

## Research Article

# A Data-Driven Method for Predicting the Cutterhead Torque of EPB Shield Machine

Kairong Hong,<sup>1</sup> Fengyuan Li,<sup>1</sup> Zhenjian Zhou,<sup>1</sup> Feng Li ,<sup>2</sup> and Xunlin Zhu <sup>2</sup>

<sup>1</sup>State Key Laboratory of Shield Machine and Boring Technology, Zhengzhou 450001, China

<sup>2</sup>School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, Henan 450001, China

Correspondence should be addressed to Xunlin Zhu; [xlzhu@zzu.edu.cn](mailto:xlzhu@zzu.edu.cn)

Received 30 August 2021; Revised 24 September 2021; Accepted 25 September 2021; Published 18 October 2021

Academic Editor: Juan L. G. Guirao

Copyright © 2021 Kairong Hong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The prediction of cutterhead torque of earth pressure balance (EPB) shield machine is mainly studied. First, the idea of shield tunneling stage division is proposed. The process of shield tunneling from start to stop (or pause) is divided into start-up and stationary driving stages. Using the change point detection method based on linear regression, the separation points between start-up stage and stationary driving stage are identified from the original construction data, and the datasets of the two stages are extracted, respectively. Then, for the start-up stage, the linear regression method is suggested for the cutterhead torque prediction, since there is a strong linear correlation between the key parameters such as the cutterhead torque and the thrust force. Meanwhile, for the stationary driving stage, considering the fact that the key parameters vary smoothly and show obvious inertia, the long short-term memory (LSTM) network method can be used to establish the relationship model between cutterhead torque and other key parameters, such as the thrust force. Through the test experiments of construction data in Zhengzhou, Luoyang, and Dalian shield projects, the results show that the proposed segmented modeling method possesses good adaptability and the cutterhead torque prediction model has high prediction accuracy.

## 1. Introduction

Torque of the cutterhead is one of the key parameters for the tunnel boring machine (TBM), which maintains the cutterhead to rotate and cut the soil continuously, and is also the basic parameter for equipment energy efficiency control and safety state monitoring. It is closely related to the size and configuration of the equipment, geomechanical parameters, shield operation parameters, etc., so its influencing factors are very complex [1, 2]. Estimation of cutterhead torque is an important basis for the equipment parameter design and construction control and is also the basic guarantee for fast, safe, and efficient tunneling.

The existing literature mainly focuses on the geometric characteristics of the components of the cutterhead, loading state, and soil constitution and establishes the model predicting the cutterhead torque based on mechanical analysis. Usually, many assumptions and approximations are needed to satisfy the ideal modeling conditions. However, the

physical and mechanical properties of soil are strongly correlated to soil composition, water content, geological structure, and other factors, so current soil strength theories can only be applied to the physical and mechanical properties of some certain type of soil under specific loading conditions, and there is no accurate constitutive relationship to describe its mechanical properties in practice [3–5]. For example, the friction coefficient between the cutterhead and the surrounding soil must be assumed artificially when calculating the torque of the cutterhead. All these assumptions would lead to inaccurate estimation of cutterhead torque [6–10].

In recent years, with the rapid development of computer and artificial intelligence technology, various data-driven methods have provided much support for the analysis of huge engineering data [11–13]. In [14], Wang et al. studied the prediction of cutterhead torque based on the real data of Pine diversion and water supply project in Jilin Province of China, where nonlinear support vector regression (NSVR) is

employed, and cutterhead torque is taken as the output and some operation parameters as inputs. However, all of these algorithm models, such as artificial neural network (ANN), support vector machine (SVM), and random forest models (RFM), are difficult to explain because of their complex nonlinear structure and are misleading to understand the relationship between the response and the predictors [15–18].

In this paper, based on the periodic characteristics of the shield tunneling process, the data of the starting and steady tunneling stages of the TBM are separated and two different regression models about the cutterhead torque and other operation parameters are proposed, according to the different distributions of the data in both the stages. The analysis of actual TBM operation data shows that the proposed data-driven piecewise modeling method for predicting the cutterhead torque has better accuracy and applicability. Specifically, the main contributions of this paper are listed as follows:

- (1) A clear data collection method is proposed, that is, data are collected according to similar geological conditions and the same type of shield machine
- (2) The change point between the start-up stage and the stationary driving stage in an excavation cycle can be identified by the change point detection method based on linear regression, and the data are separated into two parts as start-up dataset and the stationary driving dataset
- (3) After denoising and normalization, linear regression models are used on the starting dataset and the long short-term memory (LSTM) recurrent neural network [19] model is employed on the stationary driving dataset
- (4) Compared with the existing methods, the segmented regression models are established, and the proposed method takes into account not only the influence of geological conditions and type of TBM on the modeling but also the influence of different states of TBM

The rest of this paper is organized as follows. In Section 2, the necessity of data segmentation for shield construction is clarified and a method of data segmentation is provided. In Section 3, the modeling methods of start-up stage and steady driving stage are presented, respectively. In Section 4, a real data analysis is given to illustrate the effectiveness of the proposed modeling methods. Finally, Section 5 concludes the paper pointing towards future research directions.

## 2. Data Extraction and Preprocessing

In the process of shield construction, the TBM successively experiences the stages of starting, driving, suspension, driving, suspension, ..., driving, shutdown, etc. It is easy to see that the operation parameters have different distributions at the stages of starting and driving. Therefore, in order to accurately predict the cutterhead torque, it is necessary to separate the data into two sets, corresponding to the start-up stage and the driving stage, respectively.

*2.1. Necessity for the Separation of Data according to Stages.* The operation parameters present different trends when TBM works in different stages such as start-up and stationary driving. Figure 1 shows the local curves of the cutterhead rotation speed (CRS), the advance velocity (AV), the total thrust force (TTF), and the cutterhead torque (CT) (observed once per second), where the dataset is sampled from Ring 550 of Hui-Shang section of the subway tunnel project Line 4 in Zhengzhou of China. One can see that the operation parameters such as CRS, CT, and TTF grow up quickly at the start-up stage, and then the operating parameters oscillate slightly. The TBM operation data lasting no less than 120 s from the start-up to the suspension (or shutdown) stage are regarded as normal operation data, while those lasting less than 120 s from the starting to the suspension (or shutdown) stage are regarded as abnormal operation data, and the data in the suspension or shutdown stage are defined as shutdown data. The normal operation data, abnormal operation data, and shutdown data are continuous segmentally and appear alternately. As shown in Figure 1, when the shield machine is working, the total thrust force and the advance velocity will not be zero at the same time; otherwise, the shield machine will be suspended or shutdown. According to this feature, the starting and ending points of each normal operation data segment can be detected, and the normal operation data segments that meet the requirements can be extracted.

