Electronic Supplementary Materials

For 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 Institute of Geotechnical Engineering, Guangzhou, China*

S1 Methodology

S1.1 Data Embedding Layer

The Informer model does not have an RNN-like recurrent architecture. To exploit the timing information, the first layer in the Informer model is chosen to be the positional embedding. The specific operation is shown in the formulas.

$$PE_{(pos,2i)} = \sin(pos / 10000^{\frac{2i}{d\_model}}) \tag{S1}$$

$$PE_{(pos,2i+1)} = \cos(pos / 10000^{\frac{2i}{d\_model}}) \tag{S2}$$

where pos represents the position (data order) and i = 1, 2, …, $d_{model}/2$ indicates the dimension.

In addition, the model introduces temporal embedding and token embedding. Temporal embedding is a method used to encode temporal information into vector representations. This inclusion enables the model to concurrently discern the correlations between temporal and other features, thereby gaining a deeper understanding of sequence data dynamics. Token embedding serves as a technique for encoding symbols within the input data as vector representations. Different symbols may have variable associations and levels of significance among them. By integrating symbol embeddings into the model, it can more effectively grasp and leverage these associations and significance levels.

Consequently, the data embedding component within Informer can be formulated as follows:

$$\begin{aligned}\text{DataEmbed} = {} & \text{PosEmbed(input)} + \\ & \text{TokenEmbed(input)} + \text{TemEmbed(input)}\end{aligned} \tag{S3}$$

S1.2 Data Embedding Layer

The architecture of the multi-attention mechanism comprises several self-attention layers (Fig. S3). Upon reaching the decoding layer, the query (Q), key (K), and value (V) undergo initial mapping via various linear transformations. These mapped values are then concatenated to create the inputs for the multi-head attention. Subsequently, additional linear transformations are applied to yield the final output. Through this multi-head attention mechanism, the model gains the ability to concurrently attend to distinct segments of the input data, thereby augmenting the model's expressive capacity.

S2 Dataset establishment

S2.1 Input parameter selection

TBM parameters included earth pressure, rotation speed, penetration rate, torque, thrust force, and boring speed. The data were exported from the Guangzhou Metro Big Data platform as an excel file. A comprehensive literature review was conducted to elucidate the multitude of factors that affect TBM boring performance. Table S2 presents the pertinent parameters used by researchers in the prediction of TBM boring performance.

In light of the variety of input parameters selected by researchers, in this study we incorporated

stratum information as an additional input to the model during the excavation process. Our aim was to enhance the model's adaptability to diverse strata conditions and enhance its generalization capacity. The thrust force exerted by the TBM on the cutter as it traverses the strata is recognized as a representation of the TBM's excavation performance and was designated as the chosen output parameter. Fig. S5 shows the parameters selected for consideration in the model's development.

Geological parameters included several significant attributes, namely uniaxial compressive strength (UCS), shear wave velocity, surrounding rock grade, liquid limit, plastic limit, moisture content, wet density, and void ratio. UCS, shear wave velocity, and surrounding rock grade are recognized as pivotal indicators, each potentially exerting distinct effects on TBM excavation. Consequently, we selected these parameters as our input variables. It is important to note that the data for these parameters were obtained through comprehensive testing procedures conducted by the project constructor following stratum sampling.

Unlike in other studies, we chose to include some less commonly used parameters, namely, liquid limit, plastic limit, moisture content, wet density, and void ratio, as input variables. These parameters directly serve as indicators reflecting the physical and mechanical characteristics of the soil. Soils characterized by elevated liquid limit values typically exhibit heightened plasticity and fluidity, which can potentially result in mud generation during the boring process, consequently impacting TBM operations. The plastic limit of the soil may contribute to residue buildup within the cutter and guidance system, thereby affecting the TBM's excavation progress. Soil moisture content exerts a significant influence on its mechanical properties, including shear strength and compaction behavior. Excessive moisture content may render the soil muddy, thereby increasing the cutting resistance experienced by the TBM and diminishing its excavation velocity. Conversely, insufficient water content may render the soil dry and resistant to cutting, hindering TBM efficiency. Higher wet densities may render the soil harder, posing challenges for TBM cutting and potentially affecting excavation speeds. Conversely, lower wet densities may increase cutting resistance. Elevated void ratios may enhance soil compressibility but can also lead to mud and water flow during tunneling, adversely impacting TBM cutting operations. This rationale underpinned our selection of these parameters as inputs for our model.

S2.2 Data standardization

In light of the different scales within each data sample, it is imperative to standardize the data, thereby ensuring their uniform comparability and processing. To this end, the Z-Score standardization method was used, equating the feature means to 0 and the standard deviations to 1. The formula for this standardization method is shown below:

$$x = \frac{x - \mu}{\sigma} \tag{S4}$$

where μ and σ represent the mean and standard deviation, respectively, of all data.

S2.3 Data segmentation

The partitioning of a dataset constitutes a pivotal aspect of machine learning, demanding a suitable approach to ensure robust training of the model. In this study, dataset partitioning was executed via a sliding window technique, a common practice in time series models. This approach entails configuring the input sequence length as 8, the label sequence length as 1, and maintaining the prediction sequence length at 1. Consequently, the model operates on successive 8-minute time windows as input, accompanied by contiguous 1-minute time windows as target labels, facilitating predictions for the subsequent 1 minute. This arrangement empowers the model to assimilate insights from the preceding 8 minutes of thrust force data, thereby forecasting trends for the ensuing 1 minute.

