

Realtime prediction of hard rock TBM advance rate using temporal convolutional network (TCN) with tunnel construction big data

Zaobao LIU^{a,b*}, Yongchen WANG^a, Long LI^a, Xingli FANG^c, Junze WANG^a

^a Key Laboratory of Ministry of Education on Safe Mining of Deep Metal Mines, Institute of Deep Engineering and Intelligent Technology, Northeastern University, Shenyang 110819, China

^b State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu 610065, China

^c Software College, Northeastern University, Shenyang 110819, China

*Corresponding author. E-mail: liuzaobao@mail.neu.edu.cn

© Higher Education Press 2022

ABSTRACT Real-time dynamic adjustment of the tunnel bore machine (TBM) advance rate according to the rock-machine interaction parameters is of great significance to the adaptability of TBM and its efficiency in construction. This paper proposes a real-time predictive model of TBM advance rate using the temporal convolutional network (TCN), based on TBM construction big data. The prediction model was built using an experimental database, containing 235 data sets, established from the construction data from the Jilin Water-Diversion Tunnel Project in China. The TBM operating parameters, including total thrust, cutterhead rotation, cutterhead torque and penetration rate, are selected as the input parameters of the model. The TCN model is found outperforming the recurrent neural network (RNN) and long short-term memory (LSTM) model in predicting the TBM advance rate with much smaller values of mean absolute percentage error than the latter two. The penetration rate and cutterhead torque of the current moment have significant influence on the TBM advance rate of the next moment. On the contrary, the influence of the cutterhead rotation and total thrust is moderate. The work provides a new concept of real-time prediction of the TBM performance for highly efficient tunnel construction.

KEYWORDS hard rock tunnel, tunnel bore machine advance rate prediction, temporal convolutional networks, soft computing, construction big data

1 Introduction

The tunnel bore machine (TBM) has become one of the most advanced technologies for tunnel construction, due to its advantageous characteristics of fast, safe, and environment-friendly construction mode. It has been widely promoted in the excavation of ground traffic tunnels and the construction of water diversion tunnels [1]. However, the TBM has poor adaptability to complex geological conditions, which cause decision-making errors in tunnels with engineering-difficult geological conditions. It also degrades the construction efficiency [2,3]. The prediction of the TBM performance parameters

can help the TBM operator optimize the tunnel construction schedule and thus improve construction efficiency when dealing with engineering-difficult geological conditions [4]. The advance rate is one of the most important TBM performance parameters that the TBM operator optimizes to improve construction efficiency.

The advance rate however is difficult to estimate since it is affected by many different factors. Fortunately, the TBM advance rate can be predicted based on the data collected over a large number of rock-machine interaction parameters using the soft computing techniques. It has been an advantage predictive model to many engineering fields.

For example, the artificial neural network (ANN), fuzzy logic [5] and adaptive neuro-fuzzy inference

system [6] were applied to predict the compressive strength of concrete. The general regression neural network, and the extreme learning machine (ELM), were integrated to predict the ground settlement induced by tunnel excavation [7], and other geotechnical problems [8]. The support vector machine (SVM) [9], the ELM [10], particle swarm optimization (PSO)-neural network [11], were also used to predict slope stability and displacement [12]. The relevance vector machine classifier, and the cuckoo search optimization, were combined for spatial prediction of landslides [13]. The ANNs and SVMs were also used to predict unconfined compressive strength of soils stabilized [14]. It can be seen that the application of soft computing methods in engineering forecast of problems, is of a great progression in solving practical engineering problems. Different soft calculation methods have also been developed and applied to predict TBM penetration rate or advance rate.

The TBM advance rate is mainly affected by parameters of penetration rate [15]. The models including traditional or optimization regression analysis [16,17], PSO [18], SVM, and its optimization [19,20], fuzzy logic method [21], were used to predict penetration rate. The linear multiple regression, and non-linear multiple regression, were proposed to predict TBM penetration rate and advance rate [22]. The Grey Wolf optimizer-feature, weighted-multiple kernel-support vector regression technique, was used to predict the penetration rate [23]. The gene expression programming equation was developed to predict the penetration rate [17,24,25]. The risk matrix method was used to analyze the geotechnical risk and predict the advance rate [26]. Additionally, a variety of optimization algorithms have been used to optimize the SVM to predict the advance rate [27]. However, the rock-machine interaction parameters may have a certain correlation in-time series. The outputs of traditional machine learning methods are only determined by the current input without prior learning information, thus they cannot be directly applied to the real-time prediction of TBM operation parameters [28].

In addition, deep learning as a sub-field of machine learning, has been widely used to predict penetration rate and advance rate due to its mightier computing power. For example, an ANN has been developed to predict advance rate based on rock-machine interaction parameters [29,30]. Two hybrid intelligent systems, the PSO-ANN and the ICA-ANN, were established to predict penetration rate, respectively [31]. Firefly algorithm combined by ANN aimed to predict penetration rate of TBM [32]. Several other optimized ANN combination models were used to predict advance rate [33]. A variety of optimized PSO-ELMs were also used to predict advance rate [34]. The total thrust, cutterhead rotation, and rock characteristics, were used as input parameters to

realize the estimation of advance rate based on XG-Boost and Bayesian optimization. It is verified that the machine parameters have a more evident influence on the prediction of advance rate [35]. However, due to the structural limitations of traditional neural network, it has poor memory for data with temporal attributes. At present, the most widely used soft computing method for processing data with time attributes, is the recurrent neural network (RNN) [36] and long short-term memory (LSTM) network proposed on the basis of RNN to effectively resolve the problems of gradient disappearance and gradient explosion [37]. The LSTM has been applied to the prediction of TBM performance [38], tunnel lithology [39] and pore-water pressure in front of a TBM [40] based on construction big-data.

However, recent studies have shown that the temporal convolutional network (TCN) can retain longer-term memory than LSTM. It also has a simpler and more precise frame structure design. The prediction problem of long sequences is precisely compared with LSTM, and the prediction accuracy is improved to a certain extent [41]. In addition, the TBM rock-machine interaction parameters have significant volatility and complex change trends. Therefore, it is necessary to select a long sequence of rock-machine interaction parameters as the time sequence input length. TCN can have absolute accuracy advantages in its prediction after training. In response to the above, we propose to use TCN to establish a TBM advance rate prediction model, and compare it with the RNN and LSTM prediction models to verify the prediction accuracy of TCN for TBM parameters.

The structure of this paper is as follows. Section 2 introduces the principle and architecture of TCN. Section 3 introduces the engineering background and data analysis that this research relies on and the construction of the advance rate prediction model. Section 4 introduces the prediction results and discussion. Finally, the corresponding conclusions are obtained.

2 Temporal convolutional network principle and architecture

In the process of TBM excavation, interactive parameters of the rock-machine are collected over time. From the perspective of deep learning, the prediction of TBM advance rate can be regarded as a time series prediction problem. The powerful nonlinear mapping capability of TCN can be used to learn complex relationships from the vast original database.

It has the following two characteristics as a new sequence analysis model. Firstly, causality between layers of TCN is generated to reduce the occurrence of historical information omission and future data omission. Secondly,

the architecture of TCN model can be adjusted to any depth and can be mapped to the corresponding output according to the number of output channels. TCN mainly relies on the structural design of causal convolution, residual block, and dilated convolution, to improve the predictive ability of the model.

2.1 Casual convolution

Causal convolution was first proposed in the WaveNet network [42]. It is a convolution model used to deal with sequence problems. The sequence problem can be transformed into predicting the information at time t based on the sample information before time t .