Figure 2 shows the curves of typical segments of the normal operation data in Ring 550. When the shield tunnel machine starts, the key parameters (including CT, TTF, AV, and CRS) increase rapidly; then, the parameters become stationary after a period (about 100 s), which is called the stationary driving stage, and now the shield tunnel machine is in a relatively steady working state. Because of the significant difference of the operation parameters between the start-up stage and the stationary driving stage, it is necessary to establish the regression model segmentally. All the data are divided into three parts according to the start-up stage, the stationary driving stage, and other stages, where the datasets with respect to other stages contain some data that do not share the characteristics of the start-up stage and the stationary driving stage. In this way, the segmented models based on the datasets of different stages can better characterize the operation parameters in different stages and will have good performance in terms of prediction accuracy.

*2.2. Change Point Detection between the Start-Up Stage and Stationary Driving Stage.* Because of the different characteristics of the operation parameters shown in different stages, it is possible to identify the change point between the start-up stage and stationary driving stage. All the data are divided into three parts according to the start-up stage, the stationary driving stage, and other stages, and the segmented regression model based on the start-up dataset and stationary driving dataset would be established, respectively.

From Figure 2, one can see that the thrust force shows obviously different trends between the start-up stage and

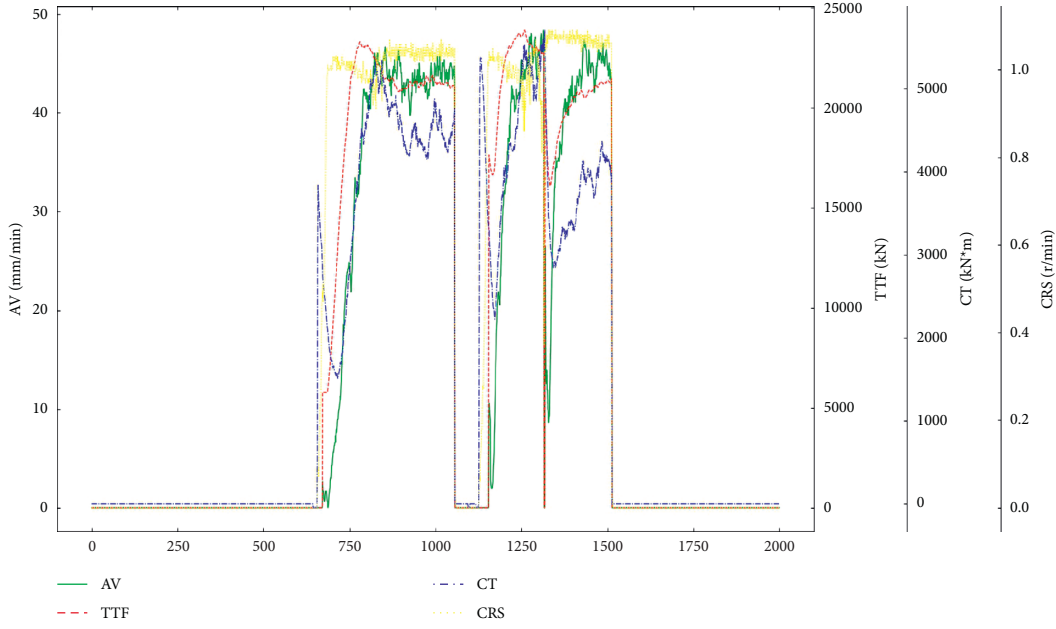


FIGURE 1: Local curves of some key TBM operation parameters in Ring 550.

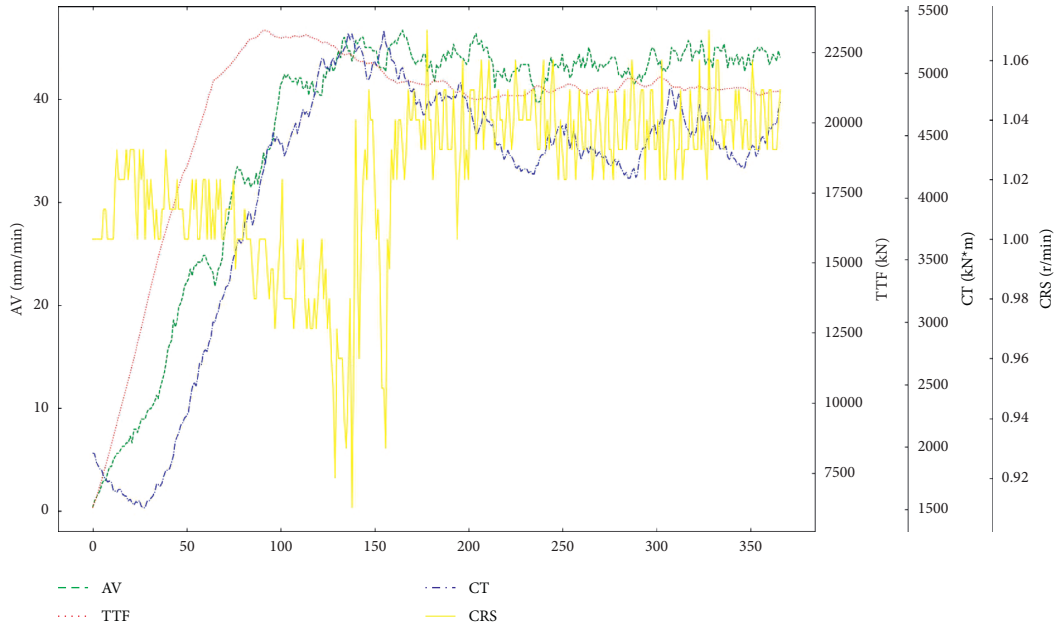


FIGURE 2: A section of normal tunneling data in Ring 550.

stationary stage, and advance velocity and the cutterhead torque display the similar phenomenon to the thrust force. Therefore, the change point detection method based on the linear regression model will be applied to the three parameters mentioned above, so as to comprehensively determine the true change point.

Change point detection based on the linear regression model can be described as follows. Let  $(t_i, y_i) (1 \leq i \leq N)$  be the observations, where  $y_i$  is observed at instant  $t_i$ , and there exist  $m (1 < m < N)$  and  $a_i$  and  $b_i (i = 1, 2)$  which attain the minimum of  $S_m$ :

$$S_m = \sum_{i=1}^m (a_1 t_i + b_1 - y_i)^2 + \sum_{i=m+1}^N (a_2 t_i + b_2 - y_i)^2. \quad (1)$$

If  $a_1 = a_2$  and  $b_1 = b_2$  do not hold simultaneously, a mutation occurs at  $m$  in the regression model, and  $m$  is called the change point.

