IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 5 June 2024, accepted 21 June 2024, date of publication 24 June 2024, date of current version 2 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3418774

RESEARCH ARTICLE

Rapid Prediction of Cutterhead Torque in Hard-Rock Tunneling Using IEWOA-TSVD-ITELM

LONG LI[1] AND ZAOBAO LIU[2,3]

[1]School of Management Science and Engineering, Shandong Technology and Business University, Yantai, Shandong 264005, China
[2]Key Laboratory of Ministry of Education on Safe Mining of Deep Metal Mines, College of Resources and Civil Engineering, Northeastern University, Shenyang 110819, China
[3]Key Laboratory of Liaoning Province on Deep Engineering and Intelligent Technology, Northeastern University, Shenyang 110819, China

Corresponding authors: Long Li (long__li@hotmail.com) and Zaobao Liu (liuzaobao@mail.neu.edu.cn)

This work was supported in part by the Doctoral Initiation Fund of Shandong Technology and Business University under Grant BS202335 and in part by the Intelligent Control and Support Software to Safely and Efficiently Operate the tunnel boring machine (TBM) Tunnels from the China Railway Engineering Equipment Group Company Ltd., Project Team for the National Basic Research Program (973 Program).

ABSTRACT Predicting cutterhead torque is essential for optimizing TBM construction strategies and minimizing jamming risks. This study presents a novel hybrid model (IEWOA-TSVD-ITELM), developed using data from 7635 tunneling cycles from the Yinsong Project, to enhance cutterhead torque prediction accuracy. The EWOA enhances its global search capability by introducing new position updating and adaptive adjustment strategies (IEWOA). In addition, by leveraging the Softsign function for the nonlinear transformation of the expected output matrix of the TELM, a third hidden layer is added to enhance the feature extraction capabilities (ITELM), whereas truncated singular value decomposition (TSVD) is employed to reduce the noise in the output matrix of the third hidden layer of the ITELM (TSVD-ITELM). Furthermore, the IEWOA optimized the number of neurons and randomly generated weights and biases in the TSVD-ITELM. This study comprehensively evaluates and compares six optimization algorithms using 25 standard test functions. Additionally, the IEWOA-TSVD-ITELM is compared with eight classical machine learning models. This study examines the impact of different timing lengths of the rising phase and rock mass grades on model performance. The results demonstrate the outstanding performance of the IEWOA as an optimization algorithm. The IEWOA-TSVD-ITELM model achieves an R^2 value of 0.644 on the test set, with an MAE of 326.623 and an RMSE of 435.821, outperforming the other algorithms. Increasing the timing length from 30 to 60 s reduces the MAE and RMSE by 11.82% and 9.56%, respectively, but the gains diminish when the timing length increases from 60 to 90 s.

INDEX TERMS TBM cutterhead torque prediction, whale optimization algorithm, truncated singular value decomposition, two-hidden-layer extreme learning machine extreme, hybrid model.

I. INTRODUCTION

Full-face tunnel boring machines (TBMs) are extensively employed in the construction of hard-rock tunnels, subways, hydropower projects, and highways because of their high excavation efficiency, notable economic benefits, and minimal surface disturbance [1], [2]. As one of the most advanced mechanical devices for long-distance tunnel construction, TBMs integrate mechanical design and manufacturing, hydraulics, electronics, sensing, information, and automation technologies [3], [4]. They can simultaneously perform multiple tasks in tunnel excavation, including rock fragmentation, muck transport, and segment installation, significantly reducing the surface environmental impact compared with traditional drilling and blasting methods. However, the rock fragmentation process in TBMs is often constrained by complex and challenging geological environments. Rock fragmentation is achieved through the synergistic rotation of the cutterhead and cutters, where the face pressure applied by the cutterhead plays a crucial role [5]. The torque of the cutterhead, which serves as the driving force for the rotation of the cutters, is a pivotal metric that

The associate editor coordinating the review of this manuscript and approving it for publication was Asadullah Shaikh.

reflects the interaction intensity between the TBM and the geological conditions [3]. Their dynamic variations and high-load characteristics provide direct insights into geological forecasting and cutter wear. In addition, the cutterhead torque is a critical reference for operators to adjust the penetration rate and cutterhead rotation speed. Therefore, predicting the cutterhead torque is essential not only for the adaptive adjustment of tunneling strategies but also for mitigating the risk of jamming and related issues.

To make the significance of this study more accessible, imagine driving a car through a tunnel where you constantly adjust the steering and speed based on the road conditions. In TBM operations, predicting the cutterhead torque is akin to forecasting these adjustments ahead of time, allowing for smoother and safer tunneling operations. This prediction helps engineers adapt their strategies dynamically, much like a driver would adjust driving based on upcoming road conditions, thereby optimizing efficiency and minimizing risks.

Models for predicting the cutterhead torque are broadly classified into theoretical models, empirical models, numerical simulations, and intelligent algorithms. Within the domain of theoretical models, linear cutting machines are predominantly used to examine the torque produced by cutter heads, particularly the torque generated by individual cutters [6]. Various factors have been established that exhibit a strong correlation with cutterhead torque, including cutterhead diameter, number of cutters, geological parameters, cutterhead opening ratio, cutter spacing, friction coefficient, and penetration [7], [8]. Notably, the geological conditions and control parameters (cutterhead rotation speed and penetration rate) serve as the principal determinants of cutterhead torque [9]. The Norwegian University of Science and Technology has developed a comprehensive hard rock TBM database to create an empirical prediction model that incorporates parameters such as the cutterhead torque [10]. Both the univariate and multivariate regression methods have been extensively applied to empirical models [11]. Given known variables, such as the cutterhead diameter, number of cutters, and geological parameters, theoretical and empirical models facilitate rapid preliminary estimates of the cutterhead torque before the initiation of TBM tunneling. However, owing to the challenges associated with the real-time acquisition of geological parameters, theoretical and empirical models often fail to meet practical construction requirements. Numerical simulations, including the discrete element method [12], [13] and the finite element method [14], have been employed to assess the cutterhead torque by simulating the intricate interaction mechanisms between the TBM cutterhead and rock, thereby playing a critical role in optimizing the TBM cutterhead design and mitigating jamming risks.

With the application of internet of things technology, TBM operational data are no longer limited to isolated datasets, and real-time data transmission and acquisition have become feasible. For example, in the Yin-Song project, the sensor parameters reached 199. These datasets cover complete tunneling cycles and downtime data under various geological conditions. Owing to the heterogeneity of geological conditions and groundwater distribution, significant fluctuations in the cutterhead torque have been observed across different tunneling cycles of TBMs [15]. These fluctuations reflect the high variance in the data, which exacerbates the difficulty in predicting the cutterhead torque. TBM tunneling cycles are typically segmented into rising and stable phases. During the stable phase, the tunneling parameters were dynamically adjusted based on the data accumulated during the rising phase [16]. Given that the rising phase generally extends beyond 4 min [17], it reduces the construction efficiency of the TBM. To address this challenge, there is an increasing focus on utilizing data from the initial 30 s of the rising phase combined with intelligent algorithms to swiftly predict the cutterhead torque of the stable phase [18]. The mechanical, electrical, and hydraulic characteristics of the TBM were incorporated as input features into these predictive models [19]. This approach aims to enhance predictive accuracy and operational efficiency by rapidly adapting to evolving tunneling conditions.

The application of intelligent algorithms for predicting cutterhead torque in TBMs is characterized by two predominant research orientations. Models such as convolutional neural networks [20], [21], long short-term memory (LSTM) networks [18], and bidirectional LSTM (BLSTM) [4] are primarily employed when continuous data from the rising phase are used as the input. These models were selected because of their ability to effectively capture complex temporal dependencies in TBM operational data. However, these methods have limitations. For example, convolutional neural networks are computationally intensive and require substantial processing power. LSTMs and BLSTMs can experience issues, such as vanishing gradients, and generally have slower training speeds. Moreover, all of these models require large amounts of labeled data for effective training.

Conversely, when mean values derived from ramp-up data are utilized as inputs, decision trees, support vector machines (SVMs), and Gaussian processes [22], along with gradient bosting decision trees [23], random forests [24], and deep neural networks [25], are the preferred models. Utilizing mean values as inputs not only simplifies the analysis process but also reduces the demands on computational resources and minimizes the impact of data noise. However, current models have certain limitations, such as the inadequate nonlinear fitting capabilities of decision trees and the enhanced accuracy of SVM in datasets characterized by small sample sizes and limited features. Studies have shown that two hidden layer extreme learning machines (TELMs) are suitable for predicting TBM tunneling parameters (such as thrust), but their performance is limited by the number of hidden layers and randomly generated weights and biases [17]. Effective hyperparameter optimization can

significantly enhance model performance by fine-tuning the parameters that govern learning algorithms; however, existing methods often fail to provide robust solutions. Therefore, there is a pressing need for more advanced and efficient optimization techniques to improve the predictive accuracy and reliability of these models.

Recently, optimization techniques such as particle swarm optimization (PSO) [26], grey wolf optimizer (GWO) [27], butterfly optimization algorithm (BOA) [28], and whale optimization algorithm (WOA) [29] have gained prominence in research. The WOA is extensively utilized in path optimization [30], engineering problem resolution [31], and fault diagnosis [32] because of its rapid convergence and strong global search capabilities. Certain scholars have further enhanced the performance of the algorithm by simplifying and refining the prey strategy, resulting in an enhanced whale optimization algorithm (EWOA) [33] that exhibits superior performance compared to other variants. However, limitations persist in the position-update strategies of the EWOA and adaptive searching, indicating substantial scope for performance enhancement.

To improve the prediction accuracy of the cutterhead torque, this study proposes a hybrid machine learning model based on the TELM framework. The feature extraction capabilities of the TELM were enhanced by initially employing the Softsign function to execute a nonlinear transformation of the outputs from the second hidden layer. This approach circumvents the establishment of rigid boundaries and mitigates loss of pertinent information. The model architecture was subsequently deepened by the addition of a third hidden layer, with a variable number of neurons employed to enhance the generalization capacity of the model. To combat the risk of overfitting, truncated singular value decomposition (TSVD) was implemented to retain the singular values of the weight matrix of the third hidden layer that contributed the most to the model, thereby reducing data noise. To address the issue of poor parameter adaptability owing to the intrinsic randomness in the generation of weights and biases in TELMs, the EWOA introduces new methods for calculating convergence factors and dynamic weight parameters, enhancing the adaptive search capabilities of the algorithm. By iteratively updating the randomly generated weights and biases in the ITELM, this approach reduces training errors.

The remainder of this paper is organized as follows: Section II introduces the extreme learning machine model, whale optimization algorithm, and evaluation metrics. Section III presents the proposed IEWOA-TSVD-ITELM. Section IV describes engineering case data and feature selection. Section V compares the performance of the optimization algorithms and the predictive results of nine machine learning models. Section VI discusses the impact of the rising timing length and rock mass grades on the model performance, as well as the data distribution before and after weight optimization. Finally, section VII concludes the study and presents the main findings.

II. METHODS

TBM construction data are continuously updated in real time, allowing the quick prediction of cutterhead torque under various geological conditions. Complex geological conditions complicate torque prediction, necessitating a robust and efficient modeling approach. Gaussian processes offer excellent uncertainty quantification but face high computational costs [34], limiting their real-time application in TBM operations. Physics-informed neural networks ensure that predictions adhere to physical laws but require significant computational resources [35], [36].

By contrast, the ELM, proposed by Huang et al., addresses the time consumption issue in iterative backpropagation training in neural networks [37], as shown in Figure 1. ELM employs a single hidden-layer feedforward neural network with randomly assigned fixed input weights and hidden-layer biases. The ELM quickly adapts to new data, meets the real-time demands of TBM operations, and provides a practical and efficient solution for predicting the cutterhead torque under complex geological conditions.

A. EXTREME LEARNING MACHINE

Given a dataset with N samples, there exists an input matrix $\boldsymbol{E} = [\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N]^{\mathrm{T}} \in \mathbf{R}^{N\times n}$ and an output matrix $\mathbf{O} = [\boldsymbol{O}_1, \boldsymbol{O}_2, \cdots, \boldsymbol{O}_N]^T \in \mathbf{R}^{N\times m}$. Between them, where $\boldsymbol{e}_j = \left[e_{j1}, e_{j2}, \ldots, e_{jn}\right]^{\mathrm{T}} \in \mathbf{R}^n$, $\mathbf{O}_j = \left[o_{j1}, o_{j2}, \ldots, o_{jm}\right]^{\mathrm{T}} \in \mathbf{R}^m$. The input feature dimension is n, and the output feature dimension is m. The ELM method initially randomizes a bias matrix $\mathbf{B} = \left[b_1, b_2, \ldots, b_q\right] \in \mathbf{R}^{N\times L}$ and a weight matrix $\mathbf{W} = \left[\boldsymbol{W}_1, \boldsymbol{W}_2, \ldots, \boldsymbol{W}_q\right] \in \mathbf{R}^{n\times L}$ to connect the input layer with the hidden layer, where L denotes the number of neurons in the hidden layer, and $\mathbf{W}_q = \left[W_{q1}, W_{q2}, \ldots, W_{qn}\right]^{\mathrm{T}} \in \mathbf{R}^n$ is the weight vector connecting n input neurons to the q hidden neurons.

The output matrix $\mathbf{H}$ can be calculated by Eq. (1).

$$\mathbf{H} = h\left(\mathbf{EW} + \mathbf{B}\right) \tag{1}$$

The output matrix $\boldsymbol{O}$ is calculated by Eq. (2).

$$\mathbf{HF} = \mathbf{O} \tag{2}$$

where $\mathbf{F} = [\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_L]^T \in \mathbf{R}^{L\times m}$ represents the output weight matrix connecting the hidden and output layers.

The output weight matrix $\mathbf{F}$ is calculated using the least-squares method, as shown in Eq. (3).

$$\mathbf{F} = \mathbf{H}^{+}\mathbf{O} \tag{3}$$

The Moore–Penrose (MP) inverse matrix $\mathbf{H}$ is represented as $\mathbf{H}^{+}$. When N is greater than L, if $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ is nonsingular, then $\mathbf{H}^{+} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}$, conversely, when N is less than L, $\mathbf{H}\mathbf{H}^{\mathrm{T}}$ is nonsingular, and then $\mathbf{H}^{+} = \mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{H}^{\mathrm{T}}\right)^{-1}$.

