Geomechanics and Engineering, *Vol. 40, No. 3 (2025) 217-227*
https://doi.org/10.12989/gae.2025.40.3.217

Rock mass classification method of TBM tunnel face based on driving parameters

Yimin Xia[1,2], Dun Wu[1], Jie Ke[1] and Laikuang Lin*[1,2]

[1]*College of Mechanical and electrical Engineering, Central South University, 932 Lushan South Road, Changsha, China*
[2]*State Key Laboratory of Precision Manufacturing for Extreme Service Performance, Central South University, 932 Lushan South Road, Changsha, China*

(*Received December 23, 2023, Revised January 28, 2025, Accepted January 31, 2025*)

Abstract. Rock mass classification of TBM tunnel is an important reference index to analyze the TBM performance and determine the support mode. It is of great significance for the safe and efficient construction of tunnel to quickly recognize the rock mass classes of tunnel face. In this paper, the data preprocessing is carried out by statistical method, and the input eigenvalues of model are selected through correlation analysis. A LIBSVM rock mass classification model is established, with the penetration rate, thrust and cutterhead torque as the model inputs, and the rock mass classes obtained by the Hydropower Classification (HC) method as the model output. The test results show that the average precision of LIBSVM model is 0.919, and the recall rate of Class II rock mass is as high as 0.980. The data overlap of rock mass of Class III, IV and V is an important factor affecting the precision of the model. Analyzing the importance of the 6 input eigenvalues of model in each two-classifier, the results show that the mean value of the penetration rate, the thrust and the cutterhead torque play a major role, and the standard deviations of their variation play a supplementary role.

Keywords: driving parameters; LIBSVM; Rock mass classification; Tunnel Boring Machine (TBM)

1. Introduction

Tunnel Boring Machine (TBM) is a kind of large-scale special engineering machine for full-section tunnel excavation, which is widely used in tunnel engineering. Compared with the drill-blasting method, TBM has higher construction efficiency and a safer working environment. However, TBM is not as adaptable to geological conditions as drill-blasting method, which requires the driver to adjust the operation parameters according to the surrounding rock in the tunneling process. Operation parameters that do not match the geological conditions will aggravate cutter wear, reduce penetration rate, and even cause tunnel face collapse and machine jams (Kim *et al*. 2022, Mahmoodzadeh *et al.* 2022, Wang *et al.* 2022). Therefore, it is of great significance to obtain the rock mass classes of the face in time to guide TBM construction.

There are many rock mass classification methods, among which the rock mass rating (RMR) system (Kim *et al.* 2020), the rock mass quality system (Q-system) (Kim *et al.* 2022) and the basic quality (BQ) method (Zhou and Yang 2021) have been widely recognized and applied. The BQ method is a method obtained by analyzing the correlation between various indicators and selecting the most representative ones after referring to many rock mass classification methods including RMR and Q-system (Zhou and Yang 2021). It combines the qualitative characteristics of rock mass basic quality with the basic quality index BQ of rock mass to classify the rock mass. The qualitative characteristics include hardness degree of rock and intactness degree of rock mass. The value of quantitative index BQ is calculated by weighting the rock saturated uniaxial compressive strength (R_C) and the rock mass intactness index (K_V).

The Hydropower Classification (HC) method (The National Standards Compilation Group of People's Republic of China 2009), widely used in hydropower and water conservancy projects in China, categorizes rock masses based on indices such as rock saturated uniaxial compressive strength, rock mass intactness index, structural plane conditions, and groundwater presence. This method provides a comprehensive assessment by integrating geological and hydrological characteristics, making it particularly suitable for projects with complex geological settings.

Many scholars have carried out relevant research on the above rock mass classification methods. Some ones have discussed the correlation between different rock mass classification methods (Fernandez-Gutierrez *et al.* 2017, Hoek and Brown 1997). And some ones have optimized the rock mass classification method to make it suitable for TBM construction (Alber 2000, Barton 1999). In view of the construction characteristics of TBM, von Preinl *et al*. (2006) proposed the rock mass excavability (RME) indicator, which can predict the rock mass excavability by quantifying TBM performance. Sapigni *et al*. (2002) analyzed the correlation between penetration rate and rock mass classification systems including RMR, Q and Q_{TBM}. Hassanpour *et al.* (2011) established an equation between

*Corresponding author, Associate Professor
E-mail: linlaikuang@csu.edu.cn

ISSN: 2005-307X (Print), 2092-6219 (Online)

rock quality designation (RQD), uniaxial compressive strength and field penetration index (FPI). The previous research was mainly based on correlation analysis and regression analysis. The influence of rock mass classification on TBM performance was analyzed by establishing explicit relations between input parameters of rock mass classification system or between them and FPI (Liu *et al*. 2017, Jain *et al.* 2016). With the development of artificial intelligence and machine learning technology, such as Classification and Regression Tree (CART) (Salimi *et al.* 2018), Artificial Neural Network (ANN) (Gholamnejad and Tayarani 2010) and Support Vector Regression (SVR) (Mahmoodzadeh *et al*. 2021) are widely used in the research of TBM performance prediction and rock mass classification. Wang Chao *et al*. (2018), combined with a large amount of tunneling data, found that the cutterhead torque has a significant correlation with the class of rock mass, the cutterhead rate and the penetration rate, and established the NSVR prediction model of cutterhead torque based on this. Salimi *et al*. (2018, 2019) compared the precision of the TBM performance prediction models established by nonlinear regression, regression trees and genetic algorithms, and obtained equations between rock property parameters, classes of rock mass and FPI.

However, traditional measurement methods can only obtain the features of the exposed surrounding rock, and cannot accurately reflect the features of the tunnel face, the timeliness of which is poor. TBM tunneling is essentially a process of interaction between rock and cutter (Yang *et al*. 2016). The driving parameters of TBM will be affected by geological conditions during the process of cutting the rock mass by the cutter, which makes it possible to get the rock mass information of the face by analyzing the driving parameters. Liu *et al*. (2020) proposed a hybrid algorithm combining Back Propagation Neural Network (BPNN) and Simulated Annealing (SA) to realize the recognition of rock mass parameters under various rock formations. These studies have successfully obtained the rock mass information of the tunnel face from the TBM driving parameters, but for the TBM driver, rock mass class is more conducive to the selection of operation parameters. Zhang *et al*. (2019) used K-mean++ algorithm to cluster the massive TBM driving data and divided them into five rock mass categories. Based on this potential rock mass classification method, a support vector classification (SVC) model with an average classification precision of 98.6% is trained. Liu *et al.* (2020) combined CART and AdaBoost algorithm to build a rock mass classification model based on HC method. The model takes the eigenvalues calculated from penetration rate, cutterhead rate, thrust and cutterhead torque as input, and obtains a classification precision of 86.5%. These models have achieved good classification accuracy, but the input parameters of the model are not discussed and filtered. By analyzing the influence of input parameters on rock mass classification and simplifying the input parameters of the model, the same classification precision can be obtained with less input parameters.