In essence, two linear regression models with different slopes are employed to represent the relationship between the operation parameter and time  $t$ , and the start-up stage and stationary driving stage are separated by the change

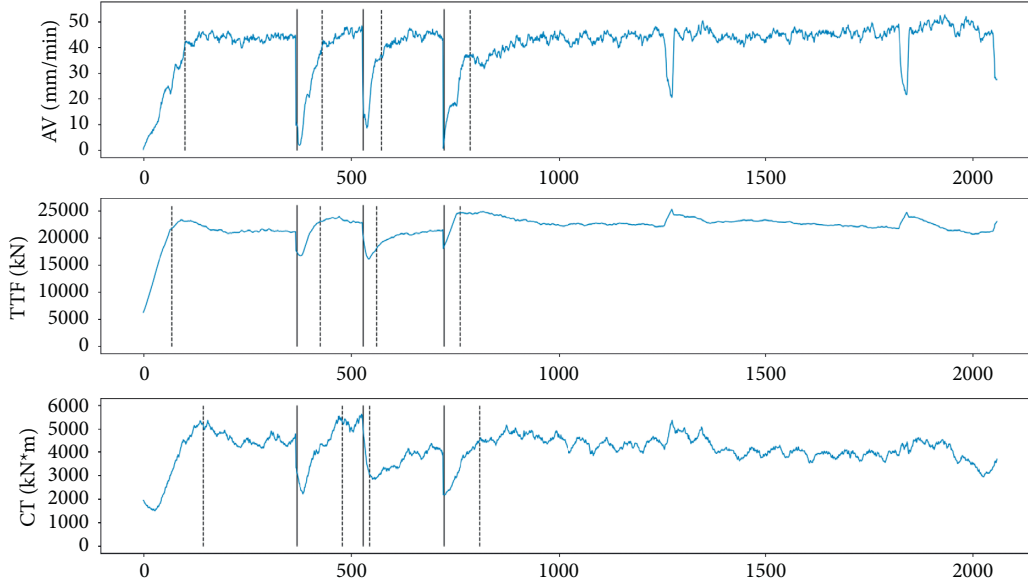


FIGURE 3: The regression change points of the normal tunneling data segments in Ring 550.

point  $m$ . Figure 3 depicts the change points of AV, TTF, and CT in Ring 550. The vertical dashed line represents the time of change point, and the vertical solid line identifies the starts and ends of the normal tunneling data. It could also be observed that the change points identified from AV, TTF, and CT differ slightly. So, the mean of the three change points is taken as the estimated change point. Figure 4 shows the estimated segmented regression lines of the three operation parameters on time, where the vertical dashed line indicates the change point.

**2.3. Denoising and Normalization.** To reduce the interference of the outliers, principal “ $3\sigma$ ” is employed to remove the data out of the interval  $[\bar{x} - 3\sigma, \bar{x} + 3\sigma]$ , where  $\bar{x}$  and  $\sigma$  represent the sample mean and standard deviation of operation parameter  $x$ , respectively. All the operation parameters are also scaled valued in the interval  $[0, 1]$ , i.e.,

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (2)$$

where  $x_{\min}$  and  $x_{\max}$  are the minimum and maximum of the sample of  $x$ . Then, the Pearson correlation coefficients are calculated and sorted to select the important parameters, which are strongly related to cutterhead torque.

### 3. Regression Model for Cutterhead Torque on Other Operation Parameters

In this section, we will establish different regression models to capture the relationship between cutterhead torque and other operation parameters based on the start-up dataset and the stationary driving dataset.

**3.1. Regression Model for Start-Up Stage.** The data collected by the State Key Laboratory of Shield Tunneling Technology of China on the intelligent big data platform of shield

tunneling construction contain more than 500 attributes. The key attributes (or parameters) include CT, propelling pressure of group A of cylinders (PPA), propelling pressure of group B of cylinders (PPB), propelling pressure of group C of cylinders (PPC), propelling pressure of group D of cylinders (PPD), AV, CRS, and rotation speed of screw conveyor (RSSC). To reduce the influence of the geological conditions and type of TBM, the construction data with the same type of TBM and similar geological conditions are sampled. For example, we sampled the dataset from the subway tunnel project Line 4 in Zhengzhou of China, Hui-Shang section, Rings 550–556, where the main geology is fine sand. The standard penetration value is 23.1, and the diameter of the TBM is 6.8 m. After normalizing of the start-up dataset, we compute the correlation coefficient matrix of the selected operation parameters, and the results are listed as follows.

The correlation coefficient between the variables (parameters)  $x_i$  and  $x_j$  is defined as

$$\text{Cor}(x_i, x_j) = \frac{\text{Cov}(x_i, x_j)}{\sqrt{\text{Var}(x_i)}\sqrt{\text{Var}(x_j)}} \quad (3)$$

Larger  $|\text{Cor}(x_i, x_j)|$  means stronger linear relationship between parameters  $x_i$  and  $x_j$ . It could be observed from Table 1 that the cutterhead torque has strong linear relationship with the advance velocity, the rotation speed of screw conveyor, and the propelling pressure of four groups of cylinders. Thus, the following linear regression model is employed to predict cutterhead torque:

$$y = \beta_0 + \sum_{i=1}^7 \beta_i x_i + \varepsilon, \quad (4)$$

where the response variable  $y$  is the value of cutterhead torque, the operation variables  $x_i$  ( $i = 1, 2, \dots, 7$ ) are the values of CRS, AV, PPA, PPB, PPC, PPD, and RSSC,

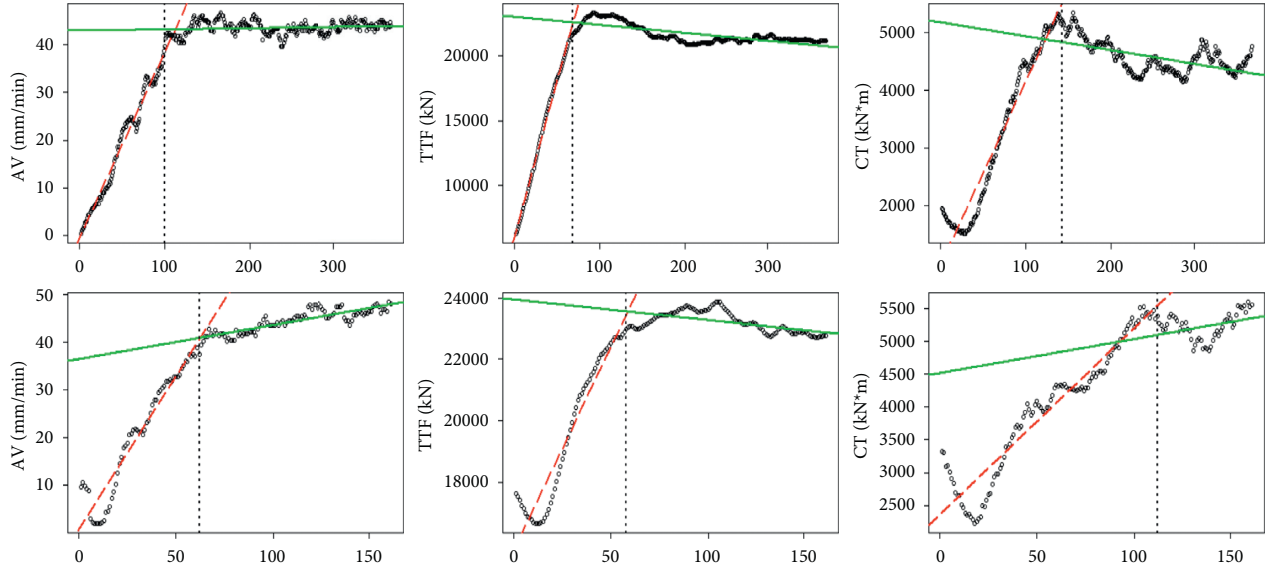


FIGURE 4: The regression model fitting of the first two normal tunneling data sections in Ring 550.

