Geomechanics and Engineering, *Vol. 39, No. 1 (2024) 27-41*
https://doi.org/10.12989/gae.2024.39.1.027

Estimation of tunnel boring machine penetration rate: Application of long-short-term memory and meta-heuristic optimization algorithms

Mengran Xu*[1,2], Arsalan Mahmoodzadeh[3], Abdelkader Mabrouk[4], Hawkar Hashim Ibrahim[5], Yasser Alashker[6,7] and Adil Hussein Mohammed[8]

[1]*College of Civil Engineering and Architecture, Xinjiang University, Urumqi 830046, Xinjiang, China*
[2]*College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310000, Zhejiang, China*
[3]*IRO, Civil Engineering Department, University of Halabja, Halabja, 46018, Iraq*
[4]*Department of Civil Engineering, College of Engineering, Northern Border University, Arar 73222, Saudi Arabia*
[5]*Department of Civil Engineering, College of Engineering, Salahaddin University-Erbil, 44002 Erbil, Kurdistan Region, Iraq*
[6]*Department of Civil Engineering, College of Engineering, King Khalid University, PO Box 394, Abha 61411 Kingdom of Saudi Arabia*
[7]*Center for Engineering and Technology Innovations, King Khalid University, Abha 61421, Saudi Arabia*
[8]*Department of Communication and Computer Engineering, Faculty of Engineering, Cihan University, Kurdistan Region, Iraq*

(*Received February 14, 2023, Revised August 17, 2024, Accepted August 27, 2024*)

Abstract. Accurately estimating the performance of tunnel boring machines (TBMs) is crucial for mitigating the substantial financial risks and complexities associated with tunnel construction. Machine learning (ML) techniques have emerged as powerful tools for predicting non-linear time series data. In this research, six advanced meta-heuristic optimization algorithms based on long short-term memory (LSTM) networks were developed to predict TBM penetration rate (TBM-PR). The study utilized 1125 datasets, partitioned into 20% for testing, 70% for training, and 10% for validation, incorporating six key input parameters influencing TBM-PR. The performances of these LSTM-based models were rigorously compared using a suite of statistical evaluation metrics. The results underscored the profound impact of optimization algorithms on prediction accuracy. Among the models tested, the LSTM optimized by the particle swarm optimization (PSO) algorithm emerged as the most robust predictor of TBM-PR. Sensitivity analysis further revealed that the orientation of discontinuities, specifically the alpha angle (α), exerted the greatest influence on the model's predictions. This research is significant in that it addresses critical concerns of TBM manufacturers and operators, offering a reliable predictive tool adaptable to varying geological conditions.

Keywords: long-short-term memory; machine learning; machine penetration rate; metaheuristic optimization; sensitivity analysis; tunnel boring

1. Introduction

Numerous digitalization methodologies have been introduced by researchers aiming to enhance the classification of rock types, facilitate access to geological data, and improve tunnel design analysis (Mahmoodzadeh *et al*. 2022a, b, c, Pan *et al.* 2021, Han *et al*. 2021). Among the most pressing engineering challenges where digital innovations can offer substantial time and cost efficiencies is tunnel construction. The increasing urban population and the consequent demand for improved intercity transportation and communication infrastructure have significantly heightened the need for tunnels. Consequently, considerable attention and resources have been directed towards the development of advanced tunneling technologies. Mechanized tunneling, particularly through the use of Tunnel Boring Machines (TBMs), represents a pivotal approach in this domain. Predicting TBM performance remains a critical area of study, with considerable scholarly focus (Flor *et al.* 2023, Ghorbani and Yagiz 2024, Gokceoglu 2022, Shan *et al.* 2022). Essential metrics for assessing machine performance during tunnel excavation include the TBM Penetration Rate (TBM-PR) and the TBM Advance Rate (TBM-AR). Precise prediction of these parameters is crucial for design optimization and for minimizing uncertainties such as construction delays and budget overruns (Jing *et al.* 2021, Li *et al.* 2024, Liu *et al*. 2022, Yalei *et al.* 2024).

In this paper, we categorize the existing methodologies for TBM performance prediction into three primary groups: (i) theoretical and empirical methods rooted in laboratory experiments, field performance data, cutting forces, and rock properties (Carter and Marinos 2020); (ii) statistical methods grounded in mathematical principles (Gong and Zhao 2009); and (iii) computational methods leveraging artificial intelligence (AI) and machine learning (ML) techniques (Koopialipoor et al., 2020).

For the first category, a notable approach to TBM-PR prediction was developed by Ozdemir (1977), who utilized large-scale laboratory cutting tests combined with multiple regression analyses, resulting in the Colorado School of Mines Model (CSM)—a pivotal predictive tool. Another significant model was proposed by the Norwegian University of Science and Technology (NTNU), which applied regression analysis to both driving parameters and

*Corresponding author, Ph.D.
E-mail: xumengran@zju.edu.cn

http://www.techno-press.org/?journal=gae&subpage=7
ISSN: 2005-307X (Print), 2092-6219 (Online)

rock mass characteristics (Lislerud 1988). The literature also discusses other conventional methods for forecasting TBM performance. For instance, Hamidi *et al.* (2010) explored the relationship between the Field Penetration Index (FPI) of TBMs and the five key parameters of the Rock Mass Rating (RMR) system, confirming the correlation between FPI, Unconfined Compressive Strength (UCS), and discontinuity orientation. Additionally, other studies have utilized classifications such as the Geological Strength Index (GSI), Q-TBM, and Rock Mass Drillability Index (RMDI) to estimate TBM performance (Frough *et al.* 2015). However, while these models—rooted in experimental and theoretical frameworks—offer valuable insights, they typically consider a limited set of factors and often fail to account for variations in operational conditions or material properties, leading to potential inaccuracies in TBM efficiency predictions (Yagiz 2015).

In response to these limitations, several researchers have turned to statistical methodologies, the second category, for TBM performance forecasting. These methods, based on rigorous mathematical models, have been both utilized and developed by various scholars. For instance, Yagiz (2008) and Yagiz *et al.* (2009) employed linear and non-linear multiple regression techniques to predict TBM-PR. Hassanpour *et al.* (2011) demonstrated significant correlations between FPI and rock mass properties such as Rock Quality Designation (RQD), RMR, UCS, and joint spacing. Their research concluded that RQD and UCS are the most influential factors in predicting FPI. Additionally, Rayatdust *et al.* (2012) used data from the Alborz service tunnel construction to develop a linear multiple regression equation for accurately predicting TBM-PR based on UCS, volumetric joint count, and the angle between weakness planes and the tunnel axis. While these statistical methods are prevalent, they are often insufficient for describing complex, non-linear systems, particularly when dealing with large and geographically dispersed datasets.

The third category encompasses AI and ML techniques, which have been increasingly employed to address the complex task of identifying the factors influencing TBM performance and its parameters. Other AI strategies, such as Particle Swarm Optimization (PSO), fuzzy logic, and Artificial Neural Networks (ANNs), have also been effectively applied to this problem (Yagiz and Karahan 2011). Salimi *et al*. (2016) utilized Adaptive Neural-Fuzzy Inference System (ANFIS) and SVR methods for TBM-PR forecasting, with SVR proving to be the more robust model. Fattahi (2016) introduced a hybrid model integrating ANFIS with fuzzy C-means clustering for TBM-PR prediction. Liu *et al.* (2019) introduced a specialized version of SVR, termed Single-Purpose SVR, which outperformed generic SVR models in predicting TBM performance. More recently, Li *et al.* (2021) proposed an LSTM neural network for predicting TBM performance metrics such as total thrust and cutterhead torque, demonstrating the importance of input parameter relevance using a random forest technique. Zhou *et al.* (2021) developed novel hybrid predictive models based on Extreme Gradient Boosting (XGB) for TBM-PR forecasting, with their investigation concluding that the PSO-XGB model provided the best predictive accuracy. Overall, AI and ML methods are exceptionally flexible, allowing researchers and engineers to develop more precise and reliable solutions for highly complex, non-linear problems (Yagiz and Karahan 2011).

This study aims to develop an LSTM-based model for predicting TBM-PR using 1125 datasets collected from the Alborz Service Tunnel in Iran. The model leverages meta-heuristic algorithms, including PSO, Grey Wolf Optimization (GWO), Imperialist Competitive Algorithm (ICA), Shuffled Frog Leaping Algorithm (SFLA), and Genetic Algorithm (GA). To prevent overfitting, the dropout method was employed during model training. The proposed hybrid prediction models achieved high-performance levels in TBM-PR estimation, as validated by multiple statistical assessment metrics. The outcomes of this study could inform the development of TBM-PR early warning systems and serve as a valuable reference for researchers interested in applying LSTM models to other geotechnical engineering challenges.

2. Methodology

2.1 LSTM algorithm

LSTM networks represent a sophisticated class of artificial neural networks, playing a pivotal role in the realms of deep learning and artificial intelligence. Unlike conventional feedforward neural networks, LSTMs integrate feedback connections, which enable them to process not only isolated data points—such as images—but also entire sequences of data, including speech and video. This characteristic renders LSTMs particularly suited for tasks that involve sequential information, such as continuous handwriting recognition, speech recognition, machine translation, robotic control, video game AI, and healthcare applications. Since its introduction, LSTM has been one of the most cited neural network architectures, fundamentally shaping the field of neural computation throughout the late 20th century (Hochreiter and Schmidhuber 1997).

