**Geomechanics and Engineering**, *Vol. 37, No. 3 (2024) 223-241*
https://doi.org/10.12989/gae.2024.37.3.223


# Computing machinery techniques for performance prediction of TBM using rock geomechanical data in sedimentary and volcanic formations

Hanan Samadi[1], Arsalan Mahmoodzadeh[1], Shtwai Alsubai[2], Abdullah Alqahtani[2], Abed Alanazi*[2] and Ahmed Babeker Elhag[3]

[1]*IRO, Civil Engineering Department, University of Halabja, Halabja, 46018, Iraq*
[2]*Department of Computer Science, College of Computer Engineering and Sciences in Al-Kharj, Prince Sattam bin Abdulaziz University, P.O. Box 151, Al-Kharj 11942, Saudi Arabia*
[3]*Department of Civil Engineering, College of Engineering, King Khalid University, Abha 61413, Saudi Arabia*

(*Received May 19, 2023, Revised March 24, 2024, Accepted April 11, 2024*)

**Abstract.** Evaluating the performance of Tunnel Boring Machines (TBMs) stands as a pivotal juncture in the domain of hard rock mechanized tunneling, essential for achieving both a dependable construction timeline and utilization rate. In this investigation, three advanced artificial neural networks namely, gated recurrent unit (GRU), back propagation neural network (BPNN), and simple recurrent neural network (SRNN) were crafted to prognosticate TBM-rate of penetration (ROP). Drawing from a dataset comprising 1125 data points amassed during the construction of the Alborze Service Tunnel, the study commenced. Initially, five geomechanical parameters were scrutinized for their impact on TBM-ROP efficiency. Subsequent statistical analyses narrowed down the effective parameters to three, including uniaxial compressive strength (UCS), peak slope index (PSI), and Brazilian tensile strength (BTS). Among the methodologies employed, GRU emerged as the most robust model, demonstrating exceptional predictive prowess for TBM-ROP with staggering accuracy metrics on the testing subset ($R^2$ = 0.87, NRMSE = 6.76E-04, MAD = 2.85E-05). The proposed models present viable solutions for analogous ground and TBM tunneling scenarios, particularly beneficial in routes predominantly composed of volcanic and sedimentary rock formations. Leveraging forecasted parameters holds the promise of enhancing both machine efficiency and construction safety within TBM tunneling endeavors.

**Keywords:** computing machinery techniques; deep learning; rock geomechanical data; TBM penetration rate

## 1. Introduction

The tunneling industry has been experiencing significant advancements in recent years, driven by technological innovation and the growing demand for underground infrastructure (He *et al.* 2023, Hu *et al.* 2023, Liu *et al.* 2023). In recent decades, there has been a rise in the construction of tunnels in urban and hard rock areas using tunnel boring machines (TBM). Tunnels for metro, rail, and water transfer purposes, typically with circular cross sections, are well-suited for mechanized excavation. An important aspect of project management is the accurate estimation of the driving and operating parameters of TBMs to ensure efficiency in pre-construction, construction, and post-construction phases, as well as effective cost management (Gokceoglu 2022, Hassanpour *et al.* 2011).

Failure to establish precise driving and operating parameters based on geological conditions along the tunnel route can lead to increased maintenance and repair costs for cutterhead's and disc cutters due to abrasion (Adoko *et al.* 2017, Gokceoglu 2022, Liu *et al.* 2015). Among these parameters, the Rate of Penetration (ROP) of TBMs is crucial for assessing the machine's performance in hard rock conditions. Accurately estimating the ROP of the TBM is essential for minimizing delays, costs, and uncertainties along the tunnel path (Ates *et al.* 2014).

Based on the literature review, there are three categories of theoretical, statistical, and empirical approaches to calculating and assessing TBM performance (Carter and Marinos 2020, Elmo and Stead 2021, Farrokh *et al.* 2012, Hassanpour *et al.* 2011, Yang *et al.* 2018). However, the use of these methods has become less popular due to their time-consuming and costly nature, as well as the complex conditions of the ground and the non-linearity factors. Therefore, to solve such problems, it is necessary to use more powerful methods, including machine learning (ML) algorithms (Yin *et al*. 2023).

Recently, ML methods have shown an acceptable ability to solve various engineering problems (Shi *et al*. 2023, Shi *et al*. 2023, Su *et al*. 2023, Yan *et al*. 2024, Zhao *et al*. 2024). In the field of mechanized excavation, several ML algorithms, including recurrent neural networks (RNN), convolutional neural networks (CNN), artificial neural networks (ANN), support vector machine (SVM), and fuzzy inference systems (FIS), have been coded and developed to predict TBM performance (Jung *et al*. 2019, Mahmoodzadeh *et al*. 2022, Shahrour and Zhang 2021, Wang *et al*. 2020, Zhou *et al*. 2021). Shao *et al*. (2013) performed least squares support vector machine (LS-SVM),

*Corresponding author, Assistant Professor
E-mail: ad.alanazi@psau.edu.sa



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

genetic programming (GP), and extreme learning machine (ELM) algorithms to predict the rate of penetration (ROP) of TBM. In other studies, the SVM and long-short-term memory (LSTM) were used to assess the operating and driving parameters of TBM (Liu *et al*. 2019,; Liu *et al*. 2015). Furthermore, Zhou *et al.* (2021) applied a hybrid predictive network, namely extreme gradient boosting (XGB), for the estimation of ROP.

This study aims to apply several supervised and unsupervised algorithms for predicting the TBM-ROP based on 1125 datasets collected from one of the most important tunneling projects in Iran, namely the Alborz service tunnel on the Tehran-Shomal motorway project (ASTSHM). The applied methods include statistical and DL algorithms. Based on the statistical methods, a new empirical formula to predict the ROP was proposed by considering the geomechanical properties of different types of rocks (volcanic and sedimentary) along the tunnel route. Also, gated recurrent units (GRU), back propagation neural networks (BPNN), and simple recurrent neural networks (SRNN) were developed to predict the TBM ROP. 80% of the dataset was considered for training networks, and the remaining dataset was applied for testing networks (see Fig. 1). As a first step, the principal component analysis (PCA) and multi-collinearity analysis were introduced and applied to generate the collected dataset, and then the correlation between dependent and independent variables was defined based on the linear and non-linear single and multi-variable regression (MVR) to develop a new empirical formula. After that, the DL algorithms were used to develop predictive networks for the prediction of ROP. At the end, a comparison between applied methods was made to introduce the robust model to predict the TBM-ROP.

The review of relevant literature shows that many researchers have developed various soft computing approaches using different databases. It is important to emphasize that the quality and quantity of the database significantly affect the performance of these models. While most researchers have used neural network models for evaluating TBM driving and operating parameters, models based on deep neural networks with standardized and normalized databases have not been developed and utilized for assessing TBM performance using geomechanical parameters such as the plane of weakness, TBM-driven direction (α), distance between planes of weakness (DPW), uniaxial compressive strength (UCS), peak slope index (PSI), and Brazilian tensile strength (BTS). Furthermore, statistical hypothesis tests like analysis of variance (ANOVA), Van-der-Waerden's and Tukey's methods, collinearity statistics, Probability-Probability and Quantile-Quantile plots, and Chi-tests have not been employed to select the relevant research hypothesis.

This study aims to develop, train, test, and analyze deep learning approaches. Expanding on the identified gap in the literature review, this research introduces the following contributions:

The novelties of this study include the creation of empirical models for TBM-ROP calculation through statistical analysis and unsupervised learning, using a robust database analyzed thoroughly by statistical methods. New empirical equations were formulated for predicting TBM-ROP based on rock geomechanical properties. The accuracy and reliability of the models were assessed using various evaluation metrics, and an in-depth analysis of statistical, empirical, and supervised learning results was conducted to identify the most effective model for TBM-ROP calculation. This work is significant as it can address operator concerns regarding TBM performance by optimizing the boring process in rock tunneling with similar geomechanical and geological characteristics. This work is significant as it can address many operator concerns regarding TBM performance by optimizing the boring process and utilizing TBMs in rock tunneling with similar geomechanical and geological characteristics (sedimentary and volcanic rocks: Sandstone, Limestone, Tuff, Gypsum, Shale).

## 2. Artificial neural network algorithms

The algorithms of artificial intelligence, in particular DL, have been utilized in tunneling projects with complex geological conditions and numerous required input factors. A concise introduction to BPNN, SRNN, and GRU is given in this section. Then, the results of the developed models to estimate the TBM-ROP are presented.

As a first step, the dataset, including inputs and output factors, should be developed (1125 data). Then the dataset should be divided into two subsets: training and testing subsets. 80% of the dataset is considered to be used to train the networks, and the remaining (20%) is applied to test the developed networks. These algorithms were trained and developed based on the geomechanical rock properties, including BTS, UCS, and PSI as inputs and the values of ROP as output. After that, a comparison between measured and predicted values should be applied to test the reliability and accuracy of the predictor networks. In this regard, several evaluation metrics and statistical indices, including RMSE, MSE, NRMSE, RRSE, RSE, MAD, MAPE, and $R^2$ have been suggested to evaluate the accuracy and efficiency of the obtained results. Finally, the models should be verified by comparing the training and testing errors.

### *2.1 Back Propagation Neural Network (BPNN)*

BPNN consists of several layers, including an input layer, hidden layer(s), and an output layer. The deviation of the predicted results from the actual data is determined in the output layer, which will propagate back towards the input layer.

The outline of the BPNN is presented in Fig. 2, which shows three interconnected layers having various weights. Both the hidden and output layers have biases with weight one. This network is performed in three phases: sending, back propagation stages, and weight updating. In the first stage, the signal is sent from the input layer to the output layer, and in the second phase, the error from the output layer to the input is propagated back. Moreover, the determined weights are being updated in the third step. This technique is more efficient than a simple neural network

