



OPEN

# A jamming risk warning model for TBM tunnelling based on Bayesian statistical methods

Shuang-jing Wang<sup>1,2</sup>, Le-chen Wang<sup>1</sup>, Lei-jie Wu<sup>1</sup>✉ & Xu Li<sup>1</sup>

This study presents a comprehensive jamming risk assessment framework for Tunnel Boring Machine (TBM) jamming accidents during excavation. Using real-time boring data and Bayesian conditional probability, a novel risk warning model is proposed to enhance safety and efficiency of tunneling projects. Through statistical analysis of excavation parameters, distinct patterns between jamming and normal excavation states are identified. A comprehensive jamming perception index ( $\eta$ ) is introduced that synthesizes multiple parameters to accurately identify jamming states with a recognition rate of 95%. This integrated approach overcomes the limitations of single-parameter analysis and provides improved accuracy in jamming risk assessment. Additionally, a quantitative model for calculating jamming probability is developed, accounting for differences in sample size between jamming and normal excavation sections. The refined model yields realistic estimates of jamming probability, with an average of 94% in jamming sections and 7% in normal excavation sections. Furthermore, geological analysis shows that the Class III surrounding rock is the most suitable for excavation and has the lowest jamming probability. This finding emphasizes the importance of considering geological conditions in excavation planning to effectively mitigate jamming risks. In conclusion, this research provides a practical framework for the prediction and management of TBM jamming accidents, contributing to enhanced safety and efficiency in tunneling projects.

**Keywords** Tunnel boring machine, Risk assessment, TBM jamming, Bayesian statistics, Boring data

Tunnel Boring Machine (TBM) exhibits excellent adaptability in the construction of hard rock tunnels, boasting advantages such as rapid excavation speed and a high degree of automation. Consequently, it has been widely employed in highway tunnels, railway tunnels, and diversion tunnels<sup>1–5</sup>. However, due to limited preliminary geological survey information, tunnel construction is fraught with potential accidents, including collapses<sup>6–8</sup>, water and mud inrushes<sup>9,10</sup>, rockbursts<sup>11,12</sup>, and large deformation of the surrounding rock<sup>13,14</sup>. Especially when the tunnel is under complex geological conditions such as large burial depth, fractured rock, and high geostress, the surrounding rock develops significant squeezing deformation. This leads to increased frictional resistance between the surrounding rock and the shield. When the thrust force of the TBM becomes insufficient to overcome this resistance, machine jamming occurs. In addition, the strength and geological characteristics of different rock strata significantly influence the risk of TBM jamming, softer and fractured formations are particularly prone to triggering such incidents. Consequently, variations in TBM operational parameters, including cutterhead torque, thrust force, and penetration rate, serve as critical indicators by reflecting the dynamic interaction between the machine and surrounding rock mass<sup>15–17</sup>.

Many scholars have conducted research on the causes of TBM jamming through approaches such as in-situ monitoring and numerical simulation. Liu et al.<sup>18,19</sup> found that with the release of geostress, the deformation of surrounding rock becomes more severe, eventually pressing against the TBM shield and increasing frictional resistance. When the machine's thrust is unable to counteract this resistance, jamming occurs. They developed a model to perceive the jamming state based on theoretical formulas and finite element simulations. Zhang et al.<sup>20</sup> analyzed the time-varying interactions between the TBM and surrounding rock, including contact, compression, and friction, using a theoretical model. The results showed that the main reasons for jamming are the increase in shield pressure and the expansion of the compression contact area. Huang et al.<sup>21</sup> achieved real-time monitoring of shield pressure by installing strain gauges, sensors, and pressure cells on the surface of the TBM shield, and combined this with computer programs to realize real-time warning of jamming. Xu et al.<sup>22</sup> and Lin et al.<sup>23</sup> summarized and analyzed 121 cases of TBM jamming and identified fault fracture zones as one

<sup>1</sup>Key Laboratory of Urban Underground Engineering of Ministry of Education, Beijing Jiaotong University, Beijing 100044, China. <sup>2</sup>China Railway Engineering Equipment Co., Ltd, Zhengzhou, China. ✉email: Leijee\_Wu@163.com

of the most common unfavorable geological conditions when jamming happens. They then proposed a method to identify fault fracture zones by integrating the microstructure, geochemistry, and mineralogy of rocks.

In the realm of TBM jamming warning research, Hasanpour et al.<sup>24</sup> employed FLAC3D to simulate to simulate TBM excavation processes, accurately estimated tunnel squeezing and shield pressure, and thereby determined the risks associated with TBM jamming. Liu et al.<sup>25</sup> conducted numerical simulations of jamming accidents, proposing comprehensive TBM jamming criteria that take into account rock strength, geostress, and surrounding rock deformation, effectively enabling the assessment of jamming risks. Yu et al.<sup>26</sup> developed a dynamic TBM performance evaluation model that incorporates both translational and angular velocity characteristics, utilizing the fuzzy analytic hierarchy process (FAHP) to assign weights within the model. Furthermore, Hasanpour et al.<sup>27</sup> leveraged artificial neural networks and Bayesian networks to analyze jamming risks, achieving notable success in predicting TBM jamming induced by soft rock squeezing.

However, these prior works have certain limitations. Some studies only qualitatively predict the risk of TBM jamming, lacking the quantitative probability of risk occurrence. Other studies rely heavily on single-parameter models or have limited interpretability. For example, certain machine learning methods, while effective in prediction, are often criticized as “black boxes” due to their complex structures and numerous parameters, making it difficult to explain the decision-making process and the contribution of each factor to the results. This limits the practical application of these models in actual TBM excavation scenarios where engineers need clear and transparent risk assessment information.

To address these research gaps, this paper proposes a Bayesian statistics framework for TBM jamming risk assessment. Unlike previous models, this framework offers high interpretability, providing a clear explanation of the underlying decision-making process and the contribution of each factor to the risk results. It can not only qualitatively determine whether there is a risk but also quantitatively provide the probability of risk occurrence. This enables engineers to make more informed decisions and take appropriate actions based on the risk probability. In addition, this framework fully utilizes the rich information from multi-source monitoring data, effectively capturing the complex relationships between various factors during the excavation process. It is expected to provide strong support for safe TBM excavation, improving excavation efficiency and reliability, and has significant theoretical and practical value for TBM excavation in complex geological conditions.

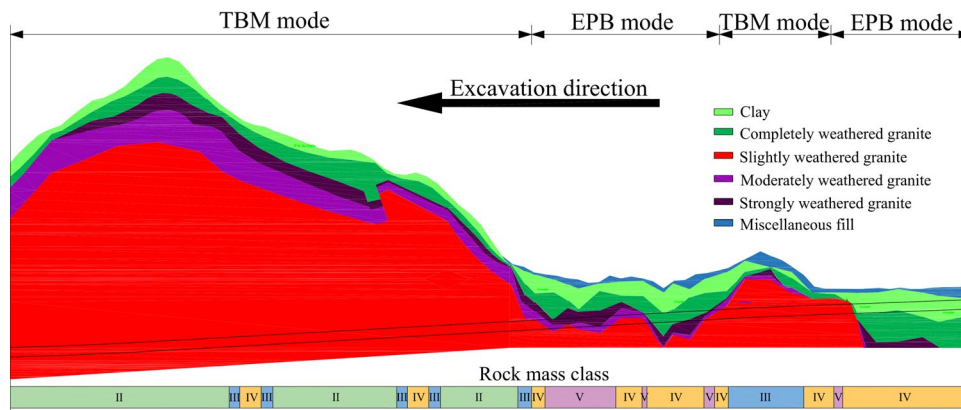
## Project brief Geological and TBM machine

The subway project is located in Shenzhen, Guangdong Province, China. The total length of the tunnel is 4546.72 m, and 1781.85 m has been excavated, with a total of 992 rings in the excavated section. The tunnel is buried at a depth of 15–525 m. The crossing strata are mainly composed of gravelly sand, clay, and weathered granite, with slightly weathered granite accounting for 65% of the total. The longitudinal profile of the excavated section is shown in Fig. 1. The strength of the strata varies significantly, with the uniaxial compressive strength (UCS) of weathered granite in the tunnel area ranging from 27 to 67 MPa, while the maximum strength of slightly weathered granite reaches approximately 106 MPa.

To overcome the limitations of low excavation efficiency in composite strata when using a single-mode shield machine, an EPB/TBM dual-mode shield machine is employed. The EPB (Earth Pressure Balance) mode is utilized for excavation in weak strata to ensure tunnel face stability and minimize ground settlement. Conversely, the TBM (Tunnel Boring Machine) mode is adopted for excavation in hard rock strata, avoiding issues such as slow excavation speed and severe secondary wear of the disc cutter that may occur in EPB mode. The main specifications of the shield machine are detailed in Table 1.

