Petroleum Science 23 (2026) 5524–5542

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Formation drillability characterization method based on multi-scale convolutional neural representation with stacking ensemble

Jian-Sheng Liu [a,b], Hua-Lin Liao [a,b,*], Zhe Huang [c], Feng-Tao Qu [d], Fang Shi [a,b], Tian-Yu Wu [a,b]

[a] State Key Laboratory of Deep Oil and Gas, China University of Petroleum (East China), Qingdao, 266580, Shandong, China
[b] School of Petroleum Engineering, China University of Petroleum (East China), Qingdao, 266580, Shandong, China
[c] Drilling Technology Research Institute, Sinopec Shengli Oilfield Service Corporation, Dongying, 257099, Shandong, China
[d] College of Petroleum Engineering, Xi'an Shiyou University, Xi'an, 710065, Shaanxi, China

ARTICLE INFO

Article history:
Received 2 August 2025
Received in revised form
6 November 2025
Accepted 29 April 2026
Available online 6 May 2026

Handling editor: Xian-Zhi Song
Edited by Jia-Jia Fei

Keywords:
Geomechanics
Rock drillability characteristics
Data-driven
Multi-scale convolutional neural network
Stacking ensemble

ABSTRACT

Formation drillability characteristics are critical for optimizing drilling performance and designing bottom-hole assemblies. Accurate estimation of drillability parameters, including uniaxial compressive strength (σ_c), drillability grade (K_d), and abrasiveness index (I_a), is fundamental for safe and efficient drilling operations. To overcome the generalization constraints of conventional single models and their inability to capture multi-scale feature responses, this study proposes a novel data-driven framework, termed the multi-scale convolutional neural representation with stacking ensemble (MSCNR-SE), for accurately estimating σ_c, K_d, and I_a. The framework uses a three-stage strategy: (1) a multi-scale convolutional module extracts features across different spatial scales; (2) an ensemble of diverse base regressors (e.g., random forest (RF), gradient boosted decision tree (GBDT)) enhances representational capacity; and (3) extreme gradient boosting (XGBoost) is employed as a meta-learner to integrate the outputs of base regressors, thereby mitigating the risk of model-specific bias. This hierarchical deep-learning ensemble effectively captures complex nonlinear relationships intrinsic to formation properties. Experimental results show that MSCNR-SE accurately models depth-aligned drillability profiles, with strong agreement between predictions and measurements (the coefficient of determination, R^2 > 0.93 on the blind testing sequence). The MSCNR-SE model achieved superior predictive performance compared with other models across all three parameters. Specifically, MSCNR-SE reduced the mean absolute error (MAE) by approximately 3%–50%, the root mean squared error (RMSE) by 4%–64%, the mean absolute percentage error (MAPE) by 3%–59%, and increased the R^2 by 1%–32%. These results highlight the superior predictive accuracy and robustness of MSCNR-SE, underscoring its potential as a practical, highly efficient data-driven solution for real-time parameter estimation, monitoring, and control in complex drilling.

1. Introduction

In recent years, with the advancement of deep and ultra-deep well drilling technologies, increasingly complex lithological formations have been encountered. However, the insufficient understanding of the formation drillability characteristics leads to the improper selection and application of rock-breaking tools and operational parameters, thereby resulting in reduced rate of penetration (ROP) (Ahmednour et al., 2025; Bai et al., 2025; Allawi et al., 2025; Shao et al., 2024; Yang et al., 2025a; Liu et al., 2025a). Formation drillability is a comprehensive indicator that characterizes the resistance of formations to bit-induced fragmentation (Guan et al., 2023; Oliver et al., 2013; Wang et al., 2024; Liu et al., 2025b). Drillability-related parameters encompass various physical and mechanical properties of the rock, such as uniaxial compressive strength (σ_c) (Sabri et al., 2025), drillability grade (K_d) (Shi et al., 2023; Yang et al., 2025), and abrasiveness index (I_a)

* Corresponding author.
E-mail address: liaohualin2003@upc.edu.cn (H.-L. Liao).

Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.04.054
1995-8226/

(Sharma et al., 2025; Hamzaban et al., 2025). Accurate estimation of these parameters serves as a critical basis for the selection of drilling tools for ROP enhancement, optimization of drilling parameters, and ROP prediction in oil and gas well operations.

There are three main methods for obtaining drillability characteristic parameters of geological formations: laboratory testing, mathematical statistics, and data-driven models. Laboratory tests mainly rely on rock mechanics standard tests (Asante-Okyere et al., 2023), such as drillability tests (Yang et al., 2025b), compressive strength tests (Li et al., 2024), cutter wear tests (Ke et al., 2023), etc., to obtain accurate results of formation drillability characteristic parameters through high-quality testing. Empirical correlations construct physical estimation models for key parameters (Heydari et al., 2025; Giro et al., 2025; Yu et al., 2025), including rock dynamic mechanical parameters, formation hardness, formation shear strength, rock internal friction angle, and bit drillability grade, and then calculate corresponding drillability characteristic parameters based on these formulas (Xiong et al., 2023). Data-driven methods mainly take seismic attributes and logging data as inputs, and utilize machine learning and deep learning algorithms to construct predictive models for formation drillability parameters (Kor et al., 2021; Ehsan et al., 2025; Zhang et al., 2025; Davoodi et al., 2023; Wang et al., 2025). Although the above methods can all obtain the drillability characteristic parameters of the formation, there are certain limitations: (1) The laboratory determination of geomechanical parameters through core testing is inherently destructive, cost-intensive, and yields only discrete data points, which limits its applicability for continuous formation characterization. (2) Mathematical statistics normally consider limited factors, resulting in deviations between the calculated results and actual values. (3) Although machine learning methods have achieved partial recognition of the drillability characteristics of geological formations, the generalization ability of individual models is often weak, the prediction risk is high, and the prediction effect is uneven.

Stacking generalization is an ensemble method that combines the advantages of multiple models through a hierarchical structure (Sharma et al., 2023; Feng et al., 2023; Gan et al., 2025). Its core principle is divided into two layers: In the first layer, multiple heterogeneous base models are trained independently to generate preliminary predictions; in the second layer, these predictions are used as new features for the meta-learner (Yu et al., 2025; Ghanitoos et al., 2024; Wang et al., 2021). The meta-model learns the inherent rules of the new features and forms the optimal combination strategy to output the final prediction (Mahmood et al., 2022; Pacis et al., 2023). Its advantage lies in the ability to effectively integrate the diversity of different models (such as complementary linear and nonlinear models), flexibly adapt to complex tasks, and significantly improve the prediction accuracy of target values (Koopialipoor et al., 2022; Srivastava and Vemavarapu, 2021).

To address the problem of insufficient attention to the relationship between surface drilling engineering parameters and formation resistance at the drill bit location in existing research, and the difficulty in accurately modeling the quantitative relationship between them, this paper proposes a multi-scale convolutional neural representation and stacking ensemble (MSCNR-SE) prediction method for formation drillability characteristic parameters. This method uses a multi-scale convolutional neural network (MSCNN) to extract and compress spatiotemporal features of complex depth sequences and high-dimensional drilling parameters, capturing the variation patterns of formation response at different scales. Subsequently, a stacking ensemble learning framework was constructed, using decision tree regression (DT), linear regression (LR), random forest (RF), gradient boosting tree (GBDT), support vector regression (SVR), multi-layer perceptron (MLP), and K-nearest neighbor regression (KNN) as heterogeneous base learners, and extreme gradient boosting model (XGBoost) as meta-learner, to achieve high-precision prediction of drillability characteristic parameters such as uniaxial compressive strength, drillability grade, and abrasiveness index of geological formations. This method fully integrates the advantages of deep learning in feature modeling and the ability of ensemble learning to improve model generalization, providing a feasible technical path for modeling the drillability characteristics of formations. It also provides important data support and decision-making basis for optimizing drilling parameters, selecting acceleration tools, and predicting drilling speed.

2. Methodology

2.1. Multi-scale convolutional neural representation and stacking ensemble

This study proposes a method that integrates deep learning and ensemble learning, leveraging the advantages of deep learning in feature extraction and the performance of ensemble learning in nonlinear modeling to achieve more accurate and robust prediction results. The method architecture is shown in Fig. 1.

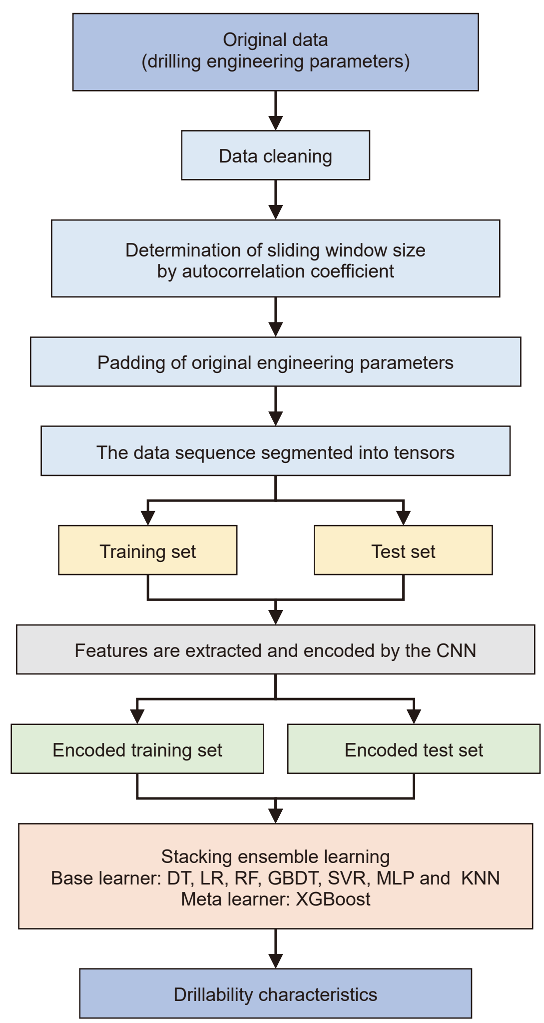

Fig. 1. Working flow of MSCNR-SE predicting drillability characteristics.

The optimal sliding window size is determined according to the autocorrelation characteristics of formation drillability parameters and drilling engineering parameter sequences along the depth axis. Subsequently, a sliding window is used to perform edge filling and tensor segmentation on the original drilling engineering parameter sequence, ensuring the integrity of the data at the beginning of the sequence. The engineering parameters with depth as the sequence axis were transformed into structured modeling samples, which were divided into training and testing sets. Each input tensor is then subjected to feature extraction and encoding by MSCNN to generate one-dimensional feature vectors, fully capturing local variation patterns in drilling engineering parameters and enhancing the diversity and robustness of feature expression. In the supervised learning stage, this study selects multiple regression models (including DT, LR, RF, GBDT, SVR, MLP, KNN) to model the extracted features, and learns the complex nonlinear mapping relationship between formation drillability characteristic parameters (including σ_c, K_d and I_a) and input features from different perspectives.

To further improve the generalization ability and prediction accuracy of the model, the stacking ensemble learning strategy is adopted to use the above regression models as the base learners and XGBoost as the meta-learner, to fuse and optimize the outputs of each base model (Alvarado-Pérez et al., 2025). This framework fully combines the advantages of deep neural network in automatic feature extraction and the ability of integrated learning in multi-model fusion and generalization, effectively alleviates the overfitting or underfitting problems that may occur in a single model, thus realizing accurate and stable prediction of formation drillability characteristic parameters.

2.2. Sequence autocorrelation analysis and sliding segmentation

(1) Autocorrelation coefficient

The selection of sliding window size is crucial in the data preprocessing stage, as its length not only needs to consider the physical meaning of geological features and the acceptable range of model computing resources, but also directly affects the ability of subsequent feature extraction modules (such as multi-scale convolutional neural networks) to capture local patterns (Sharma and Bandhu, 2024). The window size determines the range of geological information covered by each input tensor, thereby affecting the learning effect of the model on key structures such as stratigraphic units, sedimentary joints, and geological mutations. The appropriate window length should be able to characterize the spatial scale of typical stratigraphic thickness or sedimentary units, ensuring that the input sequence contains representative geological information. If the window is too large, it may introduce redundant information and increase the computational burden of the model; If it is too small, it may cause incomplete expression of geological response characteristics and reduce the discriminative ability of the model. Therefore, the window length should balance model efficiency and feature representation while preserving information integrity, aligning the data with the requirements of subsequent deep learning and ensemble modeling.

ACF is widely used in the selection of the sliding window size. ACF can quantify the correlation between logging data at different depth intervals and identify potential periodicity or significant correlation intervals by characterizing the dependencies between data. By analyzing ACF values, a sliding window length that can reflect geological structure characteristics and facilitate model learning can be selected to improve the training effectiveness and predictive ability of the model (Qu et al., 2024). Its mathematical model is:

$$\widehat{\rho}_k=\frac{\sum\limits_{d=1}^{n-k}(P_d-\overline{P})(P_{d+k}-\overline{P})}{\sum\limits_{d=1}^{n-k}(P_d-\overline{P})^2} \tag{1}$$

In the above equation, P_d represents the engineering data at depth d, and k denotes the depth interval, which refers to the length of the sliding window. $\overline{P}$ is the mean value of the corresponding data. $\widehat{\rho}_k$ is the autocorrelation coefficient of the engineering data with a lag of k, taking values within the range of $[-1,1]$. A coefficient value closer to 1 indicates a strong positive correlation of the engineering data over the interval k, whereas a value closer to -1 suggests a strong negative correlation. A value of 0 implies no correlation between the data at the given depth interval.