B. TWO-HIDDEN-LAYER EXTREME LEARNING MACHINE

To further improve the nonlinear fitting ability of the ELM, the TELM was proposed by Qu et al. in 2016 [38], as shown in Figure 2. Compared to the ELM, the TELM has two hidden

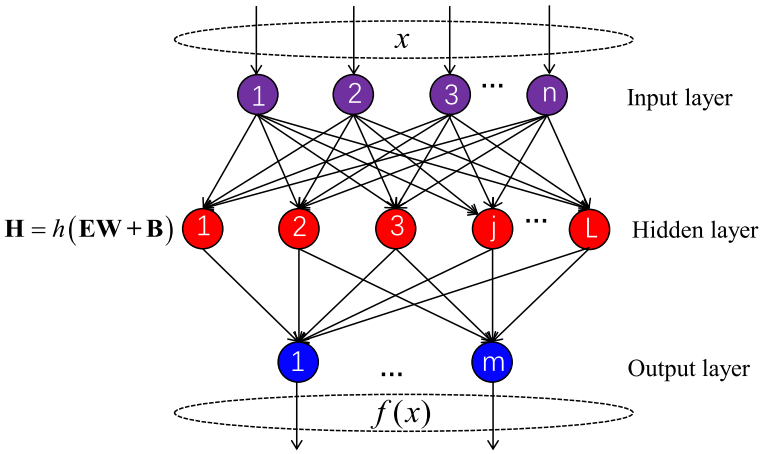


FIGURE 1. Structure of the ELM.

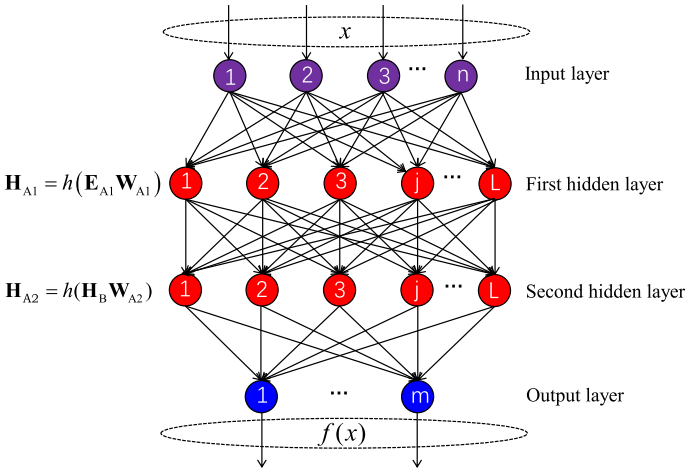


FIGURE 2. Structure of the TELM.

layers, that is, it has 2 L neurons. Given a data set containing $(\boldsymbol{e}_j, \boldsymbol{o}_j)(j = 1, 2, \ldots, N)$ with N samples, after initializing the weight matrix $\mathbf{W}_1$ and bias $\mathbf{B}_1$ of the first hidden layer, $\mathbf{W}_1$ and $\mathbf{B}_1$ are spliced to construct an augmented matrix $\mathbf{W}_{A1} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{W}_1 \end{bmatrix}$. To meet the computational requirements, the input matrix $\mathbf{E}$ and N column vectors, whose elements are all one, are concatenated, that is, $\mathbf{E}_{A1} = [\mathbf{1}\ \mathbf{E}]$. The calculation process for the output matrix $\mathbf{H}_{A1}$ of the first hidden layer is expressed in Eq. (4).

$$\mathbf{H}_{A1} = h\left(\mathbf{E}_{A1}\mathbf{W}_{A1}\right) \tag{4}$$

The calculation process of the weight matrix $\mathbf{F}_1$ is shown in Eq. (5), which is used to connect the second hidden layer and the output layer.

$$\mathbf{F}_1 = \mathbf{H}_{A1}^{+}\mathbf{O} \tag{5}$$

The expected output $\mathbf{H}_2$ of the second hidden layer can be calculated using Eq. (6).

$$\mathbf{H}_2 = \mathbf{O}\mathbf{F}_1^{+} \tag{6}$$

To implement the calculation process between the first and second hidden layers, the augmented matrices $\mathbf{W}_{A2}$ and $\mathbf{H}_B$ must be defined. $\mathbf{W}_{A2}$ was calculated from the weight matrix $\mathbf{W}_2$ of the first and second hidden layers and the bias matrix of the second hidden layer $\mathbf{B}_2$, that is $\mathbf{W}_{A2} = \begin{bmatrix} \mathbf{B}_2 \\ \mathbf{W}_2 \end{bmatrix}$. $\mathbf{H}_B$ is obtained by concatenating the output matrix of the first hidden layer and N column vectors, whose elements are all 1; that is, $\mathbf{H}_B = [\mathbf{1}\ \mathbf{H}_{A1}]$. Eq. (7) expresses the $\mathbf{H}_2$ calculation process for the expected output matrix.

$$\mathbf{H}_2 = h\left(\mathbf{H}_{A1}\mathbf{W}_2 + \mathbf{B}_2\right) \tag{7}$$

The augmented matrix $\mathbf{W}_{A2}$ can also be obtained by the inverse calculation of Eq. (8), thereby avoiding the need to randomly generate a new weight matrix $\mathbf{W}_2$ and bias matrix $\mathbf{B}_2$.

$$\mathbf{W}_{A2} = \mathbf{H}_B^{+}h^{-1}\left(\mathbf{H}_2\right) \tag{8}$$

The generalized inverse matrix of $\mathbf{H}_B$ is $\mathbf{H}_B^{+}$, $h^{-1}(\mathbf{x})$ which represents $h(\boldsymbol{x})$ the inverse activation function. Eq. (9) is the calculation process of the second hidden layer output matrix

$\mathbf{H}_{A2}$.

$$\mathbf{H}_{A2} = h(\mathbf{H}_B \mathbf{W}_{A2}) \tag{9}$$

Eq. (10) shows the calculation process for the weight matrix $\mathbf{F}_2$ that connects the second hidden layer to the output layer.

$$\mathbf{F}_2 = \mathbf{H}_{A2}^{+}\mathbf{O} \tag{10}$$

With $\mathbf{H}_{A2}^{+}$ representing the generalized inverse matrix of $\mathbf{H}_{A2}$, the final output of the TELM can be expressed as $f(\boldsymbol{x}) = \mathbf{H}_{A2}\mathbf{F}_2$.

C. IMPROVED TWO-HIDDEN-LAYER EXTREME LEARNING MACHINE

In the actual calculation process of the TELM, to avoid the null value in the result of $h^{-1}(\mathbf{H}_2)$ in Eq. (10), when the maximum value of $\mathbf{H}_2$ exceeds 1 or the minimum value is lower than -1, it needs to be normalized to the range of -0.9 to 0.9. $\mathbf{H}_{A2}$ performs the corresponding denormalization to ensure the consistency of the numerical values in the calculation process. Nonetheless, this normalization and denormalization processes may inadvertently modify the fundamental attributes of the data, thereby introducing errors and potentially compromising model performance.

To avoid these effects, the Softsign function is used to perform a nonlinear transformation on $\mathbf{H}_2$ such that the numerical interval of $\mathbf{H}_2$ is within $(-1, 1)$, as shown in Eq. (11). Through its continuous and asymptotic characteristics, the Softsign function can achieve adaptive compression of data and map values to the $(-1, 1)$ interval. This process not only avoids the setting of hard boundaries and reduces the risk of information loss but also preserves the relative size and structure of the data. Furthermore, it helps alleviate the sensitivity of the model to extreme values and improves its generalization ability and numerical stability.

$$\text{Softsign} = \frac{x}{1 + |x|} \tag{11}$$

The two hidden layers of the TELM have the same number of neurons, and this structure is limited in capturing deeper and more complex data features and relationships, particularly for highly nonlinear and multidimensional datasets. Therefore, this study added a third hidden layer based on the TELM to improve the ability of the model to extract features and define the number of transformable neurons in the third layer. After randomly initializing the weight matrix $\mathbf{W}_3$ and bias $\mathbf{B}_3$ of the third hidden layer, the calculation process for the third hidden layer output matrix $\mathbf{H}_{A3}$ is shown in Eq. (12).

$$\mathbf{H}_{A3} = h(\mathbf{H}_{A2}\mathbf{W}_3 + \mathbf{B}_3) \tag{12}$$

The weight matrix $\mathbf{F}_{A3}$, which connects the third hidden layer to the output layer, is calculated using Eq. (13).

$$\mathbf{F}_3 = \mathbf{H}_{A3}^{+}\mathbf{O} \tag{13}$$

Based on the above improvements to the TELM, we abbreviate it as ITELM, as shown in Figure 3. The increased

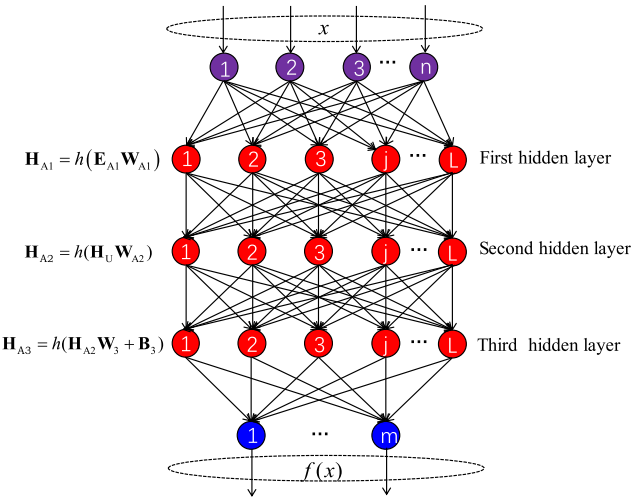


FIGURE 3. Structure of the ITELM.

number of neurons helps the TELM show better generalization capabilities. However, this design may cause the risk of overfitting and increase the instability of the model. To balance these factors, this study uses the truncated singular value decomposition (TSVD) method to process $\mathbf{H}_{A3}$, as shown in Eq. (14). This method optimizes the data representation by retaining only the top k singular values that contribute the most to the original matrix. This method maintains the dimensions of the original matrix and reduces redundancy and potential noise in the data by retaining only the singular values that contribute the most energy and zero out the remaining singular values. This process is described as follows:

$$\mathbf{H}_{A3} = \mathbf{U}\sum\mathbf{V}^T. \tag{14}$$

Given the matrix $\mathbf{H}_{A3}$, the three matrices, $\mathbf{U}$, $\sum$, and $\mathbf{V}^T$, are obtained through singular value decomposition, and $\sum$ is a diagonal matrix.

The process for calculating the total energy D_{total} is shown in Eq. (15), where σ_i is a singular value in $\sum$.

$$D_{\text{total}} = \sum\nolimits_i \sigma_i^2 \tag{15}$$

The cumulative energy proportion $D_{\text{cumulative}}(i)$ is shown in Eq. (16).

$$D_{\text{cumulative}}(i) = \frac{\sum_{j=1}^{i} \sigma_j^2}{D_{\text{total}}} \tag{16}$$

The steps to determine the number k of retained singular values based on the cumulative energy proportion and set energy threshold are shown in Eq. (17).

$$k = \min\{k | D_{\text{cumulative}}(k) \geq \text{threshold}\} \tag{17}$$

Finally, $\mathbf{H}_{A3}$ is reconstructed using the first k singular values, as expressed in Eq. (18).

$$\mathbf{H}_{A3} = \mathbf{U}\sum\nolimits_k \mathbf{V}^T \tag{18}$$

Here, $\sum_k$ is a modified diagonal matrix containing only the first k singular values; the other elements are 0 and the dimensions matching the original matrix, that is $\sum_k = \text{diag}(\sigma_1, \sigma_2, \ldots, 0, \ldots 0)$.

Based on TSVD's improvements to ITELM, we call it TSVD-ITELM. The weights and biases are randomly initialized in the third hidden layer, which may cause the TSVD-ITELM to ineffectively match the distribution of the current training data. Iteratively updating the weights and biases using optimization algorithms is an effective strategy to solve this problem. Considering the applicability of the whale optimization algorithm in many fields, such as fault diagnosis and path optimization, this study intends to use this method to solve this problem.

D. BASIC WHALE OPTIMIZATION ALGORITHM

The whale optimization algorithm (WOA) is inspired by the hunting behaviors of humpback whales. Mirjalili and Lewis used three processes to model it: (i) surround the prey, when $|G| < 1$ (exploitation stage); (ii) use a bubble net to attack the prey; and (iii) randomly search for prey, when $|G| > 1$ (exploration stage).

if $r_1 < 0.5$ and $|G| \leq 1$:

$$X_{t+1} = X^* - G \times P, \tag{19}$$

where $D = |C \times X^* - X_t|$, $G = 2 \times a \times r - 2$, $a = 2 - 2(t/t_{\max})$, and $C = 2 \times r$.

If $r_1 < 0.5$ and $|G| > 1$:

$$X_{t+1} = X_{rand} - G \times P_{rand}, \tag{20}$$

where $P_{rand} = |C \times X_{rand} - X_t|$.

If $r_1 \geq 0.5$:

$$X_{t+1} = X^* + \cos(2\pi l) \times P' \times e^{bl}, \tag{21}$$

where $P' = |X^* - X_t|$, $l = (a_2 - 1) \times rand + 1$, $a_2 = -1 - t/t_{\max}$

Factor G fluctuates randomly within the $[-2, 2]$ interval, while parameter l varies randomly between $[-1, 1]$ intervals. Variable t represents the current iteration, with $t_{\max}$ setting the maximum iteration limit. The constant b is recommended to be 1. In addition, r is a uniformly distributed random variable within the range [0, 1].

E. IMPROVED OPTIMIZATION ALGORITHM

To improve the simplicity and efficiency of the WOA, the Enhanced WOA (EWOA) was introduced [33]. This revised algorithm features two significant enhancements: firstly, it adopts a cosine function (where $1 > \cos(2\pi l) > -1$) to refine the searching and encircling phases, replacing the older criteria of $2 > A > -2$. Second, it advances the exploitation phase by assigning parameter b a random integer value within the range of 0–500, aiming for more dynamic optimization processes.

If $r_1 \geq 0.5$:

$$X_{t+1} = X^* + \cos(2\pi l) \times P' \times e^{bl}, \tag{22}$$

where $D' = |X^* - X_t|$

If $r_1 < 0.5$:

$$X_{t+1} = X^* - \cos(2\pi l) \times P_{new}, \tag{23}$$

where $P_{new} = (C \times X^* - X_t)$.