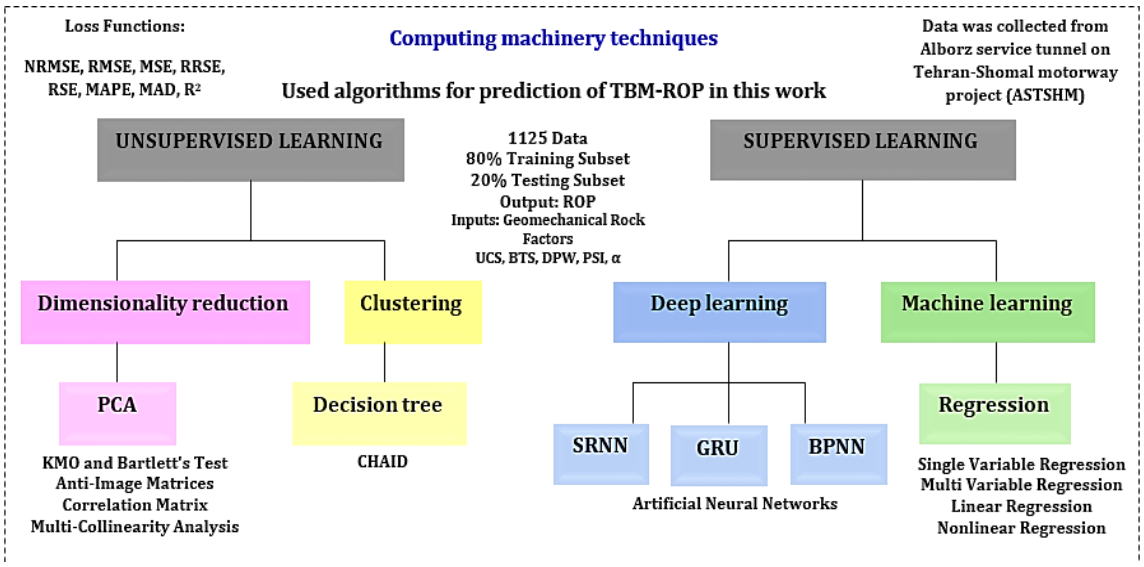



Fig. 1 Schematic of the methods applied to this study

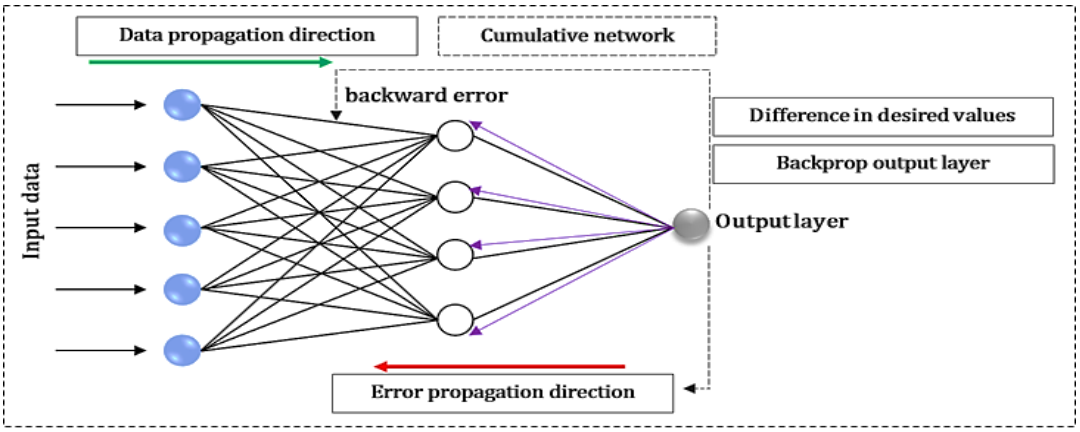



Fig. 2 The outline of the BPNN method

due to its fast processing, high flexibility, and lack of need to pay any special attention to the features during the training stage.

### *2.2 Simple Recurrent Neural Network (SRNN)*

RNN consists of two hidden states and main inputs whose output passes through nonlinear excitation functions such as tanh. In this type of network, the relationship between nodes over a time sequence (data transmitted over time) is graph-oriented. Eq. (1) indicates how to multiply the inputs in the weight matrices and calculate the output, where $b_n$ is the considered bias for the model. Therefore, independent activators first become dependent, and each layer ($x_t$) is assigned the same weight (w) and bias ($b_n$) to reduce the complexity of the RNN. By placing the previous output as the next layer input ($h_t$), a standard platform is provided to store the previous outputs ($h_{t-1}$). Finally, all the same weights and biases are combined into one recurrent unit. The whh is the current neuron weight, and the wxh is the input neuron weight. The structure and details of the SRNN is shown in Fig. 3.

### *2.3 Gated Recurrent Units (GRU)*

The structure of GRU has been proposed by Cho *et al.* (2014) to eliminate the limitations of the SRNN, such as the vanishing gradient problem and the overhead reduction in the LSTM network. To solve the vanishing gradient problem, concepts such as reset gate and update gate have been used in this network type. The general relation between these gates is in the form of Eq. (2), where W, U, and b are the specific coefficients of the gate and σ is the sigmoid function. The memory cell is also denoted by $C_t$, which is mathematically represented in Eq. (3). In this regard, $T_u$ shows the update gate, whose calculations for this gate are derived from Eq. (4).

$$h_t = \tanh(whh \times (h_{t-1}) + wxh \times xt + b_n) \quad (1)$$

$$\Gamma = \sigma(Wx < t > + Ua < t-1 > + b) \quad (2)$$

$$Ct = Tu \times \sim Ct + (1 - Tu) \times (Ct - 1) \quad (3)$$

$$Tu = \sigma(W3 \times (Ct - 1) + W4 \times Xt + bu) \quad (4)$$

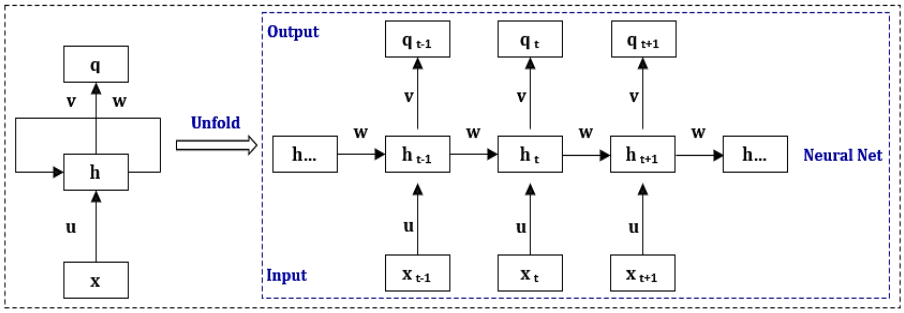



Fig. 3 The structure of SRNN

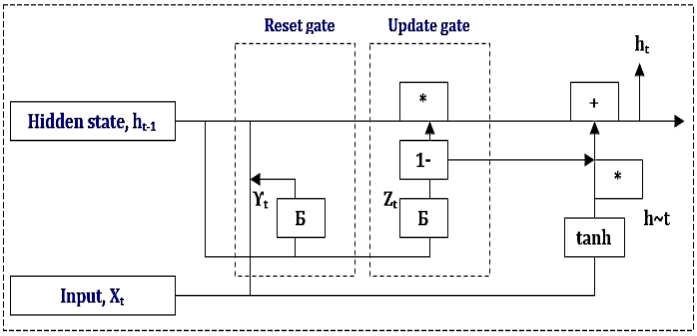



Fig. 4 The structural of GRU


Fig. 5 Open gripper TBM manufactured by Wirth

The update gate acts as a switch, which in a step of time determines the use of the previous mode, input, or combination of both. With this new feature, the network is able to affect the state of several previous time steps in the next few time steps in long sequences. In other words, the network will be able to use elements that it has kept in its memory for a long time. The reset gate may alternatively be thought of as a switch, allowing the network to choose between using all of the information from the preceding phase and skipping over some of it. When this switch is in the zero position, the network is tricked into thinking it is reading the beginning of the input sequence and is thus able to forget its current state calculation. The mathematical formula for the reset gate is in the form of Eq. (5), and the general structure of the GRU network is shown in Fig. 4.

$$\mathrm{Tr} = \sigma\,(\mathrm{W3} \times [\mathrm{Ct} - 1.\ \mathrm{Xt}] + \mathrm{br}) \qquad (5)$$

## 3. Project description

In this study, a database is collected from one of the mechanized tunneling projects in Iran. The Alborz service tunnel on the Tehran-Shomal motorway project (ASTSHM) connects the cities of Tehran and the south of the Caspian Sea. The excavation of the ASTSHM project was completed using an open gripper TBM manufactured by Wirth with a 5.2 m diameter (Fig. 5). Alborz main tunnels project, with a total length of 121 km, has over 30 twin tunnels with double lanes (length of each 6.1 km) and is one of the most important transportation projects in Iran to reduce traffic and travel time.

The geographical situation of the ASTSHM project is presented in Fig. 6. The route of the tunnel mainly passed through sedimentary and volcanic rocks such as Tuffs, Limestone, Sandstone, Conglomerates, and Andesite with

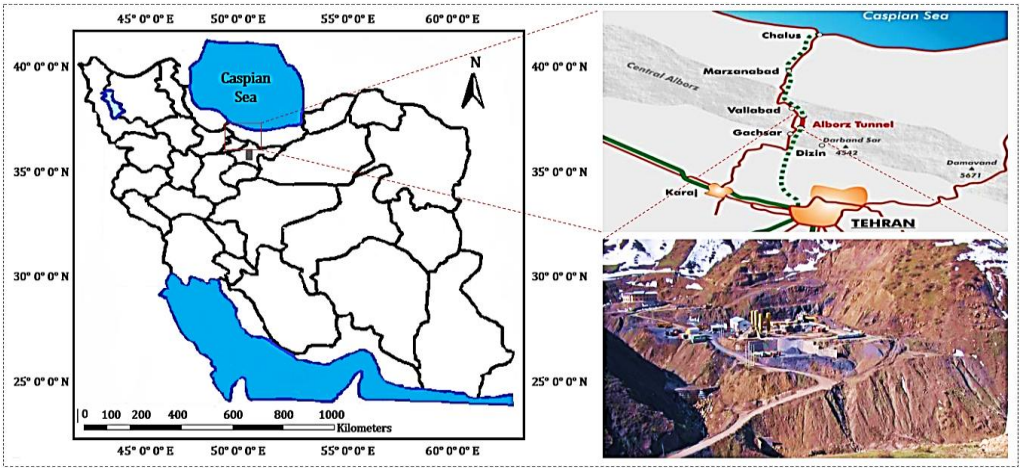



Fig. 6 Geographical situation of the Alborz service tunnel

(a) (b)

Fig. 7 (a) Heavy water ingress and (b) Water ingress zone, getting under control

Table 1 Geomechanical parameters for different engineering geological units

| Rock type | Ave. Density (kg/m$^3$) | Ave. UCS (MPa) | RQD (%) | RMR |
|---|---|---|---|---|
| Sandstone | 2700 | 120 | 50–75 | 45–55 |
| Limestone | 2650 | 85 | 45-55 | 45-50 |
| Tuff | 2650 | 70 | 50–60 | 45–50 |
| Gypsum | 2300 | 65 | 70–80 | 55–60 |
| Shale | 2500 | 55 | 55–70 | 50–60 |

Gypsum Karst. Based on the results of geomechanical observations, the compressive strength of the rocks along the tunnel varies from 20 to 120 MPa.

In a special section of the tunnel, a squeezing hazard was observed due to the high flow of water into the tunnel (Fig. 7). This section passed through the longest fault, which was located in ST5339–5361. The geotechnical profile of the ASTSHM project is indicated in Fig. 8. The main geomechanical properties of the major lithologies along the tunnel are presented in Table 1.

As the main challenges during tunnel construction, the following items should be mentioned: squeezing rock, several fault zones leading to considerable overbreak and cutter head jamming, significant water ingress, methane gas in high concentrations, and the presence of H2S gas.

## 4. Database development

In order to generate and develop the database for this study, rock geomechanical data (intact and rock mass properties) and TBM performance based on the Alborz service tunnel on the ASTSHM project were collected during the pre-construction and construction phases. The collected data points were arranged in a specific database, including 1125 data points where the information on the TBM performance and geotechnical properties was reliable (Fig. 9). The geomechanical data on intact and rock mass properties were obtained and investigated directly from the observations of the tunnel face or established using drilled boreholes and the results of laboratory tests on core samples. Pareto graphs presented in Fig. 10 show the histograms and distribution of different geomechanical properties of rocks and TBM performance recorded in the dataset. It is verified that the measurements of KMO and Bartlett's test result in a high degree of anti-image evaluation across both tests. And statistics tests are presented to the fellow participants, they find more evidence for further improvement in these models. Choosing what type of feature parameters gives better results, how many input data points need frequent reference - all these are important for high accuracy ML analysis.

