Transportation Infrastructure Geotechnology (2026) 13:56
https://doi.org/10.1007/s40515-026-00810-0

RESEARCH

# A Hybrid Intelligent Framework for the Assessment and Classification of Squeezing Potential of Rocks Around Tunnels

**Muhammad Kamran[1] · Muhammad Faizan[2] · Ridho Kresna Wattimena[3] · Danial Jahed Armaghani[4] · Panagiotis G. Asteris[5]**

Received: 6 October 2025 / Accepted: 28 January 2026 / Published online: 2 March 2026


## Abstract

Tunnel squeezing is characterized as a significant degree of distortion in the surrounding rock mass that is typically larger than the designed deformation. The squeezing potential of rocks around tunnels can result in support failures, floor heave, and even flood disasters. In this study, the squeezing potential of rocks around tunnels were estimated by employing a hybrid intelligent framework to improve the performance of a classification algorithm. A total of 139 adjacent rock-squeezing patterns were acquired from places such as China, Nepal, and India to form the empirical basis for this study. The data consists of five influential variables, i.e., strength factor, tunnel depth, rock mass quality index, tunnel equivalent diameter and support stiffness. The mechanism of prediction consisted of three steps. Firstly, factor analysis was utilized to reduce the number of influential variables. The resulting factors were then categorized using k-means clustering. Finally, a random forest algorithm was developed to predict various levels of surrounding rock squeezing potential of rocks around tunnels. The proposed hybrid intelligent framework achieved a strong predictive capability of 96%, contributing to safer and more sustainable tunneling practices by reducing operational risks and improving overall structural stability.

**Keywords** Tunnel squeezing · Factor Analysis · k-Means Clustering · Random Forest · Safety

## Nomenclature

### Abbreviations

| | |
|---|---|
| BN | Bayesian Network |
| BP | Back Propagation |

Extended author information available on the last page of the article

| | |
|---|---|
| C5.0 | Algorithm to build either a decision tree or a rule set |
| DT | Decision Tree |
| GWO | Gray Wolf Optimization |
| GWO-IPNN | Gray Wolf Optimization- Improved Probabilistic Neural Network |
| IPNN | Improved Probabilistic Neural Network |
| KNN | K-Nearest Neighbor |
| M-SVM | Multiclass Support Vector Machine |
| SVM | Support Vector Machine |
| SVM-BP | Support Vector Machine Back Propagation |
| SVM-WOA | Support Vector Machine—Whale Optimization Algorithm |
| TBM | Tunnel boring machine |
| WOA | Whale Optimization Algorithm |
| RF | Random Forest |
| ISRM | International Society of Rock Mechanics |
| NS | No Squeezing |
| LS | Light Squeezing |
| FS | Fair Squeezing |
| HS | Heavy Squeezing |
| VHS | Very Heavy Squeezing |
| CV | Coefficient of Variation |
| BQ | Rock Mass Quality Index |
| SF | Strength factor |
| SD | Standard Deviation |
| FA | Factor Analysis |
| PCA | Principle Components Analysis |
| D | The tunnel equivalent diameter |
| H | The tunnel depth |
| K | Support Stiffness |
| CV | Coefficient of variation |

**Symbols**

| | |
|---|---|
| $\varepsilon$ | Developed/computed strain |
| $\varepsilon_{cr}$ | The critical strain of the rock material (experimental value) |
| $\varepsilon_e$ | The experimental value of the elastic tangential strain of the rock material |
| $\varepsilon_f$ | The experimental value of the strain where the flowing state of rock materials take place |
| $\varepsilon_s$ | The experimental value of the strain where the rock material starts to exhibit a strain-softening behavior |
| $\varepsilon_p$ | The experimental value of strain at peak stress of the rock material |
| $D$ | The tunnel equivalent diameter |
| $H$ | The tunnel depth |
| $K$ | Support Stiffness |
| $\omega i$ | Weight of $f$ |
| $\eta_k$ | Eigen value |

| | |
|---|---|
| $N_s$ | The number of the surrounding rock squeezing in the sth cluster |
| $Pr$ | Precision |
| $Re$ | Recall |
| T | Number of tree |
| D | Training dataset |
| $D^{(i)}$ | A bootstrap Sample from D |
| V | Very small set of V |
| $f_1$, $f_2 and f_k$ | The common factors that are unattainable factors |
| $b_{11}, b_{21} and b_{n1}$ | Loading factors |
| $\omega_i$ | Variance contribution rate of the preceding-transformation or subsequent- transformation factors |
| Sd | Surrounding rock squeezing data |
| $K$ | Number of surrounding rock squeezing levels |
| $Z_{aa}$ | Indicator number accurately forecasted for the class $k$ |
| $Z_{ab}$ | The indicator number of the level a that is classified to class b. |
| $\sigma_1$, $\sigma_2$ and $\sigma_n$ | Special factors |
| $y_i$ | Eigenvalue of the correlation matrix R |
| $Y_1, Y_2 and Y_n$ | Observable random factor |
| $\eta_1, \eta_2 and \eta_1$ | Loadings |
| R | Correlation coefficient |
| L | Standard orthogonality |
| $f_i$ | The variance contribution rate of the preceding-transformation or subsequent- transformation factors |

## 1 Introduction

The construction of tunnels involves significant challenges, particularly when a tunnel boring machine (TBM) intersects highly fractured strata as the narrow overtaking maneuver distance between the TBM and the overlying bedrock, may result in the entanglement of the TBM due to the gradual radial convergence of the rock mass into the tunnel (Hoek and Guevara., 2009; Arora et al. 2021). The gradual radial convergence of the rock mass into the tunnel, which results in a significant reduction of the tunnel cross section is termed rock mass squeezing and its intensity depends on the in situ geological conditions (bedding planes, schistosity, joints and faults), groundwater flow and rock mass properties (Zhang et al. 2020). Predicting the magnitude of the radial convergence caused by squeezing is crucial in determining the suitable excavation method and the required supporting structures.

Tunneling through squeezing rock poses substantial engineering challenges that have been comprehensively reported in a variety of case studies. Kimura et al. (1987) examined analogous difficulties faced while tunneling through two significant fault zones, where rock squeezing behavior demonstrated high heterogeneity due to fault-induced instability. Goel et al. (1995) investigated the tunnel in the Himalayas, emphasizing the significant compression challenges faced during tunneling in young, geologically dynamic regions and stressing the importance of predictive models in such contexts. Panet (1996) identified two notable cases where

tunnels passed through squeezing rock, highlighting the complexities of predicting rock deformation and effectively managing these high-stress environments. Sakurai (1997) emphasized the critical insights obtained from field measurements in tunneling, particularly on the management of the dynamic behavior of squeezing rock. Panthi and Nilsen (2007) improved this understanding by conducting an uncertainty analysis on tunnel squeezing for two tunnels, emphasizing the necessity of including geomechanical uncertainties into the design process. Wang. (2020) contended that the critical element for successful tunneling in squeezing ground is the effective alleviation of actual rock pressure, which is crucial for maintaining tunnel stability under extreme stress conditions. Wang et al. (2021) examined deformation in the Tawarazaka Tunnel, highlighting the necessity for efficient technical solutions under squeezing situations. Liu et al. (2021) suggested the time-dependent characteristics of squeezing rock using field data and numerical models to improve forecasts of rock deformation during excavation. Najm and Daraei (2023) examined the challenges associated with rock squeezing in the context of the Middle East's longest highway tunnel, highlighting two primary failure mechanisms that can jeopardize structural integrity. Leone et al. (2024) investigated the impact of creep on shield tunneling, emphasizing stability issues and performance under squeezing conditions. Luo et al. (2024) developed a stress–strain zoning model to categorize substantial deformations, offering insights for risk reduction in squeezing rock situations. Wang et al. (2024) investigated the impact of tunneling on ground collapse, focusing on the mechanisms that contribute to the extent of collapsed areas. Huang et al. (2024) applied Chebyshev Inequalities to assess the risks of collapse failure behind steel structures in tunnels, providing a quantitative approach to understanding deformation and stability in tunnel construction. Hence, it is imperative to accurately evaluate the intensity of rock mass squeezing prior to excavation in order to guarantee the smooth progression of tunnel construction.

Predicting the squeezing intensity of rock masses using advanced constitutive modeling requires a deep understanding of the complex visco-elastoplastic mechanisms that drive the onset and evolution of squeezing, which remain poorly understood and insufficiently calibrated against field data ( Gao et al. 2015). This makes accurate prediction particularly challenging for engineers, as it directly impacts the choice of excavation methods and the stability of support structures (Hoek.and Marinos. 2000; Hoek. 2001; Wu et al. 2020; Liu et al. 2021; Wu et al. 2021). Developing advanced constitutive models for the prediction of the squeezing intensity of the rock mass requires a fundamental understanding of the complex visco-elastoplastic mechanisms underpinning the onset and evolution of the rock mass squeezing intensity, which are generally not yet fully understood and not sufficiently calibrated against field data (Kovári,and Staus 1996). A viable alternative is to utilize the existing measurements and observations of rock mass squeezing intensity recorded in the field to build a database to develop and train advanced non-linear computational models for the prediction of rock squeezing intensity (Zhou et al. 2021; Huang et al. 2022). Advanced soft computing models can be used to convey complex highly non-linear associations between multiple input parameters and may therefore be particularly suitable for the prediction of rock squeezing intensity which is a highly multivariant phenomenon (Hasanipanah et al. 2016; Motahedi et al., 2018; Hajijas-

sani et al., 2020;). To this end, the proposed study aimed at developing advanced soft computing models using a hybrid-intelligent-based model for the prediction of rock squeezing intensity.

Table 1 presents a summary of the various soft computing models developed to predict rock mass squeezing intensity and the prediction accuracy achieved. The prediction accuracy results presented in Table 1 show that supervised classification algorithms have significant limitations in simulating complex mechanical responses such as rock mass squeezing due to the inability to compile a considerable amount of high-quality labelled data. At this point, it is worth stressing that it is inappropriate to compare mathematical simulants based on the value of the index of accuracy. A comparison must refer to an adequate amount of data and indeed to be reliable must be based on the same database.

This limitation may be overcome if classification machine learning techniques are combined with unsupervised algorithms to increase the prediction accuracy of the developed model. In this research, the squeezing potential of rocks around tunnels was predicted by employing a hybrid intelligent framework to improve the performance of a typical classification algorithm.

## 2 Short Review of the Rock Squeezing Phenomenon

### 2.1 Rock Squeezing Phenomenon

The gradual radial convergence of the rock mass into the tunnel, which results in a significant reduction of the tunnel cross section is termed rock mass squeezing. Squeezing of the rock mass occurs when the tangential stresses around the tunnel periphery increase beyond the uniaxial compressive strength of the rock giving rise to time-depen-

**Table 1** Overview of state of supervised machine learning algorithm to forecast surrounding rock squeezing in chronological order

| Author(s) | Approaches | Datasets size | Accuracy (%) |
|---|---|---|---|
| Shafiei et al. (2012) | SVM | 198 | 84.1 |
| Feng and Jimenez (2015) | BNs | 166 | 86.65 |
| Sun et al. (2018) | M-SVM | 117 | 88.13 |
| Azizi et al. (2019) | BN's | 4 | – |
| Ghasemi and Gholizadeh (2019) | C5.0 | 115 | 94 |
| Ghasemi and Gholizadeh (2019) | KNN | 115 | 95 |
| Huang et al. (2020) | SVM-BP' | 178 | 92.11 |
| Chen et al. (2020) | DT | 154 | 93.5 |
| Zhou et al. (2021) | SVM-WOA | 114 | 95 |
| Huang et al. (2022) | GWO-IPNN | 139 | 95 |
| Liang et al. (2024) | Stacking ensemble model* | 162 | 90.54 |
| Ghasemi et al. (2024) | Stacking ensemble model* | 117 | 88.2 |

dent-large-tunnel deformations ( Dalgıç., 2002; Singh et al. 2007; Gao et al. 2015; Tran et al., 2015). Figure 1 presents typical squeezing failures from field case study.