## Dataset

Prior to March 11, 2023, the TBM recorded boring data at a frequency of 1 time per second. Starting from March 12, 2023, the recording frequency was adjusted to 1 time per 10 s. Over the course of 478 days, a total of raw



**Fig. 1.** Geological profile of the excavated area. *Note:* The rock is classified into Classes II, III, IV, and V according to the Chinese rock classification system, Similar to the RMR system<sup>28,29</sup>, the rock mass classification standard is based on rock strength, joints, water conditions, and field stress conditions<sup>30</sup>.

	Diameter (mm)	Number of disc cutters	Rated thrust force (kN)	Rated torque (kN·m)	Maximum instantaneous torque (kN·m)	Cutterhead speed (r·min <sup>-1</sup> )	Maximum penetration rate (mm·min <sup>-1</sup> )
EPB /TBM dual-mode shield	9130	57	81,853	24,212	29,054	0~1.3~3.4	80

**Table 1.** Main specification of EPB /TBM dual-mode shield machine.

data was generated, with each sample comprising 843 parameters. The raw data was organized by ring number, excluding data collected during non-boring states. Subsequently, abnormal data was cleaned using the Pauta criterion<sup>31</sup>, ensuring the integrity and reliability of the dataset for further analysis.

To more accurately analyze the interrelationship between geological conditions and tunneling parameter, seven parameters were preliminarily selected as the analysis objects based on existing research<sup>32–35,36</sup>. These parameters include: (1) Total torque of the cutterhead  $T$ , (2) Total thrust of the cutterhead  $F$ , (3) Penetration of the cutterhead  $p$ , (4) Cutterhead rotation speed  $n$ , (5) Penetration rate  $v$ , (6) Field penetration index  $FPI$ , (7) Torque penetration index  $TPI$ .

Due to the significant difference in data between TBM mode and EPB mode, as well as the occurrence of jamming events in TBM mode, subsequent analysis focuses exclusively on data from TBM mode. Based on the actual excavation situation, the dataset is divided into two categories: jamming samples and normal excavation samples. The jamming section dataset contains 743 sets of samples, while the normal excavation section dataset includes 79,531 sets of samples.

Taking some typical parameters as examples, the statistical characteristics of the two types of samples are shown in Fig. 2. The figure reveals that the statistical indicators (mean, median, and range of values) for jamming and normal excavation samples exhibit significant different. Table 2 further details the geological conditions and excavation modes corresponding to different categories of samples, providing a clear basis for comparative analysis.

### Bayesian statistical method for jamming perception

To address the limitations of existing methods, we propose a comprehensive Bayesian statistical framework for TBM jamming risk assessment. This framework integrates multiple parameters to provide a more accurate and interpretable risk assessment. Specifically, we introduce a comprehensive jamming perception index  $\eta$  that synthesizes multiple parameters to accurately identify jamming states. Additionally, we develop a quantitative model for calculating jamming probability, accounting for differences in sample size between jamming and normal excavation sections. These innovations enable more accurate and reliable risk assessment.

### The distribution dissimilarity between different categories of samples

To quantitatively analyze the difference in feature distribution between machine jamming samples (jam) and normal excavation samples (non-jam), we denote the values of feature variables  $X_i$  in non-jam and jam samples as  $x_i^{\text{non-jam}}$  and  $x_i^{\text{jam}}$ , respectively. Figure 3 illustrates the distribution of these feature variables  $X_i$  in both categories of samples, where  $x_i^c$  serving as the threshold for distinguishing between the two classes.

In the samples located to the left side of  $x_i^c$  in the figure, not only are the correctly identified jam category samples included, but also the non-jam samples that were mistakenly classified as jam. Within the non-jam category, the proportion of samples misclassified as jam is referred to as the false positive rate ( $FPR$ ), which can be estimated by the cumulative probability density of  $x_i^{\text{non-jam}} < x_i^c$ , as shown in the grey shaded area in Fig. 3, corresponding to Eq. 1.

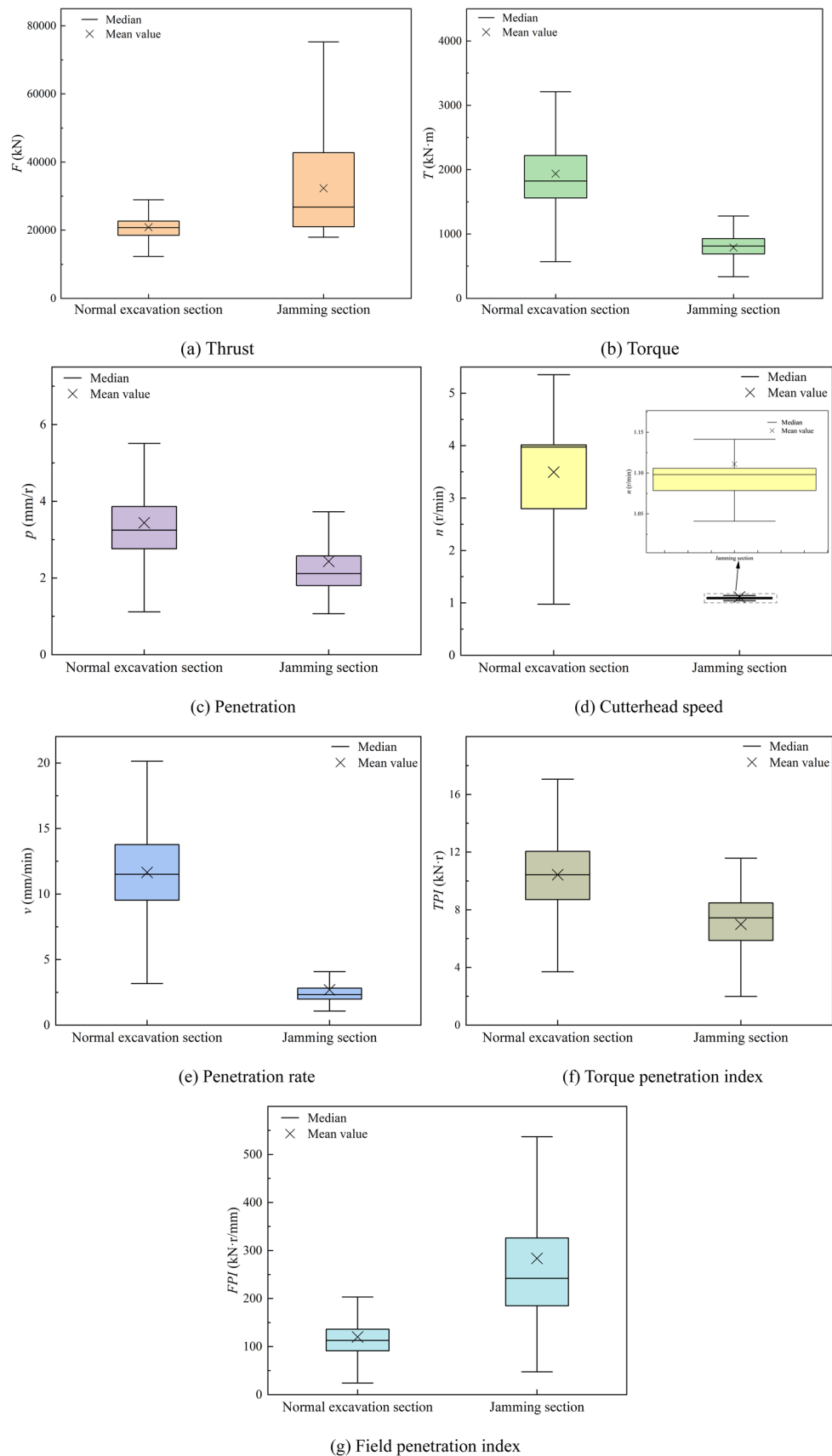
Similarly, on the right side of  $x_i^c$  in the figure, there are samples that include correctly classified non-jam category samples and jam samples that were mistakenly classified as non-jam. Among these, the proportion of jam samples misclassified as non-jam is known as the false negative rate ( $FNR$ ), which can be estimated by the cumulative probability density of  $x_i^{\text{jam}} > x_i^c$ , as shown in the red shaded area in the Fig. 3, corresponding to Eq. 2.

$$FPR = \Phi_i^{\text{non-jam}}(x_i^{\text{non-jam}}) = \int_{\min(x_i^{\text{non-jam}})}^{x_i^c} \varphi_i^{\text{non-jam}}(x_i^{\text{non-jam}}) dx \quad (1)$$

$$FNR = 1 - \Phi_i^{\text{jam}}(x_i^{\text{jam}}) = \int_{x_i^c}^{\max(x_i^{\text{jam}})} \varphi_i^{\text{jam}}(x_i^{\text{jam}}) dx \quad (2)$$

Here  $\Phi_i^{\text{jam}}$  and  $\varphi_i^{\text{jam}}$  respectively represent the cumulative probability density function (CDF) and probability density function (PDF) of feature  $X_i$  in the jam category samples;  $\Phi_i^{\text{non-jam}}$  and  $\varphi_i^{\text{non-jam}}$  respectively represent the CDF and PDF of feature  $X_i$  in the non-jam category samples.

