

A rock-machine-muck multi-source information fusion method for tunnel boring machine tunneling parameter control

Guang ZHANG^a, Dongxin LIU^a, Qiuming GONG^{a*}, Zhaocai SANG^b, Minghao SUN^a

^a Key Laboratory of Urban Security and Disaster Engineering of Ministry of Education, Beijing University of Technology, Beijing 100124, China

^b Sinohydro Bureau 3 Co., Ltd., Xi'an 710024, China

*Corresponding author. E-mail: gongqiuming@bjut.edu.cn

© The Author(s) 2026. This article is published with open access at link.springer.com and journal.hep.com.cn

ABSTRACT The selection of tunnel boring machine (TBM) tunneling parameters predominantly relies on manual experience, which hinders the achievement of optimal control of the tunneling process. This study proposes a new control method that integrates multi-source information from the rock mass, the machine, and the muck to enhance tunneling performance while ensuring construction safety, thereby providing a scientific basis for TBM control. The proposed methodology comprises three key stages. First, a multi-source heterogeneous engineering database is constructed by systematically collecting TBM tunneling parameters, muck parameters, and rock mass parameters. Subsequently, two models are developed: a rock mass grade identification model and a tunneling performance prediction model, both based on the tunneling parameters and muck characteristics from the TBM's thrust-raising phase. These models are designed to identify the rock mass grade at the tunnel face and predict key performance parameters during the stable tunneling phase, respectively. Finally, by integrating the model outputs with a statistical characterization of engineering data, a graded control strategy tailored to different rock mass grades is formulated. Engineering application verification demonstrates that the proposed method achieves robust adaptability and effective control performance. This work provides a theoretical framework and practical guidance for auxiliary intelligent TBM tunneling in complex geological conditions.

KEYWORDS TBM, multi-source information fusion, rock mass identification, performance prediction, parameter control

1 Introduction

The full-face hard rock tunnel boring machine (TBM) is widely employed in tunnel construction due to its notable advantages in safety, efficiency, and excavation quality [1–3]. The key to safe and efficient TBM tunneling lies in the optimal matching between rock mass conditions and machine control parameters. Currently, TBM operators primarily rely on empirical judgment to assess face rock mass conditions and machine performance, dynamically adjusting control parameters to maintain a relatively stable operational state. However, this experience-dependent approach often struggles to ensure both safety and high efficiency under complex geological conditions.

Therefore, developing a scientific method for TBM parameter control is of significant importance for enhancing overall tunneling performance.

During TBM operation, sensors integrated into the machine system automatically collect tunneling parameters, which, to some extent, reflect the machine's status and rock-machine interaction mechanisms. Data-driven models built using machine learning have been applied to TBM face rock mass identification, performance prediction, and parameter decision. In the area of face rock mass identification, Nie et al. [4] and Jung et al. [5] utilized Artificial Neural Networks, for real-time rock mass identification using tunneling parameters. Zhu et al. [6] employed the Gini impurity index to select key TBM parameters as inputs to an AdaCost classifier for rock

grade identification. Liu et al. [7] used eight TBM parameters, FPI, and TPI as inputs to build an AdaBoost-CART model for rock mass identification. Hou et al. [8] implemented a Stacking ensemble learning classifier for this purpose. Guo et al. [9] established a Random Forest model using the data from the thrust-raising phase of a tunneling cycle to identify rock mass grade. Yin et al. [10] developed an ensemble model with cutterhead rotational speed, thrust, penetration, and torque as inputs for rock mass grade identification. Zhang et al. [11] applied the K-means algorithm for unsupervised clustering of TBM parameters to distinguish rock mass grades, then used these results as labels to train a support vector classifier for rock mass grade identification. Regarding TBM performance prediction, thrust and cutterhead torque are key feedback parameters. Wang et al. [12] proposed an integrated framework combining feature engineering and model development to predict the average thrust and torque during the stable phase of a complete tunneling cycle. Furthermore, Li et al. [13], Zhang et al. [14], Zhang et al. [15], Fu et al. [16], and Li et al. [17] developed deep learning models based on long short-term memory (LSTM), bidirectional gated recurrent unit with attention mechanism, causal-temporal graph convolutional network, and hybrid graph neural network and convolutional neural network architectures, respectively, using data from the thrust-raising phase to predict the mean thrust and torque in the subsequent stable phase. Kong et al. [18] and Yu et al. [19] constructed a Random Forest model and a multi-channel decoupled deep neural network model, respectively, using both tunneling and geological parameters to predict stable-phase mean thrust and torque. Advancing this field, Li et al. [20] proposed a mechanics-analysis-based data conversion method to predict stable-phase thrust and torque across different projects. However, the accuracy of the aforementioned rock mass identification models typically remains below 90%, while the coefficient of determination (R^2) for performance prediction models generally falls short of 0.9. In the realm of TBM parameter control, Zhang et al. [21] analyzed correlations between rock mass and TBM parameters to establish a rock-machine interaction perception model, adjusting advance rate and cutterhead rotational speed by comparing predicted and actual stable-phase parameters. Hou et al. [22] first established a safety premise by identifying rock mass grade and predicting geological hazard risks, then built an optimization model targeting penetration rate, rock breaking energy consumption, and cutter ring wear. This framework adaptively selects efficiency-priority or safety-priority strategies for setting advance rate and cutterhead speed based on rock mass conditions. Fu et al. [23] extracted features from the TBM thrust-raising phase, constructed an Osprey-Cauchy-Sparrow Search Algorithm-Improved Temporal Convolutional Network

model for parameter prediction, and proposed a control method based on thrust-raising phase data.

Although existing research has made progress in TBM face rock mass identification, performance prediction, and parameter control, most studies rely predominantly on tunneling parameters alone. A significant limitation is the general neglect of muck, a direct byproduct and information carrier of rock-machine interaction. The failure to fully integrate multi-source information, such as muck parameters, limits model performance. Muck, resulting from TBM rock breaking, exhibits morphological characteristics influenced not only by rock strength and integrity but also by TBM tunneling parameters. Field studies have demonstrated a significant correlation between muck parameters and rock-machine interaction [24]. Therefore, muck parameters provide an objective reflection of the *in-situ* rock mass information under the applied TBM parameters. Incorporating muck parameters can significantly enhance the interpretation of rock-machine interactions and is crucial for improving the accuracy of both rock mass identification and performance prediction. Traditionally, muck parameters obtained via conveyor belt sampling and sieve analysis suffer from discontinuous data and timeliness issues. However, Gong et al. [25,26] developed a real-time muck data analysis system based on deep learning, capable of acquiring muck images and characteristic parameters in real-time during tunneling. This advancement makes real-time rock mass identification and performance prediction feasible by integrating tunneling parameters with muck characteristics. Furthermore, existing approaches often treat identification, prediction, and control as separate tasks, lacking an integrated, systematic methodology for TBM parameter regulation. Additionally, the application and validation of such models in real-world field engineering remain relatively scarce.

This study, based on the Eping Water Diversion Project, proposes an integrated TBM parameter control method that synergistically combines tunneling parameters and muck characteristics. The research framework is as follows. First, a multi-source heterogeneous engineering database is constructed by systematically collecting and processing TBM operational data, muck parameters, and rock mass parameters through dynamic monitoring of the entire TBM tunneling process. Second, machine learning and deep learning algorithms are employed to develop integrated models for TBM rock mass grade identification and tunneling performance prediction, incorporating both tunneling parameters and muck features. Finally, by statistically analyzing field tunneling data and comprehensively considering both construction safety and tunneling performance, an adaptive, graded control strategy for TBM parameters tailored to different rock mass grades is formulated. The effectiveness of the proposed methodology is validated through field application.

2 Project background

2.1 Project overview

The Eping Water Diversion Project is located in Shiyan City, Hubei Province, with a total water diversion line length of approximately 14.26 km. The diversion tunnel spans about 9.59 km, of which the TBM-bored section accounts for approximately 9.42 km. The tunnel has an excavated diameter of 4.03 m and a final lined diameter of 3.2 m, with a longitudinal gradient of 1:10000. The project employs the “Wulinghao” double-shield TBM, jointly developed by China Railway Engineering Equipment Group and Sinohydro Bureau 3 Co., Ltd. The key technical parameters of the TBM are summarized in Table 1. The TBM comprises the host and the backup system. The host system is about 10 m long, while the entire machine has a total length of approximately 428 m and a total weight of about 900 t. Its minimum horizontal turning radius is approximately 500 m, as illustrated in Fig. 1.

2.2 Integration of the tunnel boring machine muck monitoring system

The muck generated during TBM tunneling is a direct

Table 1 Main technical parameters of TBM

Technical parameter	Design value
Maximum thrust (kN)	15833
Maximum cutterhead power (kW)	4×350
Cutterhead rotating speed (r/min)	0–11.3
Rated torque (kN·m)	2248
Cylinder stroke (mm)	1300
Normal stroke (mm)	600

result of rock–machine interaction, serving as a critical indicator of both *in-situ* rock mass conditions and the appropriateness of the tunneling parameters. This study employs a self-developed TBM muck monitoring system to acquire muck parameters [25]. This system primarily consists of a data acquisition subsystem, a communication and control subsystem, a database and data processing subsystem, and a software and visualization subsystem, as illustrated in Fig. 2.

