Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering) 2025 26(3):226-237
www.jzus.zju.edu.cn; www.springer.com/journal/11582
E-mail: jzus_a@zju.edu.cn

Research Article

https://doi.org/10.1631/jzus.A2400012

Predicting tunnel boring machine performance with the Informer model: a case study of the Guangzhou Metro Line project

Junxing ZHAO[1], Xiaobin DING[1,2]✉

[1]*School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China*
[2]*Guangdong Provincial Key Laboratory of Modern Civil Engineering Technology, South China University of Technology, Guangzhou 511442, China*

Abstract: Accurately forecasting the operational performance of a tunnel boring machine (TBM) in advance is useful for making timely adjustments to boring parameters, thereby enhancing overall boring efficiency. In this study, we used the Informer model to predict a critical performance parameter of the TBM, namely thrust. Leveraging data from the Guangzhou Metro Line 22 project on the big data platform in China, the model's performance was validated, while data from Line 18 were used to assess its generalization capability. Results revealed that the Informer model surpasses random forest (RF), extreme gradient boosting (XGB), support vector regression (SVR), k-nearest neighbors (KNN), back propagation (BP), and long short-term memory (LSTM) models in both prediction accuracy and generalization performance. In addition, the optimal input lengths for maximizing accuracy in the single-time-step output model are within the range of 8–24, while for the multiple-time-step output model, the optimal input length is 8. Furthermore, the last predicted value in the case of multiple-time-step outputs showed the highest accuracy. It was also found that relaxation of the Pearson analysis method metrics to 0.95 improved the performance of the model. Finally, the prediction results were most affected by earth pressure, rotation speed, torque, boring speed, and the surrounding rock grade. The model can provide useful guidance for constructors when adjusting TBM operation parameters.

Key words: Boring machine performance; Informer model; Deep learning; Thrust force

1 Introduction

With rapid urban development, metro construction has become increasingly popular (Chen et al., 2022). In the process of metro construction, tunnel boring machines (TBMs) have been widely used because of their advantages of good safety and high efficiency (Zhang et al., 2015; Liu et al., 2016; Zheng et al., 2016). The cost of metro construction depends largely on the use time of TBMs, and reasonable control of the construction period is also directly related to cost control. Without an accurate understanding of the TBM's boring performance, it is difficult to make an accurate estimate of project completion time, which may lead to unreasonable scheduling of the construction period and TBM boring arrangements, thereby increasing the project cost. Therefore, to keep the cost and duration of a project within reasonable limits, it is necessary to predict the boring performance of the TBM before excavation (Rostami, 2016; Zhou et al., 2020; Xu et al., 2023).

In the past, scholars have explored the boring performance of TBMs mainly by conducting tests. The first model to predict boring performance was the Colorado School of Mines model (Rostami, 1997). It is an idealized empirical model that predicts boring performance based on rock properties and an idealized breaking mechanism. The model has given many scholars a method to predict boring performance. Most similar empirical models come from the analysis of parameters in real engineering (Yagiz, 2008; Hassanpour, 2018; Pan et al., 2020; Bilgin and Yüksel, 2023). Empirical models that consider field or experimental conditions have an important reference value. However, almost all TBMs operate in complex geological environments. Most empirical models are applicable only

✉ Xiaobin DING, dingxb@scut.edu.cn
Junxing ZHAO, https://orcid.org/0009-0009-8230-0165
Xiaobin DING, https://orcid.org/0000-0002-6168-4819

Received Jan. 5, 2024; Revision accepted Apr. 19, 2024;
Crosschecked Mar. 1, 2025

to specific situations and lack generalizability, restricting their widespread application.

Neural networks have a strong ability to extract features from information, and fit nonlinear relationships better than empirical models. They are widely used in civil engineering (Asteris et al., 2016, 2019, 2021a, 2021b; Psyllaki et al., 2018; Hajihassani et al., 2019; Li ZM et al., 2021; Emad et al., 2022). Lu and Shi (2023) combined the kernel extreme learning machine, variable modal decomposition, and the Levy-hunter-prey optimizer algorithm model to predict the boring speed of a TBM and found that the combined model outperformed the whale algorithm and the unimproved hunter-prey seeking algorithm. Shi et al. (2022) used a genetic algorithm to optimize a back propagation (BP) neural network to predict the tunneling speed of a shield machine.

The dynamic adaptation of TBM operational parameters in response to evolving geological conditions can be conceptualized as a time series problem in machine learning. A BP neural network works by forecasting output parameters based on prevailing input parameters. However, there are inherent constraints in its capacity to anticipate forthcoming operational parameters. Timely projections for future TBM operational parameters are necessary, as they enable proactive adjustments by operators, ensuring the TBM is maintained at an optimal operational performance level (Fig. S1 of the electronic supplementary materials (ESM)).

As a result, the recurrent neural network (RNN) model for input sequences with continuity is popular in predicting TBM boring performance. RNN-like models consider the back-and-forth relationship of the data (i.e., they consider the temporal nature of the data), and the effects of previous inputs are considered when processing the current input. Mahmoodzadeh et al. (2022) proposed a gray wolf-optimized long short-term memory (LSTM) model for TBM penetration rate prediction. It was found to be more accurate in its prediction than the LSTM model. Gao et al. (2021) obtained an accurate prediction of TBM penetration rate using an LSTM model. Li JH et al. (2021) used an LSTM neural network model to predict the cutterhead torque and total thrust.

The above studies showed the superiority of time series modeling in predicting TBM boring performance. However, the RNN model has disadvantages, such as its inability to deal with very long time series, and the problems of gradient vanishing and gradient explosion (Bengio et al., 1994). The LSTM model, as a variant of the RNN model, can effectively inhibit the gradient vanishing problem that occurs in the RNN model. However, if the time span of the TBM input data is very long, LSTM may still face the problem of gradient vanishing or gradient explosion, and has a disadvantage in parallel processing.

The Informer model is an improvement of the transformer model. It has the advantages of the transformer model and alleviates its shortcomings (Zhou et al., 2021). For example, the Informer model combines a global self-attention mechanism and a local attention mechanism to capture both global and local informations, while the traditional transformer model relies more on global information and less on local information when processing sequence data. The Informer model is processed by time-step chunking, which splits the long sequences into multiple subsequences, thereby reducing the memory requirement of the model. The problems of excessive memory consumption and computational complexity of the transformer model in dealing with long sequences are solved in the Informer model.