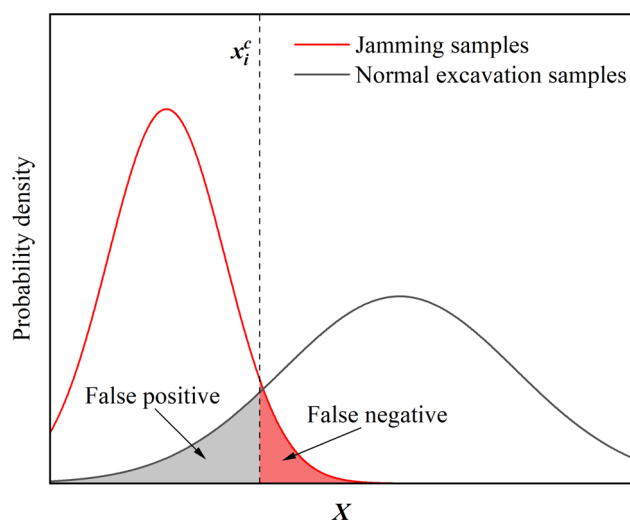
As can be seen from Fig. 3, as  $x_i^c$  increases, the  $FPR$  will also increase, while the  $FNR$  will decrease. To minimize the sum of  $FPR$  and  $FNR$ , the following relationship should be satisfied:



**Fig. 2.** Statistical characteristics of typical parameters in two types of samples.

Chainage range	Length (m)	Ring number range	Strata type	Excavation mode
8830.48 ~ 8615.55	214.93	1 ~ 120	Sandy clay	EPB
8615.55 ~ 8606.56	8.99	121 ~ 125	Completely weathered granite	
8606.56 ~ 8590	16.56	126 ~ 133	Slightly weathered granite	
8590 ~ 8575	15	134 ~ 142		
8575 ~ 8564.1	10.9	143 ~ 148		
8564.1 ~ 8520	44.1	149 ~ 173		
8520 ~ 8380	140	174 ~ 251	Moderate and slightly weathered granite	TBM
8389 ~ 8363.8	16.2	252 ~ 260		
8363.8 ~ 8355	8.8	261 ~ 265	EPB	
8355 ~ 8335	20	266 ~ 276		Completely, strongly, moderate and slightly weathered granite
8335 ~ 8300	35	277 ~ 295		Completely and strongly weathered granite
8300 ~ 8230	70	296 ~ 335		Completely weathered granite
8230 ~ 8220	10	336 ~ 340		Completely, strongly, moderate and slightly weathered granite
8220 ~ 8170	50	341 ~ 367		Moderate and slightly weathered granite
8170 ~ 8093.55	76.45	368 ~ 410		Strongly and moderate weathered granite
8093.55 ~ 8070.15	23.4	411 ~ 423		Strongly, moderate and slightly weathered granite
8070.15 ~ 8040	30.15	424 ~ 440		Strongly and moderate weathered granite
8040 ~ 8015	25	441 ~ 454		Slightly weathered granite
8015 ~ 7990	25	455 ~ 468		
7990 ~ 7845	145	469 ~ 548		
7845 ~ 7825	20	549 ~ 560		
7825 ~ 7785	40	561 ~ 582		
7785 ~ 7765	20	583 ~ 593		
7765 ~ 7535	230	594 ~ 720		
7535 ~ 7515	20	721 ~ 732		
7515 ~ 7475	40	733 ~ 754		
7475 ~ 7455	20	755 ~ 765		
7455 ~ 7048.63	406.37	766 ~ 992		

**Table 2.** Geological condition and excavation mode along chainage.



**Fig. 3.** Distribution of feature  $X_i$  in both datasets.

$$\begin{cases} \min(FPR + FNR) = \min(\oplus_i^{\text{non-jam}}(x_i^{\text{non-jam}}) - \oplus_i^{\text{jam}}(x_i^{\text{jam}}) + 1) \\ x_i^c \in [\min(x_i^{\text{non-jam}}), \max(x_i^{\text{jam}})] \end{cases} \quad (3)$$

$$\begin{cases} (\oplus_i^{\text{non-jam}}(x_i^{\text{non-jam}}) - \oplus_i^{\text{jam}}(x_i^{\text{jam}}) + 1)' = 0 \\ x_i^c \in [\min(x_i^{\text{non-jam}}), \max(x_i^{\text{jam}})] \end{cases} \quad (4)$$

$$\begin{cases} \varphi_i^{\text{non-jam}}(x_i^{\text{non-jam}}) - \varphi_i^{\text{jam}}(x_i^{\text{jam}}) = 0 \\ x_i^c \in [\min(x_i^{\text{non-jam}}), \max(x_i^{\text{jam}})] \end{cases} \quad (5)$$

When the two distributions overlap within the range  $[\min(x_i^{\text{non-jam}}), \max(x_i^{\text{jam}})]$ , the threshold  $x_i^c$  must fall within this interval. When  $x_i^c$  is positioned at the intersection point of the two distribution curves, the sum of *FPR* and *FNR* can be minimized.

When the sample size is sufficiently large, it can be assumed that  $x_i^{\text{non-jam}}$  and  $x_i^{\text{jam}}$  are approximately normally distributed, and the following relationship holds:

$$\beta_i = \frac{\mu_i^{\text{non-jam}} - x_i^c}{\sigma_i^{\text{non-jam}}} = \frac{x_i^c - \mu_i^{\text{jam}}}{\sigma_i^{\text{jam}}} \quad (6)$$

Here  $\beta_i$  represents the distribution dissimilarity of the feature  $X_i$  between the two categories samples.  $\mu_i^{\text{non-jam}}$  and  $\sigma_i^{\text{non-jam}}$  are the mean and standard deviation of  $x_i^{\text{non-jam}}$ , respectively, while  $\mu_i^{\text{jam}}$  and  $\sigma_i^{\text{jam}}$  are the mean and standard deviation of  $x_i^{\text{jam}}$ , respectively. By rearranging Eq. 6, the following relationship can be obtained:

$$x_i^c = \mu_i^{\text{jam}} + \beta_i \sigma_i^{\text{jam}} = \mu_i^{\text{non-jam}} - \beta_i \sigma_i^{\text{non-jam}} \quad (7)$$

$$\beta_i = \frac{\mu_i^{\text{non-jam}} - \mu_i^{\text{jam}}}{\sigma_i^{\text{non-jam}} + \sigma_i^{\text{jam}}} \quad (8)$$

Here the larger the value of  $\beta_i$ , the more pronounced the statistical difference between  $x_i^{\text{non-jam}}$  and  $x_i^{\text{jam}}$ . As shown in Table 3, when the  $\beta_i$  value is 0, it indicates that the two distributions have the same mean, and at this point,  $x_i^c$  cannot effectively distinguish between the two categories. In this situation, both *FPR* and *FNR* are 50.0%. When the  $\beta_i$  value is 1, and both  $x_i^{\text{non-jam}}$  and  $x_i^{\text{jam}}$  strictly follow a normal distribution, the *FPR* and *FNR* are approximately 16.0%. When the  $\beta_i$  value is 2, and both  $x_i^{\text{non-jam}}$  and  $x_i^{\text{jam}}$  strictly follow a normal distribution, the *FPR* and *FNR* are approximately 2.0%. When the  $\beta_i$  value is 3, and both  $x_i^{\text{non-jam}}$  and  $x_i^{\text{jam}}$  strictly follow a normal distribution, the *FPR* and *FNR* are approximately 0.15%.

### Qualitative model for jamming risk

Through the above process, we can establish a qualitative discrimination model based on the feature  $X_i$  to identify the risk of jamming machines. This model employs a binary classification method by setting a decision threshold  $x_i^c$  to distinguish whether a sample poses a jamming risk or not.

Specifically, if the value  $x_i$  of the sample's feature  $X_i$  is lower than the corresponding decision threshold  $x_i^c$ , the sample is judged as "jam"; conversely, if  $x_i$  exceeds the threshold  $x_i^c$ , the sample is judged as "non-jam". The mathematical expression of this model is as follows:

$$IFJ = \begin{cases} 1 & \text{if } x_i < x_i^c \\ 0 & \text{if } x_i > x_i^c \end{cases} \quad (9)$$

Here *IFJ* represents the model's identification result, with a value of 1 indicating a positive identification result (in this case, "jam"), and a value of 0 indicating a negative identification result (here, "non-jam"). Additionally, this model contains another layer of meaning: the closer the value of  $x_i$  is to the threshold  $x_i^c$ , the higher the credibility of the identification result. This implies that samples near the threshold are more uncertain in their classification compared to those further away from  $x_i^c$ .