The dataset was fragmented into distinct sequences, effectively shaping the dataset. Subsequently, 70% of the data were allocated for training, 20% for the test set, and the remaining 10% for the validation set. The exact number of data sets is shown in Table S3.

S3 Results analysis

S3.1 Model evaluation metrics

Mean Squared Error (*MSE*) and Coefficient of Determination (R^2) served as the evaluation metrics for the model. The *MSE* quantifies the average of the squared differences between the model's predicted values and the true values. A smaller *MSE* indicates a closer alignment between the model's predictions and the actual values. R^2, on the other hand, gauges the relationship between the model's predictions and the true values. R^2 values range from 0 to 1, with higher values indicating stronger predictive capabilities of the model. The expressions are provided below:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{S5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2}{\sum_{i=1}^{n}(\overline{Y}_i - Y_i)^2} \tag{S6}$$

where Y_i, $\hat{Y}_i$, and $\overline{Y_i}$ are the true value, prediction of the i-th sample, and the mean value of all data, respectively.

Moreover, for the comparative assessment of the performance of various models, we introduced a novel index termed the "a20-index," which is replaced by the symbol *a* in the following. (Asteris et al., 2021a; Asteris et al., 2021b) designed to enhance the evaluation of the performance of different models. The expression is provided below:

$$a = \frac{m}{M} \tag{S7}$$

where *M* denotes the sample size of the dataset, *m* refers to the count of samples where the ratio of observed values to predicted values lies within the range of 0.80 to 1.20. In the case of an ideal predictive model, the a20 index should attain a value of 1.

S3.2 Model optimization

The main parameters governing the predictive efficacy of the Informer model include decoder depth, encoder depth, and attention mechanism. The model uses an iterative methodology, manipulating the variables of two parameters while observing the resultant changes in another parameter.

Fig. S6 depicts the *LOSS* value across various hyperparameters. Based on the hyperparameters corresponding to the minimum error value, the optimal model hyperparameters were identified as follows: encoder depth of 5, decoder depth of 1, and the use of the Full Attention mechanism. The specific model hyperparameters are detailed in Table S4.

It is of utmost importance to implement measures to prevent the occurrence of model overfitting (Asteris et al., 2019; Armaghani and Asteris, 2021). Overfitting occurs when a machine learning model excessively learns from the training data, capturing noise or random fluctuations as if they were meaningful patterns. This can lead to poor performance when the model is applied to new data. To mitigate overfitting, we incorporated the neural network regularization technique known as ‘dropout’ into our model. Dropout serves as a mechanism to mitigate overfitting by stochastically nullifying the output of a subset of neurons during the neural network's training process. Furthermore, we used the Early Stopping technique as an additional measure. This method entails continual evaluation of the model's performance on a validation set throughout the neural network training process, ceasing training once the model's performance no longer shows improvement, thereby safeguarding against overfitting.

S4 Discussion

S4.1 Input parameter selection

Input parameters characterized by substantial correlations might not necessarily contribute to enhanced model performance. This rationale underpinned our use of Pearson's analysis, as detailed in Section 3.2.3, to selectively remove input parameters featuring correlation coefficients exceeding 0.8.

Nevertheless, the model's input parameters are limited to eight. Hence, the appropriateness of the stringency of excluding variables with correlation coefficients surpassing the 0.8 threshold needs to be considered. Consequently, we explored the possibility of relaxing the assessment criterion to 0.95, while also incorporating the penetration rate as an input parameter. The augmented dataset composition is outlined in Table S6.

The prediction outcomes of the two models are

shown in Table S7. Analysis of the table shows that the new model exhibited lower loss values and higher correlation coefficients than the original model. Even with the model's excellent performance, a subtle enhancement was observed because of the relaxation of the Pearson coefficient. This validates our hypothesis that in models with few input parameters (fewer than 8), the outcomes become more credible when the threshold for the rejection of parameters with excessively large correlation coefficients is adjusted to 0.95.

S4.2 Model hyperparameter analysis

In Section S3.2, we established the model's hyperparameters and proceeded to investigate how various hyperparameter configurations impact the model's accuracy. In this section, we analyze the mechanism by which the hyperparameters affect the performance of the model.

A diminutive encoder configuration may curtail the model's capacity to glean a sufficient amount of pertinent information to effectively represent the input data, thereby giving rise to underfitting, a condition where the model inadequately fits the training data. Conversely, an overly extensive encoder might introduce an excess of superfluous information, which can lead to a model excelling on the training data but faltering in generalization when confronted with unseen data. This was the rationale behind our decision to set the encoder depth at 5.

The primary function of the decoder is to reconstruct the encoded features to match the dimensions of the original data. An excessively large decoder can learn numerous superfluous features that might enhance the model's performance on the training data. However, this could also introduce noise and extraneous information when applied to unseen data, consequently diminishing the model's capacity for generalization. This was the rationale behind our decision to set the decoder depth at 1.

Different tasks are better suited to different attention mechanisms. In some cases, the full attention mechanism may be more suitable because it considers all parts of the input sequence, making it suitable for tasks that require global information. In other cases, the sparse attention mechanism may be more appropriate as it reduces computational complexity and is suitable for handling long sequence data, thereby improving efficiency. Therefore, the choice of attention mechanism should be balanced and determined based on task requirements and computational resources.

In the case of this model, the full attention mechanism showed superiority over the sparse attention mechanism. Fully attentive mechanisms, despite their higher parameter count and computational complexity, include a comprehensive consideration of all input locations. Consequently, they are capable of capturing all intricate relationships and dependencies within the input sequence.