The database was divided into two main categories. The

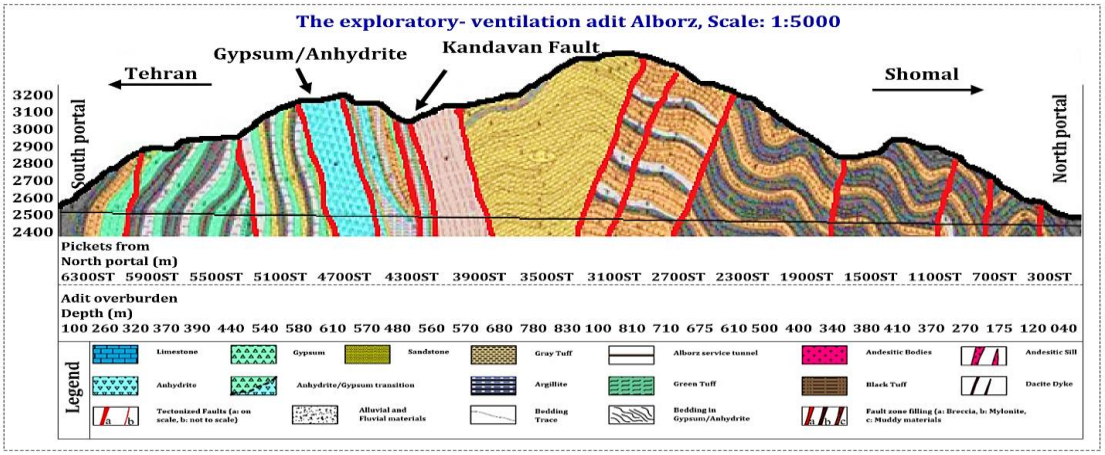



Fig. 8 Geotechnical profile of the Alborz service tunnel

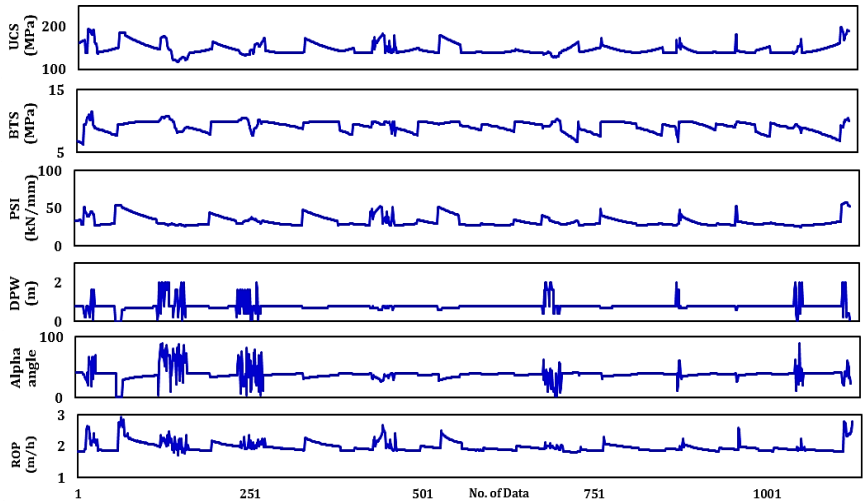



Fig. 9 The order and range of changes in defined factors in the generated database

first category (output parameter) was based on the TBM-ROP. The second section of the dataset (input factors) included geomechanical data such as α, DPW, UCS, PSI, and BTS. These parameters are essential in comprehending the behavior of the rock and determining the efficacy of TBM performance (Jung *et al.* 2019, Mahmoodzadeh *et al.* 2022, Shahrour and Zhang 2021, Wang *et al.* 2020, Zhou *et al.* 2021).

In this study, PCA and Pearson correlation analysis are used to generate and select appropriate independent parameters among all the collected data points. These methods are use multivariable statistical concepts to calculate the principal components, several original variables are correlated in hope of reducing dimensions based on imagination the most suitable generalization of what makes up datums: variation over observations (Elhaik 2022, Jolliffe and Cadima 2016, Kalnins 2022). PCA therefore was applied for output parameters or input values. Consequently, by parallelism relation coefficient $p<0.01$ conclusions are also: between TBM penetration speed and intrinsic properties such as modulus of the rock. The results of the PCA and Pearson correlation are presented in Table 2 and Figs. 11 and 12. According to these results, while parameters α and DPW might not exhibit a significant correlation with TBM-ROP, the remaining parameters (UCS, PSI, and BTS) displayed a stronger correlation, underscoring the importance of rock strength characteristics in the tunneling process. Consequently, our attention shifts towards these three influential parameters for the development of smart supervised algorithms, as they demonstrate a more pronounced impact on the output parameter during model training and testing. Furthermore, understanding the interplay between these geomechanical properties and their impact on the performance of TBM is crucial for enhancing tunneling efficiency and safety. The UCS of the rock material plays a vital role in determining the compressive strength of the formation to deformation under stress, thereby influencing the cutting capacity of the TBM. Similarly, the PSI is indicative of the rock's inherent strength and brittleness, affecting the machine's ability to penetrate the formation efficiently. Additionally, the BTS parameter provides insights into the resistance and tensile strength encountered during the tunneling process, highlighting the need for optimal cutting and thrust forces to overcome the geological challenges.

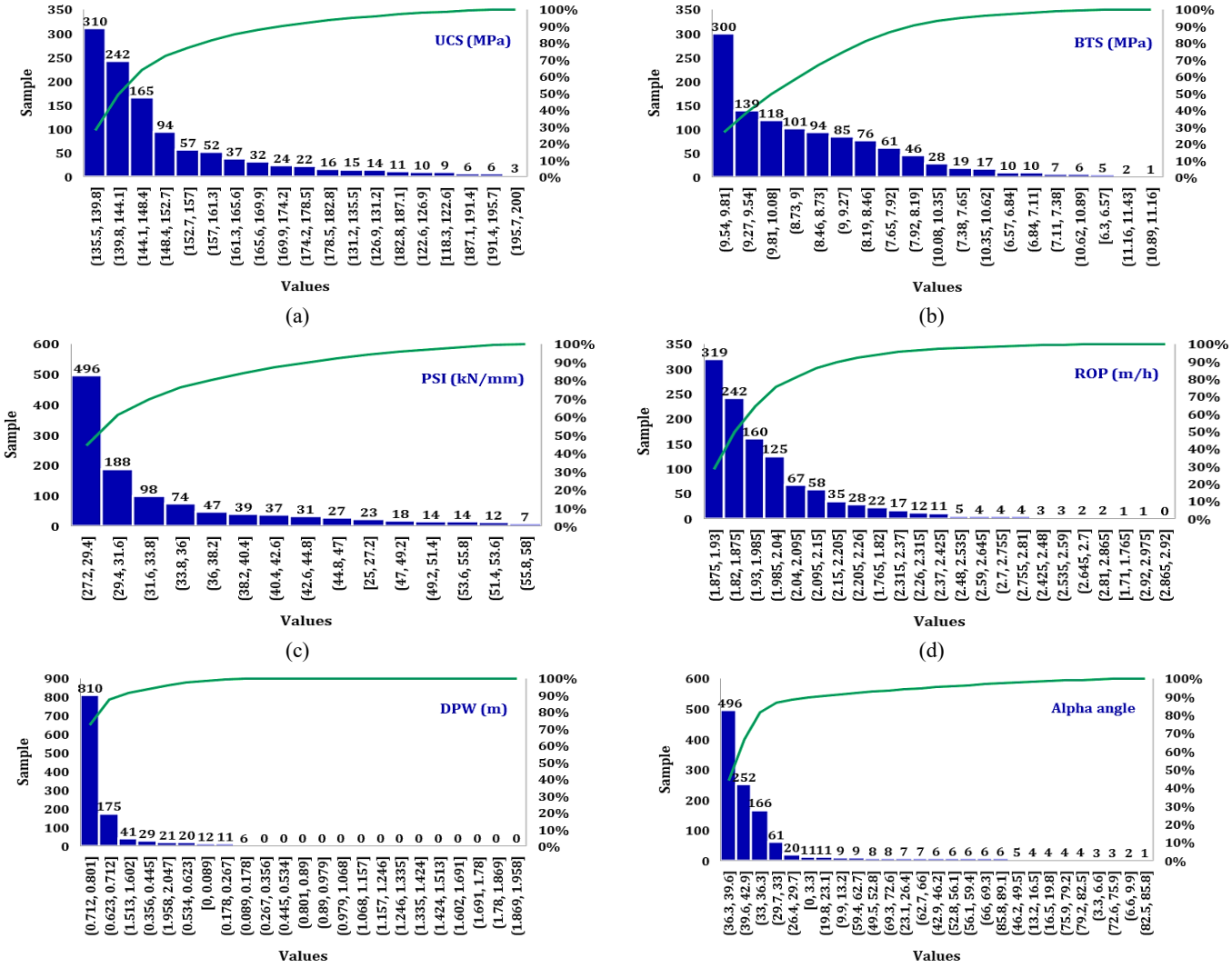
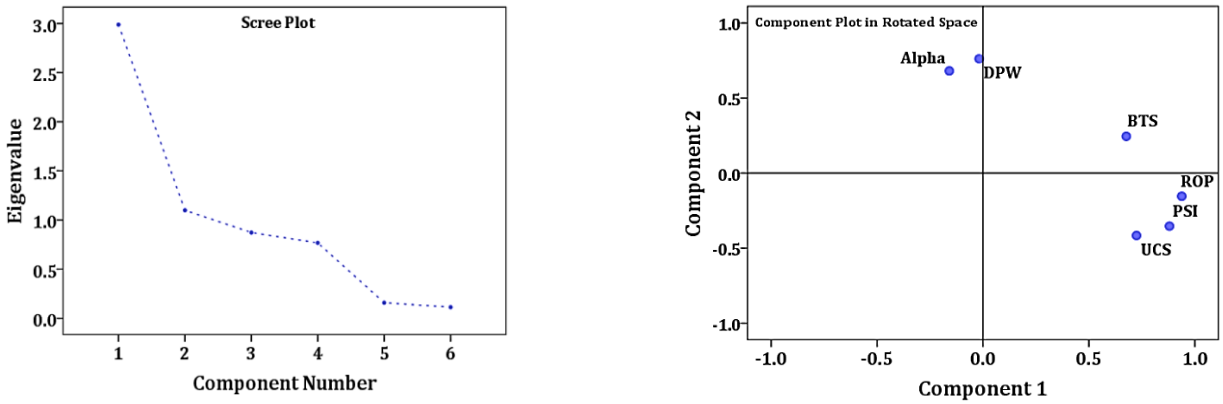



Fig. 10 Pareto graphs and distribution of input and output factors

Table 2 PCA results for the generated dataset to calculate the TBM–ROP

| KMO and Bartlett's Test | | | | | | |
|---|---|---|---|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | | | | | 0.69 |
| Bartlett's Test of Sphericity | | Approx. Chi-Square | | | | 3618.95 |
| | | df | | | | 15 |
| | | Sig. | | | | 0.00 |
| Anti-image Matrices | | | | | | |
| | | UCS | BTS | PSI | DPW | Alpha | ROP |
| Anti-image Covariance | UCS | 0.26 | 0.17 | -0.11 | 0.01 | -0.01 | -0.06 |
| | BTS | 0.17 | 0.54 | -0.04 | -0.009 | 0.02 | -0.15 |
| | PSI | -0.11 | -0.04 | 0.16 | 0.02 | 0.10 | -0.09 |
| | DPW | 0.01 | -0.009 | 0.02 | 0.91 | -0.13 | -0.01 |
| | Alpha | -0.01 | 0.02 | 0.10 | -0.13 | 0.83 | -0.06 |
| | ROP | -0.06 | -0.15 | -0.09 | -0.01 | -0.06 | 0.19 |
| Anti-image Correlation | UCS | 0.69 | 0.46 | -0.52 | 0.03 | -0.02 | -0.27 |
| | BTS | 0.46 | 0.50 | -0.13 | -0.01 | 0.03 | -0.48 |
| | PSI | -0.52 | -0.13 | 0.72 | 0.07 | 0.27 | -0.53 |
| | DPW | 0.03 | -0.01 | 0.07 | 0.84 | -0.15 | -0.02 |
| | Alpha | -0.02 | 0.03 | 0.27 | -0.15 | 0.68 | -0.16 |
| | ROP | -0.27 | -0.48 | -0.53 | -0.02 | -0.16 | 0.72 |


Fig. 11 Pearson correlation matrix between defined factors in the developed database