### Comprehensive evaluation index for jamming risk

Using Eq. 7 and 8, we can calculate the  $x_i^c$  and  $\beta_i$  values for classification using each individual indicator. Clearly, the  $\beta_i$  values for each indicator differ, indicating varying abilities of these indicators to distinguish between jam and non-jam conditions. A larger  $\beta_i$  value signifies a stronger overall recognition effect, but this does

$\beta_i$ -value	False positive rate, <i>FPR</i> (%)	False negative rate, <i>FNR</i> (%)
0	50.0	50.0
1	16.0	16.0
2	2.0	2.0
3	0.15	0.15

**Table 3.** Relationship between  $\beta_i$ -values and *FPR* and *FNR* in normal distribution.



not necessarily translate to better recognition for every specific sample. For instance, when  $\beta_1 > \beta_2$ , feature  $X_1$  outperforms feature  $X_2$  in overall recognition effectiveness, yet for certain specific samples, feature  $X_2$  may provide better discrimination than feature  $X_1$ . This indicates that a single feature can only reflect a limited aspect of the rock mass conditions, highlighting the importance of considering multiple features for robust classification.

To address this issue, we attempt to integrate all individual features for jamming risk warning to enhance overall accuracy. The specific steps are as follows:

- (1) Calculation of the weight of each individual feature.

Based on the distribution dissimilarity index  $\beta_p$ , the weight  $w_i$  of each single feature is calculated using Eq. 10:

$$w_i = \frac{\beta_i}{\sum_{i=1}^N \beta_i} \quad (10)$$

Here  $w_i$  represents the weight of feature  $X_p$ ,  $\beta_i$  is the distribution dissimilarity of feature  $X_p$  and  $N$  is the total number of features. A smaller distribution dissimilarity implies a lower weight, indicating that this feature carries less significance in the comprehensive evaluation.

- (2) Calculation of the comprehensive evaluation feature.

The comprehensive evaluation feature  $\eta$ , based on the weighted approach, can be calculated using Eq. 11:

$$\eta = \sum [1 - \oplus_i^{\text{jam}}(x_i^{\text{jam}})] w_i \quad (11)$$

Here  $\eta$  is the comprehensive feature used to evaluate jamming risk, with a value range of  $[0, 1]$ .  $x_i^{\text{jam}}$  represents the value of feature  $X_i$  in the jam category.  $\oplus_i^{\text{jam}}$  denotes the CDF of feature  $X_i$  in the jam category, calculated as follows:

$$\oplus_i^{\text{jam}}(x_i^{\text{jam}}) = \Psi\left(\frac{x_i^{\text{jam}} - \mu_i^{\text{jam}}}{\sigma_i^{\text{jam}}}\right) \quad (12)$$

Here  $\Psi$  represents the CDF of the standard normal distribution.

### Probability model for jamming risk

This section focuses on constructing a quantitative probability model based on Bayesian conditional probability theory to assess the risk of machine jamming during TBM excavation processes. By collecting real-time data, we calculate the comprehensive feature  $\eta$  of the current sample. Using the qualitative identification model, we make a preliminary judgment to determine whether the sample belongs to a “jam” situation. However, while the qualitative model identifies “jam” states, it cannot quantify the probability of jamming. To address this limitation, we establish a quantitative probability model for jamming risk warning, leveraging the comprehensive feature  $\eta$  and Bayesian theory. The derivation process is as follows:

Assuming the comprehensive feature of the current sample is  $\eta'$ , and the recognition result of the qualitative model is  $IFJ(\eta')$  (where  $IFJ=1$  for “positive” and  $IFJ=0$  for “negative”), we use the posterior conditional probability  $P(\text{jam} | IFJ(\eta'))$  to evaluate the reliability or accuracy of the qualitative model. For instance,  $P(\text{jam} | \text{positive}) = 0.9$  indicates that the credibility of the qualitative model judging the current rock mass as “positive” (jam) aligns with the actual situation (jam) at a 90.0% confidence level.

From previous analysis, we know that the comprehensive feature  $\eta$ , similar to indicators such as  $TPI$  and  $FPI$ , reflects the quality of the rock mass through its value. Therefore, we use the posterior conditional probability to estimate the probability  $R(\eta')$  of a jamming occurring in the current rock mass. This includes two events: the conditional event is the recognition result  $IFJ(\eta')$  (where  $IFJ=1$  for “positive” and  $IFJ=0$  for “negative”) of the qualitative model for the current sample, and the outcome event is that the real label of the sample is “jam”. According to Bayes’ theorem,  $P(\text{jam} | IFJ(\eta'))$  can be expressed as:

$$R(\eta') = P(\text{jam} | IFJ(\eta')) = \frac{P(\text{jam}) \times P(IFJ(\eta') | \text{jam})}{P(IFJ(\eta'))} \quad (13)$$

This contains three terms, which we will introduce one by one.

- (1)  $P(IFJ(\eta'))$  represents the ability of the qualitative model to recognize samples as “positive” or “negative”. Taking  $IFJ(\eta') = 1$  (positive) as an example, this includes two scenarios:

**Correct identification:** The real label of the sample is “jam”, and the identification result is also positive.

**False alarm:** The real label of the sample is “non-jam”, but the identification result is positive.

Similarly,  $IFJ(\eta') = 0$  (negative), it also includes two scenarios:

**Correct identification:** The real label of the sample is “non-jam”, and the identification result is also negative.

**Missing:** The real label of the sample is “jam”, but the identification result is negative.

We can express this using the following formula:

$$P(IFJ(\eta')) = P(IFJ(\eta') \cap \text{jam}) + P(IFJ(\eta') \cap \text{non-jam}) \quad (14)$$

According to the method of calculating joint probability, the two terms on the right side of Eq. 14 can be expressed in the following forms:

$$P(IFJ(\eta') \cap \text{jam}) = P(IFJ(\eta') | \text{jam}) \times P(\text{jam}) \quad (15)$$

$$P(IFJ(\eta') \cap \text{non-jam}) = P(IFJ(\eta') | \text{non-jam}) \times P(\text{non-jam}) \quad (16)$$

- (2)  $P(\text{jam})$  and  $P(\text{non-jam})$  represent the probabilities of a sample belonging to “jam” or “non-jam” categories. These probabilities can be determined in two ways:

**First approach:** Using prior probabilities. The prior probabilities are derived from the proportions of samples belonging to the “jam” and “non-jam” categories within the entire dataset. This can be formulated as follows:

$$\begin{cases} P(\text{jam}) = \frac{N_2}{N_1+N_2} \\ P(\text{non-jam}) = \frac{N_1}{N_1+N_2} \end{cases} \quad (17)$$

Here  $N_1$  and  $N_2$  represent the number of “non-jam” samples and “jam” samples, respectively, in the original dataset.

**Second approach:** Treating both categories as equally likely. This approach disregards the prior probabilities and assumes that both categories have an equal probability of appearing. This can be expressed as:

$$\begin{cases} P(\text{jam}) = 0.5 \\ P(\text{non-jam}) = 0.5 \end{cases} \quad (18)$$

In this case, the occurrence of both categories is treated as a random event with no bias toward either category.

- (3)  $P(IFJ(\eta') | \text{jam})$  represents the likelihood of the model’s prediction given the true label. These terms represent the probability that a qualitative model predicts the sample as  $IFJ(\eta')$ , given the true label of the sample. Specifically:

$P(IFJ(\eta') | \text{jam})$  is the likelihood that when the sample belongs to the “jam” category, it falls near a specific value  $\eta'$ .

$P(IFJ(\eta') | \text{non-jam})$  is the likelihood that when the sample belongs to the “non-jam” category, it falls near a specific value  $\eta'$ .

This can be expressed using the following formula:

$$\begin{cases} P(IFJ(\eta') | \text{jam}) = \varphi^{\text{jam}}(\eta') \\ P(IFJ(\eta') | \text{non-jam}) = \varphi^{\text{non-jam}}(\eta') \end{cases} \quad (19)$$

Here,  $\varphi^{\text{jam}}$  and  $\varphi^{\text{non-jam}}$  represent the prior distribution functions for samples belonging to the “jam” and “non-jam” categories, respectively. These distributions reflect the inherent characteristics of the rock mass quality indicators (e.g.,  $TPI$ ,  $FPI$ ) under each category, providing a basis for calculating the likelihood of the model’s predictions.

- (4) Calculating the probability of jamming risk using substituted terms: By substituting these terms into Eq. 13, can calculate the probability of the current tunnel face rock mass experiencing jamming risk. The final expression is as follows:

$$\begin{cases} R_1(\eta') = \frac{N_2 \cdot \varphi^{\text{jam}}(\eta')}{N_2 \cdot \varphi^{\text{jam}}(\eta') + N_1 \cdot \varphi^{\text{non-jam}}(\eta')} & \text{Using prior probabilities} \\ R_2(\eta') = \frac{\varphi^{\text{jam}}(\eta')}{\varphi^{\text{jam}}(\eta') + \varphi^{\text{non-jam}}(\eta')} & \text{Not using prior probabilities} \end{cases} \quad (20)$$

Here  $R_1$  represents the quantitative probability model that considers the influence of the category proportions in the prior distribution, while  $R_2$  represents the quantitative probability model that does not consider the influence of the category proportions in the prior distribution.