The nomenclature "long short-term memory" reflects the network's dual ability to maintain information over both short and extended durations. This dual capability mirrors the brain's mechanisms, where long-term memories are encoded through lasting changes in synaptic strengths, while short-term memories are captured through transient patterns of neural activation. Similarly, in LSTM networks, connection weights and biases adjust during training, capturing long-term dependencies, while activation patterns shift at each timestep to encode short-term information (Hochreiter and Schmidhuber 1997). The architecture of LSTM was explicitly designed to empower recurrent neural networks (RNNs) with the ability to retain short-term memory over potentially thousands of timesteps (Gers and Schmidhuber 2000).

A standard LSTM unit is composed of several key components: cells, input gates, output gates, and forget gates. The cell serves as a memory element that can store

values over arbitrary periods, while the gates regulate the flow of information into and out of the cell, effectively controlling the memory's retention and utilization (Gers and Schmidhuber 2000).

LSTM networks are exceptionally well-suited for the classification, processing, and prediction of time series data, particularly in scenarios where there are significant delays between relevant events within the sequence. The development of LSTM networks was driven by the need to address the vanishing gradient problem, a challenge that often hampers the training of conventional RNNs. Due to their architecture, LSTMs demonstrate superior performance across a range of sequence learning tasks, outperforming traditional RNNs, hidden Markov models, and other sequential processing techniques, especially in cases where the temporal gap between critical inputs is extensive (Hochreiter and Schmidhuber1997, Gers and Schmidhuber 2000).

2.2 PSO algorithm

PSO is a robust optimization technique inspired by the collective behavior of swarms. This method utilizes a population of particles to explore the solution space, each representing a candidate solution. The particles navigate the search space, progressively converging towards the optimal solution. Each particle is guided by both its own best-known position (pbest) and the global best position found by the swarm (gbest), inherently assuming its own optimum as superior to others. The movement of each particle is dynamically adjusted based on its current velocity, the distance to its pbest, and the distance to the gbest, as described by the equations provided by Kennedy and Eberhart (1995).

$$v_i^{t+1} = wv_i^t + c_1 \times rand \times (pbest_i - \chi_i^t) + c_2 \times rand \times (gbest - \chi_i^t) \quad (1)$$

$$\chi_i^{t+1} = \chi_i^t + v_i^{t+1} \quad (2)$$

In each iteration, several critical variables come into play: the weighting coefficient (w), the current position of the particle (χ_i^t), the acceleration coefficient (c_j), and a random variable rand (0,1), all of which are functions of the iteration time (t). Additionally, within this algorithm, *gbest* represents the best global position identified by the swarm thus far, while $pbest_i$ denotes the optimal position achieved by particle i at iteration t.

2.3 GWO algorithm

To mathematically represent the hierarchy of wolves, the Grey Wolf Optimizer (GWO) employs a predatory-based optimization model, wherein the leading agent, or the most proficient hunter, is designated as alpha (α) (Mirjalili *et al*. 2014). The subsequent ranks in this hierarchy are occupied by beta (β) and delta (δ) wolves, with omega (ω) representing the lowest tier. In the GWO framework, the optimization process is governed by these hierarchical classifications. The surrounding process is quantitatively described by Eq. (3) (Mirjalili *et al*. 2014).

$$pos(t+1) = pos_p(t) - a\left|c.pos_p(t) - pos(t)\right| \quad (3)$$

where pos_p represents the position of the prey and $pos_p(t)$ denotes the position of the candidate solution at iteration t. The vectors a and c are defined by Eqs. (4) and (5), respectively (Mirjalili *et al*. 2014).

$$a = 2.f.rand_1 - a \quad (4)$$

$$c = 2.rand_2 \quad (5)$$

where f decreases progressively from two to zero over the course of iterations. The random variables $rand_1$ and $rand_2$ are constrained within the interval (0,1). Initially, wolves are randomly generated by the GWO. Subsequently, these wolves are classified into one of three predefined categories, each denoted mathematically. The dynamics of the prey's pursuit can be described by the following equations (Mirjalili *et al.* 2014).

$$Pos(t+1) = \frac{pos_1 + pos_2 + pos_3}{3}$$

$$\begin{aligned} pos_1 &= pos_\alpha - a_1(d_\alpha), \\ pos_2 &= pos_\beta - a_2(d_\beta), \quad pos_3 \\ &= pos_\delta - a_3(d_\delta) \end{aligned} \quad (6)$$

$$d_\alpha = |c_1 pos_\alpha - pos|, \quad d_\beta = \left|c_2 pos_\beta - pos\right|, \quad d_\delta = |c_3 pos_\delta - pos|$$

Wolves possess the remarkable capability to track and encircle their prey. Typically, the task of hunting is undertaken by the α wolves, while those in the beta and delta stages contribute supportively. The challenge lies in the search space's incomplete understanding of the prey, which represents the best solution. Mathematically, it is posited that the wolves with the most profound knowledge of prey locations are those with extensive hunting experience. Consequently, while certain agents may need to reposition themselves to align with a specific wolf's location, the top three solutions are preserved and prioritized.

2.4 ICA algorithm

The Imperialist Competitive Algorithm (ICA) is an optimization technique inspired by swarm dynamics, wherein each member of the population represents a distinct country. Initially, the most promising countries are designated as imperialists (Atashpaz-Gargari and Lucas 2007). The algorithm then employs mechanisms of colonial integration, revolutionary dynamics, imperialist-colonial exchanges, and competition among imperialists to generate novel solutions once an empire is established.

The ICA process unfolds as follows:

1. Initialization: Create an initial population and evaluate the cost for each individual. Sort all solutions in descending order of their cost, select the best-performing ones as imperialists, and allocate the remaining countries to these imperialists.

2. Formation of the Initial Empire: Implement colony assimilation and instigate revolutions within select colonies.
3. Repositioning: If feasible, swap the positions of colonies and imperialists within each empire.
4. Imperialist Competition: Facilitate competition among imperialists to drive the optimization process.
5. Empire Disintegration: Dissolve the empire without disrupting the status of other countries.
6. Termination Check: Repeat the process if the termination criteria are not met; otherwise, conclude the search.

The objective function evaluates the true cost of each country, with lower costs indicating better solutions. Imperialists (N_{imper}) select the most cost-effective solutions, while the remaining countries are governed by colonial powers. There are N_{colon} colonies, and these are assigned to imperialists based on their relative power and cost, resulting in empires that are subsequently normalized to reflect these adjusted costs.

$$U_k = \left| \frac{C_k}{\sum_{i=1}^{N_{imper}} C_i} \right| \tag{7}$$

The normalized cost associated with imperialist i is denoted as $cost_i$ with $C_k = \max(cost_i)$ representing the maximum cost among all imperialists, and $round(U_k, N_{colon})$ transforms a fractional value into its nearest integer, determining the initial number of colonies assigned to the imperialist country as k. Consequently, C_k denotes the aggregate of colonies under the control of an imperialist.

During the assimilation phase, each colony within an empire moves towards its respective imperialist. This movement is governed by a random distribution over the interval $(0, z \times p]$, $p \in (1,2)$, where p falls within the range (1,2) and z represents the distance between the colony and the imperialist. The parameter p>1 ensures that the colony progresses in the direction of the imperialist, inducing a deviation from a straight trajectory. This deviation, denoted as φ, follows a random distribution within the range $[-\varphi, \varphi]$, where φ is a stochastic variable.

The revolution phase addresses the need for positional shifts within colonies due to unexpected changes in their attributes. In essence, alterations in a colony's characteristics (such as shifts in language or religion) can significantly affect its global position. This process in the ICA mirrors the genetic mutation observed in GA, facilitating broader exploration and mitigating premature convergence to suboptimal solutions.

Once adaptation and revolution are completed, colonies within an empire are exchanged for imperialists if their costs are lower. This process of imperialist competition is pivotal in augmenting the empire's overall strength. To assess the total power of each empire, we calculate its overall cost using the equation provided by Atashpaz-Gargari and Lucas (2007)

$$TotCost_k = cost_i + \xi \times Average[cost(colonies\ of\ empire\ k)] \tag{8}$$

where ξ is a positive integer in the range of 0 to 1, but not quite as low as zero.

Finally, the overall cost of an empire and its power are calculated using a normal total cost of an empire.

$$NormTotCost_k = \max[TotCost_k] - TotCost_k$$

$$EMP_k = \left| \frac{NormTotCost_k}{\sum_{i=1}^{N_{impre}} NormTotCost_i} \right| U_k = \left| \frac{C_k}{\sum_{i=1}^{N_{imper}} C_i} \right| \tag{9}$$

After calculating $[EMP_k - p_1, EMP_2 - p_2, \dots, EMP_{N_{imper}} - p_{N_{imper}}]$, each empire's weakest colony is given an index value between 0 and 1, representing a random integer drawn from a bag from the uniform distribution in $(0,1]$.