Table S1 of the ESM lists some of the studies that have contributed to significant progress in this field. However, a research gap exists regarding the applicability of the Informer model to the domain of TBM performance prediction. Previous studies emphasized the analysis of TBM performance prediction results solely within the confines of the test set, neglecting a comprehensive exploration of the generalization capabilities across numerous models.

Given these challenges, the primary objective of this investigation was to test the efficacy of the Informer model in forecasting TBM performance. Additionally, a comparative analysis was carried out involving the construction of models such as random forest (RF), extreme gradient boosting (XGB), support vector regression (SVR), k-nearest neighbors (KNN), BP, and LSTM. The input parameters were earth pressure, rotation speed, penetration rate, torque, thrust force, and boring speed (TBM parameters), as well as rock uniaxial compressive strength, surrounding rock grade (SRG), and liquid limit (geological parameters), with thrust force characterizing the TBM boring performance as the output. The significance of the input parameters was checked by performing a sensitivity analysis of the high-precision model. A tunnel section

of the Guangzhou Metro Line 22 was used as a case study to demonstrate the feasibility of the model. A tunnel section of Line 18 of the Guangzhou Metro was used as a case study to demonstrate the generalization ability of the model. The specific workflow diagram is shown in Fig. 1.

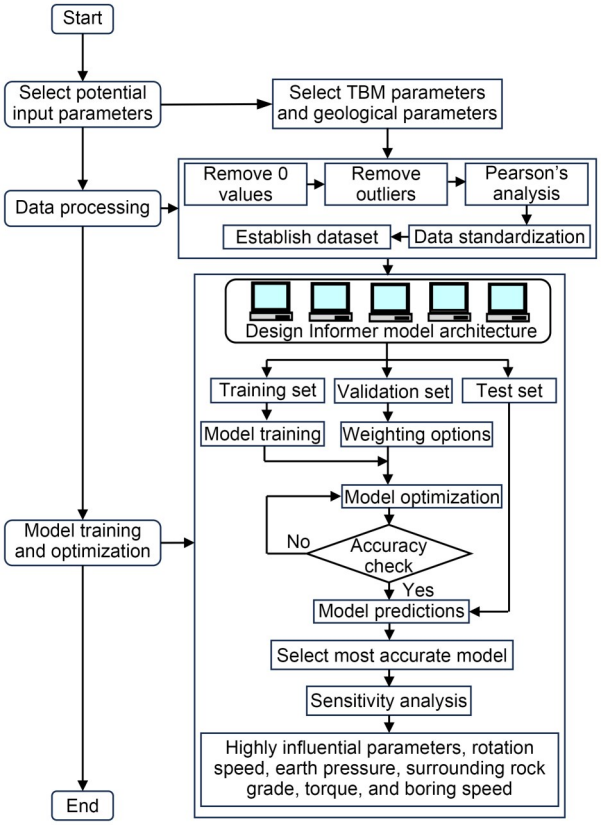


Fig. 1 Workflow diagram of this study

Compared to previous studies (Table S1), the innovations of this study were as follows:

1. For the first time, use the liquid-limit indicator of the formation as an input parameter.

2. For the first time, the Informer model was used to predict TBM performance, with a correlation coefficient of 0.99 using just over 30000 pieces of data. The trained model performed well in other projects.

3. Using only eight input parameters, the model performed better than previous models. The Informer model reduced the running time and improved the efficiency.

4. The Informer model performed better than the RF, XGB, SVR, KNN, BP, and LSTM models on different projects, proving that the model has excellent generalization ability.

2 Methodology

The Informer model proposed by Zhou et al. (2020) is an advanced time series forecasting model based on deep learning and a self-attention mechanism. It incorporates a self-attention mechanism inspired by the transformer model, enabling it to effectively capture long-term dependencies and spatio-temporal correlations within time series data.

The core idea behind the Informer model revolves around an encoder-decoder architecture (Fig. S2 of the ESM). The encoder-decoder architecture constitutes a pivotal component within the TBM-Informer model. The encoder plays a vital role in the conversion of the input sequence into an intermediate representation, typically a fixed-length vector, which encapsulates the semantic information inherent in the input sequence. In contrast, the decoder assumes the crucial responsibility of iteratively generating the output sequence, leveraging the intermediate representation crafted by the encoder as its foundational basis.

To account for the temporal aspect, the Informer model introduces a mechanism for encoding temporal features, embedding temporal information into the model. This allows the model to learn time-dependent patterns and trends, improving its accuracy in predicting future time series values.

3 Dataset establishment

3.1 Case description

The data used in the model came from the Guangzhou Metro Big Data Platform. The project that generated this data, Guangzhou Metro Line 22, is located in Panyu, Guangzhou, China. The specific route is shown in Fig. S4. Granite is the most dominant rock type in the tunnel. The length of the interval is 4987.8 m, the radius of the line plane is 900, 1100, 1120, and 2200 m, the buried body of the tunnel top is 23.4–37.0 m, and the line spacing is 7.5–10.3 m.

3.2 Data preprocessing

The data used in the model came from the shield construction monitoring system for new line construction of rail transit and had 57006 entries. Data were collected at one-minute intervals, and the collected data parameters included earth pressure and thrust. Many

useless values were included in the collected parameters, so the data had to be processed according to the TBM boring characteristics. In this section, we introduce the data processing methods.

3.2.1 TBM boring section data extraction

A TBM does not remain in a state of excavation (Fig. 2). Intervals where rotation speed (n_R), torque (M), thrust force (F_T), and boring speed (v_B) record zero values signify non-excavation states. The parameters collected encompass various operational states of the TBM, such as assembly (i.e., the installation of the cutter to the disc), advancement (TBM excavation), and cessation (TBM entering a pause state after advancing a specified distance). Consequently, significant variations are observed in metrics such as torque, rotation speed, and thrust force, partly attributable to these distinct operational phases.