Based on the structural characteristics of the “Wulinghao” double-shield TBM, the data acquisition subsystem was installed above the backup conveyor belt. This subsystem comprises functional components including a camera box, a volumetric flow meter, and a speed measurement wheel, as shown in Fig. 3. The

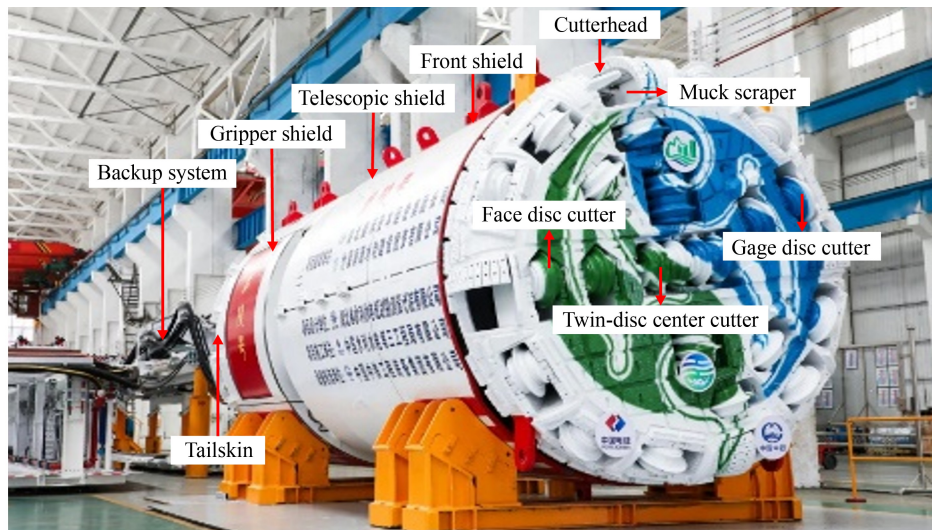


Fig. 1 Key components of double-shield TBM.

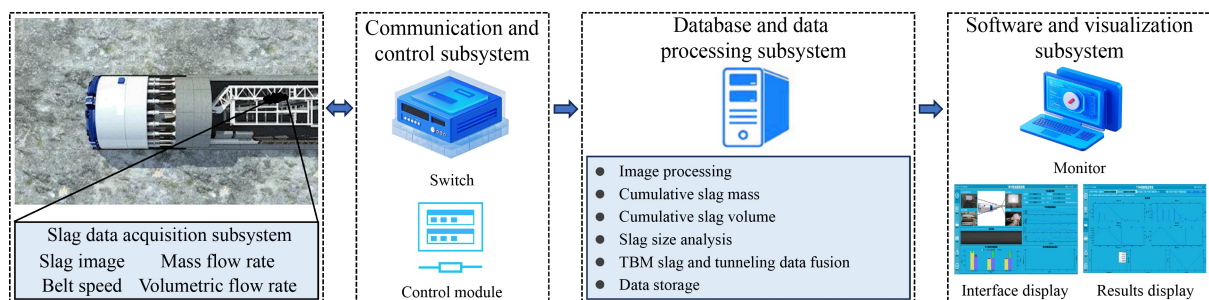


Fig. 2 Overall framework of the TBM muck monitoring system.

camera box, positioned approximately 26 m from the cutterhead and about 500 mm above the conveyor belt, continuously captures muck images. The volumetric flow meter, installed about 26.3 m from the cutterhead and roughly 350 mm above the belt, continuously scans the muck volume. The speed measurement wheel, mounted on a belt idler near the flow meter, measures the conveyor belt speed.

During muck discharge in this project, the conveyor belt speed was typically maintained at 2.5 m/s, resulting in a transport time of approximately 10 s from the cutterhead to the data acquisition subsystem. The time delay for muck transport does not significantly impact the monitoring process, as the rock mass conditions at the tunnel face remain stable over the corresponding short distance. Considering the duration of a complete TBM tunneling cycle, the system operates intermittently, with 30-s active data collection periods alternating with 5-min

pauses. During each active period, approximately 65 images are captured. Each image has a resolution of 2048×4096 pixels, with a pixel scale of 0.285 mm/pixel, corresponding to a physical capture area of $583.68 \text{ mm} \times 1167.36 \text{ mm}$.

3 Development of the multi-source database

3.1 Rock mass parameters

Rock mass parameters at the tunnel face in double-shield TBM tunnels are primarily acquired by combining muck analysis with geological mapping. Muck analysis involves collecting large rock fragments produced during TBM excavation. The lithology is identified through rock mineral analysis, while the point load strength of the rock is determined via on-site point load tests, which is subsequently converted into uniaxial compressive strength, as illustrated in Fig. 4. Geological mapping is conducted during scheduled TBM maintenance stops. The recorder enters the cutterhead area to conduct face mapping, involving photography and detailed recording of geological information. The recorded data primarily include lithology, rock weathering degree, joint parameters (spacing and orientation for each joint set), groundwater conditions, and instances of rockfall, as shown in Fig. 5.

The data for this study were collected from the TBM-driven section between chainage K9 + 359.25 m and K6 + 678.63 m (in the direction of decreasing chainage) of the diversion tunnel. This section spans approximately

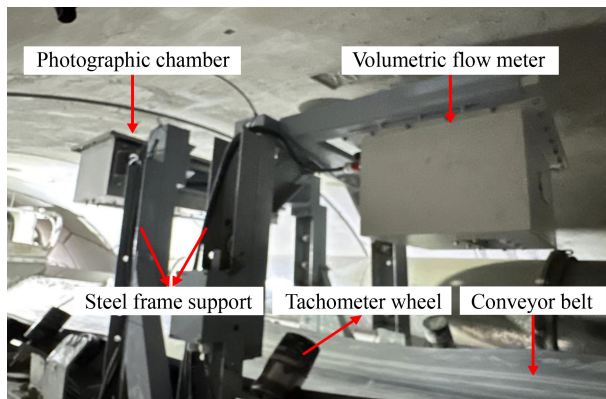


Fig. 3 Field deployment of the data acquisition subsystem.

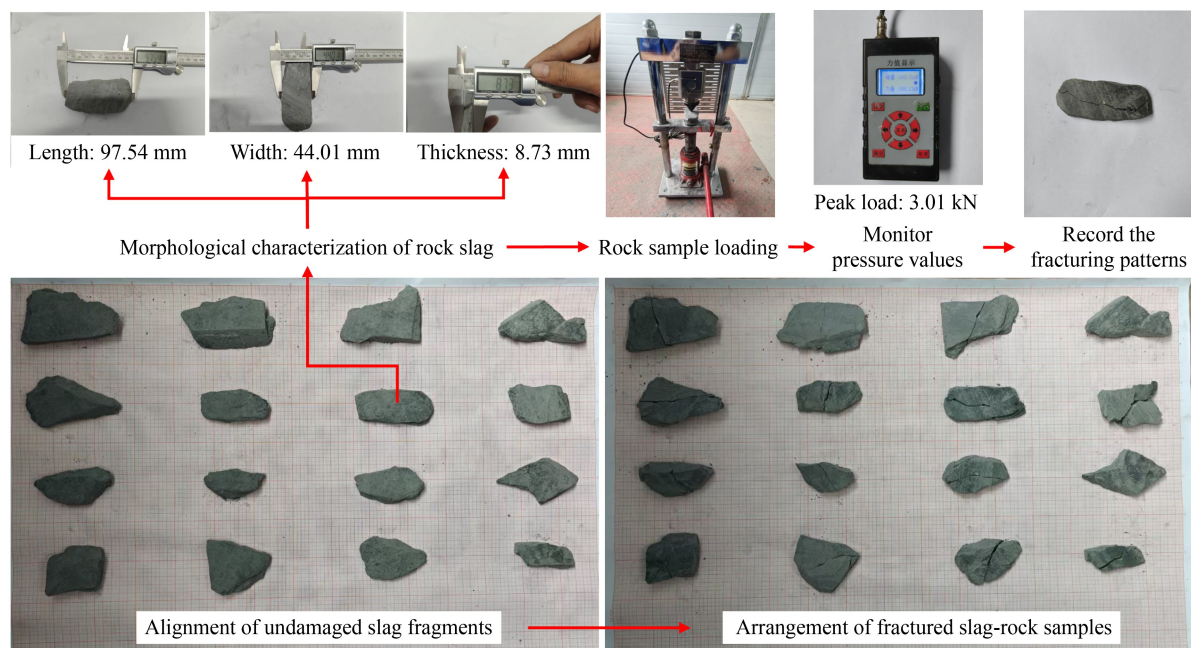


Fig. 4 Field point load test.

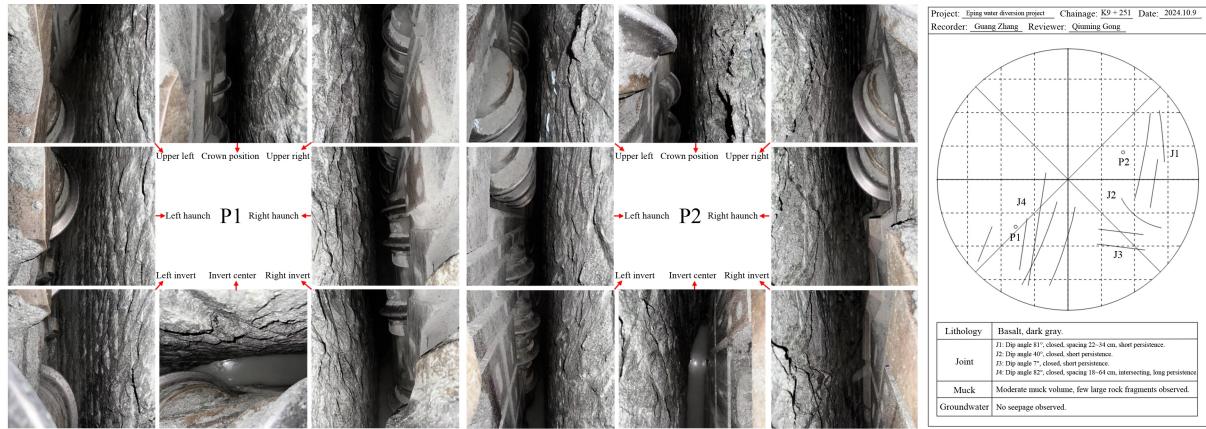


Fig. 5 Tunnel face sketch.

