DOI: https://doi.org/10.68732/90331bcebd061d78

ISBN: 978-9989-2053-7-8

Intelligent rock mass classification for different TBM projects

Chenfeng Gao, SongyuYue, Chengyao Wei & Yaoru Liu
Tsinghua University, Beijing, PR China
liuyaoru@tsinghua.edu.cn
0000-0001-6703-6604

Shaokang Hou
China Renewable Energy Engineering Institute, Beijing, PR China

Abstract

Accurate classification of surrounding rock grades during Tunnel Boring Machine (TBM) operations is essential to improve tunnelling efficiency and ensure construction safety. However, traditional classification methods face significant challenges, particularly in leveraging existing TBM engineering data to classify surrounding rock in new TBM projects—an important research focus in the field. This study proposes an intelligent cross-project TBM rock classification method based on data from two different TBM projects. Firstly, the raw TBM engineering data are preprocessed by eliminating invalid rock-breaking data, and effective tunnelling cycles are segmented to extract the stable section tunnelling data. Subsequently, considering the engineering design parameters (including tunnel diameter and number of cutters), the conversion relationships of key tunneling parameters are established based on TBM tunnelling mechanics. Adopting the data from Project A, an intelligent classification model is developed and trained with key tunnelling parameters as inputs and surrounding rock grades as outputs. Finally, the key tunnelling parameters from Project B are converted into model inputs, and the classification results are analysed and discussed. The results indicate that applying the conversion rules to the key tunnelling parameters of Project B improved the prediction accuracy of surrounding rock grades by 2%–5%.

Keywords

TBM, Rock mass classification, Stacking classifier, Tunnelling parameters conversion.

1 Introduction

Accurate identification of surrounding rock mass during Tunnel Boring Machine (TBM) operations is essential for ensuring construction safety, optimising tunnelling parameters, and improving excavation efficiency (Li et al., 2023a,b). Conventional rock mass classification methods rely heavily on geological investigations and manual interpretation, which are often discontinuous and lag behind the excavation face (Benardos et al., 2004). With the widespread application of TBMs, large volumes of operational data are continuously recorded during excavation. These data provide a valuable basis for developing data-driven approaches to real-time rock mass classification (Grima et al., 2000). In recent years, various machine learning and ensemble learning methods have been applied to TBM data for

rock mass identification and performance prediction (Armaghani et al., 2019). While promising results have been reported, most existing studies are confined to a single TBM project (Hassanpour et al., 2021). Differences in TBM design parameters, such as tunnel diameter and cutterhead configuration, significantly limit the direct transfer of trained models to new projects (Grima et al., 2000). As a result, the generalisation and applicability of data-driven rock mass classification models across different TBM projects remain challenging.

To address this limitation, this study proposes an intelligent cross-project rock mass classification method that integrates TBM tunnelling mechanics with ensemble learning. Conversion relationships for key tunnelling parameters are established to account for project-specific design differences. A stacking ensemble model is trained using data from one TBM project and then applied to another project after parameter conversion. The proposed method enhances model transferability and provides a practical solution for rock mass classification in TBM projects with limited labelled data.

2 Methodology

2.1 Overall approach

Fig. 1 illustrates the overall technical framework for intelligent cross-project rock mass classification based on TBM data. Firstly, the raw TBM excavation data are preprocessed. Key steps include the removal of invalid records, identification of complete excavation cycles, and assignment of rock mass labels to each cycle. This process produces two datasets, corresponding to Project A and Project B. Subsequently, the dataset from Project A is used for model training and optimisation. The dataset is divided into training and validation subsets based on a predefined ratio. The validation set is used for hyperparameter optimisation and construction of a high-performance stacking classification model. During generation of the test set from Project B, key parameters are converted to account for project-specific characteristics. Finally, the processed test set is input into the trained stacking model to generate classification results. This enables cross-project prediction from Project A to Project B.