In this paper, the data collected from a TBM tunneling section in western China are analyzed, and a LIBSVM based rock mass classification model is trained by using the filtered eigenvalues. These data include operating parameters, the poses of TBM and rock mass classes. And the classes of rock mass are obtained using the HC method. At last, the ranking scores of feature c_p (Xue *et al.* 2018) used in Gaussian kernel support vector machine recursive feature elimination (SVM-RFE) are calculated to analyze the impact of each input eigenvalue on the model.

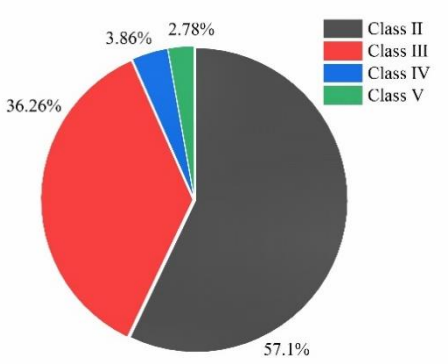


Fig. 1 Percentage distributions of rock mass classifications in the dataset

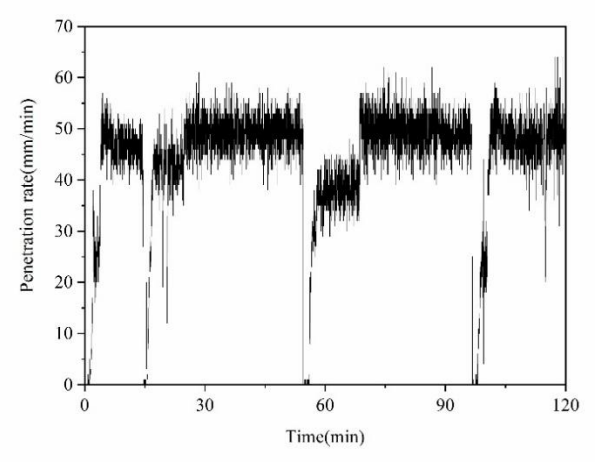
Table 1 Tunnel boring machines specifications

Parameter	Value
Cutterhead diameter (mm)	7030
TBM weight (t)	600
Number of cutters	4 center cutters (432 mm)
	29 face cutters (483 mm)
	13 edge cutters (483 mm)
Cutterhead power (kW)	350×8=2800
Cutterhead rate (rpm)	0-8
Cutterhead nominal torque (kN·m)	4410@5.5 rpm
Max thrust force of cutterhead (kN)	27489@350 bar
Thrust cylinder stroke (mm)	2000

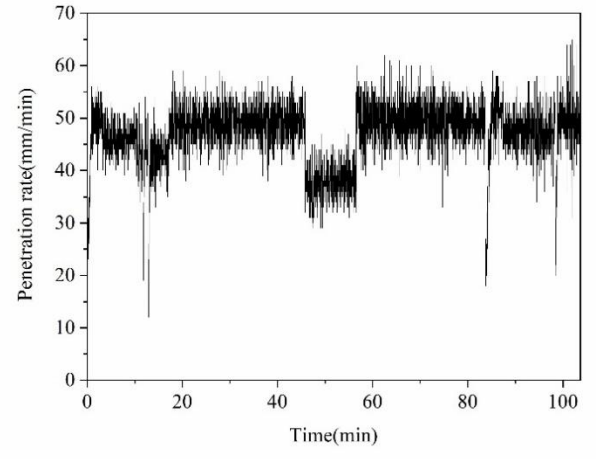
2. Establish dataset and training model

2.1 Geological conditions and TBM equipment

The data used in this paper comes from a TBM tunneling bid section from western China. This section of the tunnel is above 30 km in total, and the TBM construction section is nearly 20 km. The TBM section is located in a low mountainous area with bare bedrock. The ground along the line has an elevation of 775 to 1000 m. The terrain is not very undulating, and the relative elevation difference is generally 5 to 35 m. The local maximum elevation difference is up to 55 m. The total terrain is high in the east and low in the west. The lithology of the TBM construction section is tuffaceous sandstone of Carboniferous system and Devonian system. The rock mass is fresh, complete and stable. It is dominated by Class II and III rock masses, partially interspersed with Class IV and V rock masses. The percentage distribution of rock mass category data in the study interval is shown in Fig. 1. The



(a) Raw data



(b) Processed data

Fig. 2 Variation of penetration rate in the working state

excavation diameter of this TBM is 7.03 m, 46 tools are installed, with a total power of 4600 kW. The specific parameters are shown in Table 1. TBM will record the driving parameters generated during the excavation, with a sampling frequency of 1 Hz.

2.2 Data pre-processing

During the construction process, TBM will be suspended due to moving gripper shoe, surrounding rock support, cutter replacement, maintenance, etc., but the driving parameters during this period will still be recorded. These large amounts of non-working state data are of no value to rock mass classification and can be eliminated using the binary discriminant function proposed by Wang (2018). The driving data under normal working conditions includes two parts of data in the start-up phase and the stable phase (Zhang *et al*. 2019). Compared with the data in the start-up phase, the data in the stable phase is more stable and can better reflect the lithology of the face. According to experience, the diving parameters of 300 s length after the breakpoint caused by the use of the binary discriminant function are eliminated (Liu *et al.* 2020). Before and after the treatment, the curve of the penetration rate of the 2-hour length is shown in Fig. 2.

Table 2 Specific parameters of Gaussian distribution of driving parameters in Class II rock mass

Driving parameters	R^2	Peak numbers	μ	σ	[μ-2σ, μ+2σ]
Th(kN)	0.98186	2	17075.90	541.99	[15991.91, 18159.88]
			13966.89	2032.16	[9902.58, 18031.21]
T(kN·m)	0.94344	2	3032.28	394.41	[2243.46, 3821.09]
			1793.26	728.17	[336.92, 3249.59]
PR(mm/min)	0.9807	2	46.21	20.32	[5.57, 86.84]
			16.33	3.68	[8.98, 23.69]
PRev(mm/r)	0.96742	1	6.30	3.42	[-0.5, 13.13]
P_b(Bar)	0.94561	1	74.59	4.96	[64.67, 84.51]

As can be seen from Fig. 2, the driving parameters after processing still contain some singular values that need to be eliminated. Take the driving parameters in Class II rock mass as an example, the data distribution of thrust (Th), cutterhead torque (T), penetration rate (PR), penetration per revolution (PRev) and drive pressure of main belt conveyor (P_b) are shown in Fig. 3. They all conform to the Gaussian distribution, and the specific parameters of the Gaussian distribution function are shown in Table 2.

According to the 2 Sigma principle, only the data within the range of 2σ on both sides of the peak are retained. In Class II rock mass, the effective data interval of Th is [9902.58, 18159.88], T is [336.92, 3821.09], PR is [5.57, 86.84], PRev is [-0.5, 23.69], and P_b is [64.67, 84.50]. The effective data intervals of the above five driving parameters in 4 rock mass classes are calculated respectively, and the data points outside the interval are eliminated.