TABLE 1: Correlation coefficient matrix of some key operation parameters based on the start-up dataset.

	CRS	AV	PPA	PPB	PPC	PPD	RSSC	CT
CRS	1	-0.13643	-0.02136	-0.00857	-0.04812	-0.09561	-0.09722	0.12934
AV	-0.13643	1	0.73151	0.70609	0.80865	0.25975	0.95053	0.75640
PPA	-0.02136	0.73151	1	0.99760	0.95197	0.34319	0.73406	0.83770
PPB	-0.00857	0.70609	0.99760	1	0.94163	0.34318	0.71289	0.83657
PPC	-0.04812	0.80865	0.95197	0.94163	1	0.36540	0.79748	0.80595
PPD	-0.09561	0.25975	0.34319	0.34318	0.36540	1	0.28389	0.15211
RSSC	-0.09722	0.95053	0.73406	0.71289	0.79748	0.28389	1	0.79534
CT	0.12934	0.75640	0.83770	0.83657	0.80595	0.15211	0.79534	1

respectively. Without loss of generality, all the variables are normalized. The unknown coefficients  $\beta_i$  ( $i = 0, 1, 2, \dots, 7$ ) can easily be estimated via the least square method, and the estimators are denoted by  $\hat{\beta}_i$ . Consequently, given the operation parameters  $x_i$  ( $i = 1, 2, \dots, 7$ ), the cutterhead torque can be predicted as  $\hat{y} = \hat{\beta}_0 + \sum_{i=1}^7 \hat{\beta}_i x_i$ .

Although some machine learning algorithms such as random forest model (RFM), support vector regression (SVR), artificial neural network (RNN), and long short-term memory neural network (LSTM) could be used to predict cutterhead torque, here we prefer to apply the linear regression model (LM) for its good interpretability, less computational complexity, and especially for the strong linear correlation between cutterhead torque and other operation parameters shown in Table 1. The analysis of actual data in the next section demonstrates that the LM performs well compared with SVR and RFM.

### 3.2. Regression Model for Stationary Driving Stage.

Different from the start-up stage, the cutterhead torque and other parameters in the stationary driving stage are in a relatively steady state. Observations of each parameter can be regarded as a stationary time series, which can be verified by the augmented Dickey-Fuller (ADF) stationarity test. Let

$\{X_t | t = 0, 1, 2, \dots\}$  be a time series with expectation  $\mu_t = E(X_t)$ , variance  $\sigma_t^2 = \text{Var}(X_t)$ , and covariance  $\gamma(s, t) = \text{Cov}(X_s, X_t)$ , and if  $\mu_t = \mu$ ,  $\sigma_t^2 = \sigma^2$ , and  $\gamma(s, t) = \gamma(t - s)$ ,  $\{X_t | t = 0, 1, 2, \dots\}$  are called a stationary time series. If the  $p$  value of the ADF test for the time series is close to zero, we can conclude the time series are stationary.

For stationary time series, the lagged order  $p$  can be determined by autocorrelation function (ACF) and partial autocorrelation function (PACF). As shown in Figure 2, the cutterhead torque in the stationary driving stage could be demonstrated as a stationary time series. In addition, there is a strong correlation between cutterhead torque and other operating parameters during the stationary driving stage. Specifically, we take the dataset which is sampled from Rings 550–556 of Hui-Shang section of the subway tunnel project Line 4 in Zhengzhou of China as an example. After extraction and normalization for the stationary driving dataset, we obtain the correlation matrix among cutterhead torque and other key operation parameters shown in following table.

In Table 2, RA (rolling angle) and PA (pitch angle) are other two parameters. It can be observed that the correlation coefficients are all above 0.6 between cutterhead torque and other seven parameters, CRS, AV, PPA, PPB, PPC, PPD, and RSSC.

TABLE 2: Correlation coefficients among cutterhead torque and other key parameters on the stationary driving dataset.

	CRS	AV	PPA	PPB	PPC	PPD	RSSC	RA	PA	CT
CRS	1	0.83	0.65	0.81	0.78	0.55	0.84	0.25	0.21	0.91
AV	0.83	1	0.74	0.85	0.88	0.61	1.00	0.29	0.25	0.95
PPA	0.65	0.74	1	0.78	0.76	0.72	0.75	0.68	-0.02	0.76
PPB	0.81	0.85	0.78	1	0.84	0.58	0.85	0.52	-0.02	0.89
PPC	0.78	0.88	0.76	0.84	1	0.63	0.88	0.42	0.13	0.88
PPD	0.55	0.61	0.72	0.58	0.63	1	0.62	0.23	0.28	0.60
RSSC	0.84	1.00	0.75	0.85	0.88	0.62	1	0.29	0.26	0.95
RA	0.25	0.29	0.68	0.52	0.42	0.23	0.29	1	-0.53	0.34
PA	0.21	0.25	-0.02	-0.02	0.13	0.28	0.26	-0.53	1	0.19
CT	0.91	0.95	0.76	0.89	0.88	0.60	0.95	0.34	0.19	1

Because of the autocorrelation of cutterhead torque and strong linear correlation with other parameters, the LSTM neural network is employed to predict cutterhead torque for stationary driving stage. It is a special recursive neural network (RNN) with the hidden nodes (also called cells) incorporating three kinds of gate operations, including the input gate, the forget gate, and the output gate. By tuning the status of these gates, the information flowing among LSTM's hidden layers could be controlled and the drawback of training the traditional RNNs could be overcome, where vanishing gradients or exploding gradients usually occur when the network is used to process a long input sequence [20, 21].

In the prediction of cutterhead torque in the shield tunneling stage, the use of LSTM neural network model can not only take the influence of the previous moment of cutterhead torque into consideration but also cover the correlation between cutterhead torque and other key equipment parameters. In this way, while ensuring the accuracy of prediction, it can also enhance the adaptive ability of the model and expand the applicable scope of the model.

The LSTM network for prediction of cutterhead is set as follows:

Input: CRS, RSSC, PPA, PPB, PPC, PPD, and CT, with the training window width  $p$

Output: CT at next time instant

**3.3. Evaluation Statistic for Regression Model.**  $R^2$  is usually employed to assess the performance of the regression model, which is defined as follows:

$$R^2 = 1 - \frac{\sum_{j=1}^N (y_j - \hat{y}_j)^2}{\sum_{j=1}^N (y_j - \bar{y})^2}, \quad (5)$$

where  $y_j$  is the  $j$ th observation of the response,  $\hat{y}_j$  is its predicted value,  $\bar{y}$  is the sample mean of response,  $N$  is sample size. Closer to 1 as  $R^2$  is, the performance of the regression model would be better.

#### 4. Analysis of Actual Data

In this section, some actual data sets including subway tunnel projects in Zhengzhou, Luoyang and Dalian of China are analyzed to illustrate the effectiveness of the proposed methods and models. Firstly, the data are denoised and normalized. Secondly, start-up data set and stationary driving data set are extracted by the change point detection method based on linear regression models. Thirdly, the segmented regression models for the start-up stage and stationary driving stage are established, some other machine learning models as SVR and RFM are also compared.