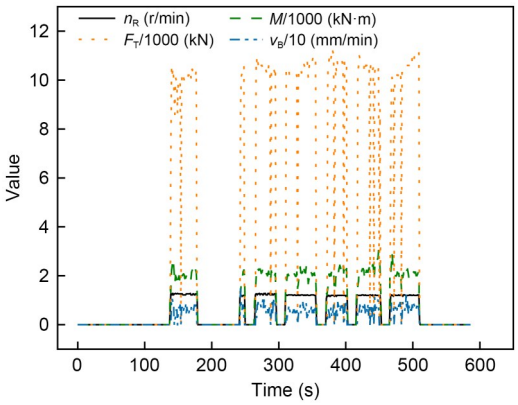


Fig. 2 TBM parameter raw data

Based on these variable operational statuses, an initial step involves the exclusion of data corresponding to the assembly and cessation phases. Subsequently, data collected during the initial 1 or 2 min of the excavation section, during which the TBM remains in an assembled state due to operational response delays in certain excavation sections, are also identified and removed.

We define a state judgment formula to determine whether the TBM is working or not by using the boring speed, rotation speed, thrust force, and torque. The specific formula is expressed as follows:

$$D(X)=f(v_B)\cdot f(n_R)\cdot f(M)\cdot f(F_T),$$
$$f(x)=\begin{cases}0, & x=0,\\ 1, & x\neq 0,\end{cases} \tag{1}$$

$$D(X)=\begin{cases}0, & \text{remove data},\\ 1, & \text{retentioned data},\end{cases} \tag{2}$$

where $D(X)$ represents the formula for determining whether the data need to be removed. If it is 0, then the data need to be removed. If it is 1, then the data do not need to be removed. X represents the value of different data.

3.2.2 TBM boring section data extraction

Fig. 3 illustrates the geological profile of the TBM tunneling. Different colors distinguish soil and rock layers, which are labeled with specific layer codes. Table 1 provides the corresponding names and SRGs for these layers. The tunnel excavation area is denoted in Fig. 3 by a series of purple rectangles, each symbolizing a ring in the TBM excavation process. In the case of a metro traversing through composite strata, their values are computed through a weighted average method.

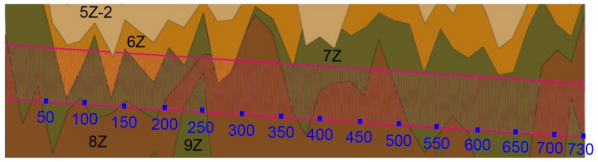


Fig. 3 Geological profile of TBM boring. References to color refer to the online version of this figure

Table 1 Layer code explanation

Code	Name	SRG
5Z-2	Hard-plasticized mixed granite residual soil	V
6Z	Mixed fully weathered granite	V
7Z	Mixed strong weathered granite	V–IV
8Z	Mixed medium weathered granite	IV
9Z	Mixed slightly weathered granite	II–III

3.2.3 TBM boring section data extraction

Input parameters characterized by strong correlations among their features may result in overfitting or model instability, thereby diminishing the model's generalization capability (Xue et al., 2023). Consequently, it is imperative to filter the preprocessed input parameters. Given the numerical nature of the variables, we used the Pearson correlation coefficient (Benesty et al., 2009) to evaluate the associations among the input parameters. Fig. 4 shows the correlation coefficients derived via the Pearson correlation coefficient method.

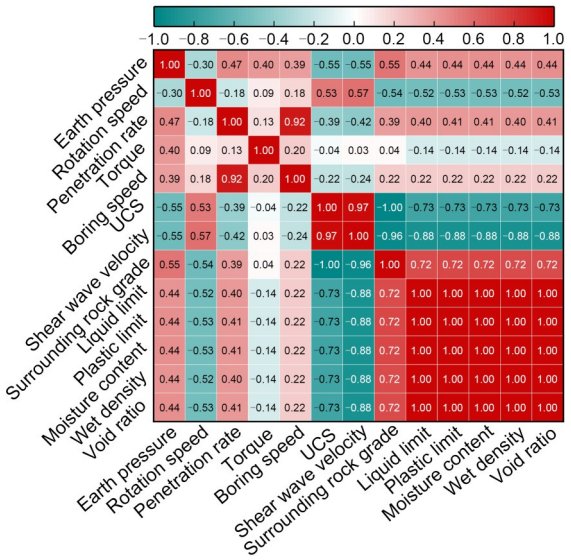


Fig. 4 Heat map of the Pearson correlation coefficient matrix. UCS represents the uniaxial compressive strength. References to color refer to the online version of this figure

The findings reveal a notable linear correlation (exceeding 0.8) between penetration and boring speed, shear wave velocity, and uniaxial compressive strength, as well as liquid limit, plastic limit, moisture content, wet density, and pore ratio. Therefore, this segment of the data needed to be excluded.

Data were acquired through the TBM's integrated sensors. Anomalies in the data can arise due to equipment or mechanical malfunctions. The presence of outliers necessitates their removal, as they do not faithfully represent the authentic state of TBM excavation. In this study, parameter outliers were identified using a boxplot method (Carter et al., 2009). Fig. 5 shows the boxplot derived from the computation of TBM data parameters, including earth pressure, rotation speed, penetration rate, torque, thrust force, and boring speed.

A boxplot is a graphical tool used for the depiction of data distributions and the identification of outliers. This technique uses quartiles and the interquartile range (R_{IQ}) as fundamental calculation metrics. Quartiles represent the four values used for partitioning a dataset. The first quartile (Q_1) designates the lower 25% segment of the sorted dataset, while the second quartile (Q_2) corresponds to the midpoint or the 50% position in the sorted dataset. The third quartile (Q_3) represents the upper 25% portion of the sorted data. The interquartile range is defined as the difference between Q_3 and Q_1, signifying the range encompassing the central 50% of the dataset.