In the proposed formation drillability parameter prediction method based on drilling engineering data, the model input is constructed in the form of tensors extracted from continuous sequences of drilling parameters using a sliding window mechanism. During the data segmentation process, the movement of the sliding window may result in insufficient historical information at the beginning of some sequences, making it impossible to form complete input windows. Consequently, some input features cannot be effectively matched with corresponding label values, thereby reducing the number of valid training samples and potentially compromising model robustness.

To mitigate this issue and retain as many available formation drillability parameter labels as possible, a one-sided padding strategy is introduced during the data preprocessing stage. Specifically, zero-padding (or equivalent value padding) is applied only at the initial positions of the raw drilling parameter sequences to ensure complete input tensors are generated. The detailed procedure is illustrated in Fig. 2.

Compared with two-sided or tail-end padding, one-sided (pre-window) padding is more reasonable and better aligned with engineering practice. In drilling operations, current engineering parameters are not only influenced by the geological characteristics of the formation at the bit location, but also significantly depend on the historical evolution of drilling parameters from the preceding section. This implies that preserving and leveraging early-stage variations in engineering parameters is crucial for accurately characterizing formation drillability behavior.

Moreover, in practical drilling scenarios, mechanical responses captured at the surface often lag behind actual changes in formation properties. This delay results in formation characteristics having a more substantial impact on future engineering parameters than on preceding ones. Therefore, from a modeling perspective, compensating for missing information at the beginning of the sequence via forward pre-sequence padding is more consistent with the physical interaction mechanisms between the drill bit and the formation than pre-filling at the sequence end.

The padding values can be selected using various strategies such as replicating the initial value, applying the sequence mean, or inserting zeros. In this study, a zero-padding strategy is adopted to neutralize the impact of the padded values and avoid introducing bias into the model training process.

As illustrated in Fig. 2, the padding and slicing process is carefully designed based on the results of autocorrelation coefficient analysis to determine the optimal sliding window size. Combined with a predefined step length, the required number of padded samples is calculated. Specifically, when the sliding window size is set to 3 and the step size is 1, two placeholder values

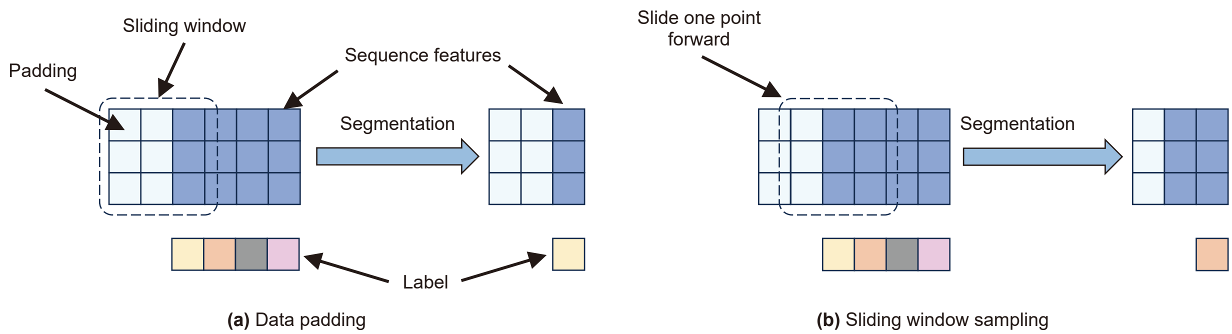

Fig. 2. Single edge padding and sliding segmentation of data sequence.

are needed at the start of the sequence to ensure the completeness of the first valid input window.

2.3. Multi-scale convolutional neural network

Convolutional neural networks (CNNs), as a leading architecture in the field of deep learning, have demonstrated significant advantages in handling complex time-series signals and extracting features from high-dimensional data. In drillability prediction, CNNs are capable of capturing both local and global spatiotemporal features from the input data through multiple convolutional layers. These layers progressively extract hierarchical features—from low-level to high-level representations—which effectively enhance the model's ability to recognize patterns within the data. Additionally, the integration of pooling layers allows the model to preserve key information while reducing feature dimensionality, thereby decreasing computational complexity and mitigating the risk of overfitting.

MSCNNs further enhance feature extraction by applying convolutional kernels of various sizes in parallel to the same input data. This architecture enables the model to capture both fine-grained local details and broader global structures across different scales. Compared to traditional single-scale convolutional architectures, multi-scale convolution is more adept at handling features with diverse spatial resolutions, which is particularly beneficial for formation drillability prediction, where geological structures may vary significantly in scale. By employing multiple kernel sizes, the network performs a more granular analysis of the input data, facilitating improved feature representation and enhancing the robustness of the model.

The multi-scale convolutional architecture proposed in this study is illustrated in Fig. 3. The figure depicts a dual-branch convolutional structure in which each branch utilizes convolutional kernels of different sizes—specifically, 2×2 and 3×3—for feature extraction. This design not only enables simultaneous multi-scale information capture but also promotes effective feature fusion across scales, thereby further improving the overall performance of the model.

(1) Convolution operation

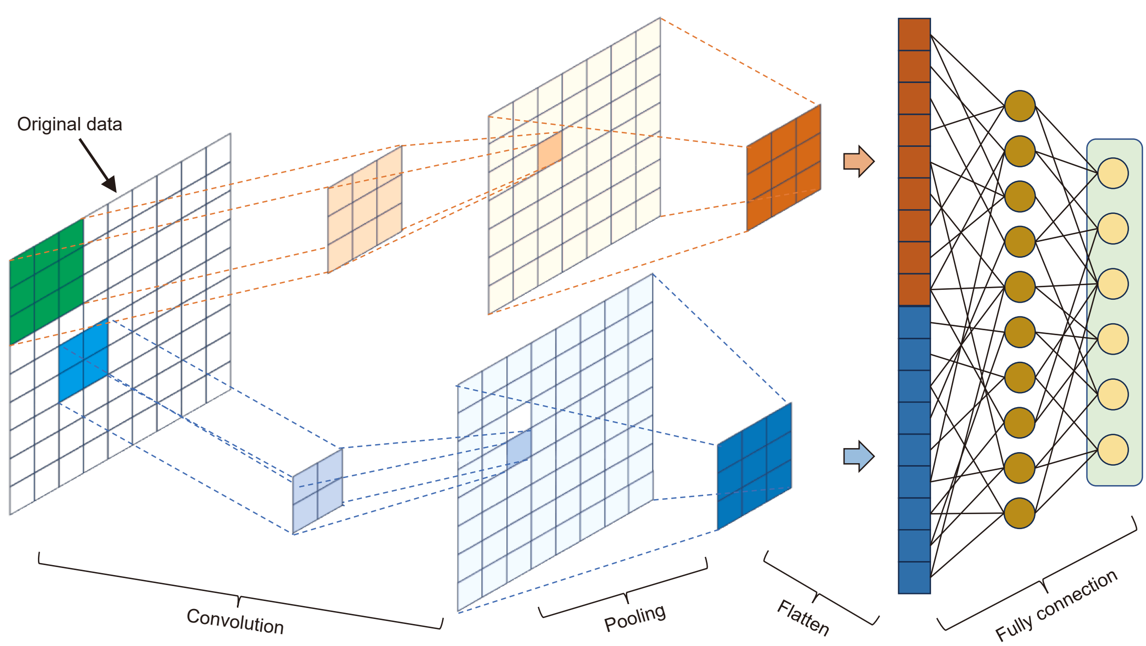

Fig. 3. Feature extraction network structure based on multi-scale convolution operation.

The convolution operation extracts local features from the input signal (e.g., logging time-series data or feature maps from previous layers) by sliding convolutional kernels across the input. The mathematical formulation of the convolution operation is given by:

$$(I*K)_{(i,j)} = \sum_{m=1}^{k} \sum_{n=1}^{k} I_{(i+m,j+n)} * K_{(m,n)} + b_j \tag{2}$$

where I denotes the input matrix, K represents the convolutional kernel, k is the width and height of the square kernel, $I*K_{(i,j)}$ is the output element at position, (m,n), b_j is the bias term, and * indicates the convolution operation.

The convolutional output is passed through a ReLU (rectified linear unit) activation function to introduce non-linearity and enable sparse activation, thereby enhancing the network's computational efficiency and representational power. The ReLU function is defined as:

$$Y_{(i,j)} = \text{ReLU}\left[(X*K)_{(i,j)}\right] = \max\left[0, (X*K)_{(i,j)}\right] \tag{3}$$

where $Y_{(i,j)}$ is an element of the convolutional output matrix.

(2) Pooling operation

Pooling layers reduce the spatial dimensionality and computational complexity of the feature maps while preserving critical information. They also enhance model robustness and reduce the risk of overfitting. The two most common pooling methods are max pooling and average pooling, defined as:

$$P_{(k,l)} = \begin{cases} \max_{(e,f)\in W} \boldsymbol{X}_{(k\cdot s+e,l\cdot s+f)} \\ \dfrac{1}{|W|} \sum_{(e,f)\in W} \boldsymbol{X}_{(k\cdot s+e,l\cdot s+f)} \end{cases} \tag{4}$$

where $P_{(k,l)}$ is the output of the pooling operation; $\boldsymbol{X}$ is the input matrix; W is the size of pooling window; s denotes the sliding stride of the pooling window; $|W|$ is the number of elements within each pooling window; e and f denote the width and height of the sliding pooling window respectively; k and l denote the row index and column index of the target element in the output pooled feature map respectively.

(3) Flattening and concatenation

Since the output from pooling layers is typically a multi-dimensional matrix, it needs to be flattened into a one-dimensional vector before being fed into the fully connected layers. The flattening process is defined as:

$$V = \text{Flatten}(P) \tag{5}$$

where P denotes the multi-dimensional feature map output from the pooling layer, and V represents the flattened vector.

The pooled feature vectors from the two branches, V_1 and V_2, are concatenated along the channel dimension to obtain the merged feature vector V_{merged}, expressed as:

$$V_{\text{merged}} = \text{Concatenate}(V_1, V_2) \tag{6}$$

(4) Fully connected operation

The fully connected layer is typically positioned at the end of the network, responsible for integrating and transforming all high-level features extracted by preceding convolutional and pooling layers into a new feature representation for subsequent classification or regression tasks. In this study, a fully connected layer is employed, and its mathematical formulation is expressed as:

$$y_g = \sum_{h=1}^{n} \omega_{gh} \cdot x_h + b_g \tag{7}$$

where y_g denotes the output of the g-th neuron; x_h represents the input from the h-th neuron; ω_{gh} is the weight connecting the h-th input neuron to the g-th output neuron; b_g is the bias term associated with the g-th neuron. The input to the first fully connected layer is the vector V_{merged}.

In the prediction of formation drillability characteristics, the spatial heterogeneity of rock properties—such as vertical anisotropy varying nonlinearly with depth and horizontal anisotropy influenced by well deviation and azimuth—plays a critical role in model accuracy. MSCNN addresses this complexity by employing convolutional kernels of multiple scales, enabling the simultaneous extraction of spatial features across different resolutions from the input data. This multiscale approach allows the model to develop a comprehensive understanding of formation property distributions. Compared with conventional single-scale convolutional methods, MSCNN demonstrates superior capability in capturing data patterns that vary across different spatial scales, thereby improving both the accuracy and robustness of feature extraction. Furthermore, MSCNN exhibits strong noise resistance, enabling it to extract meaningful information even from noisy or incomplete datasets, and effectively mitigating the risk of overfitting. These advantages make MSCNN particularly well-suited for modeling the complex and nonlinear relationships inherent in geological drilling data.

2.4. Stacking generalization

Due to inherent differences among regression models in terms of their response mechanisms to engineering parameters, generalization capabilities, and training efficiency, it is challenging for a single model to fully capture the complex nonlinear relationships among multidimensional variables in formation drillability prediction tasks. To address this limitation, a stacking ensemble learning framework is adopted in this study, integrating multiple base learners to leverage their respective strengths across different feature subspaces. This approach constructs a stronger predictive model compared to individual learners, thereby enhancing overall prediction accuracy, generalization ability, and model reliability.

The basic architecture of the stacking ensemble method is illustrated in Fig. 4. The original dataset is first partitioned into several subsets and fed into multiple base learners at the first layer, each generating its own prediction output. These first-layer outputs are subsequently used as input for a second-layer meta-learner, which is trained to produce the final prediction. If any individual base learner exhibits substantial prediction bias, the meta-learner effectively generalizes and corrects such deviations by integrating the outputs from multiple base models, leading to improved overall predictive performance.

Specifically, for a dataset $Q = \{(y_n, x_n), n = 1, \ldots, N\}$ containing N samples, where x_n is the feature vector of the n-th sample and y_n is the corresponding target value. j represents the number of features contained in each feature vector, i.e., $x_n = (x_1, x_2, \ldots, x_j)$. The data is randomly divided into K equally sized subsets $Q_1, Q_2, \ldots, Q_k$ using K-fold cross-validation. Define $Q_{-k} = Q - Q_k$, where Q_k is the testing sets and Q_{-k} is the training set of the k-th fold. For the first-level predictive models, K base learners are trained. Each base learner L_k ($k = 1, 2, \ldots, K$) is trained on the training set Q_{-k}.

