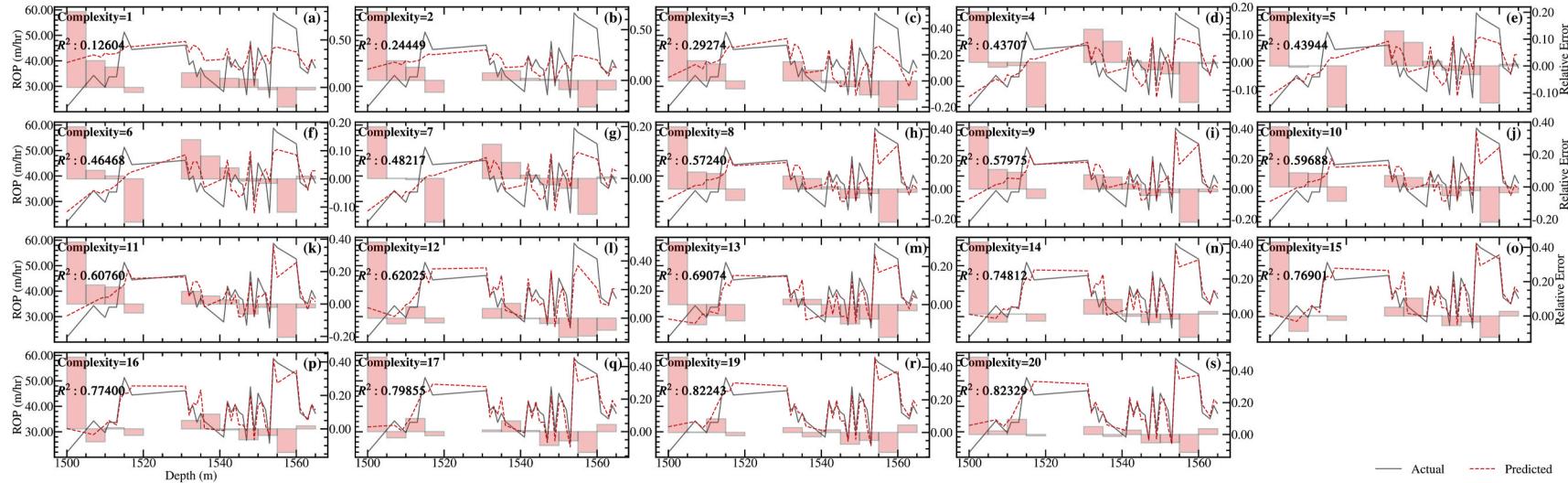


Fig. B1. Prediction performance of empirical ROP equations at different complexities.

Appendix C. Application of the DPODDPG Algorithm for Drilling Parameter Optimization

Figure C1 provides a detailed illustration of the changes in key drilling parameters, measurement parameters, and performance indicators during the drilling process in the depth interval of 865–870 m, reflecting the drilling response to softer formations and the optimization effects of the DPODDPG algorithm.

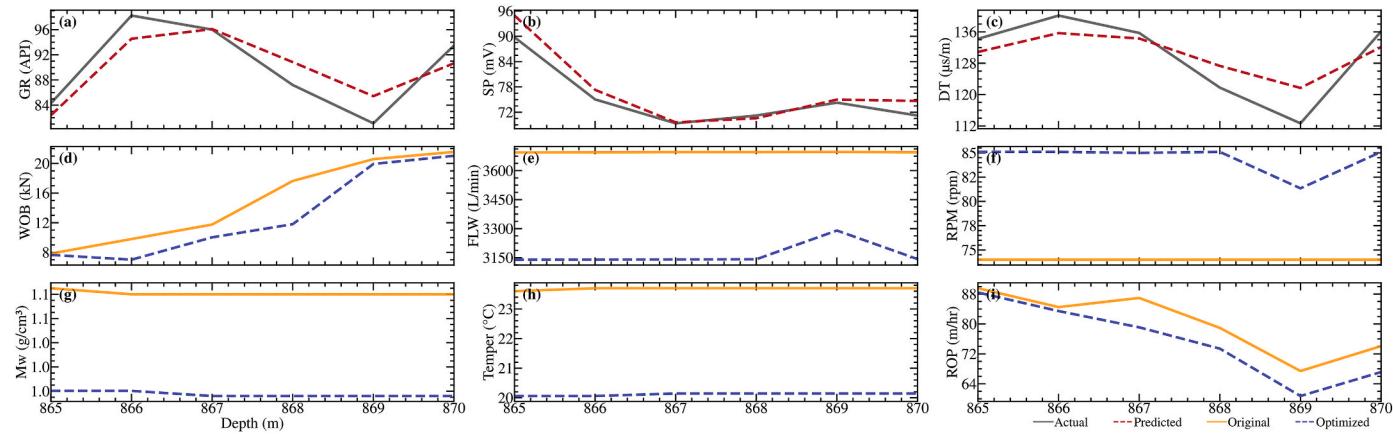


Fig. C1. Drilling performance and parameter optimization for softer formations at 865–870 m using DPODDPG algorithm.