**4.1. Separation between Start-Up and Stationary Driving Stages.** All the phenomenon of the separations between start-up and stationary driving stages are similar, so we only present Figure 5 as an example, where PDA denotes the propelling displacement of group A of cylinders. Figure 5 shows the change points by dotted vertical lines, where the data are from Ring 552 of Hui-Shang section of the subway tunnel project Line 4 in Zhengzhou of China.

**4.2. Linear Regression Model for Cutterhead in the Start-Up Stage.** Take the data from Rings 550–556 of Hui-Shang section of the tunnel project Line 4 in Zhengzhou as an example, and the data from the first six rings are used as training data and the last ring as testing data. The estimated linear regression model for cutterhead torque is obtained as follows:

$$\hat{y} = -0.13 + 0.09x_1 + 0.06x_2 - 0.74x_3 + 1.79x_4 - 0.48x_5 - 0.16x_6 + 0.54x_7, \quad (6)$$

where all the variables are normalized, and  $x_i$  ( $i = 1, 2, \dots, 7$ ) represent CRS, AV, PPA, PPB, PPC, PPD, and RSSC, respectively. Figures 6–8 show the predictive curves and true

curves on testing data by the proposed linear regression model, SVR, and RFM, respectively. The predictive curves on the nonsegmented data by the linear regression model are

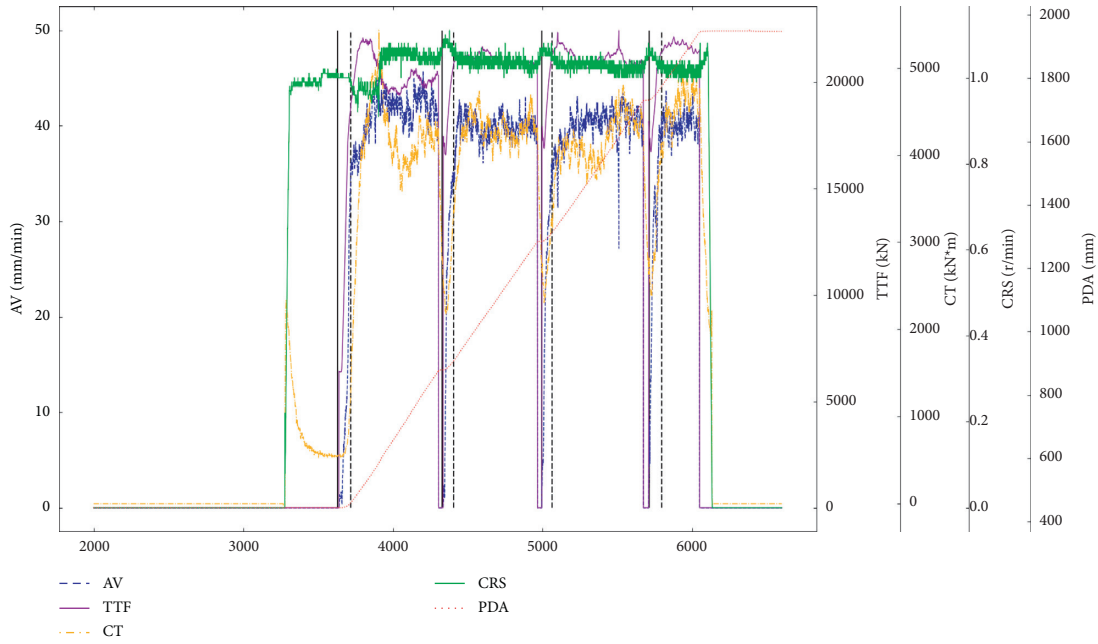


FIGURE 5: Segmentation results of construction data in ring 552 of hui-shang section, line 4 of zhengzhou.

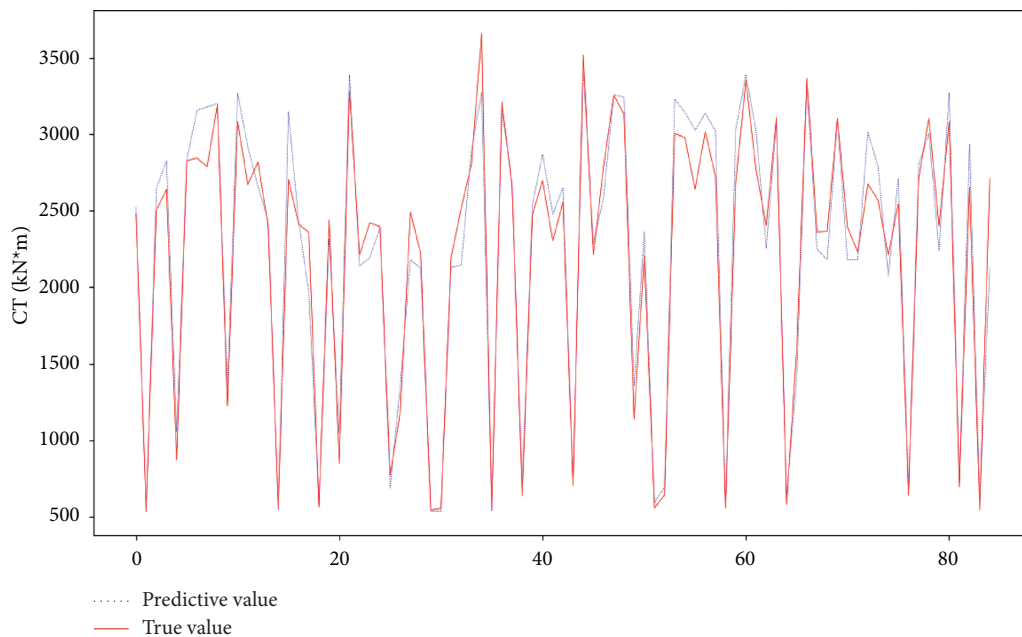


FIGURE 6: Effect of predicting cutterhead torque by using the linear regression model on the testing dataset.

depicted in Figure 9. It is easy to see that the linear regression model performs well compared with SVR and RFM for start-up datasets, while the linear regression model performs poorly for nonsegmented data. This fact also reveals that the segmented modeling is necessary for better prediction of cutterhead torque.

In addition, the comparison results for these models on different datasets are presented in Table 3. The data are processed by using Python 3.7 in a computer with Core i5-

6200U at 2.4 GHz, 8 GB RAM, Windows 10 Operation System. For the SVR, the Gaussian radial basis function kernel is employed, where  $\gamma = 0.1$ , the penalty parameter  $C = 1000$ , and  $\epsilon = 0.1$ . For RFM, the number of regression trees is selected as 13 and the maximal depth of the trees is 9.