2.5 SFLA algorithm

SFLA, an emergent and highly regarded metaheuristic, draws inspiration from the evolutionary dynamics of a collective of frogs in pursuit of the richest sources of nourishment (Eusuff *et al*. 2006). Originating from the proposal in 2003 to amalgamate the benefits of the memetic algorithm (MA) with those of PSO, SFLA represents a novel fusion of these methodologies. In SFLA, the population of solutions, referred to as frogs, is partitioned into subsets known as memeplexes (MEXs). Each memeplex encompasses a distinct culture of memes (ranging from artistic expressions and ideologies to popular phrases and technological practices) each conducting localized searches within the solution space. During the memetic evolution process, frogs within a memeplex utilize specialized techniques to probe the local solution space, facilitating the dissemination of memes among individuals. A specified number of evolutionary iterations lead to the redistribution of information across memeplexes. This iterative shuffling and localized exploration continue in tandem until the system reaches the predefined convergence criteria.

The following outlines the SFLA algorithm: Consider an initial population of P frogs, each generated randomly. For a problem with D dimensions, each frog i is represented as $Y_i = (y_{i1}, y_{i2}, \dots, y_{iD})$. After sorting these frogs based on their fitness in descending order, they are organized into subsets. Each subset, known as a MEX, contains n frogs. Consequently, the total population is partitioned into m memeplexes, where $P = m \times n$. The first frog is assigned to the first MEX, the second frog to the second MEX, the third frog to the m^{th} MEX, and the fourth frog cycles back to the first MEX. This rotational assignment continues, ensuring that each frog is distributed across memeplexes in a systematic manner.

In each MEX, the most optimal and least optimal frogs are denoted as Y_b and Y_w, respectively. The frog with the highest fitness across the entire population is identified as Y_q. To enhance the fitness of Y_w, a technique analogous to PSO is employed within each subset. The adjustment procedure is delineated as follows (Eusuff *et al.* 2006):

$$FP = rand() \times (Y_b - Y_w) \quad (10)$$

$$Y_{new} = Y_w + FP, \quad -FP_{max} \leq FP \leq FP_{max} \quad (11)$$

Here, the random value produced by $rand()$ varies between 0 and 1. FP_{max} represents the maximum positional adjustment permissible for a frog. Should a superior solution, Y_{new}, be found, it replaces the least effective solution, Y_w; otherwise, the calculations specified in Eqs. (10) and (11) are reiterated, substituting Y_b for Y_g in Eq. (10). If no improvement is achieved, replacements are made randomly. Subsequently, all memeplexes are shuffled and reorganized during local search phases to facilitate information exchange and resource reallocation for subsequent search iterations.

2.6 GA algorithm

Selecting the optimal Genetic Algorithm (GA) for a particular application is a multifaceted challenge. There is no universal solution applicable to every problem. The current landscape frequently employs various representations, including integers and floating-point numbers (Azhdari Tehrani *et al.* 2022). As the complexity of the chromosome strings—such as those containing floating-point values—increases, so does the number of potential chromosomes. To ensure that the set of possible chromosomes adequately represents the spectrum of potential solutions, various contemporary (or non-classical) GAs utilize diverse encoding methods.

Upon selecting an encoding scheme, several critical decisions must be addressed. Factors such as the fitness function, population size, crossover and mutation operators, and the evolutionary strategy must all be carefully considered. The design of the GA typically involves a combination of experience, problem-specific modeling, and experimentation with different evolutionary schemes and parameters. Examples of non-classical GAs can be found in the literature, including practical applications discussed by Azhdari Tehrani *et al*. (2022).

For illustrative purposes, consider a GA architecture that employs only traditional genetic operators:

1. Initialization: Generate an initial population of P chromosomes randomly.
2. Fitness Evaluation: Assess the fitness, F(c), of each chromosome c in the source population.
3. Evolutionary Process: Iterate the following steps until the population of P chromosomes is exhausted:
   - Select two chromosomes, c1 and c2, from the source population using proportional fitness selection.
   - Produce a child chromosome (c) from c1 and c2 through one-point crossover with a crossover rate of PC.
   - Apply a mutation rate pm to chromosome c and perform uniform mutation.
   - Add the mutated chromosome c to the subsequent generation's population.
4. Population Update: Replace the original population with the newly generated one.
5. Termination Check: Repeat step 2 if the stopping criteria have not been met.

Initial GA implementations are typically straightforward, leveraging conventional design choices and the inherent adaptability and robustness of GAs. However, refining a GA to achieve superior performance in application-specific contexts can be more challenging. For problems such as scheduling, where numerous constraints make traditional gradient-based optimization methods ineffective, GAs have demonstrated significant efficacy.

3. Database preparation

3.1 The project

The Alborz Service Tunnel, a key component of the Tehran-Shomal Road project in Iran, is integral to this research database. This motorway connects Tehran, the Iranian capital, with Chalus, situated south of the Caspian Sea, spanning a distance of approximately 121 kilometers. Currently, travel time is between five to six hours along the winding minor roads through the Alborz Mountains. Upon completion of this project, travel time will be reduced to under two hours, and overall capacity will be significantly increased.

The planned highway route includes over 30 pairs of parallel, two-lane tunnels, with the Alborz Tunnel standing out as the longest in the world, extending 6,400 meters in length and rising 2,400 meters in elevation. Refer to Fig. 1 for the location of the Alborz Service Tunnel within this extensive tunnel network.

The service tunnel, positioned centrally between the twin Alborz tunnels, facilitates maintenance and inspection of the adjoining tubes (see Fig. 2). With its construction now complete, the tunnel's northern entrance connects to Shomal-S, while the southern exit leads to Tehran-T.

The tunnel traverses a complex geological setting characterized by a diverse lithology, including andesite, tuffs, anhydrite, sandstone, and limestone. These rock types exhibit varying compressive strengths, ranging from 20 to 120 MPa, indicative of their different mechanical behaviors under stress. The lithological diversity presents significant challenges in terms of stability and support design, as each rock type responds differently to excavation and loading conditions.

A particularly critical aspect of the tunnel's geological environment is the presence of the longest fault in the area, located between ST5339 and ST5361. This fault zone is of paramount concern due to its potential to influence tunnel stability. Fault zones are often zones of weakness, characterized by fractured and crushed rock that can significantly reduce the overall strength of the surrounding rock mass. The interaction between the fault and the tunnel could lead to increased deformation, necessitating careful consideration in both the design and construction phases.

Hydrologically, the region is marked by significant groundwater inflow into the tunnel, particularly in the vicinity of the fault. The high rate of water ingress not only complicates the construction process but also exacerbates the risk of rock mass weakening due to hydrostatic pressure and water-induced chemical weathering. The inflow of

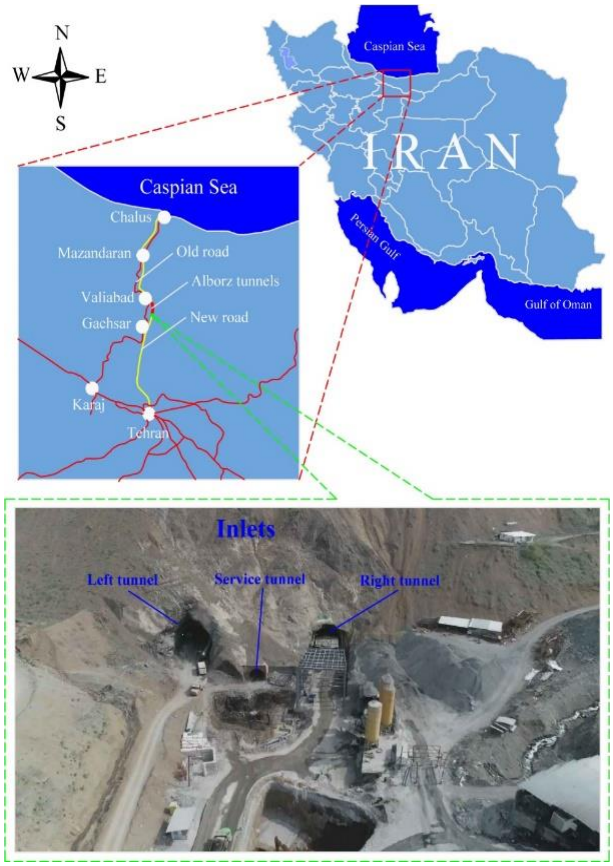

Fig. 1 The planned site of the Alborz Tunnel (Mahmoodzadeh *et al*. 2021)

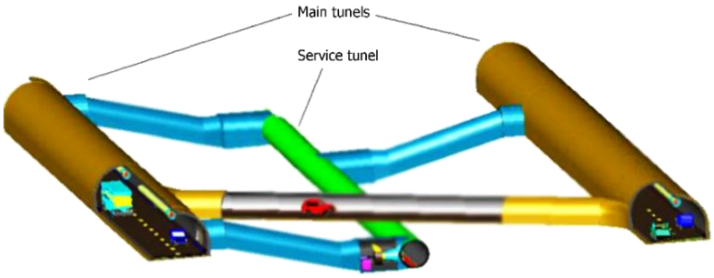

Fig. 2 Schematic of Alborz twin tunnel (Torabi *et al*. 2013)

water into the fault zone can lead to the saturation and softening of fault gouge and surrounding rock, further reducing their mechanical strength and increasing the likelihood of squeezing or collapse.