It is worth mentioning that rock mass squeezing typically occurs in great depths where the surrounding material depicts low strength. However, the deformation of the surrounding rock during the squeezing phenomena occurs gradually when the stress condition resulting from excavation exceeds the strength of the surrounding medium (Aydan et al. 1993). Specifically, squeezing of rock is essentially associated with creep caused by exceeding a limiting shear stress defined by International Society of Rock Mechanics (ISRM) (1994) standards. Exceedance of the limiting shear stress induces irreversible creep strains that build up during time at constant and eventually increasing rate giving rise to tertiary creep strains (Sterpi and Goida., 2009).

Muirwood. (1972) was one of the first to acknowledge the importance of the ratio of the uniaxial compressive strength of rock to the overburden stress termed competency index in assessing rock mass stability. Hoek. (2001) suggested that squeezing occurs when the competency index drops below 0.2. The magnitude of the radial convergence caused by squeezing is influenced by the geological conditions (bedding planes, schistosity, joints and faults), groundwater flow and the rock mass properties (Hoek and Marinos 2000). Whilst the squeezing intensity of the rock mass may vary significantly it is broadly classified into the following three failure types, which have been proposed by Aydan et al. (1993), nearly three decades ago:

a. Complete shear failure: The complete shear failure type (Fig. 2a) involves the complete process of shearing of the tunnel circumference and is generally observed in very ductile rock masses or in rock masses along existing widely spaced discontinuities.

**Fig. 1** Typical squeezing failures from field case study (Hoek. 2001)

b. Buckling failure: Buckling failure (Fig. 2b) generally occurs in metamorphic rocks (phylitte, mica-schists) or thinly bedded ductile sedimentary rocks (mudstone, shale, siltstone, sandstone, evaporitic rocks).
c. Shearing and sliding failure: Shearing and sliding failure (Fig. 2c) occurs in generally thickly bedded sedimentary rocks and involves sliding along bedding planes.

The buckling and shearing and sliding failure types generally result in extreme squeezing problems.

### 2.2 Classification and Available Methods for the Squeezing Phenomenon

Based on field measurements, Sakurai and Takeuchi (1983) suggested in their classical work, that tunnel strain levels in excess of approximately 1% may be associated with the onset of tunnel instability, on the other hand Hoek (2001) noted that some tunnels exhibited strains beyond 4% without exhibiting stability problems and suggested that the 1% strain limit proposed by Sakurai and Takeuchi (1983) should be used as a general early indication of tunnel instability without necessitating tunneling support installation. The disparity in the level of stain associated with squeezing instabilities reported in the literature necessitates the definition of squeezing intensity and the suitable strain index required to quantify the squeezing intensity. The general strain evolution of rock mass with increasing stress is shown in Fig. 3 where the stress–strain diagram of the rock material is presented including all the characteristics values of strain.

As the rock mass is progressively strained beyond the elastic region $(\varepsilon_e)$, it enters a hardening state with significant plastic strain accumulation $(\varepsilon_e \leq \varepsilon \leq \varepsilon_p)$ followed by a yielding state after the peak of the stress strain curve is exceeded $(\varepsilon_p \leq \varepsilon \leq \varepsilon_s)$. Yielding of the rock mass results in a significant weakening of the material $(\varepsilon_s \leq \varepsilon \leq \varepsilon_f)$, which ultimately gives rise to a flowing state $(\varepsilon_f \leq \varepsilon)$. Table 2 presents a summary of the various squeezing intensity classifications reported in the literature.

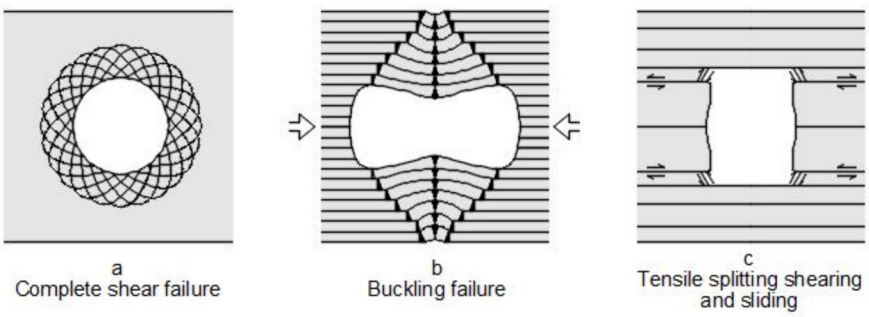


**Fig. 2** Classification of rock squeezing failure modes/patterns in tunnels

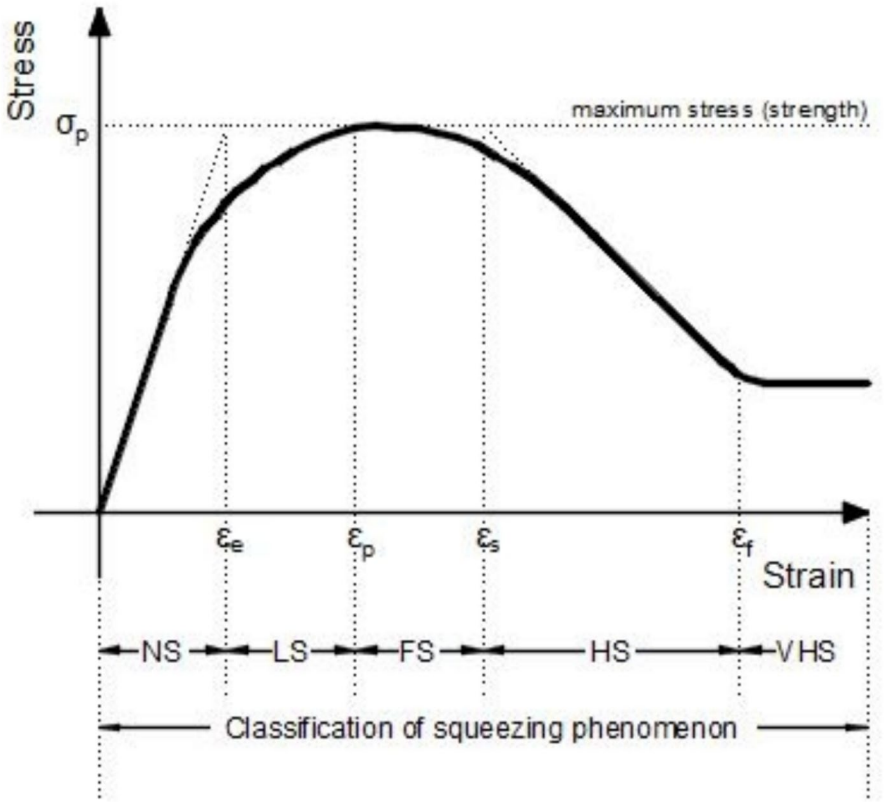


**Fig. 3** Terminology, classification and idealized stress–strain curve of rock materials

**Table 2** Classification of the squeezing phenomenon

| Class | | | Tunnel strain | | | Comments |
|---|---|---|---|---|---|---|
| Nr | Intensity | Symbol | in normalized terms | | in denormalized terms | |
| | | | Aydan et al. (1993) | Singh et al. (2007) | | |
| 1 | No squeezing | NS | $\varepsilon/\varepsilon_e \leq 1$ | $\varepsilon/\varepsilon_{cr} \leq 1$ | $\varepsilon \leq \varepsilon_e$ | elastic response where failure does not take place |
| 2 | Light squeezing | LS | $1 \leq \varepsilon/\varepsilon_e \leq 2$ | $1 \leq \varepsilon/\varepsilon_{cr} \leq 2$ | $\varepsilon_e \leq \varepsilon \leq \varepsilon_p$ | strain-hardening response where starts the development of micro cracks |
| 3 | Fair squeezing | FS | $2 \leq \varepsilon/\varepsilon_e \leq 3$ | $2 \leq \varepsilon/\varepsilon_{cr} \leq 3$ | $\varepsilon_p \leq \varepsilon \leq \varepsilon_s$ | medium rate strain-softening response; moderate depth |
| 4 | Heavy squeezing | HS | $3 \leq \varepsilon/\varepsilon_e \leq 5$ | $3 \leq \varepsilon/\varepsilon_{cr} \leq 5$ | $\varepsilon_s \leq \varepsilon \leq \varepsilon_f$ | high rate strain-softening response; great depth |
| 5 | Very heavy squeezing | VHS | $5 \leq \varepsilon/\varepsilon_e$ | $5 \leq \varepsilon/\varepsilon_{cr}$ | $\varepsilon_f \leq \varepsilon$ | rock flow; great depth |

Whereas, $\varepsilon$ is the developed/computed strain; $\varepsilon_e$ depicts the experimental elastic tangential strain of the rock material and $\varepsilon_{cr}$ defines the critical strain of the rock material (experimental value).

## 3 Material and Methods

The surrounding rock squeezing occurs as a result of ground stress, geological formations, rock mass conditions, and the disturbances effects in excavations (Xu et al., 2020). The following criteria should be met when choosing squeezing prediction variables: (1) the variables should be simple to collect; (2) they should be representative; and (3) they should be able to reflect the squeezing's features from many perspectives. Based on the literature, researchers have employed several surrounding rock squeezing influential characteristics such as strength factor (SF), tunnel depth (H), rock mass quality index (BQ)based on the literature. Besides, support stiffness ($K$) is important in managing surrounding rock deformation (Dwivedi et al. 2013) and the tunnel scale effect, namely tunnel equivalent diameter ($D$) (Goel et al. 1995).

At this point it is worth stressing the very large significance of the database that shall be used for the training and development of the soft computing-based forecasting models (). It is common phenomenon among the researchers that are involved with the development of predicting models to exhibit particular diligence and zeal in the computing method, technique and algorithm that shall be used while at the same moment exhibit less attention for the database that shall be used for the training and development of these mathematical simulants (Lu et al. 2020;). The authors of the present study, without overseeing the significance that the computing method and technique has that shall be used, believe that the reliability of the computing model is crucially dependent on the reliability of the database to be used and the degree of its capability to describe each time studied problem. With the term capable it is meant the database to be comprised from adequate number of data that cover the range of values that takes each one of the parameters infringing in the problem under study.

Adhering to the aforementioned principles, a comprehensive dataset of 139 surrounding rock squeezing patterns was compiled from various regions, including China, Nepal, and India, utilizing freely available data sources (Jiao et al. 2021). This dataset serves as the foundation for the database established in this study. The database includes five influential variables namely $H$, $BQ$, $SSR$, $K$ and $D$. In this study, the classification criteria suggested by (Hoek and Marinos. 2000) was adopted. The non-squeezing (NS) with $\varepsilon < 1\%$ is recorded as 0, minor squeezing with $1\% \leq \varepsilon < 2.5\%$ is recorded as 1 and severe-to-extreme squeezing with $\varepsilon \geq 2.5$ is recorded as 2. The distribution of various surrounding rock squeezing levels is shown in Fig. 4. Table 3 indicates the statistical description of various input variables investigated in this study.