The EWOA outperforms the WOA in terms of convergence speed and optimization efficacy; however, its performance is constrained by the update dynamics of parameter l. This parameter is crucial during the exploration phase for exploring various search areas and dictates the search trajectory during the exploitation phase, which is heavily influenced by the convergence factor a_2. Research shows that shifting a_2 from a linear to a nonlinear pattern significantly enhances the operational efficiency [39], as detailed in Eq. (24). However, this alteration resulted in an initially slow convergence rate for a_2, which then accelerated in subsequent iterations. To address this issue, we introduce an optimized updating method for a_2, as presented in Eq. (25), enabling a faster initial convergence and consistent performance in the later phases, as shown in Figure 2.

$$a_2 = 2 \times \left(e^{-\left(1 + \frac{t}{t_{\max}}\right)} - 1 \right) \tag{24}$$

$$a_2' = 2 \times \left(e^{-\left(1 + 2 \times \left(\frac{t}{t_{\max}}\right)^{0.5}\right)} - 1 \right) \tag{25}$$

To adjust the search behavior adaptively during the iterations, we introduced a dynamic weight parameter w, as shown in Eq. (26). By combining cosine and sine functions, leveraging their periodicity and phase differences, the EWOA adjusts its search strategies dynamically. The cosine function, which gradually decreases, promotes extensive global exploration early on. By contrast, the sine function, which gradually increases, supports precise local searches later. The settings for $w_{\max}$ and $w_{\min}$ ensure w varies within a predefined range.

$$w = w_{\max} \times \cos\left(\frac{\pi \times t}{t_{\max}}\right) + w_{\min} \times \sin\left(\frac{\pi \times t}{t_{\max}}\right) \tag{26}$$

To explore the weight w adjustment mechanism in the WOA, a comparative analysis of different $w_{\max}$ and $w_{\min}$ combinations revealed their effects. Figure 4 shows that reducing $w_{\max}$ and $w_{\min}$ narrows the variation range of w. Higher $w_{\max}$ and $w_{\min}$ settings expand the search range and diversity of the algorithm, whereas lower settings focus the search on specific areas during later iterations, improving efficiency and precision.

By introducing the dynamic weight parameter w, the improved calculation method is shown in Eq. (27) and Eq. (28). The calculation flowchart of the WOA and proposed IEWOA is shown in Figure 5.

If $r_1 \geq 0.5$:

$$X_{t+1} = w \times X^* + \cos(2\pi l) \times P' \times e^{bl}. \tag{27}$$

If $r_1 < 0.5$:

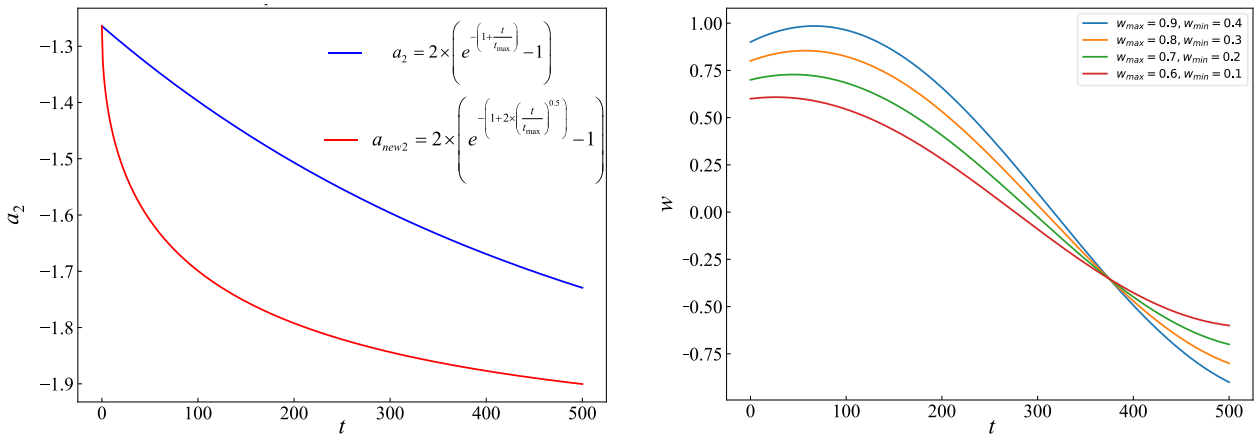
$$X_{t+1} = w \times X^* - \cos(2\pi l) \times P_{new}. \tag{28}$$

Algorithm 1 TSVD-ITELM

Input: $h(x)$: activation function, $\mathbf{E}$: input matrix with N samples, $\mathbf{O}$: output matrix with N samples

Output: $f(x) = h\{h\{[h(\mathbf{E}\mathbf{W}_1 + \mathbf{B}_1)\mathbf{W}_2 + \mathbf{B}_2]\}\mathbf{W}_3 + \mathbf{B}_3\}\mathbf{F}_{\mathbf{new}}$

1. Randomly initialize biases $\mathbf{B}_1$ and weights $\mathbf{W}_1$,define the augmented matrix $\mathbf{W}_{A1}$ and $\mathbf{E}_{A1}$
2. Calculate the matrix $\mathbf{H}_{A1} = h(\mathbf{E}_{A1}\mathbf{W}_{A1})$ and define the augmented matrix $\mathbf{H}_B$
3. Calculate matrix $\mathbf{F}_1 = \mathbf{H}_{A1}^{+}\mathbf{O}$
4. Calculate matrix $\mathbf{H}_2 = \mathbf{O}\mathbf{F}_1^{+}$
5. Calculate the augmented matrix$\mathbf{H}_2 = h(\mathbf{H}_{A1}\mathbf{W}_2 + \mathbf{B}_2)$
6. Perform nonlinear transformation on $\mathbf{H}_2 = \frac{\mathbf{H}_2}{1+|\mathbf{H}_2|}$
7. Calculate matrix $\mathbf{W}_{A2} = \mathbf{H}_B^{+}h^{-1}(\mathbf{H}_2)$
7. Calculate matrix $\mathbf{H}_{A2} = h(\mathbf{H}_B\mathbf{W}_{A2})$
8. Randomly initialize biases $\mathbf{B}_3$ and weights $\mathbf{W}_3$
9. Calculate matrix $\mathbf{H}_{A3} = h(\mathbf{H}_{A2}\mathbf{W}_3 + \mathbf{B}_3)$
10. Calculate the matrix $\mathbf{H}_{A3} = \mathbf{U}\sum_k\mathbf{V}^T$ using TSVD
11. Calculate matrix $\mathbf{F}_3 = \mathbf{H}_{A3}^{+}\mathbf{O}$



(a) Comparison of two calculation results of a_2

(b) Iterative results of different weight combinations

FIGURE 4. **Comparative results of different a_2 calculation methods and different weight combinations.**

F. EVALUATION METRICS

In this study, the prediction of cutterhead torque is formulated as a regression problem. To analyze the performance of the diverse models, the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) were selected as the primary metrics for evaluation, enabling a comprehensive assessment of model capabilities from multiple perspectives. The MAE is particularly emphasized as the main criterion during the process of optimization algorithm selection and hyperparameter tuning owing to its straightforward depiction of the average deviation between the predicted and observed values. The calculation formulas for these three evaluative metrics are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}, \tag{29}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(y_i - \hat{y}_i)|, \tag{30}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{31}$$

III. PROPOSED IEWOA-TSVD-ITELM

Figure 6 shows the predictive framework for the cutterhead torque, encompassing four modules: (1) the TBM construction data preprocessing module, (2) the proposed ITELM module, (3) the proposed IEWOA module, and (4) the prediction and evaluation of the IEWOA-TSVD-ITELM module. The specifics are as follows.

(1) TBM construction data preprocessing module: This module encompasses the segmentation of the TBM tunneling cycles, data preprocessing, and feature extraction processes.

(2) The proposed IEWOA module enhanced the optimization performance of the EWOA by introducing a new convergence factor a_2 and a methodology for calculating the dynamic weight parameters w.

(3) The proposed ITELM module: Initiates with a Softsign function for the nonlinear transformation of the expected output matrix $\mathbf{H}_2$, incorporates a third hidden layer to augment the nonlinear fitting capabilities of the TELM, and utilizes TSVD to mitigate errors caused by randomly generated weights and biases in the third hidden layer.

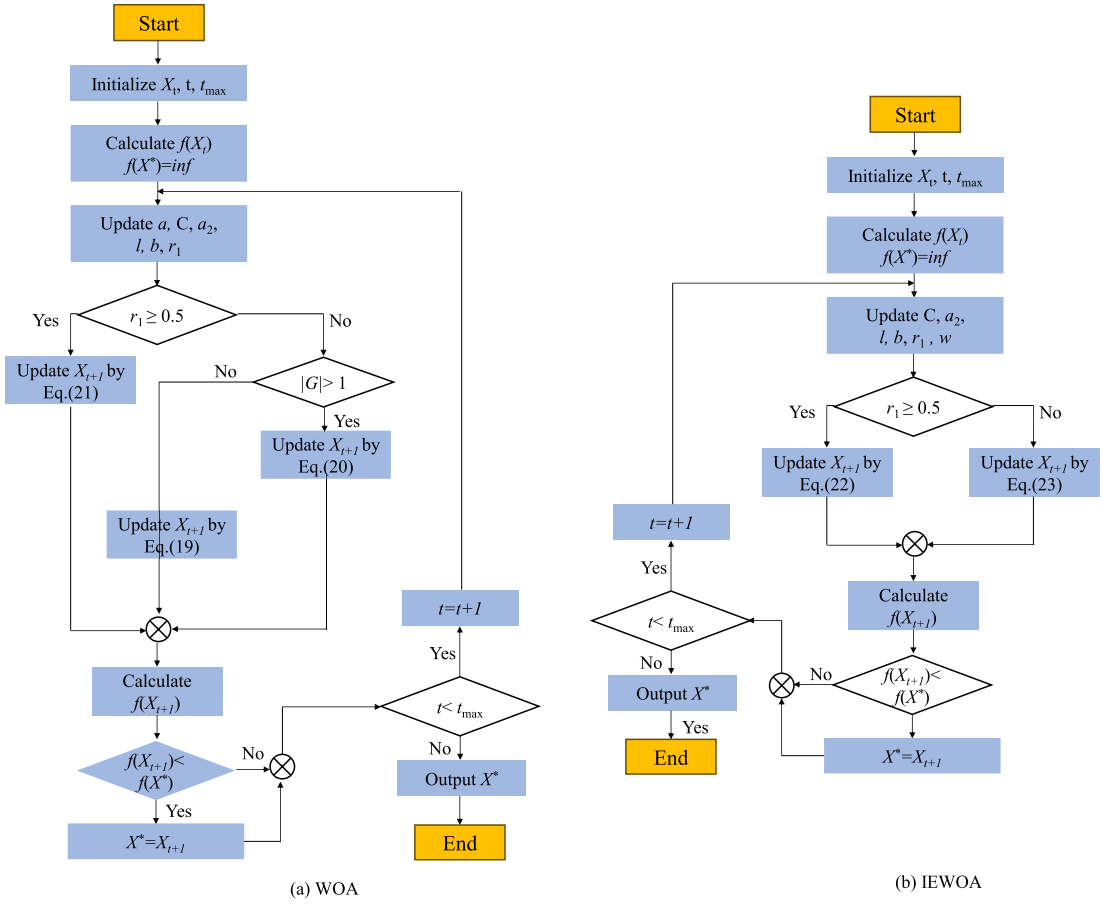


FIGURE 5. Flowchart of WOA and proposed IEWOA.

(4) Prediction and evaluation of the IEWOA-TSVD-ITELM module: Optimize the weights and biases of the ITELM using the IEWOA and evaluate the performance of various models in predicting the cutterhead torque using three metrics: the coefficient of determination (R^2), mean absolute error (MAE), and mean square error (MSE). Table 1 lists the necessary tools and techniques used in the simulation and evaluation of the proposed model.

IV. CASE DATA AND FEATURE SELECTION

The Yinsong water diversion project, a key component of China's 13th Five-Year Plan, is Jilin Province's most significant interregional water transfer project (Figure 7)s. This study focuses on the TBM-3 construction section, which begins at the Yinma River diversion and ends at the Chalu River, spanning from the 71+476 to the 51+705 milestone. Utilizing an open-type TBM for excavation, this segment achieved an effective tunneling distance of 17,488 m, with a tunnel diameter of 7.93 m. The TBM operations were extended over 728 days, with the tunnel achieving depths ranging from 85 to 260 m. According to the rock mass classification system of China Hydropower (HC) [40], the project primarily involved rock mass classes II, III, IV, and V, reflecting the diverse geological challenges of TBM construction.

The preprocessing of TBM construction data is detailed in [41]. In this project, approximately 200 parameters were collected from the TBM construction data. However, most of these parameters were not directly related to the cutterhead torque. To optimize the input features, Li et al. [18] employed a data-driven approach to identify key parameters impacting the cutterhead torque, including the elimination of constants, parameters with low variance, and those with high correlation. Consequently, 10 input features were determined for torque prediction. Li et al. [17] performed a parameter sensitivity analysis to demonstrate that the left and top shield pressures affect the TBM excavation parameters. Ultimately, 12 parameters were selected as input features, including right shield pressure (X1), top shield pressure (X2), machine conveyor motor current (X3), gripper pressure (X4), cutterhead speed setting value (X5), propel pump motor current (X6), cutterhead power (X7), cutterhead rotation speed (X8), left shield pressure (X9), propel speed potentiometer setting value (X10), gear seal pressure (X11), and propel pressure (X12). The correlation between the input features and cutterhead torque and the data distribution under different

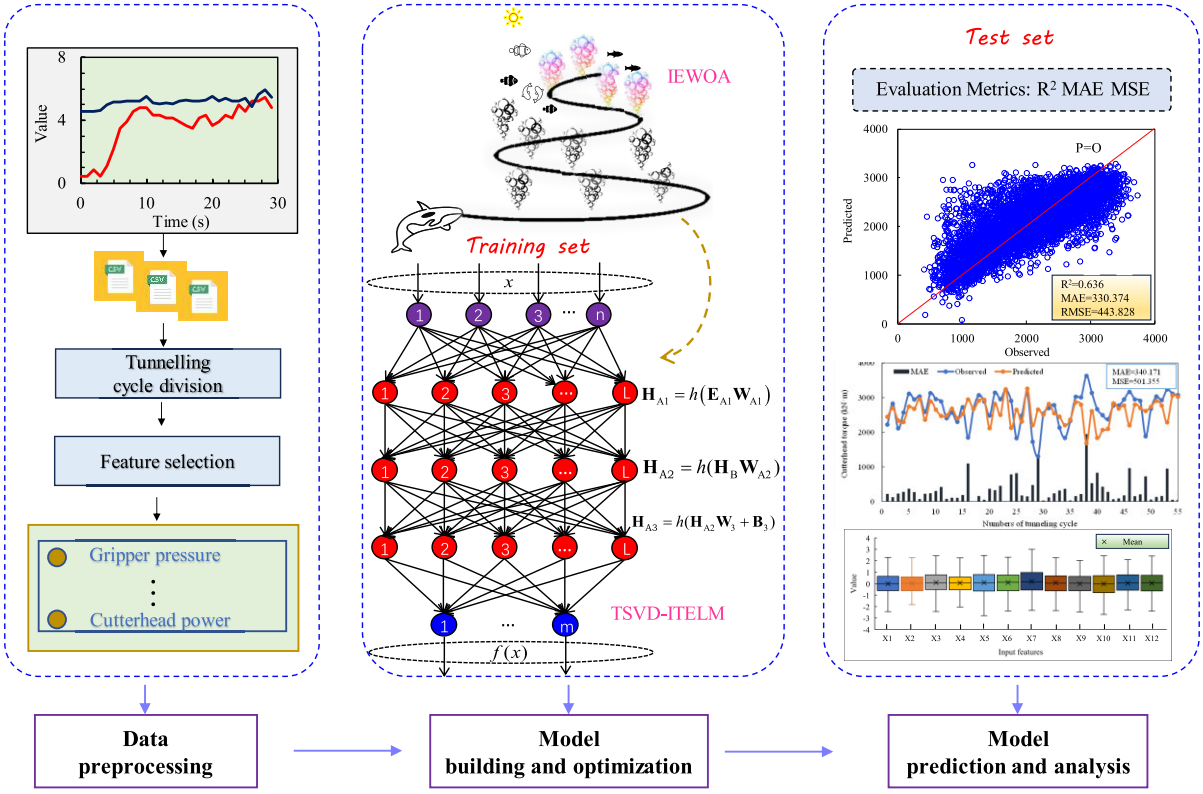


FIGURE 6. Flowchart of the developed approach for IEWOA-TSVD-ITELM prediction.

rock mass grades, taking the cutterhead rotation speed as an example, are shown in Figure 8. Although the correlation between some features is high, they have been shown to affect the prediction results of the cutterhead torque [18].