(a) (b)

Fig. 12 (a) scree plot and (b) component plot in rotated space

Table 3 Statistical analysis of the dependent and independent factors

| | Input factors | | | | | Output factor |
|---|---|---|---|---|---|---|
| | Uniaxial compressive strength | Brazilian test strength | Peak slope index | Distance between the planes of weakness | Alpha angle | ROP |
| Symbol | UCS | BTS | PSI | DPW | α | ROP |
| Unit | MPa | MPa | kN/mm | m | degrees | m/h |
| Min. | 118.30 | 6.30 | 25.00 | 0.00 | 0.00 | 1.71 |
| Max. | 199.70 | 11.40 | 58.00 | 2.00 | 89.00 | 2.95 |
| Med. | 143.25 | 9.40 | 29.80 | 0.80 | 38.50 | 1.93 |
| Ave. | 147.53 | 9.14 | 32.82 | 0.80 | 38.31 | 1.98 |
| Var. | 162.74 | 0.66 | 44.44 | 0.07 | 95.74 | 0.03 |
| STD. | 12.76 | 0.81 | 6.67 | 0.26 | 9.78 | 0.16 |
| Skew. | 1.41 | -0.81 | 1.67 | 2.14 | 1.03 | 2.28 |
| Kurt. | 2.14 | 0.38 | 2.10 | 9.60 | 9.29 | 6.72 |

After a database was built and the proper factors chosen, starters put together this database. It covers TBM-ROP and geological mechanical properties of marine volcanic and sedimentary rocks. These properties are from the UCS, BTS, and PSI tests. As can be seen, the statistical analysis, including the median (Med), maximum (Max), minimum (Min), average (Ave), standard deviation (STD), variance (Var), skewness (Skew), and kurtosis (Kurt) of dependent and independent factors is reported. Since the database is established based on one project, all the constant factors related to the physical and mechanical properties of TBM, such as the mechanical specification of cutterhead and disc cutters and backup systems, were not considered for analysis.

A normal P-P plot (probability-probability) and detrended normal P-P plot of inputs (UCS, BTS, PSI) and

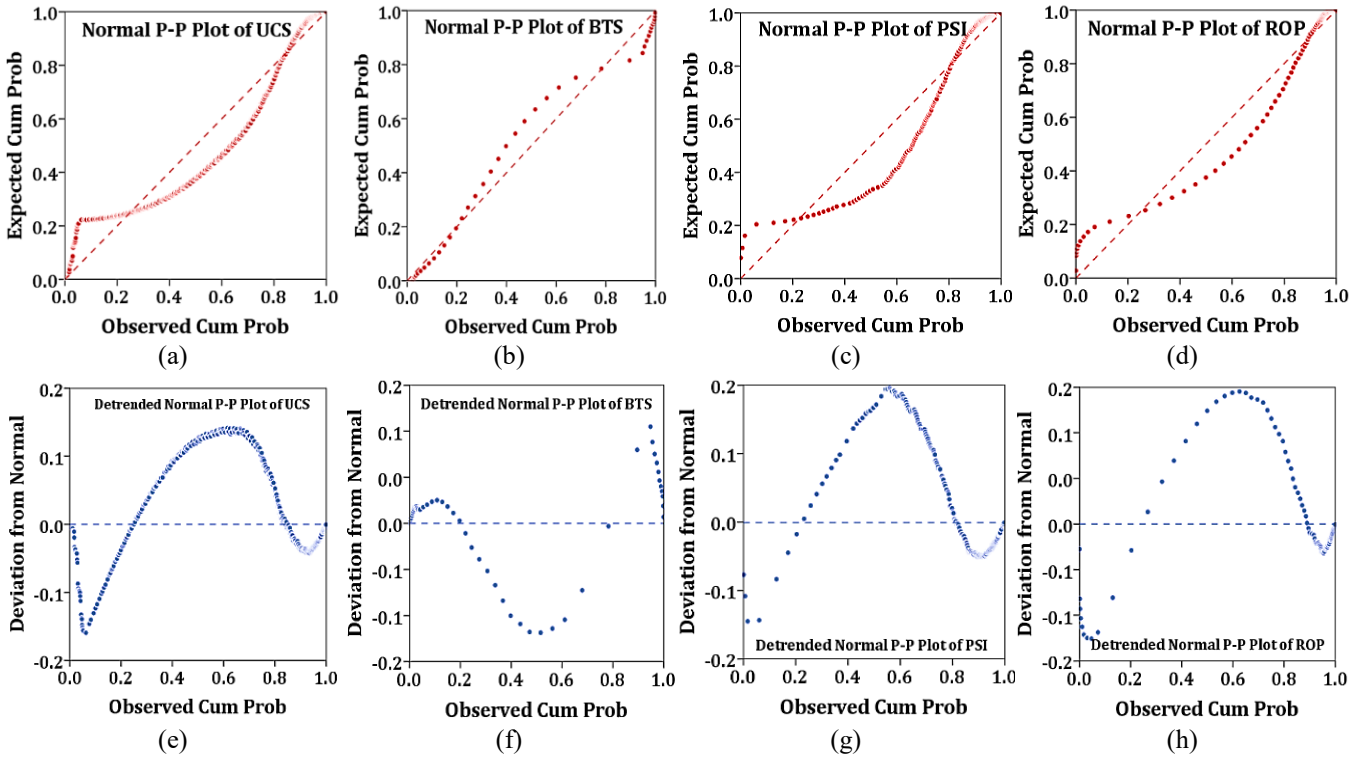



Fig. 13 The P-P and detrended normal P-P plot of dependent (ROP) and independent variables (UCS, PSI, BTS)

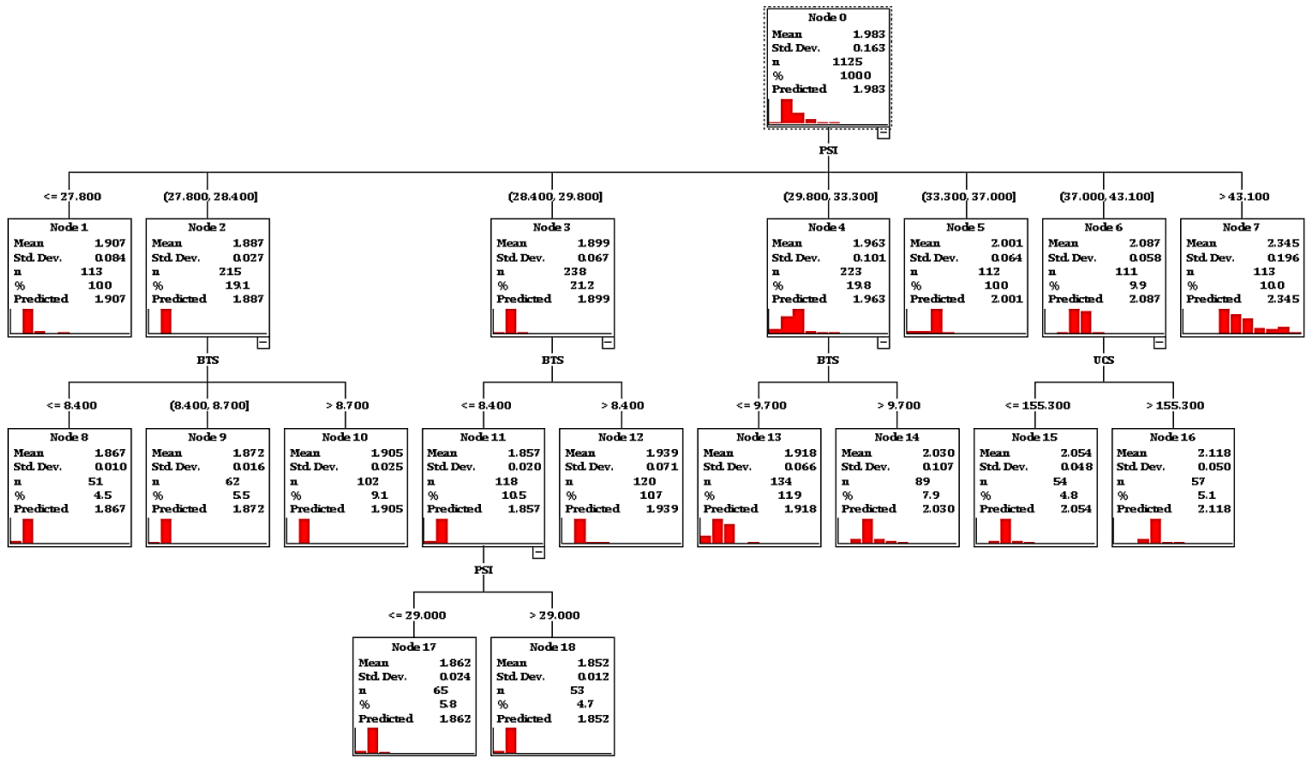



Fig. 14 A decision tree for determined factors based on the CHAID growing method

target factor (ROP) are shown in Fig. 13. A P-P plot compares the empirical cumulative distribution function. The line y = x is a good approximation to the best-fitting model. These two examples of diagrams both display credible outcomes.

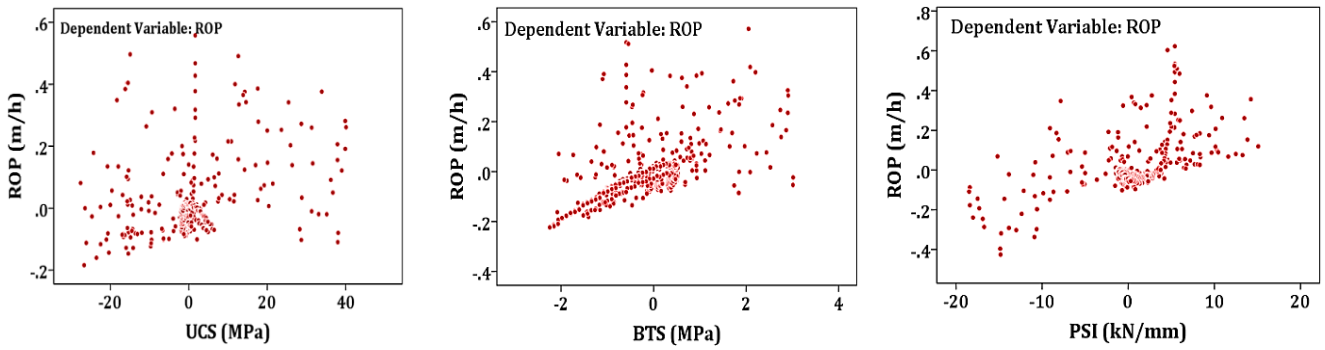
A result of the decision tree for determined factors based on the CHAID growing method is presented in Fig. 14 to identify groups and discover the relationship between parameters. A gain summary for nodes based on the application of the decision tree is presented in Table 4.