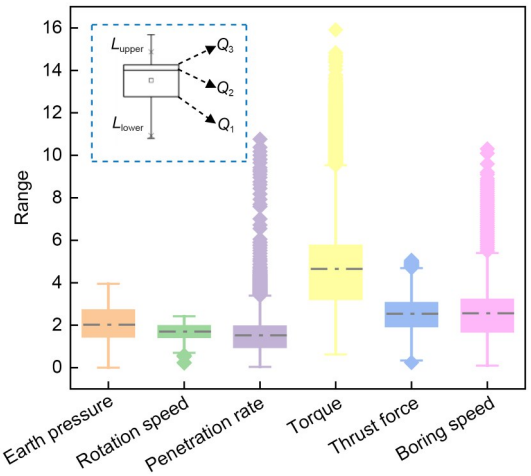


Fig. 5 Boxplot method to detect outliers in data

In accordance with Xue et al. (2023), the following formula was used to characterize outliers:

$$\begin{cases} L_{\text{upper}} = Q_3 + 1.5R_{\text{IQ}}, \\ L_{\text{lower}} = Q_1 - 1.5R_{\text{IQ}}, \end{cases} \tag{3}$$

where L_{upper} is the upper bound, and L_{lower} is the lower bound. Therefore, data points that exceed the range between L_{upper} and L_{lower} can be determined to be outliers and need to be eliminated.

Throughout the operational phase of TBM, real-time data are acquired via sensors. However, owing to the intricate engineering environment and equipment operational characteristics, the collected data often manifest various disturbances and noise, and even inadequate data collection may occur. The objective of the predictions in this study pertains to TBM performance under normal operating conditions. To address this challenge, we removed data with large intervals between boring rings and processed the data using wavelet denoising signal processing techniques. Wavelet denoising entails decomposing the original signal into low-frequency and high-frequency components through the extraction of wavelet coefficients at different scales. The low-frequency signal encapsulates the overall progress of the tunneling project, while the high-frequency signal captures signal noise. The denoised signal is derived through wavelet reconstruction of signals across diverse scales. The process of wavelet denoising unfolds as follows:

Step 1: Signal processing using wavelet transform for data contaminated by noise.

Step 2: The wavelet coefficients obtained from the transformation are subjected to processing aimed at noise removal.

Step 3: Wavelet reconstruction is undertaken across different scales to obtain the denoised signal.

Fig. 6 shows a graph comparing the thrust forces both pre- and post-denoising. The application of wavelet denoising diminishes the noise inherent in the initial thrust force, resulting in a more refined representation of the settling data and enhanced reflection of deformation characteristics. The identified model input parameters are presented in Table 2.

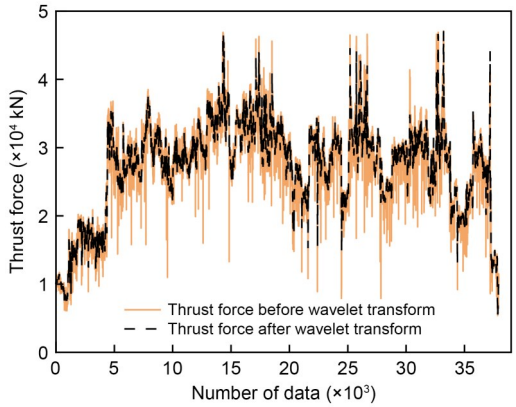


Fig. 6 Diagram of the thrust force denoising effect

4 Results analysis

Fig. S7 presents a visual contrast between the predicted thrust values and the actual thrust values, accompanied by the associated model evaluation metrics. The findings underscore the commendable thrust prediction performance achieved by the Informer model. The model shows a correlation coefficient of 0.99819 and a minimal error value of 0.001576, attesting to its good predictive capabilities. Note that, before using wavelet analysis, we obtained data with a mean square error (MSE) of 0.127341 and R^2 of 0.94373, while after using wavelet analysis, we obtained data with an MSE of 0.001576 and R^2 of 0.99819. The enhancement effect of wavelet analysis is very obvious.

The data's volatility is indicative of the TBM operating within diverse geological formations. Sudden fluctuations in data values correspond to variations in parameters and geological formations, underscoring the TBM's autonomous capability to adjust the thrust to accommodate different geological conditions. Predicting fluctuations in data poses a challenge for numerous models, including the Informer model examined in this study. Nevertheless, the model shows a robust fit in the initial stages. This observation reinforces the model's capacity to adapt to parameter and stratigraphic variations, demonstrating its resilience in handling extreme data points. Evidently, the Informer model adeptly captures the temporal dynamics of TBM thrust.

In addition, the performances of machine learning methods such as the RF, XGB, SVR, KNN, BP, and LSTM as TBM performance prediction models were compared with that of the Informer model (Table 3). Notably, the Informer model showed superior performance across the dataset. Among traditional machine learning models, the RF model showed optimal performance, whereas SVR models showed inferior outcomes.

Note that deep learning models have the capacity to use preceding thrust force as inputs for predicting subsequent outputs. This attribute contributes to the effective performance observed in LSTM and Informer models on the dataset. This observation underscores the strong guidance provided by deep time-series neural network models (such as LSTM and Informer models) in the prompt adaptation of TBM operational parameters.

Table 2 Finalize model input parameters

Parameter type	Parameter	Data type	Max value	Mean value	Min value
TBM	Earth pressure (N/mm)	Time-varying	3.83	2.31	0
	Rotation speed, n_R (r/min)	Time-varying	2.20	1.59	0.70
	Torque, M (kN·m)	Time-varying	9550	4727	660
	Boring speed, v_B (mm/min)	Time-varying	54	25	1
	Thrust force, F_T (kN)	Time-varying	46965.66	27311.46	5352.99
Geological	Uniaxial compressive strength (MPa)	Time-invariant	62.42	16.91	0.76
	SRG	Time-invariant	4.95	4.28	2.86
	Liquid limit	Time-invariant	34.31	17.87	0.00

5 Discussion

5.1 Performance of new project

In the context of machine learning algorithms, the evaluation of performance often hinges on the crucial metric of generalization ability. Models characterized by strong generalization capabilities are likely to extend well to diverse scenarios. Consequently, one of the aims of this research was to derive a generalized model tailored for the prediction of operating parameters of TBMs, transcending the constraints of singular project applications.

With the overarching objective in focus, in this section, we introduce an additional tunneling project to assess the generalization capacity of the model established in Section 4. Situated in the Panyu District of Guangzhou City, China, this project pertains to a segment of the tunnel associated with the Guangzhou Metro Line 18. The specific route is shown in Fig. S8.

The dataset used in the model originates from the shield construction monitoring system used in the construction of a new rail transit line. Comprising 35076 entries collected at one-minute intervals, the dataset encompasses parameters such as earth pressure and thrust, and aligns with that of Guangzhou Metro Line 22. Subsequent to processing, the data were integrated into the model established in this section through the training process.