In the process of tunneling, TBM will produce violent vibration, which makes collected driving parameters contain noise. It is not conducive to improve the precision of rock mass classification to use the noise data directly. So, a certain width of time window needs to be selected to extract the mean value and standard deviation of the driving parameters as the model input. The cutterhead rate is in the range of 0.5-8 rpm, so half of the longest cutter head rotation period, namely 60 s, is selected as the width of the time window. In order to balance the sample proportion of the 4 kinds of rock mass, 120,000 data points were extracted from each type of rock mass, and 6,000 samples were uniformly collected by moving window with 20 s step to form a sample set. The values 1 to 4 are used as the labels for the 4 classes of rock mass of II to V in the classification model.

Considering the specific needs of rock engineering projects, especially in water conservancy and hydropower industries, the Hydropower Classification (HC) method was developed as an engineering geological classification system. It evaluates the overall stability of the surrounding rock and provides guidance for excavation and support design (Liu *et al.* 2017).

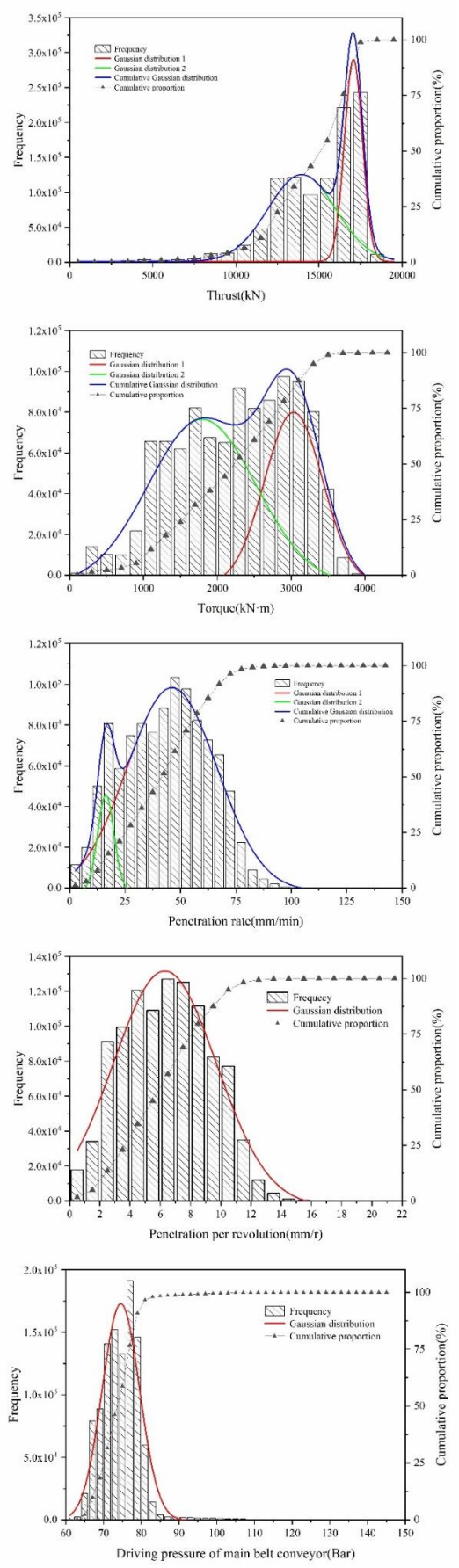


Fig. 3 Frequency distribution histogram of driving parameters in Class II rock mass

Table 3 Rock mass classification based on HC method

Class	Cumulative score T	Strength-stress ratio S
I	$H > 85$	> 4
II	$85 \geqslant H > 65$	> 4
III	$65 \geqslant H > 45$	> 2
IV	$45 \geqslant H > 25$	> 2
V	$H \leqslant 25$	-

The HC method calculates a cumulative score H to classify rock masses, incorporating five factors:

The cumulative score H is expressed as

$$H = A + B + C + D + E \tag{1}$$

where H is the rock mass rating of HC method, A, B, C, D, E are the ratings of rock strength, rock mass intactness degree, discontinuity conditions, groundwater condition and the main discontinuity plane attitude, respectively.

Rock mass classification in the HC method also considers the strength-stress ratio **(*S*)** to account for the influence of stress on the stability of surrounding rock. The strength-stress ratio is calculated as

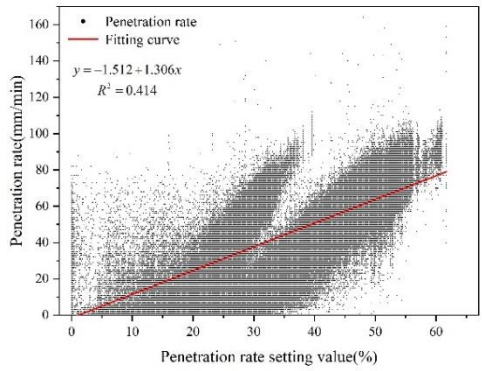
$$S = \frac{R_c \cdot K_v}{\sigma_m} \tag{2}$$

where S is the strength-stress ratio, σ_m is the maximum principal stress of surrounding rock (MPa). R_C is rock saturated uniaxial compressive strength. K_V is rock mass intactness index.

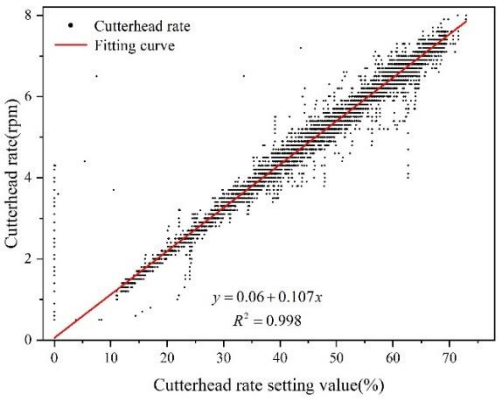
The rock mass classifications are derived from field investigations and laboratory tests conducted. Field investigations included detailed mapping of geological structures, identification of structural planes, and groundwater conditions. Rock samples were collected at regular intervals and subjected to laboratory tests to determine key HC indicators such as rock saturated uniaxial compressive strength (R_C) and rock mass intactness index (K_V). These indicators, combined with field observations of structural plane conditions, were used to classify the rock masses into Classes II to V following the HC classification guidelines. The corresponding rating criterion of rock mass classification of HC method are shown in Table 3.