2.68 km, where the surrounding rock consists predominantly of basalt and carbonaceous slate. Based on the rock mass parameters acquired from field tracking, the rock mass at the face was classified according to the “Code for Geological Investigation of Water Resources and Hydropower Engineering” [27]. This tunnel section comprises four rock mass grades, from Class II to Class V. The distribution of these classes is as follows: Class II (25.53%), Class III (51.84%), Class IV (14.88%), and Class V (7.75%).

3.2 Tunnel boring machine tunneling parameters

The TBM muck monitoring system interfaces directly with the double-shield TBM’s programmable logic controller to acquire tunneling parameters in real-time at

a sampling frequency of 1 Hz. Data are stored daily, with each record comprising 396 indicators, including tunneling control parameters, equipment performance parameters and equipment operational status parameters. The raw TBM data underwent comprehensive preprocessing, which included extraction of complete tunneling cycles, outlier detection and handling, application of a moving average filter, and segmentation into thrust-raising and stable phases. This process yields time-series data of tunneling parameters for both the thrust-raising and stable phases [28], with the specific procedure detailed in Fig. 6. Based on previous research findings and insights from experienced TBM operators [28], the time-series data for total thrust, cutterhead torque, advance rate, and cutterhead rotational speed were selected as input features for the model.

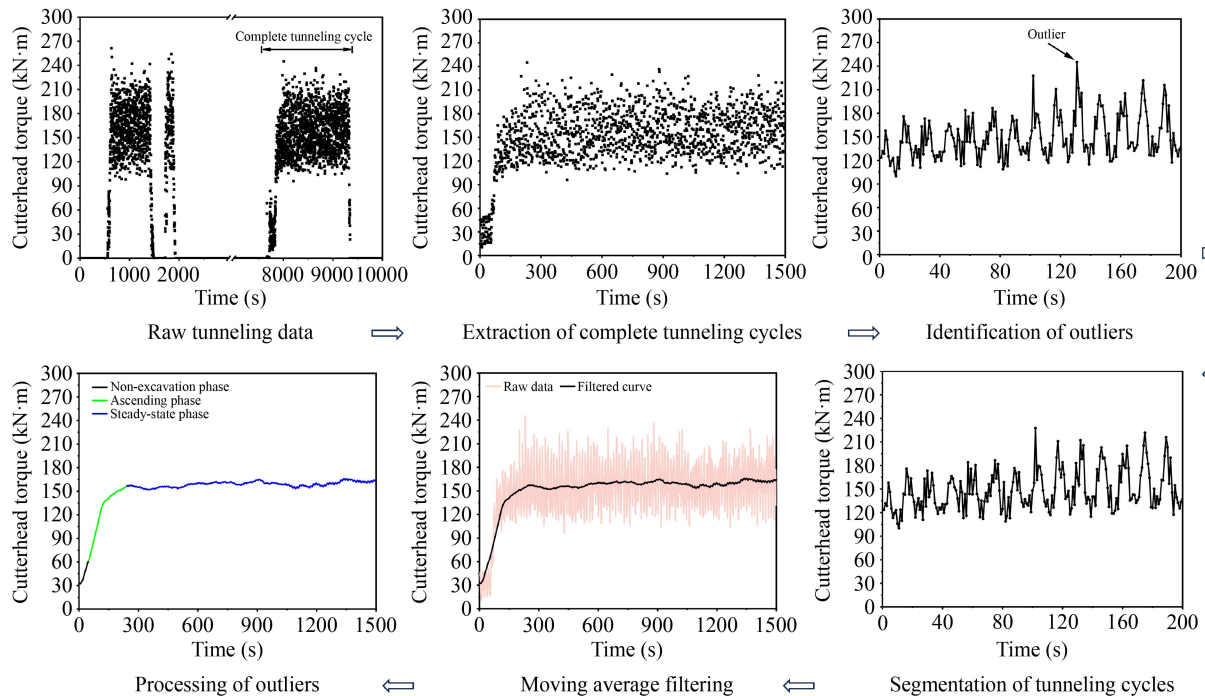


Fig. 6 Tunneling data processing procedure.

When a TBM operates in different rock masses, the total thrust and cutterhead torque during the stable phase are key parameters reflecting its tunneling performance. To provide a scientific basis for TBM operators' parameter selection decisions, the K-means clustering method was employed to classify the performance parameters of the stable phase, thereby enabling a quantitative assessment of TBM performance. The parameter settings for the K-means clustering algorithm were as follows: the number of clusters was determined as 2 based on the elbow method; to ensure reproducibility of the clustering results, a fixed random seed of 42 was set; the number of initializations was set to 10 to avoid local optima; and the maximum number of iterations was set to 300 to guarantee algorithm convergence. The resulting classification is presented in Table 2. Regarding cutterhead torque, the proportion of intervals classified as Grade A is higher than that of Grade B. This indicates that the cutterhead torque was maintained at a relatively low level under most working conditions, demonstrating superior tunneling performance of the TBM within the studied section.

Table 2 The clustering results

Parameter	Class	Range	Proportion
Thrust (kN)	A	2000–5000	48.79%
	B	5000–8000	51.21%
Cutterhead torque (kN·m)	A	90–442	60.15%
	B	442–1000	39.85%

3.3 Muck parameters

Each muck image captured by the TBM muck monitoring system was processed using image analysis algorithms to segment and identify the internal regions and outer contours of individual muck particles. This process enabled the extraction and storage of key particle size and shape parameters, including major axis length, intermediate axis length, perimeter, and area [25]. Based on the equivalence assumption with sieve analysis, the intermediate axis length of each muck particle was correlated with the sieve aperture size, and the particle mass was characterized by its projected area under the condition of equal minor axis length. In engineering application, the images from each continuous acquisition session were statistically analyzed as a complete set of sieve analysis data. In conjunction with the real-time rock mass classification results obtained from on-site synchronous tracking, the resulting particle size distribution curves were categorized according to rock mass class. For each rock mass condition, the arithmetic mean of the median particle sizes (d_{50}) of all corresponding curves was calculated, and the curve closest to this mean value was selected as the representative curve for that

class. Ultimately, representative particle size distribution curves under different rock mass conditions were generated, as shown in Fig. 7.

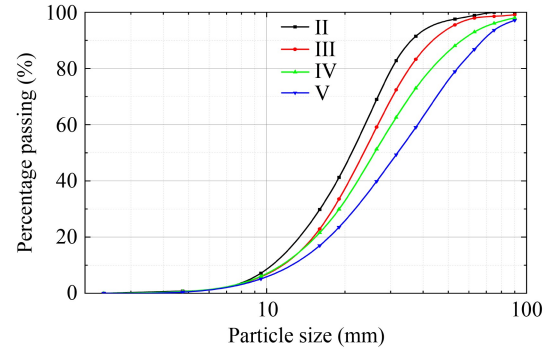


Fig. 7 Particle size distribution curve of muck fragments.

Six key parameters were calculated for each muck image set to comprehensively characterize particle size distribution, morphological features, and distribution curve properties: median particle size (d_{50}), maximum particle size (d_{max}), coefficient of uniformity (C_u), coarseness index (CI), mean aspect ratio (MAR), and coefficient of curvature (C_c) [29,30]. The corresponding calculation formulas are provided in Eqs. (1)–(6).

$$AR = L/W, \quad (1)$$

$$MAR = \frac{1}{n} \sum_{i=1}^n AR_i, \quad (2)$$

$$C_c = d_{30}^2 / (d_{10} \times d_{60}), \quad (3)$$

$$C_u = d_{60} / d_{10}, \quad (4)$$

$$X_i = \frac{W_i}{W_t} \times 100\%, \quad (5)$$

$$CI = \sum_{i=1}^n X_i, \quad (6)$$

where L and W represent the major and minor axis lengths of individual muck particles, respectively. n denotes the total number of particles in a sample set. d_{10} , d_{30} , and d_{60} represent the particle sizes at 10%, 30%, and 60% cumulative distribution on the particle size distribution curve. W_i represents the total mass of muck particles larger than the i th size fraction. W_t denotes the total mass of the muck sample. X_i represents the cumulative percentage retained on the i th sieve. The CI value was calculated using twelve specific size fractions from the distribution curve: 2.36, 4.75, 9.5, 16, 19, 26.5, 31.5, 37.5, 53, 63, 75, and 90 mm.

The median particle size (d_{50}) reflects the central tendency of the particle size distribution. The maximum particle size (d_{max}) indicates the upper limit of particle

sizes in the muck. The C_u quantifies the gradation uniformity of the particle size distribution. The CI represents the proportion of large-sized particles within the distribution. The MAR characterizes the overall morphological shape characteristics of the particles. The C_c describes the smoothness of the particle size distribution curve. Consequently, these six muck parameters were selected as input features for the model in this study.

3.4 Dataset generation

Both TBM tunneling parameters and muck characteristics reflect face rock mass conditions and equipment performance. By employing a multi-source data fusion approach, we integrated TBM tunneling parameters, muck parameters, and rock mass classification results to establish a multi-source heterogeneous engineering dataset. Specifically, a fixed-duration (10 s) delay compensation method was applied to align the muck parameters with the corresponding tunneling parameters recorded at the same moment. Discrete categorical variables for different rock mass classes were converted into binary vectors using One-hot encoding. To address scale differences among various parameters in the dataset and prevent potential adverse effects on model performance, data normalization was performed using the MinMaxScaler method. The calculation formula is as follows:

$$X_{01} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \quad (7)$$

where X_{01} is the normalized parameter, X is the original parameter, $X_{\max}$ and $X_{\min}$ are the maximum and minimum values of the original parameter, respectively.

Through the above data processing procedures, a final dataset containing 2025 samples was constructed. The distribution across rock mass grades II-V is as follows: 312, 1442, 168, and 103 samples, respectively. The main parameters of the dataset and their statistical characteristics are presented in Table 3. This curated dataset provides a solid foundation for training and testing both

the TBM rock mass identification model and the tunneling performance prediction model.