Figure C1(a)–C1(c) show a significant decrease in GR, SP, and DT, indicating the encounter of softer formations. These geological features suggest lower strength and density of the formations, which typically necessitate adjustments in drilling parameters to accommodate the formation characteristics.

In such formations, the optimized drilling parameters include: lowering the WOB, reducing FLW, increasing the RPM, and decreasing MwIN and TemIN. Figure C1(d)–C1(f) illustrate the effects of these adjustments. Lowering the WOB helps to reduce excessive pressure on the softer formation, preventing rapid wear of the drill bit while maintaining good cutting efficiency. Reducing the pump flow rate minimizes the erosion of the borehole wall while maintaining effective circulation of the drilling fluid. Increasing the RPM enhances the cutting speed of the drill bit to suit the softer formation characteristics. Lowering the mud density and temperature helps to reduce pressure on the borehole wall, maintain mud performance, and decrease thermal impact on the drill bit. Figure C1(g) further shows changes in drilling performance indicators, including a decrease in ROP. Although the ROP decreases, this adjustment aims to optimize various drilling parameters to better suit the characteristics of softer formations. These adjustments reflect the DPODDPG algorithm's effectiveness in optimizing drilling parameters in softer formations, improving the stability and efficiency of the drilling process.

Appendix D. Identified Outliers Detected Using the IQR Method

Table D1 represents data points where the drilling parameters deviated significantly from the normal range, as determined by the IQR method. Identifying and analyzing these outliers is crucial for understanding unusual drilling conditions and improving the robustness of the DPODDPG algorithm.

Table D1
Identified outliers detected using the IQR method.

Depth (m)	WOB (kN)	RPM (rpm)	FLW (L/min)	MwIN (g/cm³)	TemIN (°C)	GR (API)	SP (mV)	DT (μs/m)	ROP (m/hr)
500	0.66	49	3236	1.05	0.3	74.30219	184.2584	137.6898	72.28916
501	0.69	49	3236	1.05	0.3	82.70431	172.6425	149.9065	65.21739
502	0.68	49	3237	1.05	0.3	76.80305	161.6131	147.5279	62.5
503	0.69	49	3237	1.05	0.3	77.5096	153.1939	141.6185	61.22449
504	0.98	49	3238	1.05	0.3	76.9362	148.8712	142.6763	61.22449
505	1.07	49	3237	1.05	0.3	73.56387	145.3887	147.0804	62.5
506	0.85	47	3257	1.05	0.3	81.80946	137.9991	148.2334	65.21739
507	0.75	45	3269	1.05	0.3	82.90694	134.5095	147.6974	98.36066
508	0.76	45	3269	1.05	0.3	79.90469	132.526	153.0588	103.4483
509	0.76	45	3269	1.05	0.3	85.96244	125.9774	153.6168	101.6949
510	0.78	45	3270	1.05	0.3	85.60663	120.7985	155.1229	111.1111
511	0.8	45	3270	1.05	0.3	87.27572	116.5187	161.3833	100
512	0.81	45	3270	1.05	0.3	85.02674	114.517	153.8124	105.2632
513	0.82	45	3270	1.05	0.3	74.03695	113.8675	152.8158	98.36066
514	0.8	45	3271	1.05	0.3	77.43804	110.2072	150.2156	105.2632
515	0.81	48	3694	1.05	0.3	77.37476	106.6841	146.1191	109.0909
2451	1.23	41	1560	1.35	47.8	92.92864	76.26921	94.51273	24.19355
2452	2.57	41	1562	1.35	50.3	79.82709	79.5157	93.44398	15.38462
2453	2.49	41	1563	1.35	46.3	78.2122	81.72555	92.01565	23.07692
2454	2.21	41	1565	1.35	46.3	71.82618	80.78772	91.67333	32.08556
2455	2.14	41	1565	1.35	44.8	64.19794	78.05274	90.63791	34.88372
2456	1.35	40	1583	1.34	49.7	74.09026	63.71694	92.28022	12.73885

(continued on next page)

Table D1 (continued)