The estimated linear regression model on the dataset sampled from the subway tunnel project Line 1 in Luoyang of China, some section, Rings 104–106, is as follows:

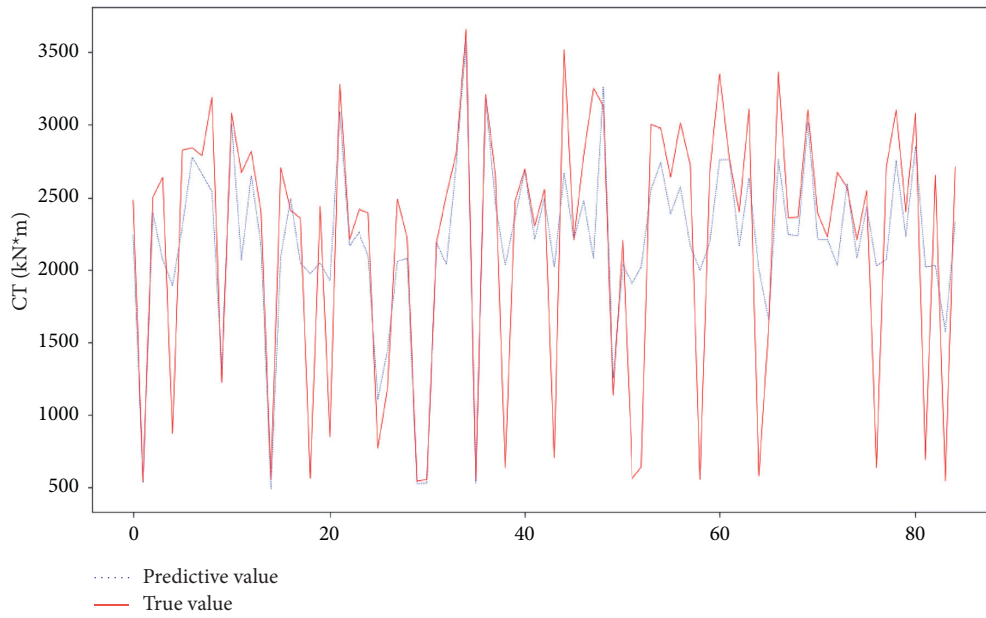


FIGURE 7: Effect of predicting cutterhead torque by using the SVR model on the testing dataset.

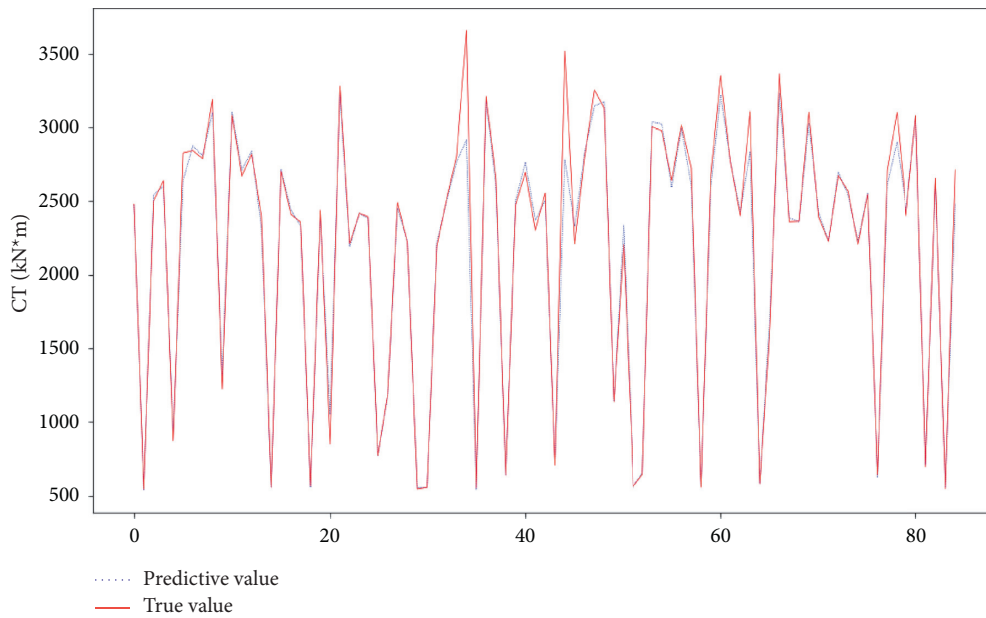


FIGURE 8: Effect of predicting cutterhead torque by using the RFM on the testing dataset.

$$\hat{y} = -0.07 + 0.40x_1 + 0.31x_2 + 0.06x_3 + 0.12x_4 - 0.04x_5 + 0.05x_6 + 0.08x_7. \tag{7}$$

Also, the estimated linear regression model on the dataset sampled from the subway tunnel project Line 5 in Dalian of China, Rings 8047-8048, is as follows:

$$\hat{y} = -0.10 - 0.14x_1 + 0.31x_2 + 0.74x_3 + 0.48x_4 - 0.56x_5 + 0.04x_6 + 0.13x_7. \tag{8}$$



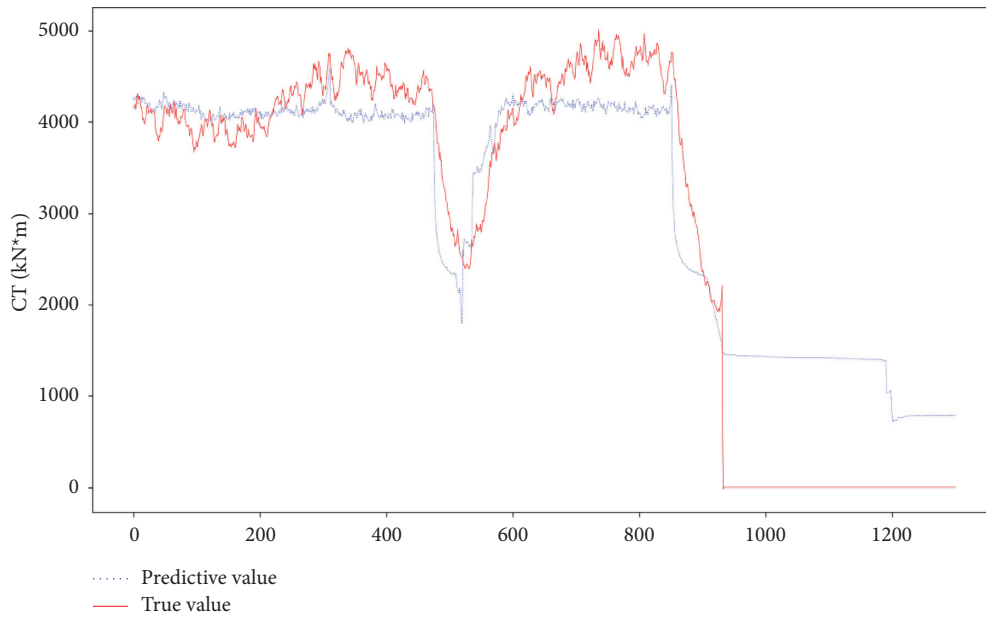


FIGURE 9: Effect of predicting cutterhead torque by using the linear regression model on the nonsegmented dataset.

TABLE 3: Comparison of scores with different models.