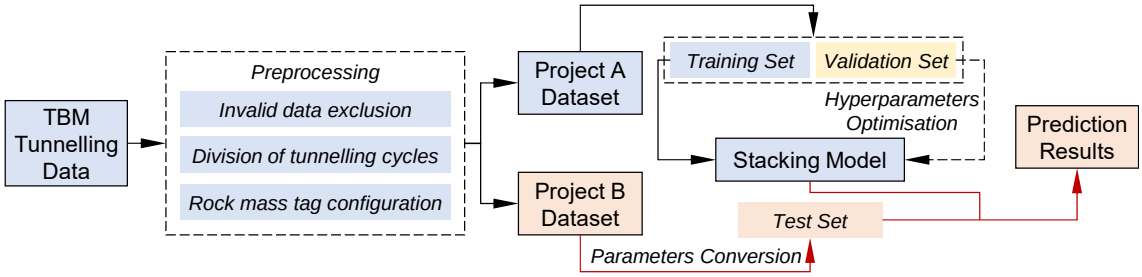


Fig. 1 Overall approach flowchart.

2.2 Data preprocessing

The TBM does not operate continuously throughout the day (Xiao et al., 2022). Consequently, the collected data contain a large number of downtime segments, as shown in Fig. 2(a). Such data are considered redundant for machine learning applications. Data preprocessing is carried out by constructing a state discrimination function, as defined in Eq. 1.

$$S = f(n) \cdot f(v) \cdot f(F) \cdot f(T), \quad f(x) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0 \end{cases} \tag{1}$$

Where S Output of the state function, where 0 indicates ineffective rock breaking data.
n Rotational speed of the TBM cutterhead, measured in revolutions per minute (r/min).
v Advance velocity of the TBM, measured in millimetres per minute (mm/min).
F Total torque generated by the cutterhead, measured in kilonewton metres (kN·m).
T Forward thrust force exerted by the TBM cylinder, measured in kilonewtons (kN).

After invalid data are removed using the binary state discrimination function, one day of TBM excavation data can be preliminarily divided into multiple tunnelling cycles.At the beginning of a tunnelling cycle, the TBM cutterhead has not yet contacted the surrounding rock. As excavation progresses, the TBM cutterhead begins to cut and break the rock mass. Consequently, all operating parameters increase rapidly. During the stable phase, the TBM advances continuously with relatively stable operating parameters. In this study, the cumulative sum method is used to delineate the stable phase, as shown in Fig. 2(b).

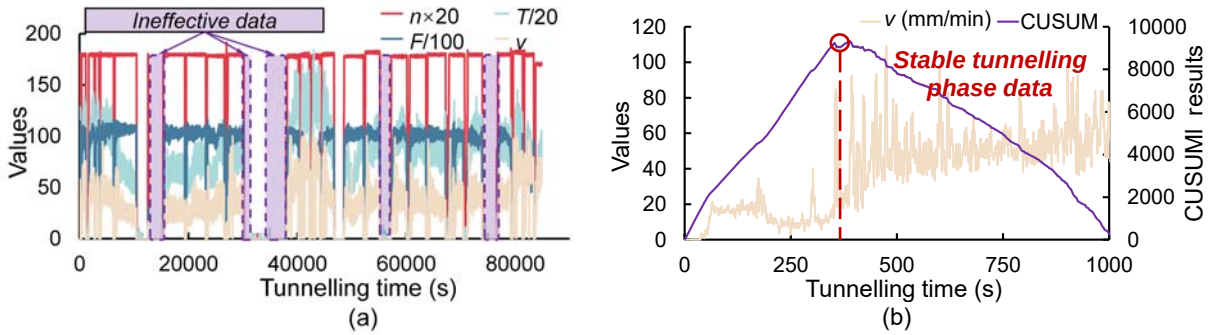


Fig. 2 Schematic diagram of TBM data preprocessing. (a) Daily tunnelling data processing: invalidation removal and tunnelling cycle division. (b) Stable tunnelling phase division within a TBM tunnelling cycle.