2.3 Establishment of the dataset

Driving parameters of TBM can be divided into 4 categories. The first category is the operating parameters, which are used to describe the operating status of TBM, including PR, PRev, Th, T, cutterhead rate, tightening pressure of gripper shoe, P_b. The PR and Th are respectively converted from the propulsion speed and pressure of the propulsion cylinder. The torque and speed of the cutterhead are converted from the torque and speed of the output end of the main drive reducer. And the pressure of gripper shoe is converted from the pressure of the rodless cavity of gripper cylinder. The second category is the variation of key parameters, which is used to judge whether the working status of TBM is abnormal and whether the geological condition has changed, such as: the variation of



(a) Correlation of penetration rate and setting value



(b) Correlation of cutterhead rate and setting value

Fig. 4 Correlation analysis between TBM tunneling parameters and setting values

the penetration rate (ΔPR), the variation of the thrust (ΔTh), the variation of the cutterhead torque (ΔT). Variation refers to the difference between two adjacent driving parameters. The third category is the setting values of operating parameter, including the setting value of penetration rate and the setting value of cutterhead rate. The fourth category is the attitude and position parameters, including head mileage, horizontal offset of head, and vertical offset of head.

The driving parameters that can be used as the input of the rock mass classification model must reflect the law of rock machine action and contain the rock mass information of the current face. The driving parameters of the third and fourth categories are not the result of the joint interaction between the TBM and the rock mass, and such parameters cannot reflect the rock mass information of the current formation. Furthermore, the tightening pressure of gripper shoe is controlled by TBM driver according to the surrounding rock condition of current tunnel section, which is not affected by the rock condition of tunnel face. In the same way, the data of PR and cutterhead rate also include the judgment of the TBM driver. To determine whether the PR and cutterhead rate can reflect the class of rock mass in the face, the correlation analysis between the PR and its setting value, and the cutterhead rate and its setting value are carried out respectively, shown in Fig. 4. It can be seen that the cutterhead rate is proportional to its setting value, and its R^2 reached 0.998. It can be considered that the cutterhead rate is completely controlled by the TBM driver and does not contain information about the rock mass of the face. On the contrary, the correlation between the PR and its setting value is very low, which indicates that the PR is affected by the rock mass of the tunnel face and contains information about the rock mass of the tunnel face. So, the PR can be used as the input of the classification model.

Table 4 Correlation coefficient between different eigenvalues and rock mass classes

Eigenvalues	$\overline{PR}$	$\overline{Th}$	$\overline{T}$	$S_{\Delta PR}$	$S_{\Delta Th}$	$S_{\Delta T}$
Pearson's r	-0.61829	-0.87339	-0.88626	-0.68707	-0.58906	-0.72175

According to the above analysis, the eigenvalues of the driving parameters in the sample set are screened, and the correlation between the eigenvalues of the remaining driving parameters and the classes of surrounding rock is analyzed. Pearson's correlation coefficient is selected for evaluation, and its mathematical expression is

$$r = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{N\sum y_i^2 - \left(\sum y_i\right)^2}} \tag{3}$$

where (x_1, y_1), (x_2, y_2), … , (x_n, y_n) are the n-th observation values of x and y. The eigenvalues with correlation coefficient greater than 0.5 in the calculation results are shown in Table 4. Finally, 6 eigenvalues, namely, the mean value of penetration rate ($\overline{PR}$), the mean value of thrust ($\overline{Th}$), the mean value of cutterhead torque ($\overline{T}$), the standard deviation of penetration rate variation ($S_{\Delta PR}$), the standard deviation of thrust variation ($S_{\Delta Th}$) and the standard deviation of cutterhead torque variation ($S_{\Delta T}$), are selected as the model inputs.

The eigenvalues contained in the sample set are normalized and linearly changed to [-1, +1] interval. Among them, the training set contains 19200 samples (4800 samples for each class), and the test set contains 4800 samples (1200 samples for each class). The specific method is to collect samples from the sample set at a fixed distance to form a test set. In this way, the samples of the test set and the training set can be avoided from falling on completely different geological sections, which will lead to too much difference between them.

2.4 Support vector classifier

Support vector classifier (SVC) is a machine learning method for data classification, which is widely used due to its strong generalization ability and high robustness (Koo *et al.* 2019, Hasan *et al.* 2017, Li *et al*. 2016). The basic idea of SVC is to create a hyperplane for decision-making in order to maximize the isolated edges between positive and negative samples. The formula for finding the hyperplane is as follows

$$\begin{aligned} &\min_{\omega,\xi} \ \frac{1}{2}\left\|\omega^2\right\| + c\sum_{i=1}^{n}\xi_i \\ &s.t. \ \ \left\|\vec{\omega}\cdot\vec{x} + b\right\| \geq 1 - \xi_i, \xi \geq 0 \end{aligned} \tag{4}$$

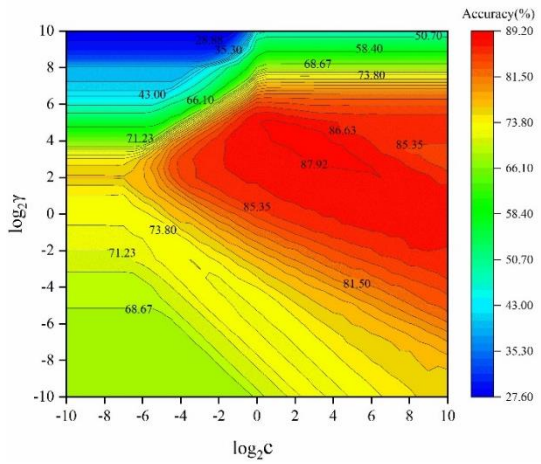


Fig. 5 10-fold cross-validation results of LIBSVM in different combinations of γ and c

where n denotes the number of training samples, c denotes the penalty coefficient, and ξ denotes the slack variable. By introducing a kernel function, SVC can also be used as a nonlinear classifier. The kernel function makes the original sample linearly separable by mapping the sample features to a higher-dimensional expression space. Among them, the Gaussian kernel function is one of the most widely used kernel functions, as shown in the following formula

$$k\left(x_i, x_j\right) = \exp\left(-\gamma \left\| x_i - x_j \right\|^2\right) \tag{5}$$

where γ denotes the parameter used to calibrate the degree of deviation and the variance of the model. In other words, γ defines the influence range of a single sample.

Because the SVC is mainly used to solve the binary classification problem, for the multi-classification problem, the SVC decomposes it into multiple binary classification problems, and realizes the multi-classification by optimizing the combination of multiple binary classification support vector classifiers. The LIBSVM used in this paper applies voting methods to deal with multi-classification problems. The basic idea is to construct $N(N\text{-}1)/2$ two-classifiers among N categories, and use the "voting method" to identify test samples (Chang and Lin 2011). After the test sample is input into the model, all the two-classifiers will classify it. If the type gets the most votes, the test sample will be considered as the type. Compared with other multi-classification SVM algorithms, LIBSVM has the advantages of fewer sub-classifiers and shorter training time.

2.5 Training and testing of the LIBSVM model

The HC classification system provides precise labels for the training and test datasets, ensuring that the LIBSVM model is trained with high-quality, accurately classified data. Each tunnel segment is categorized into one of four classes (II to V) based on HC criteria.