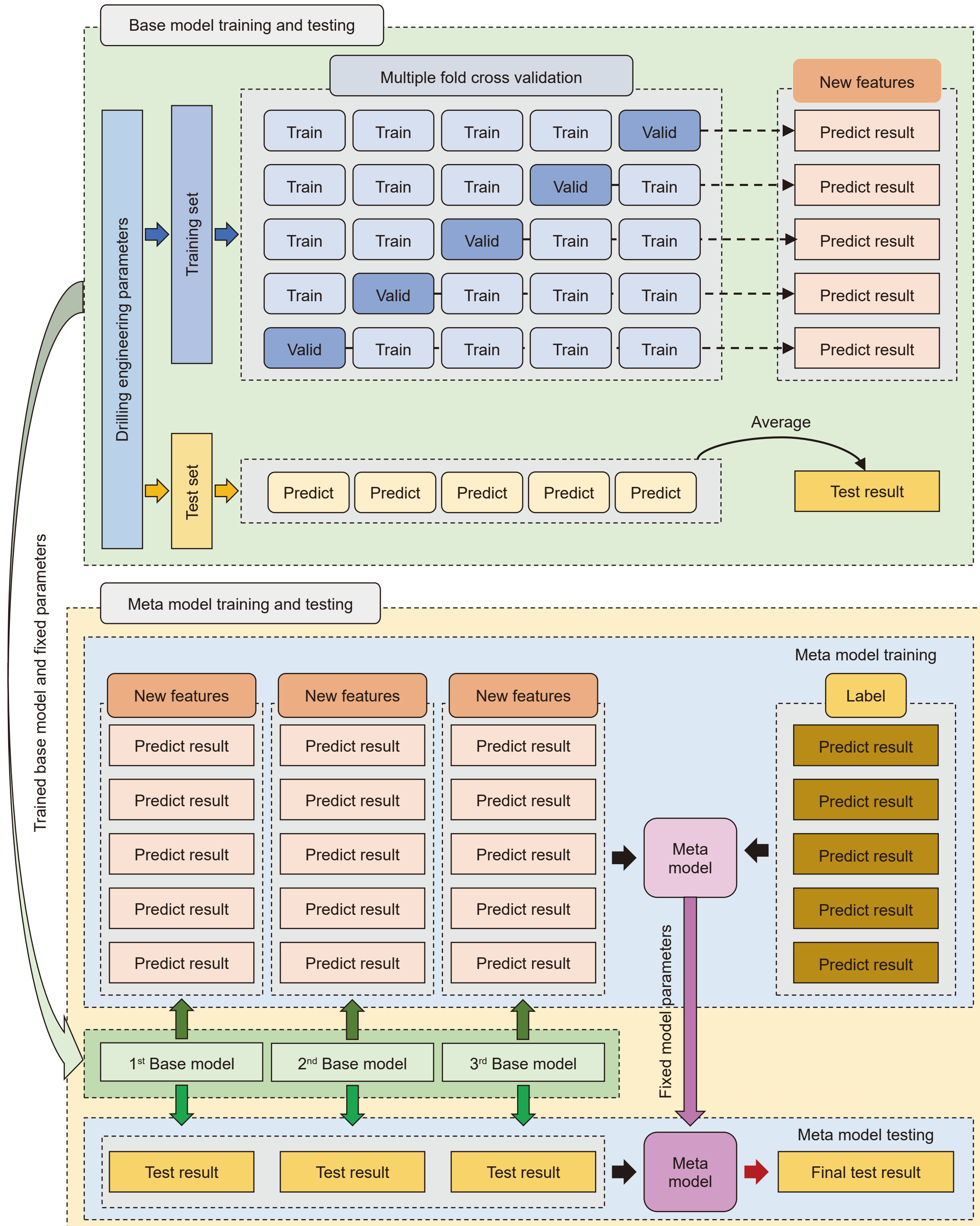

Fig. 4. Basic architecture of the stacking ensemble learning method.

The prediction by the base learner L_k on the feature vector x_n in the k-th test set Q_k is denoted as z_{kn}. After training all K base learners, their output results are aggregated into a new dataset, forming the meta-dataset Q_{new}= {(y_n, z_{1n}, …, z_{kn}), n= 1, 2, …, N}, which is used as the input for the second-level learning. The meta-learner L_{new} is trained on Q_{new}. The pseudocode for the stacking ensemble learning training process is as follows:

Input: Dataset Q= {(y_n, x_n), n= 1, 2, …, N}

Step 1: Divide the data into K roughly equal subsets Q_1, Q_2, …, Q_k, where Q_{-k}=Q − Q_k.

Step 2: Train the first-level K base learners
for k= 1 to K:
Train the first-level base learner L_k based on dataset Q_{-k}
end
Step 3: Construct a new dataset
Q_{new}= {(y_n, z_{1n}, …, z_{kn}), n= 1, 2, …, N}
Step 4: Train the second-level model L_{new} based on Q_{new}.

Output: Stacking ensemble learning model predicting formation drillability parameters.

It is noteworthy that the training set for the meta-learner is generated from the outputs of the base learners. Directly using the training set of the base learners to generate the meta-training set may cause the model to learn duplicated information, leading to severe overfitting. To mitigate the risk of overfitting, a reasonable data partitioning strategy must be applied. Based on the selected seven base learners, the original training dataset is divided into seven subsets according to time, ensuring that each subset has distinct sample IDs without overlap. For each base learner, one subset is used as the validation set and the remaining four as the training set. Each base learner outputs prediction results for its corresponding validation subset, and these seven prediction results are merged into a new dataset, which has the same size as the original dataset. Since the data blocks used for prediction are not involved in the training of the corresponding base learners, this configuration ensures that each data point is only used once for out-of-fold prediction, effectively reducing the risk of overfitting.

2.5. Multi-scale convolutional neural representation and stacking ensemble learning framework

This study proposes a stacking ensemble learning framework that integrates deep feature extraction and ensemble modeling, aiming to enhance the prediction accuracy and robustness of formation drillability parameters. In this architecture, the first layer consists of seven representative and complementary base regressors: DT, LR, RF, GBDT, SVR, MLP, and KNN. This ensemble encompasses a diverse range of algorithmic paradigms, including linear models, tree-based ensembles, kernel-based methods, deep neural networks, and instance-based learning. Such diversity ensures broad algorithmic coverage and strong model complementarity, enabling the framework to capture complex nonlinear relationships among input variables from multiple structural and theoretical perspectives, thereby improving the representation of geological engineering data.

In the second layer of the stacking architecture, XGBoost is employed as the meta-learner due to its strong generalization capabilities and high computational efficiency. As a gradient boosting-based ensemble algorithm, XGBoost integrates regularization, automatic feature selection, and robust handling of missing values. These features help suppress overfitting while improving model performance and training speed. When used as a meta-learner, XGBoost effectively aggregates and relearns the outputs of base learners by discovering latent correlations among their predictions, thereby enhancing the overall generalization ability and predictive accuracy of the ensemble. The complete model architecture is illustrated in Fig. 5, demonstrating strong adaptability, scalability, and practical applicability.

The proposed ensemble framework consists of the following key modules.

Step 1: Data input layer.

Raw drilling engineering parameters and formation drillability indices are first subjected to preprocessing, including standardization, missing value imputation, and outlier correction, to improve data quality and ensure consistency for subsequent modeling.

Step 2: MSCNN-based feature extraction and encoding.

MSCNN is employed to extract both local and global features using convolutional kernels of different sizes. Through a series of convolution and pooling operations, MSCNN captures the underlying spatial structure and nonlinear dependencies in the data while reducing feature dimensionality and preserving critical information. The extracted features are further processed through fully connected layers to encode the original input tensor into a set of discriminative deep latent features, which are then fed into the ensemble learning module.

Step 3: Base-learner training.

The deep features generated by MSCNN are independently input into multiple regression models to form the first layer of the ensemble. Each base learner applies its unique modeling strategy to fit and predict the data, enhancing the overall adaptability and representational richness of the model across diverse feature patterns.

Step 4: Stacked generalization layer (stacking fusion).

The prediction outputs from all base learners are concatenated to form a new training set, which is input into the XGBoost meta-learner for ensemble modeling. The meta-learner automatically learns the differential performance of the base models and adaptively weights and fuses their outputs, yielding a final prediction that is both more accurate and more stable.

This integrated framework leverages MSCNN's strength in data-driven feature extraction and the generalization advantages of ensemble learning, achieving a seamless unification of feature extraction, model training, and output fusion. Compared to traditional single-model approaches, this method not only reduces the dependence on manual feature engineering but also captures more comprehensive predictive information from multiple algorithms. It effectively mitigates overfitting and enhances prediction robustness. Especially when dealing with high-noise, high-dimensional, and highly nonlinear data common in geological engineering, the proposed architecture exhibits superior stability and adaptability, offering a reliable and efficient solution for the accurate prediction of formation drillability parameters.

In this study, the strategy of training three independent MSCNR-SE models is adopted, each dedicated to predicting one of the three target drillability parameters: σ_c, K_d, and I_a. This approach allows each model to specialize in capturing the unique relationships between the input features and its specific target, thereby simplifying the optimization process and enhancing prediction performance.

3. Case study

3.1. Data source

To validate the effectiveness of the MSCNR-SE model in predicting key formation drillability parameters—namely σ_c, K_d, and I_a—a case study was conducted using drilling engineering data and well-log interpretations from Huizhou Oilfield. The dataset comprises drilling engineering parameters and well-log interpretation results from nine wells, offering a comprehensive representation of the geological characteristics and physical properties of the encountered formations. The well log interpretation formulas are shown in Eqs. (8)–(14).

$$G_{\mathrm{dyn}} = 13474.45 \times \frac{\rho_b}{(\Delta t_{\mathrm{shear}})^2} \tag{8}$$

$$K_{\mathrm{dyn}} = 13474.45 \times \rho_b \left[\frac{1}{(\Delta t_{\mathrm{comp}})^2}\right] - \frac{4}{3} G_{\mathrm{dyn}} \tag{9}$$

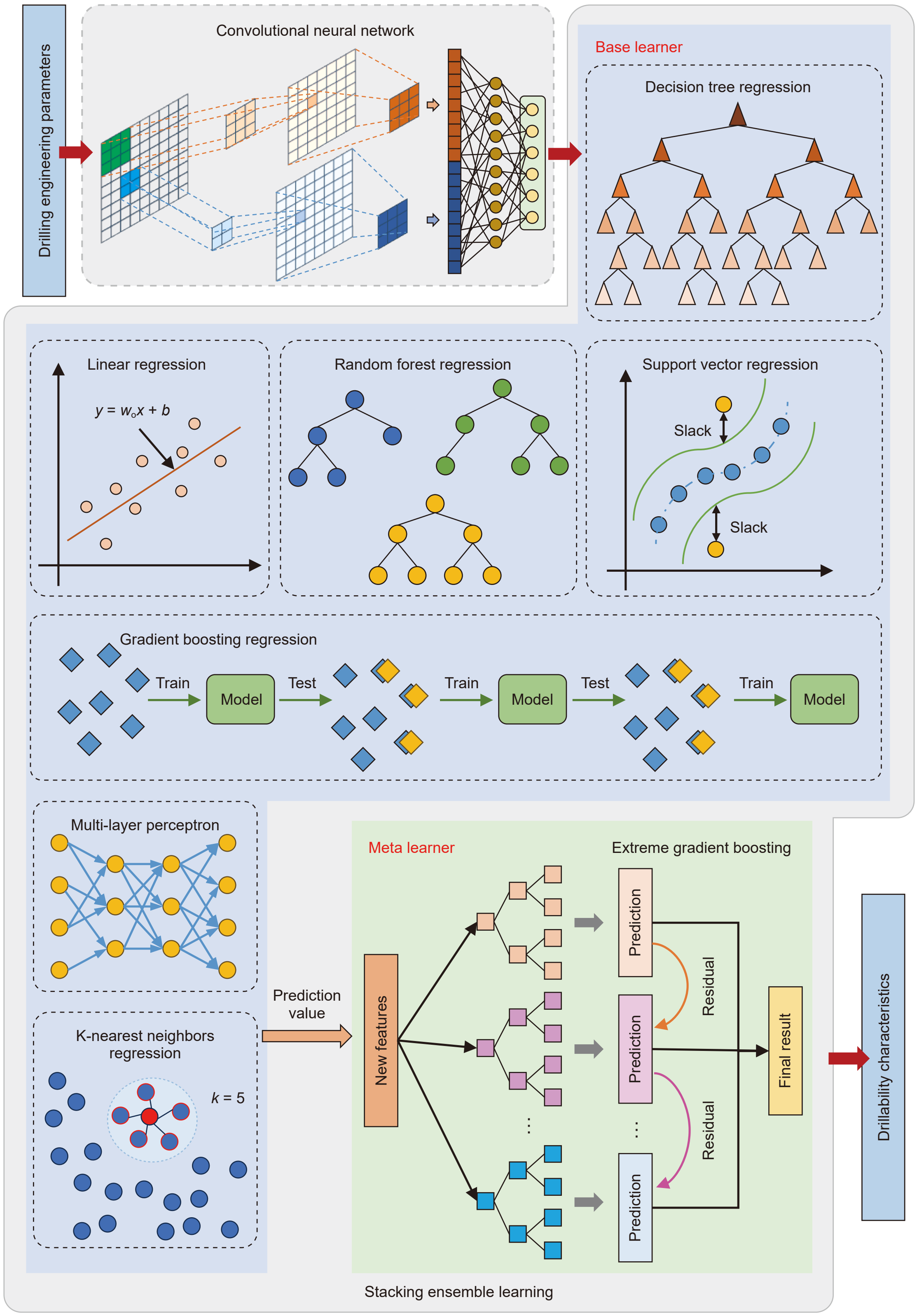

Fig. 5. Architecture of prediction method for formation drillability characteristics parameters based on MSCNR-SE.

$$E_{\mathrm{dyn}} = \frac{9G_{\mathrm{dyn}} \times K_{\mathrm{dyn}}}{G_{\mathrm{dyn}} + 3K_{\mathrm{dyn}}} \tag{10}$$

$$E_{\mathrm{sta}} = 0.0018 \times {E_{\mathrm{dyn}}}^{2.7} \tag{11}$$

$$\sigma_{\mathrm{c}} = \left(2.280 + \frac{4.1089 \times E_{\mathrm{sta}} \times 6894.76}{1000}\right) \tag{12}$$

$$I_{\mathrm{a}} = 26.392 - 0.166 \cdot \Delta t_{\mathrm{comp}} - 1.216\rho_{\mathrm{b}} \tag{13}$$

$$K_{\mathrm{d}} = 18.0602 \times \mathrm{e}^{-0.012\Delta t_{\mathrm{comp}}} + 3.2 \tag{14}$$

where ρ_{b} is the bulk density (g/cm^3); Δt_{comp} and $\Delta t_{\mathrm{shear}}$ are the compressional-wave and shear-wave slowness (μs/ft) respectively; G_{dyn}, E_{dyn} and K_{dyn} denote the dynamic shear modulus, dynamic elastic modulus and dynamic bulk modulus (Mpsi) respectively; E_{sta} is the static elastic modulus and 13,474.45 is the unit conversion coefficient.