The CUSUM calculation is based on the advance velocity v(Hou et al., 2020). The point with the maximum CUSUM value is identified as the inflection point. This point separates the parameter ascent phase from the stable phase. The corresponding time marks the start of the stable phase. The CUSUM formulation is given in Eq. 2:

$$CUSUM_i = \sum_{j=1}^{i} (\bar{x} - x_j), \quad (i = 1,2,3,\dots,N) \tag{2}$$

Where N Number of data points in the TBM tunnelling cycle.
$\bar{x}$ Mean value of all data points in the TBM tunnelling cycle.

After the stable phases are delineated, outliers are removed using the 3σ criterion based on the normal distribution. Subsequently, the mean value of each tunnelling parameter within the stable sections is calculated for model training and prediction. Different parameters exhibit large differences in magnitude. Therefore, the normal distribution normalisation is applied to normalise each parameter.

2.3 Stacking: an ensemble learning classifier

Ensemble learning classifiers differ fundamentally from single classifiers. By integrating multiple homogeneous or heterogeneous weak classifiers, ensemble learning improves robustness and generalisation performance. The Stacking algorithm integrates heterogeneous weak learners and trains its base learners in parallel (Hou et al., 2022). A meta-learner then learns from the outputs of each base learner to generate the final prediction. The Stacking ensemble learning framework consists of two tiers, namely base classifiers as Tier I and a meta-classifier as Tier II. Firstly, multiple base classifiers are trained using the original dataset. This constitutes the first-stage learning. To reduce the risk of overfitting, k-fold cross-validation is used to train each base classifier. Subsequently, the outputs of the base classifiers are combined to form a new dataset. The meta-classifier is then trained on this dataset during the second-stage learning. This yields the final classification result.

3 Cross-project TBM tunnelling parameters conversion

3.1 Key tunnelling parameters

Data collected from different TBM projects often vary owing to differences in equipment manufacturers and model versions. Data recorded during TBM operation are primarily categorised into rock-cutting parameters and equipment status parameters. Rock-cutting parameters are less influenced by equipment model. Therefore, the selected universal parameters are all rock-cutting parameters. During TBM rock cutting, five parameters are closely related to rock penetration. These parameters are n, v, T, F, and the penetration rate p. The penetration rate p denotes the advance per revolution of the cutterhead. It is measured in mm per revolution and is calculated according to Eq. 3.

$$p = v/n \tag{3}$$

In addition to the parameters directly recorded by the TBM system, two composite indicators of excavation performance are required. These indicators are the Field Penetration Index and the Torque Penetration Index, abbreviated as *FPI* and *TPI* respectively. These composite indicators reduce the influence of operator variability to some extent. They primarily reflect the excavability of the surrounding rock at the current advance position and the rock-cutting efficiency of the TBM. The calculation formulas for *FPI* and *TPI* are given in Eq. 4:

$$FPI = F/pTPI = T/p \tag{4}$$

3.2 Conversion relationships for key tunnelling parameters

By analysing the forces acting on individual cutters of the TBM cutterhead, several assumptions are introduced for the rock-breaking process. As the number of centre cutters is relatively small, all cutters on the TBM cutterhead are assumed to be identical. The forces acting on edge cutters, transition rollers, and positive rollers are assumed to be identical to those acting on the centre twin rollers. Accordingly, the tangential forcef_r^k and the normal force$\overline{f_n}$ are considered the same. The cutting tools are arranged in concentric circles. A proportional coefficient a is introduced to describe this arrangement. Based on these assumptions, calculation formulas for *T* and *F* are shown in Eq.5.

$$T = \sum_{i=1}^{N} d_i f_r^k = aDNf_r^k F = \sum_{i=1}^{N} f_r^k \approx N\overline{f_n} \tag{5}$$

Where N Number of cutting tools on the cutterhead.
D Cutterhead diameter.
d_i Distance from the *i*-th cutting tool to the centre of the cutterhead.

