IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 September 2022, accepted 17 October 2022, date of publication 21 October 2022, date of current version 1 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3216294

RESEARCH ARTICLE

Hard-Rock TBM Thrust Prediction Using an Improved Two-Hidden-Layer Extreme Learning Machine

LONG LI[1,2], ZAOBAO LIU[1,2], YUCHI LU[1,2], FEI WANG[1,2], AND SEOKWON JEON[3]

[1]Key Laboratory of Ministry of Education on Safe Mining of Deep Metal Mines, Northeastern University, Shenyang 110819, China
[2]Key Laboratory of Liaoning Province on Deep Engineering and Intelligent Technology, Northeastern University, Shenyang 110819, China
[3]Department of Energy Systems Engineering, Seoul National University, Seoul 08826, South Korea

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

This work was supported in part by the 111 Project under Grant B17009, in part by the Framework of the Sino-Franco Joint Research Laboratory on Multiphysics and Multiscale Rock Mechanics, and in part by the Intelligent Control and Support Software to Safely and Efficiently Operate TBM Tunnels from China Railway Engineering Equipment Group Company Ltd., and in part by the Project Team for the National Basic Research Program under Grant 973.

ABSTRACT It is difficult for tunnel boring machine (TBM) operators to respond for safe and high-efficient construction without accurate reference parameters such as the TBM thrust. A new hybrid model (MRFO-AT-TELM) combining an improved two-hidden-layer extreme learning machine (AT-TELM) and manta ray foraging optimization (MRFO) algorithm is proposed to predict TBM thrust with 12 selected input featuring parameters. The affine transformation (AT) activation function is used to improve the performance of TELM. Input weights and bias of AT-TELM are optimized using the MRFO algorithm. The performance of the proposed model is validated with TBM construction data collected from the Yin-Song Project in China and compared with other models. Input data of the first 30, 60, and 90 seconds of the rising period are analyzed. Results show that the proposed model is superior to the other models and with 90-second data as input outperforms that with 30 and 60-seconds data. The proposed model and the selected input features are validated in a new project. The thrust prediction model can be embedded into the TBM construction intelligence system and thus help improve construction efficiency.

INDEX TERMS Hard rock TBM, construction big data, thrust prediction, two-hidden-layer extreme learning machine.

I. INTRODUCTION

In water diversion, transport, and underground mining projects, tunnel construction is crucial [1], [2]. Tunnel boring machines (TBMs), which have benefits of higher excavation efficiency, significant economic benefits, and lower ground disturbance, have been increasingly applied in engineering practices in recent years [3], [4], [5]. However, the thrust of TBM is unstable under complex geological conditions, causing violent vibration in the cutter. This can easily lead to increased cutter wear, decreased machine availability, and possible unplanned downtime [6]. Thrust is also affected by the equipment's structure due to the complex TBM structure and the high mechanical coupling effect [7]. Thus, accurate prediction of thrust is important to guide the TBM operator to take action for safe and high-efficiency tunnel construction.

The models can be categorized as theoretical models, empirical models, numerical simulation models, and artificial intelligence models for thrust prediction. Most theoretical models are based on rock-breaking mechanisms or full-scale linear cutting tests [8]. The calculation of the thrust is based on parameters such as uniaxial compressive strength, opening ratios of the cutterhead, tensile strength, and borehole diameters [9]. Additional theoretical studies can be found in [10], [11], [12], and [13]. Theoretical modelling usually focuses on analyzing the thrust of a single cutter. Therefore, the dynamic process of rock-breaking between the TBM and surrounding rock cannot be presented [14].

The associate editor coordinating the review of this manuscript and approving it for publication was Chong Leong Gan.

Empirical models are statistical interpretations of previous TBM construction data, most of which are based on linear regression or nonlinear regression. However, the quality of data is crucial to the reliability and accuracy of empirical models [10]. Furthermore, empirical methods can only have a range of values for thrust [15].

Numerical methods have advantages in simulating tunnel excavation. The most commonly seen methods are the discrete element method (DEM), the finite difference method (FDM), and the finite element method (FEM) [16]. Compared with theoretical and empirical models, numerical methods additionally consider parameters such as elastic moduli and initial stress. However, numerical methods generally require complex computational procedures and spend much time estimating thrusts. Overall, the above models cannot predict the real-time thrust owing to the complexity and specificity of TBM construction [17].

Excavation (or construction) time required for a single segment ring in TBM tunneling can be divided into four parts: empty push period, rising period, stable period and shutdown period [14]. The rising period is a key process in which the operator controls the TBM to set the parameters of the stable period to achieve safe excavation [18]. Studies have shown that the TBM tunneling parameters of the stable period will not change significantly when geological conditions do not change significantly [19]. This provides a scientific basis for predicting the average thrust in the stable period by using the data in the rising period. On this basis, some scholars use the first 30-second data in the rising period to predict the stable period thrust [14]. However, the accuracy of predicting thrust based on the first 30-second data in the rising period is still insufficient. Because the 30-second data contains less information, resulting in poor performance of the model. Therefore, this study predicts the thrust in advance based on the first 30, 60, and 90 seconds of the rising period. For the convenience of calculation, we did not consider 30s continuous data, but averaged the data as input.

The fast-growing artificial intelligence offers a powerful tool for TBM performance prediction [20], [21], [22], [23]. For thrust prediction, fuzzy logic models [24], recurrent neural networks [15], bayesian ridge regression, back-propagation neural network [25], and long short-term memory [14] have been established [26]. Artificial intelligence models have been widely applied in thrust prediction. However, Different models have different disadvantages. For example, the fuzzy model is usually applied to incomplete data [27]. Neural networks usually require a lot of time to train the model, especially when optimizing hyperparameters. Most models take more time to optimize the hyperparameters. At present, few studies have applied extreme learning machine (ELM) to thrust prediction. Input weights and biases of ELM are set as random, which eliminates the need for back-propagation updates [28]. The hyper-parameters of ELM are only needed to optimize the neurons in the hidden layer [29]. The ELM has attracted attention due to its fast training process in many fields [29], [30]. However, the performance of ELM is still unsatisfactory with a single hidden layer network [31]. Therefore, a two-hidden-layer extreme learning machine (TELM) has been proposed [32].

However, the Hyperbolic tangent activation function used is the same in the two hidden layers of TELM, resulting in a reduced generalization performance of the model [33]. The role of the activation function is to make the machine learning model have nonlinear fitting capability. A new activation function, named affine transformation (AT) activation function, is used in the first hidden layer to improve the performance of TELM in predicting thrust. Hyperbolic tangent is still used in the activation function of the second hidden layer to ensure computational stability. The proposed model is called AT-TELM. More details will be discussed in the methodology. Randomly generated weights and biases lead to unstable model performance. The weights and biases of AT-TELM can be tuned by the optimization algorithm.

There are many algorithms used to optimize weights and biases of ELM [34], [35], such as particle swarm optimization, whale optimization algorithm, fruit fly optimization algorithm, lion swarm optimization algorithm, seagull optimization algorithm, and sparrow search algorithm. Compared with the above optimization algorithms, the manta ray foraging optimization (MRFO) algorithm [36] has the advantages of a simple structure, fewer parameters, and strong global search capability. Studies have shown that the performance of the MRFO-ELM is higher than that of other optimization algorithms [37]. Therefore, the MRFO algorithm is used to optimize the weights and biases of AT-TELM in this study.

The contributions of this study can be summarized as: (1) a new hybrid model, named MRFO-AT-TELM, is proposed; (2) the effects of different time series lengths of the rising period (such as the first 30-second, 60-second, and 90-second data) on the thrust prediction results are analyzed; (3) the transfer of the model to a new engineering project is analyzed. The rest of the paper is organized as follows. Section 2 introduces the proposed MRFO-AT-TELM model. Section 3 presents the data set and feature selection. Section 4 compares the results of different models predicting thrust. Section 5 discusses the advantages and disadvantages of the model and the limitations of this paper. Finally, Section 6 summarizes the findings.

II. METHODS

A. EXTREME LEARNING MACHINE

The method of ELM is first developed by Huang et al. [28] to avoid time-consuming backward iterative training and improve performance overall. ELM is a Single hidden Layer Feedforward Network (SLFN) with random weights and bias between the input layer and the hidden layer. ELM is widely utilized in various fields due to its incredibly easy realization, and quick training speed [38].

A dataset of N samples is given $(\boldsymbol{i}_p, \boldsymbol{o}_p)$ $(p = 1, 2, \ldots, N)$, there is an input matrix $\mathbf{I} = [\boldsymbol{i}_1, \boldsymbol{i}_2, \ldots, \boldsymbol{i}_N]^{\mathrm{T}}$ and an output matrix $\mathbf{O} = [\boldsymbol{o}_1, \boldsymbol{o}_2, \ldots, \boldsymbol{o}_N]$, where $\boldsymbol{i}_p = \left[i_{p1}, i_{p2}, \ldots, i_{pn}\right]^{\mathrm{T}}$, $\boldsymbol{o}_p = \left[o_{p1}, o_{p2}, \ldots, o_{pm}\right]^{\mathrm{T}}$ and n and m represent the

dimensions of input and output, respectively. Firstly, the ELM randomly assigns the bias $\mathbf{D} = \left[d_1, d_2, \ldots, d_q\right] \in \mathbf{R}^{N\times L}$ and the weight $\mathbf{E} = \left[\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_q\right] \in \mathbf{R}^{n\times L}$ that connects the hidden layer and the input layer, where L represents the number of hidden nodes with activation function $h(\boldsymbol{x})$, $\boldsymbol{E}_q = \left[E_{q1}, E_{q2}, \ldots, E_{qn}\right]^{\mathrm{T}}$ is the vector of link weights that connects n input nodes with the qth hidden node.

By the following equation, the hidden layer output matrix $\mathbf{M}$ is calculated.

$$\mathbf{M} = h\left(\mathbf{IE} + \mathbf{D}\right) \tag{1}$$

The following equation can express the output matrix $\mathbf{O}$.

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

where $\mathbf{F} = [\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_L]^{\boldsymbol{T}}$ represents a matrix of output weight which links the output layer to the hidden layer.

Output weight matrix $\mathbf{F}$ can be computed using the least-squares method by Eq.(3).

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

where $\mathbf{M}^{+}$ represents the Moore-Penrose (MP) generalized inverse of the matrix $\mathbf{M}$. If $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ is nonsingular, then $\mathbf{M}^{+} = \left(\mathbf{M}^{\mathrm{T}}\mathbf{M}\right)^{-1}\mathbf{M}^{\mathrm{T}}$, in this case, L is no more than N. If $\mathbf{M}\mathbf{M}^{\mathrm{T}}$ is nonsingular, then $\mathbf{M}^{+} = \mathbf{M}^{\mathrm{T}}\left(\mathbf{M}\mathbf{M}^{\mathrm{T}}\right)^{-1}$, in this case, L is greater than N.

The ELM is described in Algorithm 1 based on the discussion mentioned above.