Fig. S9 illustrates the model's performance in this new project, wherein the R^2 attained a value of 0.99843 and the MSE reached 0.001575, with a slight improvement of 0.024% in prediction accuracy. Notably, the superior performance of the trained model was shown to extend to diverse project contexts. Furthermore, the pronounced fluctuation in thrust force within the new project posed challenges for the model's predictive capabilities. However, the Informer model showed notable proficiency in learning from extreme values, underscoring its robust generalization capability and applicability across varied projects.

To assess the Informer model's generalization capability, trained RF, XGB, SVR, KNN, BP, and LSTM models were compared with the Informer model in predicting the performance of the new TBM project. The outcomes of these predictions are presented in Table 4.

Table 3 Comparison of the performances of the Informer model and other models on the test set

Model	Model parameter setting	MSE	R^2	a
RF	$N_{estimator}$=300, d_{max}=20	0.037267	0.96256	0.7155
XGB	$N_{estimator}$=400, d_{max}=7	0.062253	0.93746	0.6072
SVR	C=0.1, ε=0.1, Kernel: RBF	0.179173	0.82001	0.5320
KNN	$N_{neighbor}$=7	0.056542	0.94320	0.6666
BP	R_L=0.0001, N_{hidden}=5, Optimizer: Adam	0.341733	0.88269	0.5486
LSTM	R_L=0.0001, N_{hidden}=100, Optimizer: Adam	0.006369	0.99339	0.8533
Informer	$N_{encoder}$=5, $N_{decoder}$=1, Attention: Full	0.001576	0.99819	1.0000

$N_{estimator}$ is the number of decision trees; d_{max} is the maximum depth of each decision tree; C is the penalty parameter; ε is the tolerance to error; Kernel represents the kernel function used to map the data in a high dimensional space; RBF represents the radial basis function; $N_{neighbor}$ is the number of nearest neighbors; R_L is the learning rate; N_{hidden} is the number of neurons in the hidden layer; Optimizer represents the type of optimizer; $N_{encoder}$ is the number of encoders; $N_{decoder}$ is the number of decoders; Attention represents the type of attention mechanism; a is the accuracy

Table 4 Performances of the Informer model and other models on the new TBM project

Model	Model parameter setting	MSE	R^2	Rate of change in model accuracy	a
RF	$N_{estimator}$=200, d_{max}=20	0.393617	0.60638	−37.0%	0.6743
XGB	$N_{estimator}$=400, d_{max}=7	0.366567	0.63343	−32.4%	0.6841
SVR	C=100, ε=0.1, Kernel: RBF	0.258279	0.74172	−9.5%	0.6954
KNN	$N_{neighbor}$=9	0.621751	0.37825	−59.9%	0.6714
BP	R_L=0.0001, N_{hidden}=5, Optimizer: Adam	0.540540	0.70782	−19.8%	0.7067
LSTM	R_L=0.0001, N_{hidden}=100, Optimizer: Adam	0.110627	0.67924	−31.6%	0.8250
Informer	$N_{encoder}$=5, $N_{decoder}$=1, Attention: Full	0.001575	0.99843	0.024%	1.0000

Remarkably, most models showed suboptimal performance in novel instances. The KNN model showed the most significant disparity, featuring a prediction MSE of 0.621751, an R^2 of merely 0.38, and a 59.9% decrease in prediction accuracy. Conversely, the SVR model emerged as the top-performing traditional machine learning model, with an error of 0.258279, an R^2 of 0.74172, and only a 9.5% decrease in prediction accuracy.

The performance of the traditional models in the new project showed that most had limited generalization capacity in this particular context, rendering them less suitable for predicting TBM performance. Compared with other models, the Informer model showed a slight improvement of 0.024% in prediction accuracy. This suggests that the Informer model effectively learns the intricate relationships among data, leading to more precise predictions. This enhancement can be attributed to the capability of the Informer model in capturing temporal dependencies inherent in time series data. By leveraging the intrinsic features of time series data and using more efficient learning strategies, the Informer model has demonstrated superior performance in predictive tasks.

Evaluation of the efficacy of various models across diverse projects showed that RF, XGB, SVR, KNN, and BP models need substantial training data to develop models with robust generalization capabilities. Conversely, LSTM and Informer models can effectively leverage historical temporal patterns as inputs to predict subsequent moments. Such predictive models show outstanding performance when confronted with data exhibiting temporal correlations.

5.2 Effects of various inputs, labels, and predicted lengths

In Section S2.3 of the ESM, we delineated the methods used for predicting thrust forces by using data from the previous 8 min. This entails using the past 8 min of data for training, and the subsequent 1 min of data for supervised learning. The selection of the length of the input time assumes significant importance. An excessively short time length compels the model to emphasize partial information, whereas an unduly extended length diminishes the influence of earlier data on the model. To optimize the model's performance, time length was systematically varied within the range of 2 to 30 min, taking into consideration the distinctive characteristics of TBM boring durations. Subsequently, the model was retrained using the revised time length. The following sections discuss the effect of different time lengths (Table S5) on the model.

5.2.1 Results of single predicted length

In this section, we elucidate the effect of individual time-step outputs on the model's precision. Table 5 illustrates the various input lengths used in the model and presents the corresponding correlation coefficients achieved through model training.

Table 5 Single-time-step model length setting

Method	Input length	Label length	Pred length	MSE	R^2
Single-length output	4	1	1	0.00298	0.99657
	8	1	1	0.00158	0.99819
	16	1	1	0.00210	0.99759
	24	1	1	0.00247	0.99716
	30	1	1	0.00295	0.99660

As the input length increases, there is an enhancement in the model's precision, although the rate of change is not high. Optimal accuracy is observed within the input length range of 8–30. This trend is attributed to the model's ability to assimilate a larger volume of data as the input length increases, thereby positively influencing accuracy. However, an excessively long input length escalates the model's training complexity, subsequently increasing computational costs. In light of the precision shown by the model, we conclude that the ideal input length for a single-time-step output model falls within the range of 8–16. The model's R^2 achieved a value of 0.99 and the MSE value did not exceed 0.003.