Line	Geology	Model	Statistic			
			Training data $R^2$	Testing data $R^2$	RMSE	Time (s)
Zhengzhou, Line 4, Rings 550–556	Fine sand, penetration value 23.1, surrounding rock grade VI	LM	0.81	0.78	548.3	0.88
		SVR	0.97	0.75	544.1	2.24
		RFM	0.99	0.81	497.8	1.78
Luoyang, Line 1, Rings 104–106	Silty clay, penetration value 14.3, surrounding rock grade V	LM	0.95	0.94	122.9	0.02
		SVR	0.92	0.88	138.3	1.33
		RFM	0.99	0.96	121.9	0.42
Dalian, Line 5, Rings 47-48	Moderately weathered limestone, surrounding rock grade V	LM	0.78	0.77	259.3	0.03
		SVR	0.94	0.81	272.7	2.77
		RFM	0.99	0.79	245.2	1.36

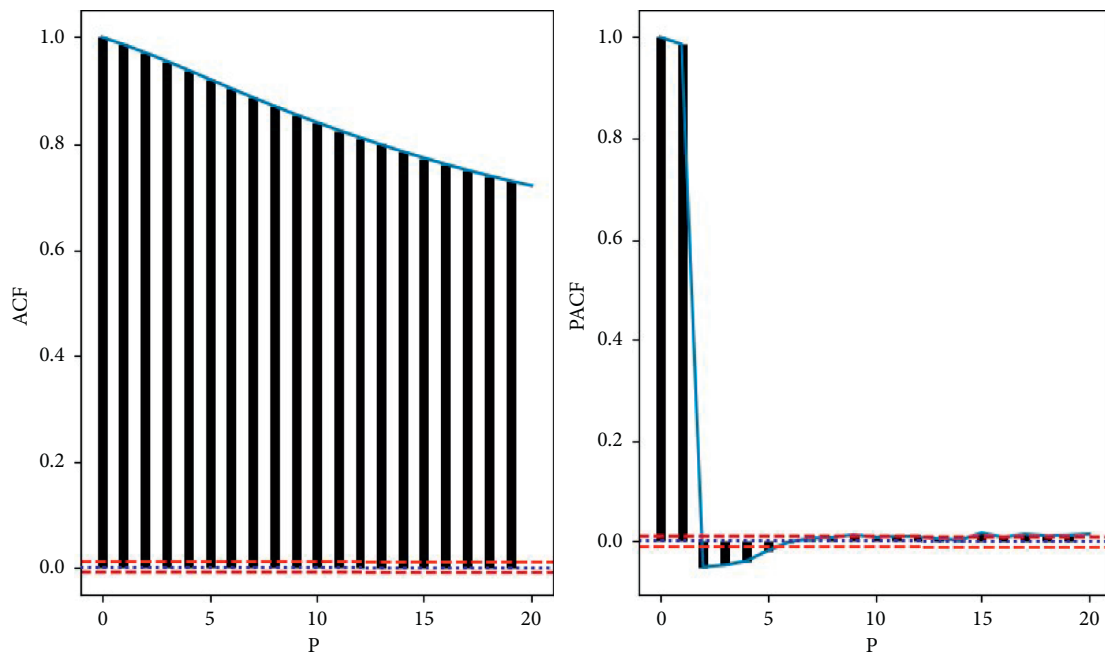


FIGURE 10: ACF and PACF charts for cutterhead series.

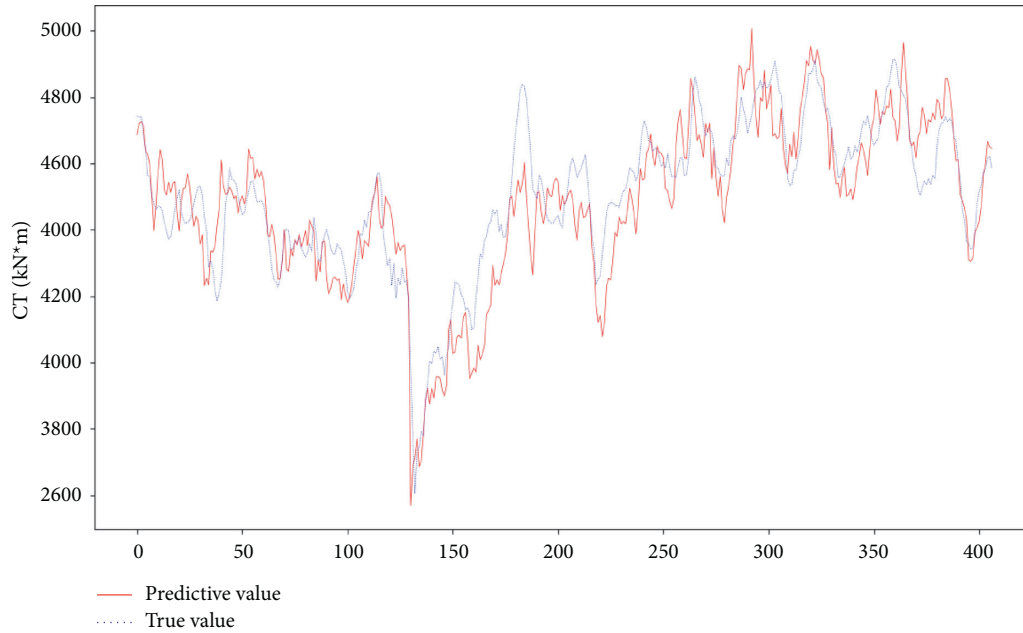


FIGURE 11: The prediction effect of LSTM model.

As shown in Table 3, the linear regression model performs better on testing data than SVR and RFM, and the latter two models overfit the training data. In addition, the time elapsed for LM is much less than that for SVR and RFM. Because of the well interpretability, facility for computation of the linear regression model, it is used to predict the cutterhead torque in the start-up stage.

**4.3. LSTM Prediction Model for Cutterhead in the Stationary Driving Stage.** Take the dataset from the subway tunnel project Line 4 in Zhengzhou of China, Hui-Shang section, Rings 520–641 as an example; after removing the outliers, the sample size of the stationary driving dataset is 248400. Among these data, sample observations of 100 rings are selected as training data and the remaining 49680 observations are taken as testing ones. The ADF test for stationary of cutterhead torque is conducted, and the  $p$  value is  $4.798 \times 10^{-24}$ , which is near to zero, so we can conclude that the cutterhead torque series are stationary. The ACF and PACF are shown in Figure 10; it can be seen that the PACF is truncated after order of  $p = 1$ . The predictive curves by LSTM and true curves of cutterhead on testing data are depicted in Figure 11, and  $R^2 = 0.97$ . The results show that the LSTM model well simulates the relationship between cutterhead torque and other operating parameters. Considering both the dependence of the series and influences of other operation parameters, the LSTM performs well in predicting cutterhead torque in the stationary driving stage.