The study focused on formations within a vertical depth range of 3000 to 4200 m, for which acoustic slowness (AC), natural gamma ray (GR), and density (DEN) logs were available. These formations cover a wide range of lithological types and drilling conditions, providing strong representativeness for model validation. All well log data were recorded at a high-resolution sampling interval of 0.1 m to ensure data continuity and accuracy. The drilling parameters were sampled at a depth interval of 0.1 m, and the well-log data, geophysical parameters, and drilling data were aligned using a depth-based matching approach. The initial fifteen parameters were selected as candidate input features, including: Standpipe pressure (SPP), flow rate (Flow), mud weight (Mw), Depth, formation fracture pressure gradient (Frac), equivalent circulating density (ECD), Torque, corrected d-exponent (D_{c}), weight on bit (WOB), temperature (Temper), revolutions per minute (RPM), and rate of penetration (ROP). In addition to the operational and hydraulic parameters, three bit-related parameters were incorporated into the input feature set to account for the influence of bit geometry on drilling performance: Cutter type (CT, coded as 0–plane, 1–stinger, 2–axe-shape, 3–tri-ridge), number of blades (NoB), and bit type (1–roller-cone bit, 2–PDC bit). However, the BHA stiffness parameter was not recorded in the available dataset and therefore could not be incorporated. A total of 1,286,458 log data points were ultimately collected, forming the experimental dataset used in this study. The key features of the model input provide basic statistics, as shown in Table 1.

3.2. Feature selection

An excessive number of input features may lead to increased model complexity and a higher risk of overfitting. To enhance model performance and practical applicability in engineering scenarios, it is crucial to identify the most relevant parameters with respect to the target variables—σ_{c}, K_{d}, and I_{a}.

These parameters help the model capture the effect of bit structure and configuration on drilling responses such as ROP, Torque, and standpipe pressure, thereby improving the physical interpretability of the MSCNR-SE framework. To evaluate the linear and monotonic associations between these features and the three target variables, as shown in Eqs. (15) and (16), the Pearson correlation coefficient (r_{p}) and Spearman's rank correlation coefficient (r_{s}) were applied. Furthermore, a composite metric termed "cumulative absolute correlation index" was introduced to quantify the overall linear association of each feature across all three target variables. This approach provides a more robust basis for feature selection, balancing predictive relevance with model generalizability.

$$r_{\mathrm{p}} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{15}$$

$$r_{\mathrm{s}} = 1 - \frac{6\sum_{i=1}^{n}{d_i}^2}{n(n^2 - 1)} \tag{16}$$

where d_i is the difference between the ranks of paired observations and n is the total number of samples.

The results of the correlation coefficient calculations are illustrated in the stacked bar chart shown in Fig. 6. It can be observed that the parameters SPP, Flow, Mw, Depth, Frac, ECD, Torque, RPM, and ROP exhibit relatively high cumulative correlation scores in both Pearson correlation coefficient and Spearman's rank correlation coefficient. This indicates that these variables possess strong overall linear associations with the drillability-related target parameters, suggesting high interpretability and predictive value.

These parameters are not only statistically significant but are also of critical importance in practical drilling operations. For example, SPP and ECD jointly reflect the annular pressure environment downhole, while Torque and ROP are closely associated with the real-time mechanical interaction between the drill bit

Table 1
Distribution of data.

Key features	Maximum	Minimum	Mean	Median	Standard deviation
σ_{c}, MPa	220.3914	25.13823	109.8478	99.46754	39.15825
K_{d}	12.82227	8.7451	11.05366	10.92942	0.732838
I_{a}	25.07646	1.695028	10.76383	9.342956	4.759214
RPM, r/min	101.00	29.00	61.95	60.00	16.38
Torque, N·m	16834.00	3375.00	9878.87	9431.00	2218.54
SPP, psi	3901.00	1172.00	3231.44	3550.00	623.16
Flow, L/min	3877.00	813.00	3321.20	3738.00	713.75
ROP, m/h	48.00	0.43	13.15	10.13	9.54
Mw, g/cm^3	1.31	0.80	1.22	1.24	0.07
ECD, g/cm^3	1.33	1.18	1.27	1.28	0.04
Frac, g/cm^3	2.33	2.14	2.23	2.23	0.05
Depth, m	4326.00	3020.00	3669.34	3664.50	378.56
D_{c}	2.32	0.69	1.20	1.16	0.23
Bit type	2.00	1.00	1.63	2.00	0.48
CT	3	0	1.59	1	1.04
NoB	6	5	5.75	6	0.43

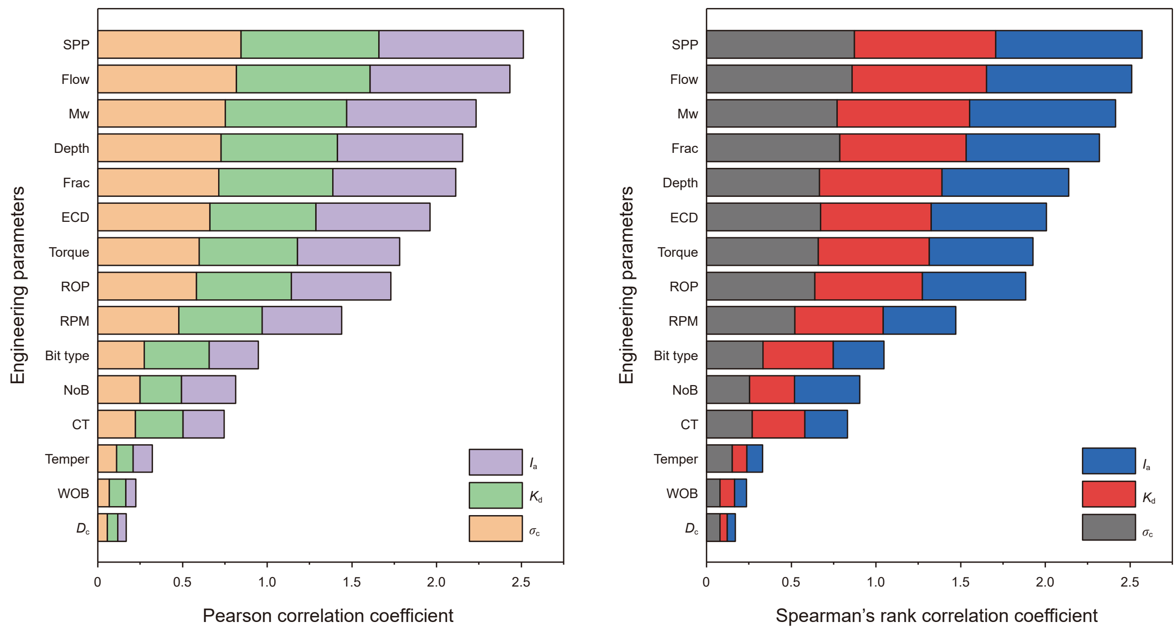

Fig. 6. Correlation coefficient stacked bar chart.

and the formation. Mw, on the other hand, plays a key role in controlling mud properties and maintaining wellbore stability. Although SPP, WOB, Flow and Torque are operational parameters, they reflect the dynamic response of the drilling system under formation resistance. When formation σ_c or abrasiveness increases, the Torque and SPP rise accordingly, while ROP and flow efficiency decrease (Hadi et al., 2019). Hence, these parameters can be treated as indirect indicators of formation drillability. In addition to the original operational parameters (SPP, WOB, Torque, Flow, ROP), the D_c was also incorporated into the feature set to enhance the model's ability to reflect formation-related characteristics.

Based on the above analysis, and considering both the measurability of the parameters during drilling and their engineering relevance, multiple parameters were ultimately selected as input features for the predictive model: bit type, NoB, CT, SPP, Flow, Mw, Depth, Frac, ECD, Torque, RPM, D_c and ROP. These selected features provide a solid data foundation for subsequent model development and application.

3.3. Selection of sliding window size

Fig. 7 illustrates the variation of the autocorrelation coefficients of selected engineering parameters (SPP, Flow, Mw, Depth) and formation drillability characteristic parameters (I_a, K_d) with respect to sampling intervals, which correspond to the lag order or sliding window size. The horizontal axis represents the lag (sampling interval), while the vertical axis indicates the corresponding autocorrelation coefficient.

As observed, the engineering parameters exhibit a clear degree of autocorrelation along the depth sequence. However, with increasing lag, the autocorrelation coefficients gradually decline and eventually stabilize, displaying a characteristic tailing-off pattern. This trend indicates that the correlation between measurements weakens as the time (or depth) lag increases, ultimately approaching a state of near-zero autocorrelation.

To ensure adequate temporal dependency within the input sequences for time-series modeling, the optimal size of the sliding window was determined based on ACF. As shown in Fig. 7, the intersection point between the ACF curve and the upper bound of the 95% confidence interval was used to identify the maximum sampling lag at which significant autocorrelation is retained. For the dataset under investigation, this intersection occurred approximately at a lag of 75, which was subsequently adopted as the sliding window size. The window size was further refined by analyzing the ACF of each engineering parameter in the training set. Specifically, the first intersection between each parameter's ACF curve and the 95% confidence bound was computed to quantify the individual temporal correlation structure. This approach enables the selection of a sliding window size that preserves the intrinsic autocorrelation characteristics of both input and output features, thereby enhancing the model's capacity to learn time-dependent patterns.

To evaluate the influence of sliding window size on model performance, five different window lengths (25, 50, 75, 100, and 125) were tested across three stratigraphic intervals: the Enping Formation, Wenchang Formation, and Pre-Paleogene Formation. The corresponding prediction metrics (R^2, MSE, RMSE, and MAE) of I_a are summarized in Table 2.

For the Enping Formation, the model achieved the highest accuracy when the window length was 50, yielding an R^2 of 0.9167, RMSE of 0.0451, and MAE of 0.0285. A moderate window (50–75) generally provided stable results, whereas excessively small (25) or large (≥100) windows led to a slight decline in accuracy, likely due to insufficient or over-smoothed temporal information capture. For the Wenchang Formation, the model achieved its optimal performance at a window length of 75, with a maximum R^2 of 0.9570 and the lowest MAE of 0.0222, indicating that this window length is conducive to improving prediction accuracy in heterogeneous deep formations. In the Pre-Paleogene strata, the best results occurred at a window size of 100, where $R^2 = 0.9136$, RMSE $= 0.0484$, and MAE $= 0.0296$. The performance variation across window sizes was relatively small ($\Delta R^2 < 0.03$), suggesting that the model maintains robustness under different temporal resolutions.

Overall, the results demonstrate that while the model exhibits consistent predictive capability across all formations, the optimal

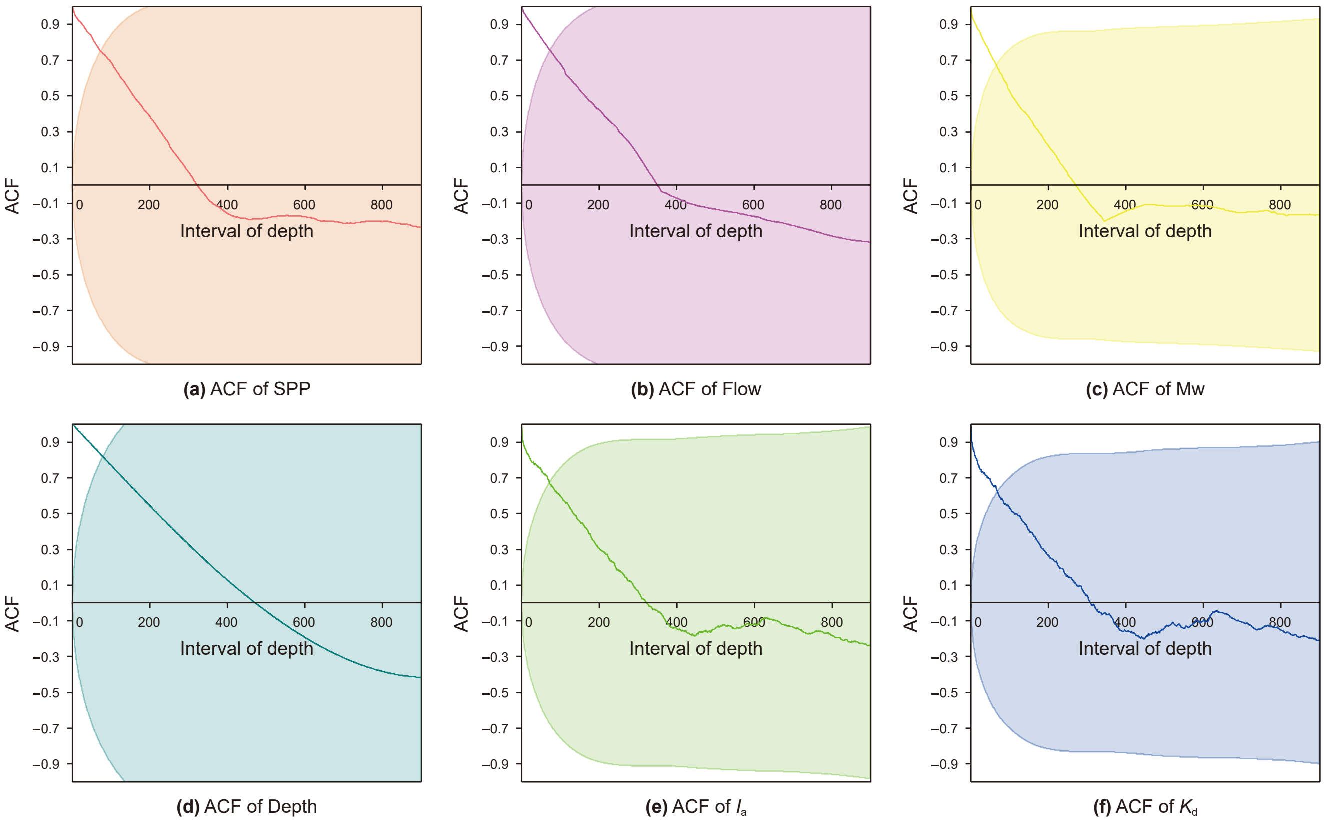

Fig. 7. The variation curves of autocorrelation coefficients (ACF) of drilling engineering parameters and formation drillability characteristics.