The Alborz Service Tunnel is being excavated at a consistent positive gradient of 1% using a 5.2-meter diameter open-gripper hard rock TBM manufactured by Wirth. The section with the greatest overburden extends over 850 meters. Excavation commenced on September 6, 2004, coinciding with the installation and commissioning of the TBM. Work officially began at TM 122 on February 6, 2005. After a rigorous 48-month excavation period, the TBM reached the S-portal heading on February 3, 2009. Out of a total of 1,459 days, 919 days (63%) were dedicated to excavation, achieving an average daily progress of 6.48 meters. The maximum daily advancement recorded was 30.47 meters, with a peak weekly progress of 110.96 meters and a maximum monthly progress of 389.43 meters.

3.2 Database

To achieve dependable forecasting, ensuring the reliability of the database is paramount and demands meticulous scrutiny. The database's integrity is crucial because even minor inaccuracies or gaps in the data can significantly skew forecasting results, thereby compromising the validity of the entire predictive model. Ensuring that the database is not only accurate but also comprehensive requires rigorous validation processes and extensive data collection to cover all relevant scenarios and parameter variations. A trustworthy database must encompass not only accurate data but also a comprehensive array of data points that span the full spectrum of parameter values influencing the problem under investigation. This means that the data should be representative of all possible conditions that could affect the outcome, ensuring that the model built upon this data is robust and generalizable across

	UCS	BTS	PSI	DPW	α	RFC	PR
UCS	1.00	0.11	0.41	-0.21	-0.25	0.03	0.32
BTS	0.11	1.00	0.40	-0.05	-0.11	-0.02	0.30
PSI	0.41	0.40	1.00	-0.23	-0.34	0.02	0.43
DPW	-0.21	-0.05	-0.23	1.00	0.23	-0.28	-0.16
α	-0.25	-0.11	-0.34	0.23	1.00	-0.01	-0.21
RFC	0.03	-0.02	0.02	-0.28	-0.01	1.00	0.02
PR	0.32	0.30	0.43	-0.16	-0.21	0.02	1.00

Fig. 3 Correlation matrix, or matrix of associations between variables

different scenarios.

Reducing data dimensionality by selecting a subset of measurable characteristics, or predictor variables, is essential in model development. By narrowing down the data to the most relevant features, the model's complexity is reduced, which can lead to more efficient computation and enhanced interpretability. This process generally involves minimizing a specified prediction error metric to choose an optimal model for a given subset. The goal is to identify the subset of variables that most accurately predicts the outcome while avoiding overfitting. The search for predictor subsets is conducted using algorithms that are constrained by limitations such as necessary or excluded features and subset size. These algorithms are designed to navigate through the vast space of possible feature combinations, considering constraints like feature inclusion requirements and the overall size of the subset, which impact the feasibility and effectiveness of the selection process.

When the original units and significance of features are crucial, and the goal is to identify an influential subset, feature selection is preferred over feature transformation. Feature selection preserves the original meaning and context of the data, which is critical when the interpretability of the results is important. Feature selection is particularly advantageous when dealing with categorical features. In cases where features represent distinct categories or classes, selecting relevant features directly helps maintain the integrity of the categorical information. Stepwise regression, a sequential feature selection strategy designed specifically for least squares fitting, is utilized to perform this optimization. This method incrementally adds or removes features based on their contribution to minimizing the least squares criterion, thus optimizing the model's predictive performance. The terms "stepwise" and "stepwisefit" are based on least-squares criteria. This approach systematically evaluates the impact of each feature, ensuring that the final model includes only those variables that significantly enhance the model's accuracy.

While generalized sequential feature selection eliminates features as they are introduced, stepwise regression can also reintegrate features previously removed. This iterative process allows for a more nuanced approach to feature selection, as it not only removes less relevant features but also considers reintroducing features if their addition improves the model's performance.

In this study, the stepwise method (SBS mode) is employed to reduce data dimensionality by concentrating on the most significant attributes. This method ensures that the final set of features is optimized for predictive accuracy while retaining interpretability. The results of this analysis identified six critical parameters (UCS, BTS, Punch Slope Index (PSI), Distance Between Planes of Weakness (DPW), Orientation of Discontinuities (α angle), and Rock Fracture Class (RFC)) as the most influential factors affecting TBM-PR. Each of these parameters was found to significantly impact the performance prediction of TBM, thereby guiding the selection of features for the predictive model. Consequently, for the database used in this research, each input training vector is of dimension 1 × 6, incorporating UCS, BTS, PSI, DPW, α, and RFC, while the output training vector is of dimension 1 × 1, comprising the TBM-PR. This configuration ensures that the model is trained on the most relevant features while predicting the target variable effectively.

Upon completion of the tunnel, researchers gained access to a comprehensive database containing 1,125 datasets. This dataset includes six input parameters and a singular output parameter, TBM-PR. Each data sample corresponds to a panel spaced approximately 6 meters apart along the tunnel's 6.4-kilometer length, yielding a total of 1,125 samples. To elucidate the relationships within the database, Fig. 3 presents the correlation matrix between input and output parameters. Pearson's correlation coefficient (R) is utilized to evaluate these associations. R, which ranges from -1 to 1, quantifies the degree of linear correlation between two variables, with -1 indicating a perfect negative correlation and 1 representing a perfect positive correlation. As illustrated in Fig. 3, the coefficients for PSI, UCS, and BTS are all below 0.43, indicating moderate correlations with TBM-PR. This suggests that while there is some correlation between these inputs and the TBM-PR outcome, it remains relatively modest. Other metrics exhibit similarly weak associations with TBM-PR.

Violin charts, which effectively illustrate data distribution across various categories and levels, are employed to present quantitative information. Fig. 4 displays the violin plots for the examined input and output parameters.

3.3 Data normalization

Data normalization is a fundamental operation in

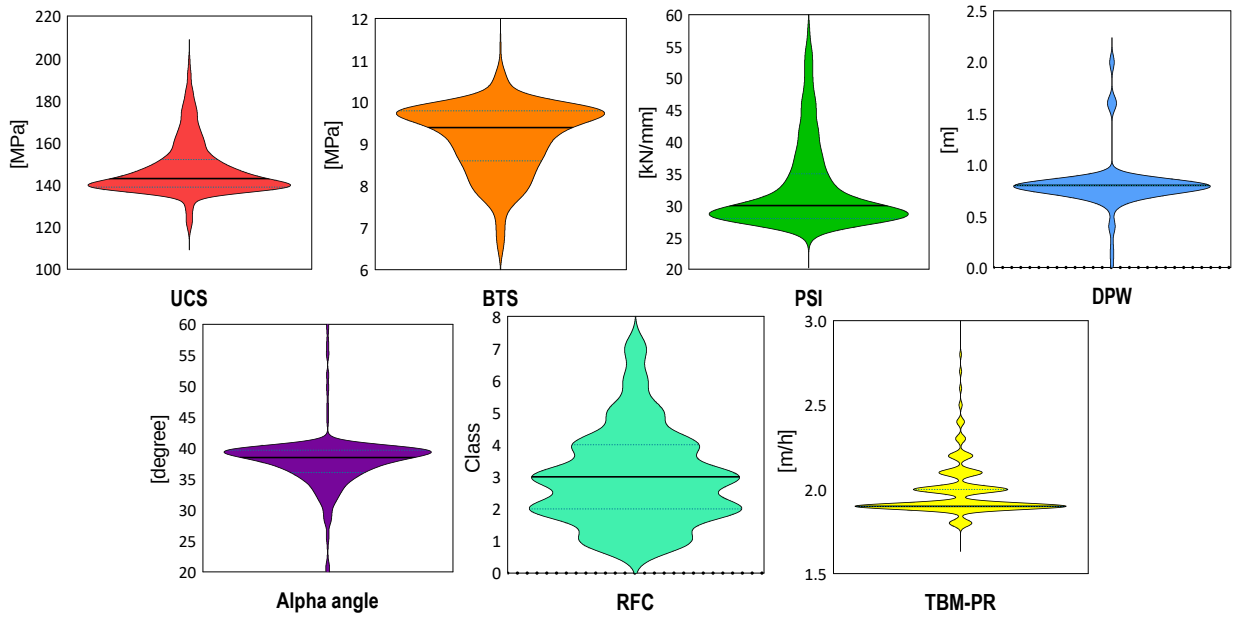

Fig. 4 Flowcharts depicting each input and output in violin form

machine learning. Given the diversity of data sources, variations in dimensions and units are often encountered. Normalizing training data accelerates model convergence by mitigating the impact of outliers and minimizing errors. The process of "data normalization" entails scaling the input features of a neural network so that they exhibit comparable means and standard deviations. This ensures that all features are uniformly scaled, facilitating more efficient and effective model training. Although data normalization does not directly prevent overfitting, normalizing our data makes the training problem easier. All data in this study are standardized in the range of zero to one using the Min-Max normalization method (see Eq. (12)).

$$x_{(0,1)} = \frac{x - Min}{Max - Min} \tag{12}$$

In this context, x represents the raw data while $x_{(0,1)}$ denotes the normalized data. Here, Min and Max refer to the minimum and maximum values within the entire dataset, encompassing both training and testing samples. Min-Max normalization has been demonstrated to be a robust technique in scientific research. However, a notable limitation is that this method may require re-normalization when new data is added, particularly if it extends beyond the original range of values. To address this issue, it is advisable to expand the Max-Min range beyond the original values. This precaution ensures that any additional data will consistently fall within the predefined Max-Min range.