Table 4 Gain summary for nodes based on the application of decision tree

| Node | Node-by-Node | | | Cumulative | | |
|---|---|---|---|---|---|---|
| | N | Percent | Mean | N | Percent | Mean |
| 7 | 113 | 10.0% | 2.3447 | 113 | 10.0% | 2.3447 |
| 6 | 111 | 9.9% | 2.0870 | 224 | 19.9% | 2.2170 |
| 14 | 89 | 7.9% | 2.0303 | 313 | 27.8% | 2.1639 |
| 5 | 112 | 10.0% | 2.0005 | 425 | 37.8% | 2.1209 |
| 12 | 120 | 10.7% | 1.9392 | 545 | 48.4% | 2.0809 |
| 13 | 134 | 11.9% | 1.9176 | 679 | 60.4% | 2.0487 |
| 1 | 113 | 10.0% | 1.9068 | 792 | 70.4% | 2.0284 |
| 10 | 102 | 9.1% | 1.9055 | 894 | 79.5% | 2.0144 |
| 9 | 62 | 5.5% | 1.8721 | 956 | 85.0% | 2.0052 |
| 8 | 51 | 4.5% | 1.8667 | 1007 | 89.5% | 1.9982 |
| 15 | 65 | 5.8% | 1.8620 | 1072 | 95.3% | 1.9899 |
| 16 | 53 | 4.7% | 1.8519 | 1125 | 100.0% | 1.9834 |

Table 5 Regression coefficients between TBM-ROP and the geomechanical properties of rocks

| No. | Parameter | $R^2$ | Reg. Type | Relationship |
|---|---|---|---|---|
| 1 | UCS | 0.51 | Linear | ROP = 0.0092 UCS + 0.6334 |
| 2 | BTS | 0.35 | Exponential | $ROP = 1.2143\ E^{0.0533\ BTS}$ |
| 3 | PSI | 0.74 | Exponential | $ROP = 1.435\ E^{0.0098\ PSI}$ |




(a) (b) (c)

Fig. 15 3D diagrams based on three main factors (UCS, BTS, PSI) and TBM-ROP as a target factor

## 5. Methodologies and analysis

### *5.1 Developing new empirical equation*

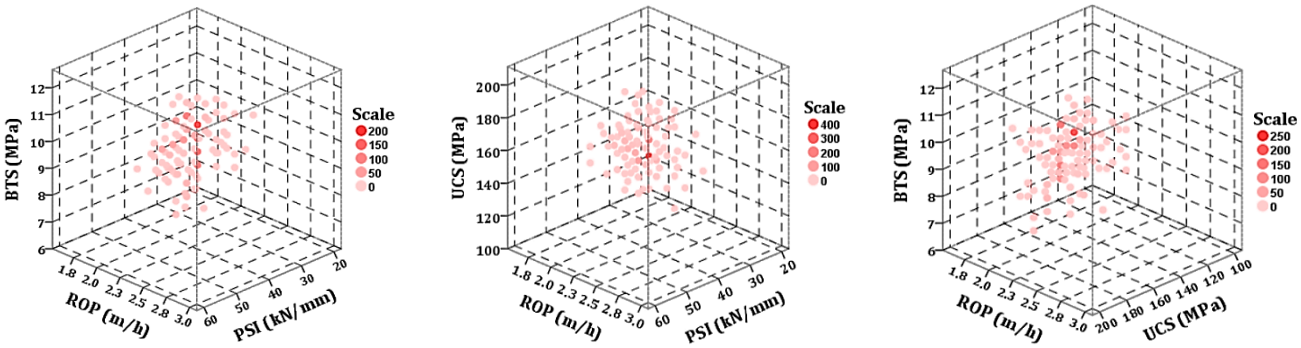
An empirical formula correlating TBM-ROP with characteristics of machine performance, as well as geological and rock strength, was introduced with the help of MVR analysis. Forward stepwise regression analysis is used to determine the impact of three geomechanical elements on TBM's performance: UCS, BTS, and PSI (Table 5). Forward stepwise regression analysis was used to determine the impact of each variable on TBM-ROP. Excluding DPW and α from the model seems to provide the best results after considering various combinations of the identified components.

Table 6 exhibits ANOVA analysis, coefficient correlations, and pertinent statistical metrics linked to the given formula for evaluating TBM-ROP. Additionally, within this table, you'll find the histogram of regression standardized residuals derived from the formulated empirical model for ROP, along with partial regression plots. New models to assess TBM-ROP may be developed using the ANOVA analysis and regression coefficients shown in Table 6 and Fig. 15. Table 6 displays the generated equation's $R^2$, which is 79%. According to the results, 79% of the variation present in the 1125 datasets is explained by the aforementioned regression model.

Fig. 16 showcases 3D diagrams illustrating the three key parameters (BTS, UCS, and PSI) crucial for TBM performance assessment. These visuals depict the scatterplot correlation between input variables and the target values through a three-dimensional perspective. Figs. 17 and 18 offer insights into the correlations and comparisons between the observed and computed ROPs,

Table 6 ANOVA and other statistical analysis of the generated model to estimate TBM-ROP

| Model Summary | | | | | |
|---|---|---|---|---|---|
| R | $R^2$ | Adjusted $R^2$ | Std. Error of the Estimate | | Durbin-Watson |
| 0.89 | 0.79 | 0.79 | 0.07 | | 0.97 |
| ANOVA | | | | | |
| | Sum of Squares | df | Mean Square | F | Sig. |
| Regression | 23.75 | 3 | 7.91 | 1450.39 | 0.00 |
| Residual | 6.11 | 1121 | 0.005 | | |
| Total | 29.87 | 1124 | | | |




(a) (b) (c)

Fig. 16 3D diagrams based on three main factors (UCS, BTS, PSI) and TBM-ROP as a target factor

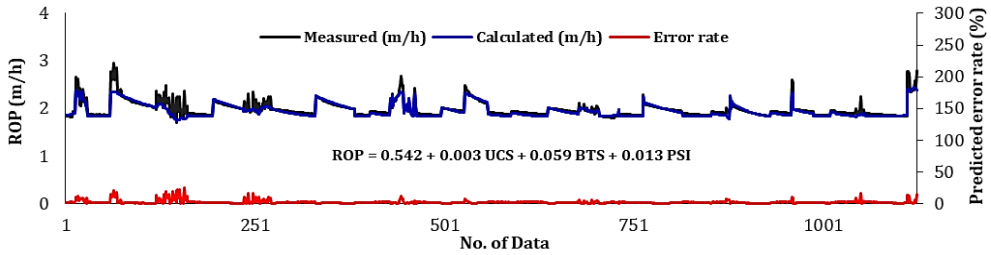



Fig. 17 Results of multi-variable regression. Comparison of actual and calculated values of TBM-ROP

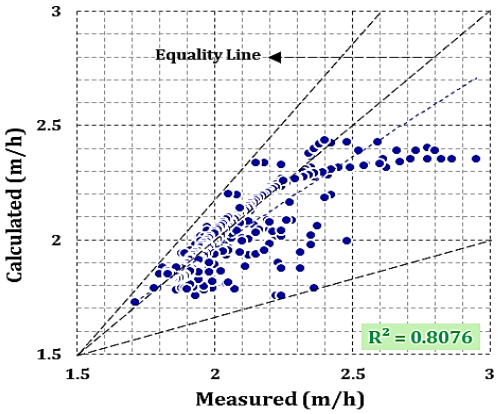
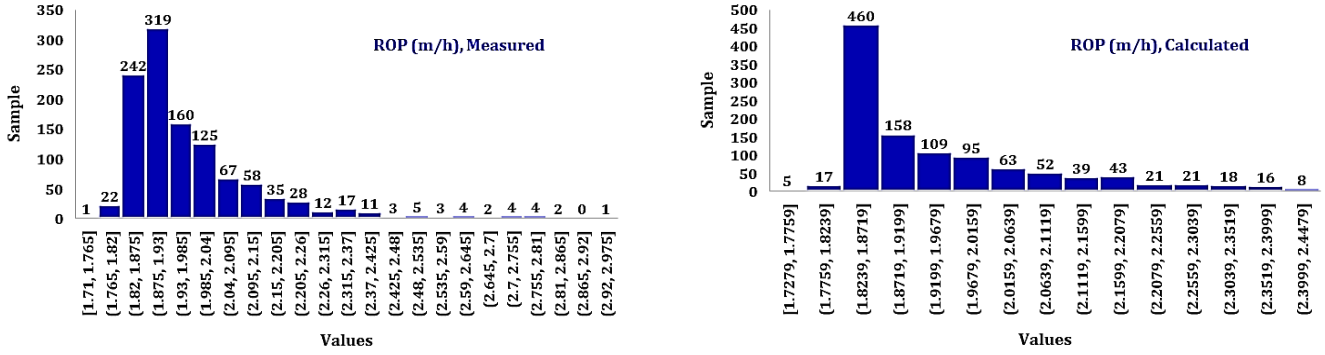



Fig. 18 Correlation between measured and calculated values based on MVR method

leveraging the developed empirical formula. Notably, the findings underscore a robust correlation between TBM-ROP, the focal parameter, and the predictive rock strength properties. The Pareto histogram of measured and calculated results is presented in Fig. 19.

The performance of the empirical model is examined through several evaluation criteria such as root mean square error (RMSE), mean square error (MSE), normalized root mean square error (NRMSE), root relative squared error (RRSE), relative squared error (RSE), mean absolute




(a) (b)

Fig. 19 Pareto graphs and distribution of measured and calculated values of TBM-ROP based on MVR analysis

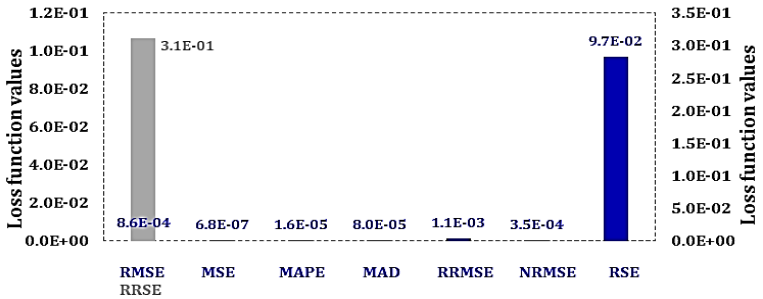



Fig. 20 The comparison between different types of evaluation metrics in terms of assessment of ROP calculated results

Table 7 The results of evaluation metrics in the development of empirical model to predict the TBM-ROP

| Evaluation metrics | Values |
|---|---|
| RMSE | 8.6E-04 |
| MSE | 6.8E-07 |
| MAPE | 1.6E-05 |
| MAD | 8.0E-05 |
| RRMSE | 1.1E-03 |
| NRMSE | 3.5E-04 |
| RSE | 9.7E-02 |
| RRSE | 3.1E-01 |
| $R^2$ | 0.79 |

deviation (MAD), and mean absolute percentage error (MAPE). Utilizing these seven factors as statistical benchmarks, we calculate the variance between predicted and actual values. The assessment metrics outcomes are consolidated in Table 7. Furthermore, Fig. 20 delineates the relationship between evaluation metric values and the trustworthiness of calculated ROP values. The range of changes in TBM-ROP and geomechanical data (UCS, BTS, PSI) and related equations are presented in Fig. 21. The summary of regression between TBM-ROP and different types of rock engineering properties based on rock fracture class is presented in Table 8. The classification of these fractures draws upon a fracture class system initially formulated by the Norwegian University of Science and Technology (NTNU) and subsequently refined by Bruland (1999). The detailed specifications for these fracture classes are outlined in Table 9. The single regression analysis presented in Table 8 highlights the crucial relationship between the ROP of a TBM and various rock engineering properties. In particular, the study focuses on the important rock mechanical parameters including UCS, BTS, and PSI of the rocks being drilled, as these properties have a significant impact on the performance and efficiency of the TBM. The data in Table 8 indicates a clear correlation between the ROP and the rock geomechanical specifications, allowing researchers and engineers to better understand and predict the behavior of TBMs in different sedimentary formations. By analyzing how these properties affect the penetration rate of the TBM, researchers can develop more accurate and efficient drilling techniques for sedimentary rocks. The correlation coefficients provided in Table 8 offer valuable insights into which rock properties have the most significant influence on the ROP of a TBM. By identifying these key factors, engineers can optimize drilling strategies and improve the overall performance of tunneling operations.

### *5.2 Results of artificial neural network algorithms*

Three distinct forms of deep ANNs have been deployed to construct networks aimed at forecasting TBM-ROP. These algorithms underwent training and refinement utilizing rock geomechanical parameters (UCS, BTS, and PSI) as inputs and TBM-ROP values as outputs. Notably, it's essential to acknowledge that the efficacy and precision of output produced by artificial intelligence networks can significantly fluctuate depending on the caliber and arrangement of the data employed for their training.