## Results and discussion

### Performance of qualitative model

Through Equations 7 and 8, Table 4 summarizes the statistical outcomes of key tunneling parameters across two sub datasets and the calculated distribution dissimilarity  $\beta$  for each parameter. The analysis reveals the following insights: The  $\beta_p$  is relatively small, indicating minimal variability. This makes it unsuitable for identifying jamming states, as there is insufficient distinction between “jam” and “non-jam” conditions. The parameter  $n$  is primarily influenced by the driver’s subjective judgment and operational habits, leading to high variability. Due to its dependence on human factors, it is also unsuitable for jamming state identification. Other five tunneling parameters ( $T$ ,  $F$ ,  $v$ ,  $TPI$ ,  $FPI$ ) exhibit significant distribution dissimilarity  $\beta$  in both “jam” and “non-jam”



Tunneling parameter	Jamming section			Normal excavation section			$\beta_i$
	Mean value $\mu_1$	Standard deviation $\sigma_1$	Coefficient of variation $c_{v1}$	Mean value $\mu_2$	Standard deviation $\sigma_2$	Coefficient of variation $c_{v2}$	
$T$	789.18	207.19	0.26	1938.08	587.47	0.30	1.45
$F$	32,303.86	14,205.25	0.44	20,863.98	4787.94	0.23	0.60
$p$	2.43	1.49	0.61	3.44	15.80	4.59	0.06
$n$	1.11	0.09	0.78	3.50	0.97	0.28	2.25
$v$	2.70	1.58	0.58	11.65	14.27	1.22	0.56
$TPI$	7.00	2.27	0.32	10.44	2.63	0.25	0.70
$FPI$	283.81	156.74	0.55	120.27	62.96	0.52	0.74

**Table 4.** Statistical analysis results of the key tunneling parameters in two datasets.

samples. This large variability makes them effective indicators for identifying jamming states. Figure 4 illustrates the distribution results of the five selected tunneling parameters ( $T$ ,  $F$ ,  $v$ ,  $TPI$ ,  $FPI$ ) in both jammed and normal excavation sections. The key observations are that the distributions of these parameters show considerable overlap between jammed and normal conditions, suggesting that they alone may not provide a clear distinction.

Figure 5 illustrates the identification performance of jamming states using the selected tunneling parameters  $T$ ,  $F$ ,  $v$ ,  $TPI$ , and  $FPI$  along with their corresponding thresholds. The samples within a 300 m interval near the jamming section are analyzed to distinguish between normal excavation and jammed conditions. Notably, significant mutations in the samples are observed in the jamming section, and the thresholds effectively separate most non-jamming state samples from jamming state samples.

For example, in Fig. 5a, the red dashed line represents the threshold for  $T$ , defined as  $T^c = 1090$ . Black sample points correspond to data from the normal excavation section, while purple points represent data from the jamming section. Visually, it is evident that using  $T^c = 1090$  as the threshold successfully identifies most jamming samples. However, some non-jamming samples are incorrectly classified as jamming. Similarly, the identification results using  $F$ ,  $v$ ,  $TPI$ , and  $FPI$  as indicators are presented in Figs. 5b–e, demonstrating consistent performance with varying degrees of accuracy and misclassification.

To comprehensively account for the influence of various parameters and further enhance identification performance, the weights of each parameter were calculated using Eq. 10, and a comprehensive perception index  $\eta$  for each sample group was derived via Eq. 11. The identification rate of the qualitative model  $IFJ$  in the jamming section dataset was computed using Eq. 9, while the false positive rate ( $FPR$ ) and false negative rate ( $FNR$ ) for each parameter's threshold and the comprehensive perception index were determined through Eqs. 1 and 2. The parameter weights and identification rates are summarized in Table 5, and the distribution of the comprehensive perception index along with a comparison of parameter identification performance is illustrated in Fig. 6.

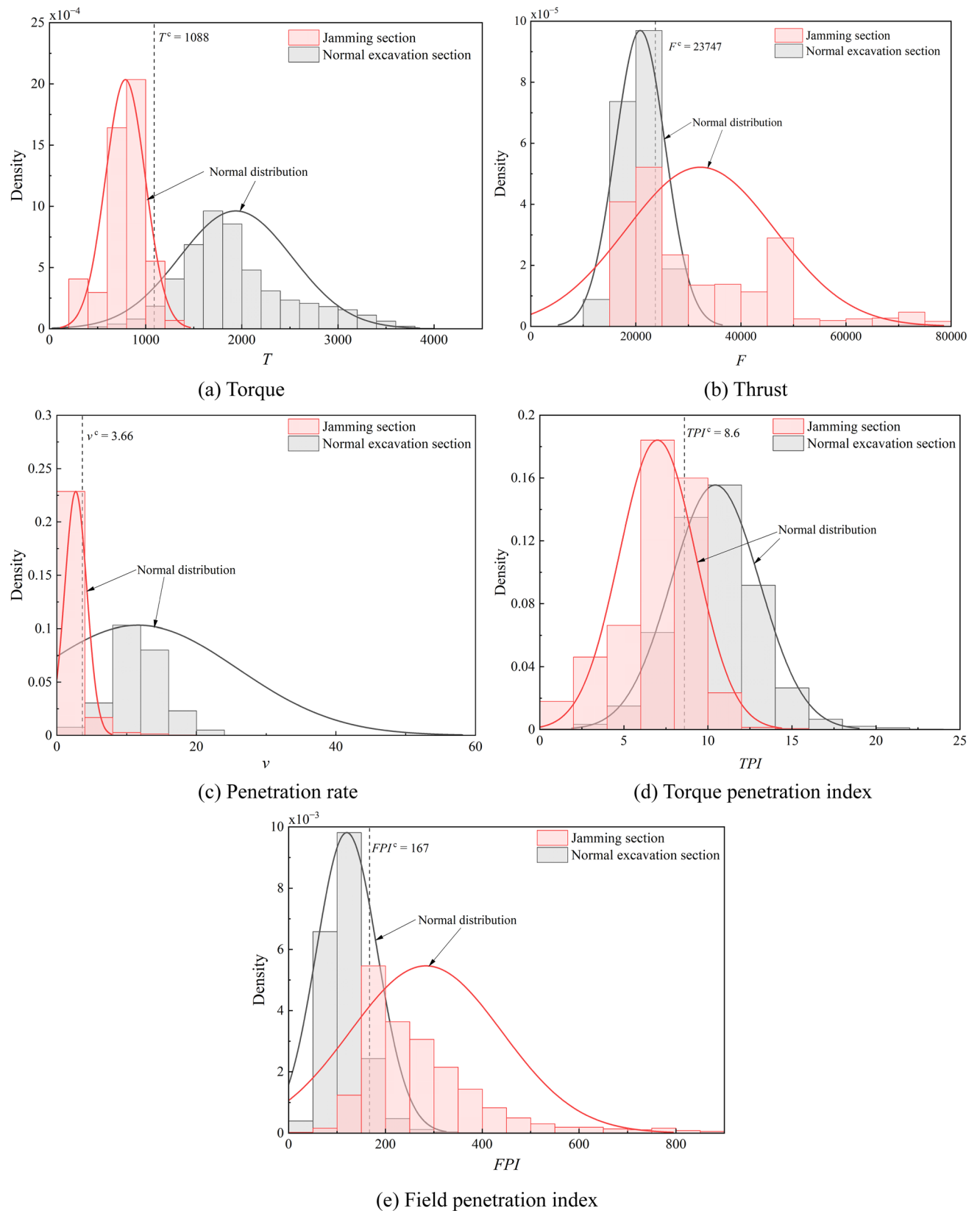
From Table 4 and Fig. 6, it is evident that the distribution dissimilarity of the comprehensive perception index  $\eta$  is significant between the jamming and normal excavation sections, demonstrating its strong ability to identify jamming states. As shown in Table 4,  $\eta$  achieves the highest  $IFJ$  identification rate among all six parameters ( $T$ ,  $F$ ,  $v$ ,  $TPI$ ,  $FPI$ , and  $\eta$ ), successfully identifying 97.8% of jamming data in the jamming section dataset with only a 2.2% false negative rate. This indicates that  $\eta$ , which integrates multiple parameters, is highly effective in distinguishing whether a given excavation data sample is in a jamming state.

Figure 6a illustrates the distribution of  $\eta$  across both datasets, revealing that nearly all samples in the jamming section dataset exceed the threshold line (Fig. 6b). A more detailed analysis in Fig. 6c shows that the qualitative model using  $\eta$  also maintains a high identification rate for normal excavation states, achieving 96.1% accuracy with a 3.9% false positive rate. Furthermore, Fig. 6d compares the  $FPR$  and  $FNR$  of each parameter's threshold, highlighting  $\eta$  as the optimal indicator with the lowest  $FPR$  and  $FNR$  among all six parameters.