4 Tunnel boring machine tunneling parameter control methodology

4.1 Framework of the tunnel boring machine tunneling parameter control method

To achieve adaptive selection and control of TBM tunneling parameters under complex rock mass conditions, this study proposes a control methodology that integrates multi-source information encompassing the rock mass, the machine, and the muck. The core objective of this methodology is to determine the optimal advance rate and cutterhead rotational speed that simultaneously ensure construction safety and maximize tunneling performance. Specifically, geological conditions are identified through the TBM rock mass identification model, reflecting safety considerations during tunneling. Concurrently, the TBM performance prediction model is employed to forecast total thrust and cutterhead torque, providing insights into equipment performance. Based on the synergistic output of these models, a graded control strategy is established: In stable rock masses (Class II and III), higher tunneling parameters are adopted to enhance tunneling performance. In unstable rock masses (Class IV and V), conservative parameters are selected to mitigate the risk of collapse. This integrated approach forms a comprehensive tunneling parameter control methodology that seamlessly combines “rock mass identification-performance prediction-parameter regulation”. The overall framework of the methodology is illustrated in Fig. 8.

4.2 Tunnel boring machine rock mass identification

4.2.1 Model development

Extreme gradient boosting (XGBoost) is an efficient

Table 3 Statistical summary of key parameters

Parameter category	Parameter	Min	Max	Mean
Tunneling parameter	thrust (kN)	2673.15	7264.09	5022.79
	cutterhead torque (kN·m)	94.01	968.40	414.73
	advance rate (mm/min)	13.25	136.01	66.53
	cutterhead rotational speed (r/min)	0.94	10.94	7.19
Muck parameter	d_{50} (mm)	17.39	67.17	32.36
	$d_{\max}$ (mm)	55.86	203.22	134.10
	C_u	1.67	4.25	2.55
	CI	277.09	768.72	555.51
	MAR	1.44	1.78	1.53
	C_c	0.80	1.24	1.04

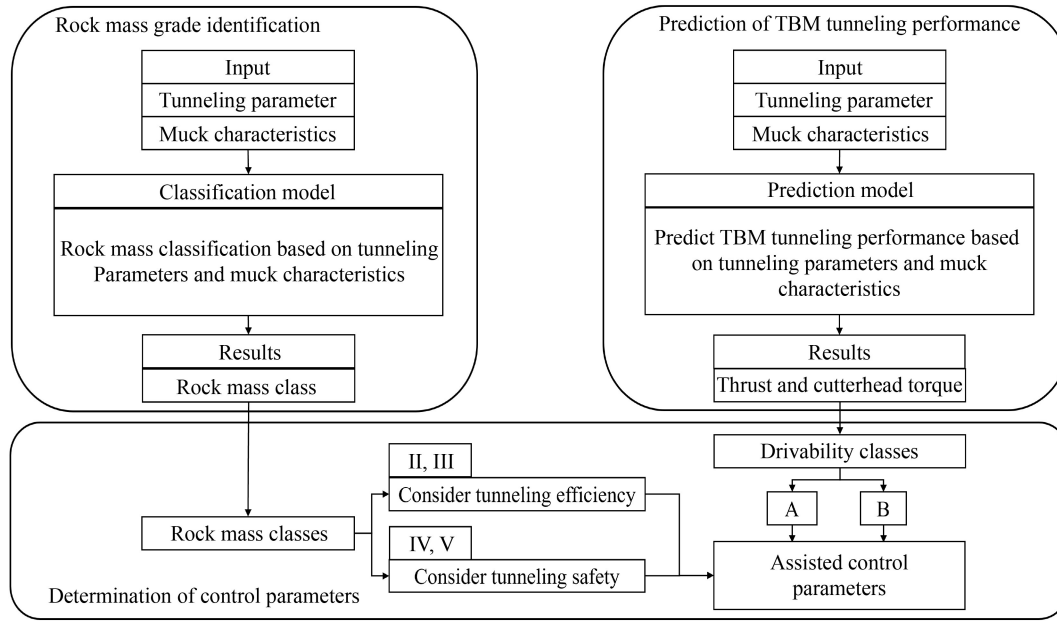


Fig. 8 Overall framework of the TBM auxiliary intelligent tunneling control method.

ensemble learning algorithm based on gradient boosting decision trees, recognized as one of the core methods for classification tasks in machine learning [31,32]. In this study, after comparing various classification models, XGBoost was selected as the core method due to its strong predictive performance, training efficiency, and generalization capability. The algorithm iteratively optimizes the weights of leaf nodes to minimize the objective function, which can be expressed as follows:

$$\begin{aligned}
 Obj(\Theta) &= \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^K \Omega(f_k) \\
 &= \sum_{i=1}^n l(\hat{y}_i, y_i) + \gamma T + \frac{1}{2} \lambda \|\omega\|^2, \quad (8)
 \end{aligned}$$

where n is the number of samples, K is the number of decision trees, $l(\hat{y}_i, y_i)$ is the loss function, $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2$ is the regularization term, y_i and $\hat{y}$ are the actual and predicted values of the i th sample, respectively, f_k is the output of the k th tree, T is the number of leaf nodes in each tree, ω represents the weights obtained after training, γ and λ are coefficients to be specified.

In this study, XGBoost models were developed to classify rock mass grades at the tunnel face using three distinct input configurations: utilizing only the tunneling parameter time series from the first 30 seconds of the thrust-raising phase, relying solely on muck characteristics, and integrating both tunneling parameters and muck characteristics. The models were implemented using the XGBoost package in Python, with their performance being highly dependent on appropriate hyperparameter selection. To enhance model performance, hyperpara-

meter optimization was conducted through Grid Search. The optimization process began by defining search ranges for key hyperparameters, including maximum tree depth, number of trees, learning rate, and subsample ratio-based on experimental experience and preliminary results. This was followed by an exhaustive evaluation of all possible hyperparameter combinations using cross-validation, where each combination was employed to train an XGBoost model for rock mass identification with accuracy as the primary evaluation metric. After comprehensive assessment, the hyperparameter configuration demonstrating optimal performance during cross-validation was selected as the final model setup.

4.2.2 Model training and evaluation

The dataset was partitioned into training, validation, and test sets in an 8:1:1 ratio to train and evaluate the XGBoost-based rock mass identification model. The model hyperparameters were configured as follows: maximum depth of 5, number of trees set to 200, learning rate of 0.01, and subsample ratio of 0.9. To comprehensively evaluate the model's predictive performance, multiple evaluation metrics commonly used in multi-class classification were employed, including Precision (P), Recall (R), $F1$ -score ($F1$), Accuracy (A), and Kappa coefficient (K). The computational expressions are provided in Eqs. (9)–(14). In the confusion matrix of the model identification results, Precision (P) represents the proportion of samples correctly predicted as positive among all samples predicted as positive. Recall (R) denotes the proportion of samples correctly predicted as positive among all samples with actual positive labels, reflecting the model's ability to identify

positive samples-the higher the recall, the stronger the model's recognition capability for positive samples. The *F1*-score is a comprehensive evaluation metric combining Precision and Recall, suitable for assessing model performance in multi-class classification tasks. Accuracy (*A*) represents the proportion of correctly classified samples among all samples, reflecting the overall classification performance of the model. The Kappa coefficient measures the agreement between model predictions and actual labels, providing a robust assessment of model reliability, particularly in scenarios with class imbalance. Its value ranges from -1 to 1 , with detailed interpretation guidelines for different agreement levels presented in Table 4. The mathematical expressions for the metrics are defined as follows:

$$P = \frac{TP}{TP + FP}, \quad (9)$$

$$R = \frac{TP}{TP + FN}, \quad (10)$$

$$F_1 = \frac{2 \times P \times R}{P + R}, \quad (11)$$

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \quad (12)$$

$$K = \frac{P_a - P_e}{1 - P_e}, \quad (13)$$

$$P_e = P_0 + P_1, \quad (14)$$

where true positive (*TP*) represents the number of positive instances correctly predicted as positive, false positive (*FP*) denotes the number of negative instances incorrectly predicted as positive, true negative (*TN*) indicates the number of negative instances correctly predicted as negative, false negative (*FN*) refers to the number of positive instances incorrectly predicted as negative, P_a represents the proportion of instances where predictions match actual labels, equivalent to accuracy, P_e indicates the probability of agreement between predictions and actual labels assuming statistical independence, P_0 denotes the product of the actual negative class proportion and predicted negative class proportion, P_1 represents the product of the actual positive class proportion and predicted positive class proportion.

Table 4 Interpretation of kappa statistic agreement levels

<i>K</i>	Agreement level
0.81–1.00	almost perfect
0.61–0.80	substantial
0.41–0.60	moderate
0.21–0.40	fair
0.01–0.20	slight
≤ 0.00	poor

Based on the TensorFlow 2.12.0 deep learning framework within the Python 3.9.13 programming environment, the experiments in this study were conducted on a hardware platform equipped with a 13th Gen Intel Core i5-13500H Central Processing Unit, 16 GB of Random Access Memory, and an NVIDIA GeForce RTX 4050 Laptop Graphics Processing Unit. Testing results indicate that the average inference time per sample for the XGBoost model on the test set is approximately 25 ms. Given that the acquisition frequency of TBM tunneling parameters is 1 Hz, this inference speed is fully capable of meeting real-time application requirements. Consequently, the model can perform rock mass classification during TBM tunneling, providing operators with timely references for parameter adjustments.

4.2.3 Analysis of rock mass identification results

Following the methodology established in Subsubsection 4.2.1, the TBM rock mass identification models with different input configurations were trained and subsequently evaluated using the test set from the constructed database. The corresponding confusion matrices are presented in Fig. 9, and the computed values for all evaluation metrics are summarized in Table 5.