Algorithm 1 ELM

Input: $h(x)$: activation function, $\mathbf{I}$: input matrix of N samples, $\mathbf{O}$: output matrix of N samples, L: the number of nodes in the hidden layer

Output: $f(\mathbf{x}) = h(\mathbf{IE} + \mathbf{D})\mathbf{F}$

1 Randomly initialize weight $\mathbf{E}$ and bias $\mathbf{D}$

2 Calculate matrix $\mathbf{M} = h(\mathbf{IE} + \mathbf{D})$

3 Calculate matrix $\mathbf{F}$ by Eq.(3)

B. TWO-HIDDEN-LAYER EXTREME LEARNING MACHINE

In 2016, Qu et al. [31] proposed the TELM algorithm. Unlike SLFN, TELM has $2L$ hidden neurons (both hidden layers each have L neurons). A dataset of N samples is given $\left(\boldsymbol{i}_p, \boldsymbol{o}_p\right)$ $(p = 1, 2, \ldots, N)$, firstly, the bias $\mathbf{D}_1$ of the first hidden layer and the weight $\mathbf{E}$ which connects the input layer and the first hidden layer are randomly initialized. The augmented matrices $\mathbf{E}_{\mathrm{A}} = \begin{bmatrix}\mathbf{D}_1 \\ \mathbf{E}\end{bmatrix}$ and $\mathbf{I}_{\mathrm{A}} = [\mathbf{1}\ \ \mathbf{I}]$ are defined, where 1 is a one-column vector of size N with scalar unit elements 1. Secondly, the output matrix $\mathbf{M}_1$ is given by Eq.(4) in the first hidden layer.

$$\mathbf{M}_1 = h\left(\mathbf{I}_{\mathrm{A}}\mathbf{E}_{\mathrm{A}}\right) \tag{4}$$

Eq.(5) can be used to compute the weight matrix $\mathbf{F}$ which connects the output layer and the second hidden layer.

$$\mathbf{F} = \mathbf{M}_1^{+}\mathbf{O} \tag{5}$$

On the second hidden layer, Eq.(6) can compute the expected output matrix $\mathbf{M}_2'$.

$$\mathbf{M}_2' = \mathbf{O}\mathbf{F}^{+} \tag{6}$$

Furthermore, the augmented matrix $\mathbf{E}_{HA} = \begin{bmatrix}\mathbf{D}_2 \\ \mathbf{E}_{\mathrm{H}}\end{bmatrix}$ is defined, where $\mathbf{E}_{\mathrm{H}}$ represents the weight matrix that connects the first and the second hidden layer and $\mathbf{D}_2$ represents the bias in the second hidden layer. We define $\mathbf{M}_{1\mathrm{A}} = [\mathbf{1}\ \mathbf{M}_1]$. The expected output $\mathbf{M}_2'$ can be also expressed as Eq.(7).

$$\mathbf{M}_2' = h\left(\mathbf{M}_1\mathbf{E}_{\mathrm{H}} + \mathbf{D}_2\right) \tag{7}$$

According to Eq.(7), the augmented matrix $\mathbf{E}_{\mathrm{HA}}$ can be calculated as

$$\mathbf{E}_{\mathrm{HA}} = \mathbf{M}_{1\mathrm{A}}^{+}h^{-1}\left(\mathbf{M}_2'\right) \tag{8}$$

where $h^{-1}(\boldsymbol{x})$ is the inverse of the activation function $h\left(\boldsymbol{x}\right)$, $\mathbf{M}_{1\mathrm{A}}^{+}$ represents the MP generalized inverse of $\mathbf{M}_{1\mathrm{A}}$. According to Eq.(9), we can calculate the second hidden layer's actual output $\mathbf{M}_2$.

$$\mathbf{M}_2 = h\left(\mathbf{M}_{1\mathrm{A}}\mathbf{E}_{\mathrm{HA}}\right) \tag{9}$$

The weight matrix which connects the output layer and the second hidden layer $\mathbf{F}_{\mathrm{new}}$ is given by Eq.(10).

$$\mathbf{F}_{\mathrm{new}} = \mathbf{M}_2^{+}\mathbf{O} \tag{10}$$

where $\mathbf{M}_2^{+}$ represents the MP generalized inverse of $\mathbf{M}_2$. After training, the output of TELM can be expressed as $f\left(\boldsymbol{x}\right) = \mathbf{M}_2\mathbf{F}_{\mathrm{new}}$.

Based on the discussion mentioned above, the TELM is described in Algorithm 2.

Algorithm 2 TELM

Input: $h(x)$: activation function, $\mathbf{I}$: input matrix of N samples, $\mathbf{O}$: output matrix of N samples, L: the number of nodes for each hidden layer

Output: $f(\mathbf{x}) = h\left\{\left[h\left(\mathbf{IE}+\mathbf{D}_1\right)\mathbf{E}_{\mathrm{H}} + \mathbf{D}_2\right]\right\}\mathbf{F}_{\mathrm{new}}$

1 Randomly initialize bias $\mathbf{D}_1$ and weight matrix E and define augmented matrices $\mathbf{E}_{\mathrm{A}}$ and $\mathbf{I}_{\mathrm{A}}$

2 Calculate matrix $\mathbf{M}_1 = h(\mathbf{I}_{\mathrm{A}}\mathbf{E}_{\mathrm{A}})$ and define an augmented matrix $\mathbf{M}_{1\mathrm{A}}$

3 Calculate matrix $\mathbf{F} = \mathbf{M}_1^{+}\mathbf{O}$

4 Calculate matrix $\mathbf{M}_2' = \mathbf{O}\mathbf{F}^{+}$

5 Calculate augmented matrix $\mathbf{E}_{\mathrm{HA}} = \mathbf{M}_{1\mathrm{A}}^{+}h^{-1}(\mathbf{M}_2')$

6 Calculate matrix $\mathbf{M}_2 = h(\mathbf{M}_{1\mathrm{A}}\mathbf{E}_{\mathrm{HA}})$

7 Calculate matrix $\mathbf{F}_{\mathrm{new}} = \mathbf{M}_2^{+}\mathbf{O}$

C. IMPROVED-TWO-HIDDEN-LAYER EXTREME LEARNING MACHINE

The input weights and biases of the extreme learning machine are randomly generated within a given range and are not uniformly distributed resulting in poor model performance [39]. A new activation function (named affine transformation, AT),

based on the maximum entropy principle, achieves that the hidden layer outputs obey a uniform distribution to improve the robustness of the model. AT activation function is built based on the Sigmoid activation function, and the formula is shown below:

$$h_n(x, s, t) = h(sx + t) \tag{11}$$

The nodes of the hidden layer network obey a zero-mean Gaussian distribution, assuming that $t = 0$, s can be expressed as:

$$s = \frac{1}{\sigma / s_e} \tag{12}$$

For the Sigmoid function $s_e = 1.67$, σ is calculated as shown below [39].

$$\sigma = \frac{\text{Median (abs (}\mathbf{V}\text{))}}{0.6745} \tag{13}$$

where $\mathbf{V} = \mathbf{I}_A \mathbf{E}_A$.

For TELM, the activation function is the same for both hidden layers. Using different activation functions can improve the performance of TELM for different hidden layers [33]. Combining multiple activation functions improve the TELM nonlinear fitting capability. For classification problems, the Sigmoid activation function is often used, $h(x) = 1/\left(1 + e^{-x}\right)$. For regression problems, the Hyperbolic tangent activation function is often used, $h(x) = \left(1 - e^{-x}\right)/\left(1 + e^{-x}\right)$. Compared with the original TELM model, the activation function of the first hidden layer is AT activation function to improve the robustness of the model. The Hyperbolic tangent activation function is used for the second hidden layer to ensure the computational stability of the model, as shown in Figure 1.

D. MANTA RAY FORAGING OPTIMIZATION (MRFO) ALGORITHM

The input weights and bias of TELM are randomly generated and limited to the range [−0.9, 0.9] [31]. Although the AT activation function can change the data distribution to improve the model performance, the randomly generated input weights and biases still lead to unstable model performance. Therefore, the weights and biases of the AT-TELM can be optimized by the MRFO algorithm to improve the model performance. The optimization algorithm can be used to obtain the optimal input weights and biases of AT-TELM under minimum error through a training iterative process. This study discusses a regression problem, so five evaluation indicators including the coefficient of determination (R^2), the mean absolute error (MAE), the root mean square error (RMSE), the mean absolute percentage error (MAPE), and the variance account for (VAF) are selected to evaluate the performance of the model. For computational convenience, the MAE is used as a fitness function to optimize the input weights and biases. The formulas of R^2, MAE, and RMSE are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}_i\right)^2} \tag{14}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left|\left(y_i - \hat{y}_i\right)\right| \tag{15}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2} \tag{16}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{17}$$

$$\text{VAF} = \left[1 - \frac{\text{var}\left(y_i - \hat{y}_i\right)}{y_i}\right] \times 100\% \tag{18}$$

where n is the number of samples, y_i is the observed value, $\hat{y}_i$ is the predicted value, $\bar{y}_i$ is the average of the observed value.

MRFO algorithm, proposed by Zhao in 2020, considers three processes, chain foraging, cyclone foraging, and somersault foraging, to overcome the problem of local optimum. MRFO algorithm is characterized by fewer parameters, high search capability, and fast convergence. The formula for the chain foraging update position is shown below (19), as shown at the bottom of the next page, where $x_i^d(t)$ denotes the position of the ith individual at t iterations of dimension d, r is a random vector with the value [0,1], α is a weight coefficient, $\alpha = 2 \cdot r \cdot \sqrt{|\log(r)|}$, $x_{\text{best}}^d(t)$ is the optimal position on the d-dimensional space at t iterations. The location update formula is shown below for cyclone foraging.

When $\frac{t}{T} >$ rand, the location update formula for cyclone foraging is (20), as shown at the bottom of the next page, where, β is a weight coefficient, $\beta = 2e^{r_1 \frac{T-t+1}{T}} \cdot \sin(2\pi r_1)$, T is the maximum number of iterations, r_1 is a random number with the value [0,1].

When $\frac{t}{T} \leq$ rand, (21), as shown at the bottom of the next page, where $x_{rand}^d(t) = \text{Lb}^d + r \cdot (\text{Ub}^d - \text{Lb}^d)$, Ub^d and Lb^d are the upper and lower limits of the search interval, respectively.

The location update formula for somersault foraging is:

$$x_i^d(t+1) = x_i^d(t) + S \cdot [r_2 \cdot x_{\text{best}}^d(t) - r_3 \cdot x_i^d(t)] \tag{22}$$

where S is a weight coefficient, r_2 and r_3 are random numbers of [0,1].

E. MRFO-AT-TELM

A New hybrid prediction model MRFO-AF-TELM is proposed. The flow chart of this study is shown in Figure 2. Different researchers have used different machine learning models to predict TBM thrust. To analyze the difference of thrust prediction between different models, classical models are chosen for comparisons such as the least absolute shrinkage and selection operator (Lasso), decision tree (DT) and support vector machine (SVM). Although ELM and the above models have different structures, it is helpful to inspire other researchers to understand the performance of different models. Feature contribution, different time series lengths of

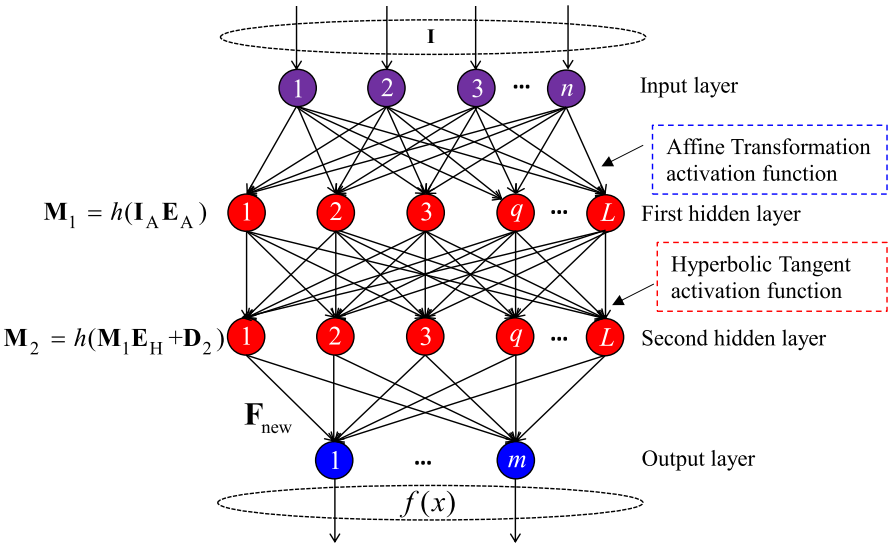


FIGURE 1. Structure of the proposed AT-TELM model.

the rising period, data distribution of weights before and after optimization, number of individuals of the MRFO algorithm, and different geological conditions on the MRFO-AT-TELM performance are discussed.