The LIBSVM model was trained and tested using the training set and test set described in Section 2.3. According to experience, the Gaussian kernel function is selected as the kernel function of LIBSVM (Shi *et al.* 2020). The selection range of kernel function parameter γ and penalty coefficient c is set to [2^{-10}, $2^{-9.5}$, 2^{-9}, … , 2^{9}, $2^{9.5}$, 2^{10}]. In order to determine the best model hyperparameters, γ and c were continuously adjusted, and a 10-fold cross-validation was conducted using the training set to calculate the average precision. The cross-validation results for the LIBSVM with different hyperparameters are shown in Fig. 5. When the kernel function parameter and penalty coefficient are $2^{2.5}$ and 2^{4}, respectively, the 10-fold cross-validation result has the highest recognition accuracy, at 89.20%.

Table 5 LIBSVM recognition results using test set

Actual class	Predicted results				Total
	Class II	Class III	Class IV	Class V	
Class II	1175	25	0	0	1200
Class III	4	1107	47	42	1200
Class IV	0	**157**	929	**114**	1200
Class V	0	2	52	1146	1200
Total	1179	1291	1028	1302	4800

In order to evaluate the training results of the rock mass classification model more comprehensively, the accuracy, precision, recall and F1-score are introduced as the evaluation indicators of the model. The calculation formulas for these four indicators are as follows

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1-score = \frac{2 \times precision \times recall}{precision + recall} \tag{9}$$

where TP denotes true positives, FP denotes false positives, FN denotes false negatives, and TN denotes true negatives.

3. Experiment results

3.1 Recognition results using test set

Tables 5 and 6 show the recognition results of the LIBSVM model on the test set. From the tables, it can be seen that the recognition effect of Class II rock mass is the best, and the four indicators are all above 0.97. And it shows that the model can effectively distinguish Class II rock mass from the other three classes of rock mass, and only a small number of data points of Class II rock mass are incorrectly identified as Class III. Although the recall of Class III rock mass and Class V rock mass are both above 0.9, the precision is both lower than 0.9. This is because the model misidentified a large number of Class IV rock mass

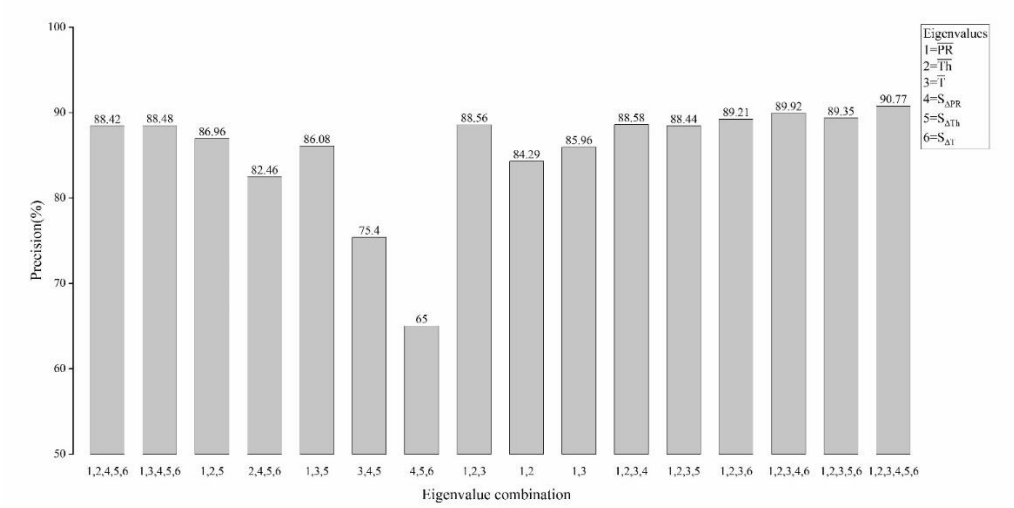


Fig. 6 Precision of different eigenvalue combinations using test set

Table 6 Accuracy, precision, recall and F1-score of LIBSVM model using test set

Actual class	Accuracy	Precision	Recall	F1-score	Support
Class II	**0.994**	**0.997**	**0.979**	**0.988**	1200
Class III	0.942	**0.857**	0.923	0.889	1200
Class IV	0.923	0.904	**0.774**	0.834	1200
Class V	0.956	**0.880**	0.955	0.916	1200
Avg./total	0.954	0.909	0.908	0.907	4800

samples as Class III and Class V. This also cause the recall of Class IV rock mass to be only 0.774. It can be found that no samples of Class IV and Class V rock mass have been wrongly identified as Class II, while a certain number of Class III, IV and V rock mass samples have been wrongly identified as the other type of them. So, the further problem of the classification model is the classification problem between Class II and III rock mass, and the classification problem between Class III, IV and V rock mass.

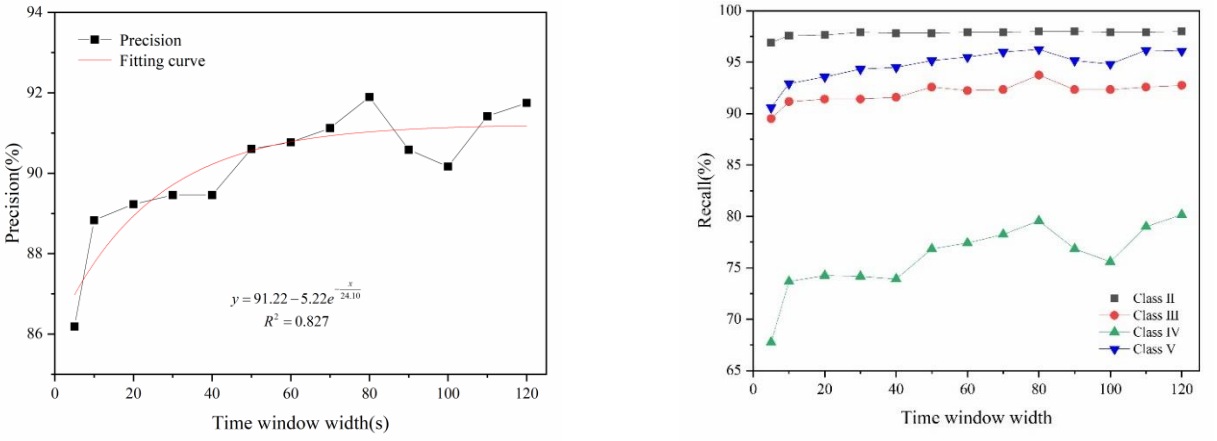
Considering that there may be redundancy in the sample set composed of all 6 eigenvalues, multiple combinations of these 6 eigenvalues are used for model training, and the classification precision obtained by using the test set is shown in Fig. 6. It can be seen that the influence of the 6 eigenvalues on the model classification results is overlapping. The recognition precision, using only the 2 eigenvalues of $\overline{PR}$ and $\overline{T}$, reached 85.96%. Only in terms of recognition precision, the effect of $\overline{T}$ and $\overline{Th}$ are similar. Compared with the eigenvalues $\overline{PR}$, $\overline{Th}$ and $\overline{T}$, the eigenvalues $S_{\Delta PR}$, $S_{\Delta Th}$ and $S_{\Delta T}$ have less effect on improving the precision of rock mass classification model.