Figure 9 and Table 5 reveal notable differences in model performance on the test set depending on the input features. The model utilizing integrated tunneling and muck parameters achieved the highest *F1*-scores across all rock mass grades (II to V), with values of 85.67%, 95.67%, 85.68%, and 85.71%, respectively. This performance significantly surpasses that of models relying on single-feature inputs. The Kappa coefficient for this integrated model reached 84.78%, which, according to standard interpretation guidelines, indicates an “almost perfect” agreement between the model predictions and the true labels. When using only tunneling parameters as input, the model performed satisfactorily for Class II and III rock masses but demonstrated markedly inferior recognition capability for Class IV and V rock masses. In the more competent Class II and III rock masses, characterized by higher strength and less joint development, the TBM operates under relatively stable conditions, resulting in tunneling parameters that exhibit clear, consistent relationships with rock mass grades. In contrast, for Class IV and V rock masses, which exhibit lower strength and poorer integrity, the operational priority shifts from maximizing efficiency to ensuring safety. This leads to the adoption of more conservative control strategies for the tunneling parameters, which consequently weakens the discernible relationship between these parameters and the rock mass grade, thereby reducing the performance of models based solely on tunneling parameters. The model based solely on muck characteristics demonstrated the lowest overall

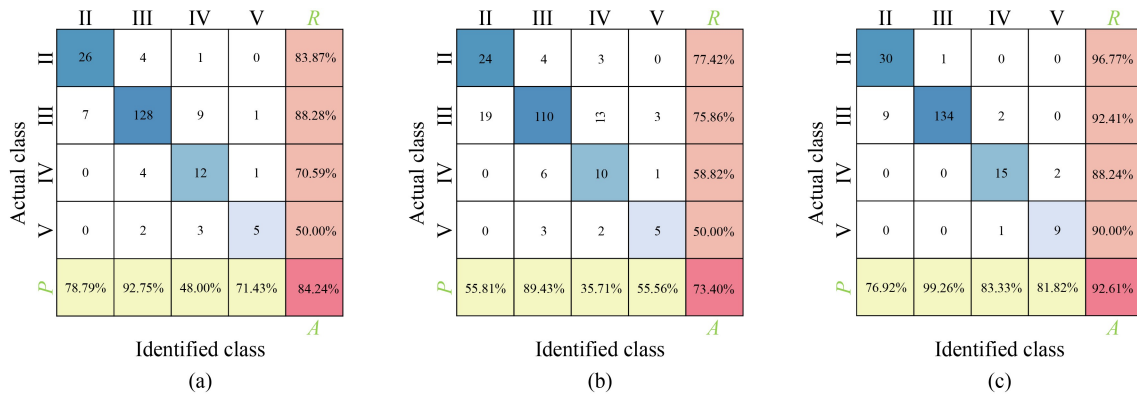


Fig. 9 Confusion matrices of the identification model with different input features: (a) tunneling parameter; (b) muck characteristic parameter; (c) integrated parameters.

Table 5 Comparison of *F1*-scores and Kappa coefficients for different models

Model input	F1				K
	II	III	IV	V	
Tunneling parameter	81.25%	90.48%	57.14%	58.82%	66.99%
Muck characteristic parameter	64.86%	82.09%	44.44%	52.63%	48.97%
Integrated parameters	85.67%	95.67%	85.68%	85.71%	84.78%

performance. This is primarily because muck characteristics mainly reflect the fragmentation outcome rather than the intrinsic rock mass properties such as joint development and rock strength. In contrast, tunneling parameters capture the dynamic response process of rock-machine interaction, encompassing mechanical feedback induced by variations in rock mass conditions. Muck parameters alone cannot incorporate critical information such as dynamic adjustments and equipment status, making it difficult to distinguish the modulation effect of TBM control parameters from the geological properties of the rock mass. This ultimately limits the model's performance under complex rock mass conditions. The findings above demonstrate that the TBM rock mass identification model integrating both tunneling parameters and muck characteristics significantly enhances the performance of the XGBoost algorithm in identifying rock mass grades.

4.3 Tunnel boring machine tunneling performance prediction

4.3.1 Model development

The Gated Recurrent Unit (GRU) neural network [28], derived as a variant of the LSTM network, offers advantages including a simpler architecture, fewer hyperparameters, and lower computational complexity. Its unit structure is illustrated in Fig. 10. The BiGRU enhances the standard GRU by incorporating a bidirectional structure, demonstrating superior capability in handling time-series problems. The architecture of a

BiGRU unit is shown in Fig. 11.

Building upon previous algorithm comparisons and experimental validation [24], this study adopts the BiGRU as the core modeling method, chosen for its ability to capture bidirectional temporal dependencies and to model both long- and short-term relationships in sequential data-properties that align well with the time-series characteristics of TBM tunneling parameters. Leveraging the bidirectional temporal modeling capability of the BiGRU neural network to process muck parameters in a time-series format. Three distinct TBM

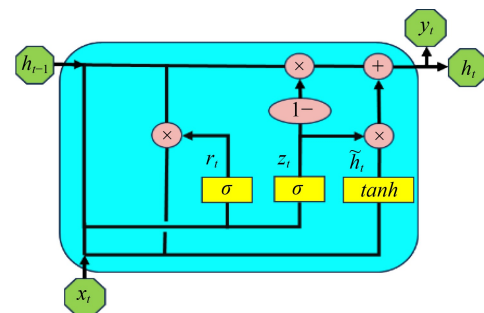


Fig. 10 GRU network unit structure.

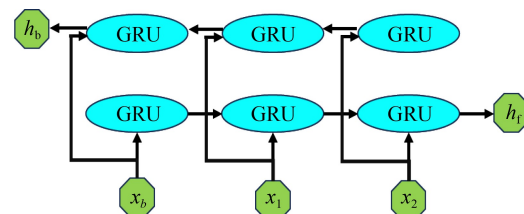


Fig. 11 BiGRU network structure.

tunneling performance prediction models were developed using different input configurations: tunneling parameter time series from the first 30 s of the thrust-raising phase, time-series formatted muck characteristic parameters, and integrated tunneling and muck parameters. The models were implemented using the Keras framework with a TensorFlow backend in Python.

4.3.2 Model training and evaluation

The dataset was partitioned into training, validation, and test sets using an 8:1:1 ratio for developing and evaluating the TBM tunneling performance prediction model. The model was trained using adaptive optimization algorithms combined with regularization techniques. The model architecture incorporates a three-layer BiGRU structure with hidden units sequentially configured as 128, 64, and 32 to enable hierarchical extraction of temporal features. Each BiGRU layer is followed by normalization and Dropout layers to enhance generalization capability. The final output layer produces predictions for both total thrust and cutterhead torque. For parameter optimization, the Adam optimizer was employed with mean squared error serving as the loss function. During training, validation loss was dynamically monitored to implement early stopping and learning rate decay mechanisms. The final model configuration parameters were set as follows: batch size of 32, maximum training epochs of 100, and $L2$ regularization coefficient of 0.001. To evaluate the model's predictive performance, two metrics commonly used in regression analysis were adopted: mean absolute error (MAE) and mean absolute percentage error ($MAPE$). The computational expressions are provided below:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (15)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i}, \quad (16)$$

where y_i represents the true value, $\bar{y}$ denotes the mean of the true values, $\hat{y}_i$ indicates the predicted value from the model, N signifies the total number of samples.

The BiGRU model was tested under the same hardware environment, and the results show that its average inference time per sample is approximately 40 ms. This inference speed also meets the real-time requirements of the 1 Hz data acquisition frequency in TBM construction, enabling simultaneous performance prediction during TBM tunneling and providing decision support for parameter regulation.

4.3.3 Analysis of tunneling performance prediction results

Following the methodology established in Subsubsection

4.3.1, the TBM tunneling performance prediction models with different input features were trained and evaluated using the test set. The prediction results under different input configurations are presented in Fig. 12, while the corresponding evaluation metric values are summarized in Table 6.

Figure 12 and Table 6 reveal significant differences in the performance of the prediction models depending on the input features. When utilizing integrated tunneling and muck parameters as input, the model demonstrated optimal performance in predicting both the total thrust and cutterhead torque of the TBM. The MAE were 203.96 kN and 33.25 kN·m, respectively, and the $MAPE$ were 4.14% and 9.46%, significantly lower than those of models relying on single-feature inputs. Furthermore, as shown in Figs. 12(e) and 12(f), the predicted curves from this integrated model align closely with the measured curves, exhibiting the smallest fluctuation range in absolute errors, which indicates excellent stability and prediction accuracy. This study demonstrates that the multi-source input integrating tunneling parameters and muck characteristics can fully exploit and utilize the dynamic response features of cutterhead-rock mass interaction embedded within the TBM operational data. This approach effectively captures the dynamic fluctuations in tunneling parameters caused by variations in rock mass conditions. Simultaneously, by incorporating the rock fragmentation characteristics represented by muck parameters, the model captures crucial supplementary information about rock mass properties. This significantly enhances the prediction accuracy and reliability of TBM tunneling performance under complex geological conditions, further validating the effectiveness of the multi-source data fusion strategy in improving the prediction robustness and engineering applicability of rock-machine interaction models.

4.4 Selection of tunnel boring machine tunneling parameters

4.4.1 Selection of tunneling parameters considering tunneling performance

Under Class II and Class III rock mass conditions, where the rock mass integrity is relatively good, the TBM tunneling process is generally stable. Consequently, the TBM operator typically selects a higher advance rate and cutterhead rotation speed to control the tunneling process, aiming to maximize TBM tunneling efficiency. The total thrust force and cutterhead torque are key parameters reflecting the rock mass boring ability and equipment tunneling performance, which are closely related to the TBM advance rate and cutterhead rotation speed. Figure 13 illustrates the variation of advance rate with the total thrust force under Class II and Class III rock mass conditions.