III. CASE DATASET AND FEATURE SELECTION

A. ENGINEERING CASE

As is shown in Figure 3, the Yin-Song water diversion project is now one of the greatest large-scale cross-regional water diversion projects. This study focuses on the TBM-3 construction part of the project, which starts with the Yinma river and ends with the Chalu river. The TBM-3 construction part lasted for 928 days, with the TBM method accounting for 728 of those days. The total length of TBM construction is 17.5 km. For geological conditions, limestones, granites, tuffaceous sandstones, diorites, and carbonaceous slates are the main lithology of the tunnel. The rock mass class is determined by China's hydropower (HC) system [40], including five classes from I to V. In this project, the rock mass class is mainly Class II, III, IV, and V.

B. TBM CONSTRUCTION DATA

The Yin-Song project collects 198 parameters and corresponding geological data at a collection frequency of 1Hz and collects for one single text up to 86,400 pieces of information per day and relevant parameters in the literature [14]. The maximum length of a single ring is 1.8 m. A total of 7,635 working rings are collected.

A single ring consists primarily of an empty period, arising period and a stable period [41]. The empty push period is where the TBM overcomes friction so that the cutter on the TBM cutter touches the tunnel face. The rising period is an adaptive adjustment of TBM under diverse geological circumstances to set optimal parameters [14]. The preprocessing process for TBM construction data is described in the literature [42]. As is shown in Figure 4, the starting point of the rising period is 94s on average, and the starting point of the stable period is 255s on average, with a difference of 161s between the starting points of the two periods. As can be seen, the operator takes much time to set the optimal parameters. Therefore, this study attempts to predict the thrust of the stable period based on the first 30-second, 60-second, and

$$x_i^d(t+1) = \begin{cases} x_i^d(t) + r \cdot [x_{\text{best}}^d(t) - x_i^d(t)] + \alpha \cdot [x_{\text{best}}^d(t) - x_i^d(t)] & i = 1 \\ x_i^d(t) + r \cdot [x_{\text{best}}^d(t) - x_i^d(t)] + \alpha \cdot [x_{\text{best}}^d(t) - x_i^d(t)] & i = 2, \ldots, N \end{cases} \tag{19}$$

$$x_i^d(t+1) = \begin{cases} x_{best}^d(t) + r \cdot [x_{\text{best}}^d(t) - x_i^d(t)] + \beta \cdot [x_{\text{best}}^d(t) - x_i^d(t)] & i = 1 \\ x_{best}^d(t) + r \cdot [x_{i-1}^d(t) - x_i^d(t)] + \beta \cdot [x_{\text{best}}^d(t) - x_i^d(t)] & i = 2, \ldots, N \end{cases} \tag{20}$$

$$x_i^d(t+1) = \begin{cases} x_{rand}^d(t) + r \cdot [x_{rand}^d(t) - x_i^d(t)] + \beta \cdot [x_{rand}^d(t) - x_i^d(t)] & i = 1 \\ x_{rand}^d(t) + r \cdot [x_{i-1}^d(t) - x_i^d(t)] + \beta \cdot [x_{rand}^d(t) - x_i^d(t)] & i = 2, \ldots, N \end{cases} \tag{21}$$

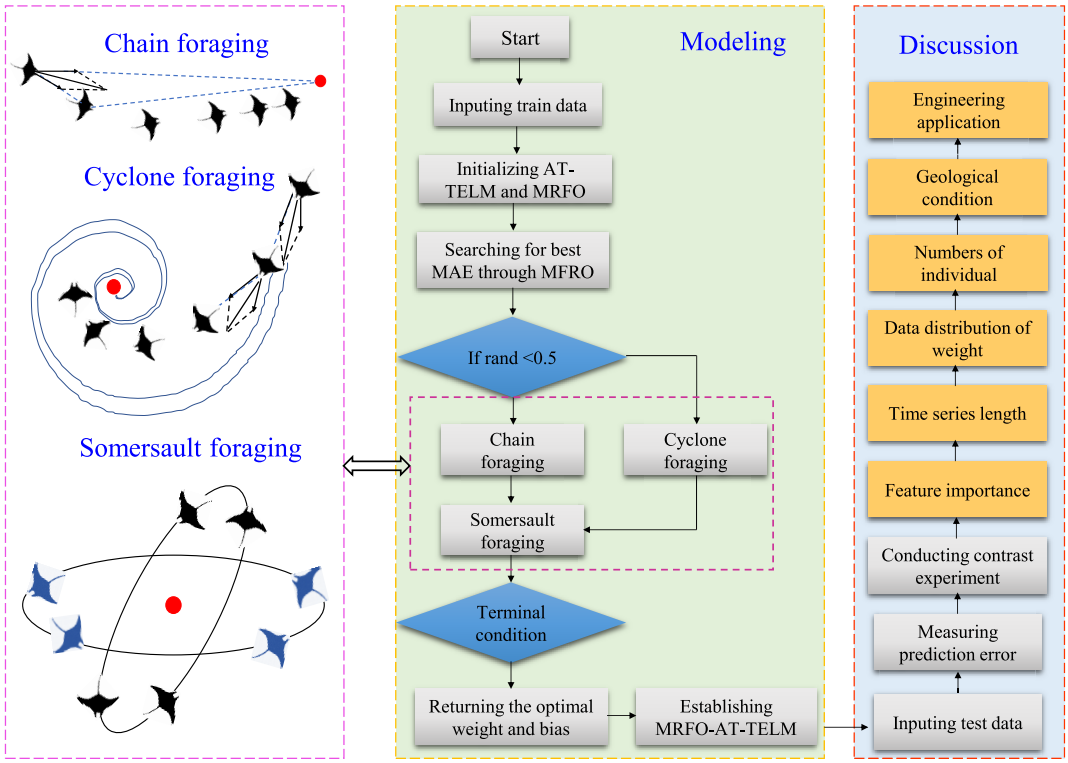


FIGURE 2. The flowchart of MRFO-AT-TELM integrating prediction and analysis.

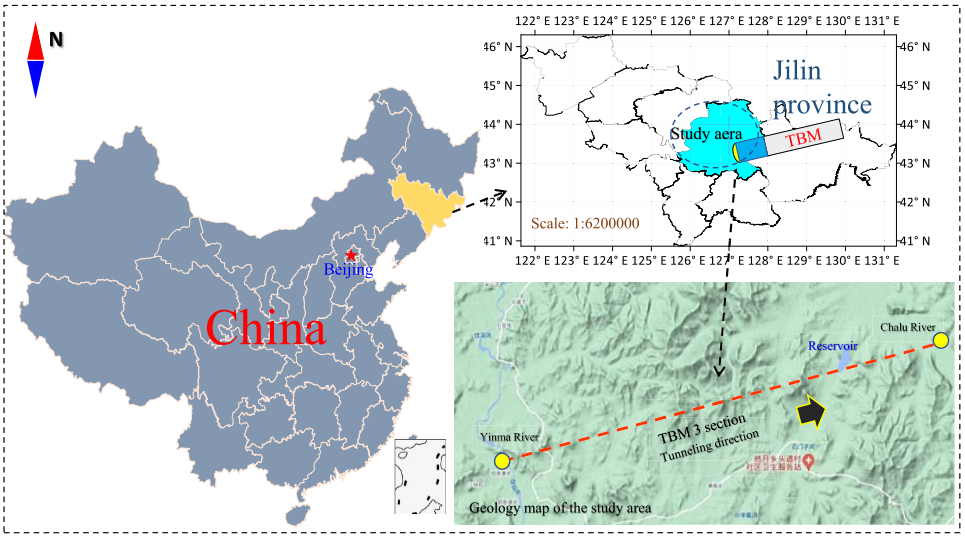


FIGURE 3. Location of the Yin-Song diversion project.

90-second data in the rising period to assist the operator in quickly adjusting for the current ring.

C. FEATURE SELECTION

The effective feature is critical for improving machine learning model prediction accuracy, especially for high-dimensional data [18]. For thrust prediction, Li et al. selected ten key features as input features based on their contribution score under the random forests (RF) model [14]: machine conveyor motor current (X_1), gripper pressure (X_2), cutterhead speed setting value (X_3), propel pump motor current (X_4), cutterhead power (X_5), cutterhead rotation speed (X_6), left shield pressure (X_7), propel speed potentiometer setting value (X_8), gear seal pressure (X_9), and propel pressure (X_{10}). In addition, the TBM construction process needs to overcome the friction between the TBM shield and the surrounding rock [43], The TBM shield is set up with three sensor parameters on the top, left and right sides respectively, so the right shield pressure (X_{11}) and the top shield pressure (X_{12}) are taken into consideration. In the end, this study

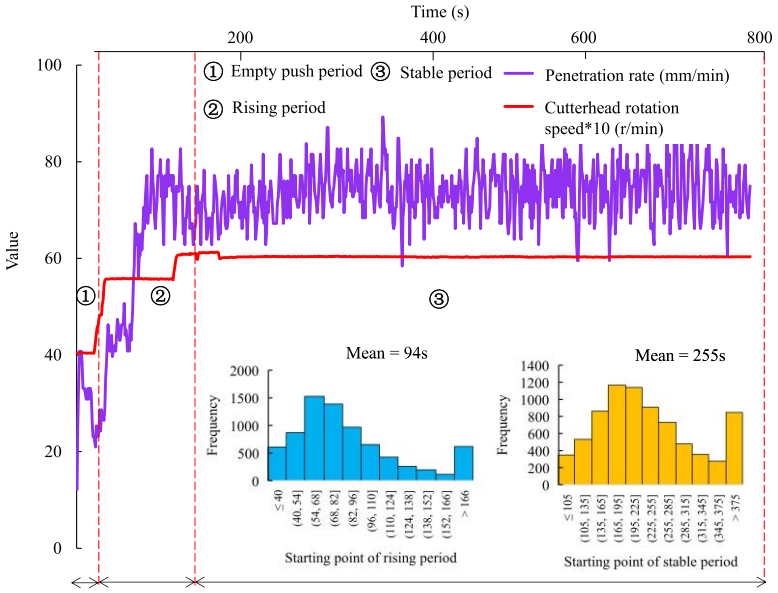


FIGURE 4. Typical characteristics of TBM ring in empty push period, rising period and stable period.

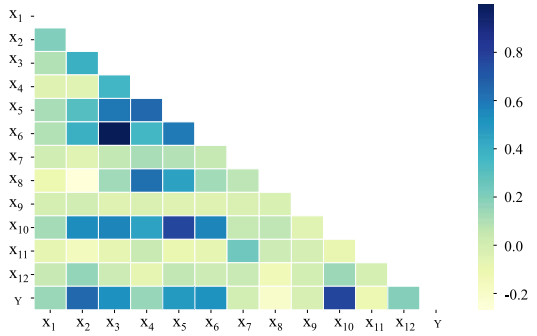


FIGURE 5. Correlation analysis of input features and.

identifies 12 input features for predicting thrust. The correlation between input features and thrust is analyzed using the Pearson correlation coefficient, as shown in Figure 5. Although there is a strong correlation between cutterhead speed setting value and cutterhead rotation speed, the two input features are useful for thrust prediction [14].