Figure 5 shows a box plot of each variable for the surrounding rock squeezing levels. The surrounding rock squeezing levels are negatively correlated with SF and BQ. The greater the readings of the variables present the higher level of surrounding rock squeezing. Furthermore, certain outliers can be seen in the entire variables of the surrounding rock squeezing dataset for each level. Furthermore, the $H$, $D$, and $K$ values overlap at each level, indicating the complexity of the surrounding rock squeezing phenomenon. As a result, the purpose of this study is to increase confidence in the surrounding rock-squeezing database by considering the effect of all relevant variables.

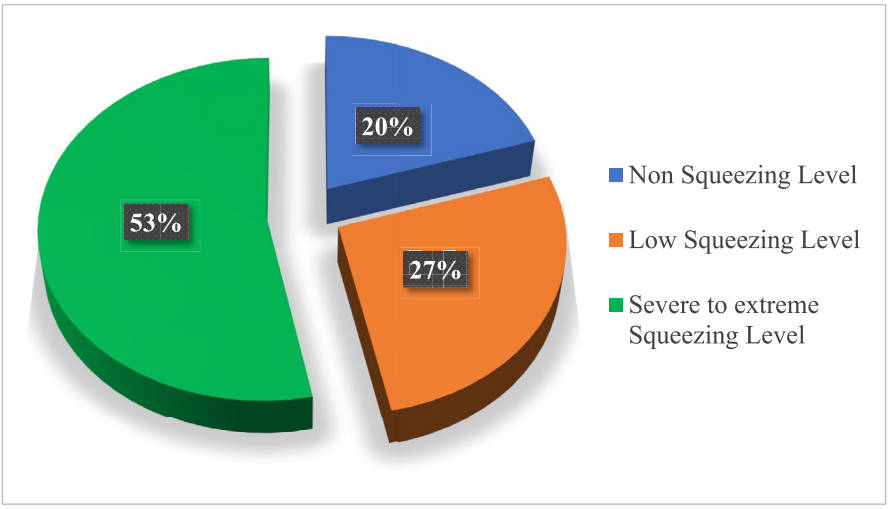


**Fig. 4** Distribution of various surrounding rock squeezing levels

**Table 3** Statistical description of the compiled rock squeezing database

| Variables | | | | Statistical Indices | | | | |
|---|---|---|---|---|---|---|---|---|
| Name | Symbol | Units | Category | Min | Max | Mean | SD | CV |
| Strength factor | SF | - | Input | 0.008 | 1.55 | 0.31 | 0.27 | 0.87 |
| Tunnel depth | $H$ | m | Input | 45.1 | 1110 | 361.38 | 214.77 | 0.59 |
| Rock mass quality index | BQ | - | Input | 70 | 441 | 216.07 | 72.58 | 0.33 |
| Tunnel equivalent diameter | $D$ | m | Input | 4 | 14 | 9.74 | 2.92 | 0.29 |
| Support stiffness | $K$ | MPa | Input | 2.98 | 1979.56 | 1110.35 | 482.01 | 0.43 |
| Squeezing classification | | - | Output | - | - | - | - | - |

SD the Standard Deviation; CV the Coefficient of variation.

## 4 Methodology

### 4.1 Factor Analysis (FA)

One of the statistical procedures that is commonly used for preprocessing in attempt to reduce the dimensionality of the data is known as factor analysis (FA). FA is a complex and multifaceted strategy for condensing a large quantity of data into a relatively small number of factors by incorporating preliminary parameters into strong connotations (Liu et al. 2018). Factor in factor analysis presents a wide variety of parameters for whom cross-correlation assists in lowering absolute divergence. While adopting the rotation process, each factor is considered as a possible explanation for the link between the two group variables.

The principle components analysis (PCA), which is employed in the majority of factor analysis methodologies, transforms a group of associated input parameters into such a random variables set. PCA is a rearrangement of the data's characteristics. The primary component is designated as the orientation in which all the broadest

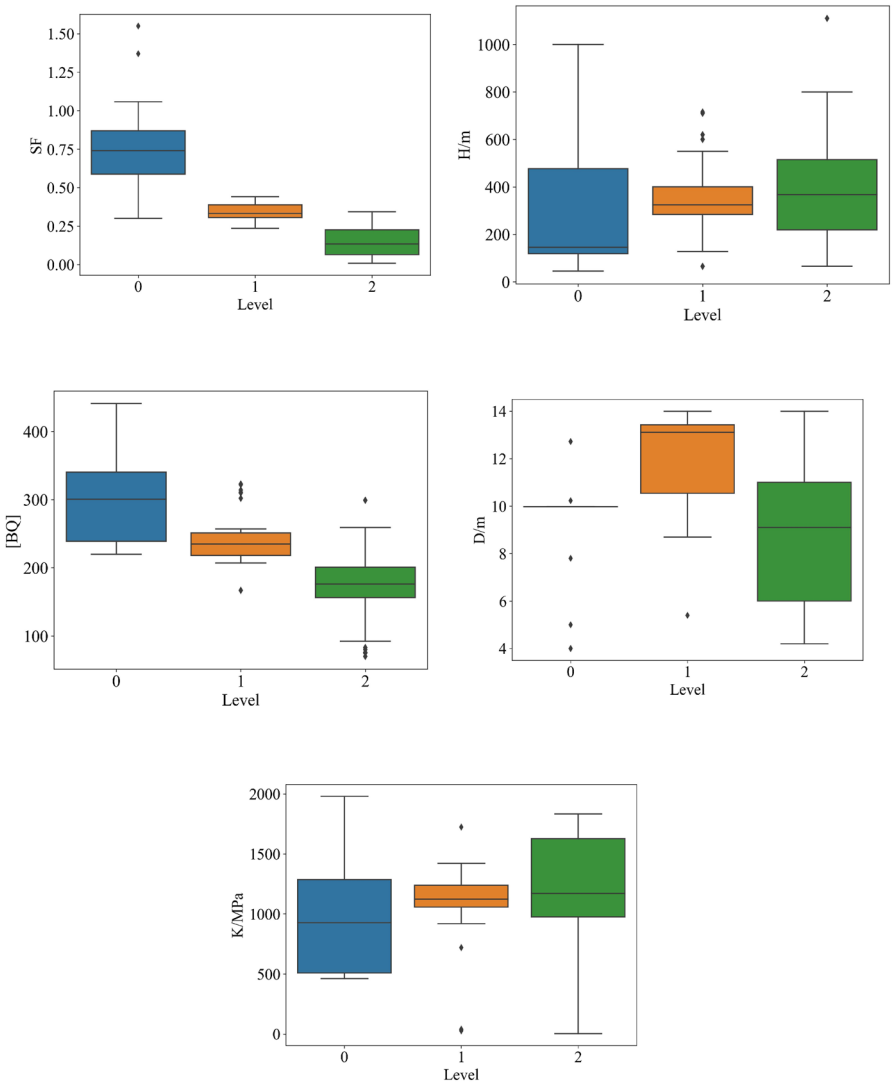


**Fig. 5** Boxplot of each influencing variables to the associated surrounding rock squeezing level

distribution among elements is observed, as well as the elements are portrayed into system of measurement all along axis, leading to the loss of complete details about the variations between both the elements from each other as well (Yun et al. 2014). The factors rotation procedure is applied in factor analysis that helps in describing the underlying fundamental significance of the principal components and even the majority of the total distribution variation. The preceding are the stages of the factor analysis technique: The entire set of data is rotated counter—clockwise, with a first axis constitute the largest dispersion and the other axes reflecting the rest. The data

closest to the source of coordinates is frequently reduced and has the potential to be reduced as a direct result of the data set being rotated.

FA has the potential to perceive underlying features, as well as the ease and efficiency with which it can be performed. The effective implementation of the procedure is reliant on the researchers' competency to provide such a comprehensive description of the entity. If the attributes in the given dataset have a similar format, the FA would then designate one factor to them, concealing the factors of a particular value to the research, and so on. Figure 6 depicts the mechanism of FA.

### 4.2 k-Means Clustering

To prevent artificial division and supervision, clustering analysis has proven to be the best option. Clustering groups, a dataset by a comparable number and keeps the groups with the highest similarity together. The dataset is divided into sections based on the intervals between the instances. In addition, the criteria of homogeneity and heteroge-

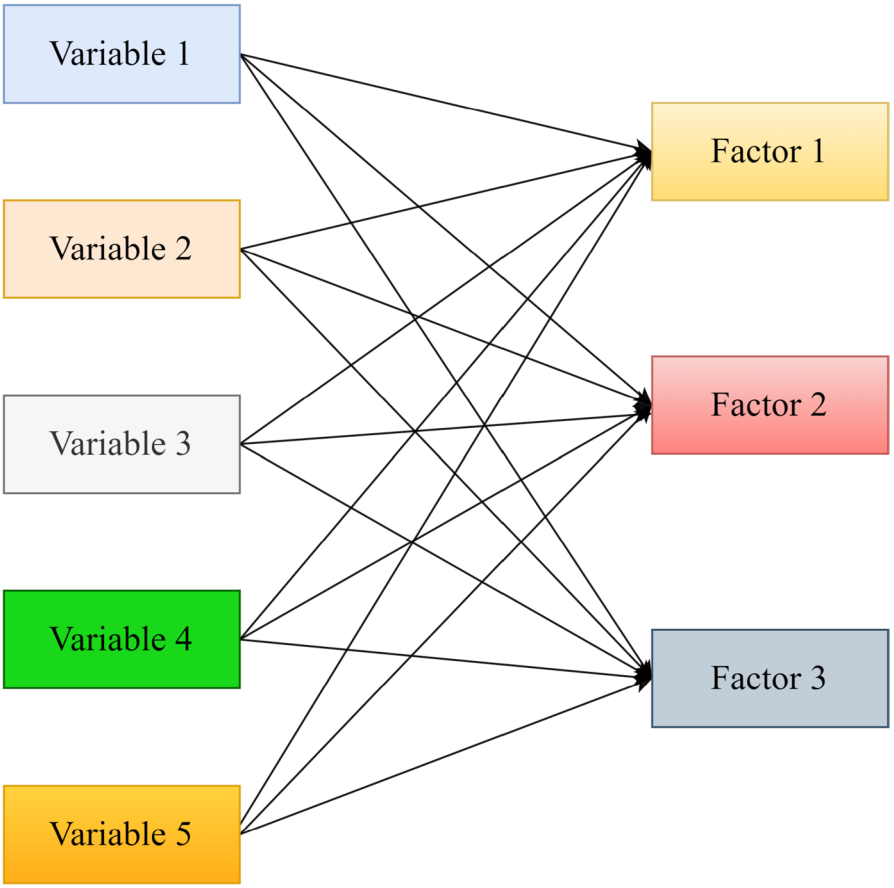


**Fig. 6** Detail mechanism of FA

neity play a significant part in the data division process. The k-Means clustering (Kamran et al., 2022a) divides n observations into k clusters, which offers a unique array of possible applications. Each observation is assigned to the cluster that has the meaning that is closest to it when using the k-Means clustering mechanism. This algorithm's underlying logic may be broken down into two independent components that each perform their own function. The first step involves selecting the k centers at random using a number for k that has previously been calculated, and the second step involves collecting all of the data objects that are located in the vicinity of the center that is the closest to them (Zhu et al. 2019). The most common grouping criterion is the sum of the squared Euclidean distances between the points in an issue. The distance that separates an instance from the geographic center of a cluster is the primary focus of this criterion. The flow chart of the k-Means clustering algorithm is illustrated in Fig. 7.