V. APPLICATION OF THE IEWOA ON BENCHMARK FUNCTIONS

A. TWENTY-FIVE BENCHMARK FUNCTIONS

To evaluate the performance of the proposed IEWOA, we tested it using the 25 well-known benchmark functions listed in Table 2. These functions comprehensively assess the ability of optimization algorithms to find the global optima under various dimensions and complexities. The performance of the IEWOA was compared with those of several standard algorithms, including PSO, GWO, WOA, LWOA, and EWOA. All the algorithms used a population size of 30 and a maximum of 500 iterations. Each benchmark function was independently run 30 times to ensure accuracy and reliability. The means and standard deviations of these runs were used for detailed comparison. This approach minimized the impact of randomness and provided stable and reliable evaluations. Experiments were conducted on a desktop computer with an Intel i5-12500 3.00 GHz processor and 16 GB of RAM, running Windows 10 (64-bit) with PyCharm as the development environment.

B. COMPARISON OF THE RESULTS OF SIX OPTIMIZATION ALGORITHMS

Table 3 presents the mean and standard deviation of the different algorithms for each test function, with the best results highlighted in bold. The results show that, except for F2, F9, F10, F13, and F25, the proposed IEWOA outperformed the other algorithms on all test functions. This demonstrates the superiority of the IEWOA. Notably, the IEWOA shows a significant performance improvement over the EWOA, highlighting the effectiveness of the innovative convergence factor a_2 and dynamic weight parameter w.

VI. RESULTS

A. PREDICTIVE OF IEWOA-TSVD-ITELM

To evaluate the performance of the IEWOA-TSVD-ITELM, the dataset was split into training and testing sets in an 80% to 20% ratio. The training set was used to obtain the optimal hyperparameter combination using optimization algorithms to achieve the best performance. The testing set was employed to assess the generalization ability of the model for unknown data. Given that the performance of the ELM is directly influenced by the number of neurons, it is necessary to optimize the number of neurons. Adjustments were made to the number of neurons for the ITELM and TSVD-ITELM using the IEWOA, with the target range set between 100 and 300, followed by the utilization of the IEWOA to optimize randomly generated weights and biases

TABLE 1. Tools and techniques used in the simulation and evaluation of the proposed model.

Technique/Tool	Description	Purpose
IEWOA (Improved EWOA)	Introduces new position updating and adaptive adjustment strategies to EWOA	Enhances global search capability and optimizes parameters such as neurons, weights, and biases in the model
TSVD (Truncated Singular Value Decomposition)	Applied to the output matrix of the third hidden layer of the ITELM	Reduces noise and improves data quality, ensuring better model stability and performance
ITELM (Improved TELM)	Uses the Softsign function for nonlinear transformation and adds a third hidden layer	Enhances feature extraction capabilities and adjusts to the nonlinearity in data
25 Standard Test Functions	Used to evaluate and compare the performance of six optimization algorithms	Assesses the robustness and effectiveness of the optimization algorithms used
Comparison with other machine learning models	Eight machine learning models are used as benchmarks	Provides a comparative analysis to evaluate the superiority of the proposed model
Analysis of timing lengths and rock mass grades	Examines different timing lengths of the rising phase (30, 60, and 90 s) and various rock mass grades (II, III, IV, V)	Investigates how varying operational and geological conditions affect model performance

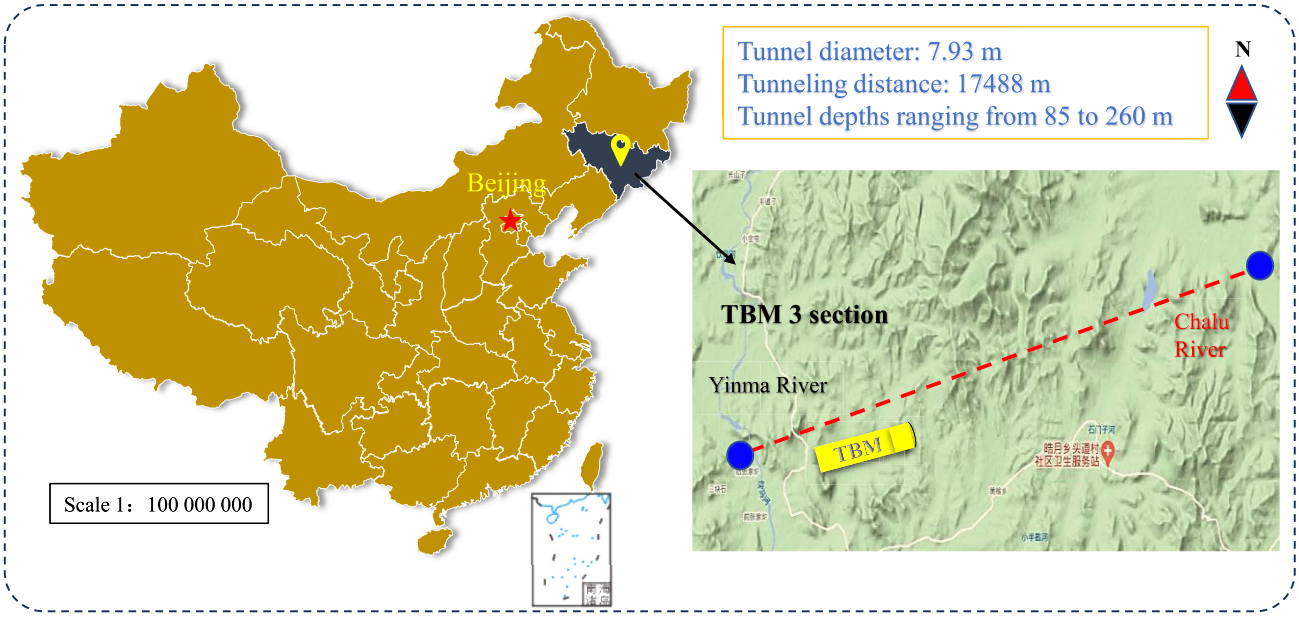
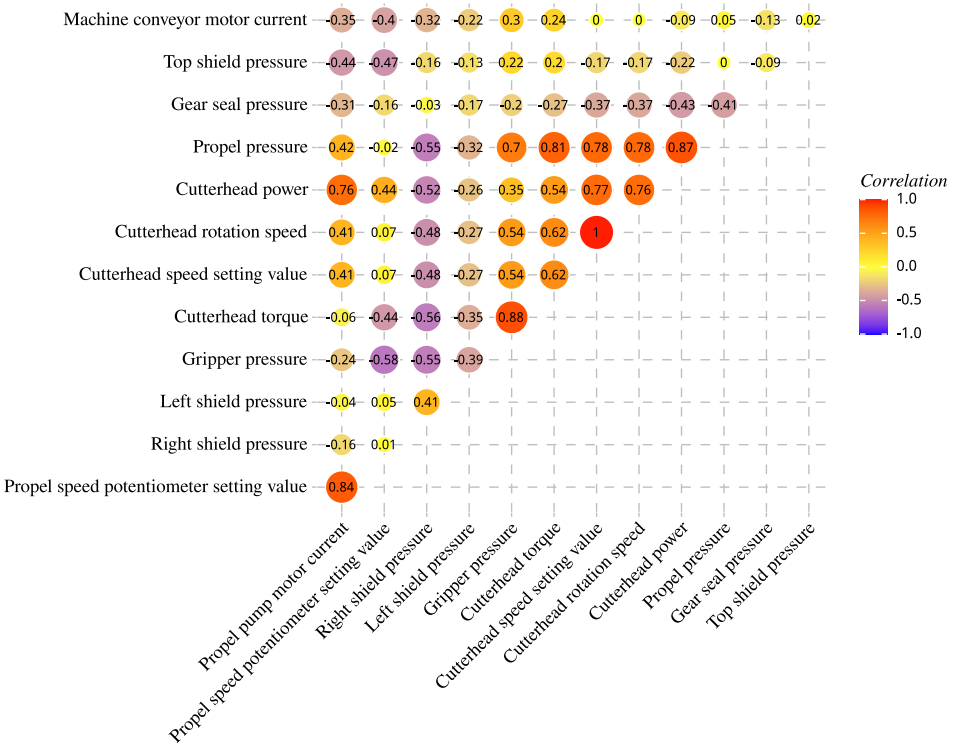


FIGURE 7. Location and planimetric map of the Yinsong diversion project.

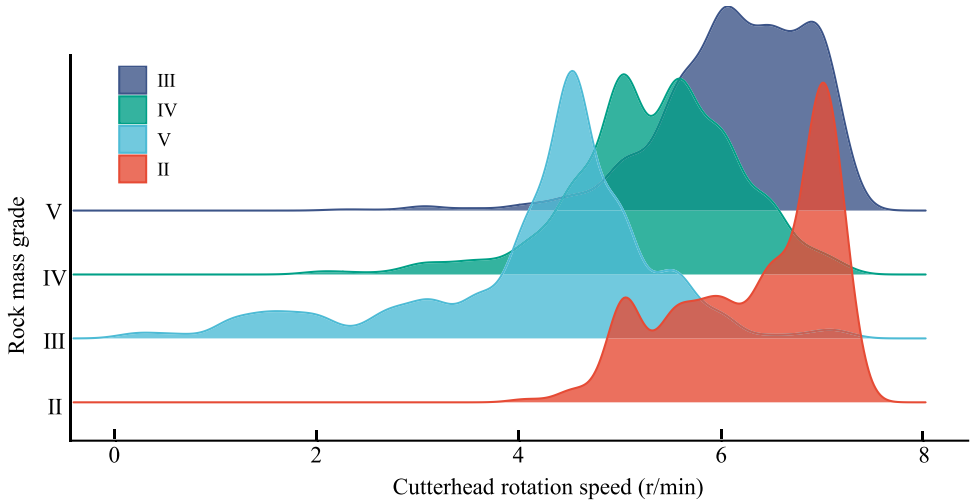
and conduct performance comparison analyses involving the ITELM, TSVD-ITELM, and IEWOA-TSVD-ITELM.

To obtain the optimal hyperparameter combination, this study employed 5-fold cross-validation [17] combined with the IEWOA, where MAE served as the fitness function to minimize errors. Considering the limitations of the computational resources, 30 particles were set with a maximum of 30 iterations. This study demonstrated that using the mean of the data from the first 90 s of the rising phase accurately predicted the mean thrust during the stable phase [17]. Therefore, utilizing the same data input, the impact of different durations of the rising phase on cutterhead torque prediction was analyzed, which is included in the discussion section.

Following the optimization of the cumulative energy ratio for TSVD-ITELM, it was ultimately set to 0.99. Compared with ITELM, TSVD-ITELM significantly enhanced the generalization capability of the test set, achieving an R^2 of 0.634 and an RMSE of 444.980, as listed in Table 4. The performance degradation in the ITELM is attributed to the



(a) Correlation analysis of input features and cutterhead torque



(b) Distribution of cutterhead rotation speed under different rock mass grades

FIGURE 8. **Feature correlation analysis and data distribution of cutterhead rotation speed.**

random generation of weights and biases in the third hidden layer, which introduces errors. TSVD-ITELM optimizes the performance by retaining only the top k singular values from the output matrix of the third hidden layer and zeroing the remaining singular values. This approach not only streamlines the data but also minimizes errors induced by randomness, thereby improving the model performance. Next, the performance of the optimized TSVD-ITELM is analyzed to further validate its effectiveness and robustness.

Compared to IEWOA-TSVD-ITELM, TSVD-ITELM exhibited superior performance on the training set, specifically reflected by an R^2 of 0.651, MAE of 319.174, and RMSE of 434.611. However, when evaluating the results on the test set, IEWOA-TSVD-ITELM demonstrated stronger generalization capabilities, with an R^2 of 0.644, MAE of 326.623, and RMSE of 435.821, as shown in Figure 9. These results indicate that the optimization of randomly generated weights and biases in the ITELM

TABLE 2. Twenty-five standard benchmark functions.