Table 2
Sensitivity analysis of sliding window length on model performance across different formations.

Window size	Formation	R^{2a}	MSE [b]	RMSE [c]	MAE [d]
25	Enping	0.9024	0.0024	0.0487	0.0316
50	Enping	0.9167	0.0020	0.0451	0.0285
75	Enping	0.9095	0.0023	0.0483	0.0275
100	Enping	0.8843	0.0032	0.0563	0.0291
125	Enping	0.9162	0.0024	0.0493	0.0248
25	Wenchang	0.9032	0.0035	0.0589	0.0366
50	Wenchang	0.9085	0.0034	0.0581	0.0353
75	Wenchang	0.9570	0.0018	0.0421	0.0222
100	Wenchang	0.9355	0.0026	0.0506	0.0272
125	Wenchang	0.8810	0.0045	0.0672	0.0326
25	Pre-Paleogene	0.8999	0.0033	0.0574	0.0384
50	Pre-Paleogene	0.9096	0.0026	0.0514	0.0344
75	Pre-Paleogene	0.8936	0.0029	0.0542	0.0342
100	Pre-Paleogene	0.9136	0.0023	0.0484	0.0296
125	Pre-Paleogene	0.8812	0.0028	0.0534	0.0303

[a] R^2 refers to coefficient of determination; [b] MSE refers to mean squared error; [c] RMSE refers to root mean squared error; [d] MAE refers to mean absolute error.

sliding window length varies slightly with formation characteristics. A window length of 50–100 provides a balanced trade-off between local feature sensitivity and temporal context representation, thus 75 being adopted as the final configuration in this study.

3.4. Details of the MSCNR-SE

(1) Construction of MSCNN

MSCNN was developed to extract key features from the raw drilling engineering parameter data, providing an effective latent spatial representation for the subsequent stacking ensemble learning model. This enables accurate prediction of σ_c, K_d, and I_a.

The hyperparameters of MSCNN were optimized using a grid-search approach combined with 5-fold cross-validation on the training dataset. The tuning focused on the parameters that significantly influence convolutional representation learning, including kernel size, stride, number of filters, activation function, and output size. The search space and tuning methods are summarized as follows:

Kernel size: [(2,2) (2,3) (3,3) (3,5) (2,5) (5,5)]— multi-scale kernels were adopted in parallel branches to capture formation responses at different depth scales. A small kernel (2,2) captures fine-scale variations in drilling parameters, which are typically associated with short lithological transitions (e.g., thin interbedded sand–mud layers, local fracture zones). A medium kernel (3,3) reflects meso-scale heterogeneities such as changes in rock compaction and mechanical strength over several meters. A large kernel (3,5) or (5,5) captures broader-scale geological trends, such as gradual changes in lithology or stress fields along depth. This multi-scale convolutional design enables the model to simultaneously extract local, intermediate, and global formation response features, which are essential for accurately modeling formation drillability under variable geological conditions.

Stride: [1, 2]—a stride of 1 was found optimal for preserving depth resolution without excessive downsampling. A max-pooling operation with a stride of 1 was employed to reduce dimensionality while retaining the most representative features from each scale. This helps enhance computational efficiency and prevents overfitting, while maintaining key geological signal characteristics.

Number of filters: [8, 16, 32, 64, 128]—increasing filters improved representational capacity up to 64, beyond which marginal gains diminished. Activation function: {ReLU, LeakyReLU,

Table 3
MSCNN architecture for feature extraction.

Name		Kernel size	Stride	Filters	Activation function	Output size
Search range		(2,2), (2,3), (3,3), (3,5), (2,5), (5,5)	1, 2	8,16, 32, 64, 128	ReLU, LeakyReLU, ELU	—
Input layer		—	—	—	—	(75, 13, 1)
Branch 1	Conv layer 1	(2, 2)	1	32	ReLU	(75, 13, 32)
	Pooling layer 1	(3, 3)	—	—	—	(25, 3, 32)
Branch 2	Conv layer 2	(3, 3)	1	32	ReLU	(75, 13, 32)
	Pooling layer 2	(3, 3)	—	—	—	(25, 3, 32)
Merge layer		—	—	—	—	(50, 3, 32)
FC layer 1		—	—	—	ReLU	512
FC layer 2		—	—	—	ReLU	256
FC layer 3		—	—	—	ReLU	128
Output layer		—	—	—	ReLU	64

ELU}—ReLU yielded the best convergence stability and computational efficiency. Output size: determined adaptively by the pooling strategy to ensure consistent dimensionality for stacking ensemble integration. Each candidate configuration was trained until the validation loss converged or met the early stopping criterion (no improvement for 20 epochs).

The hyperparameters of MSCNN designed for feature extraction is detailed in Table 3.

The network employs two parallel convolutional branches, each using a different kernel size—2×2 and 3×3—to extract local spatial features at multiple scales, thereby capturing both fine-grained details and broader patterns within the data. Each convolutional branch is followed by a pooling layer that downsamples and compresses the resulting feature maps, reducing computational complexity and mitigating overfitting while retaining essential information.

After the multi-scale feature extraction stage, the feature maps from both branches are concatenated via a merge layer to form a richer, higher-dimensional representation. At this point, the network has integrated spatial features learned at different scales into a unified latent embedding. This embedding is then passed through several FC layers, which apply nonlinear transformations, further dimensionality reduction, and feature refinement to produce the final latent spatial representation.

The multi-scale module is designed to extract high-dimensional latent features and hierarchical representations from heterogeneous data. These features serve as the input to the subsequent stacking ensemble learning model for accurate prediction of σ_c, K_d, and I_a.

(2) Regression model construction

In the model development process, a combination of randomized search and K-fold cross-validation was employed for hyperparameter optimization. This approach aims to comprehensively explore the parameter space of the model, thereby enhancing its performance and mitigating the risk of overfitting. By performing multiple rounds of random sampling within the predefined hyperparameter ranges, the approach effectively improves tuning efficiency and ensures that the performance of each regression model on the validation set is both representative and robust. The optimal hyperparameter combinations for each model, along with their performance during the training phase, are shown in Table 4.

After completing the training and parameter optimization of the base models, this study further constructs a multi-model ensemble prediction framework based on stacking ensemble learning, as proposed in Section 2.4. This framework is divided into two layers: The first layer consists of base learners, which include a variety of representative models that are widely used in regression tasks. These models include DT, LR, RF, GBDT, SVR, MLP, and KNN. Each base model is trained independently in the MSCNN-encoded feature space, extracting diverse predictions by modeling different data structures and nonlinear relationships, thereby enhancing the diversity and information richness of the ensemble model.

In the second layer, the meta-learner, the study selects XGBoost, a gradient boosting model that has demonstrated excellent performance in recent years, as the final integrator. XGBoost is known for its outstanding generalization ability, robustness to feature noise, and flexible regularization mechanisms. It effectively integrates the predictions from all base models, providing a comprehensive prediction of the target parameters, such as σ_c, K_d, and I_a.

(3) Evaluation metrics establishment

Given that this study primarily focuses on the regression prediction of formation drillability parameters, four commonly used statistical metrics, MAE, RMSE, the mean absolute percentage error (MAPE) and R^2, are selected as evaluation criteria to comprehensively assess the model's performance in continuous variable prediction. These metrics evaluate the accuracy and stability of the prediction results from different perspectives, providing a comprehensive reflection of the model's generalization ability and error characteristics (Qu et al., 2023).

MAE: Average absolute deviation between predicted and true values, reflecting overall bias (smaller value represents better accuracy).

RMSE: Deviation metric with high sensitivity to large errors and outliers.

MAPE: Relative error in percentage form, applicable for cross-unit model comparison.

R^2: Proportion of variance explained by the model (closer to 1 means better fit).

The unified expressions are as follows:

Table 4
The optimal hyperparameter combinations for each model.

Model	Key parameters
DT	random state = 42
LR	alpha = 10
RF	max depth = 10; min samples split = 10; n estimators = 50
GBDT	learning rate = 0.033; max depth = 10; n estimators = 149; subsample = 0.83
SVR	C = 100; gamma = scale; kernel = rbf
MLP	activation = ReLU; alpha = 0.017; hidden layer sizes = 50; learning rate = adaptive; learning rate init = 0.019; solver = sgd
KNN	algorithm = brute; n neighbors = 18; p = 1; weights = distance
XGBoost	colsample bytree = 0.79; gamma = 0.083; learning rate = 0.012; max depth = 3; n estimators = 443; subsample = 0.82

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \widehat{y}_i| \tag{17}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \widehat{y}_i)^2}{N}} \tag{18}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \widehat{y}_i}{y_i}\right| \times 100\% \tag{19}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2} \tag{20}$$

where y_i is the true value of the i-th sample; $\overline{y}$ denotes the mean value of the target variable across all samples; $\widehat{y}_i$ is the predicted value of the i-th sample; N is the number of samples.

3.5. Application results

The comparison between the actual and predicted values of σ_c, I_a, and K_d for a real case well (depth range 3100–4250 m) using the developed MSCNR-SE method is shown in Fig. 8.

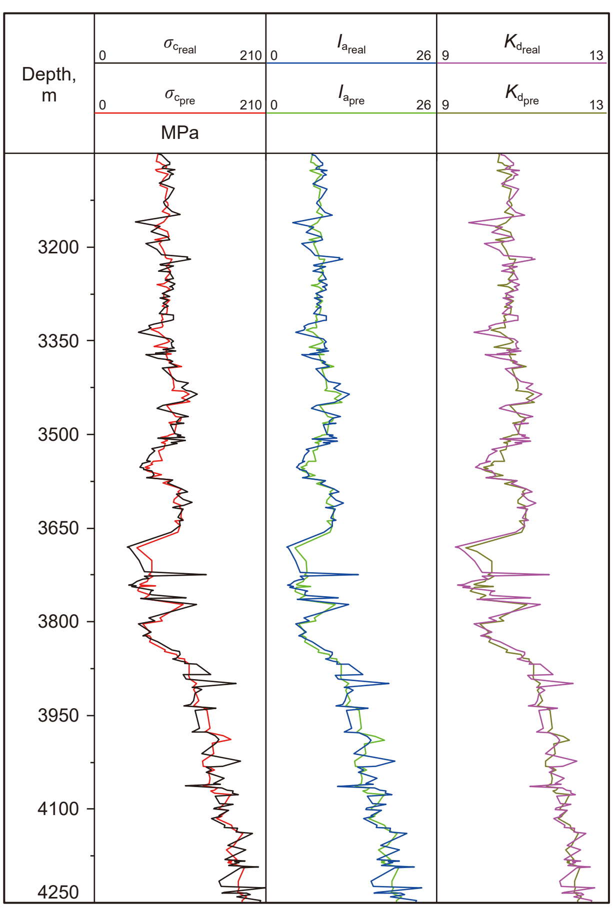

Fig. 8. Predicted and actual values of σ_c, I_a, and K_d models for the case well versus well depth.

The results presented in Fig. 8 demonstrate the comparison between the actual values and predicted values of each parameter as a function of well depth. It is clearly observed that with the gradual increase in well depth, the three indicators—σ_c, I_a, and K_d—show a distinct upward trend. This phenomenon aligns with the actual geological conditions: during the drilling process, as the wellbore penetrates deeper formations, more compact, or harder formations, the rock's strength increases, its resistance to abrasion strengthens (I_a increases), and its drillability decreases (i.e., the K_d value increases, indicating greater drilling difficulty). This gradual variation in the physical and mechanical properties of the formation with depth reveals the driving effect of the increasing complexity of the downhole environment on the variation of drilling parameters.

As shown in Fig. 8, the predicted and measured values of σ_c, I_a, and K_d exhibit strong consistency in most intervals, particularly within the upper Enping Formation (3600–3650 m) and the lower Wenchang Formation (3900–3950 m), where lithology is relatively homogeneous and the drilling response remains stable. However, noticeable deviations occur in the transitional intervals (e.g., 3660–3700 m), which correspond to abrupt lithological changes characterized by interbedded sandstone and shale. These rapid lithological transitions can cause significant variations in mechanical properties, leading to discrepancies between predicted and measured values. In addition, minor misalignments between drilling and logging data, as well as localized data anomalies (e.g., sensor noise or missing records), may contribute to residual prediction errors. Overall, the well-matched intervals correspond to geologically stable formations, whereas mismatched zones are primarily associated with complex or rapidly varying lithological conditions.