## 5. Conclusions

According to the characteristics of the key operation parameters, the idea of segmenting data is initiated and the change point detection based on the linear regression model

is employed to separate the start-up and stationary driving stages. Then, segmented regression models are established for predicting cutterhead torque. The main conclusions are briefly presented as follows:

- (1) The proposed change point detection method based on the linear regression model separates the data into start-up dataset and stationary driving dataset quite well.
- (2) In the start-up stage, CT, CRS, and TTF increase quickly and cutterhead torque has strong linear correlation with other key operation parameters. Compared with machine learning models such as SVR and RFM, the established linear model performs better. Meanwhile, the linear model merits the good interpretability and easy feasibility, which is a perfect choice to predict the cutterhead torque in the start-up stage.
- (3) In the stationary driving stage, considering the dependence of the parameter series and the influence of other operation parameters, the LSTM model is employed to predict cutterhead torque. The analysis of actual data demonstrates that the LSTM performs quite well.
- (4) The proposed segmented cutterhead torque predictive model can also be applied for early warning. In brief, first, the cutterhead torque prediction model based on the dataset from previous two rings is established; then, the cutterhead torque is predicted to judge whether the prediction error is bigger than a preset threshold value, and a warning will be issued if it is. This application is currently being tested on the intelligent big data platform of shield and TBM construction in the State Key Laboratory of Shield Machine and Boring Technology.

Next, we will further optimize the modeling process according to the test results, find effective geological condition classification methods, and explore the establishment of cutterhead torque automatic alarm system, so as to better improve the safety of shield construction.

### Data Availability

The data sampled from Rings 520–641 of Hui-Shang section of the tunnel project Line 4 in Zhengzhou, Rings 104–106 of Line 1 in Luoyang, and Rings 47–48 of Line 5 in Dalian of China are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant nos. 61873245 and 61772476, the National Key Research and Development Project under Grant no. 2018YFB1701404, the Major Scientific and Technological Innovation Project of Zhengzhou under Grant no. 188PCZX786, China Railway Tunnel Group's Scientific and Technological Innovation Plan: "Research on real-time intelligent control technology of shield attitude based on deep reinforcement learning" (Contract no. 2018-40), and China Railway Tunnel Group's Scientific and Technological Innovation Plan: "Shield active parameter optimization and passive parameter early warning based on artificial intelligence and big data analysis" (Contract no. 2018-39) and Open Project of State Key Laboratory of Shield Machine and Boring Technology (Contract no. SKST-2018-K02).

### References

- [1] H. Guan and B. Gao, "Theoretical model for estimation of cutter head torque in shield tunneling," *Chinese Journal of Southwest Jiaotong University*, vol. 43, no. 2, pp. 213–217, 2008.
- [2] J. Cao, N. Zhang, G. Liang, and L. Feng, "Analysis of thrust and torque prediction method for TBM cutter head," *Construction Mechanization*, vol. 39, no. 5, pp. 51–54, 2018.
- [3] L. Zhu, Z. Zheng, and R. Li, "Calculation of shield cutter head torque in mix ground," *Mechanical Engineering & Automation*, vol. 6, pp. 110–111, 2014.
- [4] B. Zhou and H. Zhang, "Shield machine advancing speed optimization study base on genetic algorithm," *Construction Machinery*, vol. 5, pp. 101–104, 2015.
- [5] Z. Yu, "Change rule and correlation analysis of slurry shield tunneling parameters under argillaceous siltstone formation," *Construction Technology*, vol. 47, no. 1, pp. 116–121, 2018.
- [6] W. Xu, W. Ni, H. Zhao, and S. Chen, "Research on driving parameters of large diameter slurry shield tunnel under complex interbedded strata," *Tunnel Construction*, vol. 38, no. Z2, pp. 373–378, 2018.
- [7] J. Xie, *Study on Tunneling Parameters of Mud Water Shield Tunnel with Sand and Limestone Composite Strata*, Guangzhou University, Guangzhou, China, 2018.
- [8] H. Wang and D. Fu, "A mathematical model and the related parameters for EPB shield tunneling," *China Civil Engineering Journal*, vol. 39, no. 9, pp. 86–90, 2006.
- [9] H. Zhang, X. Wu, and W. Zeng, "Study on tunneling experiment and mathematical model of EPB shield," *Chinese Journal of Rock Mechanics and Engineering*, vol. 24, no. S2, pp. 5762–5766, 2005.
- [10] H. Zhou, S. Ban, and Y. Han, "Study and application of optimal tunneling parameters of TBM," *Water Resources Development & Management*, vol. 29, no. 4, pp. 86–88, 2009.
- [11] R. A. de Assis, R. Pazim, M. C. Malavazi, P. P. d. C. Petry, L. M. E. de Assis, and E. Venturino, "A mathematical model to describe the herd behaviour considering group defense," *Applied Mathematics and Nonlinear Sciences*, vol. 5, no. 1, pp. 11–24, 2020.
- [12] L. Li, Y. Wang, and X. Li, "Tourists forecast lanzhou based on the baolan high-speed railway by the arima model," *Applied Mathematics and Nonlinear Sciences*, vol. 5, no. 1, pp. 55–60, 2020.
- [13] F. Jin, M. W. Wong, S. Liu, X. Zhang, J. Duan, and Z. Zhang, "Prediction of Chinese automobile growing trend considering vehicle adaptability based on cui," *CLawson Model, Applied Mathematics and Nonlinear Sciences*, vol. 5, no. 2, pp. 367–376, 2020.
- [14] C. Wang, G. Gong, H. Yang, J. Zhou, L. Duan, and Y. Zhang, "NSVR based predictive analysis of cutterhead torque for hard rock TBM," *Journal of Zhejiang University*, vol. 52, no. 3, pp. 479–486, 2018.
- [15] S. Lee, J. Ha, O. Na, and S. Na, "The cusum test for parameter change in time series models," *Scandinavian Journal of Statistics*, vol. 30, no. 4, pp. 781–796, 2003.
- [16] L.-J. Jing, J.-B. Li, C. Yang, S. Chen, N. Zhang, and X.-X. Peng, "A case study of TBM performance prediction using field tunnelling tests in limestone strata," *Tunnelling and Underground Space Technology*, vol. 83, pp. 364–372, 2019.
- [17] W. Sun, M. Shi, C. Zhang, J. Zhao, and X. Song, "Dynamic load prediction of tunnel boring machine (TBM) based on heterogeneous in-situ data," *Automation in Construction*, vol. 92, pp. 23–34, 2018.
- [18] A. Salimi, J. Rostami, C. Moormann, and A. Delisio, "Application of non-linear regression analysis and artificial intelligence algorithms for performance prediction of hard rock TBMs," *Tunnelling and Underground Space Technology*, vol. 58, pp. 236–246, 2016.
- [19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [20] X. Gao, M. Shi, X. Song, C. Zhang, and H. Zhang, "Recurrent neural networks for real-time prediction of TBM operating parameters," *Automation in Construction*, vol. 98, pp. 225–235, 2019.
- [21] A. Peng, D. Cao, B. Zhao, X. Yang, and M. Zhang, "Reservoir physical parameters prediction based on LSTM recurrent neural network," *Progress in Geophysics*, vol. 34, no. 5, pp. 1849–1958, 2019.