Depth (m)	WOB (kN)	RPM (rpm)	FLW (L/min)	MwIN (g/cm ³)	TemIN (°C)	GR (API)	SP (mV)	DT (μs/m)	ROP (m/hr)
2457	1.7	40	1592	1.35	51.6	108.932	38.89785	93.00847	27.77778
2458	1.86	40	1592	1.35	51.8	105.331	27.17793	87.11817	36.58537
2835	2.18	46	1973	1.36	41.8	112.2487	50.98679	87.9058	28.43602
2836	2.58	46	1975	1.36	42.1	113.8092	50.21944	83.95281	24
2837	3.33	48	1974	1.36	42.4	111.3442	52.57258	84.50381	17.34104
2838	3.39	49	1974	1.36	31	109.7193	55.34517	84.54084	15.03759
2839	2.06	49	1975	1.36	45.9	114.8947	60.12277	84.63022	27.27273
2840	2.44	49	1976	1.36	45.9	105.8848	72.88673	88.97838	24.19355
2841	2.54	49	1974	1.36	45.1	109.3882	74.67691	87.5003	18.07229
2842	1.33	49	1973	1.36	44.2	110.0747	70.21753	84.00311	20.47782
2843	1.54	49	1975	1.36	43.9	95.33975	81.18654	87.90923	25.1046
2844	2.55	49	1976	1.36	43.6	98.32629	75.79719	83.28912	17.91045
2845	3.39	49	1975	1.36	43.7	104.355	61.55425	83.78344	22.30483

Appendix E. Detailed Mathematical Formulations and Derivations

Variational Mode Decomposition of LWD Parameters

The VMD decomposes data into several Intrinsic Mode Functions (IMFs), each characterized by a specific bandwidth and single frequency band, thus optimizing the subsequent data analysis and interpretation applicability. The modal decomposition process under variational constraint is expressed by Eq. (E1) (ur Rehman and Aftab, 2019):

$$\begin{cases} \min_{\{u_k, w_k\}} \left(\sum_K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right)^* u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right) \\ \text{subject to } \sum_k u_k(t) = f(t) \end{cases} \quad (\text{E1})$$

where $f(t)$ represents the original data, a function of the time series data t ; $u_k(t)$ denotes the k -th modal component; ∂_t indicates the partial derivative with respect to time t ; w_k is the corresponding center frequency, ensuring that each modal component oscillates around its main frequency; $\delta(t)$ represents the Dirac distribution, used to simulate the instantaneous pulse response in actual physical processes; and $*$ denotes the convolution operation, employed to simulate the temporal locality of the data.

In addressing the optimization challenges posed by constraints in the VMD method, a quadratic penalty factor α and Lagrange multipliers λ were introduced to transform the original variational problem into a more tractable unconstrained optimization problem (Nazari and Sakhaii, 2020). The solution process is effectively simplified by constructing an augmented Lagrangian function, while ensuring that the constraints are satisfied. The expression for the augmented Lagrangian function is given in Eq. (E2) (ur Rehman and Aftab, 2019):

$$\begin{aligned} \mathcal{L}(u_k, w_k, \lambda) = & \alpha \sum_K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right)^* u_k(t) \right] e^{-jw_k t} \right\|_2^2 + \lambda \left(\sum_K u_k(t) - f(t) \right) \\ & + \frac{\alpha}{2} \left(\sum_K u_k(t) - f(t) \right)^2 \end{aligned} \quad (\text{E2})$$

where α denotes the coefficient of the quadratic penalty term, adjusting the penalty strength for constraint violations, and λ is the Lagrange multiplier, ensuring that the sum of the modal components precisely reconstructs the original data $f(t)$. This optimization strategy is implemented through iterative updates of parameters $u_k(t)$, w_k , and λ , employing the Alternating Direction Method of Multipliers to enhance the efficiency and stability of handling optimization tasks with separable constraints (Lian et al., 2018). By progressively optimizing these parameters, the augmented Lagrangian method not only boosts computational efficiency but also ensures precise reconstruction of the original data during multimodal data decomposition.

One-Dimensional Convolutional Neural Network (1D-CNN)

A 1D-CNN has demonstrated exceptional feature extraction capabilities in analyzing LWD data, particularly in recognizing local dependencies and dynamic changes within time-series data (Gu et al., 2018). The data were entered into the network through an input layer that adjusted the format of the raw LWD data for subsequent processing. Successive convolutional layers expand the feature representation layer-by-layer by setting the kernel size and padding strategy, ensuring the integrity and continuity of the feature maps. Subsequently, a max-pooling layer reduces the dimensionality of the data, enhances the abstraction ability of the model, reduces the computational load, and improves the generalizability of the model. This stacked convolutional structure increases the sensitivity of the network to subtle variations in the data and builds a comprehensive feature library through a continuous feature extraction process, effectively facilitating the in-depth analysis of complex drilling data (Matinkia et al., 2022).

Bidirectional Long Short-Term Memory (BiLSTM) Network

A BiLSTM is an enhanced recurrent neural network structure specifically designed to capture both forward and backward dependencies in time-series data analysis (Van Houdt et al., 2020). BiLSTM implements two independent LSTM units at each time point, processing the forward and backward sequences of the data to capture the information comprehensively (Sampaio Descovi et al., 2023).