The MSCNR-SE model accurately captures the changing trends of these geological patterns in its predictions. The predicted values closely match the actual measurements, with the variation curves largely overlapping at key fluctuation points and trend changes, demonstrating the model's strong capability in modeling formation response characteristics. This ability is primarily attributed to the model's multi-scale feature extraction module and multi-model ensemble learning strategy. Through a deep shared feature extraction mechanism, the model is able to explore the nonlinear coupling relationships between input drilling parameters and lithological responses, while the stacking ensemble mechanism further enhances the overall generalization ability, enabling the model to maintain strong adaptability and robustness to feature variations under different formation conditions.

To further quantify the predictive performance of the MSCNR-SE model in real well sections, MAE, RMSE, MAPE and R^2 were employed, as shown in Table 5.

These results confirm that the improvements in predictive accuracy are not due to random variation, but reflect a consistent enhancement achieved through the multi-scale feature extraction and stacking ensemble mechanism.

As shown in the table, in the σ_c prediction task, the model's MAE is controlled at 6.92 MPa, with an RMSE of 8.98 MPa, MAPE of only 7.60%, and an R^2 value as high as 0.94. This indicates that the model has a strong ability to characterize the trend of formation strength changes, providing a reliable basis for drill bit selection

Table 5
MSCNR-SE model performance in the case well.

Drillability characteristics	MAE	RMSE	MAPE, %	R^2
σ_c	6.92	8.98	7.60	0.94
I_a	0.84	1.04	10.19	0.94
K_d	0.14	0.18	1.33	0.93

and wellbore inclination control. In the I_a prediction, the MAE and RMSE are 0.84 and 1.04, respectively, with a MAPE of 10.19%, and an R^2 of 0.94. This suggests that the model exhibits good consistency and prediction stability in capturing the rock abrasiveness with well depth. For K_d prediction, the prediction error is minimal (MAE = 0.14, RMSE = 0.18, MAPE only 1.33%), with an R^2 value of 0.93, further reflecting the model's high sensitivity and accuracy in capturing changes in drillability grade.

These evaluation results demonstrate that the MSCNR-SE model not only possesses strong nonlinear feature modeling capabilities but also exhibits good adaptability for engineering applications. It shows high accuracy and stability in predicting multiple key rock mechanics and drilling parameters, making it an effective tool for assisting engineers in real-time formation identification, drill bit parameter optimization, and risk control for drilling through complex geological sections. It holds significant practical value and strong potential for engineering implementation.

3.6. An engineering case study: decision optimization based on MSCNR-SE

To demonstrate the practical utility of the proposed MSCNR-SE model in guiding engineering decisions, a blind test was conducted on a nearby well, W63, from the same field. Crucially, the data from Well W63 were not included in the training dataset, ensuring an unbiased evaluation of the model's generalization capability and its value as a pre-drill planning tool.

The drilling engineering parameters from the interval of 3460 to 3720 m in Well W63 were fed into the trained MSCNR-SE model as shown in Fig. 9. The model output identified a section with I_a values exceeding the threshold of 10, indicating a potentially highly abrasive formation that could lead to accelerated bit wear.

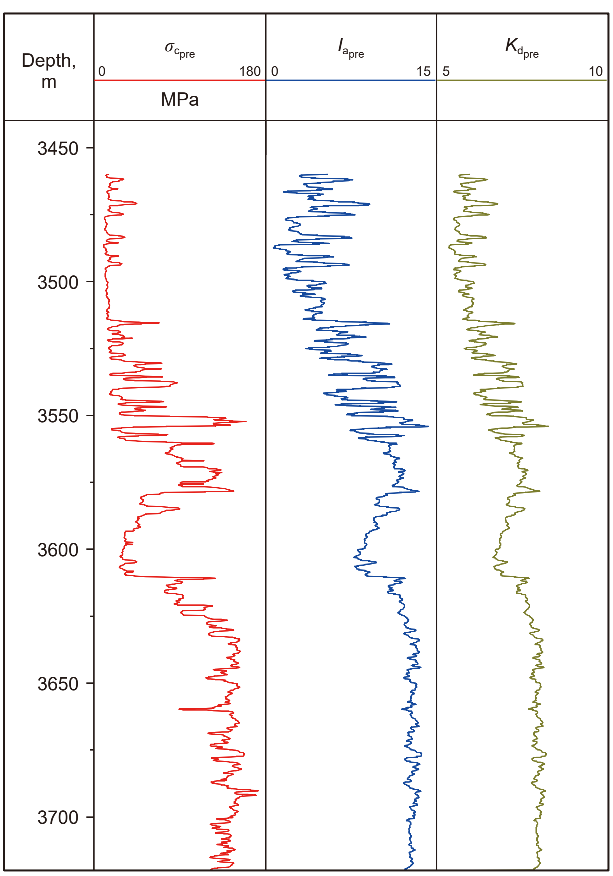

Fig. 9. Predicted values of σ_c, I_a, and K_d for Well W63 using MSCNR-SE model.

The predicted high-abrasiveness interval was consistent with the subsequent drilling response and bit wear observations. The operation initially employed a standard 6-blade PDC bit. When drilling reached approximately 3542 m, ROP dropped sharply from 8–10 to 2–4 m/h, indicating severe drilling dysfunction. Upon pulling out of the hole, the bit was found to have suffered significant outer cutter wear. The post-run wear analysis according to the IADC dull grading system was 5-3-WT-A-X-1/16-CT-PR. The total footage drilled by this bit was only 96.2 m, which was substantially lower than the average footage achieved by the same type of bit in other, less abrasive formations in the same field.

In retrospect, the model's prediction of high abrasivity would have provided a critical and actionable warning. Had this prediction been available during the well planning phase, it would have strongly justified the selection of a more robust bit design from the outset for this specific interval. In fact, the solution adopted by the field crew after the failure aligns perfectly with this logic: a 7-blade PDC bit with enhanced diamond wear coatings was run subsequently. This bit successfully navigated the challenging formation, with the ROP recovering to and stabilizing at 6–8 m/h.

This case study demonstrates that the MSCNR-SE model not only serves as an accurate predictive tool but also has practical value as a decision-support system for drilling optimization. By accurately forecasting formation abrasivity, it enables proactive bit selection, reduces the risk of premature bit failure, and contributes to improved drilling efficiency and reduced non-productive time.

4. Discussion

4.1. Analysis of the effectiveness of MSCNR-SE

To validate the effectiveness of MSCNR-SE, this section presents a systematic comparison of the MSCNR-SE ensemble model's prediction performance with that of each individual base model (DT, LR, RF, GBDT, SVR, MLP, KNN, XGBoost), on a specifically partitioned dataset. Four deep-sequence benchmarks (LSTM, GRU, 1D-CNN, TCN) have been included under identical data splits. The four evaluation metrics (MAE, RMSE, MAPE, R^2) are used as standardized evaluation criteria to ensure fairness and comparability of the results. Each model is run with its optimal hyperparameter configuration, and 5-fold cross-validation is employed to enhance the robustness of the evaluation. The results are shown in Table 6.

To validate the statistical significance of the performance improvement achieved by the proposed MSCNR-SE model, a Wilcoxon signed-rank test was conducted using paired error metrics (MAE and RMSE) between MSCNR-SE and all baseline models listed in Table 6, including both traditional machine-learning models (DT, LR, RF, GBDT, SVR, MLP, KNN, XGBoost) and deep-learning models (LSTM, GRU, 1D-CNN, and TCN).

As summarized in Table 7, the obtained p-values for both MAE and RMSE across all three drillability parameters (σ_c, I_a, and K_d) are below 0.05, confirming that the improvements achieved by MSCNR-SE are statistically significant at the 95% confidence level. These results demonstrate that the superior performance of MSCNR-SE is not due to random fluctuations but reflects a consistent enhancement in prediction accuracy and model robustness across geological formations.

As shown in the table, the MSCNR-SE model outperforms other comparative models in predicting drilling characteristic parameters (σ_c, I_a, K_d) across all evaluation metrics (MAE, RMSE, MAPE, R^2), demonstrating superior prediction accuracy and generalization ability. This result confirms the effectiveness and applicability of

Table 6
Prediction performance of MSCNR-SE and comparative models.

Model	σ_c, MPa				I_a				K_d			
	MAE	RMSE	MAPE, %	R^2	MAE	RMSE	MAPE, %	R^2	MAE	RMSE	MAPE, %	R^2
LSTM	8.85	10.42	16.15	0.82	1.03	1.34	19.65	0.88	0.24	0.34	2.84	0.87
GRU	8.31	9.44	16.74	0.81	1.23	1.36	21.65	0.87	0.26	0.37	2.92	0.86
1D-CNN	8.25	9.85	13.06	0.87	1.19	1.28	16.26	0.92	0.20	0.29	2.89	0.91
TCN	7.98	9.57	15.12	0.88	0.98	1.28	11.96	0.92	0.21	0.30	2.62	0.91
DT	8.52	10.98	8.73	0.90	1.03	1.32	11.37	0.96	0.27	0.34	1.85	0.92
LR	8.01	10.43	8.20	0.90	0.98	1.23	11.32	0.91	0.25	0.27	1.48	0.90
RF	7.75	9.60	8.00	0.91	0.96	1.20	11.35	0.96	0.19	0.27	1.49	0.93
GBDT	7.75	9.83	7.87	0.94	0.97	1.25	11.31	0.93	0.25	0.30	1.63	0.95
SVR	13.96	18.57	17.53	0.73	1.76	2.47	24.99	0.75	0.31	0.50	2.96	0.78
MLP	11.53	14.90	10.56	0.84	1.07	1.27	11.40	0.92	0.19	0.25	1.45	0.96
KNN	8.96	16.39	14.15	0.92	1.60	1.99	18.23	0.86	0.33	0.39	2.52	0.79
XGBoost	8.34	10.80	8.27	0.96	0.99	1.32	10.97	0.96	0.20	0.29	1.52	0.95
MSCNR-SE	6.97	9.06	7.64	0.97	0.93	1.06	10.24	0.97	0.18	0.18	1.37	0.98

Table 7
The Wilcoxon signed-rank test results of models in Table 6.

Drillability parameter	*p*-value (MAE)	*p*-value (RMSE)	*p*-value (MAPE)	Significance
σ_c	0.005	0.004	0.007	$p < 0.05$
I_a	0.012	0.015	0.019	$p < 0.05$
K_d	0.008	0.009	0.011	$p < 0.05$

the multi-scale feature extraction mechanism and multi-model ensemble learning strategy in modeling complex nonlinear relationships.

The MSCNR-SE model achieved superior predictive performance compared with other models (LSTM, GRU, 1D-CNN, TCN, DT, LR, RF, GBDT, SVR, MLP, KNN and XGBoost) across all three geomechanical parameters (σ_c, I_a, and K_d). Specifically, relative to both deep learning and machine learning models, MSCNR-SE reduced the MAE by approximately 3%–50%, the RMSE by 4%–64%, and the MAPE by 3%–59%. In terms of goodness of fit, the coefficient of determination (R^2) increased by 1%–32%.

From a model structure perspective, MSCNR-SE integrates MSCNN with the stacking ensemble framework. MSCNN models local features and long-range dependencies of the input drilling parameters through convolutional kernels with different receptive fields, effectively extracting multi-scale semantic information highly related to rock properties. These features are mapped to a shared latent space, serving as input for multiple base regression models. Ultimately, the XGBoost meta-learner performs ensemble optimization. This synergistic architecture, integrating the feature layer and decision layer, constructs a unified feature representation and structural modeling logic for the three prediction tasks, enhancing the model's stability and cross-task generalization ability.

The single learners are limited by their simple model structures or excessive dependence on input features, making it difficult to adequately model the complex nonlinear relationships between drilling parameters and lithological indicators in geological environments. As a result, they show a lack of consistency across multiple tasks. While ensemble learning models do have some nonlinear fitting ability, they primarily perform ensemble optimization based on shallow features, lacking effective deep abstraction and cross-scale feature capture capabilities, thus struggling to capture high-order interactions among complex multivariable data.

The predicted trends of the drilling characteristic parameters (σ_c, I_a, K_d) under different models show high consistency, with σ_c and K_d fitting performances being particularly strong. This not only reflects the inherent correlation between the three parameters in terms of rock drillability and physical properties, but also further demonstrates that the MSCNR-SE model has strong feature adaptation capability and consistent modeling ability across multiple tasks.

4.2. Comparison of different meta-learners

To evaluate the impact of different meta-learners on the predictive performance of the MSCNR-SE model, this section compares eight commonly used regression models as meta-learners within the stacking framework, while maintaining the consistency of the multi-scale feature extraction module and base regression learner structure. The selected meta-learners are namely DT, LR, RF, GBDT, SVR, MLP, KNN, and XGBoost. Comparison experiments are conducted, and scatter fitting plots along with 95% confidence intervals are drawn on the test set to evaluate their generalization ability. The results are shown in Fig. 10. The corresponding R^2 values for each meta-learner are as follows: DT (0.89), LR (0.78), RF (0.91), GBDT (0.90), SVR (0.76), MLP (0.85), KNN (0.87), and XGBoost (0.97), with XGBoost achieving the highest R^2 and exhibiting the best performance.