Table 8 Developed empirical formulas between TBM-ROP and rock engineering properties based on rock fracture class

| RFC | UCS | $R^2$ | BTS | $R^2$ | PSI | $R^2$ |
|---|---|---|---|---|---|---|
| 1 | ROP = 0.0098 UCS + 0.5344 | 0.59 | ROP = 0.1098 BTS + 0.9756 | 0.24 | ROP = 0.0204 PSI + 1.3168 | 0.75 |
| 2 | ROP = 0.0088 UCS + 0.6903 | 0.45 | ROP = 0.102 BTS + 1.0451 | 0.27 | ROP = 0.0204 PSI + 1.3168 | 0.75 |
| 3 | ROP = 0.0089 UCS + 0.6757 | 0.53 | ROP = 0.0986 BTS + 1.073 | 0.26 | ROP = 0.0207 PSI + 1.299 | 0.80 |
| 4 | ROP = 0.0099 UCS + 0.5249 | 0.63 | ROP = 0.1305 BTS + 0.7818 | 0.38 | ROP = 0.0208 PSI + 1.2975 | 0.80 |
| 5 | ROP = 0.01 UCS + 0.5143 | 0.55 | ROP = 0.1232 BTS + 0.8725 | 0.30 | ROP = 0.0218 PSI + 1.2725 | 0.71 |
| 6 | ROP = 0.0126 UCS + 0.1061 | 0.57 | ROP = 0.1116 BTS + 1.006 | 0.26 | ROP = 0.0274 PSI + 1.0856 | 0.69 |
| 7 | ROP = 0.0131 UCS + 0.0566 | 0.61 | ROP = 0.1267 BTS + 0.8483 | 0.19 | ROP = 0.0264 PSI + 1.138 | 0.80 |

Table 9 Specifications of rock fracture classes modified by Bruland (1999)

| Rock fracture class | The mean separation distance between planes of weakness (m) | Rock description |
|---|---|---|
| 1 | ∞ | A continuous expanse of rock characterized by its substantial massiveness, featuring minimal to absent joints or fissures. The interval boasts a spacing between planes of weakness exceeding 1.6 meters. |
| 2 | 1.6 | A solid rock interval with considerable massiveness, where the gap between planes of weakness measures 1.6 meters. |
| 3 | 0.8 | A rock interval exhibiting substantial massiveness, albeit relatively so, with a spacing between planes of weakness measuring 0.8 meters. |
| 4 | 0.4 | A rock interval characterized by intermittent discontinuities, featuring a spacing between planes of weakness measuring 0.4 meters. |
| 5 | 0.2 | A highly discontinuous rock mass, where the spacing between planes of weakness measures a mere 0.2 meters. |
| 6 | 0.1 | An extensively discontinuous rock mass, where the spacing between planes of weakness measures only 0.10 meters. |
| 7 | 0.05 or less | A rock mass exhibiting significant brecciation, characterized by closely spaced and intersecting weakness planes with intervals of 0.05 meters or less. This condition is often observed in areas of stress relief and fault breccia formations. |

Table 10 Hyper-parameters of the SRNN and GRU

| Hyper-parameters | Values and specifications |
|---|---|
| Optimizer | Adam |
| Epoch | 200-400 |
| RNNs | (none, 40-50) |
| Dense | (none, 1) |
| Batch-size | 64 |
| Validation split | 0.2 |
| Min. train timing | 4.82 s |
| Max. train timing | 15.38 s |

In recurrent neural networks (RNNs), the modeling process involves utilizing sequential data where the output depends on previous inputs. This allows RNNs to capture temporal dependencies. Key components include the input layer, hidden layer (s), and output layer. The hidden layer(s) have recurrent connections to retain memory of past inputs. Backpropagation, a common training algorithm in neural networks, including RNNs, calculates the gradient of the loss function with respect to the network's parameters. Parameters are updated in the opposite direction of the gradient to minimize the loss. In RNNs, backpropagation through time (BPTT) is often used, calculating gradients over multiple time steps.

The modeling process in RNNs and backpropagation neural networks involves setting up the network architecture, training the model with backpropagation, and fine-tuning parameters to optimize performance on the given task. Thus, before training the networks, the order of the collected data is changed randomly. The major hyper-parameters and structure of SRNN and GRU are summarized in Table 10.

By incorporating these key geomechanical properties into the database, the study aimed to enhance the accuracy and flexibility of regression and ML techniques in analyzing and predicting TBM performance in similar geotechnical conditions. The detailed data collection and analysis provided valuable insights into the factors influencing TBM performance in rock tunneling projects, facilitating informed decision-making and optimizing tunnel excavation processes

Figs. 22 and 23 display the outcomes from the predictor networks during both testing and training phases. These diagrams reveal a close alignment between the predicted TBM-ROP values and the actual measurements. Notably, the networks demonstrate a capability to accurately anticipate sudden shifts in target values. The correlation between the measured and predicted ROPs is depicted in Fig. 24.

Within this study, the precision of the employed supervised learning models is assessed through various

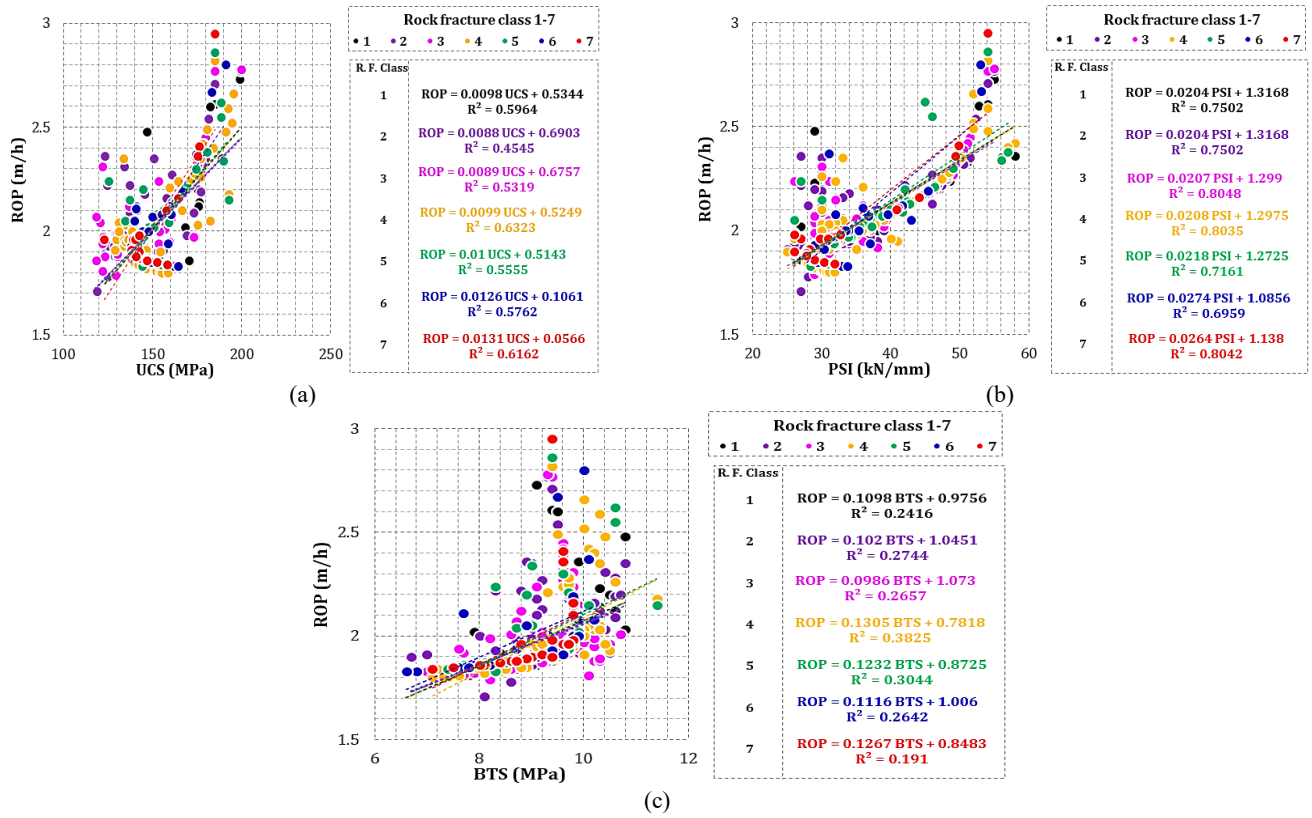


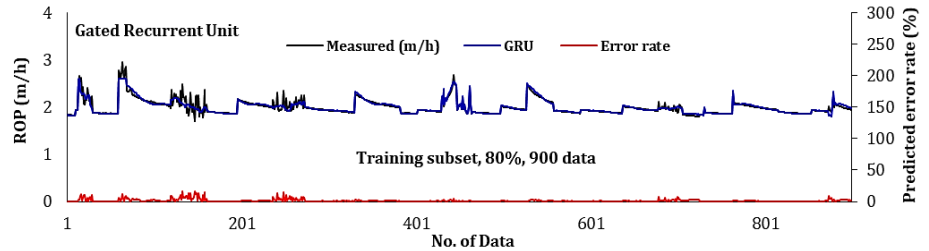

Fig. 21 The range of changes in TBM-ROP and geomechanical data (UCS, BTS, PSI) by considering seven rock fracture classes (1 to 7 classes) and related equations

Table 11 The results of evaluation metrics in the development of DL networks for prediction of TBM-ROP

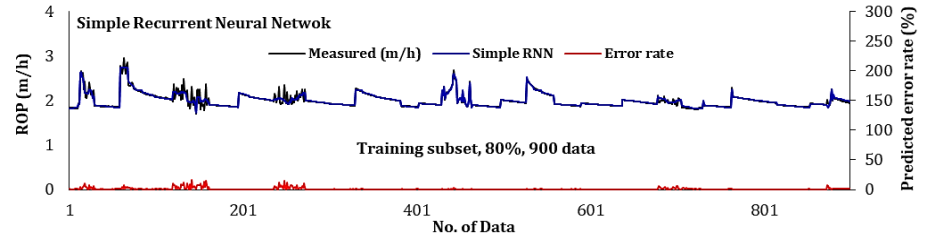
| | Loss Func. | GRU | Simple RNN | BPNN |
|---|---|---|---|---|
| Training subset | RMSE | 1.19E-04 | 5.65E-04 | 7.17E-06 |
| | MSE | 5.72E-09 | 1.37E-07 | 8.93E-10 |
| | MAPE | 2.29E-07 | 7.73E-06 | 2.31E-06 |
| | MAD | 1.90E-04 | 1.90E-04 | 1.83E-04 |
| | RRMSE | 1.55E-04 | 2.34E-04 | 2.14E-04 |
| | NRMSE | 4.59E-05 | 2.06E-04 | 2.61E-06 |
| | RSE | 2.46E-04 | 6.81E-03 | 3.74E-03 |
| | RRSE | 1.57E-02 | 8.25E-02 | 6.12E-02 |
| | Loss Func. | GRU | Simple RNN | BPNN |
| Testing subset | RMSE | 1.94E-03 | 2.53E-03 | 2.52E-03 |
| | MSE | 3.12E-06 | 6.37E-06 | 8.89E-06 |
| | MAPE | 7.48E-05 | 1.04E-04 | 1.30E-04 |
| | MAD | 2.85E-05 | 3.34E-05 | 2.50E-05 |
| | RRMSE | 1.96E-03 | 2.59E-03 | 2.97E-03 |
| | NRMSE | 6.76E-04 | 7.95E-04 | 7.99E-04 |
| | RSE | 1.07E+00 | 1.82E+00 | 2.19E+00 |
| | RRSE | 1.04E+00 | 1.35E+00 | 1.48E+00 |

evaluation metrics, statistical parameters, and loss functions, encompassing NRMSE, MSE, RRMSE, RMSE, RRSE, RSE, MAPE, and MAD. Each of these metrics serves as a statistical gauge for determining the deviation between predicted and measured values. The evaluation metric results for TBM-ROP across both training and testing subsets are consolidated in Table 11.