Causal convolution has two characteristics. The first characteristic is: The given input sequence is $[X_1, \dots, X_{T-1}, X_T, X_{T+1}, \dots, X_{T+n}]$ only, according to the time before t sequence $[X_1, \dots, X_{T-1}]$ to predict Y_T , is utilised instead of using the sequence $[X_{T+1}, X_{T+2}, \dots]$ that has not started learning. Accordingly, the predicted value at this time is only related to the information at the previous time. Secondly, the longer the historical information that needs to be memorized, the more hidden layers, will be needed.

The causal convolution model is shown in Fig. 1. It can be observed that the result of the causal convolution at a certain point in the future will take into account the information before that point in time. The advantage is that it can be used to deal with prediction problems with time series attributes. However, when the prediction result requires long-term information, more layers of convolutional layers are needed. This leads to the more common problems in deep learning such as, complex training, vanishing gradient, and poor fitting effects. Therefore, the joint application of Dilated Convolution and Residual block is required to solve the above

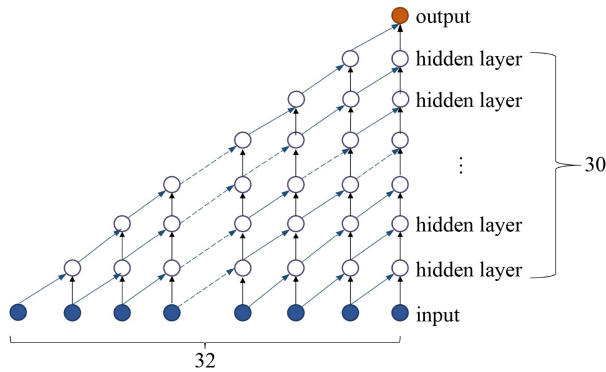


Fig. 1 Causal convolution structure.

problems.

2.2 Residual block

The conventional design idea of convolutional neural networks, is the deeper the network design, the better the effect. However, with the increase of the number of hidden layers in the conventional network (as illustrated in Fig. 1), the easier the phenomenon described in Subsection 2.1, will occur. Therefore, the fundamental to improve performance is how to solve the vanishing gradient problem, on the basis of increased network depth. In response to this phenomenon, the TCN model uses a residual module, as is shown in Fig. 2.

The residual block is a network structure widely used to solve the vanishing gradient problem primarily by making some information move onto the hidden layer. The cross-layer output information, and the information output by the hidden layer, in the normal convolution condition, are arithmetic added to obtain the output result. Establishing the identity mapping in the model enables the structure of the entire model to converge in the direction of the identity mapping. The final error rate will not increase with the increase of the depth of the hidden layer of model, as is described in Eq. (1) [41].

$$o = \text{Activation}(x + f(x)), \quad (1)$$

where x represents the module information of the identity mapping, and $f(x)$ represents the information entered into the current layer by the previous hidden layer.

2.3 Dilated convolution

The form of injecting holes into the standard hidden layer to form spaced convolution is called, dilated convolution. Dilated convolution can improve the structure of CNN, obtain a larger receptive field, and reduce the omission of information [43]. The hyper-parameter of CNN contains the size of the convolution kernel. The expansion convolution in addition, contains a dilation parameter used to indicate the size of the hole. Ordinary convolution receptive field is calculated according to Eq. (2) [44].

$$R_k = R_{k-1} + \left[(N_k - 1) \times \prod_{i=1}^{k-1} S_i \right], \quad (2)$$

where R_k represents the receptive field size of the k th layer, S_i represents the stride of the i th layer, and N_k

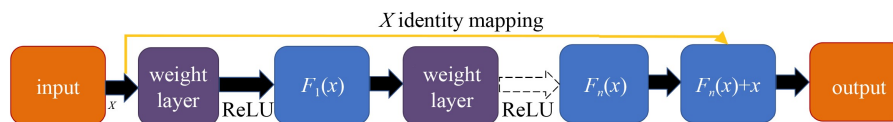


Fig. 2 Residual block structure.

represents the size of the convolution kernel of the current layer.

Dilated convolution can increase the size of the receptive field without missing pooling layer information. The structure of dilated convolution is manifested in Fig. 3. It can be observed that the dilated convolution is equivalent to inserting d holes into the convolution kernel. The size of the convolution kernel of the dilated convolution is then calculated by Eq. (3) [45].

$$N_{k,d} = (d - 1) \times (N_k - 1) + N_k, \quad (3)$$

where $N_{k,d}$ represents the Kernel Size after adding dilated convolution. The receptive field of dilated convolution has been greatly improved according to the equation. The calculation of the output result calculated by Eq. (4).

$$c_t = W_c \bigoplus_{k=0}^r x_{t \pm d}, \quad (4)$$

where W_c is the weighting matrix, and $x_{t \pm d}$ is the information of current layer input to the next layer.

The input data of this article is a 4-dimensional feature sequence $X_i = [X_1^i, X_2^i, \dots, X_{32}^i]$ with a length of 32. The stride (s) is set to 1. The size of the convolution kernel (R_k) is set to 2, and the dilated rate of the k th layer (d_k) is set to 2^{k-1} . It can be known that in the case without dilated convolution, 30 hidden layers (as is shown in Fig. 1) are required for the prediction result by taking the state of the sequence into account at all times. Regardless, 7 hidden layers are required with dilated convolution and two residual blocks (as is shown in Fig. 3).

Figure 4 shows a general flow of the TCN. In Fig. 4, M , represents the length of the input time sequence. A , represents the length of the output time sequence. N ,

represents the number of M in a single tunneling cycle. S represents the number of tunneling cycles, and T , represents the dimension of the input feature. K , represents the number of convolution kernels. The overall design framework of the model is demonstrated in Fig. 5. Firstly, we select the features and prediction parameters, and then divide the selected loop into training set, validation set, and test set. The training set is passed forward through the TCN layer and the full connection layer, and then the parameter weightings are back-propagated to update. The validation set is then used for evaluation until the model reaches the optimal effect within the selected range of hyper-parameters, and the training is stopped. Finally, the test set is used for testing.

3 Tunnel bore machine data analysis and model construction

3.1 Project background

In this paper, a granite cross section of the Yinsong water diversion project in Jilin province, is taken as the research

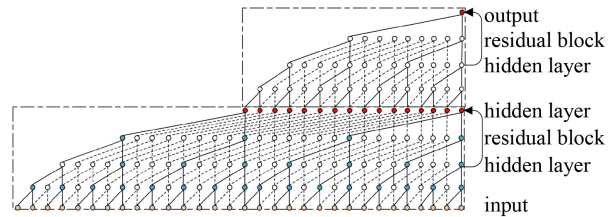


Fig. 3 One-dimensional dilated temporal convolution with two blocks.