### 4.3 Random Forest (RF)

Random Forest (RF) is a machine learning technique that employs various methods to generate an ensemble of decision trees for outcome prediction. Initially, RF was introduced as a bagging ensemble method aimed at improving the accuracy of attribute selection (Altman et al., 2017). Breiman (2001) noted that RFs constructed using random vectors during tree formation behave similarly to a kernel applied to

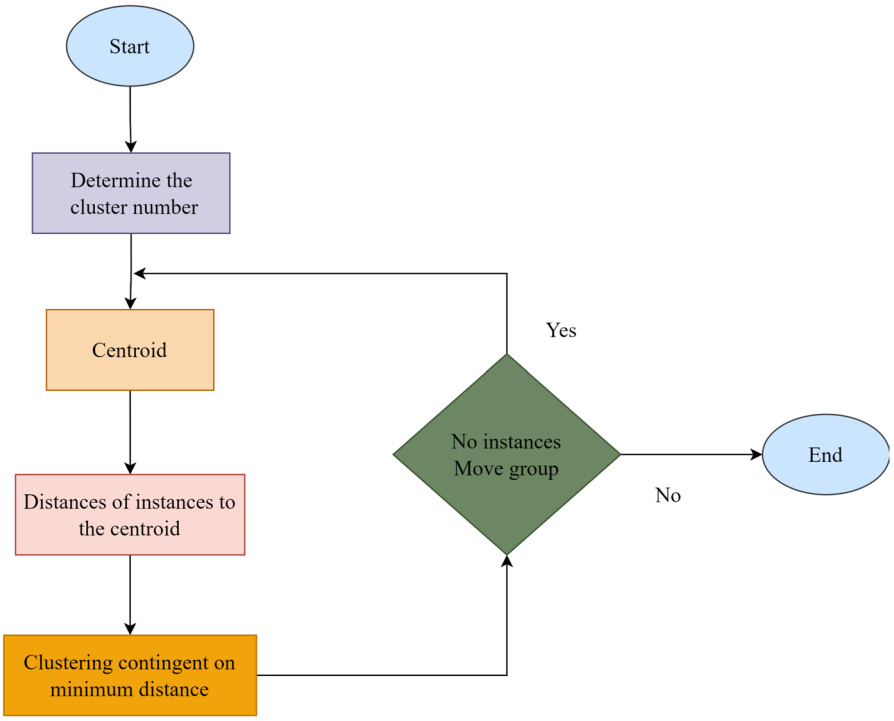


**Fig. 7** Flowchart of k-Means Algorithm

the actual margin. This approach leverages two key principles of randomization—bagging and random feature selection (Breiman, 2001). Each decision tree in the RF is built by utilizing a different subset of input variables during its construction. Because of this, the RF framework is superior to the decision tree in general because it includes the over-fitting properties of the DT into its training dataset. To put it another way, the RF surpasses the decision tree. Breiman (2001), in his foundational work on the Random Forest (RF) framework, demonstrated that beyond a certain threshold, increasing the number of trees does not enhance the model's performance. This finding highlights that an excessively large number of trees is not necessary to achieve high efficiency in RF; instead, a relatively small number of trees is enough. Figure 8 shows the detailed overview of RF algorithm. Table 4 exhibits the RF algorithm.

## 5 Results and Discussion

### 5.1 Surrounding Rock Squeezing Database Preprocessing using FA

Consider n is the number of related random variables in surrounding rock squeezing database, that has k independent factors. Equation (1) represents the FA mechanism.

$$\begin{cases} Y_1 = b_{11}f_1 + b_{12}f_2 + b_{13}f_3 + \sigma_1 \\ Y_2 = b_{21}f_1 + b_{22}f_2 + b_{23}f_3 + \sigma_2 \\ Y_n = b_{n1}f_n + b_{n2}f_n + b_{n3}f_n + \sigma_n \end{cases} \tag{1}$$

whereas $f_1$, $f_2 and f_k$ indicates the common factors that are unattainable factors. The coefficients $b_{11}, b_{21} and b_{n1}$ are loading factors, and $\sigma_1, \sigma_2 and \sigma_n$ are special factors and cannot be included in public factors. The mechanism of FA is depicted in the seven steps below.

Step 1: The initial features should be normalized in order to remove the variations in variable magnitude and size as shown in Eq. (2).

$$y_i = \frac{y_i - E(y_i)}{\sqrt{Var(y_i)}} \tag{2}$$

Step 2: Compute the normalized data correlation matrix. The correlation coefficient r has a value of between -1 and 1. Where r from -1 to 1 shows the correlation between the features becomes increasingly strong.

Step 3: Eq. (3) and (4) should be applied to compute the variance contribution and cumulative variance contribution of the public factor.

$$\frac{\eta_k}{\sum_{n=1}^{p} \eta_n} (k = 1,2,3,\ldots\ldots.p) \tag{3}$$

$$\frac{\sum_{n=1}^{k} \eta_n}{\sum_{n=1}^{p} \eta_n} (k = 1,2,3,\ldots\ldots.p) \tag{4}$$

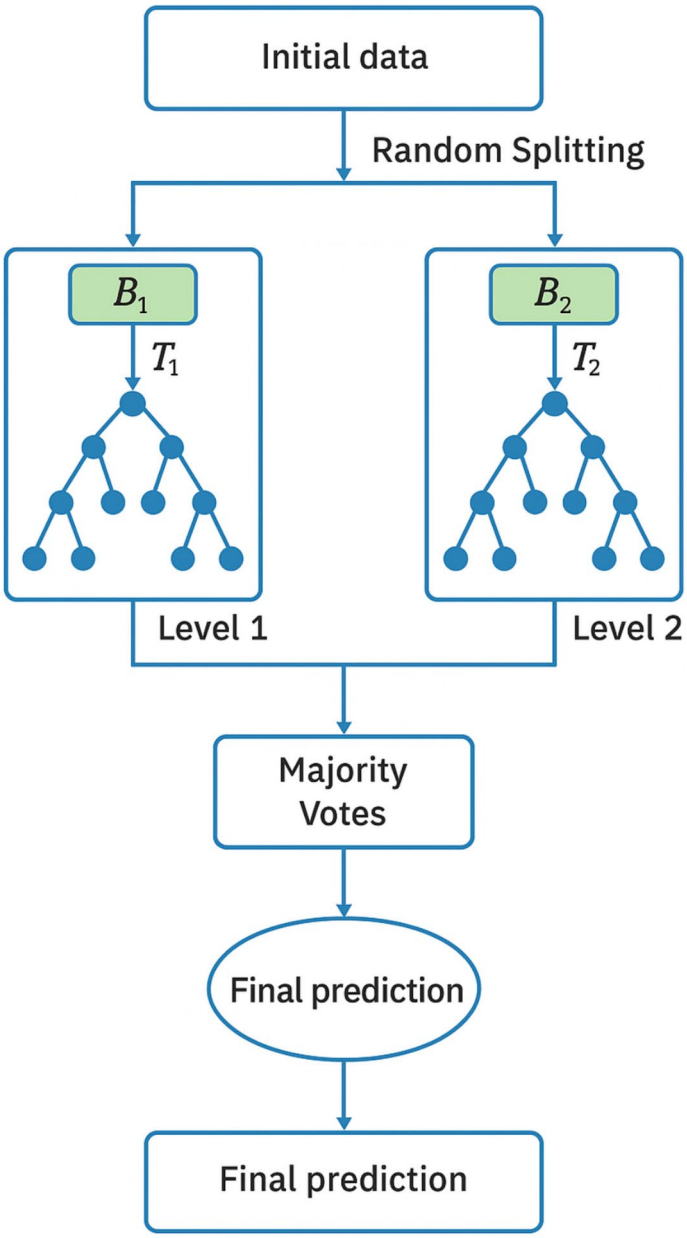


**Fig. 8** Overview of RF algorithm

**Table 4** Algorithm of Random Forest

| Input | Random Forest |
|---|---|
| Step 1 | Prerequisite: Training dataset D:=($x_1$, $y_1$), ($x_2$,$y_2$).…….($x_n$, $y_n$), Variables V and number of tree in forest T |
| Step 2 | Function Random Forest (D, V) |
| Step 3 | P←Ø |
| Step 4 | **for** I∈1,2,3… T do |
| Step 5 | $D^{(i)}$←A bootstrap Sample from D |
| Step 6 | Pi←RandomizedTreeLearn($D^{(i)}$, V) |
| Step 7 | Pi←P∪{$p_i$} |
| Step 8 | End for |
| Step 9 | Return H |
| Step 10 | End function |
| Step 11 | Function RandomizedTreeLearn (D, V) |
| Step 12 | At each node |
| Step 13 | v←very small set of V |
| Step 14 | Divide on best variable in v |
| Step 15 | Return The Learn Tree |
| Step 16 | End function |

where $\eta_k$ represent an eigenvalue.

Step 4: Choose the factors that are common. Adjust the factors to $f_1$, $f_2 and f_k$, which comprise the total quantity of data information. If the first k factors' cumulative data information approaches or exceeds 80%, they are considered to indicate the real assessment index.

Step 5: If the first factors are unable to resolve or if the factual concept of the factors is ambiguous, the factors need to be transformed to acquire a clearer pragmatic concept. Equation (5) depicts the specific transformation.

$$b_{kl} = \sqrt{\eta_k m_{kl}}\,(\mathrm{k},\ \mathrm{l}\ = 1, 2, 3, \ldots..\mathrm{p}) \tag{5}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \cdots \\ b_{p1} & b_{p2} & \cdots & b_{pk} \end{bmatrix} \tag{6}$$

$$\begin{aligned} &= \begin{bmatrix} \sqrt{\eta_1} m_{11} & \sqrt{\eta_2} m_{12} & \cdots & \sqrt{\eta_k} m_{1\mathrm{k}} \\ \sqrt{\eta_2} m_{21} & \sqrt{\eta_2} m_{22} & \cdots & \sqrt{\eta_k} m_{2\mathrm{k}} \\ \vdots & \vdots & \ddots & \cdots \\ \sqrt{\eta_1} m_{\mathrm{p1}} & \sqrt{\eta_2} m_{\mathrm{p2}} & \cdots & \sqrt{\eta_k} m_{\mathrm{pk}} \end{bmatrix} \\ &= \left(\sqrt{\eta_1} m_1, \sqrt{\eta_2} m_2, \ldots\ldots\ldots, \sqrt{\eta_k} m_{\mathrm{k}}\right) \end{aligned} \tag{7}$$

Step 6: Individual factor scores should be computed using a linear combination of the actual variables. The regression assessment should be carried out to measure the factor score.

Step 7: Comprehensive Outcomes; As can be seen in Eq. (8), the detail assessment index function is a linear composition of each factor.

$$f = \frac{\gamma_1 f_1 + \gamma_2 f_2 + \cdots + \gamma_k f_k}{\gamma_1 + \gamma_2 + \gamma_3} = \sum_{i=1}^{m} \omega_i f_i \tag{8}$$

where $\omega_i$ shows the weight of $f$ and $f_i$ indicates the variance contribution rate of the preceding-transformation or subsequent- transformation factors as depicted in Eq. (8).

The correlation heatmap between five predicting variables is shown in Fig. 9. It has been shown that the coefficient of correlation between SF and BQ is greater than 0.5, showing that there is a significant relationship. Because the framework's operating efficiency will be affected by this data redundancy, so factor analysis has been used to remove the relationship between the variables. We initially utilize Eq. 7 to normalize the original data in order to eliminate the impact of variable dimensions on the features extraction procedure.