For tunnels excavated within the same rock mass class, for example Class III, the ranges of tangential and normal forces acting on an individual disc cutter during rock fragmentation are comparable. Therefore, for two projects denoted as A and B, a conversion relationship between torque and thrust can be established, as expressed in Eq. 6.

$$\frac{T_A}{T_B} = \frac{a_A D_A N_A}{a_B D_B N_B}\frac{F_A}{F_B} = \frac{N_A}{N_B} \tag{6}$$

The maximum designed linear velocity of the cutterhead is defined by Eq. 7. When projects A and B have tunnels with comparable diameters, this maximum linear velocity can be regarded as identical. On this basis, the conversion relationship for the cutterhead rotational speed is derived in Eq. 8.

$$V = r \cdot \omega = \pi D n/60 \tag{7}$$

$$\frac{n_A}{n_B} = \frac{D_A}{D_B} \tag{8}$$

The advance velocity*v* is an active operational parameter controlled by the TBM operator. Under favourable geological conditions with medium-strength rock, the maximum *v* is commonly limited by the muck removal capacity. In hard rock conditions, where the thrust reaches its rated value, the *v* is constrained by the maximum allowable cylinder pressure. Under rated thrust conditions, achieving a large penetration per revolution is difficult, even at high rotational speeds. As a result, the *v* remains limited. Therefore, no conversion relationship is required for the advance velocity *v*. Based on the conversion relationships described above, the corresponding relationships for the *FPI* and the *TPI* are obtained, as presented in Eq. 9 and Eq. 10.

$$\frac{FPI_A}{FPI_B} = \frac{F_A/p_A}{F_B/p_B} = \frac{F_A n_A/v_A}{F_B n_B/v_B} = \frac{F_A n_A}{F_B n_B} = \frac{D_B N_A}{D_A N_B} \tag{9}$$

$$\frac{TPI_A}{TPI_B} = \frac{T_A/p_A}{T_B/p_B} = \frac{T_A n_A/v_A}{T_B n_B/v_B} = \frac{T_A n_A}{T_B n_B} = \frac{a_A N_A}{a_B N_B} \tag{10}$$

4 Practical application in engineering

4.1 Introduction to engineering Project A and B

Project A and Project B investigated in this study show a high degree of comparability in key engineering design parameters. As summarised in Table 1, both projects adopt open-type tunnel boring machines, with bore diameters of 5200 mm and 4750 mm, respectively. In addition, critical parameters including cutter number *N*, maximum thrust *F*, maximum torque *T* and maximum advance velocity *v* within comparable ranges. This strong consistency in design parameters provides basis for cross-project analogy of construction experience and for the conversion of technical parameters.

Table 1. Key design parameters for Project A and B.

Project	TBM type	*D* [mm]	*N*	F_{max} [kN]	T_{max} [kN•m]	v_{max} [mm/min]	n_{max} [r/min]	Tunnelling distance [mm]	Cutter pitch [mm]
A	Open-type	5200	34	11340	3340	120	11.45	1800	70
B	Open-type	4750	28	13608	2870	120	11.30	1800	80

With respect to geological conditions, as summarised in Table 2, the tunnel sections of both projects are predominantly excavated in Class III rock mass. Comparable proportions of Class IV and Class V rock mass are also observed. Consequently, the reliability and applicability of the comparative analysis between the two projects are further strengthened.

Table 2. Statistics on rock mass categories for the tunnelling cycles of Project A and B.