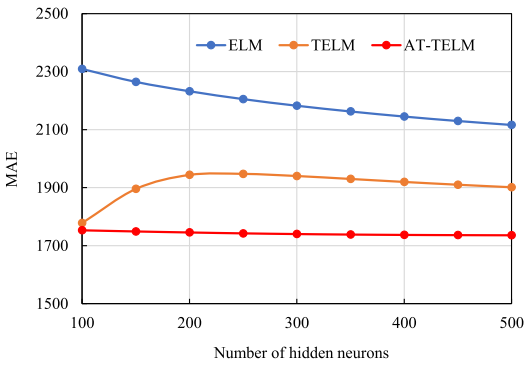
IV. ANALYSIS OF THRUST PREDICTION RESULTS

In this section, the feasibility of predicting the thrust is analyzed using different models. Before the establishment of the model, the proportion of the training set and test set is 80% and 20% respectively, and normalization of data is carried out. Generally, the training set is used to optimize the hyper-parameters of the machine learning model, and the test set is used to evaluate the performance of the model.

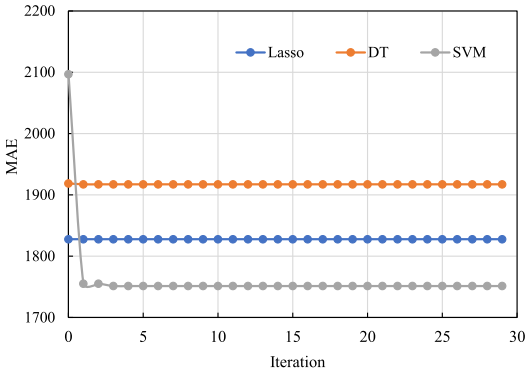
Weights and biases are the key factors affecting the performance of ELM and depend on the number of hidden neurons [44]. Therefore, ELM, TELM and AT-TELM need to optimize the number of neurons, as shown in Figure 6 (a). Weights and biases are optimized with the MRFO algorithm based on determining the optimal number of hidden neurons to further improve the performance of the AT-TELM. The effect of the number of particles of the MRFO algorithm on the predicted thrust of the AT-TELM is discussed in section 5.3.

DT, Lasso, and SVM are compared to further validate the effectiveness of the proposed method. Hyper-parameters of each method are optimized by 5-fold cross-validation [45] and MRFO algorithm, as shown in Figure 6 (b). For the Lasso, the constant that multiplies the L_1 term is optimized [42]. For the DT, the maximum depth of the tree, the minimum number of samples required to be at a leaf node, and the minimum number of samples required to split an internal node are optimized [18]. SVM selects poly kernel functions and optimizes the penalty coefficient and kernel coefficient [46]. For the implementation of Lasso, DT and SVM, the scikit-learn open-source framework is used. For the convenience of calculation, MAE is selected as the fitness function to optimize the hyper-parameters, and the maximum number of iterations is 30. All experiments are performed on a computer equipped with an Intel Core i9-9900k 3.60 GHz, 64 GB memory, and NVIDIA GeForce GTX 2080ti graphics card. Next, the performance of different models is compared for predicting the thrust using the first 30-second data.

The results of the predicted thrust based on the first 30-second data are shown in Table 1. For ease of presentation, only three evaluation indicators have been selected in the figure. MRFO-AT-TELM shows the best performance, with R^2 of 0.6398, MAE of 1683.6548, and RMSE of 2198.8981, as shown in Figure. 7. Followed by AT-TELM, its R^2 is 0.6356, MAE is 1702.6591, RMSE is 2211.6681, MAPE is 17.25%, and VAF is 63.63%. The performance of Lasso, DT, SVM, and ELM is much lower than that of TELM and AT-ELM.



(a) The optimization curves of Lasso, DT, and SVM

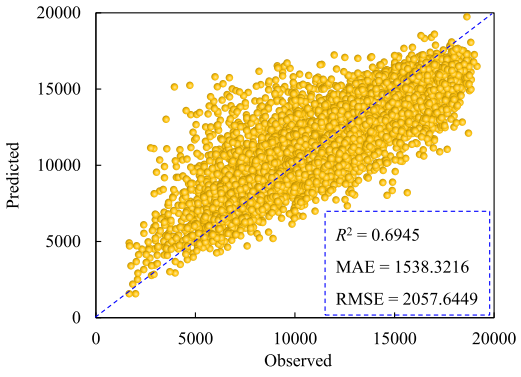


(b) The optimization iteration curves of ELM, TELM and AT-TELM

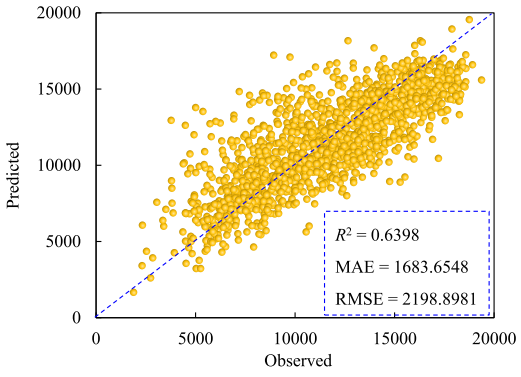
FIGURE 6. The optimization iteration curves of different models using the first 30-second data.

TABLE 1. Performance evaluation of the models using the first 30-second data.

Model	Datasets	R^2	MAE	RMSE	MAPE (%)	VAF (%)
Lasso	Training set	0.5944	1823.5353	2371.0453	21.62	59.44
	Test set	0.5985	1806.4661	2321.6967	18.81	59.88
DT	Training set	0.6169	1780.2944	2304.0928	31.28	61.69
	Test set	0.5586	1919.5329	2434.4392	19.84	55.91
SVM	Training set	0.6215	1704.1377	2290.3233	36.73	62.26
	Test set	0.6273	1721.7277	2236.7759	18.00	62.75
ELM	Training set	0.6626	1644.1836	2162.5269	29.05	66.26
	Test set	0.5892	1809.0019	2348.4225	18.37	59.08
TELM	Training set	0.6185	1749.2143	2299.3905	35.43	61.86
	Test set	0.6307	1739.0358	2226.5234	18.02	63.08
AT-TELM	Training set	0.6853	1750.6591	2088.2727	28.14	68.54
	Test set	0.6356	1702.6591	2211.6681	17.25	63.63
MRFO-AT-TELM	Training set	0.6945	1538.3216	2057.6449	24.09	69.46
	Test set	**0.6398**	**1683.6548**	**2198.8981**	**17.16**	**64.02**



(a) Training set



(b) Test set

FIGURE 7. Scatter plots comparing the predicted and observed thrust based on the first 30-second data using MRFO-AT-TELM.

The result of the predicted thrust is analyzed based on the first 60-second and 90-second data to compare the effects of different time series lengths. The result of the predicted thrust based on the first 60-second data in the rising period is shown in Table 2. It can be found that the performance of the AT-TELM is more optimal than that of Lasso, DT, SVM, ELM, and TELM. The proposed MRFO-AT-TELM has the optimal performance with R^2 of 0.7235, MAE of 1442.4626, and RMSE of 1926.8964, as shown in Figure 8. It is further demonstrated that the MRFO algorithm and AT can significantly improve the prediction performance of the TELM.

The result of the predicted thrust based on the first 90-second data in the rising period is shown in Table 3. MRFO-AT-TELM shows the best performance, with R^2 of 0.7494, MAE of 1323.0945, and RMSE of 1834.1150,

TABLE 2. Performance evaluation of the models using the first 60-second data.

Model	Datasets	R^2	MAE	RMSE	MAPE (%)	VAF (%)
Lasso	Training set	0.6810	1544.7272	2102.6336	17.10	68.10
	Test set	0.6914	1527.6354	2035.5227	16.10	69.16
DT	Training set	0.6887	1542.9864	2077.1017	28.30	68.87
	Test set	0.6555	1637.5170	2150.6215	17.10	65.56
SVM	Training set	0.6892	1470.9559	2075.5315	29.63	69.22
	Test set	0.7063	1462.0873	1985.8015	15.43	70.71
ELM	Training set	0.7105	1491.0184	2003.0348	20.21	71.06
	Test set	0.6572	1608.3412	2145.3828	16.45	65.83
TELM	Training set	0.6943	1502.3187	2058.1831	28.56	69.44
	Test set	0.7073	1486.0903	1982.3150	15.57	70.73
AT-TELM	Training set	0.7400	1377.7801	1898.3551	24.84	74.00
	Test set	0.7167	1470.4945	1950.2287	15.06	71.70
MRFO-AT-TELM	Training set	0.7459	1348.9736	1876.6792	25.41	74.59
	Test set	**0.7235**	**1442.4626**	**1926.8964**	**14.77**	**72.37**

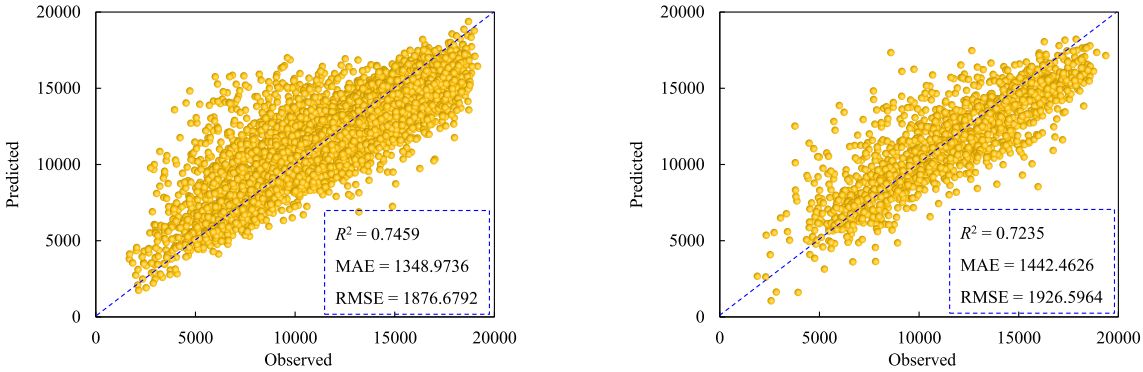


FIGURE 8. Scatter plots comparing the predicted and observed thrust based on the first 60-second data using MRFO-AT-TELM.

TABLE 3. Performance evaluation of the models using the first 90-second data.

Model	Datasets	R^2	MAE	RMSE	MAPE (%)	VAF (%)
Lasso	Training set	0.7218	1397.3607	1963.5352	15.72	72.18
	Test set	0.7233	1376.7221	1927.4670	14.57	72.35
DT	Training set	0.7255	1406.1258	1950.2917	26.35	72.55
	Test set	0.6997	1501.7840	2007.9512	15.60	70.00
SVM	Training set	0.7258	1324.4334	1949.1969	26.07	72.96
	Test set	0.7429	1326.3832	1857.7631	14.21	74.51
ELM	Training set	0.7391	1375.4630	1901.4138	17.61	73.92
	Test set	0.6868	1513.8461	2050.5055	15.66	68.75
TELM	Training set	0.7392	1348.8590	1901.2346	21.09	3.92
	Test set	0.7448	1343.1894	1850.7623	14.00	74.49
AT-TELM	Training set	0.7645	1261.7799	1806.4467	14.20	76.46
	Test set	0.7402	1337.4646	1867.5715	13.73	74.05
MRFO-AT-TELM	Training set	0.7663	1254.1431	1799.6175	16.53	76.63
	Test set	**0.7494**	**1323.0945**	**1834.1150**	**13.59**	**74.96**

as shown in Figure 9, followed by AT-TELM. A comparison of the predicted and observed values of MRFO-AT-TELM is shown in Figure 10. The test results of SVM and TELM are slightly lower than that of AT-TELM. The performance of ELM is also not as good as that of TELM and AT-TELM. In the first 60-second and 90-second data sets, the performance of Lasso is better than the DT model, and the DT model shows the worst performance.