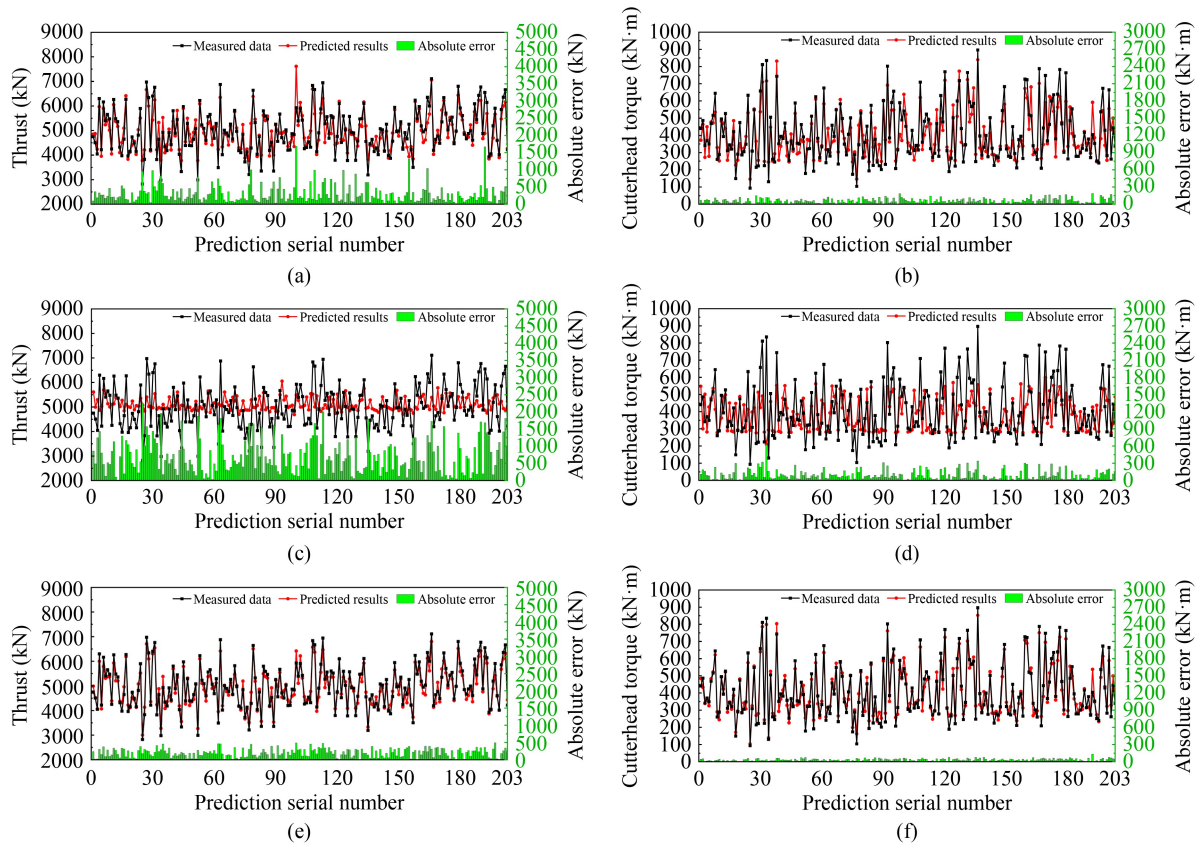


Fig. 12 Prediction results of TBM tunneling performance under different input features: (a) thrust prediction with tunneling parameters as input features; (b) cutterhead torque prediction with tunneling parameters as input features; (c) thrust prediction with muck particle parameters as input features; (d) cutterhead torque prediction with muck particle parameters as input features; (e) thrust prediction with fused parameters as input features; (f) cutterhead torque prediction with fused parameters as input features.

Table 6 Evaluation metrics of model prediction performance under different input features

Model input	Thrust (kN)		Cutterhead torque (kN·m)	
	MAE (kN)	MAPE (%)	MAE (kN·m)	MAPE (%)
Tunneling parameter	304.47	6.36	61.21	16.53
Muck characteristic parameter	718.60	15.28	99.47	26.02
Integrated parameter	203.96	4.14	33.25	9.46

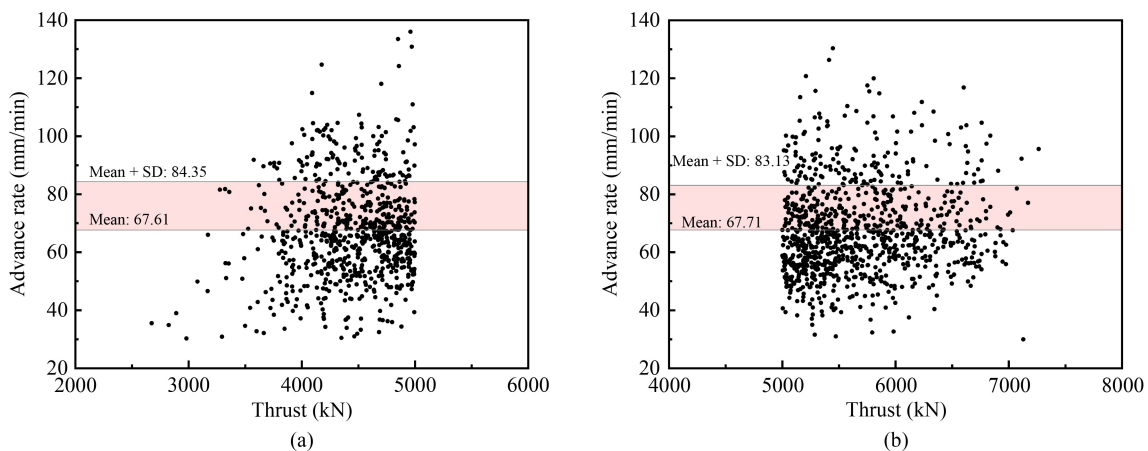


Fig. 13 Variation of penetration rate with total thrust force for Class II and Class III rock masses (A/B thrust classification threshold: 5000 kN based on clustering results in Table 2): (a) thrust of Class A; (b) thrust of Class B.

According to the statistical analysis results, the average advance rate and its standard deviation in the Grade A thrust range are 67.61 and 16.74 mm/min, respectively, while those in the Grade B thrust range are 67.71 and 15.42 mm/min, respectively. This indicates that the average advance rates in the two thrust grades are similar, but the standard deviation of the advance rate in the Grade B thrust range is smaller, suggesting less fluctuation in advance rate and a more stable tunneling process. Furthermore, the thrust force in the Grade A range is relatively low, yet its advance rate is essentially equivalent to that in the Grade B range. This reflects that the rock mass conditions corresponding to the Grade A range are more favorable for TBM tunneling. In contrast, the thrust force in the Grade B range is higher, indicating that the rock mass integrity or strength of the strata being excavated is greater. The cutterhead must overcome higher rock mass shear and compressive strengths during excavation, requiring increased thrust to maintain an advance rate comparable to the Grade A range, thereby achieving stable and efficient tunneling.

Based on the characteristics of the data statistical distribution and engineering experience, and aiming to improve tunneling performance while ensuring TBM tunneling safety, the safe and efficient advance rate range is defined based on the statistical data of advance rate. The average advance rates from both the Grade A and Grade B thrust ranges are used jointly to define the lower threshold for efficient tunneling performance, which are 67.61 and 67.71 mm/min, respectively. To enhance engineering applicability and robustness, the lower boundary of the range is uniformly set at 70 mm/min. This value can, with a high degree of confidence, cover the initial state of the efficient tunneling mode under Class II and Class III rock mass conditions. The determination of the upper limit is based on the more stable Grade B thrust range, where the sum of its average and standard deviation is 83.13 mm/min. This value is

defined as the upper boundary. It can cover at least 90% of the stable and efficient tunneling data in the population while controlling the Type I error rate, and simultaneously effectively eliminate the influence of non-representative data included due to the higher data dispersion in the Grade A thrust range. Therefore, considering the robustness of the strategy and incorporating a safety margin, the upper limit is set at 80 mm/min. This ensures tunneling performance while significantly reducing the construction risks associated with excessive tunneling speed. In summary, the control parameter range for the advance rate during efficient TBM tunneling in Class II and Class III rock masses is determined to be 70–80 mm/min.

4.4.2 Selection of tunneling parameters considering tunneling safety

Under Class IV and Class V rock mass conditions, the strength and integrity of the rock mass decrease significantly. The TBM tunneling process becomes susceptible to the stability of the rock mass, where the risk of face collapse becomes the core issue restricting safe and efficient TBM tunneling. Therefore, the TBM operator adopts a more conservative control strategy by appropriately reducing the advance rate and cutterhead rotation speed to maintain face stability and control surrounding rock deformation. Simultaneously, the operator monitors the variations in total thrust force and cutterhead torque in real-time to assess the dynamic response characteristics of the rock mass and potential instability risks. Based on this, this study adopts total thrust force and cutterhead torque as safety condition monitoring indicators, and achieves safe TBM tunneling through reasonable control of advance rate and cutterhead rotation speed. Figure 14 shows the variation of advance rate with total thrust force under Class IV and Class V rock mass conditions.

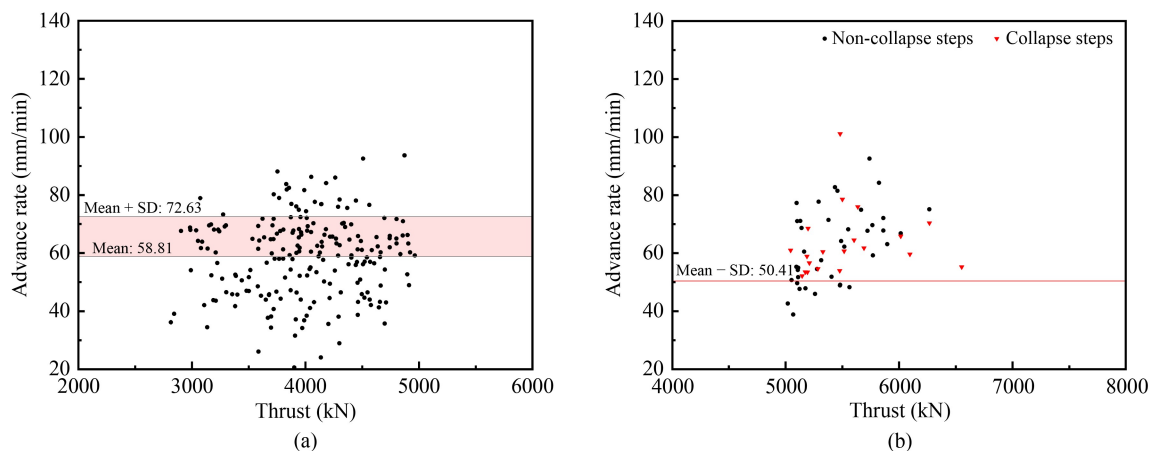


Fig. 14 Variation of penetration rate with total thrust force for Class IV and Class V rock masses (A/B thrust classification threshold: 5000 kN based on clustering results in Table 2): (a) thrust of Class A; (b) thrust of Class B.