Table S1 Reference review

Reference	Model	Factor	Aim of prediction
(Samadi et al., 2023)	Support vector machine (SVM) Multilayer perceptron neural network (ANN-MLP) Takagi–Sugeno fuzzy model (TS fuzzy)	1. Speed 2. RPM 3. Thrust 4. RPM screw conveyor 5. RPM screw conveyor 6. Torque screw conv	Earth pressure values
(Yu et al., 2023b)	Sparrow search algorithm-Extreme Gradient Boosting (SSA-XGB) Whale optimization algorithm-Extreme Gradient Boosting (WOA-XGB) Random forest (RF) Support vector machine (SVM) Artificial neural network (ANN)	1. Earth pressure 2. Thrust 3. Cutterhead torque 4. Cutterhead speed 5. Cohesion 6. Internal friction angle 7. Compression modulus 8. Ratio of boulder 9. Uniaxial compressive strength 10. Rock quality designation	Penetration rate
(Flor et al., 2023)	Artificial neural network (ANN) Long Short-Term Memory (LSTM)	1. Cutterhead torque 2. Cutterhead power 3. RPM 4. Specific energy 5. Total absorbed power	Penetration rate
(Gokceoglu et al., 2023)	Artificial neural network (ANN) Random forest (RF) K-means clustering	1. α angle 2. Uniaxial compressive strength 3. Weathering degree 4. Water conditions 5. Cerchar abrasivity index 6. RPM 7. Torque 8. Thrust	1. Penetration rate 2. ROP
(Xu et al., 2023)	Long Short-Term Memory (LSTM)	1. RPM 2. Thrust 3. Tunnel burial depth 4. Groundwater table 5. Curvature 6. Modulus of compressibility 7. Equivalent bearing capacity	1. Advance rate 2. Cutterhead torque
(Yu et al., 2023a)	Support vector regression (SVR) K-means clustering Random forest (RF) AdaBoost Extreme Gradient Boosting (XGB) Multi-channel decoupled deep neural network (MD-DNN)	1. Cutterhead speed 2. Tunneling speed 3. Geological type	1. Thrust 2. Torque
(Shan et al., 2023)	Recurrent Neural Network (RNN) variant	1. Thrust 2. Torque 3. Face pressure 4. Revolutions per minute 5. Cover depth 6. Water table 7. Ground conditions	Penetration rate
(Shan et al., 2022)	Recurrent Neural Network (RNN) Long Short-Term Memory (LSTM)	Penetration rate	Penetration rate
(Zhang et al., 2022b)	Convolutional neural networks-Bi-Long Short-Term Memory-Attention	1. Cutter head 2. Thrust 3. Torque	Advance rate

	(CNN-Bi-LSTM-Attention)	4. Penetration rate 5. Chamber earth pressure	
(Li et al., 2022a)	Extreme Learning Machine-Ant Lion Optimizer (ELM-ALO) Extreme Learning Machine-Loin Swarm Optimization (ELM-LSO) Extreme Learning Machine-Seagull Optimization Algorithm (ELM-SOA)	1. Rock quality designation 2. Uniaxial compressive strength 3. Rock mass rating 4. Brazilian tensile strength 5. Thrust 6. Revolution	Advance rate
(Yang et al., 2022)	Grey Wolf Optimizer-Feature Weighted-Multiple Kernel-Support Vector Regression (GWO-FW-MKL-SVR) Biogeography-Based Optimization-Multiple Kernel-Support Vector Regression (BBO-FW-MKL-SVR) Multiple Kernel-Support Vector Regression (MKL-SVR) Support Vector Regression (SVR)	1. Earth pressure 2. Torque 3. Speed 4. Cohesion 5. Friction angle 6. Compression modulus 7. Ratio of boulder 8. Uniaxial compressive strength 9. Rock quality designation	Penetration rate
(Zhang et al., 2022a)	Particle Swarm Optimization-Bi-Directional Long Short-Term Memory (PSO-Bi-LSTM) Bi-Directional Long Short-Term Memory (Bi-LSTM)	Includes 21 highly correlated parameters such as thrust, torque, speed, etc.	1. Power 2. Torque 3. Speed 4. Thrust
(Bai et al., 2021)	Linear Regression (LR) Decision Tree Regression (DTR) Support Vector Regression (SVR) Gradient Boosting Regression (GBR)	1. Penetration rate 2. Bentonite volume 3. Screw conveyor speed 4. Face pressure 5. Cutterwheel torque 6. Thrust	1. Thrust 2. Face pressure 3. Cutterwheel torque
(Harandizadeh et al., 2021)	Adaptive Neuro-Fuzzy Inference System-Polynomial Neural Network-Imperialism Competitive Algorithm (ANFIS-PNN-ICA)	1. Rock quality designation 2. Rock mass rating 3. Brazilian tensile strength 4. Weathering zone 5. Uniaxial compressive strength 6. Revolution 7. Thrust	Penetration rate
(Zhou et al., 2021b)	Bayesian Optimization- Extreme Gradient Boosting (BO-XGB) Extreme Gradient Boosting (XGB)	1. Brazilian tensile strength 2. Uniaxial compressive strength 3. Rock mass rating 4. Rock quality designation 5. Weathering zone 6. Trust force per cutter 7. Revolution per minute	Advance rate
This paper	Back Propagation (BP) Extreme Gradient Boosting (XGB) Random Forest (RF) Support Vector Regression (SVM) K-Nearest Neighbors (KNN) Long Short-Term Memory (LSTM) Informer	1. Earth pressure(EP) 2. Rotation speed(RS) 3. Torque(TQ) 4. Boring speed(BR) 5. Thrust force (TF) 6. Uniaxial compressive strength (UCS) 7. Surrounding rock grade(SRG) 8. Liquid limit(LL)	Thrust force