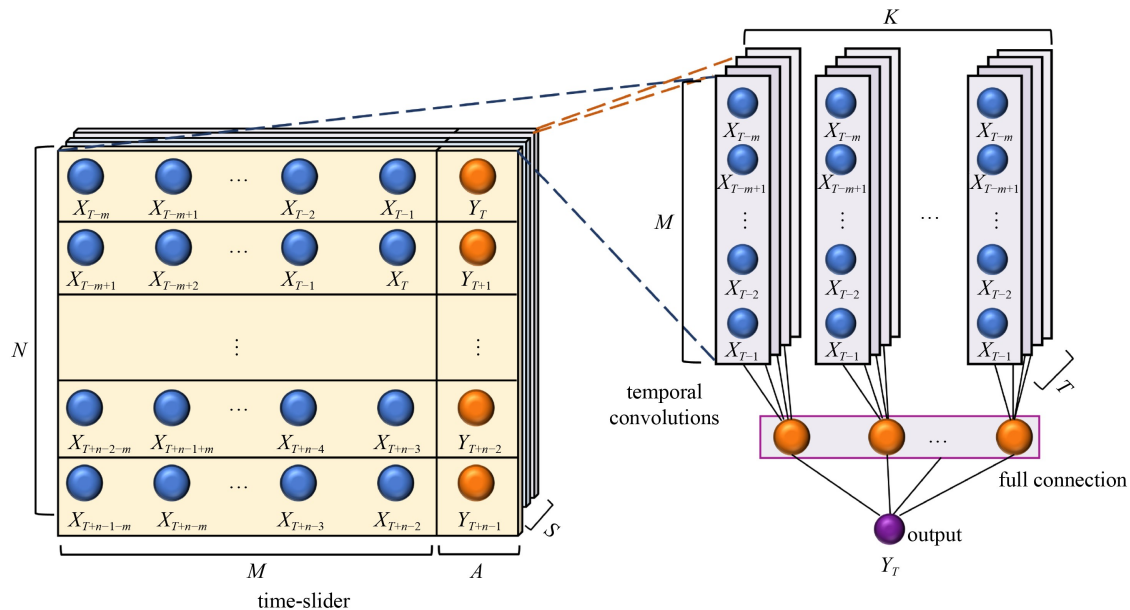


Fig. 4 Flow chart of TCN.

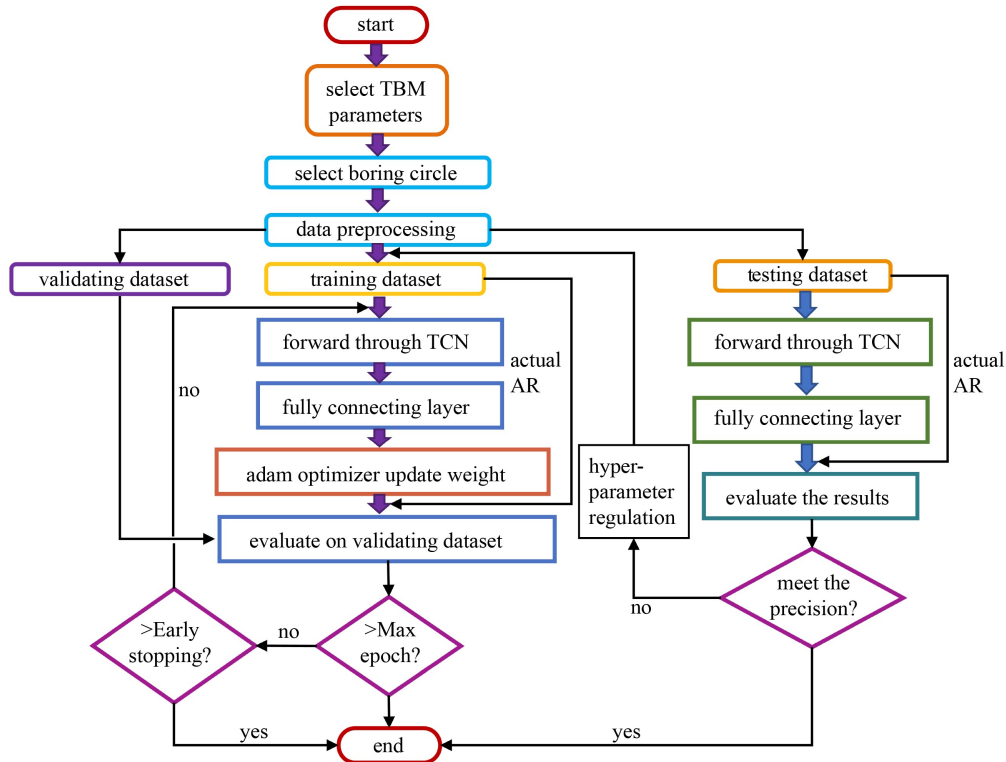


Fig. 5 Flow chart of advance rate prediction model.

object. A prediction model of advance rate based on TCN is established. The total length of the diversion tunnel project is 263.02 km. There are four sections using a total of 3 TBMs in construction. The length of the four sections is 19.7 km in total. The construction length of the TBM reaches 17.5 km. The tunnel excavation diameter is 7930 mm. The buried depth of the tunnel is 85 to 260 m. The TBM is equipped with 56 hobs. The maximum advance rate is 120 mm/min. The driving power is 3500 kW. The rated propulsion is 23260 kN, and the effective tunneling days is 728 d. The TBM information collection frequency is 1 Hz, and the total data volume reaches 4.08 billion rows with 199 parameters (such as advance rate, penetration rate, etc.). The data set used in this study adopts the pile ranging from k52+960 to k53+535.

3.2 Data analysis and preprocessing

The data is composed of multiple cycles. Each cycle can be divided into four periods over time, by visualizing the established TBM rock-machine interaction parameter database, as shown in Fig. 6. The cutterhead rotates but the TBM does not have any displacement when the TBM is in the initial period. Further analysis shows that the cutterhead has a forward displacement, but only overcomes the friction force generated on the surface of the tunnel wall, and the cutterhead does not generate thrust. The remaining two periods are the increase period, where most of the parameters are in a sharp increase period and

the stable period, where most of the parameters are floating in a small range.

From the above analysis, the increase and stable period occupy most of the time of the cycle. It can be also known that TBM does not touch the tunnel face in empty-void period. It is meaningless to predict the advance rate of empty-void period. Therefore, this study establishes the prediction model for the TBM advance rate in the increase period and the stable period. The research has intercepted 800-second continuous time series samples from each cycle. Each sample includes an increase period and a stable period as a typical cycle with sufficient time series length to train the model. The change trend of some parameters in a typical tunneling cycle is demonstrated in Fig. 7.

Pearson correlation coefficient, is widely used to measure the degree of correlation between two variables. The value range is between -1 and 1 , and is calculated by Eq. (5). In this study, the correlation degree of the research parameters is divided into 4 levels. $|R| < 0.4$ is a weak correlation. $0.4 < |R| < 0.6$ is a moderate correlation. $0.6 < |R| < 1.0$ is a strong correlation. $R = 1$ is a perfect positive correlation. $R = -1$ is a perfect negative correlation. The correlation between the two variables can only be caused by accidental factors, so we have to judge the significance level of the correlation between the two variables. It indicates that there is a significant linear correlation between the two variables, when $p < 0.05$.

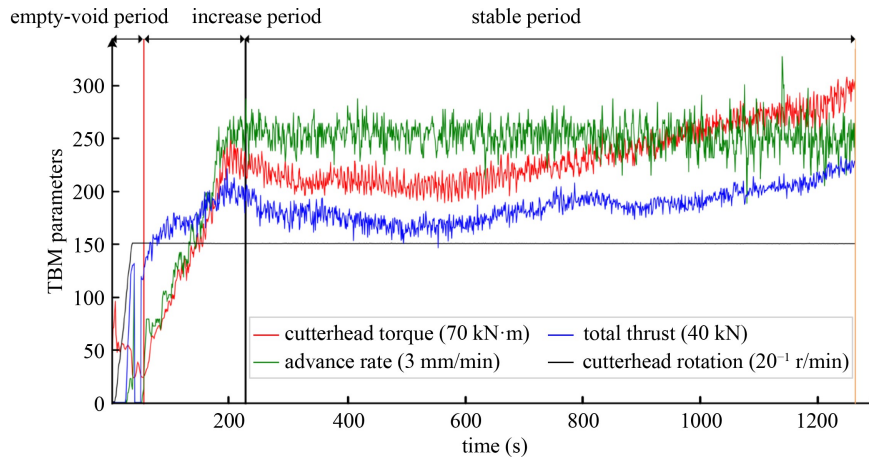


Fig. 6 Variation trend of some parameters in complete driving cycle.