Project	III	IV	V	All
A	158	43	19	220
B	1471	434	90	1995

4.2 Establishment of the rock mass classification model

In stacking ensemble learning classifiers, the selection of appropriate base classifiers and a suitable meta-classifier is critical for achieving reliable predictive performance. In this study, four widely used and representative classifiers were selected, namely Random Forest, Gradient Boosting Decision Tree, Extreme Gradient Boosting, and Light Gradient Boosting Machine. These four models were employed as base classifiers, while Random Forest was adopted as the meta-classifier of the stacking framework. Data preprocessing was performed for TBM Project A and B. Six feature parameters were extracted from the stable operating phases, including n, v, T, F,TPI, and FPI. For each parameter, the mean value and variance were calculated, resulting in a total of twelve input variables for the Stacking model. The rock mass category was defined as the model output. Owing to the limited size of the dataset, the training and validation sets were divided at a ratio of eight to two. Model hyperparameters were subsequently optimised using five-fold cross-validation.

4.3 Evaluation of modelprediction results

Model performance was evaluated using the standard classification metrics of accuracy and F_1 score (Gao et al., 2019). *ACC* represents the proportion of correctly classified samples relative to the total number of samples. The F_1 is defined as the harmonic mean of precision *PRE* and recall *REC*, providing a comprehensive evaluation of the balance between classification accuracy and completeness. The calculation formulas for *ACC* and F_1 are presented in Eq. 11 and Eq. 12.

$$F_1 = \frac{2 \times PRE \times REC}{PRE + REC} \quad ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

$$PRE = \frac{TP}{TP + FN} \quad REC = \frac{TP}{TP + FP} \tag{12}$$

Where TP True positives, samples that are actually positive and correctly predicted as positive.
TN True negatives, samples that are actually negative and correctly predicted as negative.
FP False positives, samples that are actually negative but predicted as positive.
FN False negatives, samples that are actually positive but predicted as negative.

As illustrated in Fig. 3, after applying the conversion of design parameters for Project B, the predictive performance of all classification models shows a systematic improvement.

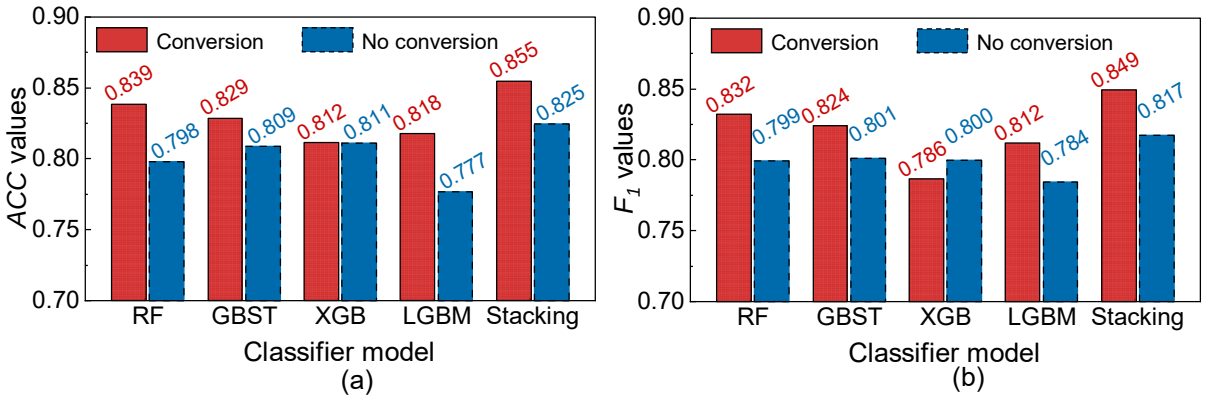


Fig. 3 Evaluation results for each classifier model. (a) *ACC*for each classifier model. (b) $F_1$for each classifier model.

In terms of *ACC*, the stacking ensemble model exhibits the most pronounced enhancement, with the accuracy value increasing from 0.825 to 0.855, as shown in Fig. 3(a). This corresponds to a relative improvement of approximately 3.6%. With respect to composite performance metrics, the RF model is taken as an example. Its F_1 score increases from 0.799 to 0.832, representing a relative improvement of approximately 4.1%, as shown in Fig. 3(b). Overall, the parameter conversion leads to an average

performance improvement of approximately 2% to 5% across all models. These results demonstrate that the key parameter conversion method is both effective and broadly applicable, and that it provides a more reliable performance benchmark for rock mass prediction in Project B.