5.2.2 Results of long predicted length

In this section, we assess the predictive proficiency of the Informer model across multiple time steps. Table 6 shows the temporal extent configurations of

Table 6 Multiple-time-step model length setting

Method	Input length	Label length	Pred length	MSE	R^2
Multi-length output	8	4	4	0.00842	0.99031
	16	4	4	0.00923	0.98939
	24	4	4	0.00946	0.98912
	30	4	4	0.01297	0.98509

the multi-step prediction model and the correlation coefficients derived from model training.

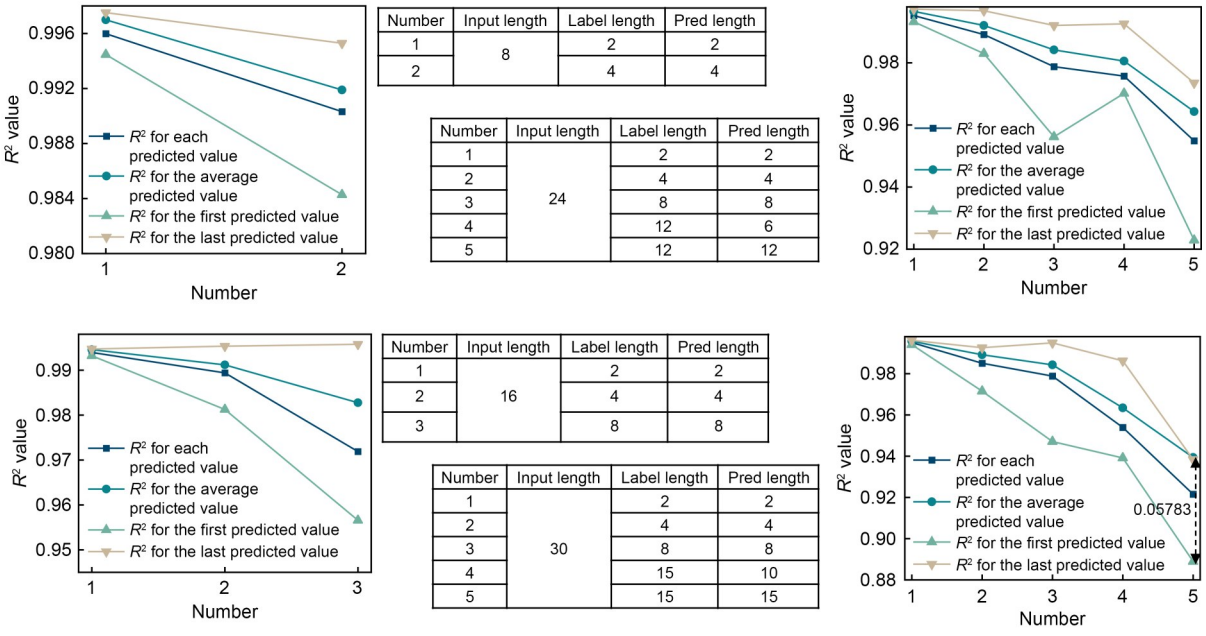
As the input length increases, the change in model accuracy becomes less prominent, ultimately stabilizing around 0.98. The increase in input length facilitates the assimilation of a greater amount of information by the model. However, this increase does not result in a notable improvement in accuracy and so does not justify the increase in training time. The input length of 8, as shown in this study, achieves an equilibrium between maintaining a reasonably elevated prediction accuracy and minimizing the duration of training.

5.2.3 Comparison of the accuracy of different predicted values

As illustrated in Fig. S10, within the model with a prediction length of 2, the thrust force receives a maximum of two predictions (i.e. y_2) and a minimum of one prediction (i.e. y_1). Consequently, the model's accuracy evaluation is contingent on a comprehensive comparison between all predicted values and the ground truth. However, it is important to acknowledge that diverse task specifications necessitate distinct criteria for prediction values. For certain task requisites, the ultimate prediction outcome may necessitate only the most precise value. Thus, it becomes imperative to investigate the model's accuracy under various output parameters. Within this study, metrics including the R^2 value, R^2 value of the average predicted values, R^2 value of the first predicted value, and R^2 of the final predicted value were selected for comparative analysis. The relevant parameter definitions are shown in Fig. S11.

Fig. 7 presents a juxtaposition of the R^2 across diverse length combinations as outlined in the table. Evidently, the initial prediction by the model shows the lowest R^2, while conversely, the final prediction shows the highest R^2. Moreover, the R^2 of the model demonstrates a more rapid decrease with an increase in prediction length. The smallest disparity observed among the various R^2 values is merely 0.01006, while the highest disparity can extend up to 0.05783. The TBM excavation process is inherently subject to environmental fluctuations and multiple contributing factors. Consequently, the parameters collected by the TBM show a higher correlation with data obtained from several temporally proximate time steps, both preceding and succeeding.

Based on this data characteristic, we make a reasonable speculation. The length of the prediction is too long, and the model needs to consider too much earlier data. Consequently, this proliferation of historical data dilutes the correlation with the present data, resulting in a less accurate initial prediction by the model. At the same time, with the continual sliding of the temporal window, the association between the prevailing



Number	Input length	Label length	Pred length
1	8	2	2
2		4	4

Number	Input length	Label length	Pred length
1	24	2	2
2		4	4
3		8	8
4		12	6
5		12	12

Number	Input length	Label length	Pred length
1	16	2	2
2		4	4
3		8	8

Number	Input length	Label length	Pred length
1	30	2	2
2		4	4
3		8	8
4		15	10
5		15	15

Fig. 7 Comparison of different predicted values of R^2

data and its historical counterparts intensifies progressively. This progression ensures a higher R^2 value for the ultimate prediction from the model.

5.3 Sensitivity analysis

Sensitivity analysis (Yang and Zhang, 1997) involves the assessment of the dependency between predicted values and input parameters. This analysis is instrumental in identifying the specific parameters that influence the thrust force of the TBM. The mathematical formulation of this process is as follows:

$$R_{ij}=\frac{\sum_{k=1}^{n}(x_{ik}x_{jk})}{\sum_{k=1}^{n}x_{ik}^2\sum_{k=1}^{n}x_{jk}^2}, \tag{4}$$

where x_i and x_j represent the input and output parameters, respectively, n represents the number of datasets, R_{ij} represents the correlation between the data, and the value of R_{ij} is between 0 and 1. The closer the value of R_{ij} is to 1, the stronger the correlation between the input and output parameters.