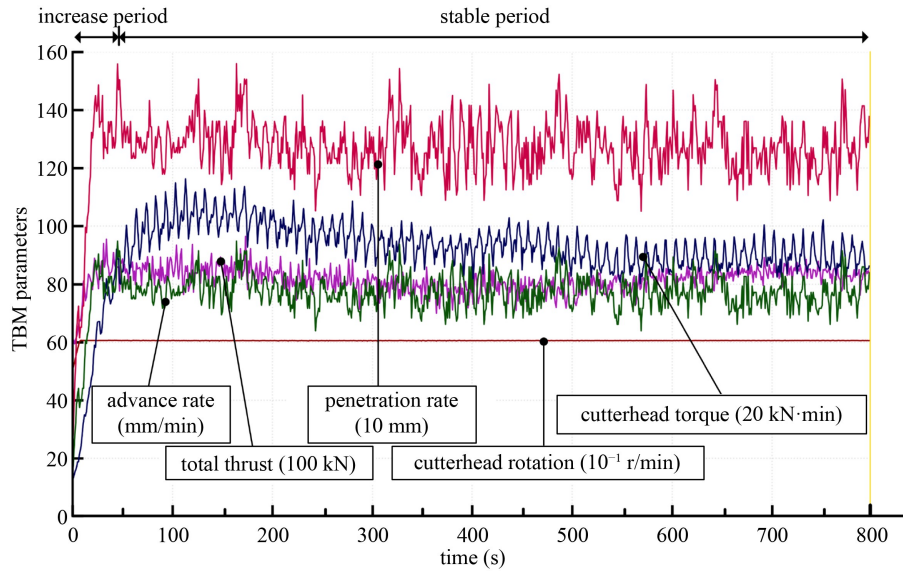


Fig. 7 Variation trend of some parameters in typical driving cycle.

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}. \quad (5)$$

The correlation analysis between the characteristic parameter in the first seconds and the current moment is carried out, and the result is displayed in Fig. 5. The Pearson correlation coefficients of penetration, and cutterhead torque, are 0.9 and 0.78 respectively, and the correlation is strong. The Pearson correlation coefficient values of cutter head speed, and total propulsion, are 0.47 and 0.54, respectively, and the correlations are both moderate. The significance level p-values are all 0. Therefore, there must be a significant linear relationship between the selected feature parameters and the prediction parameters, and it can be seen from Fig. 5 that there is also a nonlinear relationship.

It can be seen from Fig. 8 that the unit dimensions

between different rock-machine interaction parameters of TBM are different. The numerical range between the parameters is very large. It is unreasonable to directly input heterogeneous data with a large difference in numerical range with different features into the neural network without processing. Larger gradient updates may occur, which can cause the model to fail to reach a state of convergence. The sample features should have the characteristics of homogeneity to make it easier for the training of the model to achieve satisfactory test results. Most of the values are needed in the range of 0 to 1, and the values of all features should be in the same range before input to the network model.

The normalization of the data is carried out. The data range of different interaction parameters of TBM that differed greatly is eliminated. The influence of sharp numerical changes under the same dimensions is eliminated. The effects of drastic numerical changes at the same scale are resolved. The interaction parameters of

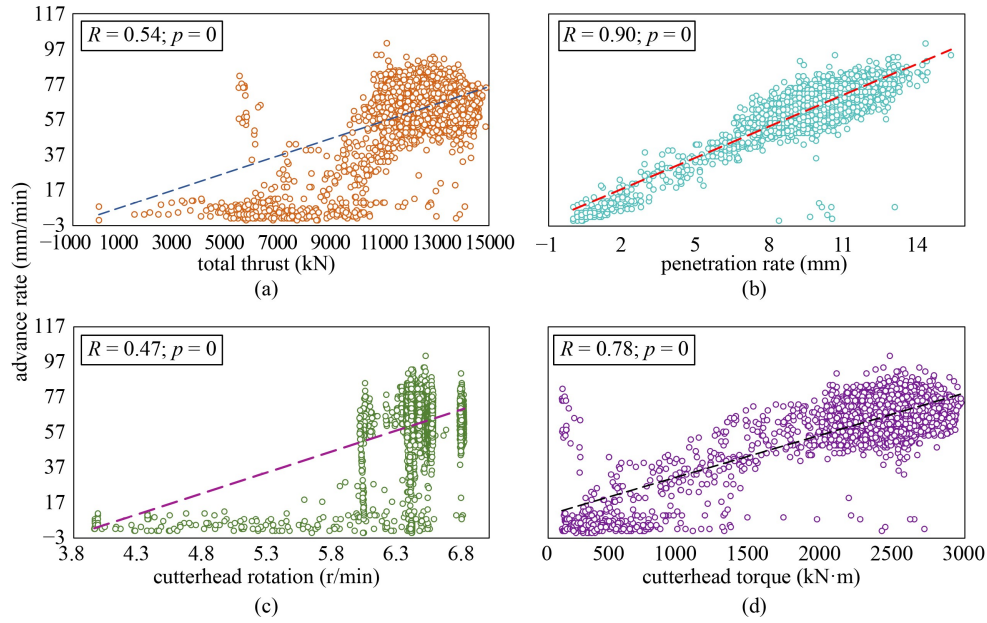


Fig. 8 Correlation analysis between characteristic parameters and advance rate: (a) total thrust; (b) penetration rate; (c) cutterhead rotation and (d) cutterhead torque.

TBM machines with different dimensions can be compared and weighted. Standard score normalization is used to scale the 4 rock-machine interaction parameters, and the parameter mean is 0 and the variance is 1, which is achieved by Eq. (6) [46]:

$$x' = \frac{x - \mu}{\sigma}, \quad (6)$$

where x is the initial input data, x' is the normalized input data, μ and σ are the mean and standard deviation of the original data, respectively.

3.3 Construction of advance rate prediction model

In this study, all datasets are divided into a training set, and a test set. The training set and validation set of the training model, used randomly selected data from 80% of 188 cycles of data. The validation set used 40 cycles of the data and test set of the model used 20% of 47 cycles of data. The TCN based on Keras framework is used to develop the TBM advance rate prediction model. The prediction model is mainly composed of TCN layer, ReLU layer, and output layer. The frequency of information collection of TBM is 1 Hz. The TBM data is changeable. This research uses the time sequence length in the 32 seconds to scroll to predict the parameters of the next second to enable the model to learn enough deformation trends with appropriate calculation. Therefore, the time sliding window is set to 32. After multiple trainings, it can be seen that TCN selects 125 convolution kernels with a size of 2 to have better results. Therefore, the values of parameters such as M, A, N, S, T, and K are 32, 1, 768, 235, 4, and 125 in Fig. 4, respectively.

The activation function of the hidden layer uses the ReLU activation function calculated according to Eq. (7) [47].

$$\text{ReLU}(x) = \max(0, x). \quad (7)$$

In addition, although we have done normalization in the initial period of data input, it is likely to change the data distribution after the input data undergoes non-linear processing. Inclusively, the multi-layer operation of the network, the data is being distributed more and more widely. Therefore, Batch normalization [48] is added between the TCN and ReLU layers, to speed up the training and improve the generalization ability of the network. The principle is demonstrated as follow. μ, σ : functions of x , analogous to responses; γ, β : parameters to be learned, analogous to weightings.

Input: Values of x over a mini-batch: $B = \{x_1, x_2, \dots, x_m\}$;
Parameters to be learned: γ, β .

For $i = 1$ to m , do:

- 1) calculate the mini-batch mean: $\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i$;
- 2) calculate mini-batch variance: $\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2$;
- 3) normalize: $\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \tilde{O}}}$;
- 4) scale transformation and scale shift: $y_i \leftarrow \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i)$;
- 5) return the learned parameter γ and β .