Table S2 Literature research

References	Main input parameters
Marilena et al. (Cardu et al., 2021)	Cutter diameter
Yu et al. (Yu, et al., 2023a)	Cutterhead speed, tunneling speed, geological type
Cao et al. (Jinpu et al., 2022)	penetration, rotation speed, advance speed, cutter brake pressure, shoe pump pressure, advance pump pressure, and cutter power, torque, thrust
Ma et al. (Ma et al., 2022)	Penetration rate, total thrust, revolutions per minute, uniaxial compressive strength, volumetric joint count
Li et al. (Li et al., 2022b)	Gripper pressure, cutterhead power, cutterhead rotation speed, and propel pressure
Chen et al. (Xu et al., 2021)	Torque, cutter head power, motor power, motor current
Zhou et al. (Zhou et al., 2020)	Rotation speed, thrust, torque, penetration, boring speed
Hou et al. (Hou et al., 2020)	Torque, penetration, cutter power, boring speed, thrust
Chen et al. (Chen et al., 2019)	Cutter torque, thrust, cutter power, feed frequency
O. Acaroglu et al.. (Acaroglu, 2011)	Disc cutter diameter，disc cutter width， cutting spacing，penetration of discs，UCS，Brazilian tensile strength

Table S3 Data set segmentation

Type	Total data	Training set	Validation set	Test set
Number	37825	26551	3795	7588

Table S4 Finalized model hyperparameters

Type	Encoder attention layer depth	Number of Decoder attention layers	Type of attention	Dimension of Model	Batch_size	Seq_len	Label_len	Pred_len
Informer	5	1	Full	1024	32	8	1	1

Table S5 Model Length Setting

	Input length	Label length	Pred length
Single time-step output	4	1	1
	8	1	1
	16	1	1
	24	1	1
	30	1	1
Multiple time-step output	4	2	2
	8	2	2
		4	4
	16	2	2
		4	4
		8	8
	24	2	2
		4	4
		8	8
		12	6
		12	12
	30	2	2
		4	4
		8	8
		15	10
		15	15

Table S6 Model input parameter enhancement set

Parameters type	Parameters	Data type	Unit
TBM parameters	Earth pressure	Time-varying parameter	N/mm
	Rotation speed		Rpm
	Torque		kN*m
	Penetration rate		mm/r
	Boring speed		mm/min
	Thrust force (output)		kN
Geological parameters	Uniaxial compressive strength (UCS)	Time-invariant parameters	MPa
	Surrounding rock grade		N/A
	Liquid limit		N/A

Table S7 Model performance comparison

Type	*MSE*	R^2
Informer-original model	0.001576	0.99819
Informer-new model	0.001333	0.99847

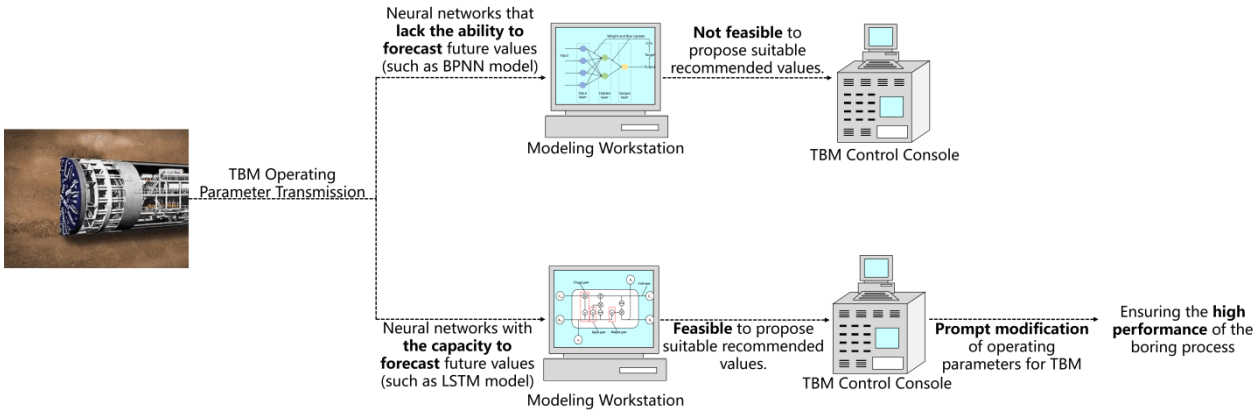

Fig. S1 Response of different models to parameter feedback and adjustment during TBM operation