The results of the sensitivity analysis are shown in Fig. 8. Two baseline values, 0.95 and 0.77, are discernible. Notably, the input parameters affecting the prediction results, in descending order of influence, are rotation speed, earth pressure, SRG, torque, and boring speed. The parameters that can be adjusted by operators during operation include rotation speed and boring speed. These parameters show correlations with the output parameters exceeding 0.95, indicating a robust association between these variables. Conversely, the uniaxial compressive strength and liquid limit of the rock show weaker correlations with the output parameters. Consequently, we surmise that the main factors influencing TBM performance in hard strata are the TBM-related parameters. These findings can serve as valuable guidance for operators when adjusting TBM parameters.

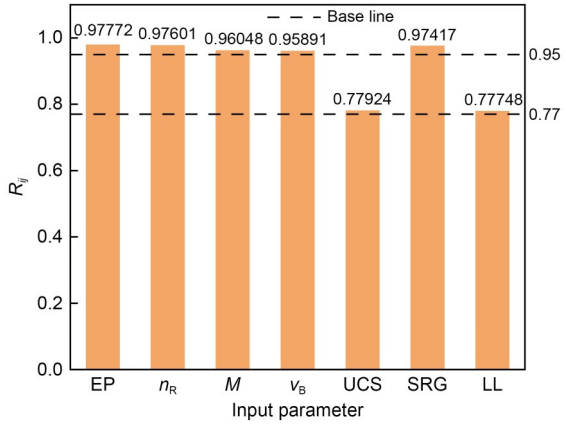


Fig. 8 Sensitivity analysis results. EP represents the earth pressure; LL represents the liquid limit

5.4 Limitations and future work

In this study, we achieved satisfactory experimental results, but there were also some limitations to our research. Firstly, our data size was relatively small, making it difficult for the trained models to be applicable in broader domains. Secondly, the numerical range of the data was limited (Table 4), posing a challenge to the performance of the models under extreme values, which requires further investigation. Lastly, there were certain difficulties in the practical application of the models. Currently, the assessment of TBM performance often cannot rely solely on big data analysis. We are working to overcome these challenges and endeavor to apply our research findings to practical engineering projects.

In the future, we will build upon the results of this study by training the models with more data to improve their generalization ability and by applying the models to practical engineering applications.

6 Conclusions

This research used the Informer model to forecast the performance of a TBM, based on data derived from specific intervals along Guangzhou Metro Lines 18 and 22. The proposed predictive model focuses mainly on elucidating the impact of TBM operational parameters and stratum properties on TBM performance. The model showed a prediction accuracy exceeding 0.99 for thrust force in two projects.

1. RF, XGB, SVR, KNN, BP, and LSTM models were used as comparison models. The models performed well on the test set of the original project, but poorly on the new project, and lacked some generalization ability.

2. Single-time-step and multiple-time-step outputs had distinct impacts on the model. Prudent consideration should be given when determining suitable temporal intervals for the input parameters. In this

investigation, an input length ranging from 8 to 24 proved suitable for single-time-step outputs, whereas an input length of 8 was deemed appropriate for multiple-time-step outputs. The model correlation coefficients reached 0.99 and 0.98, respectively.

3. Correlation coefficients were calculated for the model's predicted values, including the average, first, and last predicted values. Notably, the last predicted value showed a stronger correlation with the true value.

4. Given the limited number of input parameters in the model, the exclusion threshold in the Pearson analysis method was adjusted from 0.80 to 0.95. This adaptation resulted in the new model demonstrating superior performance on novel data compared to the previous model.

5. The sensitivity analysis revealed that the input parameters affecting the prediction results, in descending order of influence, were cutterhead rotation speed, earth pressure, SRG, and torque. These parameters had correlation coefficients exceeding 0.95. Conversely, the impact of rock uniaxial compressive strength and liquid limit on the predicted results was comparatively minor, about 0.77.

In our view, applying the obtained models to real-world engineering remains challenging within the scope of this research domain. Continuously refining models through practical application and achieving a high level of generalization poses a formidable task. Consequently, in future research endeavors, we advocate prioritizing the effective implementation of acquired models in practical engineering contexts. These findings can provide valuable guidance to operators in fine-tuning tunneling machine parameters.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 41827807) and the Foundation of the Guangdong Provincial Key Laboratory of Modern Civil Engineering Technology (No. 2021B1212040003), China.

Author contributions

Junxing ZHAO: conceptualization, methodology, data curation, writing–original draft, and writing–review & editing. Xiaobin DING: conceptualization, writing–review & editing, supervision, and funding acquisition.

Conflict of interest

Junxing ZHAO and Xiaobin DING declare that they have no conflict of interest.

References

Asteris PG, Tsaris AK, Cavaleri L, et al., 2016. Prediction of the fundamental period of infilled RC frame structures using artificial neural networks. *Computational Intelligence and Neuroscience*, 2016:5104907. https://doi.org/10.1155/2016/5104907

Asteris PG, Nozhati S, Nikoo M, et al., 2019. Krill herd algorithm-based neural network in structural seismic reliability evaluation. *Mechanics of Advanced Materials and Structures*, 26(13):1146-1153. https://doi.org/10.1080/15376494.2018.1430874

Asteris PG, Lemonis ME, Le TT, et al., 2021a. Evaluation of the ultimate eccentric load of rectangular CFSTs using advanced neural network modeling. *Engineering Structures*, 248:113297. https://doi.org/10.1016/j.engstruct.2021.113297

Asteris PG, Skentou AD, Bardhan A, et al., 2021b. Soft computing techniques for the prediction of concrete compressive strength using non-destructive tests. *Construction and Building Materials*, 303:124450. https://doi.org/10.1016/j.conbuildmat.2021.124450

Benesty J, Chen JD, Huang YT, et al., 2009. Pearson correlation coefficient. *In*: Cohen I, Huang YT, Chen JD, et al. (Eds.), Noise Reduction in Speech Processing. Springer, Berlin, Germany, p.1-4. https://doi.org/10.1007/978-3-642-00296-0_5