Output: $\{y_i = \text{BN}_{\gamma, \beta}(x_i)\}$.

In addition, the mean square error is selected as the loss function of the model, as is calculated in Eq. (8).

$$L(x, y) = \frac{\sum_{i=1}^n (x_i - y_i)^2}{n}, \quad (8)$$

where x_i is the target value, and y_i is the predicted value.

Adam optimization algorithm [49] is used to update the weighting, which can effectively calculate and accelerate the training process. The hyper-parameter search adjustment is realized by the hyperband algorithm [50]. It is verified that if Learning rate and Learning rate decay of Adam optimizer are set to 9×10^{-4} and 0.3 respectively, the model fits, will be better. Early stopping is used to reduce the over-fitting of the model and improve the generalization [51] according to the performance of the verification set and the dropout regularization method [52]. The main hyper-parameters of TCN, RNN, LSTM model and their values are shown in Table 1.

4 Results and discussion

4.1 Prediction by TCN-based model

The mean absolute error (*MAE*), root mean square error (*RMSE*), and mean absolute percentage error (*MAPE*), given in Eq. (9), are used to evaluate the deviation between the predicted values and the measured value. The coefficient of determination R^2 is used to indicate the fitting effect of the model. Thereby, the predictive performance and generalization of TCN-based model are expressed.

$$\begin{aligned} MAE &= \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \\ RMSE &= \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}, \\ MAPE &= \frac{100\%}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i}, \\ R^2 &= 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (y_i - \bar{y})^2}, \end{aligned} \quad (9)$$

where \hat{y}_i is the predicted value, y_i is the measured value, \bar{y} is the sample average, and N is the number of samples. *RMSE* focuses on the overall data set deviation. Since the actual measured value of the advancing speed in this study contains zero values, *MAPE* cannot be used directly as a metric in this study. The zero value is eliminated and then the comparative evaluations of different models are carried out. The higher the value of R^2 is, the closer the

predicted value of the model will be to the measured value.

Figure 9 shows the prediction accuracy of advance rate by TCN. As shown in Fig. 9, the trend of the predicted value is basically the same as the measured value of the advance rate. Figure 9 also shows the absolute error between the predicted value and the measured value of the advance rate.

The error variation range is wide due to the large error of individual timing points, but the overall error variation range is mainly between 0 to 4 mm/min. Large errors generally occur when advance rate mutations occur. The deviation range between the predicted and the measured value during the stable period, especially in the later period, is relatively small. This may be due to the fact that the historical information available to the model gradually increases as the rolling input timing increases, in each cycle during model training. At the same time, as shown in Fig. 5, the contact time between the cutterhead and the tunnel face increases during the increase period of the advance rate. In this phase, the values of the machine parameters change significantly, which makes the prediction more complicated.

The *RMSE*, *MAE*, *MAPE*, and R^2 are used to evaluate the performance of the TCN real-time prediction models for advance rate and the results are given in Table 2. The

Table 1 Hyper-parameters of the advance rate prediction model

hyper-parameter	candidate value	TCN value	RNN value	LSTM value
$n_{\text{neurons}}^a)$	[25, 125, 225, 325, 425]	125	325	325
$k_s^b)$	[2, 3, 4]	2	—	—
$d_s^c)$	[[1, 2, 4, 8]; [1, 2, 4, 8, 16]; [1, 2, 4, 8, 16, 24]]	[1, 2, 4, 8, 16]	—	—
$n_{\text{stacks}}^d)$	[1, 2]	2	—	—
n_{dropout}	[0.2, 0.3, 0.4, 0.5]	0.3	0.3	0.3
learning rate	[0.003, 0.001, 0.0009, 0.0003, 0.0001]	0.0009	0.001	0.0003

Notes: a) n_{neurons} : number of filters used in the TCN, number of nodes used in the RNN; b) k_s : size of the kernel used in each convolutional layer; c) d_s : dilated convolutions; d) n_{stacks} : number of stacks of residual blocks to use; e) n_{dropout} : probability of each filter being removed.

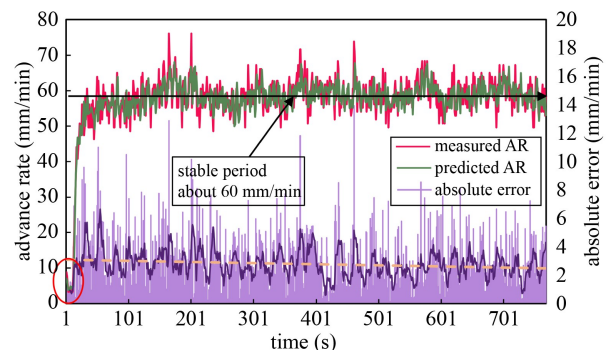


Fig. 9 Comparison of measured value and TCN predicted value of advance rate.

results on the test data set are, respectively, 4.3403 mm/min, 3.3349 mm/min, 6.182%, and 0.7213 for the *RMSE*, *MAE*, *MAPE*, and R^2 . The *RMSE* and *MAE* have the same dimension as the propulsion speed, and the average value of the advance rate, is usually in the range of 30 to 100 mm/min. Therefore, the values of the *RMSE* and the *MAE* of this model are within the acceptable range. Since the *MAPE* is less than 10%, the *MAPE* is also within the acceptable range. From the evaluation result of the test set R^2 , it can be seen that there is no under-fitting or over-fitting, which indicates the model fits well.

4.2 Comparison and analysis of temporal convolutional network-based and recurrent neural network-based models

In this study, the prediction effect of TCN on TBM advance rate is verified by the comparison with the prediction of RNN and LSTM. The database is divided four times in the same proportion (according to Subsection 3.3) for training and testing separately, to verify the prediction accuracy of the model on different training sets, and test sets. R^2 is used to evaluate the performance of the TCN, LSTM and RNN real-time prediction models for advance rate and the results are given in Table 3. As shown in Table 3, TCN model outperforms LSTM and RNN model in predicting TBM advance rate in all datasets, indicating that TCN has higher accuracy in predicting TBM advance rate in this database. However, different datasets have different precision performance, and the performance of the three prediction models in dataset 3 is better than that of other datasets. The representation of each model in dataset 3 is expanded to analyze the data representation of each model more deeply.

The fitting of the predicted and measured values of the advance rate of the three models on the same randomly selected test set, is visualized by Figs. 9 and 10. It can be

observed from Figs. 9 and 10 that the predicted trends of the three models in the stable period of the advance rate are in good agreement with the change trends of the measured values.

The prediction of the three models has a large deviation when the mutation value of the advance rate is high. That may be because after the model reaches a certain mapping range, the range of some mutation values is excessive, resulting in low prediction accuracy. In addition, there is a buffer period in about ten seconds at the beginning of the increase period, and the predictive ability of TCN at this period is significantly better than that of RNN and LSTM. However, the overall predictions of the three models all have high prediction accuracy.

Figure 11 demonstrates the absolute error distribution between the predicted and measured advance rate of the three prediction models. To better compare the prediction errors of each model, the periodic moving average error line is adopted in Fig. 11. According to the periodic moving average error in Fig. 11, it can be clearly observed that the absolute error generated by RNN is the largest in the whole cycle. The absolute error generated by TCN is the smallest in most positions. It is clearer that TCN has better performance for the prediction of advance rate.

The initial stage of the increase period and the stable period are visualized to compare advance rate measured and predicted value of TCN, LSTM, and RNN in more detail respectively, as is shown in Fig. 12. It is found that the forecasting trends of TCN, RNN and LSTM on advance rate are basically the same. The research finds that, the individual position shows a trend that is contrary to the change by comparing predicted values with measured values, which may be related to the auto-correlation of the advance rate. However, the overall agreement is good, and the predicted value based on the TCN prediction model tends to change the curve of the measured value, especially in the early stage of the increase period.