Function	Range
$F_1(x)=\sum_{i=1}^{D}\lvert x_i\rvert+\prod_{i=1}^{D}\lvert x_i\rvert$	[-10, 10]
$F_2(x)=\sum_{i=1}^{D}\left[100\left(x_{i+1}-x_i^{2}\right)^{2}+\left(x_i-1\right)^{2}\right]$	[-30, 30]
$F_3(x)=\max_{1\le i\le D}\lvert x_i\rvert$	[-100, 100]
$F_4(x)=\sum_{i=1}^{D}ix_i^{4}+random[0,1)$	[-1.28, 1.28]
$F_5(x)=\sum_{i=1}^{D}\left(\sum_{j=1}^{i}x_j\right)^{2}$	[-100, 100]
$F_6(x)=\left(\lfloor\lvert x_i+0.5\rvert\rfloor\right)^{2}$	[-100, 100]
$F_7(x)=\sum_{i=1}^{11}\left[a_i-\frac{x_1(b_i^{2}+b_ix_2)}{b_i+b_ix_3+x_4}\right]^{2}$	[-5, 5]
$F_8(x)=\left(\frac{1}{500}+\sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}(x_i-a_{ij})^{6}}\right)^{-1}$	[-65.536, 65.536]
$F_9(x)=\left(x_2-\frac{5.1x_1^{2}}{4\pi^{2}}+\frac{5x_1}{\pi}-6\right)^{2}+10\left(1-\frac{1}{8\pi}\right)\cos(x_1)+10$	[-5, 10]
$F_{10}(x)=c_i\exp\left(-\sum_{j=1}^{n}a_{ij}\left(x_j-p_{ij}\right)^{2}\right)$	[0, 1]
$F_{11}(x)=-\sum_{i=1}^{m}\frac{1}{c_i+\sum_{j}^{D}\left(x_j-a_{ij}\right)^{2}}$	[0, 10]
$F_{12}(x)=4x_1^{2}-2.1x_1^{4}+\frac{1}{3}x_1^{6}+x_1x_2-4x_2^{2}+4x_2^{4}$	[-5, 5]
$F_{13}(x)=-\sum_{i=1}^{m}c_i\exp\left(-\sum_{j=1}^{n}a_{ij}\left(x_j-p_{ij}\right)^{2}\right)$	[0, 1]
$F_{14}(x)=\sum_{i=1}^{D}x_i^{2}+\left(\sum_{i=1}^{D}0.5ix_i\right)^{2}+\left(\sum_{i=1}^{D}0.5ix_i\right)^{4}$	[-5, 10]
$F_{15}(x)=\sum_{i=1}^{D}ix_i^{2}$	[-10, 10]
$F_{16}(x)=\sum_{i=1}^{D}\lvert x_i\sin(x_i)+0.1x_i\rvert$	[-10, 10]
$F_{17}(x)=\sum_{i=1}^{D}\sin(x_i)\sin^{2m}\left(\frac{ix_i^{2}}{\pi}\right)$	[0,]
$F_{18}(x)=\exp\left(-0.5\sum_{i=1}^{D}x_i^{2}\right)$	[-1, 1]
$F_{19}(x)=0.5+\frac{\sin^{2}\left(\sqrt{x_1^{2}+x_2^{2}}-0.5\right)}{[1+0.001(x_1^{2}+x_2^{2})]^{2}}$	[-100, 100]
$F_{20}(x)=x_1^{2}+10^{6}\sum_{i=2}^{D}x_i^{2}$	[-100, 100]
$F_{21}(x)=x_1^{2}+2x_2^{2}-0.3\cos\left(3\pi x_1\right)-0.4\cos\left(3\pi x_2\right)+0.7$	[-50, 50]
$F_{22}(x)=\sum_{i=1}^{D}\left(10^{6}\right)^{\frac{i-1}{D-1}}x_i^{2}$	[-100, 100]

TABLE 2. *(Continued.)* **Twenty-five standard benchmark functions.**

$F_{23}(x) = -\dfrac{1+\cos\left(12\sqrt{x_1^2+x_2^2}\right)}{0.5\left(x_1^2+x_2^2\right)+2}$	[-5.12, 5.12]
$F_{24}(x) = 0.1\sum_{i=1}^{D}\cos\left(5\pi x_i\right) - \sum_{i=1}^{D} x_i$	[-1, 1]
$F_{25}(x) = \dfrac{\pi}{D}\left[10\sin^2(\pi y_1 - 1)^2\right]\left[1+10\sin^2(\pi y_{1+1}) + (yD-1)^2\right]$	[-10, 10]

model using the IEWOA significantly reduces the prediction errors.

B. COMPARISON OF DIFFERENT MODEL RESULTS

To evaluate the performance of the proposed model, several classical machine learning models were incorporated for comparison, including ELM, TELM, least absolute shrinkage and selection operator (Lasso), decision tree, support vector machine (SVM), and light gradient boosting machine (LightGBM). During the modeling process, the Lasso, decision Tree, and SVM were implemented using the scikit-learn library [42], whereas the LightGBM regression model was developed using the Lightbm library [41]. Subsequently, the Lasso, SVM, and decision tree, and LightGBM are each briefly introduced.

Lasso is used to enhance prediction accuracy and interpretability in torque prediction. By imposing a constraint on the sum of the absolute values of the coefficients, Lasso encourages sparsity, effectively reducing the number of predictors and focusing on the most relevant features among the 12 selected features. SVM is employed for regression in torque prediction, constructing hyperplanes in a high-dimensional space to predict continuous values accurately. The algorithm identifies critical data points, or support vectors, ensuring robust and precise regression outcomes even with complex feature sets. The decision tree predicts the torque result by recursively splitting the data based on 12 selected features. This tree-like structure helps efficiently segment the data, making the prediction process transparent and interpretable, although it may require techniques like pruning to avoid overfitting. LightGBM is well-suited for torque prediction due to its efficiency and scalability in handling large datasets. Despite its computational intensity, LightGBM achieves high prediction accuracy by growing trees leaf-wise and employing advanced optimization techniques, making it effective for processing the 12 selected features with substantial predictive power.

The performances of the ELM and TELM models are directly related to the number of neurons. In general, an increase in the number of neurons enhances the performance of the model. However, an excessive number of neurons can lead to model overfitting [38]. Through a comparative analysis, the optimal number of neurons for both the ELM and TELM models was determined to be 350. For models such as Lasso, decision tree, SVM, and LightGBM, the IEWOA and five-fold cross-validation were employed to optimize the hyperparameters, aiming to identify the combinations that yield the best predictive performance. The ranges of the hyperparameters optimized for each model are listed in Table 5.

As shown in Table 6, among the six compared models, TELM exhibits the best performance for the test set, with an R^2 value of 0.632, MAE of 330.784, and RMSE of 443.382. This is closely followed by the LASSSO, which achieves an R^2 of 0.596, an MAE of 358.853, and an RMSE of 464.587. By contrast, the SVM performs the poorest, with an R^2 of only 0.297, MAE of 507.651, and RMSE of 612.804, clearly indicating that the SVM is not suitable for predicting the cutterhead torque. Additionally, the training duration for SVM extends up to 15 min, far exceeding that of other models, such as Lasso, which requires only 0.36 min, LightGBM at 11 min, and decision tree at 0.38 min. LightGBM revealed significant overfitting when the predictions on the training and test sets were compared. Meanwhile, ELM, owing to its structure containing only one hidden layer, underperforms compared to TELM. The prediction results of the six models are shown in Figure 10.

Compared with the TELM, the ITELM demonstrated a suboptimal performance on the test set. This shortfall is attributed to the random generation of weights and biases in the third hidden layer of the ITELM, which disrupts effective computation with the output matrix from the second hidden layer. Although an additional hidden layer was incorporated, incongruent parameters failed to improve model performance. This underscores the potential of the TSVD and IEWOA optimization strategies to substantially enhance the effectiveness of the ITELM.

Compared with the TELM, the ITELM shows a suboptimal performance on the test set. This deficiency arises from the random generation of weights and biases in the third hidden layer of the ITELM, leading to a disruption in the effective computation of the output matrix from the second hidden layer. Despite the addition of an extra hidden layer, the incongruent parameters failed to enhance model performance. This highlights the potential of the TSVD and IEWOA optimization strategies to substantially improve the effectiveness of the ITELM.

TABLE 3. Statistical results of six optimization algorithms for 25 benchmark functions.

F	-	PSO	GWO	WOA	LWOA	EWOA	IEWOA
F1	Mean	3.75E+01	2.90E-16	1.18E-54	5.30E-02	6.08E-198	**0.00E+00**
	Std	3.86E+01	1.59E-16	4.09E-54	3.15E-02	**0.00E+00**	**0.00E+00**
F2	Mean	3.07E+02	2.70E+01	1.16E+01	**2.51E-02**	1.61E-01	5.81E-02
	Std	2.89E+02	6.60E-01	1.32E+01	**3.15E-02**	2.90E-01	7.79E-02
F3	Mean	8.52E+00	7.70E-07	2.80E-09	2.49E-03	6.77E-194	**0.00E+00**
	Std	1.53E+00	8.00E-07	1.33E-08	1.99E-03	**0.00E+00**	**0.00E+00**
F4	Mean	5.88E-02	2.12E-03	6.93E-04	6.90E-03	2.25E-04	**1.59E-04**
	Std	1.91E-02	1.43E-03	8.09E-04	5.13E-03	2.17E-04	**1.22E-04**
F5	Mean	2.90E+03	3.49E-05	2.85E+03	5.35E+03	**0.00E+00**	**0.00E+00**
	Std	1.46E+03	1.19E-04	6.28E+03	5.45E+03	**0.00E+00**	**0.00E+00**
F6	Mean	9.00E-01	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**
	Std	1.14E+00	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**
F7	Mean	4.51E-03	4.51E-03	**1.12E-03**	1.14E-03	5.05E-03	1.17E-03
	Std	1.13E-02	7.93E-03	2.73E-03	7.09E-04	7.29E-03	**4.28E-04**
F8	Mean	9.98E-01	4.23E+00	3.32E+00	1.98E+00	1.89E+00	**1.06E+00**
	Std	1.11E-16	4.02E+00	3.37E+00	1.96E+00	1.36E+00	**3.42E-01**
F9	Mean	**3.98E-01**	3.98E-01	3.98E-01	3.98E-01	3.98E-01	3.99E-01
	Std	**0.00E+00**	2.05E-06	2.28E-05	9.70E-06	4.48E-04	2.28E-03
F10	Mean	**-3.86E+00**	**-3.86E+00**	-3.81E+00	-3.84E+00	-3.79E+00	-3.74E+00
	Std	**2.55E-15**	3.02E-03	6.25E-02	5.59E-02	8.95E-02	1.56E-01
F11	Mean	-7.83E+00	-9.28E+00	-7.96E+00	-1.01E+01	-1.02E+01	**-1.03E+01**
	Std	3.44E+00	2.54E+00	2.82E+00	1.37E+00	3.00E-01	**2.11E-01**
F12	Mean	**-1.03E+00**	**-1.03E+00**	**-1.03E+00**	**-1.03E+00**	**-1.03E+00**	**-1.03E+00**
	Std	**6.08E-16**	3.43E-08	2.53E-09	5.33E-06	1.15E-05	6.55E-04
F13	Mean	-3.28E+00	**-3.28E+00**	-3.05E+00	-3.17E+00	-2.92E+00	-2.72E+00
	Std	**6.31E-02**	6.47E-02	1.30E-01	1.07E-01	1.88E-01	4.21E-01
F14	Mean	1.36E+02	1.32E-07	3.76E+02	5.86E+01	**0.00E+00**	**0.00E+00**
	Std	9.53E+01	2.96E-07	1.59E+02	4.15E+01	**0.00E+00**	**0.00E+00**
F15	Mean	1.34E+01	2.42E-28	9.12E-83	8.05E-04	**0.00E+00**	**0.00E+00**
	Std	3.40E+01	2.38E-28	4.88E-82	1.02E-03	**0.00E+00**	**0.00E+00**
F16	Mean	2.65E-01	3.89E-04	9.94E-54	1.78E-01	3.26E-196	**0.00E+00**
	Std	1.15E+00	4.91E-04	5.26E-53	7.97E-01	**0.00E+00**	**0.00E+00**
F17	Mean	**-1.80E+00**	**-1.80E+00**	-1.76E+00	**-1.80E+00**	-1.65E+00	-1.74E+00
	Std	**8.88E-16**	4.68E-06	1.75E-01	1.89E-05	2.78E-01	1.77E-01
F18	Mean	**-1.00E+00**	**-1.00E+00**	**-1.00E+00**	**-1.00E+00**	**-1.00E+00**	**-1.00E+00**
	Std	3.44E-05	2.03E-17	0.00E+00	1.45E-05	**0.00E+00**	**0.00E+00**
F19	Mean	**2.46E-03**	**2.46E-03**	3.72E-03	**2.46E-03**	**2.46E-03**	**2.46E-03**
	Std	**5.17E-17**	3.23E-10	4.74E-03	3.80E-09	1.82E-13	1.03E-07
F20	Mean	3.79E+05	5.68E-22	2.29E-76	8.75E+01	**0.00E+00**	**0.00E+00**
	Std	2.32E+05	5.85E-22	1.23E-75	8.74E+01	**0.00E+00**	**0.00E+00**
F21	Mean	**0.00E+00**	**0.00E+00**	**0.00E+00**	4.84E-05	**0.00E+00**	**0.00E+00**
	Std	**0.00E+00**	**0.00E+00**	**0.00E+00**	7.01E-05	**0.00E+00**	**0.00E+00**
F22	Mean	7.39E+00	1.34E-30	1.31E-83	9.37E-06	**0.00E+00**	**0.00E+00**
	Std	1.18E+01	1.83E-30	5.89E-83	8.13E-06	**0.00E+00**	**0.00E+00**
F23	Mean	**-1.00E+00**	-9.94E-01	-9.85E-01	-9.79E-01	**-1.00E+00**	**-1.00E+00**
	Std	**0.00E+00**	1.91E-02	2.70E-02	3.00E-02	**0.00E+00**	**0.00E+00**
F24	Mean	-2.90E+00	-3.00E+00	-3.00E+00	-3.00E+00	**-3.00E+00**	**-3.00E+00**
	Std	1.17E-01	2.15E-16	0.00E+00	4.16E-04	**0.00E+00**	**0.00E+00**
F25	Mean	2.46E-02	4.02E-02	1.24E-03	**5.57E-06**	1.53E-04	3.60E-05
	Std	9.74E-02	2.07E-02	1.06E-03	**4.82E-06**	1.99E-04	8.16E-05

The performance of the Lasso model, second only to TELM, indicates that the Lasso regularization strategy prevents overfitting while maintaining good predictive performance when handling high-dimensional data. The Lasso,

TABLE 4. Prediction results of ITELM, TSVD-ITELM, and IEWOA-TSVD-ITELM in the datasets.

Model	Datasets	R^2	MAE	RMSE
ITELM	Training set	0.651	319.174	434.611
	Test set	0.627	326.504	446.204
TSVD-ITELM	Training set	0.634	333.772	444.980
	Test set	0.634	333.618	442.049
IEWOA-TSVD-ITELM	Training set	0.636	330.374	443.828
	Test set	**0.644**	**326.623**	**435.821**

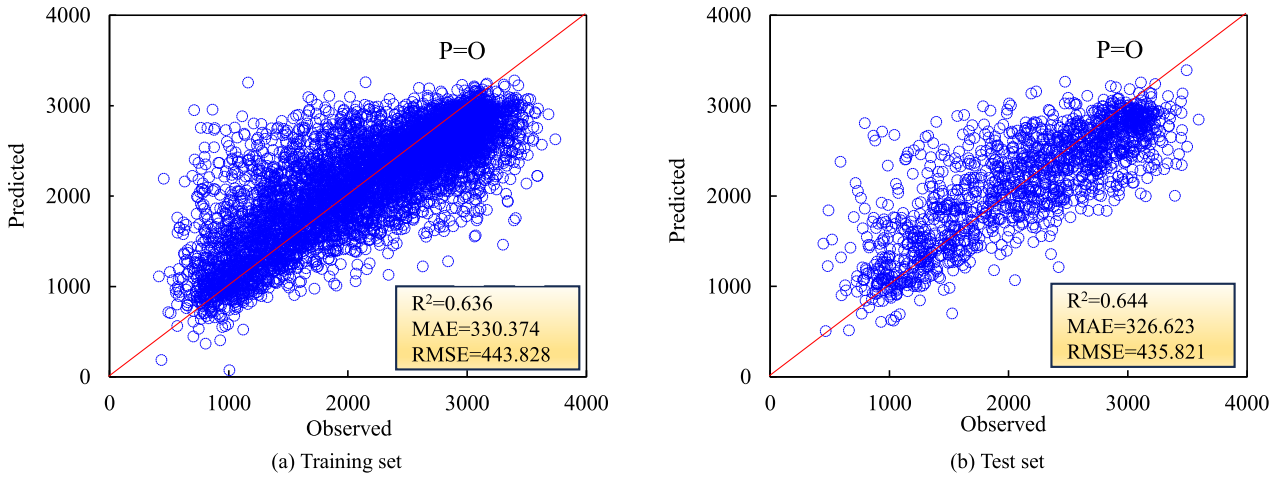


FIGURE 9. Scatter plots comparing the predicted and observed cutterhead torque for IEWOA-TSVD-ITELM.