The percentage of variance attributed to each of the selected factors is the explained variance ratio. In order to pick the number of factors to include in our framework, we have summed each factor's explained variance ratio till it reaches a total of roughly 0.9 or 90 percent. Equation 2 has been applied in the study to normalize the initial

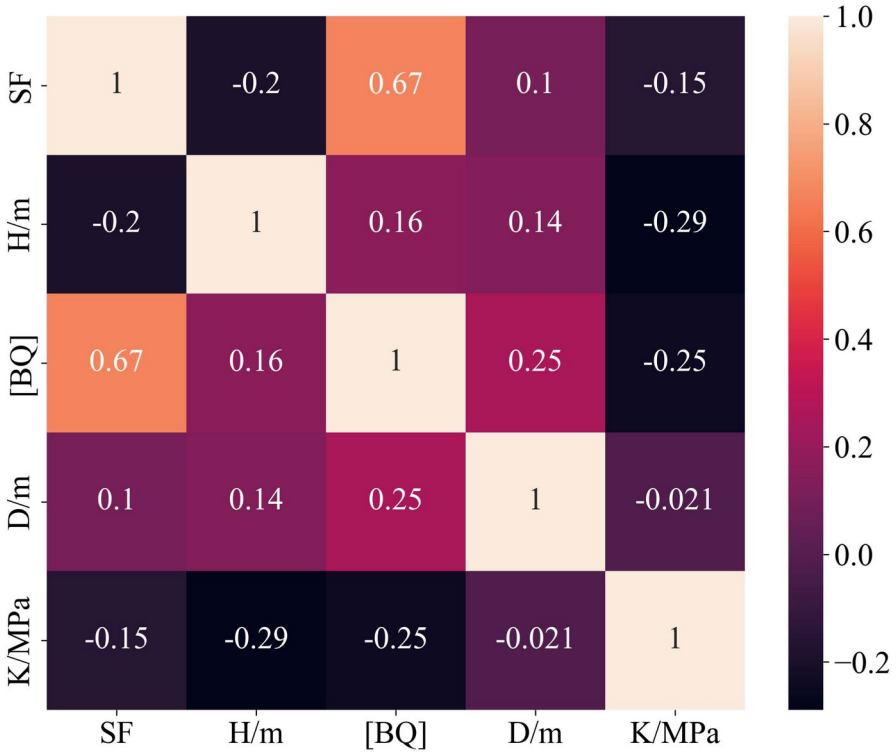


**Fig. 9** Correlation heatmap of initial surrounding rock squeezing dataset

dataset. The relationship between the explained variance ratio and the factor analysis index is depicted in Fig. 10. According to Fig. 10, the cumulative contribution rate of the first three factors is 97 percent, suggesting that the first three factors include 97 percent of the original data, satisfying the factor analysis selection criteria. As a result, the three factors are chosen to replace the five original prediction variables. In order to complete the FA, a Jupyter notebook was utilized to call the Scikit-learn module. Figure 11 illustrates the 3D visualization plot of FA based surrounding rock squeezing data.

### 5.2 k-Means Clustering on FA Based Surrounding Rock Squeezing Database

The k-means algorithm is a well-known data categorization approach that has been broadly utilized in data mining and knowledge discovery. The surrounding rock squeezing data set $Sd=(Sd_1, Sd_2,\ldots, Sd_n)$ is partitioned into T subsets $D=(D_1, D_2,\ldots, D_T)$ using this approach. The aggregating findings should adhere to the following guidelines:

$$\begin{cases} Sd = \cup_{i=1}^{T} D_i, D_i \neq \varnothing\ (i = 1,2,3,\ldots\ldots\ldots,T) \\ D_i \cap D_j \neq \varnothing (i,j = 1,2,3,\ldots\ldots\ldots T; i \neq j) \end{cases}$$

The essential methods are included in a standard k-Means clustering approach for categorizing surrounding rock squeezing.

1. Distribute the surrounding rock squeezing to *K* clusters at random.

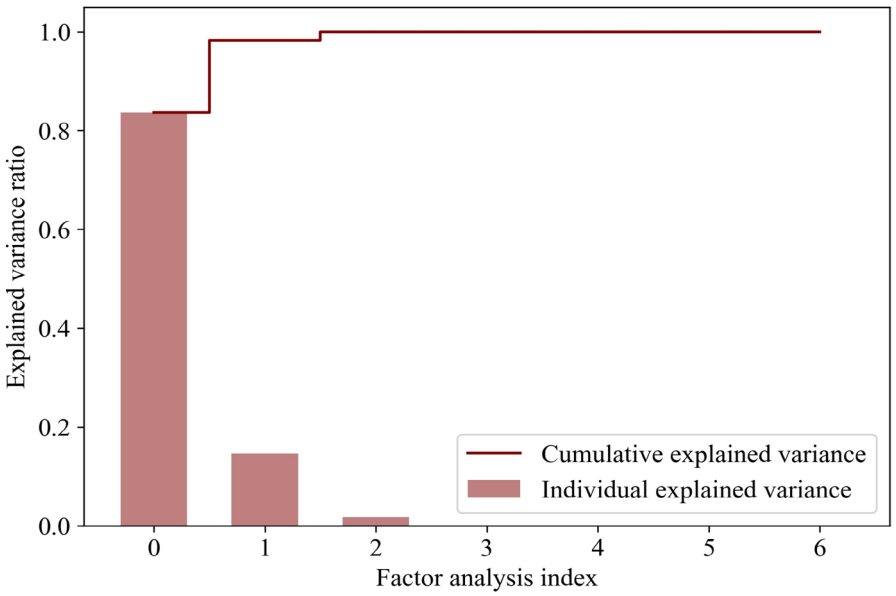


**Fig. 10** Relationship between the explained variance ratio and the factor analysis index

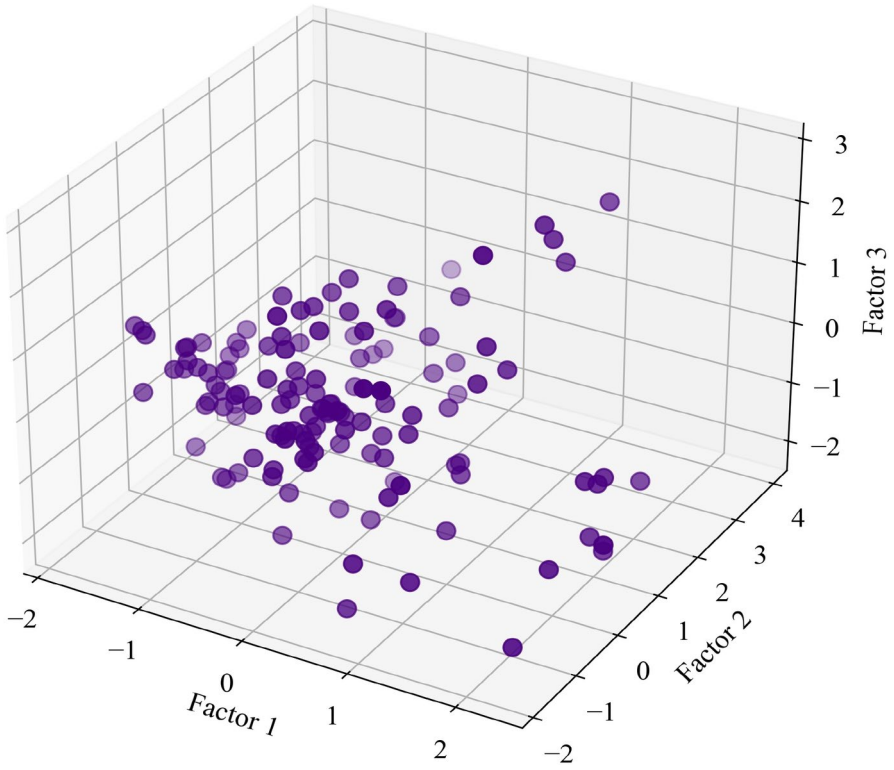


**Fig. 11** 3D visualization plot of FA based surrounding rock squeezing data

2. Using the preliminary clustering result, compute the center $C_j$ of each cluster: where $n_j$ is the number of surrounding rock squeezing in the *jth* cluster.

$$K = \frac{1}{N_S} \sum_{E_i \in D_j} E_j \tag{9}$$

where $N_S$ is the number of the surrounding rock squeezing in the $s_{th}$ cluster.

3. Calculate the distance between each surrounding rock squeezing and the centers of *k* clusters and assign each discontinuity to the cluster to which it is closest.
4. Using Eq. (9), update the center of each cluster.
5. Find the cost function by multiplying the distance between the $i_{th}$ surrounding rock squeezing and the $j_{th}$ cluster center by $D(E_i, C_j)$ is depicted as Eq. (10).

$$F = \sum_{S=1}^{k} \sum_{E_i \in D_j} D(E_i, C_j) \tag{10}$$

6. To minimize the cost function, repeat steps (3), (4), and (5).
7. When no surrounding rock squeezing can be redistributed to various clusters, the iteration process will terminate.

In order to perform k-Means clustering, a Jupyter notebook has been used with the Scikit-learn module. The categorization of cluster monitoring has been demonstrated by Kamran et al., (2024b). The closeness and separation of the elements must therefore be balanced for the silhouette mechanism to operate. The silhouette coefficient can be used to demonstrate that the FA data have been properly grouped, indicating that the elements are organized into the categories to which they correspond. The silhouette coefficient is a metric for assessing the clustering identification which has been used to identify the proper k in the clusters. In the k-Means clustering stage, the optimal number of clusters was set to three (k=3), aligning with the three observed intensity levels: non-squeezing, minor, and severe-to-extreme squeezing. This clustering was performed on the standardized dataset to ensure that the partition was based on the underlying structural patterns of the tunnel squeezing data. This study computed numerous steps, as illustrated in Fig. 12. A silhouette coefficient has been employed as an index for evaluating k-Means clustering (Kamran et al., 2023a). After the 10th iteration in the FA derived surrounding rock squeezing dataset, the maximum silhouette coefficient of 0.57 was attained in this study.

The precision, recall, and $F_1$-score metrics have been used to assess the classification accuracy of machine learning algorithms (Kamran et al. 2022b). Precision can accurately anticipate datasets; recall interprets the ability to accurately predict actual features to the highest degree; and the $F_1$-score provides a universal metric that incorporates both recall and precision performance. As a result, the aforementioned evaluation indices are used to measure the framework's execution. Assume that Eq. (11) defines the confusion matrix (10). A confusion matrix is commonly used as a benchmark for demonstrating the evaluation of a classification framework on a testing instance with known true values.

$$Z = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1k} \\ Z_{21} & Z_{22} & \cdots & Z_{2k} \\ \vdots & \vdots & \ddots & \cdots \\ Z_{b1} & Z_{b2} & \cdots & Z_{bk} \end{bmatrix} \tag{11}$$

whereas $k$ depicts the number of surrounding rock squeezing levels, $Z_{aa}$ is variables number accurately forecasted for the class $k$, and $Z_{ab}$ depicts the variables number of the level $a$ that is classified to class $b$.