4. Evaluation metrics

Statistical assessment criteria were used to assess the model's performance. These criteria are the coefficient of determination (R^2), the mean squared error (MSE), the variance account for (VAF), and the a^{20}_index.

$$R^2 = 1 - \frac{sum\ squared\ regression\ (SSR)}{sum\ of\ squares\ total\ (SST)} \tag{13}$$

$$\text{MSE} = \left(\frac{1}{n}\right)\sum_{i=1}^{n}(y_i - y_i')^2 \tag{14}$$

$$\text{VAF} = 1 - \left[\frac{var(y_i - y_i')}{var(y_i)}\right] \times 100\% \tag{15}$$

$$a^{20}\_index = \frac{m^{20}}{M} \tag{16}$$

where, y_i is the actual value, y_i' is the predicted value, $\bar{y}_i$ and $\bar{y}_i'$ are the means of actual and predicted values, n is the number of samples, M is the number of dataset samples and m^{20} is the number of samples with a value of rate experimental value/predicted value between 0.80 and 1.20.

5. Modeling procedure

In the context of this research, the process of training intelligent algorithms for predicting TBM-PR involves a series of meticulously planned stages, ensuring the development of a robust and reliable predictive model. Fig. 5 is a summary of the whole analytical and modeling process used in this study. The four essential phases of this study's methodology are laid out in this figure. It consists of four stages:

- Data Collection and Cleaning:

The first stage of the training process involves the meticulous gathering of data from the tunnel site. Given that the accuracy of the predictive model hinges on the quality of the input data, this phase focuses on obtaining comprehensive datasets that accurately represent the various factors influencing TBM-PR. Once the data is collected, it

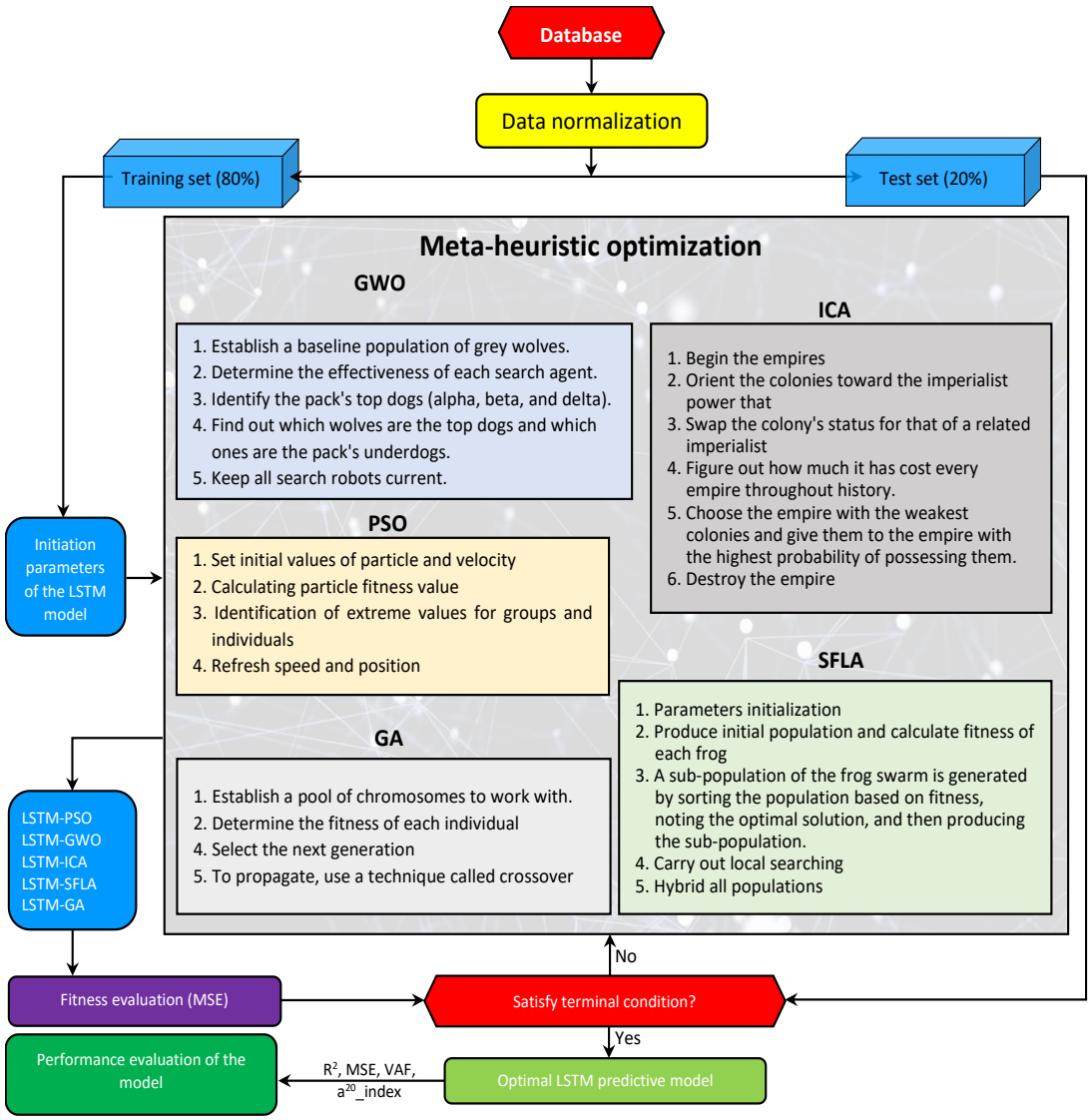

Fig. 5 Comprehensive analytical and modeling approach used in this paper

undergoes a rigorous cleaning process to ensure its integrity and suitability for model training. This step involves identifying and handling missing values, outliers, and any inconsistencies that could skew the results. By refining the data, the researchers ensure that the input to the model is as accurate and representative as possible, which is crucial for the subsequent stages of model development.

– Model Construction and Validation:

With a clean and reliable dataset in hand, the next stage involves constructing the predictive model. The research proposes the use of a LSTM network, a type of RNN particularly well-suited for handling sequential data and capturing temporal dependencies. The LSTM model is trained on the prepared datasets, allowing it to learn from the patterns and relationships inherent in the data.

To enhance the performance of the LSTM model, the study integrates meta-heuristic algorithms such as PSO, GWO, ICA, SFLA, and GA. These algorithms are employed to optimize the model hyper-parameters, improving the model's ability to generalize and make accurate predictions. During this phase, the model undergoes iterative training and validation, where it is continuously refined and tested against a validation set to ensure its robustness and accuracy.

– Model Assessment:

Once the model is trained and validated, it moves into the assessment phase. Here, the model's performance is rigorously evaluated using various metrics to determine its predictive accuracy and generalization capability. This step is crucial in ensuring that the model is not only accurate but also reliable in predicting TBM-PR under different conditions. The assessment helps identify any potential weaknesses in the model, allowing for further refinements if necessary.

– Data Analysis:

The final phase involves a detailed analysis of the model's predictions in relation to the actual outcomes. This analysis provides insights into the model's effectiveness and highlights the key factors influencing TBM-PR. By understanding the model's strengths and limitations,

Table 1 The value/type for each of the LSTM hyper-parameters optimized using different optimization algorithms.

Optimization algorithm	PSO	GWO	ICA	SFLA	GA
Number of hidden layers	3	3	2	4	3
Hidden neurons for each hidden layer	1st layer: 64 2nd layer: 32 3rd layer: 16	1st layer: 64 2nd layer: 32 3rd layer: 16	1st layer: 16 2nd layer: 8	1st layer: 64 2nd layer: 32 3rd layer: 16 4th layer: 8	1st layer: 32 2nd layer: 16 3rd layer: 8
Batch size	32	16	16	8	16
Dropout rate	0.3	0.4	0.4	0.3	0.4
Activation function	ReLu	ReLu	Exponential	ReLu	tanh
Learning rate	0.08	0.07	0.06	0.07	0.06
Momentum	0.75	0.72	0.78	0.74	0.81
Input time series length	12	10	10	8	12

researchers can make informed decisions about its practical application in real-world scenarios. Additionally, this phase may reveal areas where the model could be improved or extended, guiding future research efforts.

6. Results and discussion

Model hyper-parameters are pivotal in artificial intelligence and deep learning due to their profound impact on model performance. Recent studies indicate that with meticulous tuning and refinement, the performance of LSTM models can rival that of more complex LSTM variants (Aufa *et al.* 2020). The primary advantage of LSTM-based meta-heuristic optimization models is their algorithmic framework, which utilizes an exploration-exploitation strategy to converge on the optimal global solution. In this research, we enhance the LSTM model's hyper-parameters for predicting TBM-PR through meta-heuristic algorithms including PSO, GWO, ICA, SFLA, and GA.