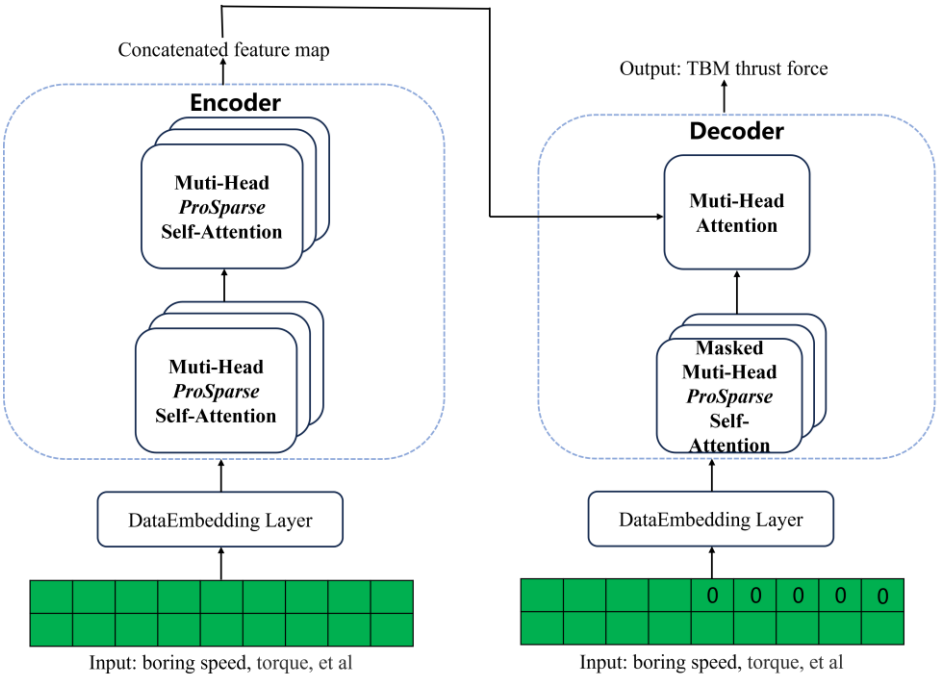

Fig. S2 Architecture of the informer model (Adapted from (Zhou, et al., 2021a))

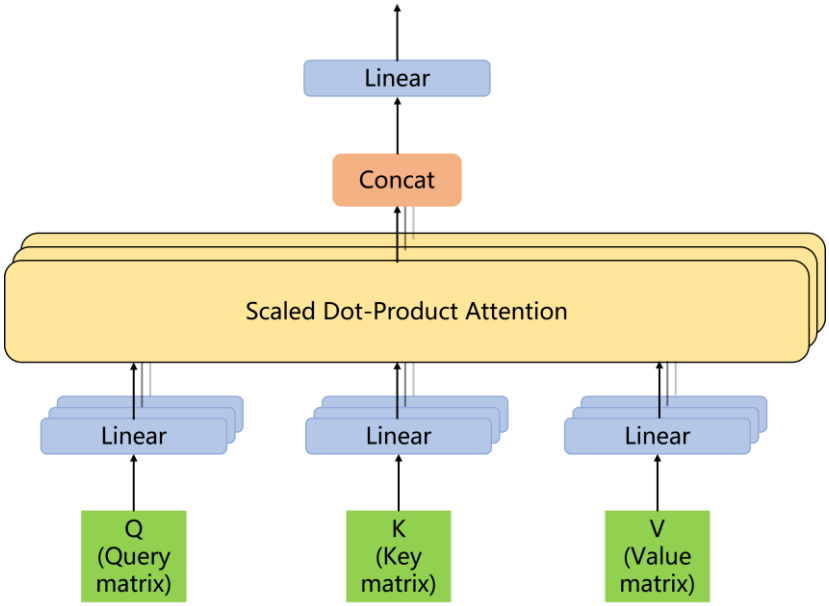

Fig. S3 Architecture of the informer model (Adapted from (Zhou, et al., 2021a))

Fig. S4 Guangzhou metro line 22 location (studied section)