The prediction accuracy of TCN, RNN, and LSTM, on the same training set and test data set, is compared through performance evaluation indicators to further prove the prediction performance of the advance rate prediction model proposed in this study. The results are illustrated in Table 4. The TCN, LSTM, and RNN, developed in this research all showed good generalization

Table 2 Performance of TCN-based model on different data sets

evaluation index	training & validation set	training set	test set
<i>RMSE</i> (mm/min)	4.0894	4.0393	4.3403
<i>MAE</i> (mm/min)	3.1582	3.1282	3.3349
<i>MAPE</i>	5.612%	5.556%	6.182%
R^2	0.6910	0.6891	0.7213

Table 3 Models prediction results of advance rate on different datasets based on R^2

model	dataset 1		dataset 2		dataset 3		dataset 4		average	
	TR ^{a)}	TS ^{b)}	TR	TS	TR	TS	TR	TS	TR	TS
TCN	0.653	0.673	0.681	0.710	0.689	0.721	0.658	0.689	0.670	0.698
LSTM	0.628	0.605	0.711	0.706	0.663	0.712	0.650	0.669	0.653	0.673
RNN	0.617	0.596	0.602	0.586	0.637	0.693	0.626	0.617	0.621	0.623

Notes: a) TS: testing; b) TR: training.

on the test set, and no over-fitting occurred. In the training set and the test set, the prediction accuracy of LSTM is improved more than that of RNN in all aspects, by comparing the three evaluation indexes of *RMSE*, *MAE*, and R^2 . This is related to the memory processing ability of the RNN for the sequence samples. LSTM is an improved recurrent neural network based on RNN. For samples with longer time series, the memory ability has

been improved by controlling various gating structures, and has a good performance.

It is obvious that the performance of the RNN is superior by comparing the *MAPE* of the two models. This may be because the prediction performance of the RNN is better than that of the LSTM when the actual measured value of the advance rate is small. Under all the evaluation criteria, the predictive ability of the TCN model is better than that of the other two models. The coefficient of determination R^2 is the largest with 0.7213, and the model fitting is the best. Especially compared with RNN, the effect is improved more obviously. The TCN time series convolutional network is suitable for predicting TBM advance rate to a certain extent.

The number of optimal cycles of the three models in the test set (47 cycles) is counted under each indicator to prove the generality of the performance of the TCN model. This operation is performed for all four evaluation indicators. According to Fig. 13, the TCN *RMSE*, *MAE* minimum cycle number, and R^2 maximum cycle number

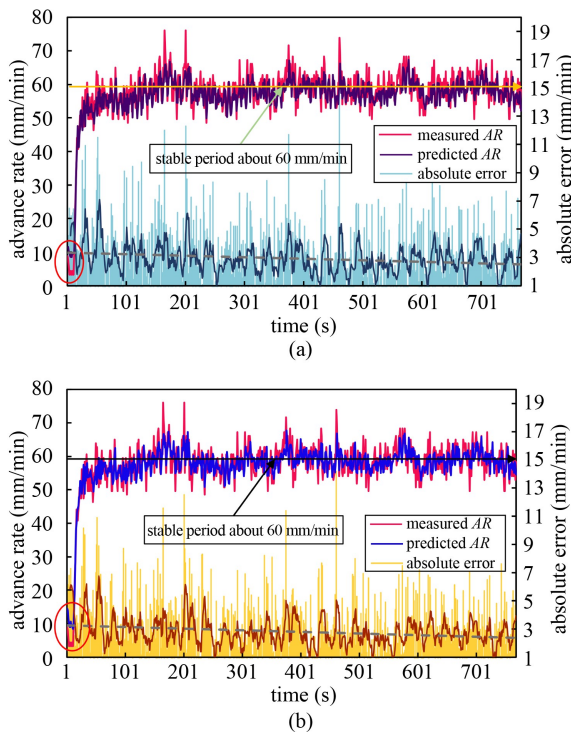


Fig. 10 Comparison of the predicted and measured advance rate based on RNNs: (a) RNN model and (b) LSTM model.

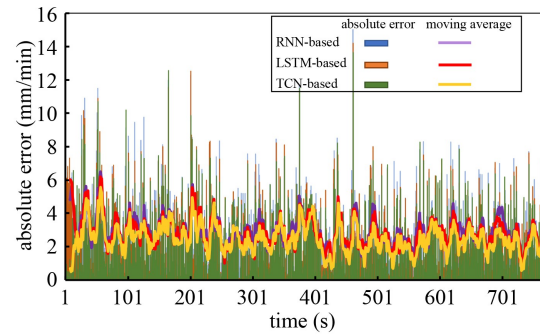


Fig. 11 Comparison of absolute errors of all models.

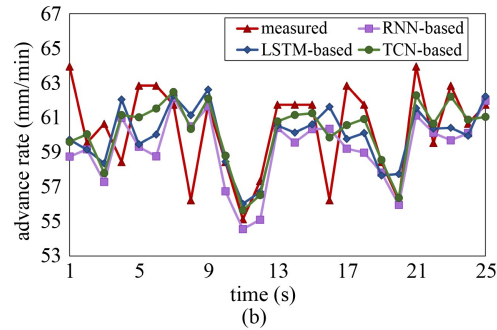
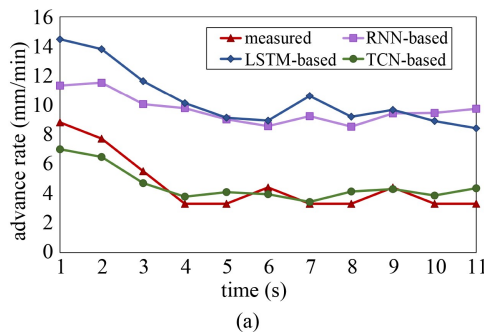


Fig. 12 Partial comparison between the measured and the predicted value of advance rate: (a) early period and (b) stable period.

Table 4 Comparison of performance evaluation indexes of prediction models

model	<i>RMSE</i> (mm/min)		<i>MAE</i> (mm/min)		<i>MAPE</i>		R^2	
	TR	TS	TR	TS	TR	TS	TR	TS
TCN	4.0393	4.3403	3.1282	3.3349	5.556%	6.182%	0.6891	0.7213
LSTM	4.2242	4.4246	3.2564	3.4888	5.954%	6.864%	0.6636	0.7115
RNN	4.3947	4.5676	3.3669	3.3978	5.929%	6.808%	0.6366	0.6925

are all 37, and the minimum cycle number of *MAPE* is 35, which shows the values of TCN model are better than those of the other two models by absolute advantage. It is further confirmed that TCN model is better than RNN and LSTM model in predicting the advance rate in the whole test set.

Figure 14 shows the correlation and linear relationship between the measured value of advance rate and each model's predicted value of advance rate. The TCN model is more convergent to the measured value, and the slope of its linear expression is 0.9837, which tends to 1 compared with the 1.0506 of LSTM, and 1.042 of RNN models.

Therefore, the prediction performance of the TCN model is optimal on both a single cycle and the whole test set according to the above analysis. The results show that the TCN has advantages in predicting the TBM advance rate. Thus, it is suggestable to use the TCN model for TBM advance rate prediction.