In the BiLSTM network, each direction's LSTM unit comprises forget, input, and output gates, which collectively regulate the update mechanism of the network's internal state. The forget gate filters out the information to be removed from the cell state, the input gate determines which new information must be stored in the cell state, and the candidate memory generates a potential state for this new information, potentially integrating it into the current cell state. The output gate manages the transformation of cell-state information to the output state. The network processes the input data in parallel using both forward and backward LSTM units, combining the outputs from both directions at each time step to form a comprehensive output. The mathematical expressions for the forget gate f_t^{fw} , input gate i_t^{fw} , output gate o_t^{fw} , and candidate memory \tilde{c}_t^{fw} of the forward long short-term memory units are presented in Eq. (E3) (Hochreiter and Schmidhuber, 1997):

$$\begin{cases} f_t^{fw} = \sigma(W_f^{fw} \cdot h_{t-1}^{fw} + U_f^{fw} \cdot x_t + b_f^{fw}) \\ i_t^{fw} = \sigma(W_i^{fw} \cdot h_{t-1}^{fw} + U_i^{fw} \cdot x_t + b_i^{fw}) \\ c_t^{fw} = \tanh(W_c^{fw} \cdot h_{t-1}^{fw} + U_c^{fw} \cdot x_t + b_c^{fw}) \\ o_t^{fw} = \sigma(W_o^{fw} \cdot h_{t-1}^{fw} + U_o^{fw} \cdot x_t + b_o^{fw}) \\ c_t^{fw} = f_t^{fw} \odot c_{t-1}^{fw} + i_t^{fw} \odot \tilde{c}_t^{fw} \\ h_t^{fw} = o_t^{fw} \odot \tanh(c_t^{fw}) \end{cases} \quad (\text{E3})$$

where h_{t-1}^{fw} represents the previous time step's hidden state; x_t is the input at the current time step; σ denotes the sigmoid activation function, which regulates the intensity of the gating signals to maintain their range between 0 and 1; \tanh , the hyperbolic tangent activation function, compresses the candidate memory values within the range of -1 to 1; W and U are weight matrices corresponding to the linear transformations of the hidden state and input data, respectively; b is the bias vector, used to adjust the outcome of the linear transformations; and \odot represents the Hadamard product, i.e., the element-wise multiplication of vectors.

The backward long short-term memory units function similarly to the forward LSTM units, but process data in reverse from the end of the sequence to its beginning. In this configuration, the state updates of the backward LSTM units depend on subsequent time-step data, enabling the model to process information from future to past. This mechanism provides the model with backward temporal dependencies, allowing it to capture the influence of future information on past states, thereby enhancing its understanding of and predictive capabilities for time-series data (Qifeng et al., 2021).

Multihead Attention (MA) Mechanism

The MA mechanism enhances the model's ability to analyze complex time-series data, particularly in capturing dependencies in LWD data through parallel input processing. With this mechanism, the input data undergo a series of independent linear transformations to generate multiple sets of query (Q), key (K), and value (V) vectors in different representation spaces. These transformations enable the model to explore data attributes across multiple abstraction levels simultaneously, with each attention head independently capturing specific data features and integrating them to form a comprehensive output. Subsequently, each head evaluated the relative importance of parts of the input data by calculating the dot products of Q and K. The scores were then normalized using the softmax function to form weights and combined with the V vectors to produce a weighted output. Before performing the attention operations, each head applied a linear transformation to the input data to optimize the data-processing workflow. The mathematical expression of the dot-product attention mechanism is given by Eq. (E4) (Shaw et al., 2018):

$$\begin{cases} Q_i = QW_i^Q, K_i = KW_i^K, V_i = VW_i^V \\ \text{Attention}(Q_i, K_i, V_i) = \text{softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V \end{cases} \quad (\text{E4})$$

where W_i^Q , W_i^K , and W_i^V are transformation matrices that map the input data to specific subspaces for the operation of that head, d_k is the dimension of the key vectors used to scale the dot products to avoid issues of gradient vanishing because of overly large computational results, and the outputs of all independent heads are concatenated and processed through an additional linear transformation to integrate into a final output vector, as illustrated in Eq. (E5) (Qin et al., 2023):

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O \quad (\text{E5})$$

where W^O is the output linear transformation matrix used to merge the outputs of all the heads into a unified representation.

Unit Cost per Foot Function

In drilling operations, the cost per foot function is a core indicator for assessing economic efficiency, focusing on accurately measuring costs and efficiency during drilling. The expression for cost per foot is given in Eq. (E6) (Wang et al., 2015):

$$C_{pm} = (C_b + C_r(t_t + t)) / H \quad (\text{E6})$$

where C_{pm} is the cost per foot, (CNY/m); C_b is the cost of the drill bit, (CNY); C_r is the rig operation fee (CNY/h); t is the drilling time, (hr); t_t is the time required to make trips and connections, (hr) and H is the total footage drilled, (m); The wear rate of the drill bit is given by Eq. (E7), illustrating the wear dynamics of the drill bit teeth.