It is crucial to acknowledge that deep learning approaches may suffer from overfitting, particularly when working with limited datasets. The 900 datasets used in this study for training are significantly fewer compared to those typically employed in language and image recognition tasks. The risk of overfitting in ANNs is exacerbated by the vast number of parameters relative to the limited training data. To mitigate overfitting and ensure accurate predictions, regularization techniques are essential. Dropout is a highly effective regularization method designed to counteract overfitting. It functions by diminishing the co-adaptability of neurons and preventing excessive reliance on specific neurons. During each training iteration, neurons are randomly deactivated with a probability p according to a Bernoulli distribution, resulting in an attrition rate of 1- p. The relevant formulae for implementing dropout are detailed below.

Without dropout

$$\tilde{p}_t = w_\circ h_t + b_\circ \tag{17}$$

With dropout

$$\tilde{p}_t = w_\circ \tilde{h}_t + b_\circ = w_\circ (r_t \odot h_t) + b_\circ \tag{18}$$

where $\tilde{p}_t$ represents the output of the model before it is passed through the time-t active function, and h_t represents the output vector from the hidden layer. Weights $w_\circ$ and bias $b_\circ$ reflect the connections between the hidden and output layers. Following the completion of dropout, the output vector of the hidden layer is denoted by the letter $\tilde{h}_t$. r_t is the Bernoulli-distributed random vector that represents the hidden layer's output after a dropout. The result of the model is

$$p_t = f(\tilde{p}_t) \tag{19}$$

where p_t is the predicted value and $f()$ is the activation function for the layer at the output of the model.

6.1 Models' performance evaluation

To evaluate the efficacy of the LSTM-based models comprehensively, the dataset was meticulously segmented into three distinct subsets: 70% was designated for training purposes, 20% was reserved for testing, and the remaining 10% was allocated for validation. This stratified approach ensures that the model is trained on a substantial portion of the data, while a separate, unbiased dataset is used for testing and validating the model's predictive accuracy.

During the training phase, a rigorous process was undertaken to optimize the LSTM model's hyper-parameters, utilizing advanced meta-heuristic algorithms. These algorithms were employed to refine several critical hyper-parameters, including the number of hidden layers, the number of neurons within each hidden layer, the batch size, the dropout rate, the choice of activation function, and the length of the input time series. The optimization process aimed to enhance the model's performance by systematically adjusting these hyper-parameters to achieve the most effective configuration.

The resultant optimized values for these hyper-parameters, which were derived from the application of various meta-heuristic optimization techniques, are meticulously documented in Table 1. This table provides a

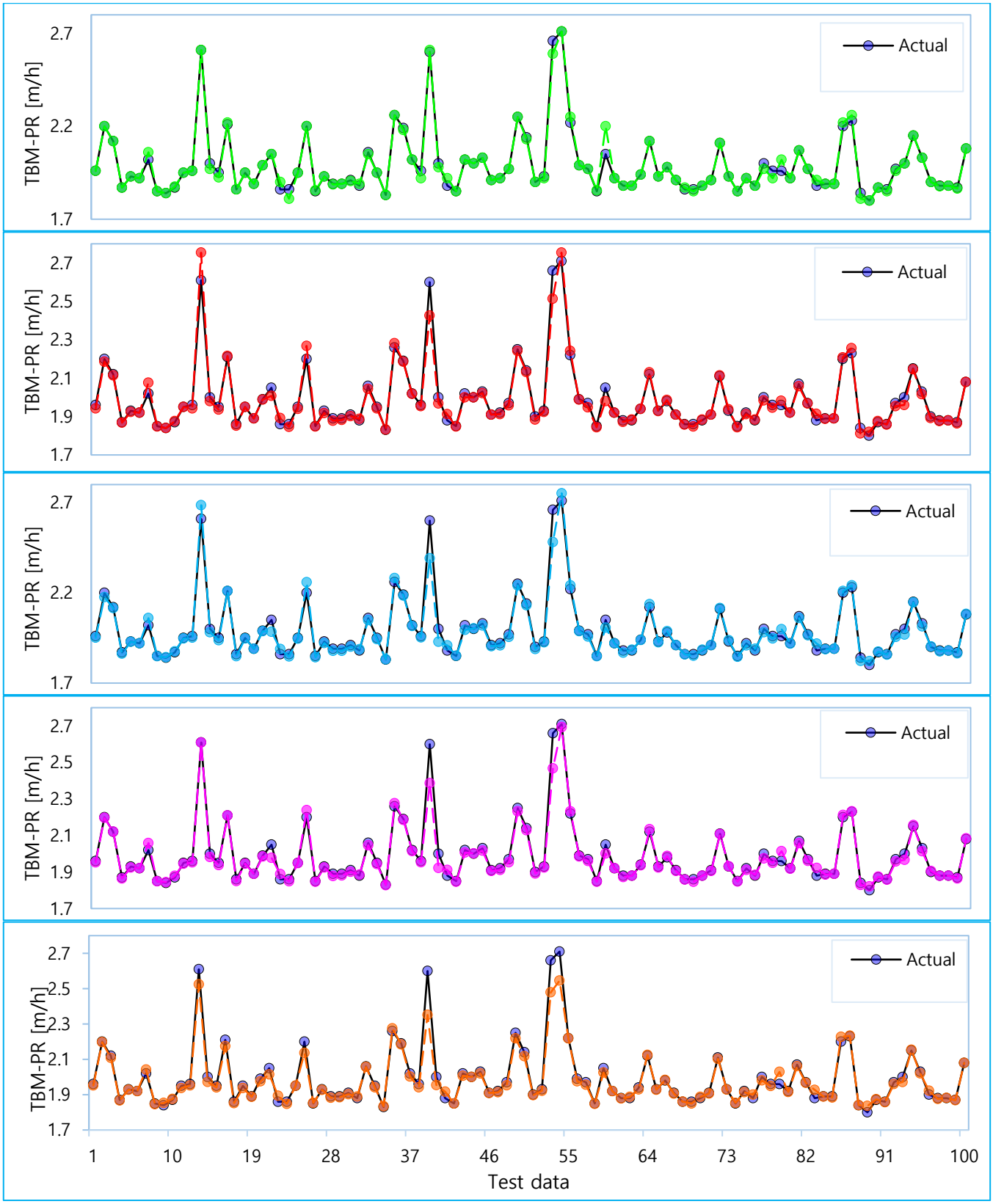

Fig. 6 Prediction results made by the optimized models

comprehensive overview of the specific settings determined for each hyper-parameter, highlighting the impact of different optimization algorithms on the model's configuration.

Upon completion of the training phase, the performance of the trained models was assessed using the reserved test data. Fig. 6 provides a visual representation comparing the TBM-PR values predicted by each model against the actual

Table 2 Analysis of the hybrid models' results in the two test and validation phases

	Method	R^2	Score	MSE	Score	VAF [%]	Score	a^{20}_index	Score	Rank
Test data	LSTM-PSO	0.9833	5	0.0015	5	94.75	5	1	1	16
	LSTM-GWO	0.9654	4	0.0023	4	91.64	4	1	1	13
	LSTM-ICA	0.9630	3	0.0027	3	89.87	3	1	1	10
	LSTM-SFLA	0.9624	2	0.0030	2	87.88	2	1	1	7
	LSTM-GA	0.9581	1	0.0031	1	85.46	1	1	1	4
Validation data	LSTM-PSO	0.9851	5	0.0013	5	95.12	5	1	1	16
	LSTM-GWO	0.9670	4	0.0020	4	93.21	4	1	1	13
	LSTM-ICA	0.9659	3	0.0025	3	91.10	3	1	1	10
	LSTM-SFLA	0.9650	2	0.0027	2	88.42	2	1	1	7
	LSTM-GA	0.9592	1	0.0030	1	86.38	1	1	1	4

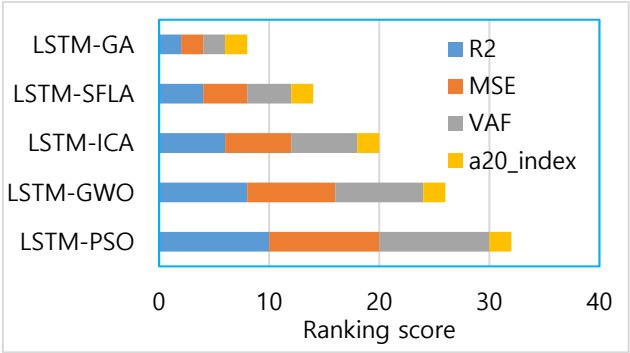

Fig. 7 Sum of test and validation ranking results

observed values. The analysis depicted in the figure demonstrates that all models produced results that were significantly satisfactory, indicating that the models performed well in terms of accuracy and reliability.

Table 2 presents the statistical indices employed to evaluate the performance of the models on both the test and validation datasets. This table provides a detailed comparison of the hybrid models, ranked according to their effectiveness in predicting TBM-PR, from highest to lowest. The results underscore that all five hybrid models exhibit commendable predictive performance. Nevertheless, the LSTM-PSO model stands out as the most proficient in generating accurate forecasts.

In addition, Table 2 incorporates a ranking system for each model, which facilitates further examination and comparison of their predictive capabilities. The aggregate ranking results, which combine both test and validation performances, are visually represented in a stacked graph format in Fig. 7. This visualization highlights that the LSTM-PSO model emerges as the most reliable and accurate across both training and validation phases. Consequently, the hybrid models are ranked in the following order of prediction accuracy: LSTM-PSO, LSTM-GWO, LSTM-ICA, LSTM-SFLA, and LSTM-GA.