Compared with TELM and AT-TELM, the performance of ELM still has greater limitations. Besides, the performance

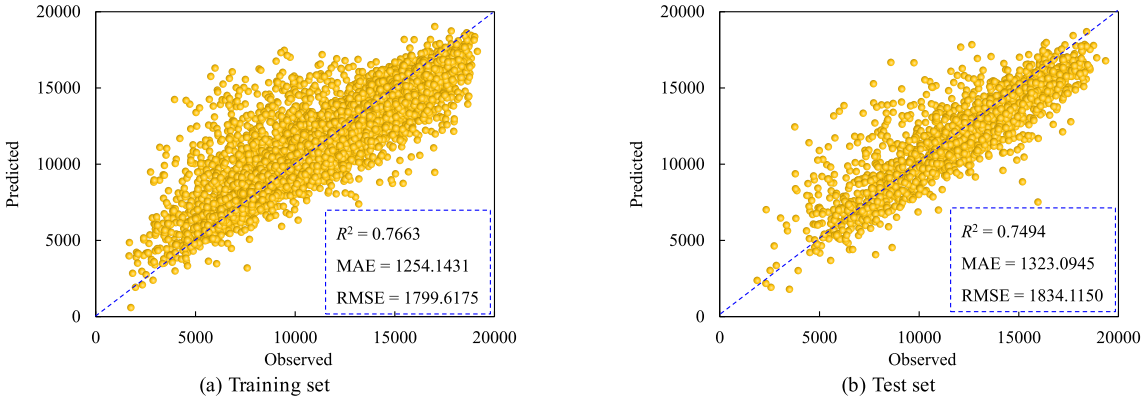


FIGURE 9. Scatter plots comparing the predicted and observed thrust based on the first 90-second data using MRFO-AT-TELM.

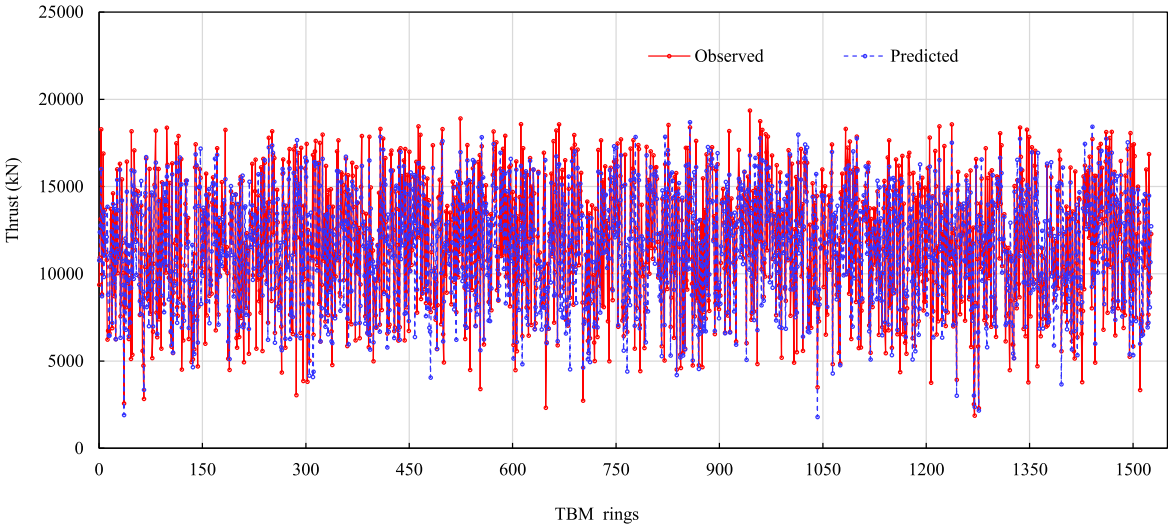


FIGURE 10. Comparison of measured and predicted thrust based on the first 90-second data using MRFO-AT-TELM.

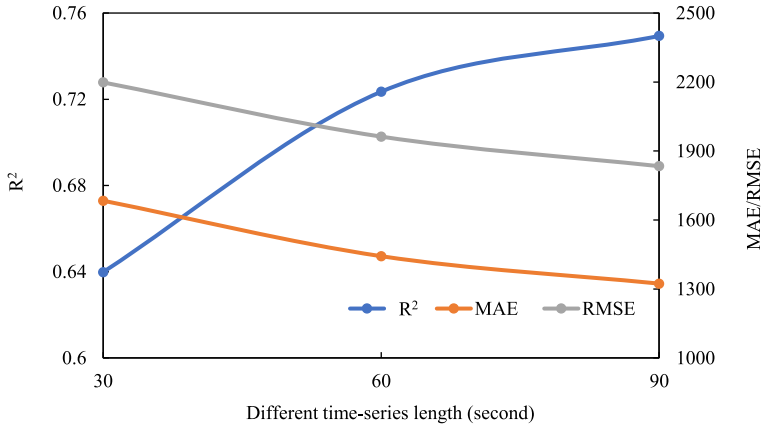
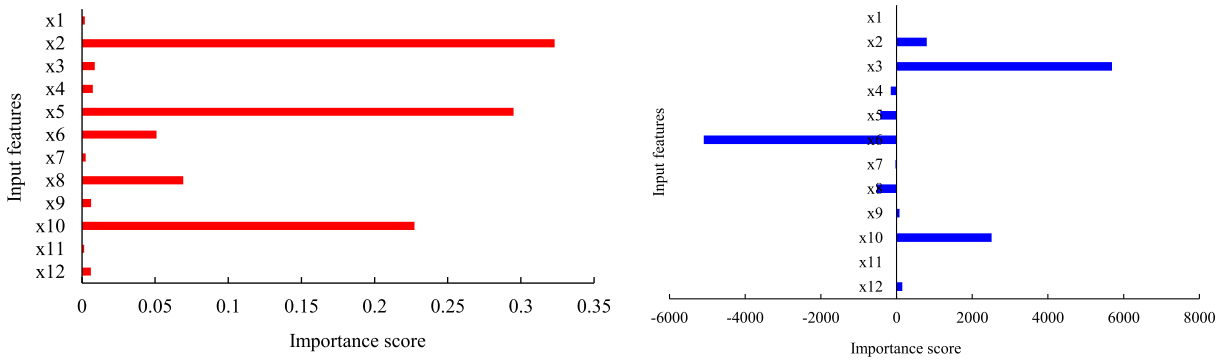


FIGURE 11. Performance evaluation of different time-series length.

of DT is the worst, followed by Lasso. It is optimal to predict the thrust based on the first 90-second data in the rising period. In terms of predicting the thrust based on the first 30-second data in the rising period, its accuracy is still insufficient. The reason is that the TBM construction data is not sufficient to predict thrust more accurately in the initial TBM rock-breaking process. In addition, the performance of all models is better as the time series length of the rising period increases, but the evaluation indicators' growth rate is reduced, as shown in Figure 11. Compared to the first



(a) Feature importance of DT

(b) Feature importance of Lasso

FIGURE 12. Importance score of input features on the thrust prediction.

30-second data, the MRFO-AT-TELM improved R^2 by 13% and reduced MAE by 14% using the first 60-second data. Compared to the first 60-second data, the MRFO-AT-TELM improved R^2 by 3% and reduced MAE by 8% based on the first 90-second data. Overall, the first 90-second data in the rising period can assist the operator in quickly evaluating thrust in advance.

V. DISCUSSION

The discussion section contains the following parts: the importance of the input features is analyzed; the data distribution of AT-TELM's weights after optimization is analyzed; the effect of different MRFO individual numbers on the optimized AT-TELM performance is analyzed; the performance of MRFO-AT-TELM under different lithology and rock mass classes is analyzed; Finally, the proposed model is migrated to a new project to validate the feasibility of the proposed model.

A. INPUT FEATURE IMPORTANCE ANALYSIS

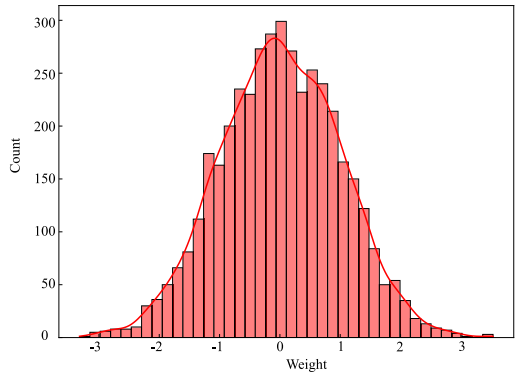
Analyzing the input features is critical for predicting thrust, which helps to explain the key input features. As mentioned in the 3.3 section, Li et al. considered that propel pressure (X_{10}), gear seal pressure (X_9), propel speed potentiometer setting value (X_8), left shield pressure (X_7), cutterhead rotation speed (X_6) are the top 5 most important input features based on RF [47]. More details are introduced in [14]. The prediction result of RF is based on the output of a set of DT jointly, and the training time is much longer than that of DT. Therefore, DT is utilized to analyze the importance score of input features [48].

For DT, gripper pressure (X_2), cutterhead power (X_5), propel pressure (X_{10}), propel speed potentiometer setting value (X_8) and cutterhead rotation (X_6) have the largest contribution scores, as shown in Figure 12 (a). The results of RF and DT are similar despite the difference in datasets and features selected. Both methods yielded propel pressure (X_{10}), propel speed potentiometer setting value (X_8) and cutterhead rotation speed (X_6) as important input features. The correlation analysis found a strong linear relationship between the input features and the thrust. Therefore the Lasso linear model is considered to analyze the importance score of the input features. Lasso incorporates L_1 regularization, which changes the coefficients of irrelevant input features to zero [49], and is widely used for feature analysis. As shown in Figure 12 (b), cutterhead rotation speed (X_6), cutterhead speed setting value (X_3), propel pressure (X_{10}), gripper pressure (X_2) and propel speed potentiometer setting value (X_8) are the top 5 input features with the highest importance score.

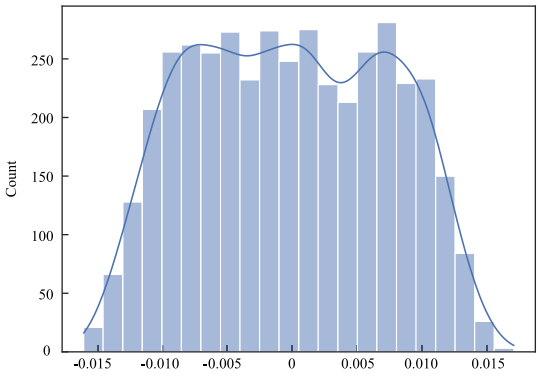
The feature importance scores obtained by different machine learning methods are different; propel pressure (X_{10}) and propel speed potentiometer setting value (X_8) are key input features; the importance score of machine conveyor motor current (X_1), propel pump motor current (X_4), left shield pressure (X_7), right shield pressure (X_{11}) is small. Compared with the left/right shield pressure, the top shield pressure has an important influence on the thrust prediction, which is caused by the larger contact area between the top shield of TBM and the surrounding rock during TBM construction.

B. DATA DISTRIBUTION OF WEIGHTS BEFORE AND AFTER OPTIMIZATION

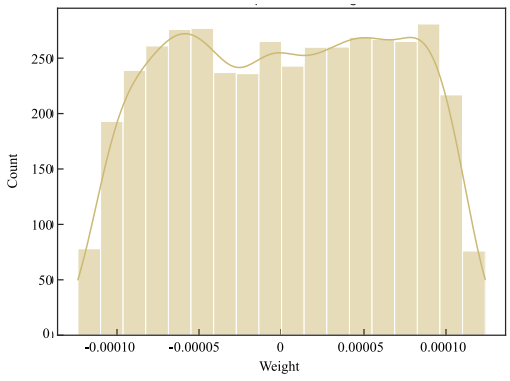
To compare the effect of AT on TELM, TELM and AT-TELM are optimized with the MRFO algorithm separately, and the data distribution of the optimized weights is analyzed. After determining the optimal number of neurons, the weights of TELM and AT-TELM are randomly generated. As shown in Figure 13 (a), the randomly generated weights are a normal distribution with data in the range $[-3,3]$. As shown in Figure 13 (b), the weights of optimized MRFO-TELM range from $[-0.015,0.015]$. Compared to weights of TELM, input weights of MRFO-TELM have a smaller range. Compared with MRFO-TELM, the weight data distribution of MRFO-AT-TELM is a uniform distribution, while the optimized AT-TELM has a smaller range of weights, as shown in Figure 13 (c). AT changes the weight data distribution of TELM to obey a uniform distribution, which can improve the model performance during the optimization process. [39].