### Performance of quantitative model

In the case studied in this article, the sample size of the jamming section is 743, while the sample size of the normal excavation section is 79,531. Following Eq. 17, the proportion of jamming samples  $P(Jam) = 0.009$  and the proportion of normal excavation samples  $P(Non-jam) = 0.991$ . By inputting the statistical results of  $\eta$  into Eq. 20, the probability of jamming corresponding to each group of samples can be calculated. Figure 7 illustrates the trend of jamming risk near 100 m in the area where the accident occurred, and Table 5 presents the jamming probabilities for all data samples.

As shown in Table 5, without considering the influence of sample size, the average probability of jamming in the jamming section is 77%, while in the normal excavation section it is 2%. However, after accounting for the significant difference in sample size, the average probability of jamming in the jamming section increases to 94%, and in the normal excavation section it rises to 7%. This indicates that the disparity in sample size between the jamming and normal excavation sections leads to an underestimation of jamming risk when sample size is not considered. In reality, jamming accidents are not sudden but develop over time<sup>22</sup>. As depicted in Fig. 7a, the jamming probability appears relatively stable with minimal fluctuations due to the imbalanced dataset, resulting in a lack of sensitivity in the calculated risk  $R$ . When a jamming accident occurs, the probability abruptly increases to a very high value. However, after incorporating the influence of sample size, the jamming probability dynamically reflects real-time changes in the data, as shown in Fig. 7b. Thus, the calculated results of  $R_2$ , which account for sample size, align more closely with actual conditions.



**Fig. 4.** Probability density distribution of tunneling parameters.

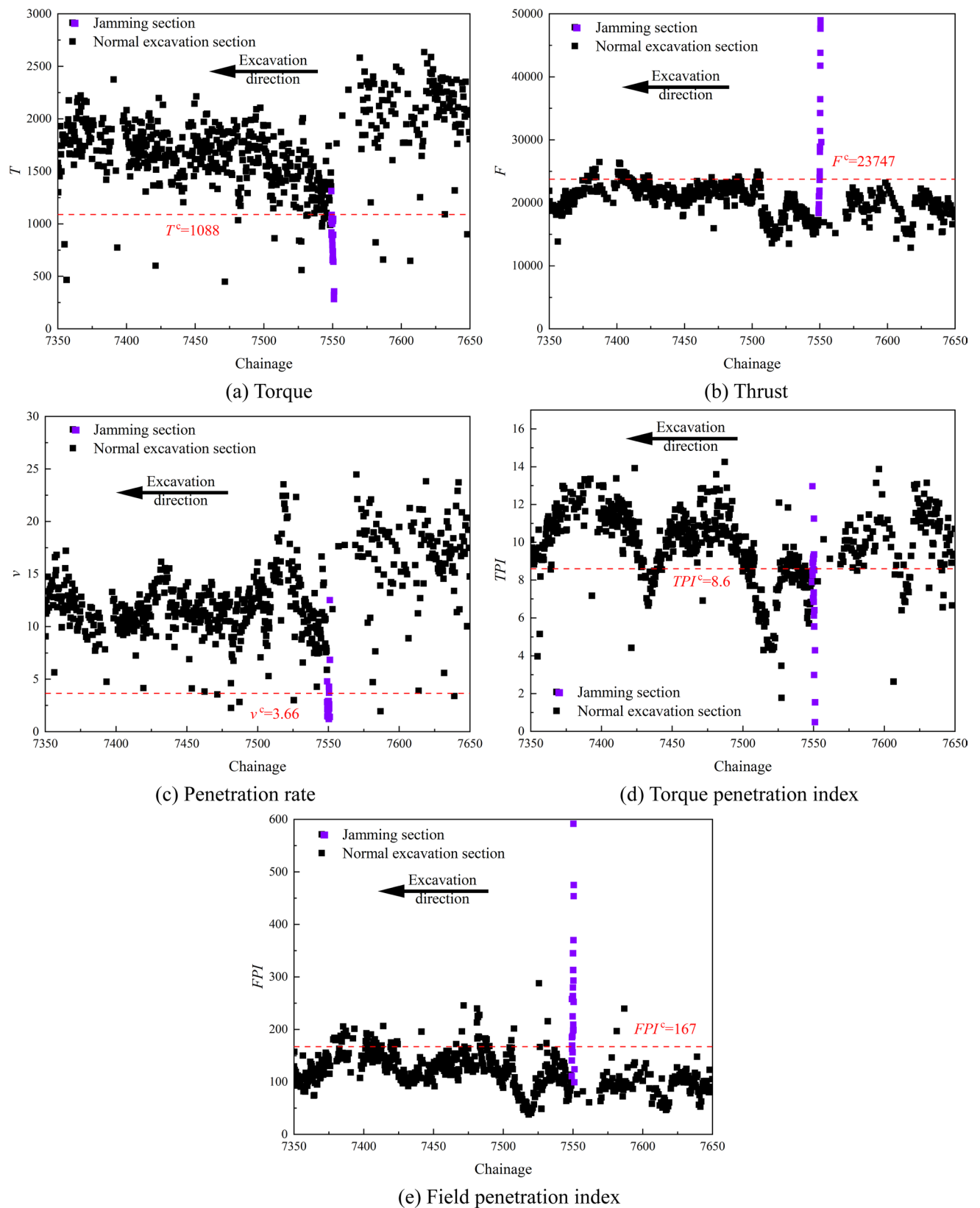
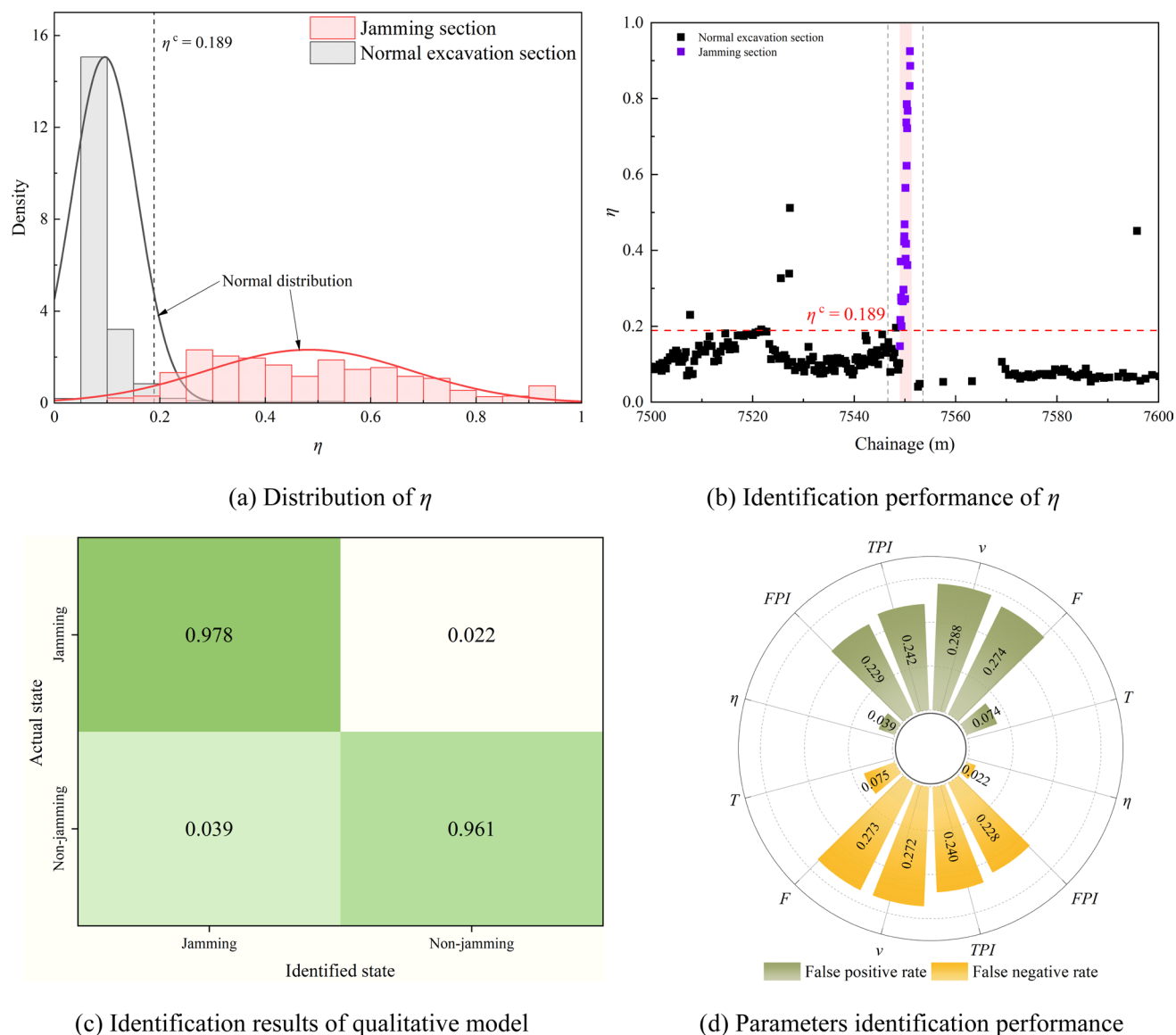


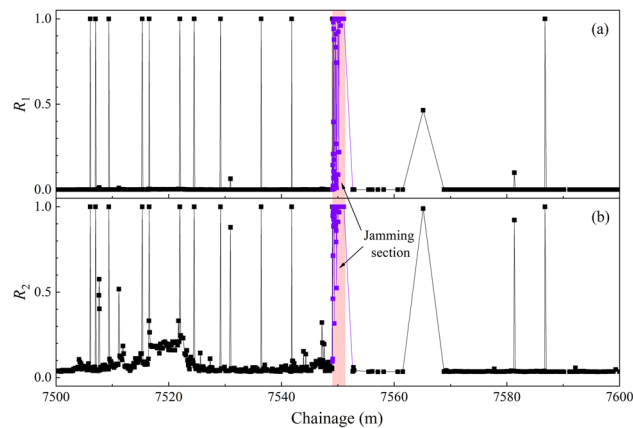
Fig. 5. Identification performance of the threshold.