Equations (12) to (14) are used to calculate the precision, recall, and $F_1$-score measure for each surrounding rock squeezing level based on the confusion matrix.

$$Pr = \frac{Z_{aa}}{\sum_{a=1}^{E} Z_{ab}} \tag{12}$$

$$Re = \frac{Z_{aa}}{\sum_{a=1}^{E} Z_{ab}} \tag{13}$$

$$F_1 - score = \frac{2 * Pr * Re}{Pr + Re} \tag{14}$$

To further investigate the performance of RF, the accuracy is represented by Eq. (15)

$$Accuracy = \frac{1}{\sum_{a=1}^{E} \sum_{b=1}^{k} Z_{ab}} \sum_{a=1}^{E} Z_{aa} \tag{15}$$

### 5.3 Performance Evaluation of the Proposed Frameworkusing RF

To ensure a representative distribution of squeezing categories, a stratified train-test split was employed, allocating 80% of the dataset for model training and the remaining 20% for model testing. On the testing dataset, the prediction results of RF and

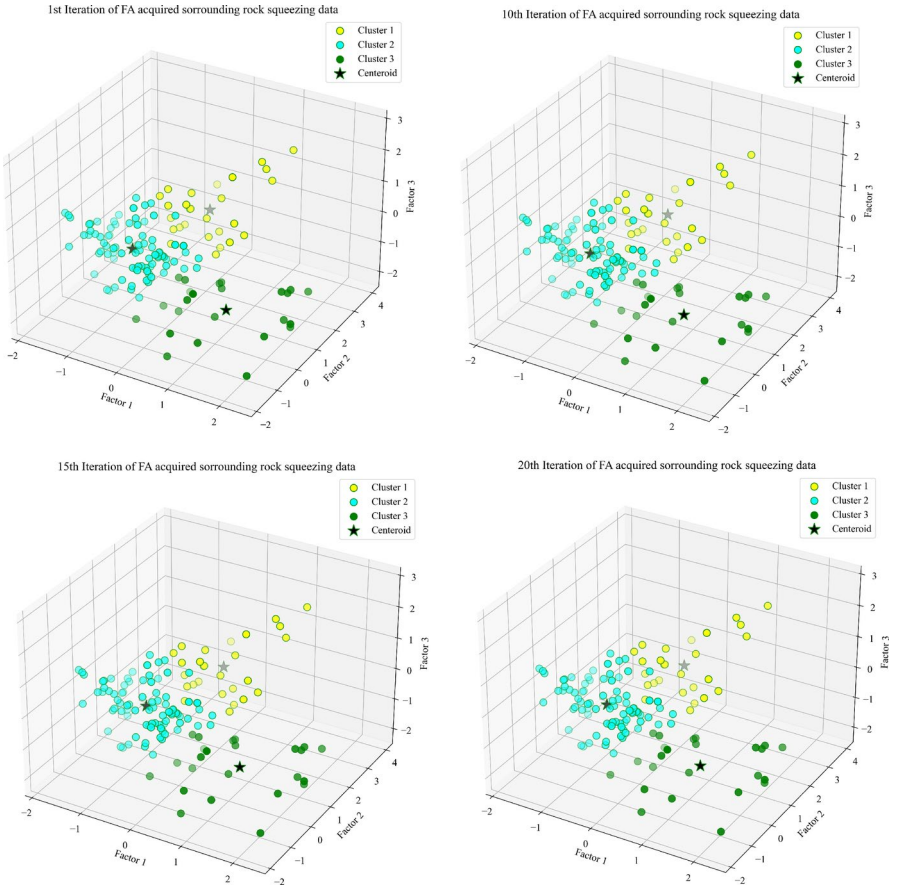


**Fig. 12** k-Means Clustering visualization applied to the FA-derived database for rock-squeezing classification

hybrid model (FA + k-Means Clustering and RF algorithms were obtained. A stratified fivefold cross validation has been carried out on the RF and hybrid model. Several evaluation matrices were used in this work to estimate the outcomes of both the RF and hybrid models. The testing dataset classification report was generated using the Python programming language. As demonstrated in Tables 5 and 6, the classification report provide an overview of the RF and the hybrid models suggested framework's performance on the surrounding rock squeezing intensity level dataset. Equation (12) was used to calculate precision values.

The RF classification results indicate that the proposed framework achieves an overall accuracy of 71% in predicting the rock-squeezing phenomenon. The RF model (Table 5) performs particularly well in the Severe-to-Extreme Squeezing category, attaining a high recall of 0.93 and an F1-score of 0.87. In contrast, the Non-Squeezing category shows a perfect precision of 1.00 but a notably low recall of 0.33, suggesting that many non-squeezing cases were misclassified. These findings highlight the need for further refinement of the hybrid model to improve its ability to distinguish rock-squeezing conditions more effectively.

When compared to minor squeezing intensity level, the precision value for non-squeezing and severe-to-extreme squeezing levels produced better results in the hybrid model (as shown in Table 6). For the non-squeezing level, the accuracy value was 100, the precision value for the minor squeezing level was 96, and the precision value for the severe-to-extreme squeezing level was 100. The recall value for each surrounding rock squeezing intensity level was calculated using Eq. (13). Based on recall value, it was evident that minor squeezing and severe-to-extreme intensity levels performed better than non-squeezing intensity levels. The Recall values of 75, 100, and 100 for non-squeezing, minor squeezing, and severe-to-extreme squeezing levels were obtained respectively.

Equation (14) was used in this study to calculate the F1-score for each corresponding surrounding rock squeezing intensity level. The F1-score for severe-to-extreme squeezing level outperformed non-squeezing and minor squeezing intensity level. The F1-score values for non-squeezing, minor squeezing, and severe-to-extreme

**Table 5** Classification report of random forest algorithm based on rock squeezing data

| Surrounding rock squeezing Intensity level | Random Forest Algorithm | | |
|---|---|---|---|
| | Precision | Recall | $F_1$-score |
| Non-squeezing | 100 | 33 | 50 |
| Minor squeezing | 44 | 57 | 50 |
| Severe-to-extreme squeezing | 82 | 93 | 87 |

**Table 6** Classification report of hybrid FA–k-Means–RF model based on rock squeezing data

| Surrounding rock squeezing Intensity level | Hybrid FA–k-Means–RF model | | |
|---|---|---|---|
| | Precision | Recall | $F_1$-score |
| Non-squeezing | 100 | 75 | 86 |
| Minor squeezing | 96 | 100 | 98 |
| Severe-to-extreme squeezing | 100 | 100 | 100 |

squeezing levels were 86, 98 and 100 respectively. Equation 15 was used in this study to determine the entire accuracy of the proposed model on the testing dataset. The hybrid model integrating FA and k-Means clustering with the RF algorithm achieved a predictive accuracy of 96% on the testing dataset. This performance represents a significant improvement over the standalone RF model and exceeds the benchmarking accuracies reported in the existing literature (Table 1), demonstrating the framework's superior capability in predicting surrounding rock squeezing intensity levels.

The accuracy of the hybrid intelligent frameworkis assessed as a whole, with recall and precision determined separately for each surrounding rock squeezing intensity level. For our framework, we use a macro average and weighted average of precision, recall, and F1-score for the surrounding rock squeezing phenomenon.

The scores that make up the macro-average are the arithmetic means of all of the rock-squeezing intensity levels in the surrounding area. As a consequence of this, the macro-average precision is determined by taking the arithmetic mean of the three separate degrees of precision that are related with the three distinct levels of rock squeezing in the surrounding area. The macro-average recall illustrates the three different levels of recall for the rock squeezing that occurs in the surrounding area. While the macro-average F1-score is the arithmetic mean of the F1-scores of three distinct levels of rock squeezing in the surrounding area. As a result, the overall average for precision, recall, and F1-score was respectively 99, 92, and 95 for the hybrid intelligent framework. The cumulative total of all levels' scores is equal to the weighted average score, which is calculated by multiplying each level's proportional weight by itself. Therefore, the weighted averages of precision, recall, and F1-score were 97, 96, and 96 of the proposed hybrid intelligent framework, respectively, for precision, recall, and F1-score.

In addition, as illustrated in Figs. 13 and 14, confusion matrices for the RF and hybrid framework were created. The samples number successfully predicted by the confusion matrix is shown by the values on the main diagonal. Most of the surrounding rock squeezing samples were correctly classified using the hybrid intelligent framework as compared to RF. Only six levels were mispreidcted by RF algorithm, whereas the hybrid intelligent framework misclassified just one surrounding rock squeezing level the entire database. According to the findings, the hybrid intelligent framework performed well in predicting the intensity levels of surrounding rock squeezing. Figure 15 further illustrates the ROC–AUC results, demonstrating the hybrid intelligent framework's strong capability, with AUC values of 1.00 for Level 0, 0.99 for Level 1, and 0.98 for Level 2. The steep ascent of each ROC curve toward the top-left corner confirms that the hybrid intelligent framework consistently maintains high true-positive rates while effectively minimizing false positives across all squeezing levels.

While FA effectively simplifies the input variable space by grouping correlated features into latent factors, k-Means clustering is subsequently applied to categorize the observations themselves based on these factors. This two-step unsupervised approach ensures that the RF model is trained on distinct, geomechanically consistent data clusters, thereby enhancing the model's ability to generalize squeezing potential across different tunnelling.

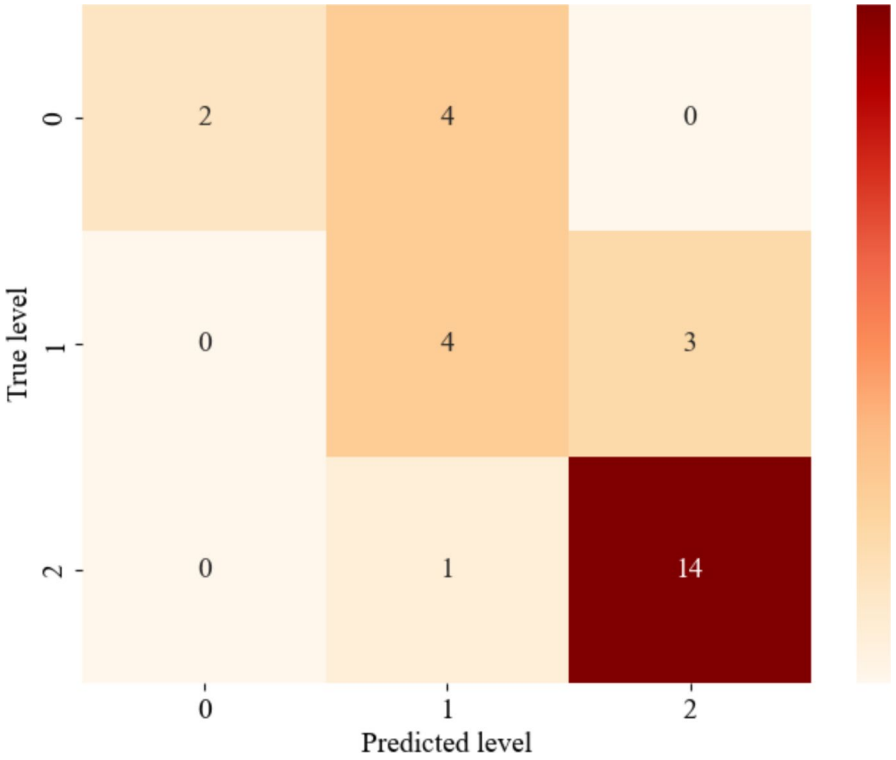


**Fig. 13** Confusion matrix of RF algorithm across the three squeezing levels

Understanding the risks associated with tunneling is essential for planning the excavation method and constructing the infrastructure needed to support underground entrances. It is vital that rock engineers understand the many types of squeezing circumstances in order to design fully functional yielding supports. There are already a wide variety of qualitative and quantitative methods available for evaluating squeezing conditions. In the current work, a hybrid intelligent system was used to determine the rock-squeezing levels in the vicinity, which improved the performance of a classification method. The results of the outcome measures reveal that the suggested methodology offers a high probability of detecting tunnels being squeezed by surrounding rocks, allowing for minimization the risk of mining operations.

## 6 Limitations and Future Works

To the best information of the authors, the database utilized in the present study is among the largest database regarding the classification of the rock materials squeezing phenomenon. However, we would like to point out certain limitations that govern the optimally developed model in our work. More specifically, that the reliability of the proposed model is valid for the values of parameters between the smaller and largest value for each one of the parameters of the database. (input parameters) (Table 3).