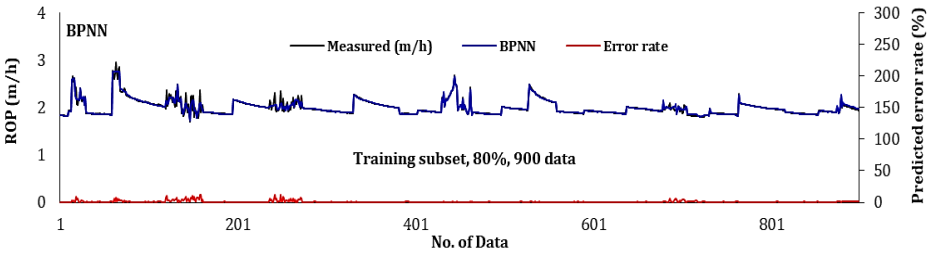
The consistent and dependable values of various evaluation metrics underscore the robust performance of the networks in accurately predicting TBM-ROP by leveraging defined rock engineering properties. This strong performance further reinforces the validity and applicability of the developed models in enhancing the understanding and optimization of TBM operations in rock tunneling




(a)

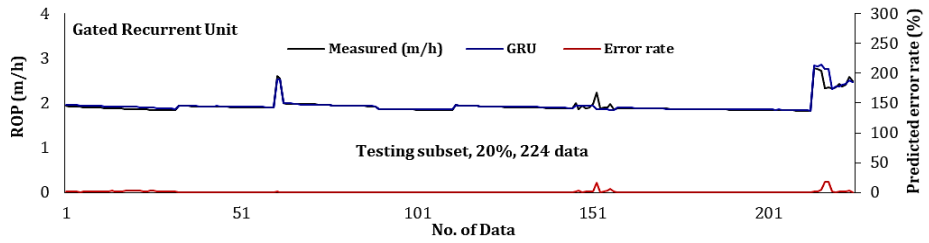



(b)

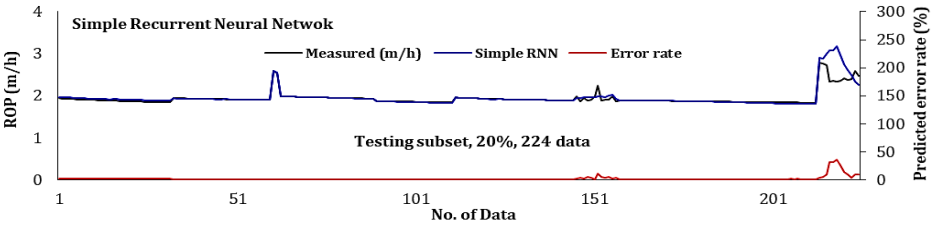



(c)

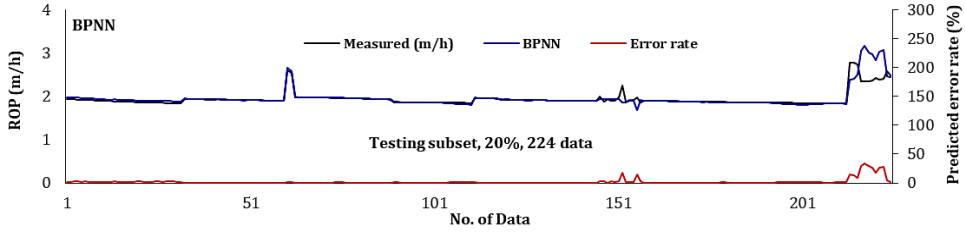
Fig. 22 Comparison between the predicted values of TBM-ROP and the measured ones based on the training dataset




(a)




(b)




(c)

Fig. 23 Comparison between the predicted values of TBM-ROP and the measured ones based on the testing dataset

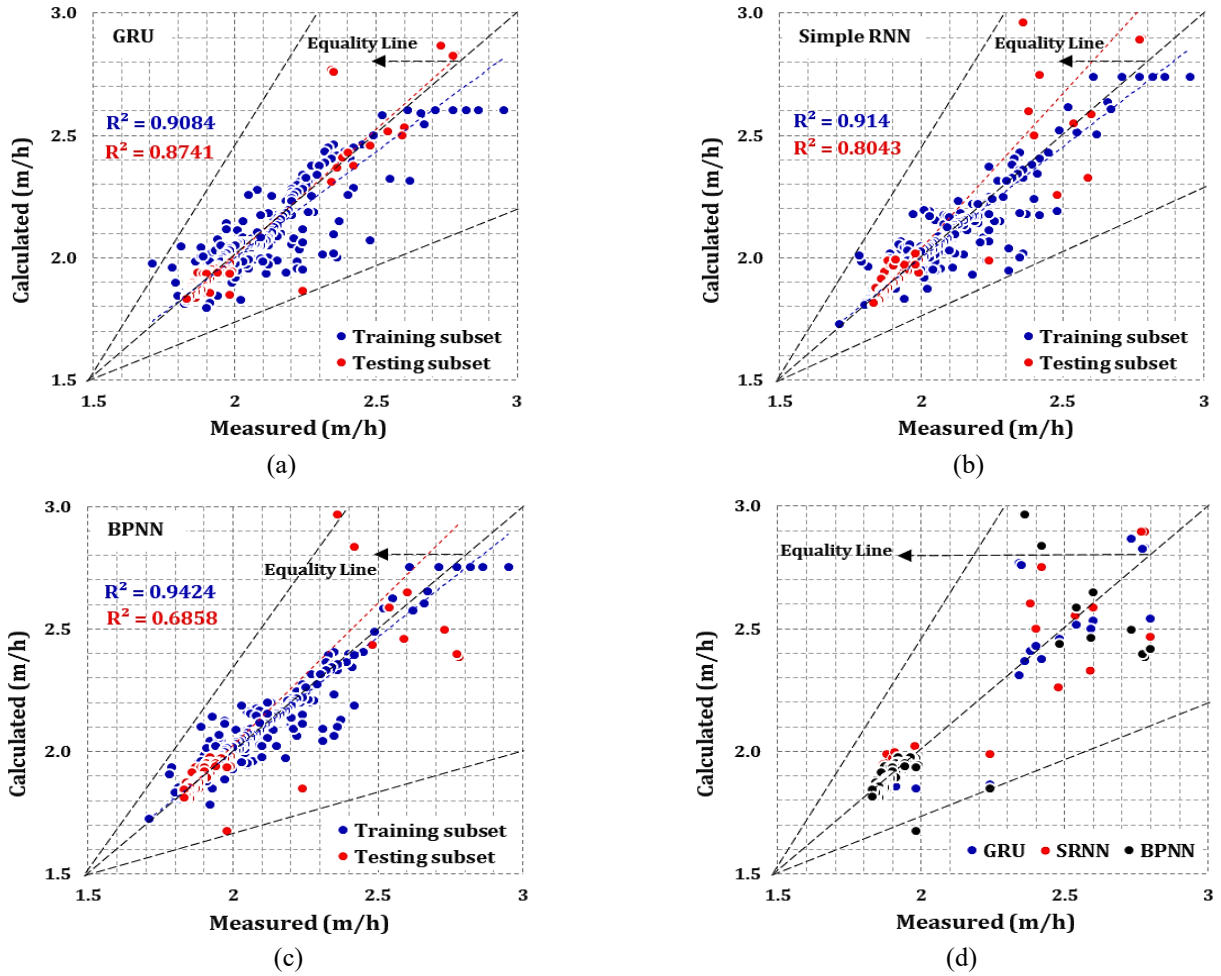



Fig. 24 The correlation between the measured and predicted TBM-ROP.

projects. It's evident that the values of various evaluation metrics are dependable, signaling the effective performance of the networks in forecasting TBM-ROP using specified rock engineering properties.

### *5.3 Comparison of applied methods*

The comparison of the MVR and DL networks using evaluation metrics for the evaluation of TBM-ROP values is presented in Fig. 25. As can be seen, all the models show accuracy and reliability in the obtained results. The value of $R^2$ for the developed statistical and DL networks is above 80%. It indicates that the applied models explain more than 80% of the total variance of the 1125 datasets. Also, the values of the evaluation metrics for all models are strongly close to zero, which shows reliable results. Among all the applied models, the GRU network is the most robust model to predict the TBM-ROP with high accuracy.

## 6. Conclusions

In this study, the assessment of TBM-ROP based on different rock geomechanical factors was investigated using supervised and unsupervised algorithms. A dataset was collected and generated from the Alborz service tunnel on the Tehran-Shomal motorway project (ASTSHM); 80% of the dataset was used for training the deep neural networks, and 20% of the dataset was used for testing the networks. The input factors included several rock engineering properties such as UCS, BTS, PSI, RFC, DPW, and α. The accuracy and performance of the developed methods were evaluated using nine evaluation metrics, including RMSE, MSE, RRMSE, NRMSE, RRSE, RSE, MAPE, MAD, and $R^2$. The values of evaluation metrics were strongly close to zero, which showed the accuracy of the results (SRNN: MSE = 6.37E-06, MAD = 3.34E-05; GRU: MSE = 3.12E-06, MAD = 2.85E-05; BPNN: MSE = 8.89E-06, MAD = 2.50E-05; MVR: MSE = 6.8E-07, MAD = 8.0E-05). The results indicated that among the four applied models, the GRU network was the most robust model in predicting the TBM-ROP and produced the most accurate and reliable results. This study also revealed that the three intact rock mass engineering properties, including BTS, UCS, and DPW, have a significant impact on the TBM-ROP. It should be noted that the applied models and the correspondence results in this work to predict the TBM-ROP are based on a database of volcanic and sedimentary rocks