Index	Jamming section	Normal excavation section
$R_1$	0.77	0.02
$R_2$	0.94	0.07

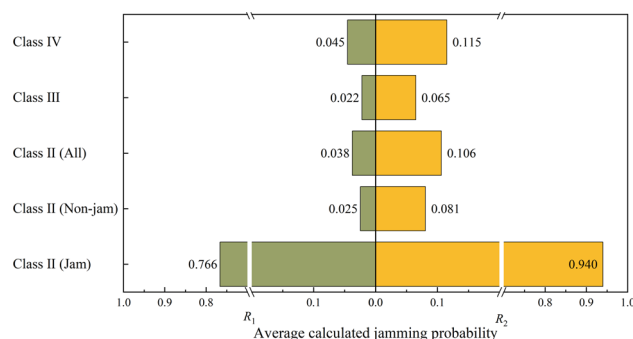
**Table 5.** Average calculation result of risk probability.**Fig. 6.** Display of the identification performance of various parameters.**Performance of quantitative model in different rock mass**

As shown in Fig. 1, the TBM mode is applied in hard rock strata, including classes II, III, and IV, with the jamming accident occurring in class II surrounding rock. The jamming probability varies during excavation in different classes of surrounding rock, as illustrated in Fig. 8. Rock class is directly related to its strength, with the following hierarchy: Class II > Class III > Class IV in terms of strength. The wear rate of disc cutters increases with rock strength<sup>37</sup>, and excessive wear can lead to abnormal conditions or even jamming accidents.

From Fig. 8, it is evident that the average jamming probability in class II surrounding rock is approximately 3% higher than in classes III and IV. Even after excluding data from the jamming section, the average jamming probability in class II remains about 0.5% higher than in classes III and IV. Additionally, the average probability in class IV is 0.02% higher than in class III, likely due to the higher prevalence of weak and fractured rock masses in lower-strength rocks, which increase the likelihood of jamming accidents. Thus, class III rock is identified as



**Fig. 7.** Probability calculation results of jamming risk: **a**  $R_1$  model; **b**  $R_2$  model.



**Fig. 8.** Average calculated jamming probability under different rock class via model.

the most suitable stratum for excavation. It effectively reduces jamming probability and minimizes disc cutter wear rates, thereby enhancing overall excavation efficiency.

### Practical application in TBM operations

Our comprehensive Bayesian statistical framework demonstrates significant improvements in jamming risk assessment. The comprehensive jamming perception index  $\eta$  achieves a 95% recognition rate, significantly higher than single-parameter models. The refined jamming probability model, which accounts for sample size differences, provides more accurate estimates of jamming probability, with an average of 94% in jamming sections and 7% in normal excavation sections. These results highlight the effectiveness of our proposed framework. In practical TBM operations, our model can be used in real-time to assist engineers in risk assessment and decision-making:

- (1) **Real-time Data Monitoring** The model calculates the comprehensive jamming perception index  $\eta$  and jamming probability  $R_2$  in real-time based on monitoring data. Engineers can obtain these indicators through the TBM control system to promptly understand the current jamming risk.
- (2) **Early Warning System Integration:** The model can be integrated into the TBM's early warning system. When the jamming probability exceeds a preset threshold, the system can automatically trigger an alarm to alert engineers to take appropriate measures.
- (3) **Decision Support:** The quantitative risk assessment provided by the model helps engineers make more informed decisions. For example, if the jamming probability is high, engineers can adjust the TBM operating parameters or take preventive measures.
- (4) **Construction Plan Optimization:** By analyzing the jamming risk under different geological conditions, engineers can optimize the construction plan and choose safer construction routes and methods.

### Conclusions

This study presents a comprehensive Bayesian statistical framework for TBM jamming risk assessment, which effectively enhances the accuracy and reliability of jamming risk assessment through a comprehensive multi-parameter jamming perception index  $\eta$  and a jamming probability model that accounts for sample size differences. The main conclusions are as follows:

- (1) By introducing the comprehensive jamming perception index  $\eta$ , we overcame the limitations of single-parameter analysis and significantly improved the recognition rate of jamming states. The experimental results show that the recognition rate of  $\eta$  reaches 95%, which is much higher than that of single-parameter models. This demonstrates the significant advantages of multi-parameter integrated assessment in jamming risk assessment.
- (2) By developing a new jamming probability model  $R_2$  that accounts for the differences in sample sizes between jamming and normal excavation samples, we made the model more practical and accurate in real applications. In actual jamming accidents, the average jamming probability in jamming sections is 94%, and in normal excavation sections, it is 7%. This indicates that considering sample size differences is crucial for improving model performance.
- (3) Through geological analysis, we found that Class III surrounding rock is the most suitable stratum for excavation, with the lowest jamming probability. This finding highlights the importance of considering geological conditions in excavation planning to effectively reduce jamming risks and improve construction efficiency.
- (4) Our Bayesian statistical method not only provides a clear decision-making process and the contribution of each factor to the risk results but also quantitatively offers the probability of jamming events. This quantitative assessment capability is crucial for site managers, as it can provide them with more specific and reliable references for their decision-making.

Although this study has yielded valuable findings regarding the TBM jamming risk assessment, several aspects still warrant further investigation. Future work will focus on the following priorities: First, we will collect additional independent datasets to verify the model's applicability and generalizability across diverse geological conditions and engineering scenarios. Second, we will conduct further optimization of the model to enhance its performance in addressing imbalanced data and adapting to complex geological environments.

### Data availability

The datasets used and analyzed during the current study available from the corresponding author on reasonable request.