(a) Randomly generated weights

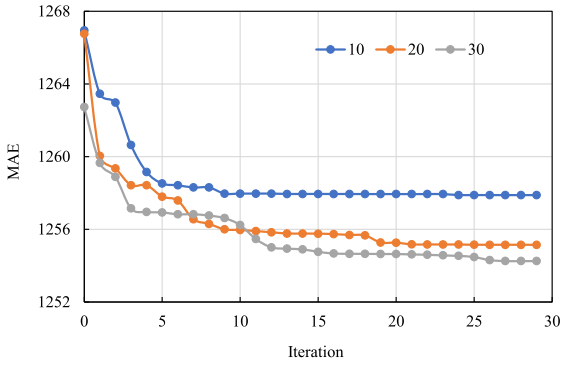


(b) Weights of optimized MRFO-TELM

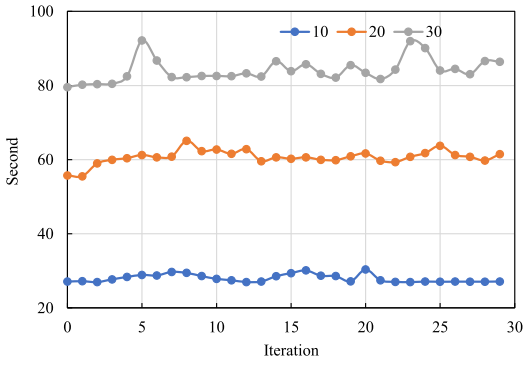


(c) Weights of optimized MRFO-AT-TELM

FIGURE 13. Data distribution of weights before and after optimization.

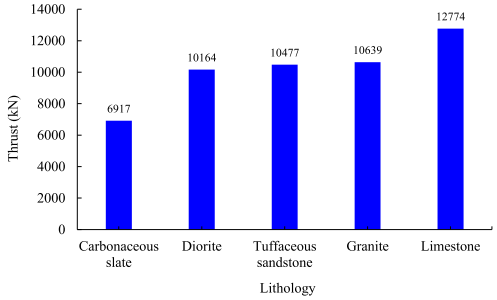


(a) Curves of MAE

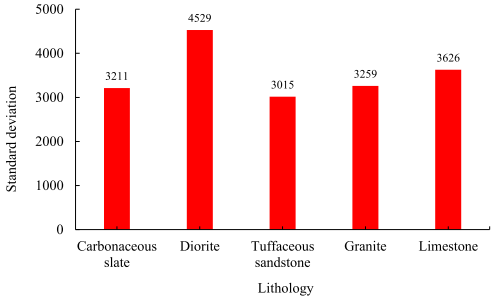


(b) Curves of time cost

FIGURE 14. Training process using MRFO-AT-TELM with different numbers of individual.

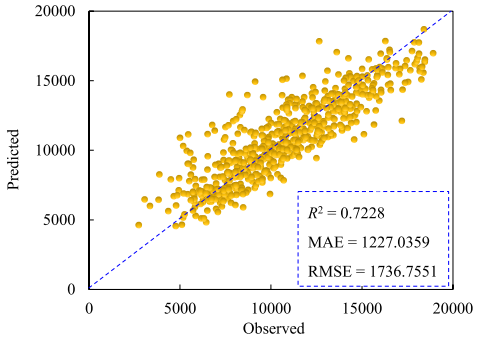


(a) Average

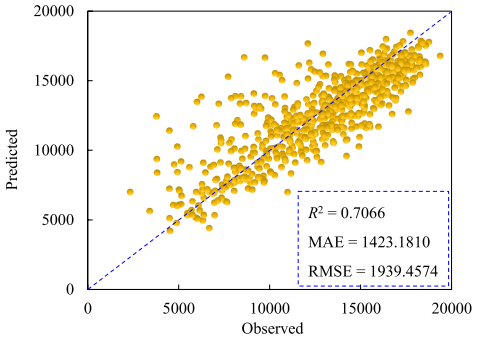


(b) Standard deviation

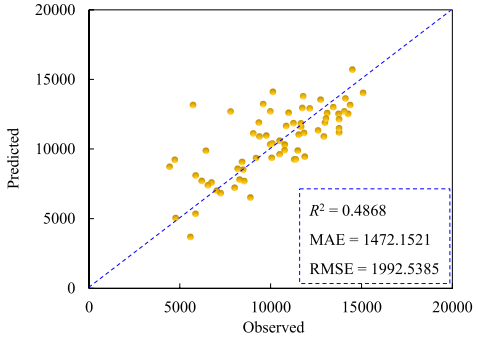
FIGURE 15. Average and standard deviation of thrust in different lithology.



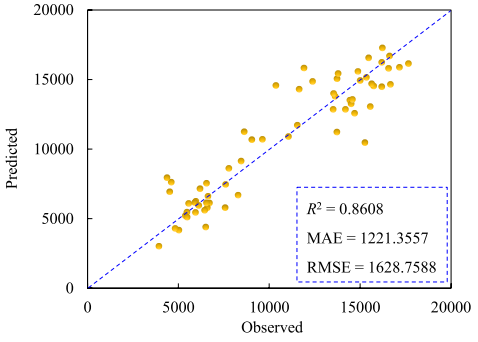
(a) Performance of MRFO-AT-TELM in granite



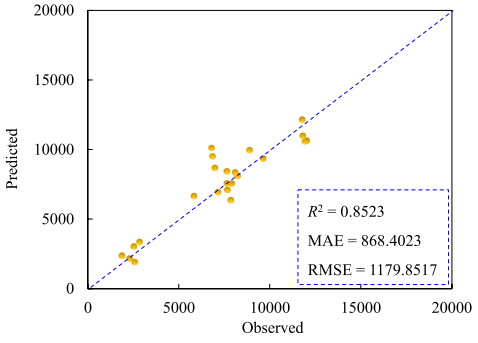
(b) Performance of MRFO-AT-TELM in limestone



(c) Performance of MRFO-AT-TELM in tuffaceous sandstone

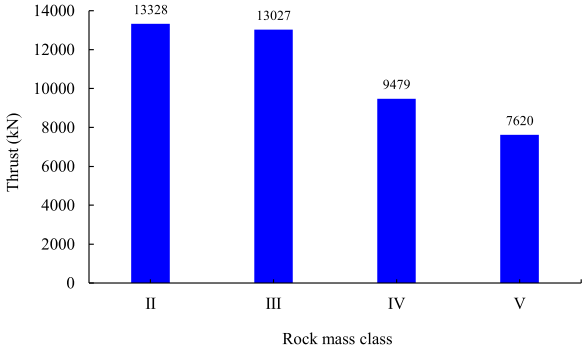


(d) Performance of MRFO-AT-TELM in diorite

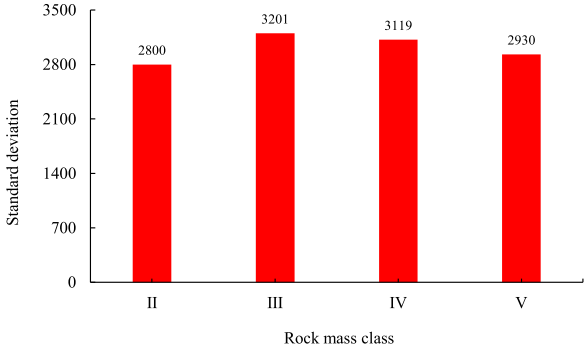


(e) Performance of MRFO-AT-TELM in carbonaceous slate

FIGURE 16. Observed and predicted values for test set in different lithology.

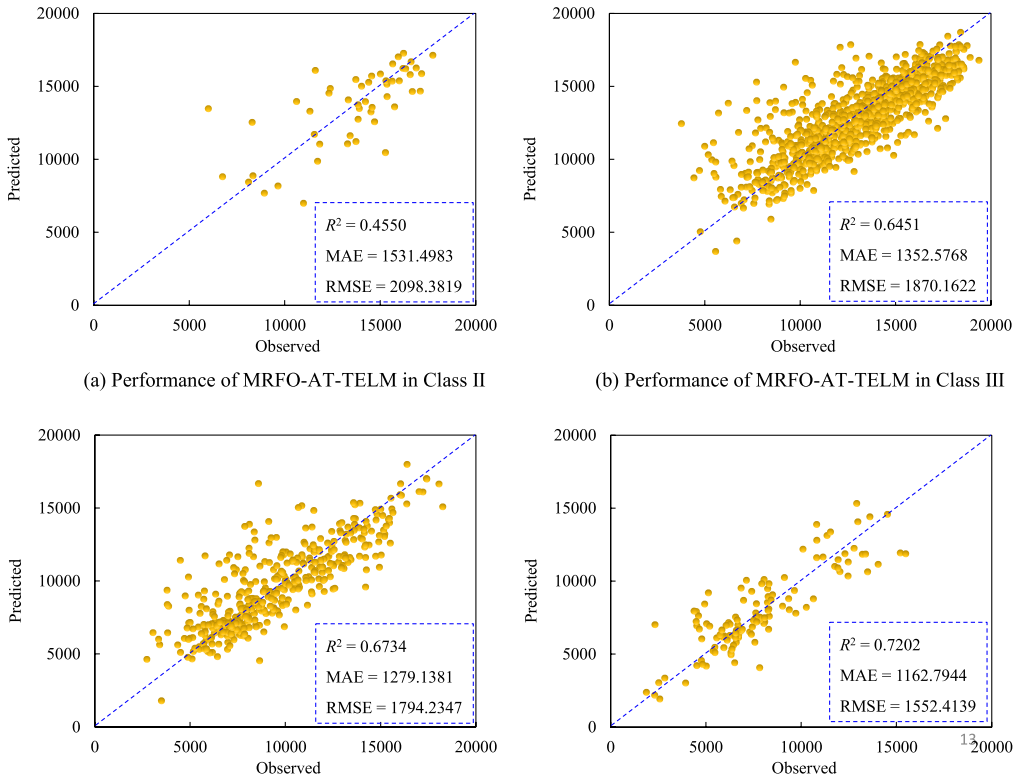


(a) Average of thrust in different rock mass class



(b) Standard deviation of thrust in different rock mass class

FIGURE 17. Average and standard deviation of thrust in different rock mass class.



(a) Performance of MRFO-AT-TELM in Class II

(b) Performance of MRFO-AT-TELM in Class III

(c) Performance of MRFO-AT-TELM in Class IV

(d) Performance of MRFO-AT-TELM in Class V

FIGURE 18. Observed and predicted values for test set in different rock mass class.