Based on the results, LR and SVR exhibit notable modeling limitations in multi-scale deep feature fusion tasks with high-dimensional, strongly nonlinear relationships. Their ability to model complex feature interactions is limited, resulting in relatively poor predictive performance. This is particularly evident when dealing with the highly nonlinear and strongly coupled mapping relationships between geological parameters and drilling response data, where linear models struggle to capture the inherent characteristics, as reflected in the significantly lower R^2 values.

In contrast, models with stronger nonlinear modeling capabilities, such as DT, RF, GBDT, and KNN, perform better in the ensemble framework, achieving $R^2 \geq 0.88$. These models show good responsiveness to complex feature inputs, particularly RF and GBDT, which exhibit stronger robustness and generalization performance in handling heterogeneous feature distributions and high-variance scenarios.

MLP, as a feed-forward neural network, possesses some depth representation capability, with an R^2 of 0.84. While it outperforms

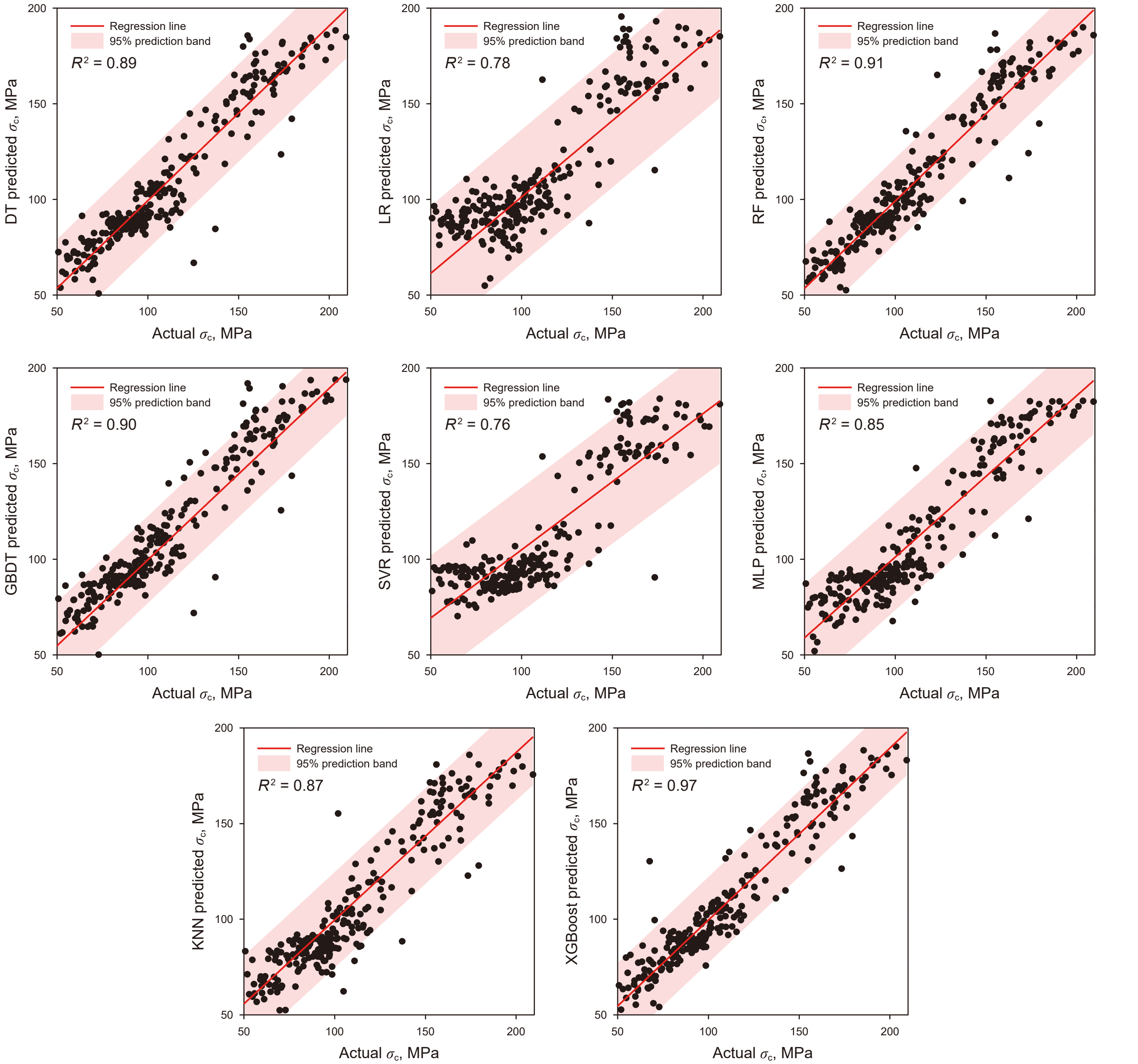

Fig. 10. Comparison of prediction performance for σ_c among different meta-learners.

linear models, it still faces limitations in stability and training efficiency due to sensitivity to parameter scales and sample sizes.

It is noteworthy that XGBoost, a gradient boosting-based ensemble algorithm, offers high flexibility in feature selection, weak learner optimization, and loss function design. It effectively models nonlinear coupling relationships and high-order interaction information between features. Additionally, its regularization strategy controls model complexity, effectively mitigating overfitting issues, making it especially suitable for high-dimensional and multi-source input tasks. In the stacking framework, XGBoost as the meta-learner is able to fully integrate the outputs of base models and multi-scale convolutional features, learning the underlying distribution structure and combinatorial patterns between them, thereby significantly improving overall predictive performance.

In conclusion, XGBoost outperforms other candidate meta-learners in both predictive accuracy and generalization ability, confirming its adaptability and advantages in the MSCNR-SE ensemble structure. Therefore, it is selected as the meta-learner for this model, providing critical support for high-precision prediction of drilling characteristics.

4.3. Analysis of latent space representations in MSCNR-SE

In this study, MSCNR-SE utilizes MSCNN to extract features from multi-dimensional drilling parameters, generating latent space representations of various dimensions to enhance the model's ability to perceive multi-scale information. This section systematically evaluates the impact of the latent space representation's dimension on model performance using a new dataset.

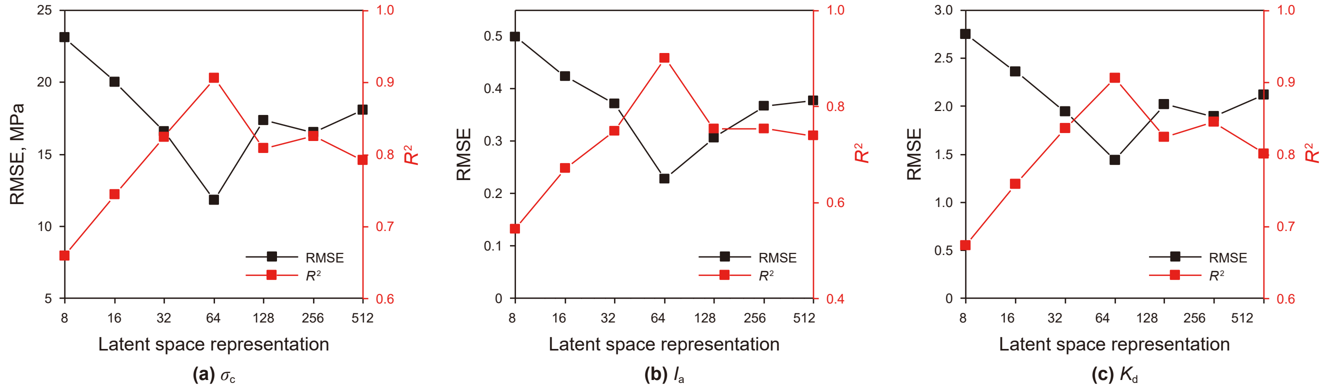

Fig. 11. Prediction performance of MSCNR-SE under different latent space representations.

Different numbers of latent nodes (8, 16, 32, 64, 128, 256, 512) were selected for the experiment, and the model's prediction results for the three parameters were evaluated by calculating RMSE and R^2 at each dimensional level. The results are shown in Fig. 11.

As shown in Fig. 11, with the gradual increase in the latent space dimension, the overall model performance significantly improves during the early stages (latent dimensions ranging from 8 to 64). This improvement occurs because increasing the dimension enhances the feature expression ability within the latent space, allowing the model to better capture the complex and non-linear relationships between drilling parameters and rock-drilling characteristic parameters. However, when the latent space dimension reaches 64, the model achieves optimal performance in the σ_c, K_d, and I_a prediction tasks, characterized by the lowest RMSE and the highest R^2. This suggests that, at this dimension, the model strikes an ideal balance between feature abstraction capability and generalization ability.

Further increasing the latent space dimensions to 128, 256, and 512 results in a noticeable local decline and fluctuations in model performance. This trend reflects that excessively high feature dimensions introduce a large number of redundant or inefficient features, leading the model to focus on noise information that lacks discriminative power, thereby diverting the learning focus. Additionally, as the feature space dimension increases, overfitting risk significantly escalates, particularly when training data is limited. The model may learn excessive detailed patterns from the training set but perform poorly in terms of generalization to the test set or real-world conditions, manifested as unstable prediction errors and increased performance fluctuations. Therefore, in the early stages of modeling, this study selects a latent space dimension of 64 for input feature compression.

4.4. Feature importance evaluation based on SHAP values

Applying SHAP (Shapley additive explanations) to analyze feature importance helps elucidate the prediction mechanism of complex models, quantify the contribution of each input feature to the target variable, identify key influencing factors, and reveal nonlinear effects and feature interactions. Fig. 12 presents the SHAP feature importance rankings for the σ_c, I_a, and K_d, offering useful insights into model refinement and the interpretability of the model's prediction logic.

Based on the SHAP summary plot for σ_c, Depth is identified as the most critical feature, followed by SPP, Flow, Frac, and Torque. These parameters are strongly associated with geological conditions, drilling hydraulics, and rock–bit interaction resistance. ROP, ECD, RPM, D_c, and Mw exhibit moderate contributions, while bit type and NoB show limited importance, indicating that σ_c is primarily governed by formation properties. From an operational perspective, greater attention should be given to depth and key drilling parameters. For model optimization, high-importance features should be retained, whereas low-importance ones may be screened or enhanced through feature engineering.

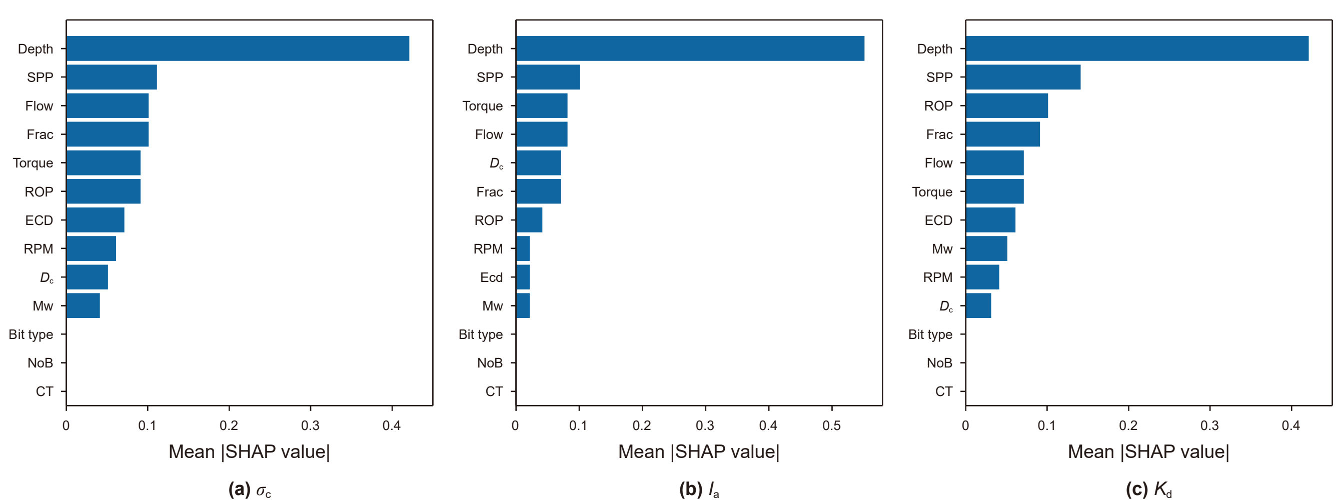

Fig. 12. SHAP feature importance for predicting σ_c, I_a, and K_d.

For I_a, Depth again emerges as the dominant feature, suggesting that deeper formations are generally denser and contain more abrasive minerals. SPP, Torque, Flow, and D_c are of moderate importance, reflecting the influence of drilling-fluid pressure, bit–rock friction, and contact force on abrasiveness. Frac, ROP, RPM, ECD, and Mw are less influential, whereas bit type, NoB, and CT contribute minimally. Compared with K_d and σ_c, Torque and D_c play a more pronounced role in I_a prediction, indicating that abrasiveness is more dependent on the direct mechanical interaction between the bit and rock.

In K_d analysis, Depth remains the most important feature, implying that drillability generally decreases with increasing depth. SPP ranks second, representing the effect of the hydraulic system on the drilling process. Moderately important features include ROP, Frac, Flow, and Torque, among which ROP serves as a direct indicator of drillability and contributes most to prediction accuracy. Features with lower importance are ECD, Mw, RPM, and D_c, while bit type, NoB, and CT show the least significance. Overall, K_d prediction is primarily driven by geological factors (Depth), real-time drilling responses (ROP), hydraulic system parameters (SPP, Flow, ECD, Mw), and mechanical variables (Torque, RPM, D_c). Compared with σ_c and I_a, ROP exhibits the highest importance for K_d, whereas pressure- and mechanically-related parameters show task-specific variations, and bit characteristics contribute marginally across all three predictions.