Bengio Y, Simard P, Frasconi P, 1994. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157-166. https://doi.org/10.1109/72.279181

Bilgin N, Yüksel A, 2023. The effect of EPB face pressure on TBM performance parameters in different geological formations of Istanbul. *Tunnelling and Underground Space Technology*, 138:105184. https://doi.org/10.1016/j.tust.2023.105184

Carter NJ, Schwertman NC, Kiser TL, 2009. A comparison of two boxplot methods for detecting univariate outliers which adjust for sample size and asymmetry. *Statistical Methodology*, 6(6):604-621. https://doi.org/10.1016/j.stamet.2009.07.001

Chen XS, Fu YB, Chen X, et al., 2022. Progress in underground space construction technology and technical challenges of digital intelligence. *China Journal of Highway and Transport*, 35(1):1-12 (in Chinese). https://doi.org/10.19721/j.cnki.1001-7372.2022.01.001

Emad W, Mohammed AS, Bras A, et al., 2022. Metamodel techniques to estimate the compressive strength of UHPFRC using various mix proportions and a high range of curing temperatures. *Construction and Building Materials*, 349:128737. https://doi.org/10.1016/j.conbuildmat.2022.128737

Gao BY, Wang RR, Lin CJ, et al., 2021. TBM penetration rate prediction based on the long short-term memory neural network. *Underground Space*, 6(6):718-731. https://doi.org/10.1016/j.undsp.2020.01.003

Hajihassani M, Abdullah SS, Asteris PG, et al., 2019. A gene expression programming model for predicting tunnel

convergence. *Applied Sciences*, 9(21):4650. https://doi.org/10.3390/app9214650

Hassanpour J, 2018. Development of an empirical model to estimate disc cutter wear for sedimentary and low to medium grade metamorphic rocks. *Tunnelling and Underground Space Technology*, 75:90-99. https://doi.org/10.1016/j.tust.2018.02.009

Li JH, Li PX, Guo D, et al., 2021. Advanced prediction of tunnel boring machine performance based on big data. *Geoscience Frontiers*, 12(1):331-338. https://doi.org/10.1016/j.gsf.2020.02.011

Li ZM, Yazdani Bejarbaneh B, Asteris PG, et al., 2021. A hybrid GEP and WOA approach to estimate the optimal penetration rate of TBM in granitic rock mass. *Soft Computing*, 25(17):11877-11895. https://doi.org/10.1007/s00500-021-06005-8

Liu QS, Huang X, Gong QM, et al., 2016. Application and development of hard rock TBM and its prospect in China. *Tunnelling and Underground Space Technology*, 57:33-46. https://doi.org/10.1016/j.tust.2016.01.034

Lu ZP, Shi KB, 2023. A novel VMD-LHPO-KELM machine learning-based TBM boring parameter prediction. *Earth Science Informatics*, 16(3):2925-2938. https://doi.org/10.1007/s12145-023-01043-2

Mahmoodzadeh A, Nejati HR, Mohammadi M, et al., 2022. Forecasting tunnel boring machine penetration rate using LSTM deep neural network optimized by grey wolf optimization algorithm. *Expert Systems with Applications*, 209:118303. https://doi.org/10.1016/j.eswa.2022.118303

Pan YC, Liu QS, Liu Q, et al., 2020. Full-scale linear cutting tests to check and modify a widely used semi-theoretical model for disc cutter cutting force prediction. *Acta Geotechnica*, 15(6):1481-1500. https://doi.org/10.1007/s11440-019-00852-4

Psyllaki P, Stamatiou K, Iliadis I, et al., 2018. Surface treatment of tool steels against galling failure. *MATEC Web of Conferences*, 188:04024. https://doi.org/10.1051/matecconf/201818804024

Rostami J, 1997. Development of a Force Estimation Model for Rock Fragmentation with Disc Cutters Through Theoretical Modeling and Physical Measurement of Crushed Zone Pressure. PhD Thesis, Colorado School of Mines, Colorado, USA.

Rostami J, 2016. Performance prediction of hard rock tunnel boring machines (TBMs) in difficult ground. *Tunnelling and Underground Space Technology*, 57:173-182. https://doi.org/10.1016/j.tust.2016.01.009

Shi QH, Song PF, Tan ZW, et al., 2022. GA-BP neural network prediction model for tunneling speed of shield machine with composite formation dual mode (TBM-EPB). Proceedings of the International Conference on Computational Infrastructure and Urban Planning, p.1-4. https://doi.org/10.1145/3546632.3546633

Xu QH, Huang X, Zhang BG, et al., 2023. TBM performance prediction using LSTM-based hybrid neural network model: case study of Baimang River tunnel project in Shenzhen, China. *Underground Space*, 11:130-152. https://doi.org/10.1016/j.undsp.2022.11.002

Xue YD, Luo W, Chen L, et al., 2023. An intelligent method for TBM surrounding rock classification based on time series segmentation of rock-machine interaction data. *Tunnelling and Underground Space Technology*, 140:105317. https://doi.org/10.1016/j.tust.2023.105317

Yagiz S, 2008. Utilizing rock mass properties for predicting TBM performance in hard rock condition. *Tunnelling and Underground Space Technology*, 23(3):326-339. https://doi.org/10.1016/j.tust.2007.04.011

Yang Y, Zhang Q, 1997. A hierarchical analysis for rock engineering using artificial neural networks. *Rock Mechanics and Rock Engineering*, 30(4):207-222. https://doi.org/10.1007/BF01045717

Zhang L, Gao JR, Zhang B, et al., 2015. Application status and prospects for the casting support tunnelling system using TBM. *Modern Tunnelling Technology*, 52(5):24-31 (in Chinese). https://doi.org/10.13807/j.cnki.mtt.2015.05.004

Zheng YL, Zhang QB, Zhao J, 2016. Challenges and opportunities of using tunnel boring machines in mining. *Tunnelling and Underground Space Technology*, 57:287-299. https://doi.org/10.1016/j.tust.2016.01.023

Zhou J, Yazdani Bejarbaneh B, Jahed Armaghani D, et al., 2020. 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*, 79(4):2069-2084. https://doi.org/10.1007/s10064-019-01626-8

Electronic supplementary materials

Sections S1–S4, Tables S1–S7, Figs. S1–S11