According to the statistical analysis results, the average advance rate and its standard deviation in the Grade A thrust range are 58.81 and 13.82 mm/min, respectively, while those in the Grade B thrust range are 62.79 and 12.38 mm/min, respectively. The average advance rates in Grades A and B are similar, but the standard deviation in Grade A is larger, reflecting slightly greater fluctuation in its advance rate. However, combined with engineering practice and construction risk analysis, the Grade B thrust range corresponds to a higher total thrust force. The shear and compression actions exerted by the TBM cutterhead on the rock mass are stronger, disrupting the *in-situ* stress equilibrium of the rock mass and potentially inducing local collapse or face instability. The risk points of collapse within the Grade B thrust range in Fig. 14 further validate this correlation. Specifically, when the thrust force exceeds a certain critical threshold, the self-stabilizing capacity of the surrounding rock is insufficient, leading to deformation accumulation until instability occurs. To optimize the control parameters while ensuring tunneling safety, this study proposes a graded control strategy based on the characteristics of the data statistical distribution and the construction risk threshold: For the Grade A thrust range, the same method used for Class II and III rock masses is applied to determine the advance rate range. The calculated lower and upper limits for the Grade A thrust range under Class IV and V rock mass conditions are 58.81 and 72.63 mm/min, respectively, which are finalized as 60–70 mm/min. Although this advance rate is lower than the efficient range for Class II and III rock masses, the corresponding total thrust force is less than 5000 kN. This helps to maximize the stability of the rock mass while achieving both safe and efficient tunneling. For the Grade B thrust range, a more conservative control strategy is necessary to prevent the risk of rock mass instability induced by the higher thrust force. Therefore, the upper limit for the advance rate in the Grade B range is set based on the difference between the average and the standard deviation, which is 50.41 mm/min, and rounded to 50 mm/min based on the principle of engineering robustness. The setting of this control value has clear data support and engineering pertinence: as shown in Fig. 14, the collapse points recorded during TBM tunneling are all distributed above this control line, indicating that this advance rate control value can effectively distinguish between stable and potentially unstable conditions, verifying the rationality of using this statistical value as the upper speed limit. Furthermore, referring to the technical specifications of this TBM, the theoretical operational range for advance rate under Class IV and V rock mass conditions is typically recommended as 30–90 mm/min. However, the statistical analysis in this study shows that the risk of collapse increases significantly when the TBM advance rate exceeds 50 mm/min. To balance tunneling efficiency

and construction safety, the final controlled advance rate range for the Grade B thrust range is defined as 30–50 mm/min.

Similarly, the cutterhead rotation speed of the TBM can be determined based on the cutterhead torque. By analyzing the variation patterns of cutterhead rotation speed with cutterhead torque under different rock mass grades and cutterhead torque grades, reasonable values of TBM cutterhead rotation speed are defined, as shown in Table 7.

Table 7 Control parameters for TBM auxiliary intelligent tunneling

Rock mass condition	Grade	Advance rate (mm/min)	Cutterhead rotational speed (r/min)
II, III	A	70–80	8–9
	B		
IV, V	A	60–70	6–7
	B	30–50	2–4

4.5 On-site engineering verification of the method

The TBM operator sets and adjusts control parameters such as advance rate and cutterhead rotation speed based on the machine's operating state during the initial start-up phase of a tunneling cycle. First, data from this start-up phase is used to determine the rock mass grade and the grade levels of the total thrust force and cutterhead torque. Then, based on the established standards, the advance rate and cutterhead rotation speed are set. The following three field cases demonstrate this process.

In the tunneling cycle from K6 + 668.78 m to K6 + 668.21 m, the actual rock mass grade was Class IV, and the TBM rock mass identification model identified the rock mass as Class IV. The TBM tunneling performance model predicted a total thrust force of 4591.47 kN and a cutterhead torque of 406.65 kN·m. According to the classification standards in Table 2, both parameters were classified as Grade A. Consequently, referring to Table 7, the recommended ranges were an advance rate of 60–70 mm/min and a cutterhead rotation speed of 6–7 r/min. During actual tunneling, the operator set the advance rate and cutterhead rotation speed to 70.71 mm/min and 6.38 r/min, respectively, with corresponding measured values of 4666.55 kN for total thrust and 433.36 kN·m for cutterhead torque. The absolute error between predicted and measured thrust was 75.08 kN, with an *MAPE* of 1.61%, and the absolute error for torque was 26.71 kN·m, with an *MAPE* of 6.16%. Both *MAPE* values were lower than the average *MAPE* of the integrated model on the test set, which was 4.14% for thrust and 9.46% for torque. The parameters selected by the operator were close to those recommended by the proposed method, with the advance rate slightly exceeding the upper limit while the cutterhead rotation speed fell within the recommended range. This resulted in a tunneling step length

approaching the normal value of 0.6 m, achieving safe, efficient, and stable TBM advancement in this cycle.

In the tunneling cycle from K6 + 659.79 m to K6 + 659.46 m, the actual rock mass grade was Class IV, and the predicted rock mass grade was IV, with predicted total thrust force and cutterhead torque of 5443.04 kN and 589.53 kN·m, respectively. According to Table 2, these were classified as Grade B, indicating a risk of surrounding rock instability in this tunneling step. Therefore, referring to Table 7, the recommended advance rate was 30–50 mm/min and the cutterhead rotation speed was 2–4 r/min. However, the operator set the actual advance rate and cutterhead rotation speed to 56.49 mm/min and 5.89 r/min, respectively, with corresponding measured values of 5413.72 kN for total thrust and 605.22 kN·m for cutterhead torque. The absolute error between predicted and measured thrust was 29.32 kN, with an *MAPE* of 0.54%, and the absolute error for torque was 15.69 kN·m, with an *MAPE* of 2.59%. Both *MAPE* values were well below the average *MAPE* of the integrated model. The control parameters chosen significantly exceeded the upper limits of the recommended ranges, subsequently triggering a collapse incident. This resulted in a tunneling step length noticeably shorter than the normal step, consistent with the analysis prediction that excessive parameters would aggravate surrounding rock disturbance. This case conversely demonstrates that deviation from the control parameter ranges significantly increases engineering construction risks.

In the tunneling cycle from K6 + 657.23 m to K6 + 656.71 m, the actual rock mass grade was Class IV, and the predicted rock mass grade was IV, with predicted total thrust force and cutterhead torque of 5160.86 kN and 379.42 kN·m, respectively. According to Table 2, the total thrust force was classified as Grade B and the cutterhead torque as Grade A, indicating potential surrounding rock instability risk in this step. Therefore, referring to Table 7, the recommended advance rate range was 30–50 mm/min and the cutterhead rotation speed was 6–7 r/min. The operator set the actual advance rate and cutterhead rotation speed to 44.93 mm/min and 5.55 r/min, respectively, with corresponding measured values of 5246.55 kN for total thrust and 398.08 kN·m for cutterhead torque. The absolute error between predicted and measured thrust was 85.69 kN, with an *MAPE* of 1.63%, and the absolute error for torque was 18.66 kN·m, with an *MAPE* of 4.69%. Both *MAPE* values were close to or below the average *MAPE* of the integrated model. The advance rate selected during actual tunneling fell within the recommended range, while the cutterhead rotation speed was slightly below the recommended lower limit. This resulted in a tunneling step length approaching the normal value, effectively reducing disturbance to the surrounding rock and avoiding a

collapse incident, further verifying the effectiveness of the proposed TBM tunneling parameter control method.

It is noteworthy that while the specific values of TBM tunneling parameters may vary across projects with different geological conditions and TBM types, the framework of the TBM tunneling parameter control method demonstrates universal applicability across different engineering contexts.

5 Conclusions

To address the challenge of adaptive selection of TBM parameters under varying rock mass conditions, this study proposed a TBM tunneling parameter control method based on multi-source information encompassing rock-machine-muck. Relying on the Eping Water Diversion Project, we conducted on-site tracking, collection, and processing of TBM tunneling parameters, muck parameters, and rock mass parameters, thereby constructing a multi-source heterogeneous engineering database. On this basis, a rock mass identification model based on XGBoost and a tunneling performance prediction model based on BiGRU were developed, forming an integrated TBM control methodology that combines rock mass identification, performance prediction, and parameter regulation. The main conclusions are as follows.

1) Through dynamic monitoring of the entire TBM construction process, tunneling parameters, muck images, and rock mass parameters were systematically collected, establishing a multi-source heterogeneous engineering database. This database integrates real-time TBM tunneling parameters, muck characteristics, face mapping data, and rock strength test results. Effective fusion of these multi-source, multi-modal data was achieved through data preprocessing, providing a solid data foundation for the intelligent perception of rock-machine interaction and the selection of tunneling parameters.

2) The multi-source data model integrating both tunneling parameters and muck characteristics significantly outperformed models using a single data source. For the TBM rock mass identification model, the model integrating tunneling and muck parameters achieved an overall accuracy of 92.61% and a Kappa coefficient of 84.78%. Notably, it achieved *F1*-scores exceeding 85% specifically under Class IV and V rock mass conditions, demonstrating strong generalization capability and reliability. In the TBM tunneling performance prediction model, the model incorporating tunneling and muck parameters reduced the *MAPE* for predicting total thrust force and cutterhead torque to 4.14% and 9.46%, respectively, indicating a significant improvement in prediction accuracy. This validates the enhanced capability of multi-source data fusion to accurately capture the dynamic response of tunneling performance under complex geological conditions.