TABLE 5. Optimized hyperparameters and optimal hyperparameters of six models.

	Hyperparameters	Values	Type	Optimal hyperparameter
ELM	Number of neurons	[150,200,250,300,350]	Integer	350
TELM	Number of neurons	[150,200,250,300,350]	Integer	350
Lasso	L_1 term	[1,500]	Integer	92
LightGBM	Maximum depth	[1,20]	Integer	14
	Number of estimators	[50,300]	Integer	236
	Number of leaves	[1,200]	Integer	99
	Learning rate	[0.01,0.5]	Continuous	0.49
Decision tree	Maximum depth	[1,20]	Integer	18
	Minimum number of samples when split	(1,20]	Integer	10
	Minimum number of samples per leaf	(1,20]	Integer	20
Support vector machine	Kernel	RBF	Categorical	RBF
	Penalty coefficient	[1,100]	Continuous	62

being a linear model, reduces model complexity and prevents overfitting by incorporating an L_1 regularization term, thus performing well with high-dimensional data. Its linear nature simplifies the optimization process, resulting in a shorter optimization time (0.36 min). The SVM model performed the worst in this study, probably because it struggled with nonlinear and high-dimensional data. Additionally, the optimization process for the SVM requires computing the kernel matrix, which is computationally intensive for large datasets, resulting in a training time of 15 min. These factors limit the applicability of SVM for real-time predictions. Although LightGBM performed well on the training set, it showed overfitting on the test set. This is likely because LightGBM, although highly capable of learning complex data patterns, tends to outperform the training data, leading to insufficient generalization. The complex structure of the

TABLE 6. Prediction results of six models in training set and test set.

Model	Datasets	R^2	MAE	RMSE
ELM	Training set	0.637	332.499	443.111
	Test set	0.592	349.281	466.963
TELM	Training set	0.661	319.914	428.569
	Test set	0.632	330.786	443.382
Lasso	Training set	0.555	380.638	490.563
	Test set	0.596	358.853	464.587
LightGBM	Training set	0.999	0.958	1.351
	Test set	0.583	350.621	471.855
Decision tree	Training set	0.692	301.751	408.402
	Test set	0.569	358.374	479.956
Support vector machine	Training set	0.274	520.159	626.746
	Test set	0.297	507.651	612.804

LightGBM and the optimization of multiple parameters require more computational resources and time, resulting in a longer optimization time (11 min).

VII. DISCUSSION

A. ANALYSIS ON THE IMPACT OF RISING PHASE TIMING LENGTH ON IEWOA-TSVD-ITLEM PERFORMANCE

To explore the relationship between the duration of the rising phase and the predictive capability for the cutterhead torque, this section presents a detailed analysis. As shown in Figure 11 (a), the timing length of the rising phase typically exceeded 90 s for most tunneling cycles. However, approximately 600 tunneling cycles exhibited a rising phase duration of less than 60 s, prompting further investigation into the impact of shorter timing lengths on the accuracy of the cutterhead torque predictions. Consequently, this section examines the specific effects of varying the timing lengths of the rising phase (30, 60, and 90 s) on the performance of the IEWOA-TSVD-ITELM.

As shown in Figure 12, as the timing length of the rising phase increases, both MAE and RMSE exhibit declining trends in the training and test datasets, thereby confirming the positive effect of extending the timing length of the rising phase to enhance the model predictive performance. A comparison between the prediction results for the 30–60 s and 60–90 s intervals revealed notable differences in the extent of error reduction. Specifically, when the timing length of the input duration was increased from 30 to 60 s, the MAE and RMSE of the test dataset decreased by 11.82% and 9.56%, respectively, whereas increasing the timing length from 60 to 90 s resulted in reductions of 6.51% and 4.71% in the MAE and RMSE, respectively. However, as the timing length continued to increase, the rate of improvement in the model performance decreased.

These findings provide invaluable guidance for tunnel boring operations and have significant practical application potential. By flexibly adjusting the duration of the rising phase during construction, contractors can predict the cutterhead torque more effectively, thereby optimizing the boring parameters and enhancing the overall operational efficiency. In particular, contractors can establish detailed operating protocols, adjust the rising phase duration to at least 60 s, and extend it to 90 s when necessary to accommodate varying geological conditions and construction environments. This flexible adjustment can significantly improve the prediction accuracy, reduce mechanical failures, and minimize unplanned downtime.

B. EFFECT OF DIFFERENT ROCK MASS GRADES

The variation in the rock mass grade significantly affects the efficiency of the TBM operation and cutter wear [41]. To further explore the impact of different rock mass grades on the model performance, this study utilized the IEWOA-TSVD-ITELM for analysis, as shown in Figure 13. Given the higher proportions of Classes III and IV in the sample set, 150 predictions related to the cutterhead torque during the tunneling cycles were selected for visualization to clearly demonstrate the model performance. Comparisons of MAE and RMSE revealed that IEWOA-TSVD-ITELM exhibited the smallest prediction errors under Class V, with values of 281.135 and 395.351, respectively, indicating the high adaptability and accuracy of the model under complex geological conditions. Additionally, the model performed relatively well under Class III with MAE and RMSE values of 321.022 and 422.640, respectively. By contrast, under Classes II and IV, the model displayed greater prediction errors, particularly under Class II, where the MAE and RMSE reached 340.171 and 501.355, respectively, indicating a weaker predictive performance for these rock mass grades. Overall, comparisons of model performance revealed a consistent trend between the predicted and actual values, demonstrating that the IEWOA-TSVD-ITELM provides effective predictions across different rock mass grades.

By incorporating the characteristics of different rock mass grades and prediction errors of the model, construction strategies can be optimized to enhance overall efficiency

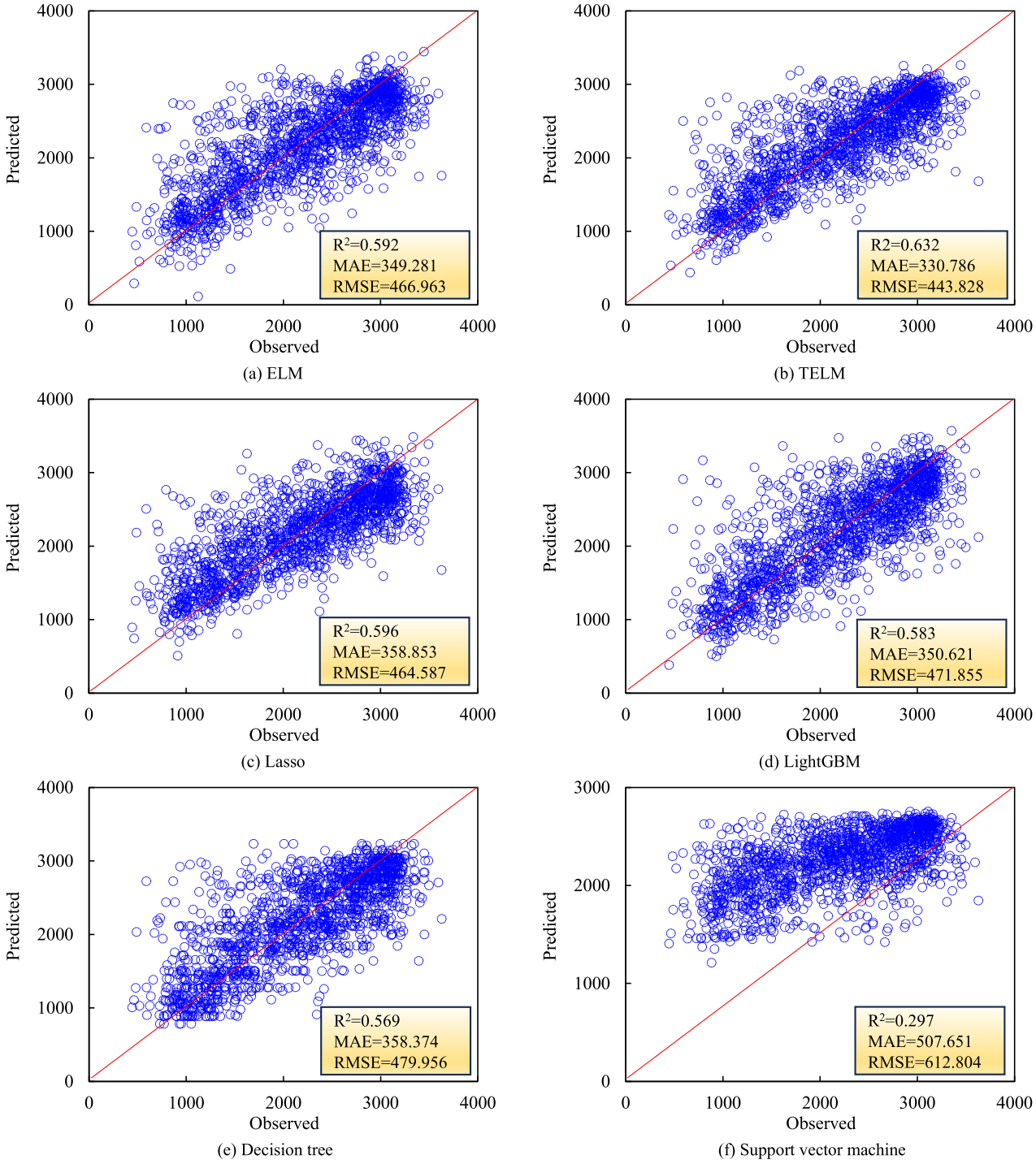
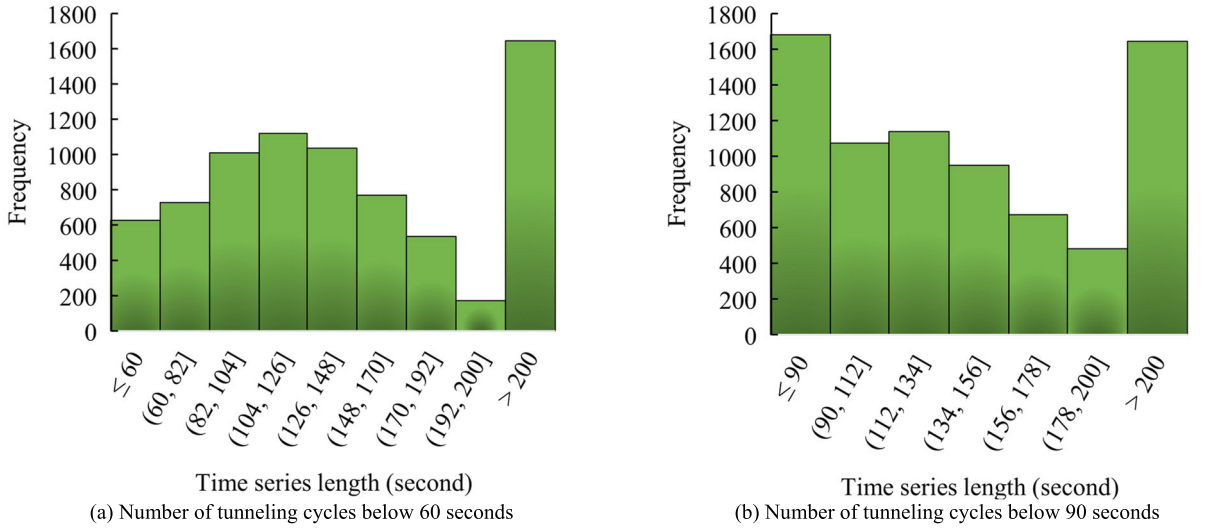


FIGURE 10. Scatter plots comparing the predicted and observed cutterhead torque for six models.

and safety. For example, in the case of Class II, the model exhibited larger prediction errors, indicating a lower accuracy under harder rock conditions; therefore, a more conservative operational strategy should be adopted. Conversely, for Class V, the model demonstrated the highest prediction accuracy, indicating reliable predictions under softer rock conditions, which increased the operational efficiency.

C. DISTRIBUTION OF WEIGHTS BEFORE AND AFTER OPTIMIZATION

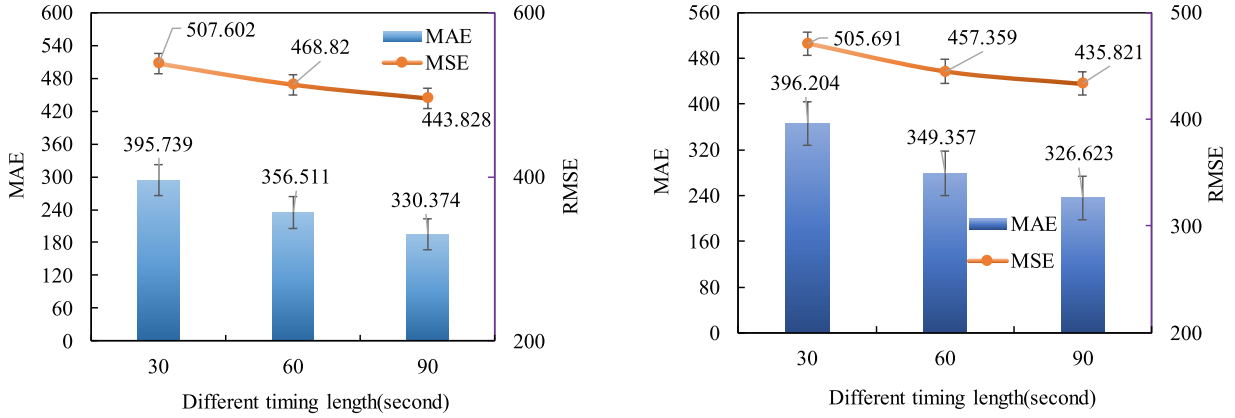
This study aims to address the limitations imposed by the random initialization of weights and biases in traditional TSVD-ITELM by proposing an approach that utilizes the IEWOA to update weights and biases, thereby minimizing errors in the training set and enhancing the generalization ability and predictive accuracy of the model. The core



(a) Number of tunneling cycles below 60 seconds

(b) Number of tunneling cycles below 90 seconds

FIGURE 11. Statistical analysis of rising phase timing length.