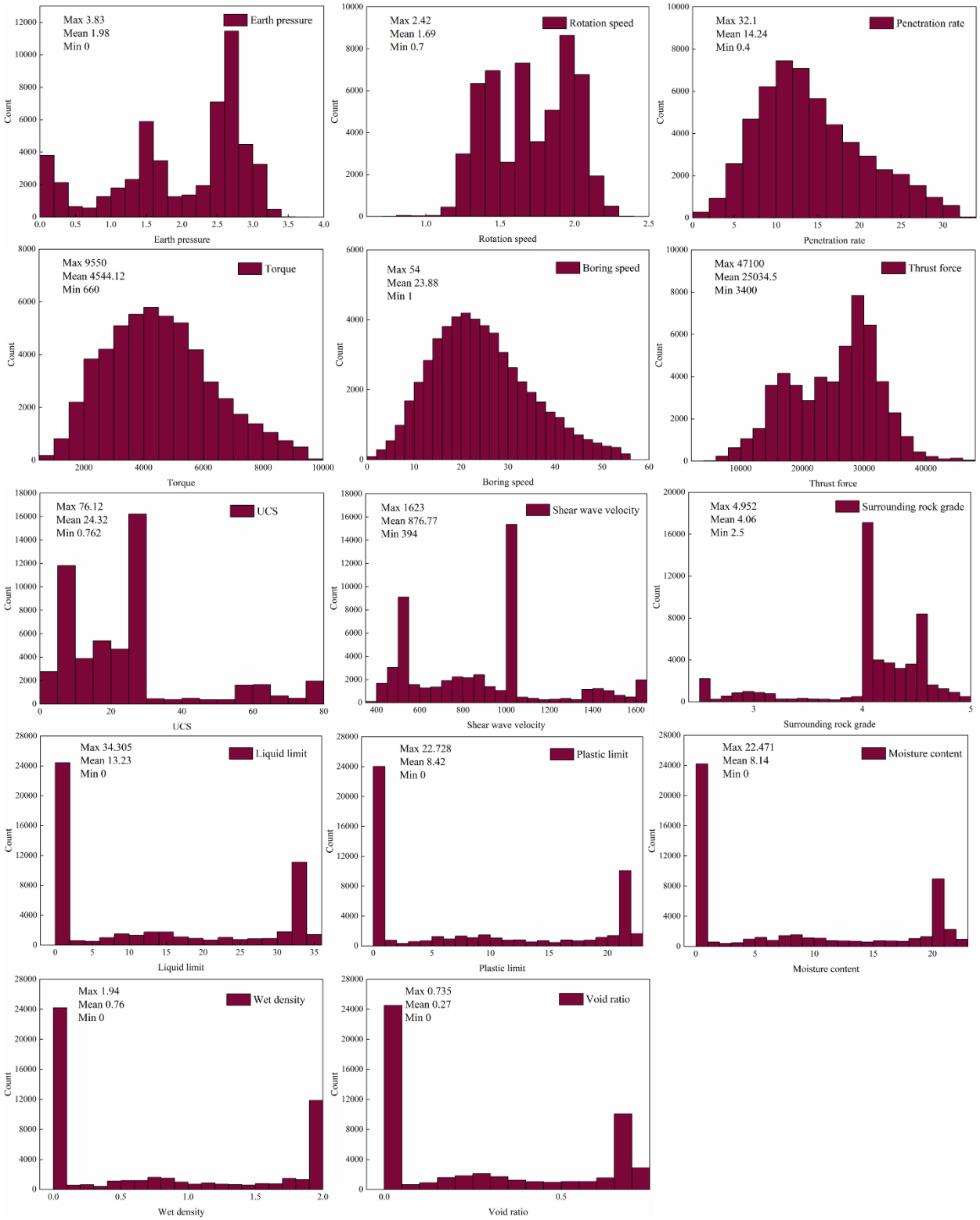

Fig. S5 Information on model input parameters

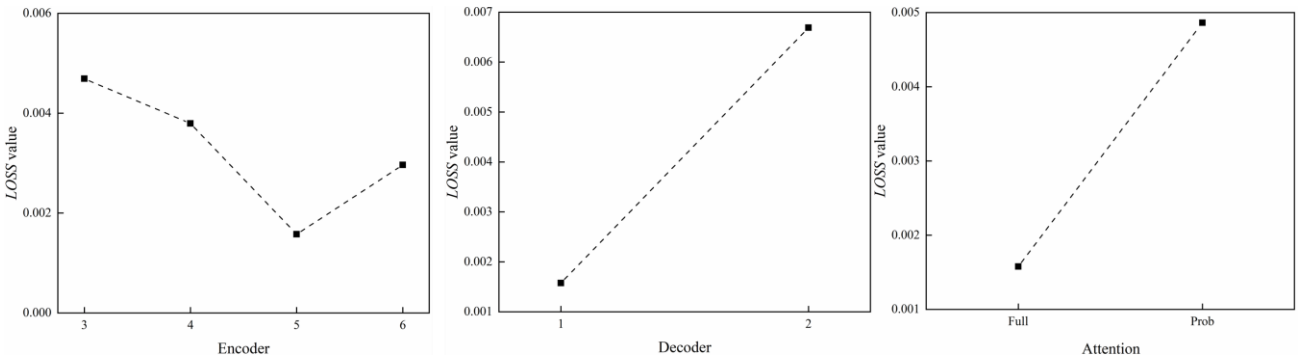

Fig. S6 Effect of different hyperparameters on model *LOSS* value

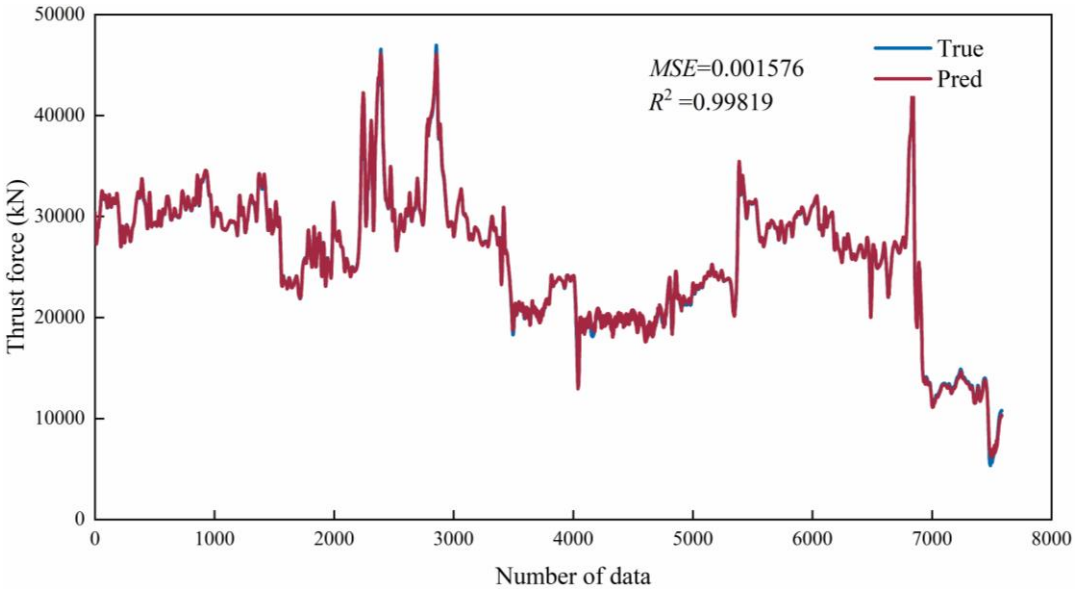

Fig. S7 Comparison of predicted values with true values

Fig. S8 Guangzhou Metro Line 18 Section Location

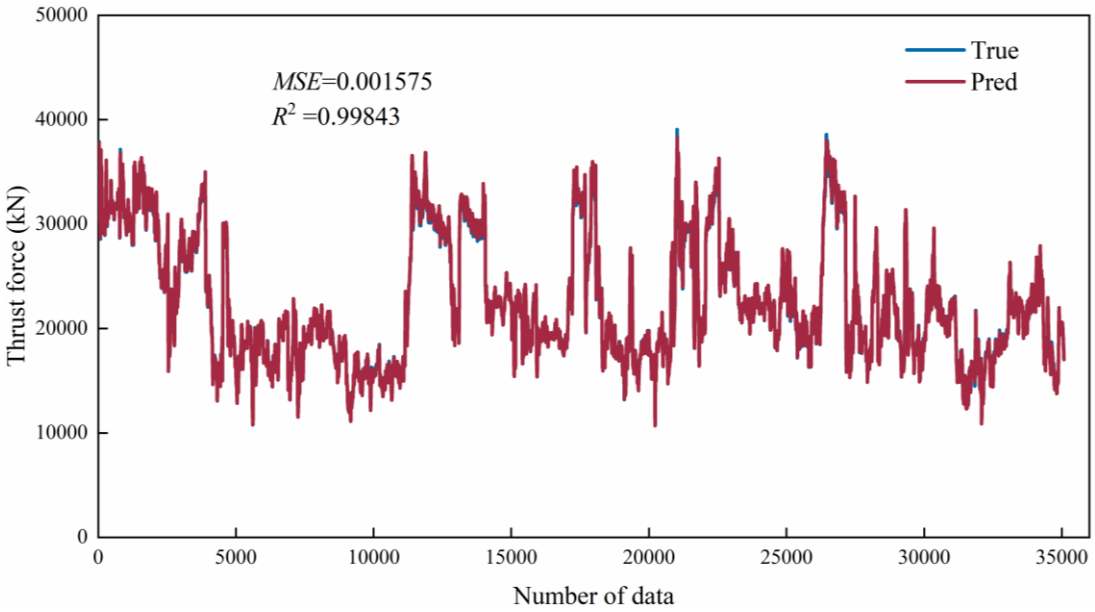

Fig. S9 Predictive performance of the model for a new project

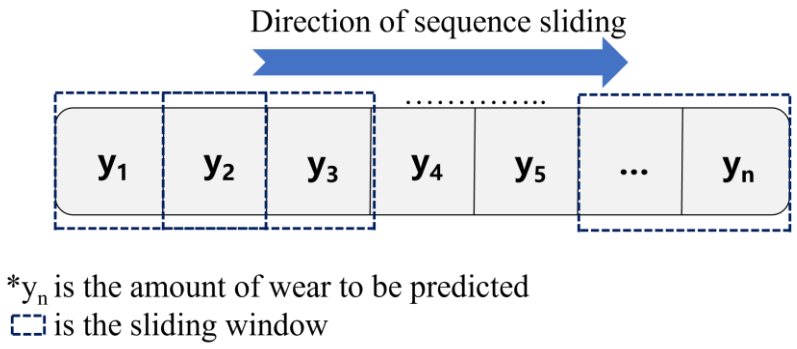

Fig. S10 Illustration of the prediction pattern

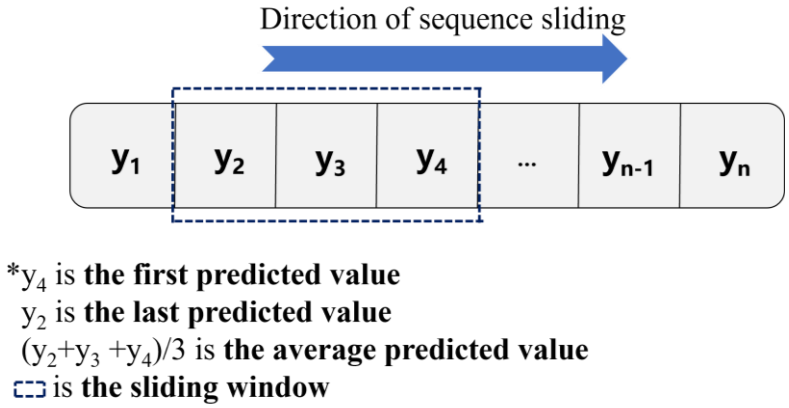

Fig. S11 Definition of relevant parameters

Reference

Acaroglu O, 2011. Prediction of thrust and torque requirements of tbms with fuzzy logic models. *Tunnelling and Underground Space Technology*, 26(2):267-275. https://doi.org/https://doi.org/10.1016/j.tust.2010.10.001

Armaghani DJ, Asteris PG, 2021. A comparative study of ann and anfis models for the prediction of cement-based mortar materials compressive strength. *Neural Computing and Applications*, 33(9):4501-4532. https://doi.org/10.1007/s00521-020-05244-4

Asteris PG, Armaghani DJ, Hatzigeorgiou GD, et al., 2019. Predicting the shear strength of reinforced concrete beams using artificial neural networks. *Computers and Concrete*, 24(5):469-488. https://doi.org/10.12989/cac.2019.24.5.469

Asteris PG, Koopialipoor M, Armaghani DJ, et al., 2021a. Prediction of cement-based mortars compressive strength using machine learning techniques. *Neural Computing and Applications*, 33(19):13089-13121. https://doi.org/10.1007/s00521-021-06004-8

Asteris PG, Lourenço PB, Hajihassani M, et al., 2021b. Soft computing-based models for the prediction of masonry compressive strength. *Engineering Structures*, 248:113276. https://doi.org/https://doi.org/10.1016/j.engstruct.2021.113276