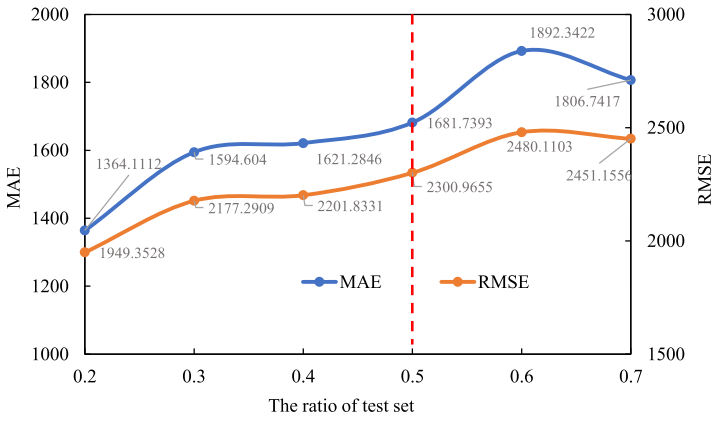


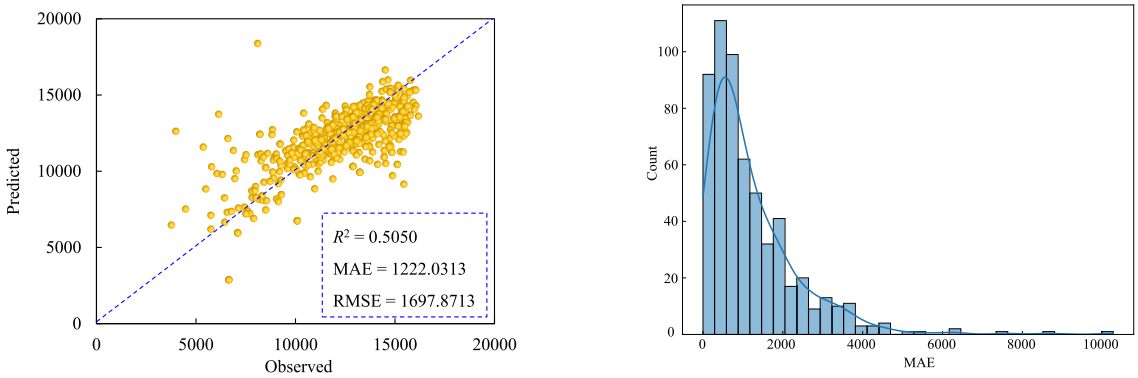
FIGURE 19. Predicting the unexcavated part of the tunnel based on the excavated data for the Yin-Song project.

C. EFFECT OF THE NUMBER OF INDIVIDUALS ON THE MRFO-AT-TELM PREDICTION PERFORMANCE

The weights and biases of AT-TELM need to be optimized by the MRFO algorithm to improve the performance. The number of individuals in the MRFO algorithm affects the performance and time cost of AT-TELM. As shown in Figure 14, when the number of individuals is set to 10, the minimum value of MAE is 1257, and the total time is 838 seconds; when the number of individuals is set to 20, the minimum value of MAE is 1255, and the total time is 1819 seconds; when the number of individuals is set to 30, the minimum value of MAE is 1254, and the total time is 2522 seconds. When the iteration period is set to 30, the training error (MAE) is small and satisfies the training requirements.

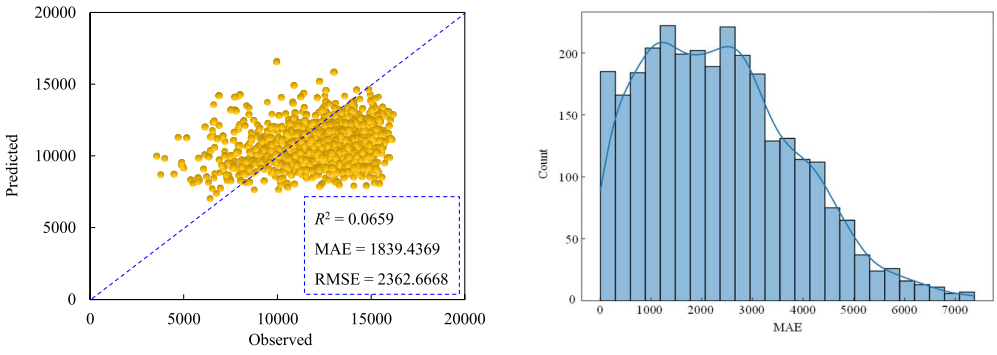
D. EFFECT OF GEOLOGICAL CONDITIONS

As mentioned before, the lithology mainly contains limestones, granites, tuffaceous sandstones, diorites, and carbonaceous slates in the Yin-Song project. Firstly, the performance of the predicted thrusts under different lithologies is analyzed using MRFO-AT-TELM. The average can reflect the overall trend of the data, and the standard deviation can reflect the



(a) Scatter plots (b) Error distribution

FIGURE 20. Predicted results of the test set in the new project.



(a) Performance of MRFO-AT-TELM in new projects (b) Error distribution of MRFO-AT-TELM in new projects

FIGURE 21. MRFO-AT-TELM predicting the thrust of the new project based on Yin-Song project.

dispersion of the data. Therefore, the average and standard deviation are selected as the measurement indicators. As is shown in Figure 15, the thrust is the largest in limestones and the smallest in carbonaceous slates with 12,774 kN and 6,917 kN, respectively. But the average does not differ much under the remaining three lithologies. The standard deviation shows that the thrust has the largest dispersion in the diorite. Next, the performance of MRFO-AT-TELM is analyzed to predict thrust under different lithological.

As is shown in Figure 16, the performance of MRFO-AT-TELM in tuffaceous sandstones is poor. Comparative analysis shows that MRFO-AT-TELM has the best performance in limestones, followed by carbonaceous slates.

The rock mass class has a significant impact on TBM construction [50], for which sensitivity analysis is necessary. As is shown in Figure 17, the thrust decreases as the class of the rock mass class increases. The average thrust reaches 13,328 kN in Class II, while the average thrust is only 7620 kN in Class V. The thrust should be reduced under Class IV and V to reduce the surrounding rock disturbance. Besides, the thrust should be increased appropriately to improve construction efficiency in better geological conditions.

As is shown in Figure 18, the performance of MRFO-AT-TELM in different rock mass classes is analyzed. The prediction result of the model in Class V is the best, with R^2 of 0.7202, MAE of 1162.7944, and RMSE of 1552.4139. For Class III and Class IV, the model still obtains better prediction results. It's worth mentioning that the performance of the model is the weakest in Class II. It has limited application in tuffaceous sandstones and Class II.

E. ENGINEERING APPLICATION

The data set is divided into a training set and a test set without being randomly partitioned to investigate whether the data from the excavated part of the project can predict the unexcavated part of the project. The test set represents the unexcavated part of the tunnel. The scale of the test set is set from 0.2 to 0.7 to test the performance of the MRFO-AT-TELM. As shown in Figure 19, as the ratio of test set increases, the prediction error increases significantly. When the ratio of the test set exceeds 0.5 (the number of TBM rings is 3816), the error grows rapidly and the MAE approaches 2000 kN. By comparing the prediction results of MRFO-AT-TELM for thrust in different lithologies and rock mass classes, the established dataset should contain complex geological conditions to meet the generalization capability of the model.

A total of 2919 TBM rings are collected from the new project to verify the feasibility of the proposed model. In the

new project, only machine conveyor motor current (X_1), gripper pressure (X_2), cutterhead rotation speed (X_6), left shield pressure (X_7), propel pressure (X_{10}), right shield pressure (X_{11}) and top shield pressure (X_{12}) are obtained. The training and test sets are randomly divided. A total of 584 TBM rings are obtained for the test set. As shown in Figure. 20 (a), the R^2 of MRFO-AT-TELM is 0.5050, the MAE is 1222.0313, and the RMSE is 1697.8713. As shown in Figure. 20 (b), the range of error is mainly from 0 to 2000, and the model has a good performance. The performance of the proposed model in predicting thrust is reduced without acquiring completely 12 input features from the new project.

The MRFO-AT-TELM thrust prediction model is established to predict the thrust of the new project based on the Yin-Song project to explore the feasibility of model transfer. As shown in Figure 21, the R^2 of the established MRFO-AT-TELM in the new project is 0.0659, the MAE is 1839.4369, and the RMSE is 2362.6668, with a large error. The rock mass class in the new project are mainly Class II. The data of Class II is less in the Yin-song project, and 283 TBM rings are not enough to train the model. Further analysis shows that the average thrust of the thrust project in category II is 13328 kN, and the standard deviation is 2800. The average thrust in the new project is 12317 kN with a standard deviation of 2321. Overall, the data distributions of the two projects are similar. It can be concluded that the lack of Class II data in the Yin-Song project is the reason for the poor performance of MRFO-AT-TELM.

Based on the above analysis, the proposed model and the selected input features are feasible. The number of TBM rings obtained for the new project is not sufficient to predict the unexcavated part of the tunnel. The number of rings should be greater than 3800 and contain data for different lithologies and rock mass classes.

F. LIMITATION

Only three time-series lengths (30s, 60s, and 90s) in the rising period are selected as input for thrust prediction. Without obtaining the most adequate number of TBM rings (over 3800 rings), thrust prediction is difficult. The applicability of the model under adverse geological conditions is not discussed. Different parameters are collected for different TBM tunnels, and models trained for developed projects are difficult to apply directly to new projects.

VI. CONCLUSION

A novel machine learning method is proposed for predicting TBM thrust, and its performance is evaluated by using the TBM construction data collected from the Yin-Song Water Diversion Project in Jilin Province of China. Based on the TBM rock-breaking construction data in the rising period, the 12 parameters are chosen as the featuring input, and the TBM thrust in the stable period is chosen as the output. Time series of different lengths of the ascent period are considered as input, which is respectively the first 30-second, 60-second, and 90-second. The main findings of this paper are as follows.

(1) The performance of the model is improved with the increasing length of the rising period, but the growth rate of evaluation indicators is significantly reduced. Compared with the first 30-second data and 60-second data, the prediction accuracy based on the first 90-second data is the highest. The suggested MRFO-AT-TELM outperforms the other machine learning methods. The prediction performance of SVM is better than that of DT, ELM, and Lasso, but significantly worse than that of TELM and AT-TELM.

(2) Among the 12 input features, propel pressure and propel speed potentiometer setting value have the greatest contribution score. Besides, top shield pressure has an important influence on thrust prediction and should be considered. Compared with MRFO-TELM, the optimized MRFO-AT-TELM has a uniform distribution of data weights and a smaller range of weight.

(3) Different geological conditions have great differences in thrust. The TBM thrust has the largest average in limestones, the smallest average in carbonaceous slates, and the largest standard deviation in diorites. The thrust decreases with the increase of the rock mass class. Besides, the adaptability of thrust prediction based on MRFO-AT-TELM is different under different geological conditions, e.g. good performance in diorite and Class V, but poor performance in tuffaceous sandstone and Class II.

(4) When predicting the unexcavated part of the tunnel based on the excavated part, the number of rings should be greater than 3800 and contain data for different lithologies and rock mass classes. Since model transfer is difficult, and in addition to satisfying the similarity of input features, as much TBM construction data as possible is collected for Class II, IV and V.

The proposed model (MRFO-AT-TELM) will be extended as the TBM construction data are continuously collected at different tunnels. In the next step, the model will be incorporated into the TBM operating system in different tunnels for applicability validation.

DECLARATION

The authors declare that there is no conflict of interest in this article.

REFERENCES

[1] K. Elbaz, S.-L. Shen, W.-J. Sun, Z.-Y. Yin, and A. Zhou, "Prediction model of shield performance during tunneling via incorporating improved particle swarm optimization into ANFIS," *IEEE Access*, vol. 8, pp. 39659–39671, 2020.

[2] S. Leng, J.-R. Lin, Z.-Z. Hu, and X. Shen, "A hybrid data mining method for tunnel engineering based on real-time monitoring data from tunnel boring machines," *IEEE Access*, vol. 8, pp. 90430–90449, 2020.

[3] M. Wei, Z. Wang, X. Wang, J. Peng, and Y. Song, "Prediction of TBM penetration rate based on Monte Carlo-BP neural network," *Neural Comput. Appl.*, vol. 33, no. 2, pp. 603–611, Jan. 2021.

[4] M. Hu, B. Li, B. Zhang, R. Wang, and L. Chen, "Improved SVR method for predicting the cutting force of a TBM cutter using linear cutting machine test data," *KSCE J. Civil Eng.*, vol. 25, no. 11, pp. 4425–4442, Nov. 2021.