(a) Tunneling cycle statistics under 60 seconds

(b) Tunneling cycle statistics under 90 seconds

FIGURE 12. Comparison of prediction results using different timing lengths in the rising phase as input.

concept involves adjusting the distribution of weights during the updating process using an optimization algorithm. Specifically, this study elucidates the changes in weights before and after iterations of the IEWOA-TSVD-ITELM. Before optimization, the weights of the model are normally distributed approximately between −3 and 3, as shown in Figure 14(a). After iterative optimization with IEWOA, the distribution of weights is narrowed to between −1 and 1 and tends toward a more uniform distribution, as shown in Figure 14(b). This transformation in the weight distribution significantly improved the predictive performance of the model.

Cao et al. [43] introduced a novel activation function based on scaling and translation (affine transformation activation function), designed to adjust the weights and biases in the hidden layer to approximate a uniform distribution, thereby enhancing the performance of the TSVD-ITELM. Furthermore, the weight distribution achieved by the proposed IEWOA is similar to the functionality of the affine transformation activation function, highlighting the critical role of optimizing the weight distribution in enhancing the performance of the TSVD-ITELM.

As shown in Figs. 15 (a) and (b), this study conducted a comprehensive analysis of the corresponding weights for the 12 input features. Taking X7 (cutterhead power) as an example, prior to weight optimization, this feature exhibited the broadest data range among all the input features. After the optimization process, the fluctuations in the data related to the weights of the cutterhead power decreased significantly. Furthermore, both the mean and median values show notable changes after optimization. During the optimization process of the TSVD-ITELM, the model utilizes error feedback for adaptive weight adjustment, significantly enhancing the overall performance of the model. This analysis provides a new perspective for understanding and improving the dynamism of weight adjustment mechanisms in the TSVD-ITELM.

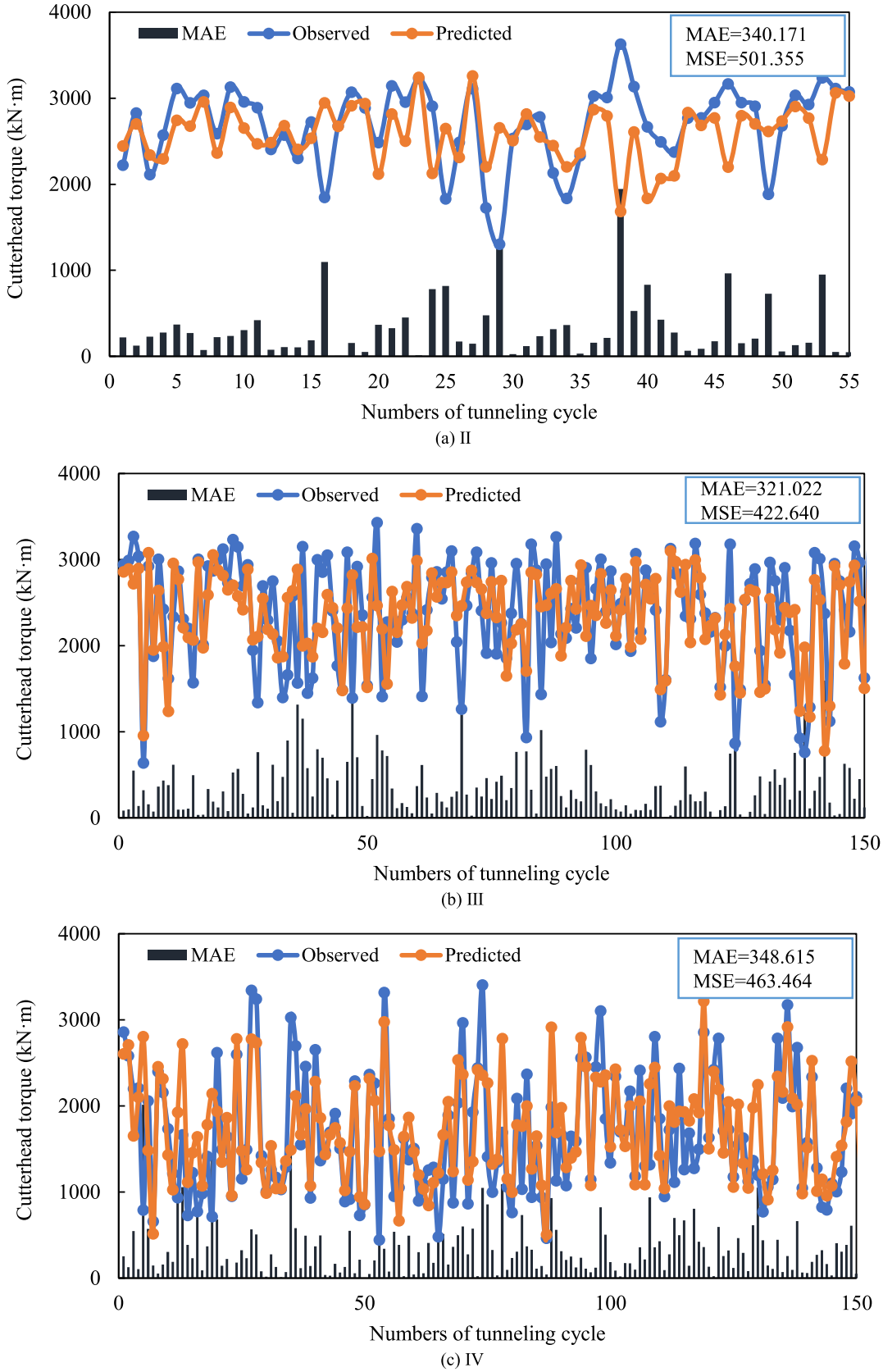


FIGURE 13. Prediction results of IEWOA-TSVD-ITELMA under different rock mass grades.

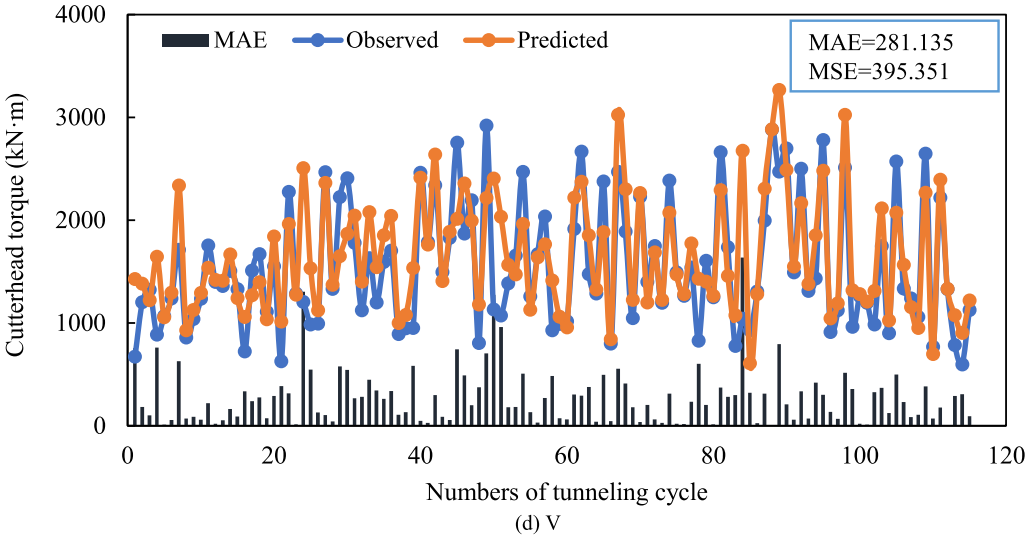


FIGURE 13. *(Continued.)* **Prediction results of IEWOA-TSVD-ITELMA under different rock mass grades.**

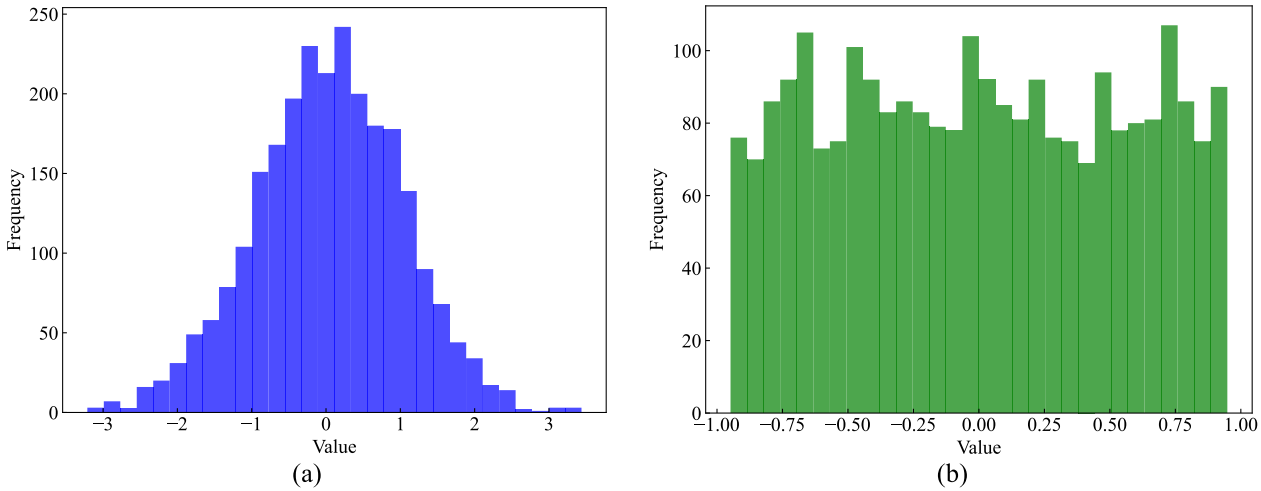


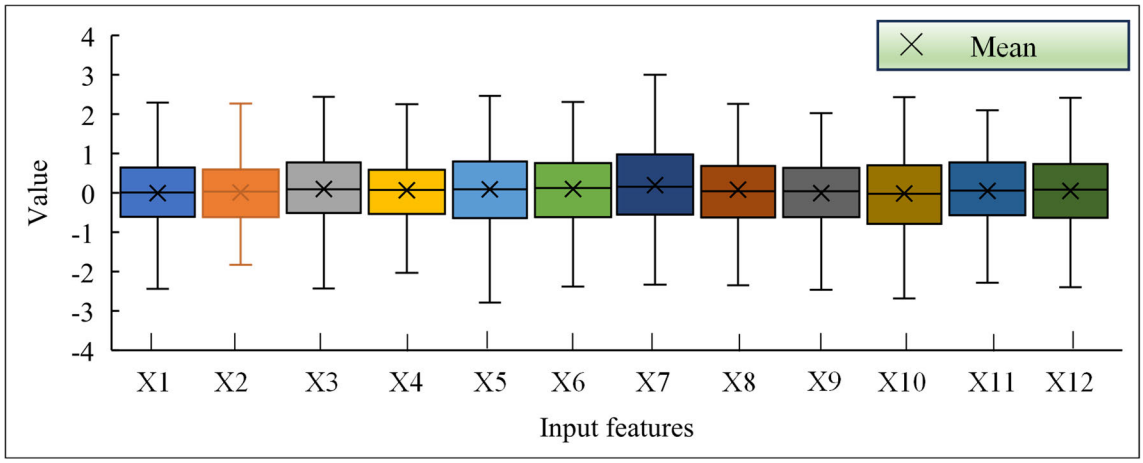
FIGURE 14. **Comparison of results before and after the IEWOA-TSVD-ITELMA weight optimization.**

D. FURTHER DISCUSSION

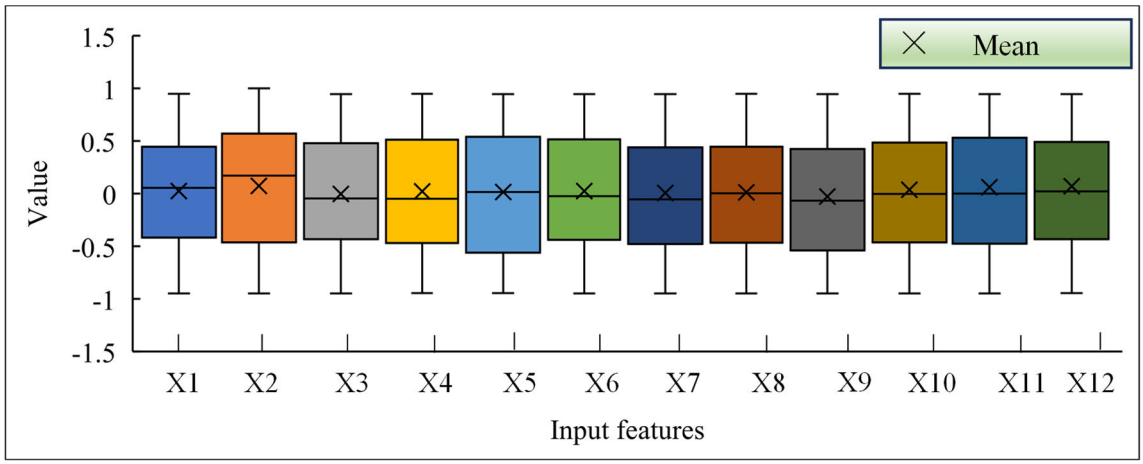
In this study, a hybrid model, IEWOA-TSVD-ITELM, surpassed traditional machine learning models in predicting the cutterhead torque for TBM construction. This study provides invaluable guidance for tunnel boring operations. By flexibly adjusting the rising phase duration between 60 and 90 s based on geological conditions, contractors can optimize boring parameters, improve prediction accuracy, reduce mechanical failures, and minimize downtime. The model performance varies with the rock mass grades, suggesting conservative strategies for harder conditions (Class II) and efficient operations for softer conditions (Class V). Incorporating this model into control systems allows real-time adjustments of construction parameters, such as the cutterhead rotation speed, which significantly reduces downtime and wear, thus lowering maintenance costs. Additionally, by adjusting the input data length, the model enhanced anomaly detection, such as cutter head jamming, thereby enabling timely preventive actions. Overall, the IEWOA-TSVD-ITELM model ensured construction safety and significantly enhanced the economic and safety outcomes of engineering projects.

E. LIMITATION AND FUTURE WORK

This study utilized the mean values of the rising phase data but did not investigate the impact of temporal models on the cutterhead torque (such as LSTM). Its primary objective was to simplify the training time and provide a method for rapidly predicting the cutterhead torque. This method has not yet been applied to other engineering projects because of the significant differences in the cutterhead diameter, number of cutters, and tunneling parameters, requiring model retraining for each new project. Additionally, this study failed to explore the effectiveness of the model under extreme conditions (e.g., fault impacts), primarily because collecting and annotating relevant station information is extremely difficult. Future studies will focus on overcoming these



(a) Before the weight optimization of TSVD-ITELM



(b) After the weight optimization of TSVD-ITELM

FIGURE 15. **Comparison of results before and after the optimization of weights corresponding to different input features.**

limitations by gathering diverse TBM engineering data and establishing datasets that are adaptable to extremely complex conditions, thereby enhancing the generalization capabilities of the model and providing more precise guidance for TBM construction.