$$\frac{dh}{dt} = \frac{A_f(a_1n + a_2n^3)}{Z_2 - Z_1W(1 + C_1h)} \quad (E7)$$

Eq. (E7) is transformed into a relationship between time dt and wear dh , as shown in Eq. (E8):

$$dt = \frac{(Z_2 - Z_1W)(1 + C_1h)}{A_f(a_1n + a_2n^3)} dh \quad (E8)$$

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

In conjunction with the previously mentioned wear rate equation, empirical ROP equations, $v_{pc}(w, n, q, k_d)$, developed using the MESR-DS algorithm, were employed to establish a mathematical model correlating the drill bit penetration H with the operational time t , as shown in Eq. (E9):

$$\begin{cases} H = \int_0^{h_f} v_{pc}(w, n, q, k_d) \cdot \frac{(Z_2 - Z_1W)(1 + C_1h)}{A_f(a_1n + a_2n^3)} dh \\ t = \int_0^{h_f} \frac{(Z_2 - Z_1W)}{A_f(a_1n + a_2n^3)} \left(h + \frac{C_1}{2}h^2 \right) dh \end{cases} \quad (E9)$$

where h_f denotes the amount of wear corresponding to the drill-bit lifespan. Consequently, the cost per foot function is reformulated as a function of the WOB and the RPM, as expressed in Eq. (E10):

$$C_{pm} = \frac{C_r \left[t_E \cdot A_f(a_1n + a_2n^3) + \left(h_f + \frac{C_1}{2}h_f^2 \right) \right]}{v_{pe}(w, n, q, k_d) \cdot \left(h_f + \frac{C_1}{2}h_f^2 \right)} \quad (E10)$$

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

MSE Function

The MSE model correlates the energy required to fracture a unit volume of rock with the rock-breaking efficiency of the drill bit and serves as a critical metric for quantifying drilling efficiency (Teale, 1965). The MSE is calculated using parameters such as the ROP, WOB, PRM, torque, and drill bit diameter. A higher MSE value indicates a lower drilling efficiency and poorer adaptability of the drill bit to the formation, suggesting that the drilling parameters require optimization. The ideal MSE calculation is given by Eq. (E11) (Dupriest and Koederitz, 2005):

$$MSE = \frac{4W}{\pi D_b^2} + \frac{480nT_b}{D_b^2 P} \quad (E11)$$

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

In practice, real torque values at the drill bit bottom are often unavailable, necessitating calculations using the drill bit's sliding friction coefficient and WOB. The torque during drilling can be expressed using a double integral, as depicted in Eq. (E12) (Guo et al., 2021):

$$T_b = \frac{1}{1000} \int_0^{D_b/2} \int_0^{2\pi} r^2 \frac{4\mu W}{\pi D_b^2} dr d\theta = \frac{\mu W D_b}{3000} \quad (E12)$$

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

Considering factors such as friction and vibrations, the actual energy utilization rate in drilling is typically between 30% and 40% (Dupriest and Koederitz, 2005). The required MSE is approximately three times the rock strength. Defining the effective energy utilization rate as E_f , and using the ROP empirical equations $v_{pc}(w, n, q, k_d)$ developed through the MESR-DS algorithm, the modified MSE model MSE_m is given in Eq. (E13):

$$MSE_m = E_f \left(\frac{4W}{\pi D_b^2} + 0.16 \frac{n\mu W}{D_b v_{pc}(w, n, q, k_d)} \right) \quad (E13)$$

Data availability

Data will be made available on request.

References

- Ahmed, O.S., Adeniran, A.A., Samsuri, A., 2019. Computational intelligence based prediction of drilling rate of penetration: a comparative study. *J. Petrol. Sci. Eng.* 172, 1–12.
- Arulkumaran, K., Deisenroth, M.P., Brundage, M., Bharath, A.A., 2017. Deep reinforcement learning: a brief survey. *IEEE Signal Process. Mag.* 34 (6), 26–38.