3.2 Recognition results under different widths of the time window

The width of the time window for extracting the data points from the sample set affects the quality of the data points to a certain extent and should be analyzed. Considering that the cutterhead rate is 0.5 to 8 rpm, the width of time window is within the range of 10 to 120 s and the interval is 10 s. The time window width of 5 s is also considered. The experimental results are shown in Fig. 7. In general, the precision of the model increases as the time window width increases, and fluctuates around the fitted curve. When the window width reaches 80 s, the model precision reaches the highest, at 91.90%. Combining with the changing law of the recall of various rock mass, it can be seen that this kind of fluctuation mainly comes from Class IV rock mass. With the increase of time window width, the recall of Class II rock mass hardly changed, and the recall rate of Class III rock mass and Class IV rock mass steadily increased, reaching the highest at 80 s.

The recognition results of each sample under different time windows are shown in Fig. 8. The abscissa in the figure is the serial number of samples in the test set, in order from Class II to Class V, and the number of rock mass samples of each Class is 1,200. It can be seen from the figure that the recognition result of the Class II rock mass is the best, which only has a small number of samples incorrectly identified as Class III rock mass. No samples of Class IV and Class V rock mass are mistakenly identified as Class II. However, the recognition result of Class IV rock mass is the worst, which has a large number of samples mistakenly identified as Class III and Class V. Among the samples of Class III rock mass that are incorrectly identified, the most are incorrectly identified as Class V. It can be considered that the Class II rock mass samples only overlap with the Class III rock mass samples, while the Class III, IV and V rock mass samples overlap each other.

There is a total of 841 samples that have been misidentified, including 44 Class II rock mass samples, 177 Class III rock mass samples, 457 Class IV rock mass samples, and 163 Class V rock mass samples. It can be seen from the figure that no matter how long the time window is selected, some samples will always be misidentified. There are a total of 254 such samples, as shown in Table 7. Among them, the Class IV rock mass samples are the most, most of which have been misidentified as Class III. And the samples of Class II and Class V are the least. The samples



(a) Precision of LIBSVM (b) Recall of different class rock mass

Fig. 7 Recognition results of LIBSVM under different time window width

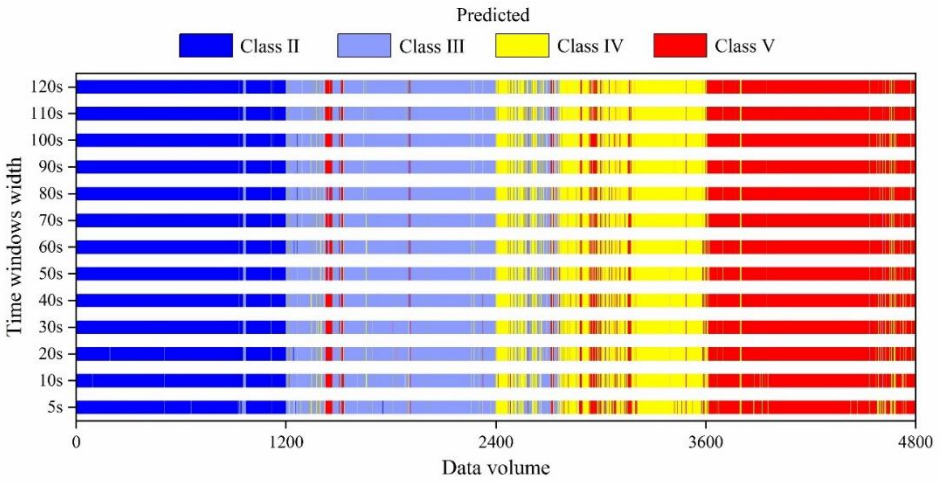


Fig. 8 Recognition results of each sample under different time windows using test set

Table 7 The predicted results of the sample which is always wrongly identified

Actual class	Definite predicted results			Uncertain results	Total
	Class III	Class IV	Class V		
Class II	19	0	0	0	**19**
Class III	0	5	34	7	46
Class IV	**113**	0	53	0	**166**
Class V	0	21	0	2	**23**
Total	132	26	87	9	**254**

Table 8 The ranking score of input features of LIBSVM in different classifiers

Eigenvalues	Classifiers					
	II vs III	II vs IV	II vs V	III vs IV	III vs V	IV vs V
$\overline{PR}$	33.54	**12.81**	6.04	133.42	**83.78**	207.61
$\overline{Th}$	**40.13**	3.27	0.07	**160.94**	57.51	**230.90**
$\overline{T}$	**42.10**	3.29	0.94	106.99	50.33	183.97
$S_{\Delta PR}$	21.45	**13.95**	**8.94**	18.78	6.90	31.66
$S_{\Delta Th}$	1.10	11.22	4.39	105.13	10.67	138.50
$S_{\Delta T}$	4.31	7.24	3.46	72.48	30.96	129.23
Total	142.64	51.79	**23.83**	597.74	240.14	**921.87**

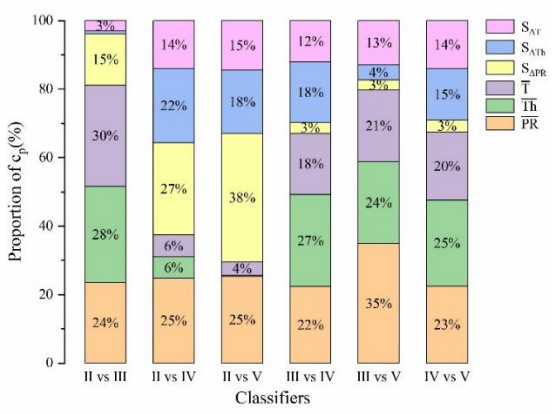


Fig. 9 Proportion of c_p in different classifiers

which have been wrongly identified as a certain class can be regarded as falling into that class. In other words, their samples overlap with other classes of samples and are inseparable. They account for 5.10% of the test set, of which Class III, IV and V rock mass samples account for 4.71% of the test set. So, it can be considered that the data overlap of Class III, IV and V is an important reason for further improving the precision of the model.