5. Conclusions

This study proposes a framework for predicting rock-drilling characteristic parameters that integrates multi-scale convolutional feature extraction (MSCNN) with stacking ensemble learning. The architecture utilizes multi-scale convolutional kernels to mine local-global features from engineering data, captures complex nonlinear relationships using seven heterogeneous base models (including DT, RF, GBDT, etc.), and integrates the prediction results of base models using the XGBoost meta-learner. This forms a complete modeling mechanism of “feature extraction-ensemble prediction-stacking optimization”.

The MSCNR-SE ensemble learning model built in this study can accurately capture the significant trends in σ_c, K_d, and I_a variations with well depth in an actual application in Huizhou Oilfield. The predicted values are highly consistent with the actual measurements ($R^2 > 0.93$ for all), with the MAE for σ_c being 6.92 MPa (MAPE = 7.60%) and the MAPE for K_d being only 1.33%.

The MSCNR-SE model effectively combines the strengths of multiple differentiated models, leading to improved prediction accuracy compared to single models. In the prediction of K_d, compared to SVR, the MSCNR-SE model reduces MAE and RMSE by approximately 41.94% and 64.01%, respectively, and increases R^2 by 25.64%. Compared to GBDT, MAE and RMSE are reduced by 28.02% and 40.21%, respectively, while R^2 increases by 3.16%.

By extracting multi-dimensional drilling parameters through the multi-scale convolutional neural network and fusing them into latent space representations, the MSCNR-SE model enhances its ability to perceive multi-source and multi-scale information. The model achieves optimal performance in σ_c, I_a, and K_d predictions when the latent space dimension is 64.

CRediT authorship contribution statement

Jian-Sheng Liu: Writing – original draft, Visualization, Methodology, Data curation. **Hua-Lin Liao:** Supervision, Resources, Methodology, Funding acquisition. **Zhe Huang:** Writing – review & editing, Methodology. **Feng-Tao Qu:** Writing – original draft, Methodology. **Fang Shi:** Writing – review & editing. **Tian-Yu Wu:** Writing – review & editing.

Declaration of competing interest

The authors 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

This paper acknowledges the support of the National Science and Technology Major Project (No. 2025ZD1401900, No. 2025ZD1401902).

References

Ahmednour, O., Chen, D., Liu, J., et al., 2025. Comparative analysis of rate of penetration prediction and optimization in deep wells using real-time continuous stacked generalization ensemble learning: A case study in Shunbei. Geoenergy Sci. Eng. 247, 213674. https://doi.org/10.1016/j.geoen.2025.213674.

Allawi, R.H., Al-Mudhafar, W.J., Abbas, M.A., et al., 2025. Leveraging boosting machine learning for drilling rate of penetration (ROP) prediction based on drilling and petrophysical parameters. Artificial Intell. Geosci. 6 (1), 100121. https://doi.org/10.1016/j.aiig.2025.100121.

Alvarado-Pérez, J.C., Garcia, M.A., Puig, D., 2025. Online dimensionality reduction through stacked generalization of spectral methods with deep networks. Mach. Learn. 114 (5), 125. https://doi.org/10.1007/s10994-024-06715-8.

Asante-Okyere, S., Marfo, S.A., Ziggah, Y.Y., 2023. Estimating total organic carbon (TOC) of shale rocks from their mineral composition using stacking generalization approach of machine learning. Upstream Oil and Gas Technology 11, 100089. https://doi.org/10.1016/j.upstre.2023.100089.

Bai, J., Wang, S., Liu, L., et al., 2025. Knowledge-based machine learning for real-time rock strength testing while drilling: bridging simulation and reality. Measurement 246, 116664. https://doi.org/10.1016/j.measurement.2025.116664.

Davoodi, S., Mehrad, M., Wood, D.A., et al., 2023. Predicting uniaxial compressive strength from drilling variables aided by hybrid machine learning. Int. J. Rock Mech. Min. Sci. 170, 105546. https://doi.org/10.1016/j.ijrmms.2023.105546.

Ehsan, M., Manzoor, U., Chen, R., et al., 2025. Pore pressure prediction based on conventional well logs and seismic data using an advanced machine learning approach. J. Rock Mech. Geotech. Eng. 17 (5), 2727–2740. https://doi.org/10.1016/j.jrmge.2024.09.049.

Feng, Z., Gani, H., Damayanti, A.D., et al., 2023. An explainable ensemble machine learning model to elucidate the influential drilling parameters based on rate of penetration prediction. Geoenergy Sci. Eng. 231 (Part A), 212231. https://doi.org/10.1016/j.geoen.2023.212231.

Gan, C., Wang, Y., Cao, W.H., et al., 2025. Real-time formation drillability sensing-based hybrid online prediction method for the rate of penetration (ROP) and its industrial application for drilling processes. Control Eng. Pract. 164, 106487. https://doi.org/10.1016/j.conengprac.2025.106487.

Ghanitoos, H., Goharimanesh, M., Akbari, A., 2024. Prediction of drill penetration rate in drilling oil wells using mathematical and neurofuzzy modeling methods. Energy Rep. 11, 145–152. https://doi.org/10.1016/j.egyr.2023.11.006.

Giro, J.P., Zanella, A., Mourgues, R., et al., 2025. A novel adaptable analog material for simulating transversely isotropic rocks: Mechanical properties and applications. Int. J. Rock Mech. Min. Sci. 192, 106145. https://doi.org/10.1016/j.ijrmms.2025.106145.

Guan, Y., Zhu, H., Liu, Q., et al., 2023. A new model for evaluating rock drillability considering the rock plasticity and chip hold down effect caused by hydrostatic column pressure under high confining pressure. Geoenergy Sci. Eng. 227, 211806. https://doi.org/10.1016/j.geoen.2023.211806.

Hadi, F., Altaie, H., AlKamil, E., 2019. Modeling rate of penetration using artificial intelligent system and multiple regression analysis. In: Abu Dhabi International Petroleum Exhibition & Conference. https://doi.org/10.2118/197663-MS

Hamzaban, M.T., Rostami, J., Farrokhzadeh, S., et al., 2025. A critique on the correlations between rock abrasiveness, hardness, and mechanical characteristics: Implication of LCPC abrasiveness and breakability coefficients. Rock Mech. Rock Eng. 58 (2), 2487–2529. https://doi.org/10.1007/s00603-024-04184-y.

Heydari, K.E., Baghbanan, A., Hashemolhosseini, H., et al., 2025. Drilling-based rock strength estimation: A validated model for Mohr-Coulomb and a new theoretical-empirical approach for Hoek-Brown parameters. Results Eng. 28, 107388. https://doi.org/10.1016/j.rineng.2025.107388.

Ke, X., Sun, J., Yang, X., et al., 2023. Cutting mechanism of a special 3D concave-shaped PDC cutter applicable to the Weiyuan shale. J. Pet. Explor. Prod. Technol. 13 (6), 1435–1451. https://doi.org/10.1007/s13202-023-01621-6.

Koopialipoor, M., Asteris, P.G., Mohammed, A.S., et al., 2022. Introducing stacking machine learning approaches for the prediction of rock deformation. Transp. Geotechn. 34, 100756. https://doi.org/10.1016/j.trgeo.2022.100756.
Kor, K., Ertekin, S., Yamanlar, S., et al., 2021. Penetration rate prediction in heterogeneous formations: A geomechanical approach through machine learning. J. Petrol. Sci. Eng. 207, 109138. https://doi.org/10.1016/j.petrol.2021.109138.
Li, H., Zhong, R., Pel, L., et al., 2024. A new volumetric strain-based method for determining the crack initiation threshold of rocks under compression. Rock Mech. Rock Eng. 57 (2), 1329–1351. https://doi.org/10.1007/s00603-023-03619-2.
Liu, L., Ge, Z., Zhou, Z., et al., 2025a. A novel theory of abrasive water jet rock breaking under triaxial stress and the evolution mechanisms of rock damage. Int. J. Rock Mech. Min. Sci. 195, 106258. https://doi.org/10.1016/j.ijrmms.2025.106258.
Liu, Y., Liu, S., Zhang, J., et al., 2025b. Real-time estimation of geomechanical characteristics using drilling parameter data and LWD. Geoenergy Sci. Eng. 244, 213450. https://doi.org/10.1016/j.geoen.2024.213450.
Mahmood, J., Mustafa, G., Ali, M., 2022. Accurate estimation of tool wear levels during milling, drilling and turning operations by designing novel hyperparameter tuned models based on LightGBM and stacking. Measurement 190, 110722. https://doi.org/10.1016/j.measurement.2022.110722.
Oliver, G.M., Ly, C.V., Spence, G., et al., 2013. A new approach to measuring rock properties data from cores & cuttings for reservoir & completions characterization: An example from the Bakken Formation. In: SPE Annual Technical Conference and Exhibition, Brisbane, Australia.
Pacis, F.J., Ambrus, A., Alyaev, S., et al., 2023. Improving predictive models for rate of penetration in real drilling operations through transfer learning. J. Comput. Sci. 72, 102100. https://doi.org/10.1016/j.jocs.2023.102100.
Qu, F., Liao, H., Liu, J., et al., 2023. Establishment of data-driven multi-objective model to optimize drilling performance. Geoenergy Sci. Eng. 231, 212295. https://doi.org/10.1016/j.geoen.2023.212295.
Qu, F., Liao, H., Liu, J., et al., 2024. A novel well log data imputation methods with CGAN and swarm intelligence optimization. Energy 293, 130694. https://doi.org/10.1016/j.energy.2024.130694.
Sabri, M.S., Verma, A.K., Singh, T.N., 2025. Direct and indirect methods for uniaxial compressive strength estimation in various geo-environments: A review. J. Rock Mech. Geotech. Eng. 17 (12), 8241–8275. https://doi.org/10.1016/j.jrmge.2025.01.050.
Shao, R., Wang, H., Xiao, L., 2024. Reservoir evaluation using petrophysics informed machine learning: A case study. Artificial Intelligence in Geosciences 5, 100070. https://doi.org/10.1016/j.aiig.2024.100070.
Sharma, A., Burak, T., Nygaard, R., et al., 2025. Hybrid ROP modeling: Combining analytical and data-driven approaches for drilling. Geoenergy Sci. Eng. 251, 213877. https://doi.org/10.1016/j.geoen.2025.213877.
Sharma, G., Bandhu, K.C., 2024. Optimizing window size determination for improved forecast accuracy through pattern sequence similarity detection. In: 2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET). https://doi.org/10.1109/ACROSET62108.2024.10743809.
Sharma, S.R., Singh, B., Kaur, M., 2023. A novel approach of ensemble methods using the stacked generalization for high-dimensional datasets. IETE J. Res. 69 (10), 6802–6817. https://doi.org/10.1080/03772063.2022.2028582.
Shi, X., Chen, S., Meng, Y., et al., 2023. Suggestions for improvement and optimization of rock drillability determination methods. Acta Pet. Sin. 44 (9), 1562. https://doi.org/10.7623/syxb202309013 (in Chinese).
Srivastava, G.K., Vemavarapu, M.S.R.M., 2021. Drillability prediction in some metamorphic rocks using composite penetration rate index (CPRI) – an approach. Int. J. Min. Sci. Technol. 31 (4), 631–641. https://doi.org/10.1016/j.ijmst.2021.05.010.
Wang, M., Mitri, H.S., Zhao, G., et al., 2025. Data-based assessment of rock strengths and cuttability using the monitored parameters while drilling, tunneling, and mining. Tunn. Undergr. Space Technol. 165, 106901. https://doi.org/10.1016/j.tust.2025.106901.
Wang, S., Potyondy, D.O., Chu, W., et al., 2024. Investigation of meso-mechanical properties of Jinping dolomitic marble based on flat-joint model. J. Rock Mech. Geotech. Eng. 17 (6), 3464–3478. https://doi.org/10.1016/j.jrmge.2024.05.020.
Wang, X., Zhang, H., Wu, H., et al., 2021. Experimental study on the influence of components on rock drillability of conglomerate. In: 55th U.S. Rock Mechanics/Geomechanics Symposium. ARMA-2021-1113.
Xiong, C., Huang, Z.W., Shi, H.Z., et al., 2023. Performances of a stinger PDC cutter breaking granite: Cutting force and mechanical specific energy in single cutter tests. Pet. Sci. 20 (2), 1087–1103. https://doi.org/10.1016/j.petsci.2022.10.006.
Yang, H., Li, Z., Gao, L., et al., 2025a. Real-time inversion of formation drillability and concurrent speedup strategies for microdrilling time optimization. SPE J. 30 (4), 1806–1821. https://doi.org/10.2118/225417-PA.
Yang, Y., Cen, X., Ni, H., et al., 2025b. A highly accurate and robust prediction framework for drilling rate of penetration based on machine learning ensemble algorithm. Geoenergy Sci. Eng. 244, 213423. https://doi.org/10.1016/j.geoen.2024.213423.
Yang, H., Xian, X., Zhou, P., et al., 2025. Bit selection method based on the rock drillability characteristics and principal component analysis of rock-breaking indicators. Heliyon 11 (3), e42232. https://doi.org/10.1016/j.heliyon.2025.e42232.
Yu, X., He, M., Wang, H., et al., 2025. Determination of brittleness and anisotropy for rocks under different confining pressures using digital drilling. Geoenergy Sci. Eng. 246, 213645. https://doi.org/10.1016/j.geoen.2025.213645.
Zhang, D., Ma, T., Liu, Y., et al., 2025. Heterogeneous multi-task learning approach for rock strength prediction in real-time during drilling based on mud log data. Geoenergy Sci. Eng. 246, 213608. https://doi.org/10.1016/j.geoen.2024.213608.