- Bäck, T., Schwefel, H.-P., 1993. An overview of evolutionary algorithms for parameter optimization. *Evol. Comput.* 1 (1), 1–23.
- Barbosa, L.F.F., Nascimento, A., Mathias, M.H., de Carvalho Jr, J.A., 2019. Machine learning methods applied to drilling rate of penetration prediction and optimization—A review. *J. Petrol. Sci. Eng.* 183, 106332.
- Borozdin, S., et al., 2020. Drilling problems forecast system based on neural network. In: SPE Annual Caspian Technical Conference. SPE, D023S013R008.
- Bourgoyne Jr, A.T., Young Jr, F., 1974. A multiple regression approach to optimal drilling and abnormal pressure detection. *Soc. Petrol. Eng. J.* 14 (4), 371–384.
- Buchinsky, M., 1998. Recent advances in quantile regression models: a practical guideline for empirical research. *J. Hum. Resour.* 38–126.
- Chen, X., et al., 2024. A real-time drilling parameters optimization method for offshore large-scale cluster extended reach drilling based on intelligent optimization algorithm and machine learning. *Ocean Engineering* 291, 116375.
- Coello, C.A.C., Montes, E.M., 2002. Constraint-handling in genetic algorithms through the use of dominance-based tournament selection. *Adv. Eng. Inf.* 16 (3), 193–203.
- Cranmer, M., 2023. Interpretable machine learning for science with PySR and SymbolicRegression. *jl. arXiv preprint arXiv:2305.01582*.
- Delahaye, D., Chaumatanan, S., Mongeau, M., 2019. Simulated annealing: from basics to applications. *Handbook of metaheuristics* 1–35.
- Delgado-Bonal, A., Marshak, A., 2019. Approximate entropy and sample entropy: a comprehensive tutorial. *Entropy* 21 (6), 541.
- Ding, W., Hou, S., Tian, S., Liang, S., Liu, D., 2023. A Bayesian optimized variational mode decomposition-based denoising method for measurement while drilling signal of down-the-hole drilling. *IEEE Trans. Instrum. Meas.* 72, 1–14.
- Dragomiretskiy, K., Zosso, D., 2013. Variational mode decomposition. *IEEE Trans. Signal Process.* 62 (3), 531–544.
- Duhamel, P., Vetterli, M., 1990. Fast Fourier transforms: a tutorial review and a state of the art. *Signal Process.* 19 (4), 259–299.
- Dupriest, F.E., Koederitz, W.L., 2005. Maximizing drill rates with real-time surveillance of mechanical specific energy. In: SPE/IADC Drilling Conference and Exhibition. SPE, SPE-92194-MS.
- Ebrahimi, N., Pflughoefl, K., Soofi, E.S., 1994. Two measures of sample entropy. *Stat. Probab. Lett.* 20 (3), 225–234.
- Eckel, J.R., 1967. Microbit studies of the effect of fluid properties and hydraulics on drilling rate. *J. Petrol. Technol.* 19 (4), 541–546.
- François-Lavet, V., Henderson, P., Islam, R., Bellemare, M.G., Pineau, J., 2018. An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning* 11 (3–4), 219–354.
- García, S., et al., 2015. Dealing with missing values. Data preprocessing in data mining 59–105.
- Gavidia, J.C.R., et al., 2024. Bridging the gap: integrating static and dynamic data for improved permeability modeling and super k zone detection in vuggy reservoirs. *Geoenergy Science and Engineering* 241, 213152.
- Govindan, R., Wilson, J., Eswaran, H., Lowery, C., Preißl, H., 2007. Revisiting sample entropy analysis. *Phys. Stat. Mech. Appl.* 376, 158–164.
- Gu, J., et al., 2018. Recent advances in convolutional neural networks. *Pattern Recogn.* 77, 354–377.
- Guo, H.J., Luo, H.d., Zhan, G.d., Wang, B.d., Zhu, S., 2021. A real-time friction prediction model for service drill string based on machine learning methods coupling with mechanical mechanism analysis. In: SPE Middle East Oil and Gas Show and Conference. SPE, D031S033R002.
- Guria, C., Goli, K.K., Pathak, A.K., 2014. Multi-objective optimization of oil well drilling using elitist non-dominated sorting genetic algorithm. *Petrol. Sci.* 11, 97–110.
- Harvey, A., Koopman, S.J., 2000. Signal extraction and the formulation of unobserved components models. *Econom. J.* 3 (1), 84–107.
- Hegde, C., Daigle, H., Gray, K.E., 2018. Performance comparison of algorithms for real-time rate-of-penetration optimization in drilling using data-driven models. *SPE J.* 23 (5), 1706–1722.
- Henderson, P., et al., 2018. Deep reinforcement learning that matters. In: Proceedings of the AAAI Conference on Artificial Intelligence.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780.
- Iversen, F., et al., 2006. Monitoring and control of drilling utilizing continuously updated process models. In: SPE/IADC Drilling Conference and Exhibition. SPE, SPE-99207-MS.
- Jahandideh, A., Jafarpour, B., 2020. Closed-loop stochastic oilfield optimization for hedging against geologic, development, and operation uncertainty. *Comput. Geosci.* 24 (1), 129–148.
- Kendall, H., Goins Jr, W., 1960. Design and operation of jet-bit programs for maximum hydraulic horsepower, impact force or jet velocity. *Transactions of the AIME* 219 (1), 238–250.
- Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of ICNN'95-international Conference on Neural Networks. IEEE, pp. 1942–1948.
- Khaleel, A.A., Adnan, M.S., Alhamd, S.J., 2021. Estimation of Bourgoyne and young model coefficients to predict optimum drilling rates and bit weights using genetic algorithms—a case study of the Fahia oil field in Iraq. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing, 012154.
- Kim, C., Han, K., 2020. Symbolic regression based on parallel Genetic Programming. *Journal of Digital Convergence* 18 (12), 481–488.
- Li, H., Wang, H., Wang, L., Zhou, X., 2020a. A modified Boltzmann Annealing Differential Evolution algorithm for inversion of directional resistivity logging-while-drilling measurements. *J. Petrol. Sci. Eng.* 188, 106916.