3) Based on the results of TBM rock mass identification and tunneling performance prediction, combined with the statistical characteristics of the data and engineering experience, a graded control strategy suitable for different rock mass conditions was proposed. For Class II and III rock masses, aiming to improve tunneling performance, the control parameter ranges are an advance rate of 70–80 mm/min and a cutterhead rotation speed of 8–9 r/min. For the Grade A thrust range in Class IV and V rock masses, targeting both safety and efficiency, the recommended ranges are an advance rate of 60–70 mm/min and a cutterhead rotation speed of 6–7 r/min. For the high-risk Grade B thrust range in these poorer rock masses, aiming to mitigate collapse risk, the recommended advance rate range is 30–50 mm/min and the cutterhead rotation speed is 2–4 r/min. Field applications demonstrated the method's good adaptability and effective control, providing a reliable basis for achieving adaptive control of TBMs under varying rock mass conditions.

The core contribution of this research lies in overcoming the limitations of traditional single-data-source models by establishing an integrated TBM parameter control method encompassing “Rock Mass Identification-Performance Prediction-Parameter Regulation.” This provides a theoretical foundation and practical guidance for enhancing TBM tunneling safety and performance optimization in complex geological conditions. It should be noted that this study is based on the Eping Water Diversion Project, where the geological conditions primarily consist of basalt and carbonaceous slate. The proposed method has not been fully validated under extreme geological conditions such as karst formations, fault fracture zones, or high *in-situ* stress. Furthermore, the control parameter ranges were constructed using statistical methods and represent empirically derived intervals, which have certain limitations when addressing dynamic factors such as cutter wear. Future research will involve collecting additional TBM tunneling data under extreme geological conditions to validate the adaptability of the proposed method in karst formations, fault fracture zones, and high *in-situ* stress environments. In parallel, future work will explore online learning and adaptive control mechanisms based on multi-modal data: specifically, incremental learning algorithms will be introduced to enable real-time integration of new operational data and dynamic model updates; additionally, cutter wear will be incorporated as an optimization objective within the control model, facilitating dynamic parameter optimization and self-tuning through online feedback. These enhancements aim to improve the real-time adaptive control capability of the method under diverse geological conditions.

Acknowledgements We would like to express our sincere gratitude to the management of Eping Water Diversion Project from Sinohydro Bureau 3 Co., Ltd. for their assistance in field data acquisition.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Competing interests The authors declare that they have no competing interests.

References

1. Liu Z B, Wang Y C, Li L, Fang X L, Wang J Z. Realtime prediction of hard rock TBM advance rate using temporal convolutional network (TCN) with tunnel construction big data. *Frontiers of Structural and Civil Engineering*, 2022, 16(4): 401–413
2. Wu F, Gong Q M, Li Z G, Huang L, Sun Q W, Liu Q Y. Study on tunnelling performance of dual-mode shield TBM by cutterhead working performance and tunnelling difference comparison: A case in Shenzhen metro line 12. *Rock Mechanics and Rock Engineering*, 2023, 56(7): 5315–5333
3. Wu L J, Li X, Yuan J D, Wang S J. Real-time prediction of tunnel face conditions using XGBoost random forest algorithm. *Frontiers of Structural and Civil Engineering*, 2023, 17(12): 1777–1795
4. Nie S W, Xue L, Jia G P, Ma Y, Chen J, Hou J Z. Identification of surrounding rock in TBM excavation with deep neural network. In: 2019 2nd International Conference on Artificial Intelligence and Big Data. Chengdu: IEEE, 2019, 251–255
5. Jung J H, Chung H, Kwon Y S, Lee I M. An ANN to predict ground condition ahead of tunnel face using TBM operational data. *KSCE Journal of Civil Engineering*, 2019, 23(7): 3200–3206
6. Zhu M Q, Zhu H H, Wang X, Cheng P P. Study on CART-based ensemble learning algorithms for predicting TBM tunneling parameters and classing surrounding rockmasses. *Chinese Journal of Rock Mechanics and Engineering*, 2020, 39(9): 1860–1871 (in Chinese)
7. Liu Q S, Wang X Y, Huang X, Yin X. Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data. *Tunnelling and Underground Space Technology*, 2020, 106: 103595
8. Hou S K, Liu Y R, Yang Q. Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning. *Journal of Rock Mechanics and Geotechnical Engineering*, 2022, 14(1): 123–143
9. Guo D, Li J H, Jiang S H, Li X, Chen Z Y. Intelligent assistant driving method for tunnel boring machine based on big data. *Acta Geotechnica*, 2022, 17(4): 1019–1030
10. Yin X, Liu Q S, Huang X, Pan Y C. Perception model of

- surrounding rock geological conditions based on TBM operational big data and combined unsupervised-supervised learning. *Tunnelling and Underground Space Technology*, 2022, 120: 104285
11. Zhang Q L, Liu Z Y, Tan J R. Prediction of geological conditions for a tunnel boring machine using big operational data. *Automation in Construction*, 2019, 100: 73–83
 12. Wang X, Zhu H H, Zhu M Q, Zhang L Y, Ju J W. An integrated parameter prediction framework for intelligent TBM excavation in hard rock. *Tunnelling and Underground Space Technology*, 2021, 118: 104196
 13. Li J H, Li P X, Guo D, Li X, Chen Z Y. Advanced prediction of tunnel boring machine performance based on big data. *Geoscience Frontiers*, 2021, 12(1): 331–338
 14. Zhang L M, Guo J, Fu X L, Tiong R L K, Zhang P H. Digital twin enabled real-time advanced control of TBM operation using deep learning methods. *Automation in Construction*, 2024, 158: 105240
 15. Zhang Q L, Yang B Y, Zhu Y W, Guo C, Jiao C, Cai A M. Prediction method of TBM tunneling parameters based on Bi-GRU-ATT model. *Advances in Civil Engineering*, 2022, 2022: 3743472
 16. Fu X L, Pan Y, Zhang L M. A causal-temporal graphic convolutional network (CT-GCN) approach for TBM load prediction in tunnel excavation. *Expert Systems with Applications*, 2024, 238: 121977
 17. Li H B, Zeng Z G, Li X, Yao M. Graph-convolutional neural networks for predicting tunnel boring machine performance. *Automation in Construction*, 2025, 180: 106436
 18. Kong X X, Ling X Z, Tang L, Tang W C, Zhang Y F. Random forest-based predictors for driving forces of earth pressure balance (EPB) shield tunnel boring machine (TBM). *Tunnelling and Underground Space Technology*, 2022, 122: 104373
 19. Yu H G, Qin C J, Tao J F, Liu C L, Liu Q S. A multi-channel decoupled deep neural network for tunnel boring machine torque and thrust prediction. *Tunnelling and Underground Space Technology*, 2023, 133: 104949
 20. Li X, Li H B, Du S Z, Jing L J, Li P Y. Cross-project utilisation of tunnel boring machine (TBM) construction data: A case study using big data from Yin-Song diversion project in China. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 2023, 17(1): 127–147
 21. Zhang N, Li J B, Jing L J, Li P Y, Xu S T. Study and application of intelligent control system of TBM tunneling parameters. *Tunnel Construction*, 2018, 38(10): 1734–1740 (in Chinese)
 22. Hou S K, Liu Y R, Yu J L, Zhang R J, Cheng L, Gao C F. Intelligent decision-making for TBM tunnelling control parameters using multi-objective optimization. *Journal of Rock Mechanics and Geotechnical Engineering*, 2025, 17(5): 2943–2963
 23. Fu K, Xue Y G, Qiu D H, Wang P. Short-impending decision-making for TBM tunneling control based on loading phase data. *Tunnelling and Underground Space Technology*, 2025, 164: 106819
 24. Gong Q M, She Q R, Hou Z S, Jiang H T, Peng Q. Experimental study of TBM penetration in marble rock mass under high geostress. *Chinese Journal of Rock Mechanics and Engineering*, 2010, 29(12): 2522–2532 (in Chinese)
 25. Gong Q M, Zhou X X, Liu Y Q, Han B, Yin L J. Development of a real-time muck analysis system for assistant intelligence TBM tunnelling. *Tunnelling and Underground Space Technology*, 2021, 107: 103655
 26. Zhou X X, Gong Q M, Liu Y Q, Yin L J. Automatic segmentation of TBM muck images via a deep-learning approach to estimate the size and shape of rock chips. *Automation in Construction*, 2021, 126: 103685
 27. GB 50487-2008. Code for Engineering Geological Investigation of Water Resources and Hydropower. Beijing: China Planning Press, 2009
 28. Zhang G, Gong Q M, Xie X F, Pei C Y, Shang C. Interval prediction of TBM parameters in stable excavation sections based on Bootstrap-COA-BiGRU model. *Modern Tunnelling Technology*, 2025, 62(2): 121–131 (in Chinese)
 29. Gu C C, Bao W J, Zhang X D, Chen F Y, Jiang H X, Liu S Y. The mechanism of rock-breaking using disc cutter assisted by pre-slotting. *Engineering Fracture Mechanics*, 2025, 322: 111165
 30. Huang L, Gong Q M, Wang J, Ma H S, Zhou X X, Xie X F, Song H J. A real-time rock mass class identification model of the tunnel face based on TBM tunneling and the corresponding muck characteristic parameters. *International Journal of Rock Mechanics and Mining Sciences*, 2025, 188: 106057
 31. Chen T Q, Guestrin C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco: ACM, 2016, 785–794
 32. Nie Y Q, Zhang Q, Hou L L, Du L J, Zhao X L, Xue Y J, Lin Z C, Cao X X, Wang Z X. TBM rock mass classification using XGBoost and interpretable machine learning. *Advanced Engineering Informatics*, 2025, 66: 103459