3.3 Feature importance analysis

In order to evaluate the results more comprehensively, it is necessary to analyze the role of 6 input features in the LIBSVM rock mass classification model. The ranking scores of features c_p are calculated to evaluate the importance of each feature in the two-classifiers. The c_p denotes the fluctuation of the error of the support vector fitting hyperplane caused by removing the p-th feature. The larger the c_p, the greater the impact of this feature on the classification effect of the classifier, and the more important this feature is. The results of ranking scores of input features in different classifiers of LIBSVM model are shown in Table 8. In general, the importance of each feature in these 6 two-classifiers is different. In the classifier for Class II and III rock mass, the $\overline{PR}$ and $\overline{Th}$ are more important. In the classifier for Class II and IV rock mass, the $\overline{PR}$ and $S_{\Delta PR}$ are more important. In the classifier for Class II and V rock mass, the $S_{\Delta PR}$ is more important. In the classifier for Class III and IV rock mass, the $\overline{Th}$ is more important. In the classifier for Class III and V rock mass, the $\overline{PR}$ is more important. In the classifier for Class IV and V rock mass, the $\overline{Th}$ is more important. It can be seen that the fluctuation of the hyperplane fitting error of the two-classifiers involving Class II rock mass is relatively small. And the fluctuation of the hyperplane fitting error of the two-classifier for Class IV and V are the largest, followed by the two-classifier for Class III and IV. This shows that the data points of Class II rock mass are easy to distinguish, while the data points of Class III, IV and V are mixed with each other, especially for the data points of Class IV that are not easy to distinguish from Class III and V.

Further, the proportion of c_p in the 6 classifiers is plotted, as shown in Fig. 9. It can be seen that the two-classifier for Class II and IV and the two-classifier for Class II and V are different from the other 4 classifiers. The sum of the c_p proportions of $S_{\Delta PR}$, $S_{\Delta Th}$ and $S_{\Delta T}$ exceeds 60%, which plays a decisive role. This shows that the model distinguishes Class II rock mass from Class IV and V rock mass on the standard deviation of ΔPR, ΔTh and ΔT, that is, the fluctuation of the change rate of the driving parameters. And the sum of the c_p proportions of the two input features ($\overline{PR}$ and $S_{\Delta PR}$) derived from the PR exceeds 50%, which means PR is the most important index for the model to distinguish Class II rock mass from Class IV and V rock mass. In the two-classifier for Class III and IV and the two-classifier for Class IV and V, the other five input features except for $S_{\Delta PR}$ play an indispensable role. This is because these data points are not easy to distinguish and need to be identified by as many features as possible.

4. Conclusions

In this paper, a method for rock mass classification of TBM tunnel face is proposed, based on the driving parameters generated in the tunneling process. The correlation analysis is performed to filter model input. And the mean value of the penetration rate, thrust and cutterhead torque and the standard deviation of their variation are found to be strongly correlated with the Class of rock mass. The LIBSVM model is trained with different combinations of eigenvalues. When all 6 eigenvalues are used as model input, the model recognition precision is 90.77%. The model has the best recognition effect on Class II rock mass, with all indicators above 97%. And the model has the worst recognition effect on Class IV rock mass, with a recall of 0.774. These samples are misidentified as Class III and V. The relationship between the width of time window and the model recognition precision and the recall of various rock mass is studied. The results show that the recognition precision of the model increases with the increase of the time window width. And the precision fluctuates in a certain range. When the window width is 80 s, the model precision is the highest. The recall of Class II rock mass is hardly affected by the time window width. The recall of Class III and IV rock mass increase slightly with the increase of the time window width. The change curve of recall of Class V with time window width is similar to the recognition precision of model, which indicates that the recognition error of model mainly comes from Class V rock mass. At last, the role of 6 features in the rock mass classification model is analyzed.

Although satisfactory classification precision has been achieved, some shortcoming of this study have to be acknowledged. First, the lithology of the tunnel section from which the driving parameters come is relatively single, mainly tuffaceous sandstone. The rock mass classification model trained from the data has certain limitations. Second, the precision of LIBSVM rock mass classification model can be improved by taking the eigenvalues of driving

parameters over a period of time as input. However, the influence of time window width on the model is failed to reveal, and the suggestion of time window width selection isn't put forward. Third, the importance of the 6 input features to each classifier is analyzed, but the specific influence ways of the input features on model are not pointed out. Fourth, the random sampling method used during model training addressed the data imbalance issue and enhanced the model's learning capability. Future research will explore more sophisticated data balancing techniques or domain adaptation methods to further align the model training process with practical field scenarios. In order to solve these shortcomings, the driving parameters will be obtained from more types of lithologic tunnel sections in the future research to improve the generality of the model. Considering the "black box" characteristic of LIBSVM model, the influence of rock mass class change on tunneling parameters will be analyzed in combination with some other methods to obtain more specific rules.

Acknowledgments

The research described in this paper was financially supp orted by the National Key Research and Development Program of China (Granted No. 2022YFC3802300) and the Key Project of Science and Technology of Hunan Province (Granted No. 2019GK1010).

References

Alber, M. (2000), "Advance rates of hard rock TBMs and their effects on project economics", *Tunn. Undergr. Sp. Tech.*, **15**(1), 55-64. https://doi.org/10.1016/S0886-7798(00)00029-8.

Barton, N. (1999), "TBM performance estimation in rock using QTBM", *Tunnel Tunn. Int.*, **31**(9), 30-34.

Chang, C.C. and Lin, C.J. (2011), "LIBSVM: a library for support vector machines", *ACM T. Intell. Syst. Tech.*, **2**(3), 1-27. https://doi.org/10.1145/1961189.1961199.

Fernandez-Gutierrez, J.D., Perez-Acebo, H. and Mulone-Andere, D. (2017), "Correlation between Bieniawski's RMR index and Barton's Q index in fine-grained sedimentary rock formations", *Informes de la Constr.*, **69**(547), e205. https://doi.org/10.3989/id54459.

Gholamnejad, J. and Tayarani, N. (2010), "Application of artificial neural networks to the prediction of tunnel boring machine penetration rate", *Min. Sci. Tech.*, **20**(5), 727-733. https://doi.org/10.1016/S1674-5264(09)60271-4

Hasan, M.A., Ahmad, S. and Molla, M.K.I. (2017), "Protein subcellular localization prediction using multiple kernel learning based support vector machine", *Molecular Biosyst.*, **13**(4), 785-795. https://doi.org/10.1039/c6mb00860g.

Hassanpour, J., Rostami, J. and Zhao, J. (2011), "A new hard rock TBM performance prediction model for project planning", *Tunn. Undergr. Sp. Tech.*, **26**(5), 595-603. https://doi.org/10.1016/j.tust.2011.04.004.

Hoek, E. and Brown, E.T. (1997), "Practical estimates of rock mass strength", *Int. J. Rock Mech. Min. Sci.*, **34**(8), 1165-1186. https://doi.org/10.1016/S0148-9062(97)00305-7.

Jain, P., Naithani, A.K. and Singh, T.N. (2016), "Estimation of the performance of the tunnel boring machine (TBM) using uniaxial compressive strength and rock mass rating classification (RMR)–A case study from the Deccan traps, India", *J. Geol. Soc. India*, **87**(2), 145-152. https://doi.org/10.1007/s12594-016-0382-0.

Kim, J.W., Chong, S.H. and Cho, G.C. (2022), "Probabilistic Q-system for rock classification considering shear wave propagation in jointed rock mass", *Geomech. Eng.*, **30**(5), 449-460. https://doi.org/10.12989/gae.2022.30.5.449.