Received: 12 August 2024; Accepted: 8 September 2025

Published online: 10 October 2025

### References

1. Armaghani, D. J., Mohamad, E. T., Narayanasamy, M. S., Narita, N. & Yagiz, S. Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition. *Tunn. Undergr. Space Technol.* **63**, 29–43. <https://doi.org/10.1016/j.tust.2016.12.009> (2017).
2. Gao, L. & Li, X. B. Utilizing partial least square and support vector machine for TBM penetration rate prediction in hard rock conditions. *J. Cent. South Univ.* **22**, 290–295. <https://doi.org/10.1007/s11771-015-2520-z> (2015).
3. Mahdevari, S., Shahriar, K., Yagiz, S. & Shirazi, M. A. A support vector regression model for predicting tunnel boring machine penetration rates. *Int. J. Rock Mech. Min.* **72**, 214–229. <https://doi.org/10.1016/j.ijrmm.2014.09.012> (2014).
4. Zheng, Y. L. & He, L. TBM tunneling in extremely hard and abrasive rocks: Problems, solutions and assisting methods. *J. Cent. South Univ.* **28**, 454–480. <https://doi.org/10.1007/s11771-021-4615-z> (2021).
5. Zhou, J. et al. Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate. *Eng. Appl. Artif. Intel.* **97**, 104015. <https://doi.org/10.1016/j.engappai.2020.104015> (2021).
6. Chen, Z. Y., Zhang, Y. P., Li, J. B., Li, X. & Jing, L. J. Diagnosing tunnel collapse sections based on TBM tunneling big data and deep learning: A case study on the Yinsong Project. *China. Tunn. Undergr. Space Technol.* **108**, 103700. <https://doi.org/10.1016/j.tust.2020.103700> (2021).
7. Guo, D. et al. Advance prediction of collapse for TBM tunneling using deep learning method. *Eng. Geol.* **299**, 106556. <https://doi.org/10.1016/j.enggeo.2022.106556> (2022).
8. Hou, S. K. & Liu, Y. R. Early warning of tunnel collapse based on Adam-optimised long short-term memory network and TBM operation parameters. *Eng. Appl. Artif. Intel.* **112**, 104842. <https://doi.org/10.1016/j.engappai.2022.104842> (2022).
9. Wang, X. T. et al. An interval risk assessment method and management of water inflow and inrush in course of karst tunnel excavation. *Tunn. Undergr. Space Technol.* **92**, 103033. <https://doi.org/10.1016/j.tust.2019.103033> (2019).
10. Xue, Y. G. et al. Water and mud inrush hazard in underground engineering: Genesis, evolution and prevention. *Tunn. Undergr. Space Technol.* **114**, 103987. <https://doi.org/10.1016/j.tust.2021.103987> (2021).
11. Gong, Q. M., Yin, L. J., Wu, S. Y., Zhao, J. & Ting, Y. Rock burst and slabbing failure and its influence on TBM excavation at headrace tunnels in Jinping II hydropower station. *Eng. Geol.* **124**, 98–108. <https://doi.org/10.1016/j.enggeo.2011.10.007> (2012).
12. Naji, A. M., Emad, M. Z., Rehman, H. & Yoo, H. Geological and geomechanical heterogeneity in deep hydropower tunnels: A rock burst failure case study. *Tunn. Undergr. Space Technol.* **84**, 507–521. <https://doi.org/10.1016/j.tust.2018.11.009> (2019).
13. Yang, S. Q., Tao, Y., Xu, P. & Chen, M. Large-scale model experiment and numerical simulation on convergence deformation of tunnel excavating in composite strata. *Tunn. Undergr. Space Technol.* **94**, 103133. <https://doi.org/10.1016/j.tust.2019.103133> (2019).
14. Zhao, K. et al. Computational modelling of the mechanised excavation of deep tunnels in weak rock. *Comput. Geotech.* **66**, 158–171. <https://doi.org/10.1016/j.compgeo.2015.01.020> (2015).
15. Wang, B., Li, Q. K., Xu, Z. F., Niu, S. B. & Nie, X. T. A safety risk assessment method for TBM tunnel construction based on attribute interval identification theory. *Sci. Rep.* **15**, 7673. <https://doi.org/10.1038/s41598-025-92375-0> (2025).
16. Sharafat, A., Latif, K. & Seo, J. Risk analysis of TBM tunneling projects based on generic bow-tie risk analysis approach in difficult ground conditions. *Tunn. Undergr. Space Technol.* **111**, 103860. <https://doi.org/10.1016/j.tust.2021.103860> (2021).
17. Zhang, Y. et al. Study on warning method for fault rockburst in deep TBM tunnels. *Rock Mech. Rock Eng.* **57**, 5557–5574. <https://doi.org/10.1007/s00603-024-03830-9> (2024).
18. Liu, Q. S., Huang, X., Shi, K. & Liu, X. W. Jamming mechanism of full face tunnel boring machine in over thousand-meter depths. *J. China Coal Soc.* **38**(1), 78–84. <https://doi.org/10.13225/j.cnki.jccs.2013.01.026c> (2013).
19. Liu, Q. S., Huang, X., Shi, K. & Zhu, Y. G. The mechanism of TBM shield jamming disaster tunneling through deep squeezing ground. *J. China Coal Soc.* **39**(S1), 78–84. <https://doi.org/10.13225/j.cnki.jccs.2012.1382> (2014).



20. Zhang, J. Z. & Zhou, X. P. Time-dependent jamming mechanism for Single-Shield TBM tunneling in squeezing rock. *Tunn. Undergr. Space Technol.* **69**, 209–222. <https://doi.org/10.1016/j.tust.2017.06.020> (2017).
21. Huang, X. et al. Development and in-situ application of a real-time monitoring system for the interaction between TBM and surrounding rock. *Tunn. Undergr. Space Technol.* **81**, 187–208. <https://doi.org/10.1016/j.tust.2018.07.018> (2018).
22. Xu, Z. H., Yu, T. F., Lin, P., Wang, W. Y. & Shao, R. Q. Integrated geochemical, mineralogical, and microstructural identification of faults in tunnels and its application to TBM jamming analysis. *Tunn. Undergr. Space Technol.* **128**, 104650. <https://doi.org/10.1016/j.tust.2022.104650> (2022).
23. Lin, P., Yu, T. F., Xu, Z. H., Shao, R. Q. & Wang, W. Y. Geochemical, mineralogical, and microstructural characteristics of fault rocks and their impact on TBM jamming: A case study. *Bull. Eng. Geol. Environ.* **81**, 64. <https://doi.org/10.1007/s10064-021-02548-0> (2022).
24. Hasanpour, R., Schmitt, J., Ozelik, Y. & Rostami, J. Examining the effect of adverse geological conditions on jamming of a single shielded TBM in Uluabat tunnel using numerical modeling. *J. Rock Mech. Geotech.* **9**(6), 1112–1122. <https://doi.org/10.1016/j.jrmge.2017.05.010> (2017).
25. Liu, L. P., Wang, X. G., Li, C. B. & Tian, Z. H. Jamming of the double-shield tunnel boring machine in a deep tunnel in Nyingchi, Tibet Autonomous Region. *China. Tunn. Undergr. Space Technol.* **131**, 104819. <https://doi.org/10.1016/j.tust.2022.104819> (2023).
26. Yu, H. D., Li, Y. Y. & Li, L. Evaluating some dynamic aspects of TBMs performance in uncertain complex geological structures. *Tunn. Undergr. Space Technol.* **96**, 103216. <https://doi.org/10.1016/j.tust.2019.103216> (2020).
27. Hasanpour, R., Rostami, J., Schmitt, J. & Ozelik, Y. Prediction of TBM jamming risk in squeezing grounds using Bayesian and artificial neural networks. *J. Rock Mech. Geotech.* **12**(1), 21–31. <https://doi.org/10.1016/j.jrmge.2019.04.006> (2020).
28. Bieniawski, Z. T. *Engineering rock mass classifications: a complete manual for engineers and geologists in mining, civil, and petroleum engineering* (Wiley, Hoboken, 1989).
29. Sen, Z. & Sadagah, B. H. Modified rock mass classification system by continuous rating. *Eng. Geol.* **67**(3–4), 269–280. [https://doi.org/10.1016/S0013-7952\(02\)00185-0](https://doi.org/10.1016/S0013-7952(02)00185-0) (2003).
30. Liu, Q. S., Liu, J. P., Pan, Y. C., Kong, X. X. & Hong, K. R. A case study of TBM performance prediction using a Chinese rock mass classification system: Hydropower classification (HC) method. *Tunn. Undergr. Space Technol.* **65**, 140–154. <https://doi.org/10.1016/j.tust.2017.03.002> (2017).
31. Li, J. B. et al. Feedback on a shared big dataset for intelligent TBM Part I: Feature extraction and machine learning methods. *Undergr. Space* **11**, 1–25. <https://doi.org/10.1016/j.undsp.2023.01.001> (2023).
32. Hassanpour, J., Rostami, J., Khamchayan, M., Bruland, A. & Tavakoli, H. R. TBM performance analysis in pyroclastic rocks: A case history of Karaj Water conveyance tunnel. *Rock Mech. Rock Eng.* **43**, 427–445. <https://doi.org/10.1007/s00603-009-0060-2> (2010).
33. Jing, L. J. et al. A case study of TBM performance prediction using field tunnelling tests in limestone strata. *Tunn. Undergr. Space Technol.* **83**, 364–372. <https://doi.org/10.1016/j.tust.2018.10.001> (2019).
34. Li, J. B. et al. Feedback on a shared big dataset for intelligent TBM Part II: Application and forward look. *Undergr. Space* **11**, 26–45. <https://doi.org/10.1016/j.undsp.2023.01.002> (2023).
35. Li, X., Wu, L. J., Wang, Y. J. & Li, J. H. Rock fragmentation indexes reflecting rock mass quality based on real-time data of TBM tunneling. *Sci. Rep.* **13**, 10420. <https://doi.org/10.1038/s41598-023-37306-7> (2023).
36. Wu, L. J., Li, X., Yuan, J. D. & Wang, S. J. Real-time prediction of tunnel face conditions using XGBRF algorithm. *Front. Struct. Civ. Eng.* **17**, 1777–1795. <https://doi.org/10.1007/s11709-023-0044-4> (2023).
37. Wu, L. J., Wang, L. C., Wang, S. J., Wang, Y. & Li, X. Prediction of tunnel face rock mass classification using an ensemble model enhanced by feature cross based on TBM boring data. *Tunn. Undergr. Space Technol.* **162**, 106647. <https://doi.org/10.1016/j.tust.2025.106647> (2025).

## Acknowledgements

This work was financially supported by grants from the Key Research Project of China Railway Hi-Tech Industry Corporation Limited (Grant No. JBGS-2024-01). We sincerely give our thanks to the data support from China Railway Engineering Equipment Group Corporation. We also thank the anonymous reviewers for their helpful remarks.

## Author contributions

S.-J.W.: Software, Conceptualization, Project administration, Resources. L.-C.W.: Writing, Experiment, Data curation, Visualization. L.-J.W.: Writing, Conceptualization, Data curation, Supervision. X.L.: Writing—Review & Editing, Validation.

## Declarations

## Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Additional information

**Correspondence** and requests for materials should be addressed to L.-J.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2025