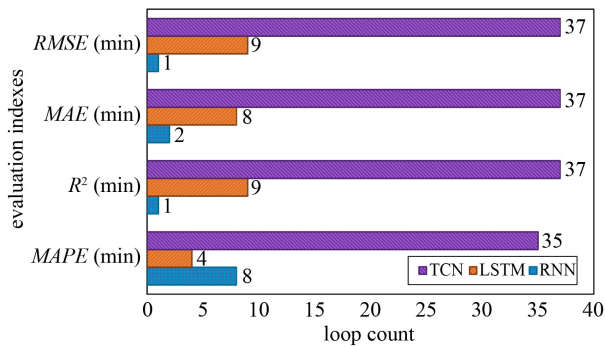


Fig. 13 Optimal cycles of evaluation indexes of each model.

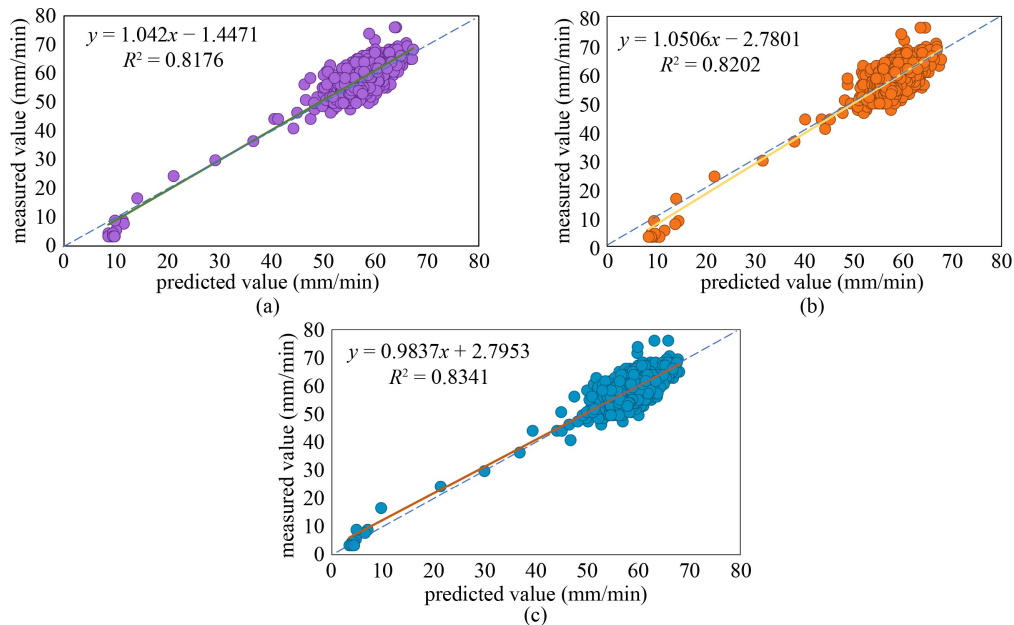


Fig. 14 Fitting of predicted and measured advance rate: (a) RNN model; (b) LSTM model; (c) TCN model.

5 Conclusions

Based on the study of real-time prediction of TBM advance rate by the TCN model, a conclusion can be drawn as follows.

1) The TCN-based model proposed in this paper has satisfactory predictive performance in TBM advance rate prediction. On the test set, the *RMSE* is 4.3403 mm/min; the *MAPE* is 6.182%; and the R^2 is 0.7213. The method can effectively predict the TBM advance rate and assist the driver in adjusting the tunneling parameters of the TBM.

2) The TCN model outperforms the RNN and LSTM models in predicting the TBM advance rate. It is found that the *MAPE* of the three models are all less than 7%, and they all show good generalization. However, the TCN prediction model with a *MAPE* value of 6.182% has a higher accuracy in predicting the advance rate compared with the RNN and LSTM models with a *MAPE* of 6.808% and 6.864%, respectively. It illustrates that TCN model is more suitable for the prediction of TBM advance rate considering its time-series characteristics.

3) The penetration rate and cutterhead torque of the current moment have significant influence on the TBM advance rate at the next moment. The influence of the cutterhead rotation and total thrust on the TBM advance rate is moderate.

The TCN-based model for real-time prediction of the TBM advance rate in this paper is of great potential with significance in guiding the TBM operator in adjusting the TBM tunnel construction control. However, the applicability of the model in varying geological condition is still under investigation with more data and validation

results going forward. In future, with collection of real-time TBM interaction parameters, the TCN model will be improved and incorporated into the TBM operating system for real-time modeling and control of TBMs.

Data Availability Statement Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request, including the model code developed by the authors, and the TBM data collected.

Acknowledgements The authors acknowledge the support of intelligent control and support software to safely and efficiently operate TBM tunnels from China Railway Engineering Equipment Group Co., Ltd. and the project team for the National Basic Research Program (973 program). Supports from National Natural Science Foundation of China (Grant No. 11902069), Sichuan University, State Key Lab Hydraul & Mt River Engn (No. SKHL1915), and the Research Project of China Railway First Survey and Design Institute Group Co., Ltd (No. 19-15 and No. 20-17-1) are also acknowledged. The work is partially supported by the 111 Project (B17009) and under the framework of Sino-Franco Joint Research Laboratory on Multiphysics and Multiscale Rock Mechanics.

References

1. Qi M. Promoting TBM in tunnel construction in China. *Tunnel Construction*, 2014, 34(11): 1019–1023 (in Chinese)
2. Shang Y, Yang Z, Qi Z, Sun Y, Shi Y, Yuan G. Retrospective analysis of TBM accidents from its poor flexibility to complicated geological conditions. *Chinese Journal of Rock Mechanics and Engineering*, 2007, 26(12): 2404–2411 (in Chinese)
3. Li L, Liu Z, Zhou H, Qi W, Zha W. Weighted voting model for advanced intelligent perception of tunnel faults based on TBM rock-machine information. *Chinese Journal of Rock Mechanics and Engineering*, 2020, 39(S2): 3401–3411 (in Chinese)
4. Liu Q, Liu J, Pan Y, Kong X. Research advances of tunnel boring machine performance prediction models for hard rock. *Chinese Journal of Rock Mechanics and Engineering*, 2016, 35(S1): 2766–2786 (in Chinese)
5. Khademi F, Akbari M, Jamal S M, Nikoo M. Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete. *Frontiers of Structural and Civil Engineering*, 2017, 11(1): 90–99
6. Sharafati A, Naderpour H, Salih S Q, Onyari E, Yaseen Z M. Simulation of foamed concrete compressive strength prediction using adaptive neuro-fuzzy inference system optimized by nature-inspired algorithms. *Frontiers of Structural and Civil Engineering*, 2021, 15(1): 61–79
7. Chen R, Zhang P, Wu H, Wang Z, Zhong Z. Prediction of shield tunneling-induced ground settlement using machine learning techniques. *Frontiers of Structural and Civil Engineering*, 2019, 13(6): 1363–1378
8. Liu Z, Shao J, Xu W, Wu Q. Indirect estimation of unconfined compressive strength of carbonate rocks using extreme learning machine. *Acta Geotechnica*, 2015, 10(5): 651–663
9. Lin S, Zheng H, Han C, Han B, Li W. Evaluation and prediction of slope stability using machine learning approaches. *Frontiers of Structural and Civil Engineering*, 2021, 15(4): 821–833
10. Liu Z, Shao J, Xu W, Chen H, Zhang Y. An extreme learning machine approach for slope stability evaluation and prediction. *Natural Hazards*, 2014, 73(2): 787–804
11. Gordan B, Jahed Armaghani D, Hajihassani M, Monjezi M. Prediction of seismic slope stability through combination of particle swarm optimization and neural network. *Engineering with Computers*, 2016, 32(1): 85–97
12. Liu Z, Shao J, Xu W, Chen H, Shi C. Comparison on landslide nonlinear displacement analysis and prediction with computational intelligence approaches. *Landslides*, 2014, 11(5): 889–896
13. Hoang N D, Tien Bui D. A novel relevance vector machine classifier with cuckoo search optimization for spatial prediction of landslides. *Journal of Computing in Civil Engineering*, 2016, 30(5): 04016001
14. Tabarsa A, Latifi N, Osouli A, Bagheri Y. Unconfined compressive strength prediction of soils stabilized using artificial neural networks and support vector machines. *Frontiers of Structural and Civil Engineering*, 2021, 15(2): 520–536
15. Jing L, Li J, Zhang N, Chen S, Yang C, Cao H. A TBM advance rate prediction method considering the effects of operating factors. *Tunnelling and Underground Space Technology*, 2021, 107: 103620
16. Hassanpour J, Rostami J, Khamehchiyan M, Bruland A. Developing new equations for TBM performance prediction in carbonate-argillaceous rocks: A case history of Nowsood water conveyance tunnel. *Geomechanics and Geoengineering*, 2009, 4(4): 287–297
17. Samaei M, Ranjbarnia M, Nourani V, Zare Naghadehi M. Performance prediction of tunnel boring machine through developing high accuracy equations: A case study in adverse geological condition. *Measurement*, 2020, 152: 107244
18. Yagiz S, Karahan H. Prediction of hard rock TBM penetration rate using particle swarm optimization. *International Journal of Rock Mechanics and Mining Sciences*, 2011, 48(3): 427–433
19. Fattahi H, Babanouri N. Applying optimized support vector regression models for prediction of tunnel boring machine performance. *Geotechnical and Geological Engineering*, 2017, 35(5): 2205–2217
20. Mahdevari S, Shahriar K, Yagiz S, Akbarpour Shirazi M. A support vector regression model for predicting tunnel boring machine penetration rates. *International Journal of Rock Mechanics and Mining Sciences*, 2014, 72: 214–229
21. Ghasemi E, Yagiz S, Ataei M. Predicting penetration rate of hard rock tunnel boring machine using fuzzy logic. *Bulletin of Engineering Geology and the Environment*, 2014, 73(1): 23–35
22. Armaghani D J, Yagiz S, Mohamad E T, Zhou J. Prediction of TBM performance in fresh through weathered granite using empirical and statistical approaches. *Tunnelling and Underground Space Technology*, 2021, 118: 104183
23. Yang H, Wang Z, Song K. A new hybrid grey wolf optimizer-feature weighted-multiple kernel-support vector regression technique to predict TBM performance. *Engineering with Computers*, 2020: 1–17
24. Jahed Armaghani D, Faradonbeh R S, Momeni E, Fahimifar A,