[5] L. Wang, W. Sun, Y. Long, and X. Yang, "Reliability-based performance optimization of tunnel boring machine considering geological uncertainties," *IEEE Access*, vol. 6, pp. 19086–19098, 2018.

[6] M. Parise, J. De Waele, and F. Gutierrez, "Engineering and environmental problems in karst—An introduction," *Eng. Geol.*, vol. 99, no. 3, pp. 91–94, Jun. 2008.

[7] Q. Zhang, Y. Kang, Z. Zheng, and L. Wang, "Inverse analysis and modeling for tunneling thrust on shield machine," *Math. Problems Eng.*, vol. 2013, Dec. 2013, Art. no. 975703.

[8] Y. Zhao, Q. Gong, Z. Tian, S. Zhou, and H. Jiang, "Torque fluctuation analysis and penetration prediction of EPB TBM in rock–soil interface mixed ground," *Tunnelling Underground Space Technol.*, vol. 91, Sep. 2019, Art. no. 103002.

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

[10] C. Balci, "Correlation of rock cutting tests with field performance of a TBM in a highly fractured rock formation: A case study in Kozyatagi-Kadikoy metro tunnel, Turkey," *Tunnelling Underground Space Technol.*, vol. 24, no. 4, pp. 423–435, Jul. 2009.

[11] Q. Zhang, C. Qu, Z. Cai, Y. Kang, and T. Huang, "Modeling of the thrust and torque acting on shield machines during tunneling," *Automat. Construct.*, vol. 40, pp. 60–67, Apr. 2014.

[12] 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.

[13] Q. Geng, Z. Wei, H. Meng, and F. J. Macias, "Mechanical performance of TBM cutterhead in mixed rock ground conditions," *Tunnelling Underground Space Technol.*, vol. 57, pp. 76–84, Aug. 2016.

[14] 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.

[15] X. Gao, M. Shi, X. Song, C. Zhang, and H. Zhang, "Recurrent neural networks for real-time prediction of TBM operating parameters," *Automat. Construct.*, vol. 98, pp. 225–235, Feb. 2019.

[16] 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.

[17] P. Cachim and A. Bezuijen, "Modelling the torque with artificial neural networks on a tunnel boring machine," *KSCE J. Civil Eng.*, vol. 23, no. 10, pp. 4529–4537, Oct. 2019.

[18] 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," *Automat. Construct.*, vol. 125, May 2021, Art. no. 103647.

[19] L.-J. Jing, J.-B. Li, C. Yang, S. Chen, N. Zhang, and X.-X. Peng, "A case study of TBM performance prediction using field tunnelling tests in limestone strata," *Tunnelling Underground Space Technol.*, vol. 83, pp. 364–372, Jan. 2019.

[20] E. Ghasemi, S. Yagiz, and M. Ataei, "Predicting penetration rate of hard rock tunnel boring machine using fuzzy logic," *Bull. Eng. Geol. Environ.*, vol. 73, no. 1, pp. 23–35, 2014.

[21] Q. Wang, X. Xie, and I. Shahrour, "Deep learning model for shield tunneling advance rate prediction in mixed ground condition considering past operations," *IEEE Access*, vol. 8, pp. 215310–215326, 2020.

[22] J. Zhao, M. Shi, G. Hu, X. Song, C. Zhang, D. Tao, and W. Wu, "A data-driven framework for tunnel geological-type prediction based on TBM operating data," *IEEE Access*, vol. 7, pp. 66703–66713, 2019.

[23] Z. Liu, Y. Wang, L. Li, X. Fang, and J. Wang, "Realtime prediction of hard rock TBM advance rate using temporal convolutional network (TCN) with tunnel construction big data," *Frontiers Struct. Civil Eng.*, vol. 16, no. 4, pp. 401–413, Apr. 2022.

[24] O. Acaroglu, "Prediction of thrust and torque requirements of TBMs with fuzzy logic models," *Tunnelling Underground Space Technol.*, vol. 26, no. 2, pp. 267–275, Mar. 2011.

[25] R. Hasanpour, J. Rostami, J. Schmitt, Y. Ozcelik, and B. Sohrabian, "Prediction of TBM jamming risk in squeezing grounds using Bayesian and artificial neural networks," *J. Rock Mech. Geotech. Eng.*, vol. 12, no. 1, pp. 21–31, Feb. 2020.

[26] 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.

[27] M. Khashei, M. Bijari, and G. A. Raissi Ardali, "Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs)," *Neurocomputing*, vol. 72, nos. 4–6, pp. 956–967, Jan. 2009.

[28] 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, 2006.

[29] Z. Liu, J. Shao, W. Xu, and Q. Wu, "Indirect estimation of unconfined compressive strength of carbonate rocks using extreme learning machine," *Acta Geotechnica*, vol. 10, no. 5, pp. 651–663, Oct. 2015.

[30] J. Zeng, B. Roy, D. Kumar, A. S. Mohammed, D. J. Armaghani, J. Zhou, and E. T. Mohamad, "Proposing several hybrid PSO-extreme learning machine techniques to predict TBM performance," *Eng. Comput.*, pp. 1–17, Jan. 2022.

[31] 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.

[32] B. T. Le, D. Xiao, Y. Mao, D. He, J. Xu, and L. Song, "Coal quality exploration technology based on an incremental multilayer extreme learning machine and remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4192–4201, Jul. 2019.

[33] Y. Fu, L. Wan, D. Xiao, and B. T. Le, "Remote sensing inversion of saline and alkaline land based on reflectance spectroscopy and D-TELM algorithm in Wuyuan areas," *Infr. Phys. Technol.*, vol. 109, Sep. 2020, Art. no. 103367.

[34] B. R. Murlidhar, D. Kumar, D. J. Armaghani, E. T. Mohamad, B. Roy, and B. T. Pham, "A novel intelligent ELM-BBO technique for predicting distance of mine blasting-induced flyrock," *Natural Resour. Res.*, vol. 29, no. 6, pp. 4103–4120, Apr. 2020.

[35] C. Li, J. Zhou, M. Khandelwal, X. Zhang, M. Monjezi, and Y. Qiu, "Six novel hybrid extreme learning machine–swarm intelligence optimization (ELM–SIO) models for predicting backbreak in open-pit blasting," *Natural Resour. Res.*, vol. 31, pp. 3017–3039, Jun. 2022.

[36] W. Zhao, Z. Zhang, and L. Wang, "Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103300.

[37] W. Wang and J. Wang, "Determinants investigation and peak prediction of CO_2 emissions in China's transport sector utilizing bio-inspired extreme learning machine," *Environ. Sci. Pollut. Res.*, vol. 28, no. 39, pp. 55535–55553, Jan. 2021.

[38] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: A survey," *Int. J. Mach. Learn. Cybern.*, vol. 2, no. 2, pp. 107–122, Jun. 2011.

[39] 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.

[40] Q. S. Liu, J. P. Liu, Y. C. Pan, X. X. Kong, and K. R. 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] Z. Chen, Y. Zhang, J. Li, X. Li, and L. Jing, "Diagnosing tunnel collapse sections based on TBM tunneling big data and deep learning: A case study on the Yinsong project, China," *Tunnelling Underground Space Technol.*, vol. 108, Feb. 2021, Art. no. 103700.

[42] L. Li, Z. Liu, H. Zhou, J. Zhang, W. Shen, and J. Shao, "Prediction of TBM cutterhead speed and penetration rate for high-efficiency excavation of hard rock tunnel using CNN-LSTM model with construction big data," *Arabian J. Geosci.*, vol. 15, no. 3, p. 280, Jan. 2022.

[43] R. Hasanpour, J. Rostami, M. Thewes, and J. Schmitt, "Parametric study of the impacts of various geological and machine parameters on thrust force requirements for operating a single shield TBM in squeezing ground," *Tunnelling Underground Space Technol.*, vol. 73, pp. 252–260, Mar. 2018.

[44] Z. Liu, J. Shao, W. Xu, H. Chen, and Y. Zhang, "An extreme learning machine approach for slope stability evaluation and prediction," *Natural Hazards*, vol. 73, no. 2, pp. 787–804, Sep. 2014.

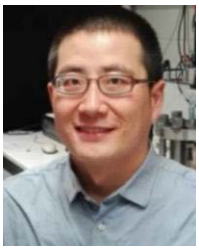
[45] A. Mahmoodzadeh, M. Mohammadi, H. Hashim Ibrahim, S. Nariman Abdulhamid, H. Farid Hama Ali, A. Mohammed Hasan, M. Khishe, and H. Mahmud, "Machine learning forecasting models of disc cutters life of tunnel boring machine," *Automat. Construct.*, vol. 128, Aug. 2021, Art. no. 103779.

[46] J. Zhou, Y. Qiu, S. Zhu, D. J. Armaghani, C. Li, H. Nguyen, and S. Yagiz, "Optimization of support vector machine through the use of Metaheuristic algorithms in forecasting TBM advance rate," *Eng. Appl. Artif. Intell.*, vol. 97, Jan. 2021, Art. no. 104015.

[47] D. Kim, K. Pham, J.-Y. Oh, S.-J. Lee, and H. Choi, "Classification of surface settlement levels induced by TBM driving in urban areas using random forest with data-driven feature selection," *Automat. Construct.*, vol. 135, Mar. 2022, Art. no. 104109.

[48] T. Tirelli and D. Pessani, "Importance of feature selection in decision-tree and artificial-neural-network ecological applications. Alburnus alburnus alborella: A practical example," *Ecol. Inform.*, vol. 6, no. 5, pp. 309–315, Sep. 2011.
[49] R. Tibshirani, "Regression shrinkage and selection via the lasso: A retrospective," *J. Roy. Stat. Soc., B (Stat. Methodol.)*, vol. 73, no. 3, pp. 273–282, 2011.
[50] Q. Liu, X. Wang, X. Huang, and X. Yin, "Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data," *Tunnelling Underground Space Technol.*, vol. 106, Dec. 2020, Art. no. 103595.

LONG LI received the master's degree in mining engineering from Northeastern University, Shenyang, in 2019, where he is currently pursuing the Ph.D. degree in mining engineering. His research interests include TBM big data analysis and the application of machine learning in tunnels.



ZAOBAO LIU received the Ph.D. degree in civil engineering from the University of Lille 1–Sciences and Technologies, in 2013. He was at the Lille Laboratory of Mechanics, from 2013 to 2018. He was granted the "Overseas Hundred Program" of Northeastern University, as a Full Professor, in 2018. His research experiences in more than ten research projects during the last ten years allow him authoring more than 60 refereed international journal articles and conference proceedings. His main research interests include multiscale approaches for multi-physics problems in rock mechanics, and intelligence prediction and safety control of engineering failure and hazards.

YUCHI LU was born in Chaoyang, China, in 1997. He received the bachelor's degree in urban underground space engineering from the Shenyang University of Technology, in 2020. He is currently pursuing the master's degree in civil engineering with Northeastern University, Shenyang. His research interest includes the application of artificial intelligent in TBM tunnels.

FEI WANG received the Ph.D. degree in geotechnical engineering from TU Bergakademie Freiberg, Germany. He is specialized in research on multi-physics coupling of rock materials under conditions that characterize the deep rock mass engineering, tunnel and monument fires. His research interests include thermo-hydro-mechanical couplings in rocks, continuum and discontinuum numerical modeling, deep geological storage, geothermal exploitation, and tunnel construction.

SEOKWON JEON received the Ph.D. degree in mining engineering from the University of Arizona. His research experiences in more than six research projects. He is a Full-Time Professor with the Department of Energy Systems Engineering, Seoul National University (SNU). His main research interests include mechanized tunneling, explosive blasting, tunnel design and stability analysis, rock fracture mechanics and fragmentation, ground subsidence and its prevention at abandoned mines, and geological disposal of radioactive waste.