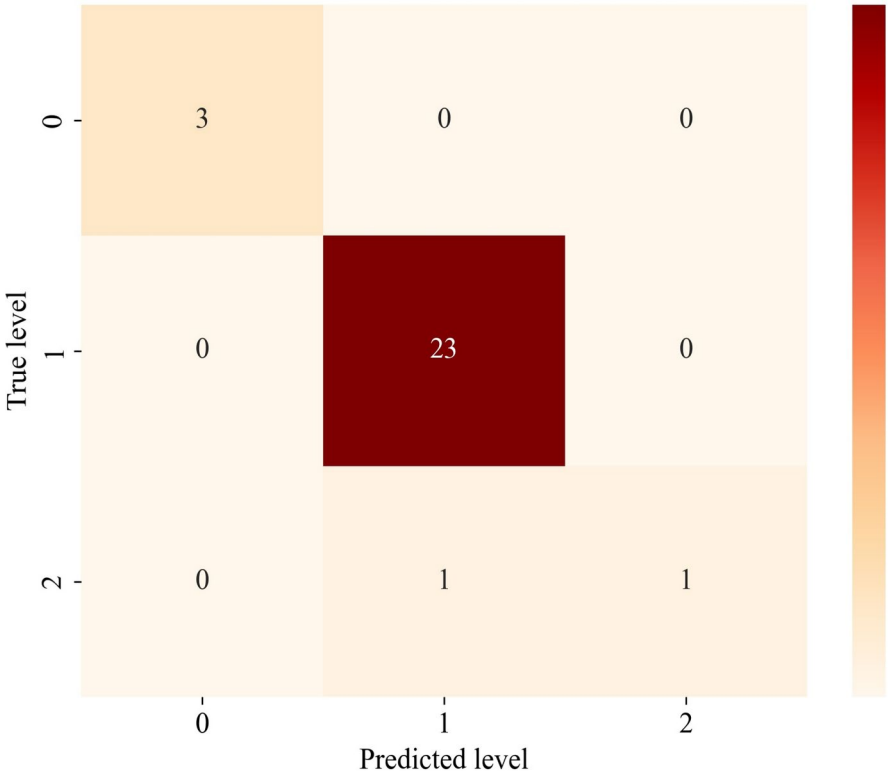


**Fig. 14** Confusion matrix of hybrid FA–k-Means–RF model across the three squeezing levels

In addition, the limitation of the proposed hybrid framework is that the application of FA transforms the raw physical variables into a lower-dimensional latent space. Consequently, while this improves the clustering performance and reduces multicollinearity, it prevents the execution of traditional feature importance or sensitivity analysis on the original parameters. Future research could explore local surrogate models to bridge this gap in interpretability.

Besides, based on the great interest of the researchers involved with the present problem as well the field engineers the immediate intentions of the authors regarding future work, is the updating of the database (Fig. 5).

## 7 Conclusion

The primary objective of this study is to explore the feasibility of utilizing a hybrid intelligent system to accurately assess the squeezing potential of rocks around tunnels. This assessment will enable informed decisions to be made on the planning of support requirements prior to the commencement of mining operations. Initially, a total of 139 patterns were compiled from tunnels using simple, convenient, and widely available input variables. Several visual representations were generated to

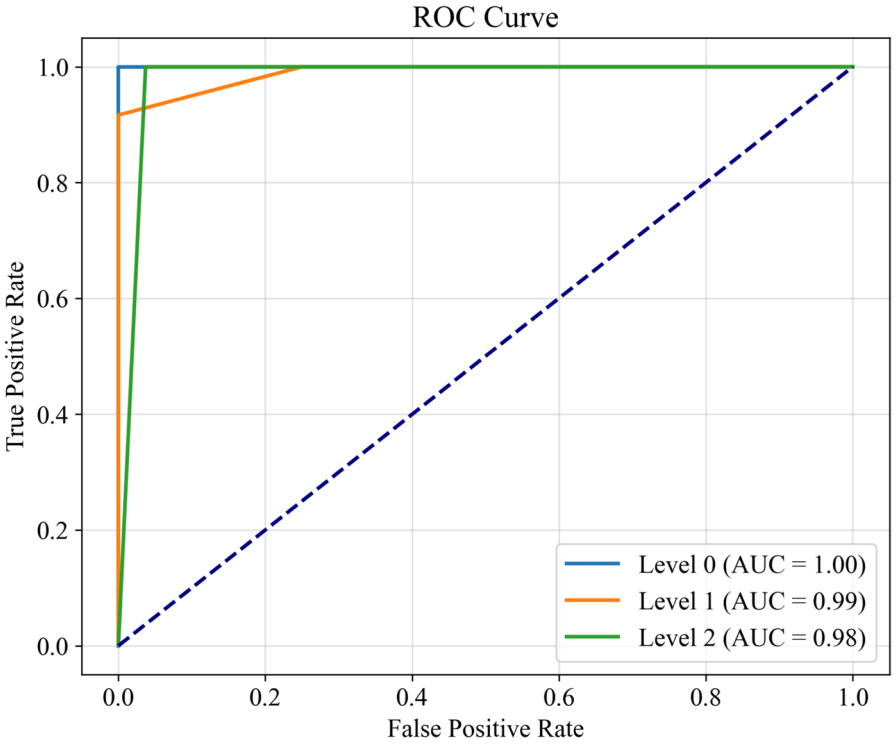


**Fig. 15** ROC curves for the hybrid FA–k-Means–RF model across the three squeezing levels

determine the relationship between various input variables and the associated rock squeezing level. All the input variables were found to have an impact on the surrounding rock squeezing level. The prediction mechanism was developed in three steps. Firstly, factor analysis was used to reduce the number of variables into a manageable one. Secondly, the acquired factors were classified into multiple groups using k-Means clustering. Finally, to anticipate varying levels of surrounding rock squeezing, a supervised random forest algorithm was developed.

The variables employed in this study indicate some of the intrinsic components that influence rock mass stress and strength. The supplementary internal and external parameters, including the tunnel facility's layout and configuration, construction circumstances, excavation methods, topographical and geological conditions, the elevated wall impacts, and the correlation of cave systems should be considered for inclusion in future surrounding rock squeezing designs.

This advancement is essential for promoting intelligent, safe, and sustainable mining practices, as it allows industry professionals to precisely evaluate the risks related to the assessment and classification of rock squeezing potential around tunnels and to make informed decisions before mining operations commence. Proactive steps are crucial for reducing operating risks, improving structural stability, and protecting the well-being of workers and the environment. More sophisticated modeling techniques should be used to more effectively manage difficult geological

conditions in order to further enhance these procedures. This will facilitate the generation of sufficient data for machine learning algorithms, enabling accurate predictions of surrounding rock squeezing potential and ultimately resulting in more efficient mining operations.

In this research, FA is chosen from the context of data dimensionality reduction processing. However, data sources require various dimensionality reduction strategies. The effects of dimensionality reduction approaches such as principal component analysis, T-distributed stochastic neighbor embedding, singular value decomposition, among others, on machine learning models should be investigated further.

**Author Contributions** MK, DJA and PGA contributed to the conception and design. MK, MF, RKW, DJA and PGA contributed to the development of software. MK, DJA and PGA wrote the first draft of the manuscript and reviewed the manuscript. All authors read and approved the final manuscript.

**Funding** Open access funding provided by HEAL-Link Greece. No funding was obtained for this study.

**Data Availability** The raw/processed data required to reproduce these findings will be made available on request.

## Declarations

**Conflict of interest** The authors declare no competing interest.



## References

Altman, N., Krzywinski, M.: Ensemble methods: bagging and random forests. Nat. Methods **14**, 933–934 (2017). https://doi.org/10.1038/nmeth.4438

Arora, K., Gutierrez, M., Hedayat, A., Cruz, E.C.: Time-dependent behavior of the tunnels in squeezing ground: an experimental study. Rock Mech. Rock Eng. **54**, 1755–1777 (2021). https://doi.org/10.1007/s00603-021-02370-w

Asteris, P.G., Mamou, A., Hajihassani, M., Hasanipanah, M., Koopialipoor, M., Le, T.T., Kardani, N., Armaghani, D.J.: Soft computing based closed form equations correlating L and N-type Schmidt hammer rebound numbers of rocks. Transp. Geotech. **29**, 100588 (2021). https://doi.org/10.1016/j.trgeo.2021.100588

Aydan, Ö., Akagi, T., Kawamoto, T.: The squeezing potential of rocks around tunnels; theory and prediction. Rock Mech. Rock Eng. **26**, 137–163 (1993). https://doi.org/10.1007/bf01023620

Azizi, F., Koopialipoor, M., Khoshrou, H.: Estimation of rock mass squeezing potential in tunnel route (case study: Kerman water conveyance tunnel). Geotech. Geol. Eng. **37**, 1671–1685 (2019). https://doi.org/10.1007/s10706-018-0714-5

Bernard, S., Adam, S., Heutte, L.: Dynamic random forests. Pattern Recognit. Lett. **33**, 1580–1586 (2012)

Chen, Y., Li, T., Zeng, P., Ma, J., Patelli, E., Edwards, B.: Dynamic and probabilistic multi-class prediction of tunnel squeezing intensity. Rock Mech. Rock Eng. **53**, 3521–3542 (2020). https://doi.org/10.1007/s00603-020-02138-8

Dalgıç, S.: Tunneling in squeezing rock, the Bolu tunnel, Anatolian Motorway, Turkey. Eng. Geol. **67**, 73–96 (2002). https://doi.org/10.1016/s0013-7952(02)00146-1

Dwivedi, R.D., Singh, M., Viladkar, M.N., Goel, R.K.: Prediction of tunnel deformation in squeezing grounds. Eng. Geol. **161**, 55–64 (2013). https://doi.org/10.1016/j.enggeo.2013.04.005

Feng, X., Jimenez, R.: Predicting tunnel squeezing with incomplete data using Bayesian networks. Eng. Geol. **195**, 214–224 (2015). https://doi.org/10.1016/j.enggeo.2015.06.017

Gao, F., Stead, D., Kang, H.: Numerical simulation of squeezing failure in a coal mine roadway due to mining-induced stresses. Rock Mech. Rock Eng. **48**, 1635–1645 (2015). https://doi.org/10.1007/s00603-014-0653-2

Ghasemi, E., Gholizadeh, H.: Prediction of squeezing potential in tunneling projects using data mining-based techniques. Geotech. Geol. Eng. **37**, 1523–1532 (2019). https://doi.org/10.1007/s10706-018-0705-6

Ghasemi, E., Hassani, S., Kadkhodaei, M.H., Bahri, M., Romero-Hernandez, R., Mascort-Albea, E.J.: An intelligent approach to predict the squeezing severity and tunnel deformation in squeezing grounds. Transp. Infrastruct. Geotechnol. **11**, 1–25 (2024). https://doi.org/10.1007/s40515-024-00434-2

Gioda, G., Cividini, A.: Numerical methods for the analysis of tunnel performance in squeezing rocks. Rock Mech. Rock Eng. **29**, 171–193 (2013). https://doi.org/10.1007/bf01042531

Goel, R.K., Jethwa, J.L., Paithankar, A.G.: Tunnelling through the young Himalayas—a case history of the Maneri-Uttarkashi power tunnel. Eng. Geol. **39**, 31–44 (1995). https://doi.org/10.1016/0013-7952(94)00002-j