Bai X-D, Cheng W-C, Li G, 2021. A comparative study of different machine learning algorithms in predicting epb shield behaviour: A case study at the xi'an metro, china. *Acta Geotechnica*, 16(12):4061-4080. https://doi.org/10.1007/s11440-021-01383-7

Cardu M, Catanzaro E, Farinetti A, et al., 2021. Performance analysis of tunnel boring machines for rock excavation. *Applied Sciences*, 11:2794. https://doi.org/10.3390/app11062794

Flor A, Sassi F, La Morgia M, et al., 2023. Artificial intelligence for tunnel boring machine penetration rate prediction. *Tunnelling and Underground Space Technology*, 140:105249. https://doi.org/https://doi.org/10.1016/j.tust.2023.105249

Gokceoglu C, Bal C, Aladag CH, 2023. Modeling of tunnel boring machine performance employing random forest algorithm. *Geotechnical and Geological Engineering*, 41(7):4205-4231. https://doi.org/10.1007/s10706-023-02516-3

Harandizadeh H, Armaghani DJ, Asteris PG, et al., 2021. Tbm performance prediction developing a hybrid anfis-pnn predictive model optimized by imperialism competitive algorithm. *Neural Computing and Applications*, 33(23):16149-16179. https://doi.org/10.1007/s00521-021-06217-x

Hou S, Liu Y, Zhang K, 2020. Prediction of tbm tunnelling parameters based on ipso-bp hybrid model. *Yanshilixue Yu Gongcheng Xuebao/Chinese Journal of Rock Mechanics and Engineering*, 39(8):1648-1657. https://doi.org/10.13722/j.cnki.jrme.2019.1084

Jinpu C, Fang L, Zhifu S, 2022. A lstm-based model for tbm performance prediction and the effect of rock mass grade on prediction accuracy. *Tumu Gongcheng Xuebao/China Civil Engineering Journal*, 55:92-102. https://doi.org/10.15951/j.tmgcxb.2022.S2.06

Li C, Zhou J, Tao M, et al., 2022a. Developing hybrid elm-alo, elm-lso and elm-soa models for predicting advance rate of tbm. *Transportation Geotechnics*, 36:100819. https://doi.org/https://doi.org/10.1016/j.trgeo.2022.100819

Li L, Liu Z, Lu Y, et al., 2022b. Hard-rock tbm thrust prediction using an improved two-hidden-layer extreme learning machine. *IEEE Access*, 10:112695-112712. https://doi.org/10.1109/ACCESS.2022.3216294

Ma T, Jin Y, Liu Z, et al., 2022. Research on prediction of tbm performance of deep-buried tunnel based on machine learning. 12(13):6599.

Samadi H, Hassanpour J, Rostami J, 2023. Prediction of earth pressure balance for epb-tbm using machine learning algorithms. *International Journal of Geo-Engineering*, 14(1):21. https://doi.org/10.1186/s40703-023-00198-7

Shan F, He X, Jahed Armaghani D, et al., 2022. Success and challenges in predicting tbm penetration rate using recurrent neural networks. *Tunnelling and Underground Space Technology*, 130:104728. https://doi.org/https://doi.org/10.1016/j.tust.2022.104728

Shan F, He X, Armaghani DJ, et al., 2023. Effects of data smoothing and recurrent neural network (rnn) algorithms for real-time forecasting of tunnel boring machine (tbm) performance. *Journal of Rock Mechanics and Geotechnical Engineering*, https://doi.org/https://doi.org/10.1016/j.jrmge.2023.06.015

Xu C, Liu X, Wang E, et al., 2021. Prediction of tunnel boring machine operating parameters using various machine learning algorithms. *Tunnelling and Underground Space Technology*, 109:103699. https://doi.org/https://doi.org/10.1016/j.tust.2020.103699

Xu Q, Huang X, Zhang B, 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/https://doi.org/10.1016/j.undsp.2022.11.002

Yang H, Wang Z, Song K, 2022. A new hybrid grey wolf optimizer-feature weighted-multiple kernel-support vector regression technique to predict tbm performance. *Engineering with Computers*, 38(3):2469-2485. https://doi.org/10.1007/s00366-020-01217-2

Yu H, Qin C, Tao J, et al., 2023a. A multi-channel decoupled deep neural network for tunnel boring machine torque and thrust prediction. *Tunnelling and Underground Space Technology*, 133:104949. https://doi.org/https://doi.org/10.1016/j.tust.2022.104949

Yu Z, Li C, Zhou J, 2023b. Tunnel boring machine performance prediction using supervised learning method

and swarm intelligence algorithm. 11(20):4237.

Zhang Q, Zhu Y, Ma R, et al., 2022a. Prediction method of tbm tunneling parameters based on pso-bi-lstm model. *Frontiers in Earth Science*, 10 https://doi.org/10.3389/feart.2022.854807

Zhang Y, Chen J, Han S, et al., 2022b. Big data-based performance analysis of tunnel boring machine tunneling using deep learning. 12(10):1567.

Zhou J, Qiu Y, Zhu S, et al., 2021b. Estimation of the tbm advance rate under hard rock conditions using xgboost and bayesian optimization. *Underground Space*, 6(5):506-515. https://doi.org/https://doi.org/10.1016/j.undsp.2020.05.008

Zhou X, Gong Q, Yin L, et al., 2020. Predicting boring parameters of tbm stable stage based on blstm networks combined with attention mechanism. *Yanshilixue Yu Gongcheng Xuebao/Chinese Journal of Rock Mechanics and Engineering*, 39:3505-3515. https://doi.org/10.13722/j.cnki.jrme.2019.1158