Fig. 8 employs a Taylor diagram to illustrate the accuracy of each model in predicting TBM-PR during both validation and testing phases, using metrics such as standard deviation and correlation coefficient. This schematic representation demonstrates that while all models exhibit close performance levels and high forecasting accuracy, the LSTM-PSO model is distinguished by its superior reliability and efficiency in estimating TBM-PR

6.2 Sensitivity analysis

Encountering a new dataset for the first time can be an overwhelming experience, particularly when many features are presented without accompanying descriptions. Determining where to commence analysis can be challenging. A prudent initial step is to employ a feature utility metric, which quantifies the degree of association between a feature and a specific target variable. This approach allows us to narrow our focus to the most influential features, thereby optimizing our analytical efforts and ensuring that our time is effectively utilized.

In this context, mutual information (MI) serves as an invaluable statistic. While both correlation and MI assess relationships between variables, correlation is limited to identifying linear associations, whereas MI can uncover any form of dependency. MI is especially useful during the early stages of feature selection when the choice of the model remains undetermined. It is adept at detecting various types of relationships and is known for its ease of use, computational efficiency, theoretical robustness, and rapid performance.

MI is grounded in the concept of uncertainty reduction.

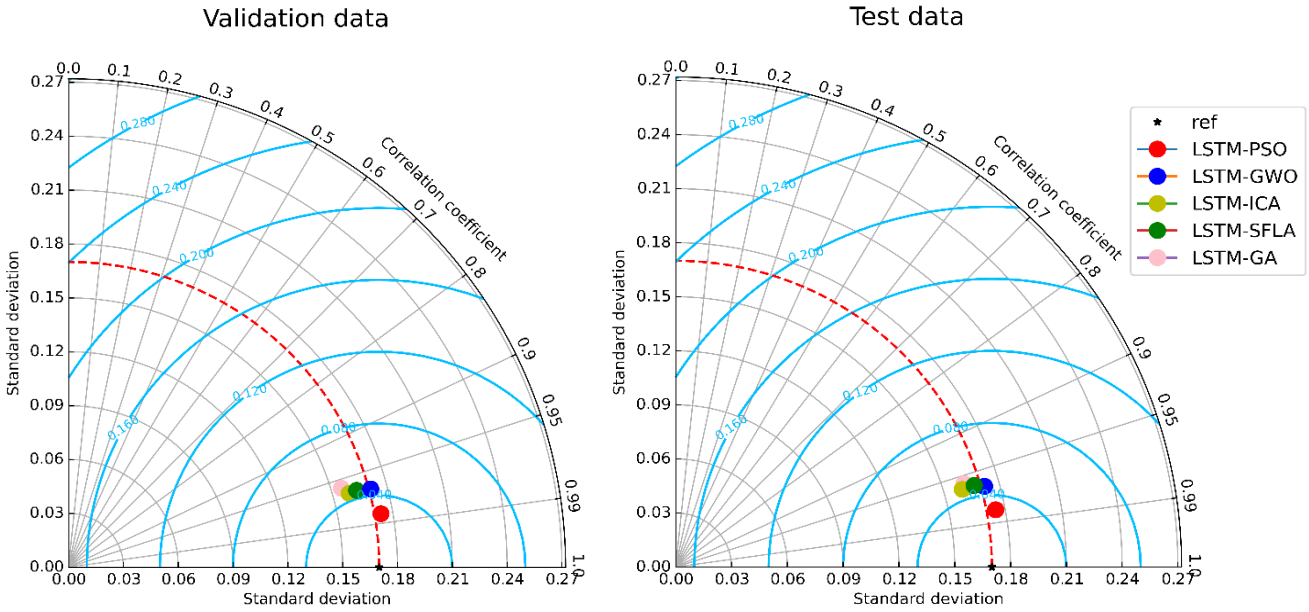

Fig. 8 Prediction performance of the models using the Taylor diagram in the testing and validation phases

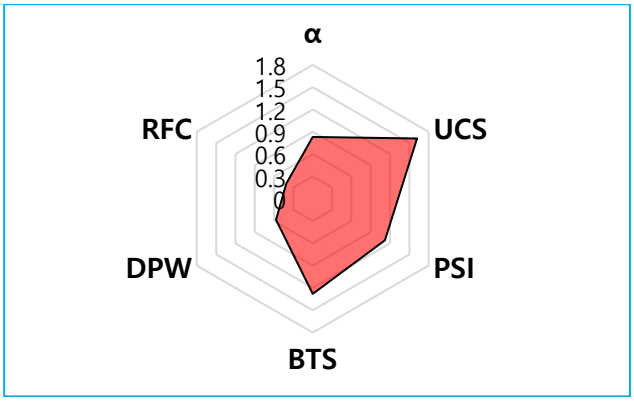

Fig. 9 The sensitivity score of each of the input parameters on the models' output calculated by the MI method

Specifically, it measures the extent to which knowledge of one variable reduces the uncertainty about another. Essentially, MI evaluates how much more confident we would be in achieving our objective if the value of a particular feature were known.

Fig. 9 illustrates the sensitivity scores of each input parameter as computed by the MI method, presented through a radar diagram. It is evident that the parameter UCS receives the highest score, indicating its significant influence. As depicted in Fig. 9, the parameters UCS, BTS, PSI, DPW, and RFC have notable effects on TBM-PR, with respective scores of 1.62, 1.28, 1.12, 0.83, 0.57, and 0.41. Consequently, these parameters are crucial in the analysis of TBM-PR. Among the six input parameters, UCS exerts the greatest impact on TBM-PR, whereas RFC has the least influence.

6.3 Limitations and future studies

Predicting TBM-PR presents a multifaceted challenge fraught with uncertainties, influenced by a myriad of factors such as geological conditions, excavation methodologies, rock strength, in-situ stress parameters, excavation dimensions, and machine specifications. In the current study, however, the focus is limited to six geological variables for constructing the machine learning models used for TBM-PR forecasting. While it is well-established that TBM-specific parameters (such as cutter arrangement, cutter type, and machine specifications) exert a considerable impact on TBM-PR, these factors were excluded from this study. This omission is due to the fact that the data utilized pertain exclusively to a tunnel excavated using a specific TBM, rendering the machine-related parameters effectively constant. Consequently, the model developed herein is not applicable for predicting the PR of varying TBM types. To develop a robust model capable of accommodating diverse geological conditions and various machine types, it is imperative to integrate extensive data from multiple tunnels, excavated under different geological settings and employing different TBMs.

7. Conclusions

The findings of this study yield the following critical conclusions:

– Meta-heuristic algorithms have demonstrated exceptional efficacy in the optimization of hyper-parameters within LSTM models. These algorithms, through their advanced search mechanisms and heuristic-driven strategies, have been proven to enhance the model's predictive accuracy by fine-tuning hyper-parameters such as the number of hidden layers, batch size, and dropout rates, thus leading to more robust and reliable model performance in complex prediction tasks.

- The comparative efficacy of machine learning models in forecasting TBM-PR, as assessed through rigorous testing and validation, ranks as follows: LSTM-PSO, LSTM-GWO, LSTM-ICA, LSTM-SFLA, and LSTM-GA. This hierarchical ranking reflects the varying levels of accuracy and robustness provided by each model, with LSTM-PSO emerging as the most proficient in delivering high-precision predictions, whereas LSTM-GA, although effective, ranks lower in performance metrics.
- Sensitivity analysis, conducted using the MI test on the dataset utilized in this research, revealed that parameters UCS, BTS, PSI, DPW, and RFC have a pronounced impact on TBM-PR prediction accuracy, with respective sensitivity scores of 1.62, 1.28, 1.12, 0.83, 0.57, and 0.41. This analysis underscored that UCS has the most significant influence on TBM-PR outcomes, whereas RFC showed the least influence, indicating differential weights in parameter significance which can guide focused data collection and feature selection for future model improvements.
- The study advocated for the deployment of the LSTM-PSO hybrid model as the most accurate and reliable approach for TBM-PR prediction. This recommendation is based on empirical evidence demonstrating the superior performance of the LSTM-PSO model in predicting TBM-PR compared to other models, highlighting its robustness in handling complex predictive tasks and its potential for broad applicability in similar predictive scenarios.
- This research is of substantial significance as it addresses a myriad of concerns pertinent to machine design and operational effectiveness across diverse geological conditions. By providing actionable insights and refined methodologies for TBM-PR prediction, the study offers invaluable guidance for manufacturers and operators, facilitating more informed decision-making and optimized performance in real-world tunneling operations.

Acknowledgments

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP.2/163/45. The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number "NBU-FFR-2024-2507-03".

References

Aufa, B.Z., Suyanto, S. and Arifianto, A. (2020), "Hyperparameter Setting of LSTM-based Language Model using Grey Wolf Optimizer", *Proceedings of the 2020 International Conference on Data Science and Its Applications (ICoDSA)*, 1-5. https://doi.org/10.1109/ICoDSA50139.2020.9213031.

Atashpaz-Gargari, E. and Lucas, C. (2007), "Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition", *Proceedings of the 2007 IEEE Congress on Evolutionary Computation*, 2007. https://doi.org/10.1109/CEC.2007.4425083.