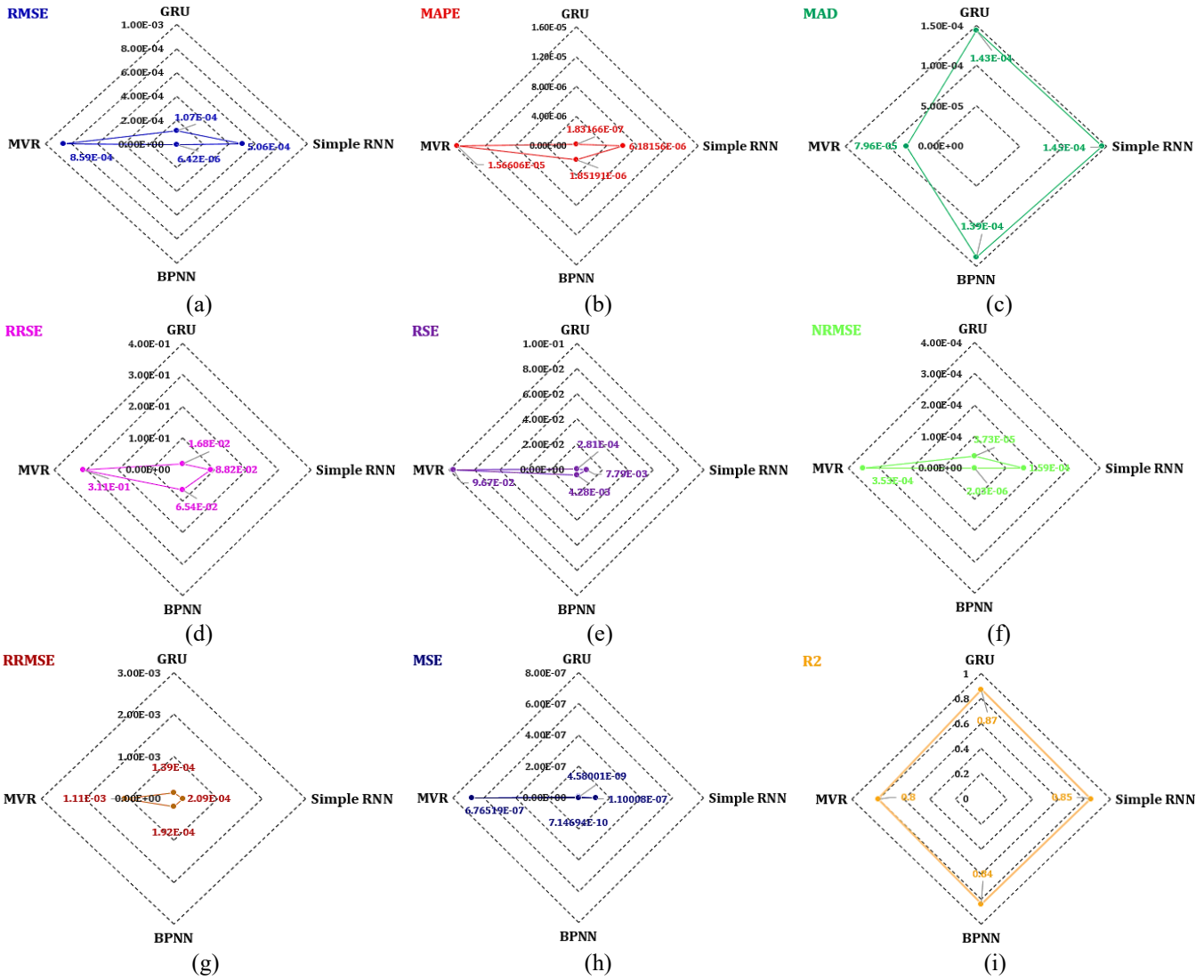



Fig. 25 Performance comparison of the DL and MVR models for evaluation the results of TBM-ROP

The significance of this work is that it can address many of the operator's concerns about TBM performance by controlling the best boring process and using the TBMs in rock tunneling with similar geomechanical and geological properties (sedimentary rocks: Sandstone, Limestone, Tuff, Gypsum, Shale).

For more clarification and indicating the limitation of this study it should be mention that the sedimentary and volcanic rock units in this study include sandstone, limestone, tuff, gypsum, and shale. To develop and present supervised models for predicting TBM performance, access to necessary training data is crucial. Failure to acquire this data may require alternative techniques like numerical simulations. As we progress into the data-driven era, improving supervised models for predicting TBM performance by harnessing more comprehensive datasets is essential. Acknowledging the limited studies on ML methods' predictive capabilities for ROP based on rock geomechanical data is important. Enhancing the predictive capabilities of ML methods must address these limitations. Many studies focus on specific conditions, resulting in models lacking comprehensiveness. Enriching existing databases rather than changing previous models is more beneficial. This would aid in developing a versatile model accurately estimating TBM performance for different rock types. However, work is still needed to create a model effectively capturing rock geomechanical factors of various rock types with diverse strength based on boring machine data. Constructing such an ML-based model requires a comprehensive database, necessitating global research collaboration.

This research underscores the constraints linked to the database utilized for assessing TBM performance. To augment the inclusivity and relevance of ML-driven and unsupervised models in this investigation, expanding the database holds paramount importance. The authors are proactively enhancing the database by incorporating data gleaned from supplementary laboratory tests, in-situ analyses, TBM records, and encompassing a broader spectrum of conditions and tunneling endeavors. The development of ML and DL models grounded on this expanded dataset is anticipated to yield notably more

impactful outcomes. Moreover, these proposed models could prove applicable to analogous ground and TBM tunneling scenarios.

## Acknowledgments

This study is supported via funding from Prince Satam bin Abdulaziz University project number (PSAU/2024/R/1445).

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large Groups RGP. 2/357/44.

## References

Adoko, A.C., Gokceoglu, C. and Yagiz, S. (2017), "Bayesian prediction of TBM penetration rate in rock mass", *Eng.Geol.*, **226**, 245-256. https://doi.org/10.1016/j.enggeo.2017.06.014.

Ates, U., Bilgin, N. and Copur, H. (2014), "Estimating torque, thrust and other design parameters of different type TBMs with some criticism to TBMs used in Turkish tunneling projects", *Tunn. Undergr. Sp. Tech.*, **40**, 46-63. https://doi.org/10.1016/j.tust.2013.09.004

Carter, T.G. and Marinos, V. (2020), "Putting geological focus back into rock engineering design", *Rock Mech. Rock Eng.*, **53**(10), 4487-4508. https://doi.org/10.1007/s00603-020-02177-1.

Cho, K., Merrienboer, B. van, Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014), "Learning phrase representations using RNN encoder-decoder for statistical machine translation", *Computation and Language*. https://doi.org/10.48550/arXiv.1406.1078.

Elhaik, E. (2022), "Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated", *Scientific Reports*, **12**(1), 14683. https://doi.org/10.1038/s41598-022-14395-4.

Elmo, D. and Stead, D. (2021), "The role of behavioural factors and cognitive biases in rock engineering", *Rock Mech. Rock Eng.*, **54**(5), 2109-2128. https://doi.org/10.1007/s00603-021-02385-3.

Farrokh, E., Rostami, J. and Laughton, C. (2012), "Study of various models for estimation of penetration rate of hard rock TBMs", *Tunn. Undergr. Sp. Tech.*, **30**, 110-123. https://doi.org/10.1016/j.tust.2012.02.012.

Gokceoglu, C. (2022), "Assessment of rate of penetration of a tunnel boring machine in the longest railway tunnel of Turkey", *SN Appl. Sci.*, **4**(1), 19. https://doi.org/10.1007/s42452-021-04903-y.

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

He, H., Wang, S., Shen, W. and Zhang, W. (2023), "The influence of pipe-jacking tunneling on deformation of existing tunnels in soft soils and the effectiveness of protection measures", *Transport. Geotech.*, **42**, 101061. https://doi.org/10.1016/j.trgeo.2023.101061.

Hu, D., Li, Y., Yang, X., Liang, X., Zhang, K. and Liang, X. (2023), "Experiment and application of NATM tunnel deformation monitoring based on 3D laser scanning", *Struct. Control Health Monit.*, **2023**, 1-13. https://doi.org/10.1155/2023/3341788.

Jolliffe, I.T. and Cadima, J. (2016), "Principal component analysis: a review and recent developments", *Philos. T. Roy. Soc. A: Math. Phys. Eng. Sci.*, **374**(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202.

Jung, J.H., Chung, H., Kwon, Y.S. and Lee, I.M. (2019), "An ANN to predict ground condition ahead of tunnel face using TBM operational data", *KSCE J. Civil Eng.*, **23**(7), 3200-3206. https://doi.org/10.1007/s12205-019-1460-9.

Kalnins, A. (2022), "When does multicollinearity bias coefficients and cause type 1 errors? A reconciliation of Lindner, Puck, and Verbeke (2020) with Kalnins (2018)", *J. Int. Business Studies*, **53**(7), 1536-1548. https://doi.org/10.1057/s41267-022-00531-9.

Liu, B., Wang, R., Guan, Z., Li, J., Xu, Z., Guo, X. and Wang, Y. (2019), "Improved support vector regression models for predicting rock mass parameters using tunnel boring machine driving data", *Tunn. Undergr. Sp. Tech.*, **91**, 102958. https://doi.org/10.1016/j.tust.2019.04.014.

Liu, J., Ren, J. and Guo, W. (2015), "Thrust and torque characteristics based on a new cutter-head load model", *Chinese J. Mech. Eng.*, **28**(4), 801-809. https://doi.org/10.3901/CJME.2015.0504.066.

Liu, W., Liang, J. and Xu, T. (2023), "Tunnelling-induced ground deformation subjected to the behavior of tail grouting materials", *Tunn. Undergr. Sp. Tech.*, **140**, 105253. https://doi.org/10.1016/j.tust.2023.105253.

Mahmoodzadeh, A., Nejati, H.R., Mohammadi, M., Hashim Ibrahim, H., Rashidi, S. and Ahmed Rashid, T. (2022), "Forecasting tunnel boring machine penetration rate using LSTM deep neural network optimized by grey wolf optimization algorithm", *Exp. Syst. Appl.*, **209**, 118303. https://doi.org/10.1016/j.eswa.2022.118303.

Shahrour, I. and Zhang, W. (2021), "Use of soft computing techniques for tunneling optimization of tunnel boring machines", *Underground Space*, **6**(3), 233-239. https://doi.org/10.1016/j.undsp.2019.12.001.

Shao, C., Li, X. and Su, H. (2013), Performance prediction of hard rock TBM based on extreme learning machine, 409-416. https://doi.org/10.1007/978-3-642-40849-6_40.

Shi, M.L., Lv, L. and Xu, L. (2023a), "A multi-fidelity surrogate model based on extreme support vector regression: fusing different fidelity data for engineering design", *Eng. Comput.*, **40**(2), 473-493. https://doi.org/10.1108/EC-10-2021-0583.

Shi, M., Hu, W., Li, M., Zhang, J., Song, X. and Sun, W. (2023b), "Ensemble regression based on polynomial regression-based decision tree and its application in the in-situ data of tunnel boring machine", *Mech. Syst. Signal Pr.*, **188**, 110022. https://doi.org/10.1016/j.ymssp.2022.110022.

Su, Y., Wang, J., Li, D., Wang, X., Hu, L., Yao, Y. and Kang, Y. (2023), "End-to-end deep learning model for underground utilities localization using GPR", *Automat. Constr.*, **149**, 104776. https://doi.org/10.1016/j.autcon.2023.104776.

Wang, X., Lu, H., Wei, X., Wei, G., Behbahani, S.S. and Iseley, T. (2020), "Application of artificial neural network in yunnel engineering: A systematic review", *IEEE Access*, **8**, 119527-119543. https://doi.org/10.1109/ACCESS.2020.3004995.

Yan, T., Xu, R., Sun, S.H., Hou, Z.K. and Feng, J.Y. (2024), "A real-time intelligent lithology identification method based on a dynamic felling strategy weighted random forest algorithm", *Petroleum Sci.*, **21**(2), 1135-1148. https://doi.org/10.1016/j.petsci.2023.09.011.

Yang, S.Q., Chen, M., Fang, G., Wang, Y.C., Meng, B., Li, Y.H. and Jing, H.W. (2018), "Physical experiment and numerical modelling of tunnel excavation in slanted upper-soft and lower-hard strata", *Tunn. Undergr. Sp. Tech.*, **82**, 248-264. https://doi.org/10.1016/j.tust.2018.08.049.

Yin, H., Wu, Q., Yin, S., Dong, S., Dai, Z. and Soltanian, M.R. (2023), "Predicting mine water inrush accidents based on water level anomalies of borehole groups using long short-term

memory and isolation forest", *J. Hydrology*, **616**, 128813. https://doi.org/10.1016/j.jhydrol.2022.128813.

Zhao, N., Li, D.Q., Gu, S.X. and Du, W. (2024), "Analytical fragility relation for buried cast iron pipelines with lead-caulked joints based on machine learning algorithms", *Earthq. Spectra*, **40**(1), 566-583. https://doi.org/10.1177/87552930231209195.

Zhou, J., Qiu, Y., Armaghani, D.J., Zhang, W., Li, C., Zhu, S. and Tarinejad, R. (2021), "Predicting TBM penetration rate in hard rock condition: A comparative study among six XGB-based metaheuristic techniques", *Geosci. Frontiers*, **12**(3), 101091. https://doi.org/10.1016/j.gsf.2020.09.020.

*IC*