5 Conclusion

This study proposes an intelligent cross-project rock mass classification method for Tunnel Boring Machine operations by integrating tunnelling mechanics with ensemble learning. Based on TBM data from two engineering projects with comparable design parameters and geological conditions, key tunnelling parameters were first preprocessed and extracted from stable excavation phases. Conversion relationships for torque T, thrust F, rotational speed n, and composite indices FPI and TPI were then established to account for project-specific design differences. Using data from Project A, a stacking ensemble classification model was developed to predict rock mass categories. After applying the proposed parameter conversion method, the converted tunnelling data from Project B were successfully used as model inputs. The results show that parameter conversion leads to a systematic improvement in prediction performance across all classification models, with ACC and F_1 score increasing by approximately 2% to 5%. These findings demonstrate that the proposed parameter conversion approach effectively enhances the transferability of TBM data-driven models between different projects. The method provides a reliable and practical solution for intelligent rock mass classification in new TBM projects where labelled data are limited, and it offers valuable support for construction decision-making and risk management in mechanised tunnelling.

Acknowledgements

This study was supported by the State Key Laboratory of Hydroscience and Engineering (Grant No. sklhse-TD-2025-D01) and the National Natural Science Foundation of China (Grant No. 52179105).

References

Armaghani, D.J., Mohamad, E.T., Narayanasamy, M.S., Narita, N., & Yagiz, S. (2017). Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition. Tunnelling and Underground Space Technology, 63, 29-43.

Benardos, A.G., & Kaliampakos, D.C. (2004). Modelling TBM performance with artificial neural networks. Tunnelling and Underground Space Technology, 19(6), 597-605.

Gao, X., Shi, M., Song, X., Zhang, C., & Zhang, H. (2019). Recurrent neural networks for real-time prediction of TBM operating parameters. Automation in Construction, 98, 225-235.

Grima, M.A., Bruines, P.A., & Verhoef, P.N.W. (2000). Modeling tunnel boring machine performance by neuro-fuzzy methods. Tunnelling and Underground Space Technology, 15(3), 259-269.

Hassanpour, J., Rostami, J., & Zhao, J. (2011). A new hard rock TBM performance prediction model for project planning. Tunnelling and Underground Space Technology, 26(5), 595-603.

Hou, S.K., Liu, Y.R., & Zhang, K. (2020). Prediction of TBM tunnelling parameters based on IPSO-BP hybrid model. Chinese Journal of Rock Mechanics and Engineering, 39(8), 1648–1657 (in Chinese).

Hou, S.K., Liu, Y.R., & Yang, Q. (2022). 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, 14(1), 123-143. .

Li, J.B., Chen, Z.Y., Li, X., Jing, L.J., Zhangf, Y.P., Xiao, H.H., Wang, S.J., Yang, W.K., Wu, L.J.,& Li, P.Y. (2023). Feedback on a shared big dataset for intelligent TBM Part I: Feature extraction and machine learning methods. Underground Space, 11, 1-25.

Li, J.B., Chen, Z.Y., Li, X., Jing, L.J., Zhang, Y.P., Xiao, H.H., Wang, S.J., Yang, W.K., Wu, L.J.,& Li, P.Y. (2023). Feedback on a shared big dataset for intelligent TBM Part II: Application and forward look. Underground Space, 11, 26-45.

Xiao, H.H., Yang, W.K., Hu, J., Zhang, Y.P., Jing, L.J., & Chen, Z.Y. (2022). Significance and methodology: Preprocessing the big data for machine learning on TBM performance. Underground Space, 7(4), 680-701.