VIII. CONCLUSION

This study introduces a novel hybrid machine learning model based on the TELM framework to enhance cutterhead torque prediction accuracy. The model leverages the Softsign function for nonlinear transformation in the second hidden layer, preserving crucial information. A third hidden layer with variable neurons improves generalization, while TSVD reduces data noise and overfitting by retaining key singular values. EWOA enhances adaptive search capabilities with innovative methods for calculating convergence factors and dynamic weight parameters, iteratively updating weights and biases to reduce training errors.

The proposed IEWOA further enhances EWOA's performance by incorporating new position-updating and adaptive adjustment strategies. Comparative analysis of 25 benchmark functions reveals that IEWOA excels in 20, outperforming PSO, GWO, WOA, LWOA, and EWOA. The IEWOA-TSVD-ITELM model achieves an R^2 of 0.644, MAE of 326.623, and RMSE of 435.821 on the test set. TSVD enhances ITELM by retaining key singular values, while IEWOA optimizes weights and biases to minimize prediction errors.

Among the models tested, the Lasso model ranks second due to its effective L_1 regularization, preventing overfitting and maintaining predictive accuracy with high-dimensional data, optimized in just 0.36 minutes. In contrast, the SVM model struggles with nonlinear, high-dimensional data, taking 15 minutes for training due to kernel matrix computation. LightGBM overfits the test set and requires 11 minutes for optimization, indicating insufficient generalization.

These results underscore the applicability of the IEWOA-TSVD-ITELM model in predicting cutterhead torque. Extending the timing from 30 to 60 seconds significantly reduces MAE by 11.82% and RMSE by 9.56%, aiding real-time decision-making. However, further extension to 90 seconds shows diminishing returns, suggesting an optimal timing length for balancing accuracy and computational efficiency.

The model performs best under Class V geological conditions, achieving the lowest prediction errors (MAE=281.135, RMSE=395.351), which is crucial for efficient TBM operation in softer rocks. Class III conditions also perform well (MAE=321.022, RMSE=422.640). However, the model shows increased errors under Class II and IV conditions, indicating the need for careful consideration of its applicability and potential risks during construction.

The IEWOA facilitates iterative optimization, effectively narrowing the weight distribution range within the IEWOA-TSVD-ITELM model, leading to near-uniform distribution. These adjustments significantly improve the model's predictive accuracy. Analysis of the weight distribution across the 12 input features indicates that the model adaptively adjusts weights using error feedback.

In summary, while the IEWOA-TSVD-ITELM model demonstrates strong performance in predicting cutterhead torque, it faces increased errors in Class II and IV geological conditions. Future research should focus on enhancing generalization techniques and developing more efficient algorithms to reduce training time and computational demands, thereby improving the model's applicability and reliability in diverse construction scenarios.

DECLARATIONS

The authors declare no conflicts of interest.

DATA AVAILABILITY

Authors do not have permission to share data.

REFERENCES

[1] J. Fu, D. Wu, H. Lan, Z. Ji, W. Li, and Y. Xia, "Online monitoring and analysis of TBM cutter temperature: A case study in China," *Measurement*, vol. 174, Apr. 2021, Art. no. 109034.

[2] M. Samaei, M. Ranjbarnia, V. Nourani, and M. Z. Naghadehi, "Performance prediction of tunnel boring machine through developing high accuracy equations: A case study in adverse geological condition," *Measurement*, vol. 152, Feb. 2020, Art. no. 107244.

[3] Y. Jin, C. Qin, J. Tao, and C. Liu, "An accurate and adaptative cutterhead torque prediction method for shield tunneling machines via adaptative residual long-short term memory network," *Mech. Syst. Signal Process.*, vol. 165, Feb. 2022, Art. no. 108312.

[4] X. Huang, Q. Zhang, Q. Liu, X. Liu, B. Liu, J. Wang, and X. Yin, "A real-time prediction method for tunnel boring machine cutter-head torque using bidirectional long short-term memory networks optimized by multi-algorithm," *J. Rock Mech. Geotechnical Eng.*, vol. 14, no. 3, pp. 798–812, Jun. 2022.

[5] D. J. Armaghani, M. Koopialipoor, A. Marto, and S. Yagiz, "Application of several optimization techniques for estimating TBM advance rate in granitic rocks," *J. Rock Mech. Geotechnical Eng.*, vol. 11, no. 4, pp. 779–789, Aug. 2019.

[6] L.-C. Deng, Q.-W. Zhuang, X.-Z. Li, and Y.-X. Yuan, "Development and application of a full-scale mechanical rock-cutting platform for measuring the cutting performance of TBM cutter," *Measurement*, vol. 204, Nov. 2022, Art. no. 112036.

[7] H. Shi, H. Yang, G. Gong, and L. Wang, "Determination of the cutterhead torque for EPB shield tunneling machine," *Autom. Construct.*, vol. 20, no. 8, pp. 1087–1095, Dec. 2011.

[8] U. Ates, N. Bilgin, and H. Copur, "Estimating torque, thrust and other design parameters of different type TBMs with some criticism to TBMs used in Turkish tunneling projects," *Tunnelling Underground Space Technol.*, vol. 40, pp. 46–63, Feb. 2014.

[9] Q. Zhang, C. Qu, Y. Kang, G. Huang, Z. Cai, Y. Zhao, H. Zhao, and P. Su, "Identification and optimization of energy consumption by shield tunnel machines using a combined mechanical and regression analysis," *Tunnelling Underground Space Technol.*, vol. 28, pp. 350–354, Mar. 2012.

[10] A. Lislerud, "Hard rock tunnel boring: Prognosis and costs," *Tunnelling Underground Space Technol.*, vol. 3, no. 1, pp. 9–17, Jan. 1988.

[11] E. Avunduk and H. Copur, "Empirical modeling for predicting excavation performance of EPB TBM based on soil properties," *Tunnelling Underground Space Technol.*, vol. 71, pp. 340–353, Jan. 2018.

[12] L. Faramarzi, A. Kheradmandian, and A. Azhari, "Evaluation and optimization of the effective parameters on the shield TBM performance: Torque and thrust—Using discrete element method (DEM)," *Geotechnical Geological Eng.*, vol. 38, no. 3, pp. 2745–2759, Jun. 2020.

[13] M. D. Han, Z. X. Cai, C. Y. Qu, and L. S. Jin, "Dynamic numerical simulation of cutterhead loads in TBM tunnelling," *Tunnelling Underground Space Technol.*, vol. 70, pp. 286–298, Nov. 2017.

[14] C. Su, Y. Wang, H. Zhao, P. Su, C. Qu, Y. Kang, T. Huang, Z. Cai, and L. Wang, "Analysis of mechanical properties of two typical kinds of cutterheads of shield machine," *Adv. Sci. Lett.*, vol. 4, no. 6, pp. 2049–2053, Jul. 2011.

[15] S. Sissins and C. Paraskevopoulou, "Assessing TBM performance in heterogeneous rock masses," *Bull. Eng. Geol. Environ.*, vol. 80, no. 8, pp. 6177–6203, Aug. 2021.

[16] Z. Liu, L. Li, X. Fang, W. Qi, J. Shen, H. Zhou, and Y. Zhang, "Hard-rock tunnel lithology prediction with TBM construction big data using a global-attention-mechanism-based LSTM network," *Autom. Construct.*, vol. 125, May 2021, Art. no. 103647.

[17] L. Li, Z. Liu, Y. Lu, F. Wang, and S. Jeon, "Hard-rock TBM thrust prediction using an improved two-hidden-layer extreme learning machine," *IEEE Access*, vol. 10, pp. 112695–112712, 2022.

[18] J. Li, P. Li, D. Guo, X. Li, and Z. Chen, "Advanced prediction of tunnel boring machine performance based on big data," *Geosci. Frontiers*, vol. 12, no. 1, pp. 331–338, Jan. 2021.

[19] C. Qin, G. Shi, J. Tao, H. Yu, Y. Jin, J. Lei, and C. Liu, "Precise cutterhead torque prediction for shield tunneling machines using a novel hybrid deep neural network," *Mech. Syst. Signal Process.*, vol. 151, Apr. 2021, Art. no. 107386.

[20] C. Xu, X. Liu, E. Wang, and S. Wang, "Prediction of tunnel boring machine operating parameters using various machine learning algorithms," *Tunnelling Underground Space Technol.*, vol. 109, Mar. 2021, Art. no. 103699.

[21] J. Jin, Q. Jin, J. Chen, C. Wang, M. Li, and L. Yu, "Prediction of the tunnelling advance speed of a super-large-diameter shield machine based on a KF-CNN-BiGRU hybrid neural network," *Measurement*, vol. 230, May 2024, Art. no. 114517.

[22] K. Glab, G. Wehrmeyer, M. Thewes, and W. Broere, "Predictive machine learning in Earth pressure balanced tunnelling for main drive torque estimation of tunnel boring machines," *Tunnelling Underground Space Technol.*, vol. 146, Apr. 2024, Art. no. 105642.

[23] X. Wang, H. Zhu, M. Zhu, L. Zhang, and J. W. Ju, "An integrated parameter prediction framework for intelligent TBM excavation in hard rock," *Tunnelling Underground Space Technol.*, vol. 118, Dec. 2021, Art. no. 104196.

[24] W. Sun, M. Shi, C. Zhang, J. Zhao, and X. Song, "Dynamic load prediction of tunnel boring machine (TBM) based on heterogeneous in-situ data," *Autom. Construct.*, vol. 92, pp. 23–34, Aug. 2018.

[25] H. Yu, C. Qin, J. Tao, C. Liu, and Q. Liu, "A multi-channel decoupled deep neural network for tunnel boring machine torque and thrust prediction," *Tunnelling Underground Space Technol.*, vol. 133, Mar. 2023, Art. no. 104949.

[26] H. Jiang, Z. He, G. Ye, and H. Zhang, "Network intrusion detection based on PSO-XGBoost model," *IEEE Access*, vol. 8, pp. 58392–58401, 2020.

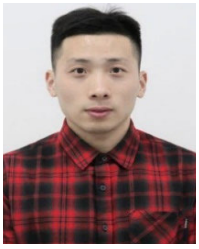
[27] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.

[28] Z. Benmamoun, W. Fethallah, M. Ahlaqqach, I. Jebbor, M. Benmamoun, and M. Elkhechafi, "Butterfly algorithm for sustainable lot size optimization," *Sustainability*, vol. 15, no. 15, p. 11761, Jul. 2023.

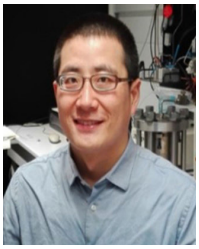
[29] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016.

[30] Q. Han, X. Yang, H. Song, and W. Du, "Multi-objective ship path planning using non-dominant relationship-based WOA in marine meteorological environment," *Ocean Eng.*, vol. 266, Dec. 2022, Art. no. 112862.

[31] H. Chen, Y. Xu, M. Wang, and X. Zhao, "A balanced whale optimization algorithm for constrained engineering design problems," *Appl. Math. Model.*, vol. 71, pp. 45–59, Jul. 2019.
[32] J. Li, W. Chen, K. Han, and Q. Wang, "Fault diagnosis of rolling bearing based on GA-VMD and improved WOA-LSSVM," *IEEE Access*, vol. 8, pp. 166753–166767, 2020.
[33] M. H. Qais, H. M. Hasanien, and S. Alghuwainem, "Enhanced whale optimization algorithm for maximum power point tracking of variable-speed wind generators," *Appl. Soft Comput.*, vol. 86, Jan. 2020, Art. no. 105937.
[34] J. Donnelly, A. Daneshkhah, and S. Abolfathi, "Forecasting global climate drivers using Gaussian processes and convolutional autoencoders," *Eng. Appl. Artif. Intell.*, vol. 128, Feb. 2024, Art. no. 107536.
[35] J. Donnelly, A. Daneshkhah, and S. Abolfathi, "Physics-informed neural networks as surrogate models of hydrodynamic simulators," *Sci. Total Environ.*, vol. 912, Feb. 2024, Art. no. 168814.
[36] J. Donnelly, S. Abolfathi, and A. Daneshkhah, "A physics-informed neural network surrogate model for tidal simulations," in *Proc. ECCOMAS*, 2023, pp. 836–844.
[37] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
[38] B. Y. Qu, B. F. Lang, J. J. Liang, A. K. Qin, and O. D. Crisalle, "Two-hidden-layer extreme learning machine for regression and classification," *Neurocomputing*, vol. 175, pp. 826–834, Jan. 2016.
[39] M. M. Saafan and E. M. El-Gendy, "IWOSSA: An improved whale optimization salp swarm algorithm for solving optimization problems," *Exp. Syst. Appl.*, vol. 176, Aug. 2021, Art. no. 114901.
[40] Q. Liu, J. Liu, Y. Pan, X. Kong, and K. Hong, "A case study of TBM performance prediction using a Chinese rock mass classification system—Hydropower classification (HC) method," *Tunnelling Underground Space Technol.*, vol. 65, pp. 140–154, May 2017.
[41] L. Li, Z. Liu, J. Shen, F. Wang, W. Qi, and S. Jeon, "A LightGBM-based strategy to predict tunnel rockmass class from TBM construction data for building control," *Adv. Eng. Informat.*, vol. 58, Oct. 2023, Art. no. 102130.
[42] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.
[43] J. Cao, K. Zhang, H. Yong, X. Lai, B. Chen, and Z. Lin, "Extreme learning machine with affine transformation inputs in an activation function," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 7, pp. 2093–2107, Jul. 2019.



LONG LI received the Ph.D. degree in mining engineering from Northeastern University, in 2023. His research interests include TBM big data analysis and the application of machine learning in tunnels.



ZAOBAO LIU is currently the Vice Dean of the School of Resources and Civil Engineering, Northeastern University, Shenyang, China. He is also the Executive Deputy Director of the Key Laboratory of Ministry of Education on Safe Mining of Deep Metal Mines, China. His research interests include high-temperature rock mechanics, seepage mechanics of unconventional geological reservoirs, geological disposal of high-level radioactive waste, and deep energy mining. He has been awarded the Changjiang Scholars Program and the First-Class Prize of the Science and Technology Progress Award of the Chinese Society of Rock Mechanics and Engineering (CSRME). He has been the coordinating investigator of dozens of research funds from the Natural Science Foundation of China and China Ministry of Science and Technology, Liaoning Province. He serves as the Vice Chairperson for the Multi-field Coupling Sub-Society of CSRME, a member for the International Society for Rock Mechanics and Rock Engineering (ISRM), and the Editorial Board Member for several journals.