Kim, K.Y., Jo, S.A., Ryu, H.H. and Cho, G.C. (2020), "Prediction of TBM performance based on specific energy", *Geomech. Eng.*, **22**(6), 489-496. https://doi.org/10.12989/gae.2020.22.6.489.

Kim, Y., Hong, J., Shin, J. and Kim, B. (2022), "Shield TBM disc cutter replacement and wear rate prediction using machine learning techniques", *Geomech. Eng.*, **29**(3), 249. https://doi.org/10.12989/gae.2022.29.3.249.

Koo, B., La, S., Cho, N.W. and Yu, Y. (2019), "Using support vector machines to classify building elements for checking the semantic integrity of building information models", *Automat. Constr.*, **98**, 183-194. https://doi.org/10.1016/j.autcon.2018.11.015.

Li, Y.B., Xu, M.Q., Wei, Y., Huang, W. and Huang, W.H. (2016), "A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree", *Measurement*, **77**, 80-94. https://doi.org/10.1016/j.measurement.2015.08.034.

Liu, B., Wang, R., Zhao, G., Guo, X., Wang, Y., Li, J. and Wang, S. (2020), "Prediction of rock mass parameters in the TBM tunnel based on BP neural network integrated simulated annealing algorithm", *Tunnelling and Underground Space Technology*, **95**, 103103. https://doi.org/10.1016/j.tust.2019.103103.

Liu, Q.S., Liu, J.P., Pan, Y.C., Kong, X.X. and Hong, K.R. (2017), "A case study of TBM performance prediction using a Chinese rock mass classification system–Hydropower Classification (HC) method", *Tunn. Undergr. Sp. Tech.*, **65**(5), 140-154. https://doi.org/10.1016/j.tust.2017.03.002.

Liu, Q.S., Wang, X.Y., Huang, X. and Yin, X. (2020), "Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data", *Tunn. Undergr. Sp. Tech.*, **106**, 103595. https://doi.org/10.1016/j.tust.2020.103595.

Liu, Q., Liu, J., Pan, Y., Kong, X. and Hong, K. (2017), "A case study of TBM performance prediction using a Chinese rock mass classification system–Hydropower Classification (HC) method", *Tunn. Undergr. Sp. Tech.*, **65**, 140-154. https://doi.org/10.1016/j.tust.2017.03.002.

Mahmoodzadeh, A., Nejati, H.R., Mohammadi, M., Ibrahim, H.H., Mohammed, A.H. and Rashidi, S. (2022), "Assessment of wall convergence for tunnels using machine learning techniques", *Geomech. Eng.*, **31**(3), 265-279. https://doi.org/10.12989/gae.2022.31.3.265.

Mahmoodzadeh, A., Nejati, H.R., Ibrahim, H.H., Ali, H.F.H., Mohammed, A.H., Rashidi, S. and Majeed, M.K. (2022), "Several models for tunnel boring machine performance prediction based on machine learning", *Geomech. Eng.*, **30**(1), 75-91. https://doi.org/10.12989/gae.2022.30.1.075.

Salimi, A., Faradonbeh, R.S., Monjezi, M. and Moormann, C. (2018), "TBM performance estimation using a classification and regression tree (CART) technique", *Bull. Eng. Geol. Environm.*, **77**(1), 429-440. https://doi.org/10.1007/s10064-016-0969-0.

Salimi, A., Rostami, J. and Moormann, C. (2019), "Application of rock mass classification systems for performance estimation of rock TBMs using regression tree and artificial intelligence algorithms", *Tunn. Undergr. Sp. Tech.*, **92**, 103046. https://doi.org/10.1016/j.tust.2019.103046.

Salimi, A., Rostami, J., Moormann, C. and Hassanpour, J. (2018), "Examining feasibility of developing a rock mass classification for hard rock TBM application using non-linear regression,

regression tree and generic programming", *Geotech. Geol. Eng.*, **36**(2), 1145-1159. https://doi.org/10.1007/s10706-017-0380-z.

Sapigni, M., Berti, M., Bethaz, E., Busillo, A. and Cardone, G. (2002), "TBM performance estimation using rock mass classifications", *Int. J. Rock Mech. Min. Sci.*, **39**, 771-788. https://doi.org/10.1016/S1365-1609(02)00069-2.

Shi, Y.P., Xia, Y.M., Zhang, Y.M. and Yao, Z.W. (2020), "Intelligent identification for working-cycle stages of excavator based on main pump pressure", *Automat. Constr.*, **109**, 102991. https://doi.org/10.1016/j.autcon.2019.102991.

The National Standards Compilation Group of People's Republic of China. (2009), *GB50487-2008 Code for Engineering Geological Investigation of Water Resources and Hydropower*, China Planning Press, Beijing, China.

Von Preinl, Z.T.B., Celada, T.B., Fernández, J.M.G. and Alvarez, H.M. (2006), "Rock mass excavability indicator: New way to selecting the optimum tunnel construction method", *Tunn. Undergr. Sp. Tech.*, **21**(3-4), 237. https://doi.org/10.1016/j.tust.2005.12.016.

Wang, C., Gong, G.F., Yang, H.Y., Zhou, J.J., Duan, L.W. and Zhang, Y.K. (2018), "NSVR based predictive analysis of cutterhead torque for hard rock TBM", *J. Zhejiang Univ. (Engineering Science)*, **52**(3), 479-486. https://doi.org/10.3785/j.issn.1008-973X.2018.03.009.

Wang, P., Xue, Y., Su, M., Qiu, D. and Li, G. (2022), "A TBM tunnel collapse risk prediction model based on AHP and normal cloud model", *Geomech. Eng.*, **30**(5), 413-422. https://doi.org/10.12989/gae.2022.30.5.413

Xue, Y.T., Zhang, L. and Wang, B.J. (2018), "Nonlinear feature selection using Gaussian kernel SVM-RFE for fault diagnosis", *Appl. Intelligence*, **48**(10), 3306-3331. https://doi.org/0.1007/s10489-018-1140-3.

Yang, H.Q., Wang, H. and Zhou, X.P. (2016), "Analysis on the Rock-Cutter Interaction Mechanism During the TBM Tunneling Process", *Rock Mech. Rock Eng.*, **49**(3), 1073-1090. https://doi.org/10.1007/s00603-015-0796-9.

Zhang, Q.L., Liu, Z.Y. and Tan, J.R. (2019), "Prediction of geological conditions for a tunnel boring machine using big operational data", *Automat. Constr.*, **100**, 73-83. https://doi.org/10.1016/j.autcon.2018.12.022.

Zhou, J. and Yang, X.A. (2021), "Deformation behavior analysis of tunnels opened in various rock mass grades conditions in China", *Geomech. Eng.*, **26**(2), 191-204. https://doi.org/10.12989/gae.2021.26.2.191

CC