- Li, Y., She, L., Wen, L., Zhang, Q., 2020b. Sensitivity analysis of drilling parameters in rock rotary drilling process based on orthogonal test method. *Engineering geology* 270, 105576.
- Lian, J., Liu, Z., Wang, H., Dong, X., 2018. Adaptive variational mode decomposition method for signal processing based on mode characteristic. *Mech. Syst. Signal Process.* 107, 53–77.
- Liu, W., Cao, S., Chen, Y., 2016. Applications of variational mode decomposition in seismic time-frequency analysis. *Geophysics* 81 (5), V365–V378.
- Liu, Z., et al., 2023. Deep-sea rock mechanics and mining technology: state of the art and perspectives. *Int. J. Min. Sci. Technol.* 33 (9), 1083–1115.
- Loshchilov, I., Hutter, F., 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.
- Magnusson, L.V., Olsson, J.R., Tran, C.T.T., 2023. Recurrent neural networks for oil well event prediction. *IEEE Intell. Syst.* 38 (2), 73–80.
- Matias, J., de Castro Oliveira, J.P., Le Roux, G.A., Jäschke, J., 2021. Real-time optimization with persistent parameter adaptation applied to experimental rig. *IFAC-PapersOnLine* 54 (3), 475–480.
- Matinkia, M., et al., 2022. Developing a new model for drilling rate of penetration prediction using convolutional neural network. *Arabian J. Sci. Eng.* 47 (9), 11953–11985.
- McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5, 115–133.
- Mehrad, M., Bajolavid, M., Ramezanizadeh, A., Neycharan, J.G., 2020. Developing a new rigorous drilling rate prediction model using a machine learning technique. *J. Petrol. Sci. Eng.* 192, 107338.
- Mnih, V., et al., 2015. Human-level control through deep reinforcement learning. *nature* 518 (7540), 529–533.
- Mohammadizadeh, S., et al., 2023. Assessing cavitation erosion on solid surfaces using a cavitation jet apparatus. *HOLOS* 5 (39).
- Mohammadizadeh, S., Moghaddam, M.A., Talebbeydokhti, N., 2021. Analysis of flow in porous media using combined pressurized-free surface network, 24 (10), 1–15.
- Nazari, M., Sakhaii, S.M., 2020. Successive variational mode decomposition. *Signal Process.* 174, 107610.
- Osarogiagbon, A.U., Oloruntobi, O., Khan, F., Venkatesan, R., Butt, S., 2020. Gamma ray log generation from drilling parameters using deep learning. *J. Petrol. Sci. Eng.* 195, 107906.
- Prechelt, L., 2002. Early Stopping-But when?. *Neural Networks: Tricks of the Trade*. Springer, pp. 55–69.
- Qifeng, S., Na, L., Youxiang, D., Hongqiang, L., Haiquan, T., 2021. Logging-while-drilling formation dip interpretation based on long short-term memory. *Petrol. Explor. Dev.* 48 (4), 978–986.
- Qin, C., et al., 2023. Geological information prediction for shield machine using an enhanced multi-head self-attention convolution neural network with two-stage feature extraction. *Geosci. Front.* 14 (2), 101519.
- Rangel Gavidia, J.C., et al., 2023. Utilizing integrated artificial intelligence for characterizing mineralogy and facies in a pre-salt carbonate reservoir, Santos Basin, Brazil, using cores, wireline logs, and multi-mineral petrophysical evaluation. *Geoenergy Science and Engineering* 231, 212303.
- Regression, Q., 2017. *Handbook of Quantile Regression*. CRC Press, Boca Raton, FL, USA.
- Richman, J.S., Lake, D.E., Moorman, J.R., 2004. Sample Entropy, Methods in Enzymology. Elsevier, pp. 172–184.
- Sampaio Descovi, C., Carlos Zuffo, A., Mohammadizadeh, S., Murillo Bermúdez, L.F., Alfonso Sierra, D., 2023. Utilizing long short-term memory (LSTM) networks for river flow prediction in the BRAZILIAN pantanal basin. *HOLOS* 5 (39).
- Shaw, P., Uszkoreit, J., Vaswani, A., 2018. Self-attention with relative position representations. *arXiv preprint arXiv:1803.02155*.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 15 (1), 1929–1958.
- Sui, D., Nybo, R., Azizi, V., 2013. Real-time optimization of rate of penetration during drilling operation. In: 2013 10th IEEE International Conference on Control and Automation (ICCA). IEEE, pp. 357–362.
- Sutton, R.S., Barto, A.G., 2018. Reinforcement learning: An introduction.
- Tan, H., 2021. Reinforcement learning with deep deterministic policy gradient. In: 2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA). IEEE, pp. 82–85.
- Teale, R., 1965. The concept of specific energy in rock drilling. In: International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts. Elsevier, pp. 57–73.
- Udrescu, S.-M., Tegmark, M., 2020. AI Feynman: a physics-inspired method for symbolic regression. *Sci. Adv.* 6 (16), eaay2631.
- ur Rehman, N., Aftab, H., 2019. Multivariate variational mode decomposition. *IEEE Trans. Signal Process.* 67 (23), 6039–6052.
- Van Houdt, G., Mosquera, C., Nápoles, G., 2020. A review on the long short-term memory model. *Artif. Intell. Rev.* 53 (8), 5929–5955.
- Vásconez Garcia, R.G., et al., 2024. Geological insights from porosity analysis for sustainable development of santos basin's presalt carbonate reservoir. *Sustainability* 16 (13), 5730.
- Wan, X., Wang, W., Liu, J., Tong, T., 2014. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med. Res. Methodol.* 14, 1–13.
- Wang, D., et al., 2015. Technical economical analysis. The China Continental Scientific Drilling Project: CCSO-1 Well Drilling Engineering and Construction 349–359.
- Wang, Y., Gao, D., 2022. Study on the marine environment limiting conditions of deepwater drilling for natural gas hydrate. *Appl. Energy* 312, 118802.
- Waughman, R.J., Kenner, J.V., Moore, R.A., 2002. Real-time specific energy monitoring reveals drilling inefficiency and enhances the understanding of when to pull worn PDC bits. In: SPE/IADC Drilling Conference and Exhibition. SPE, SPE-74520-MS.

- Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research* 30 (1), 79–82.
- Wu, G., Mallipeddi, R., Suganthan, P.N., Wang, R., Chen, H., 2016. Differential evolution with multi-population based ensemble of mutation strategies. *Inf. Sci.* 329, 329–345.
- Wu, J., He, H., Peng, J., Li, Y., Li, Z., 2018. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. *Applied energy* 222, 799–811.
- Xi, Y., et al., 2023. Numerical simulation of rock-breaking and influence laws of dynamic load parameters during axial-torsional coupled impact drilling with a single PDC cutter. *Petrol. Sci.* 20 (3), 1806–1827.
- Xi, Y., Wang, W., Zha, C., Li, J., Liu, G., 2022. Numerical investigations on rock breaking mechanism and parameter influence of torsional percussive drilling with a single PDC cutter. *J. Petrol. Sci. Eng.* 210, 110077.
- Xue, Y.-J., Cao, J.-X., Wang, D.-X., Du, H.-K., Yao, Y., 2016. Application of the variational-mode decomposition for seismic time-frequency analysis. *IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens.* 9 (8), 3821–3831.
- Yan, X., Liu, Y., Zhang, W., Jia, M., Wang, X., 2020. Research on a novel improved adaptive variational mode decomposition method in rotor fault diagnosis. *Appl. Sci.* 10 (5), 1696.
- Yang, J., 2022. Jetting Installation Method of Oil-Gas Well Conductor, Installation Methods of Offshore Oil-Gas Well Conductor. Springer, pp. 107–162.
- Yimwadsana, B., 2023. Leveraging machine learning for estimating relationship model through empirical scientific data. In: 2023 27th International Computer Science and Engineering Conference (ICSEC). IEEE, pp. 358–361.
- Yun, X., Jian, R., 2020. Features method for selecting VMD parameters based on spectrum without modal overlap. *J. Phys. Conf.*, 012002. IOP Publishing.
- Zeiler, M.D., 2012. Adadelta: an adaptive learning rate method. *arXiv preprint arXiv: 1212.5701*.
- Zha, C., et al., 2017. Combined percussive-rotary drilling to increase rate of penetration and life of drill bit in drilling hard rock formation. *Chem. Technol. Fuels Oils* 53, 254–262.
- Zhang, Q.-Z., Wei-Xiao, L., Sha, L.-X., 2017. Drilling parameters optimization based on chaotic multi-objective particle swarm optimization algorithm. In: 2017 International Conference on Electrical Engineering and Automation Control. ICEEAC.
- Zhang, Y., Tino, P., Leonardis, A., Tang, K., 2021. A survey on neural network interpretability. *IEEE Transactions on Emerging Topics in Computational Intelligence* 5 (5), 726–742.
- Zhang, Y.X., et al., 2011. Experimental study on the influence of nozzle diameter on abrasive jet cutting performance. *Adv. Mater. Res.* 337, 466–469.
- Zheng, J., Li, Z.L., Dou, B., Lu, C., 2019. Multi-objective cellular particle swarm optimization and RBF for drilling parameters optimization. *Math. Biosci. Eng.: MBE* 16 (3), 1258–1279.
- Zurek, S., Guzik, P., Pawlak, S., Kosmider, M., Piskorski, J., 2012. On the relation between correlation dimension, approximate entropy and sample entropy parameters, and a fast algorithm for their calculation. *Phys. Stat. Mech. Appl.* 391 (24), 6601–6610.