- Tahir M. Performance prediction of tunnel boring machine through developing a gene expression programming equation. *Engineering with Computers*, 2018, 34(1): 129–141
25. Zare Naghadehi M, Samaei M, Ranjbarnia M, Nourani V. State-of-the-art predictive modeling of TBM performance in changing geological conditions through gene expression programming. *Measurement*, 2018, 126: 46–57
 26. Moradi M R, Farsangi M A E. Application of the risk matrix method for geotechnical risk analysis and prediction of the advance rate in rock TBM tunneling. *Rock Mechanics and Rock Engineering*, 2014, 47(5): 1951–1960
 27. Zhou J, Qiu Y, Zhu S, Armaghani D J, Li C, Nguyen H, Yagiz S. Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate. *Engineering Applications of Artificial Intelligence*, 2021, 97: 104015
 28. Gao X, Shi M, Song X, Zhang C, Zhang H. Recurrent neural networks for real-time prediction of TBM operating parameters. *Automation in Construction*, 2019, 98: 225–235
 29. Benardos A G, Kaliampakos D C. Modelling TBM performance with artificial neural networks. *Tunnelling and Underground Space Technology*, 2004, 19(6): 597–605
 30. Zhou J, Bejarbaneh B Y, Armaghani D J, Tahir M M. Forecasting of TBM advance rate in hard rock condition based on artificial neural network and genetic programming techniques. *Bulletin of Engineering Geology and the Environment*, 2019, 79(2): 2069–2084
 31. Armaghani D J, Mohamad E T, Narayanasamy M S, Narita N, Yagiz S. Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition. *Tunnelling and Underground Space Technology*, 2017, 63: 29–43
 32. Koopialipour M, Fahimifar A, Ghaleini E N, Momenzadeh M, Armaghani D J. Development of a new hybrid ANN for solving a geotechnical problem related to tunnel boring machine performance. *Engineering with Computers*, 2020, 36(1): 345–357
 33. Armaghani D J, Koopialipour M, Marto A, Yagiz S. Application of several optimization techniques for estimating TBM advance rate in granitic rocks. *Journal of Rock Mechanics and Geotechnical Engineering*, 2019, 11(4): 779–789
 34. Zeng J, Roy B, Kumar D, Mohammed A S, Armaghani D J, Zhou J, Mohamad E T. Proposing several hybrid PSO-extreme learning machine techniques to predict TBM performance. *Engineering with Computers*, 2021: 1–17
 35. Zhou J, Qiu Y, Zhu S, Armaghani D J, Mohamad E T. Estimation of the TBM advance rate under hard rock conditions using XGBoost and Bayesian optimization. *Underground Space*, 2021, 6(5): 506–515
 36. Rumelhart D E, Hinton G E, Williams R J. Learning representations by back propagating errors. *Nature*, 1986, 323(6088): 533–536
 37. Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 1994, 5(2): 157–166
 38. Chen H, Xiao C, Yao Z, Jiang H, Guan Y. Prediction of TBM tunneling parameters through an LSTM neural network. In: 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO). Dali: IEEE, 2019: 702–707
 39. Liu Z, Li L, Fang X, Qi W, Shen J, Zhou H, Zhang Y. Hard-rock tunnel lithology prediction with TBM construction big data using a global-attention-mechanism-based LSTM network. *Automation in Construction*, 2021, 125: 103647
 40. Qin S, Xu T, Zhou W. Predicting pore-water pressure in front of a TBM using a deep learning approach. *International Journal of Geomechanics*, 2021, 21(8): 04021140
 41. Bai S, Kolter J Z, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. 2018, arXiv:1803.01271
 42. Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, Kalchbrenner N, Senior A, Kavukcuoglu K. Wavenet: A generative model for raw audio. 2016, arXiv:1609.03499
 43. Yu F, Koltun V. Multi-scale context aggregation by dilated convolution. 2015, arXiv: 1511.07122
 44. Zhou F, Jin L, Dong J. Review of convolutional neural network. *Chinese Journal of Computers*, 2017, 40(6): 1229–1251 (in Chinese)
 45. Strubell E, Verga P, Belanger D, McCallum A. Fast and accurate entity recognition with iterated dilated convolutions. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Honolulu, HI: Association for Computational Linguistics, 2017: 2670–2680
 46. Raschka S. *Python Machine Learning*. Birmingham: Packt Publishing Ltd., 2015
 47. Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th International Conference on Machine Learning*. Haifa: Omnipress, 2010: 807–814
 48. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: *Proceedings of the 32nd International Conference on Machine Learning*. Lille: PMLR, 2015: 448–456
 49. Kingma D, Ba J. Adam: A method for stochastic optimization. 2014, arXiv:1412.6980
 50. Jamieson K, Talwalkar A. Non-stochastic best arm identification and hyperparameter optimization. In: *The 19th International Conference on Artificial Intelligence and Statistics*. Cadiz: PMLR, 2016: 240–248
 51. Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks. In: 2013 IEEE international conference on acoustics, speech and signal processing. Las Vegas, NV: IEEE, 2013: 6645–6649
 52. Baldi P, Sadowski P. The dropout learning algorithm. *Artificial Intelligence*, 2014, 210: 78–122