Azhdari Tehrani, H., Ramezaninejad, S., Mardani, M., Shokouhi, S., Darnahal, M. and Hakamifard, A. (2022), "Hematologic malignancies and COVID-19 infection: a monocenter retrospective study", *Health Sci. Rep.*, **5**, e638. https://doi.org/10.1002/hsr2.638

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

Eusuff, M., Lansey, K. and Pasha, F. (2006), "Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization", *Eng. Optim.*, **38**, 129-154. https://doi.org/10.1080/03052150500384759.

Frough, O., Torabi, S.R. and Yagiz, S. (2015), "Application of RMR for estimating rock-mass–related TBM utilization and performance parameters: A case study", *Rock Mech. Rock Eng.*, **48**, 1305-1312. https://doi.org/10.1007/s00603-014-0619-4.

Fattahi, H. (2016), "Adaptive neuro fuzzy inference system based on fuzzy c–means clustering algorithm, a technique for estimation of TBM penetration rate", *Int. J. Optim. Civil Eng.*, **6**, 159-171. http://ijoce.iust.ac.ir/article-1-243-en.html.

Flor, A., Sassi, F., La Morgia, M., Cernera, F., Amadini, F., Mei, A. and Danzi, A. (2023), "Artificial intelligence for tunnel boring machine penetration rate prediction", *Tunn. Undergr. Sp. Tech.*, **140**, 105249. https://doi.org/10.1016/j.tust.2023.105249.

Ghorbani, E. and Yagiz, S. (2024), "Estimating the penetration rate of tunnel boring machines via gradient boosting algorithms", *Eng. Appl. Artif. Intell.*, **136**, 108985. https://doi.org/10.1016/j.engappai.2024.108985.

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.

Gong, Q.M. and Zhao, J. (2009), "Development of a rock mass characteristics model for TBM penetration rate prediction", *Int. J. Rock Mech. Min. Sci.*, **46**, 8-18. https://doi.org/10.1016/j.ijrmms.2008.03.003.

Gers, F.A., Schmidhuber, J. and Cummins, F. (2000), "Learning to Forget: Continual Prediction with LSTM", *Neural Comput.*, **12**, 2451-2471. https://doi.org/10.1162/089976600300015015.

Hochreiter, S. and Schmidhuber, J. (1997), "Long short-term memory", *Neural Comput.*, **9**(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Han, W., Jiang, Y., Zhang, X., Koga, D. and Gao, Y. (2021), "Quantitative assessment on the reinforcing behavior of the CFRP-PCM method on tunnel linings", *Geomech. Eng.*, **25**(2), 123-134. https://doi.org/10.12989/gae.2021.25.2.123.

Hamidi, K.J., Shahriar, K., Rezai, B. and Bejari, H. (2010), "Application of fuzzy set theory to rock engineering classification systems: An illustration of the rock mass excavability index", *Rock Mech. Rock Eng.*, **43**, 335-350. https://doi.org/10.1007/s00603-009-0029-1.

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

Jing, L., Li, J., Zhang, N., Chen, S., Yang, C. and Cao, H. (2021), "A TBM advance rate prediction method considering the effects of operating factors", *Tunn. Undergr. Sp. Tech.*, **107**, 103620. https://doi.org/10.1016/j.tust.2020.103620.

Kennedy, J. and Eberhart, R. (1995), "Particle swarm optimization," Proceedings of ICNN'95", *Proceedings of the International Conference on Neural Networks*, **4**, 1942-1948. https://doi.org/10.1109/ICNN.1995.488968.

Koopialipoor, M., Fahimifar, A., Ghaleini, E.N., Momenzadeh, M. and Jahed Armaghani, D. (2020), "Development of a new hybrid ANN for solving a geotechnical problem related to tunnel boring machine performance", *Eng. with Comput*., **36**, 345-357. https://doi.org/10.1007/s00366-019-00701-8.

Li, L., Liu, Z., Fang, X. and Qi, W. (2024), "Multi-step real-time prediction of hard-rock TBM penetration rate combining temporal convolutional network and squeeze-and-excitation block", *Scientific Reports*, **14**(1), 14326. https://doi.org/10.1038/s41598-024-65351-3.

Liu, Z., Wang, Y., Li, L., Fang, X. and Wang, J. (2022), "Realtime prediction of hard rock TBM advance rate using temporal convolutional network (TCN) with tunnel construction big data", *Front. Struct. Civil Eng.*, **16**(4), 401-413. https://doi.org/10.1007/s11709-022-0823-3.

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.

Li, J., Li, P., Guo, D., Li, X. and Chen, Z. (2021), "Advanced prediction of tunnel boring machine performance based on big data", *Geosci. Front*., **12**, 331-338. https://doi.org/10.1016/j.gsf.2020.02.011.

Lislerud, A. (1988), "Hard rock tunnel boring: Prognosis and costs", *Tunn. Undergr. Sp. Tech*., **3**, 9-17. https://doi.org/10.1016/0886-7798(88)90029-6.

Mahmoodzadeh, A., Taghizadeh, M., Mohammad, A.H., Ibrahim, H.H., Samadi, H., Mohammadi, M. and Rashidi, S. (2022a), "Tunnel wall convergence prediction using optimized LSTM deep neural network", *Geomech. Eng*., **31**(6), 545-556. https://doi.org/10.12989/gae.2022.31.6.545.

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

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

Mirjalili, S., Mirjalili, S.M. and Lewis, A. (2014), "Grey wolf optimizer", *Adv. Eng. Software*, **69**, 46-61. https://doi.org/10.1016/j.advengsoft.2013.12.007.

Mahmoodzadeh, A., Mohammadi, M., Hashim Ibrahim, H., Nariman Abdulhamid, S., Farid Hama Ali, H., Mohammed Hasan, A., Khishe, M. and Mahmud, H. (2021), "Machine learning forecasting models of disc cutters life of tunnel boring machine", *Automat. Constr.*, **128**, 103779. https://doi.org/10.1016/j.autcon.2021.103779.

Ozdemir, L. (1977), "Development of Theoretical Equations for Predicting Tunnel Borability", Ph.D. Thesis. T-1969, Colorado School of Mines, Golden, CO, USA. http://hdl.handle.net/11124/78500.

Pan, Q., Chen, Z., Wu, Y., Dias, D. and Oreste, P. (2021), "Probabilistic tunnel face stability analysis: A comparison between LEM and LAM", *Geomech. Eng.*, **24**(4), 399-406. https://doi.org/10.12989/gae.2021.24.4.399.

Rayatdust, H., Shahriar, K., Ahangari, K. and Kamali-Bandpey, H. (2012), "A statistical model for prediction TBM performance using rock mass characteristics in the TBM driven Alborz tunnel project", *Am. J. Appl. Sci*., **4**, 5048-5054. https://www.researchgate.net/publication/236011428.

Shan, F., He, X., Jahed Armaghani, D., Zhang, P. and Sheng, D. (2022), "Success and challenges in predicting TBM penetration rate using recurrent neural networks", *Tunn. Undergr. Sp. Tech.*, **130**, 104728. https://doi.org/10.1016/j.tust.2022.104728.

Salimi, A., Rostami, J., Moormann, C. and Delisio, A. (2016), "Application of non-linear regression analysis and artificial intelligence algorithms for performance prediction of hard rock TBMs", *Tunn. Undergr. Sp. Tech.*, **58**, 236-246. https://doi.org/10.1016/j.tust.2016.05.009.

Torabi, S.R., Shirazi, H., Hajali, H. and Monjezi, M. (2013), "Study of the influence of geotechnical parameters on the TBM performance in Tehran–Shomal highway project using ANN and SPSS", *Arabian J. Geosci.*, **6**(4), 1215-1227. https://doi.org/10.1007/s12517-011-0415-3.

Yalei, Y., Lijie, D., Rong, T., Fei, W. and Huilan, Z. (2024), "Prediction of TBM penetration rate for different surrounding rocks and cutter head diameters", *Heliyon*, **10**(12), e33174. https://doi.org/10.1016/j.heliyon.2024.e33174.

Yagiz, S. and Karahan, H. (2015), "Application of various optimization techniques and comparison of their performances for predicting TBM penetration rate in rock mass", *Int. J. Rock Mech. Min. Sci*., **80**, 308-315. https://doi.org/10.1016/j.ijrmms.2015.09.019.

Yagiz, S. (2008), "Utilizing rock mass properties for predicting TBM performance in hard rock conditions", *Tunn. Undergr. Sp. Tech.*, **23**, 326-339. https://doi.org/10.1016/j.tust.2007.04.011.

Yagiz, S., Gokceoglu, C., Sezer, E. and Iplikci, S. (2009), "Application of two non-linear prediction tools to the estimation of tunnel boring machine performance", *Eng. Appl. Artif. Intell*., **22**, 808-814. https://doi.org/10.1016/j.engappai.2009.03.007.

Yagiz, S. and Karahan, H. (2011), "Prediction of hard rock TBM penetration rate using particle swarm optimization", *Int. J. Rock Mech. Min. Sci*., **48**, 427-433. https://doi.org/10.1016/j.ijrmms.2011.02.013.

Zhou, J., Qiu, Y., Jahed Armaghani, D., 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. Front*., **12**, 101091. https://doi.org/10.1016/j.gsf.2020.09.020.

CC