Hajihassani, M., Kalatehjari, R., Marto, A., Mohamad, H., Khosrotash, M.: 3D prediction of tunneling-induced ground movements based on a hybrid ANN and empirical methods. Eng. Comput. **36**, 251–269 (2020). https://doi.org/10.1007/s00366-018-00699-5

Hasanipanah, M., Noorian-Bidgoli, M., Jahed Armaghani, D., Khamesi, H.: Feasibility of PSO-ANN model for predicting surface settlement caused by tunneling. Eng. Comput. **32**, 705–715 (2016). https://doi.org/10.1007/s00366-016-0447-0

Hoek, E.: Big tunnels in bad rock. J. Geotech. Geoenviron. Eng. **127**, 726–740 (2001). https://doi.org/10.1061/(asce)1090-0241(2001)127:9(726)

Hoek, E., Guevara, R.: Overcoming squeezing in the Yacambú Quibor tunnel, Venezuela. Rock Mech. Rock Eng. **42**, 389–418 (2009). https://doi.org/10.1007/s00603-009-0175-5

Hoek, E., Marinos, P.: Predicting tunnel squeezing problems in weak heterogeneous rock masses. Tunnels Tunnell. Int. **32**, 45–51 (2000). https://doi.org/10.1201/9780203963586.ch3

Huang, Z., Liao, M., Zhang, H., Zhang, J., Ma, S.: Predicting the tunnel surrounding rock extrusion deformation based on SVM-BP model with incomplete data. Modern Tunnelling Technology **57**, 17–24 (2020). https://doi.org/10.21203/rs.3.rs-379738/v1

Huang, X., Yin, X., Liu, B., Ding, Z., Zhang, C., Jing, B., Guo, X.: A gray wolf optimization-based improved probabilistic neural network algorithm for surrounding rock squeezing classification in tunnel engineering. Front. Earth Sci. **10**, 857463 (2022). https://doi.org/10.3389/feart.2022.857463

Huang, X.C., Wang, G.L., Chen, Q.N., Zhang, W.: Collapse failure assessment of geomaterials behind steel structure in tunnels using the Chebyshev inequalities. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part a: Civil Engineering **10**, 06024002 (2024). https://doi.org/10.1061/ajrua6.rueng-1268

ISRM: Comments and recommendations on design and analysis procedures for structures in argillaceous swelling rock. International Journal of Rock Mechanics and Mining Sciences 31, 5 (1994). https://doi.org/10.1016/0148-9062(94)90155-4

Jiao, Y.Y., Ou, G., Wang, H., Zhang, G.H., Zou, J.P., Tan, F.: Prediction of tunnel squeezing based on evidence theory. Journal of Applied Basic Engineering Sciences **29**, 1156–1170 (2021). https://doi.org/10.1016/j.tust.2021.104019

Kamran, M., Shahani, N.M.: Decision support system for the prediction of mine fire levels in underground coal mining using machine learning approaches. Min. Metall. Explor. **39**, 591–601 (2022). https://doi.org/10.1007/s42461-022-00569-1

Kamran, M., Ullah, B., Ahmad, M., Sabri, M.M.S.: Application of KNN-based isometric mapping and fuzzy c-means algorithm to predict short-term rockburst risk in deep underground projects. Front. Public Health **10**, 1023890 (2022). https://doi.org/10.3389/fpubh.2022.1023890

Kamran, M., Chaudhry, W., Wattimena, R.K., Rehman, H., Martyushev, D.A.: A multi-criteria decision intelligence framework to predict fire danger ratings in underground engineering structures. Fire **6**, 412 (2023). https://doi.org/10.3390/fire6110412

Kamran, M., Jiskani, I.M., Wang, Z., Zhou, W.: Decision intelligence-driven predictive modelling of air quality index in surface mining. Eng. Appl. Artif. Intell. **133**, 108399 (2024). https://doi.org/10.1016/j.engappai.2024.108399

Kimura, F., Okabayashi, N., Kawamoto, T.: Tunnelling through squeezing rock in two large fault zones of the Enasan tunnel II. Rock Mech. Rock Eng. **20**, 151–166 (1987). https://doi.org/10.1007/bf01020366

Kovári, K., Staus, J.: Basic considerations on tunnelling in squeezing ground. Rock Mech. Rock Eng. **29**, 203–210 (1996). https://doi.org/10.1007/bf01042533

Leone, T., Nordas, A.N., Anagnostou, G.: Effects of creep on shield tunnelling through squeezing ground. Rock Mech. Rock Eng. **57**, 351–374 (2024). https://doi.org/10.1007/s00603-023-03505-x

Liang, M., Peng, H., Xie, W., Yu, B., Han, Y., Zhu, M., Huang, N.: Dynamic multiclass prediction of tunnel squeezing intensity with stacking model and Markov process. Tunnell. Undergr. Space Technol. **146**, 105632 (2024). https://doi.org/10.1016/j.tust.2024.105632

Liu, W., Zhao, T., Zhou, W., Tang, J.: Safety risk factors of metro tunnel construction in China: an integrated study with EFA and SEM. Saf. Sci. **105**, 98–113 (2018). https://doi.org/10.1016/j.ssci.2018.01.009

Liu, W., Chen, J., Luo, Y., Chen, L., Shi, Z., Wu, Y.: Deformation behaviors and mechanical mechanisms of double primary linings for large-span tunnels in squeezing rock: a case study. Rock Mech. Rock Eng. **54**, 2291–2310 (2021). https://doi.org/10.1007/s00603-021-02402-5

Lu, S., Koopialipoor, M., Asteris, P.G., Bahri, M., Armaghani, D.J.: A novel feature selection approach based on tree models for evaluating the punching shear capacity of steel fiber-reinforced concrete flat slabs. Materials **13**, 3902 (2020). https://doi.org/10.3390/ma13173902

Luo, H., Wang, Z., Liu, K., Qiao, L., Qing, L.: Stress-strain state zoning model and novel large deformation classification method for squeezing tunnels. Eng. Fail. Anal. **164**, 108711 (2024). https://doi.org/10.3390/ma13173902

Muirwood, A.M.: Tunnels for roads and motorways. Q. J. Eng. Geol. **5**, 119–120 (1972)

Najm, S.J., Daraei, A.: Forecasting and controlling two main failure mechanisms in the Middle East's longest highway tunnel. Eng. Fail. Anal. **146**, 107091 (2023). https://doi.org/10.1016/j.engfailanal.2023.107091

Panet, M.: Two case histories of tunnels through squeezing rocks. Rock Mech. Rock Eng. **29**, 155–164 (1996). https://doi.org/10.1007/bf01032652

Panthi, K.K., Nilsen, B.: Uncertainty analysis of tunnel squeezing for two tunnel cases from Nepal Himalaya. Int. J. Rock Mech. Min. Sci. **44**, 67–76 (2007). https://doi.org/10.1016/j.ijrmms.2006.04.013

Sakurai, S.: Lessons learned from field measurements in tunnelling. Tunnell. Undergr. Space Technol. **12**, 453–460 (1997). https://doi.org/10.1016/s0886-7798(98)00004-2

Sakurai, S., Takeuchi, K.: Back analysis of measured displacements of tunnels. Rock Mech. Rock Eng. **16**, 173–180 (1983). https://doi.org/10.1007/bf01033278

Singh, M., Singh, B., Choudhari, J.: Critical strain and squeezing of rock mass in tunnels. Tunn. Undergr. Space Technol. **22**, 343–350 (2007). https://doi.org/10.1016/j.tust.2006.06.005

Sun, Y., Feng, X., Yang, L.: Predicting tunnel squeezing using multiclass support vector machines. Adv. Civ. Eng. **2018**, 4543984 (2018). https://doi.org/10.1155/2018/4543984

Tran Manh, H., Sulem, J., Subrin, D., Billaux, D.: Anisotropic time-dependent modeling of tunnel excavation in squeezing ground. Rock Mech. Rock Eng. **48**, 2301–2317 (2015). https://doi.org/10.1007/s00603-015-0717-y

Wang, J.: The key way is to release the genuine rock pressure—discussion on problems of tunnelling in squeezing ground. Modern Tunnelling Technology **57**, 1–11 (2020)

Wang, X., Iura, T., Jiang, Y., Wang, Z., Liu, R.: Deformation and mechanical characteristics of tunneling in squeezing ground: a case study of the west section of the Tawarazaka Tunnel in Japan. Tunn. Undergr. Space Technol. **109**, 103697 (2021). https://doi.org/10.1016/j.tust.2020.103697

Wang, G., Huang, X., Bu, D., Dai, Z.: How large is the collapsed area of ground collapse induced by tunnelling. Geomech. Geoengin. **19**, 1–14 (2024). https://doi.org/10.1080/17486025.2024.2377558

Wu, G., Chen, W., Tan, X., Zhao, W., Jia, S., Tian, H., Yang, J.: Performance of new type of foamed concrete in supporting tunnel in squeezing rock. Int. J. Geomech. **20**, 04019173 (2020). https://doi.org/10.1061/(asce)gm.1943-5622.0001543

Wu, K., Shao, Z., Qin, S., Wei, W., Chu, Z.: A critical review on the performance of yielding supports in squeezing tunnels. Tunn. Undergr. Space Technol. **115**, 103815 (2021). https://doi.org/10.1016/j.tust.2021.103815

Xu, Z.H., Wang, W.Y., Lin, P., Nie, L.C., Wu, J., Li, Z.M.: Hard-rock TBM jamming subject to adverse geological conditions: influencing factor, hazard mode and a case study of Gaoligongshan Tunnel. Tunn. Undergr. Space Technol. **108**, 103683 (2021). https://doi.org/10.1016/j.tust.2020.103683

Yun, H.B., Park, S.H., Mehdawi, N., Mokhtari, S., Chopra, M., Reddi, L.N., Park, K.T.: Monitoring for close proximity tunneling effects on an existing tunnel using principal component analysis technique with limited sensor data. Tunn. Undergr. Space Technol. **43**, 398–412 (2014). https://doi.org/10.1016/j.tust.2014.06.003

Zhang, J., Li, D., Wang, Y.: Predicting tunnel squeezing using a hybrid classifier ensemble with incomplete data. Bull. Eng. Geol. Environ. **79**, 3245–3256 (2020). https://doi.org/10.1007/s10064-020-01747-5

Zhou, J., Zhu, S., Qiu, Y., Armaghani, D.J., Zhou, A., Yong, W.: Predicting tunnel squeezing using support vector machine optimized by whale optimization algorithm. Acta Geotech. **16**, 1–24 (2021). https://doi.org/10.1007/s11440-022-01450-7

Zhu, X., Jin, X., Jia, D., Sun, N., Wang, P.: Application of data mining in an intelligent early warning system for rock bursts. Processes **7**, 55 (2019). https://doi.org/10.3390/pr7020055



## Authors and Affiliations

**Muhammad Kamran[1] · Muhammad Faizan[2] · Ridho Kresna Wattimena[3] · Danial Jahed Armaghani[4] · Panagiotis G. Asteris[5]**

✉ Panagiotis G. Asteris
panagiotisasteris@gmail.com

Muhammad Kamran
m.kamran@utas.edu.au

Muhammad Faizan
bi91xd@student.sunderland.ac.uk

Ridho Kresna Wattimena
rkw@mining.itb.ac.id

Danial Jahed Armaghani
danial.jahedarmaghani@uts.edu.au

[1] School of Engineering, University of Tasmania, Hobart, TAS 7001, Australia

[2] University of Sunderland, Sunderland, UK

[3] Bandung Institute of Technology, Bandung, Indonesia

[4] School of Civil and Environmental Engineering, University of Technology Sydney, Ultimo, Sydney 2007, Australia

[5] School of Pedagogical and Technological Education, Marousi, Greece