

# Real-time Early Warning of Clogging Risk in Slurry Shield Tunneling: A Self-updating Machine Learning Approach

Qiang Wang<sup>a</sup>, Xiongyao Xie<sup>a</sup>, and Yu Huang<sup>a</sup>

<sup>a</sup>Department of geotechnical engineering, Tongji University

E-mail: [wangqiang\\_tju@qq.com](mailto:wangqiang_tju@qq.com), [xiexiongyao@tongji.edu.cn](mailto:xiexiongyao@tongji.edu.cn), [yhuang@tongji.edu.cn](mailto:yhuang@tongji.edu.cn)

**Abstract –**

Clogging is one of the main risks when slurry shield tunneling in the mixed ground condition containing clayey soils. Severe consequences, such as instability in the excavation face and high cutter wear, may occur if the shield machine operators don't take specific measures to eliminate clogging. Therefore, early warning of clogging during one ring excavation becomes essential for the safety of tunneling. The currently available methods to judge the clogging risks focus mainly on field engineer experience, which seems arbitrary sometimes. In this paper, an automatic self-updating machine learning approach is proposed to realize the real-time early warning of clogging. More specifically, the random forest is employed with several minutes (e.g. 2 min at the beginning of one ring excavation) of tunneling parameters as input. When one ring has been finished, it will become a new training sample to update the model via randomized parameter optimization. With the case study of Nanning metro line 1, it's found that the self-updating mechanism is beneficial for better judgment of clogging, and 4 minutes tunneling parameters (24 samples) are suitable for early warning. The model can achieve an accuracy of 95% in the mixed ground condition. Meanwhile, in comparison with the other machine learning approaches is also discussed. With the training data set updating mechanism, the RF model can use less tunneling data to realized the clogging prediction. According to the feature importance result, the variation of cutterhead torque is essential for clogging prediction.

**Keywords –**

Shield tunneling; Clogging; Early warning; Random forest; Self-updating

## 1 Introduction

Shield tunneling in clayey soils is frequently obstructed by clogging problems, both for earth pressure balanced (EPB) tunnel boring machine (TBM) and slurry

pressure balanced (SPB) tunnel boring machine [1,2]. The clogging problem will trigger risks during tunneling construction, such as slower tunneling efficiency[3], instability of tunnel face[4], higher wear of cutterhead[5], etc. As a typical kind of clayey soil, mudstone has a high potential to result in a clogging problem[4]. Several projects in China have been encountered with clogging problems in mudstone rich area, for example, Wuhan Sanyang cross-river road tunnel[6], the metro tunnel line 1&2 in Nanning city[7], Nanchang metro line 1[8], Nanjing Yangze river tunnel[9]. The filed experience indicates that it's difficult to maintain the normal tunneling state in mudstone rich area, especially for the mixed ground containing mudstone.

Clogging is induced by the stickiness of the excavated clayey soil, which can be influenced both by the clay mineralogy and the slurry flow behavior [2,10]. There are several laboratory tests have been presented for evaluating the clogging potential of one certain kind of soil and slurry flow behavior [2,10–12], which brings out some directly or indirectly methodologies for clogging judgment. These clogging elevation methods rely on the soil properties and most of them focused on either sand or clayey soils, which has a limitation in mixed ground conditions[2,13]. Moreover, there is still a lack of early warning methods for clogging during the tunneling process.

The criteria to evaluate clogging potential mainly focused on the soil properties but pay little attention to the slurry or foam properties well as shield driver operations in shield tunneling process[14]. When SPB shield tunneling in the mixed ground containing mudstone, it is difficult to determine whether clogging occurs only rely on the geological conditions. Therefore it is crucial to develop an early warning approach for clogging both based on the geological conditions and tunneling parameters. The data-driven approach, such as the random forest (RF) method, seems appropriate in the tunneling process. The clogging state can be regarded as an abnormal tunneling situation, thus it can be considered as a binary classification problem. RF algorithm for the development of descriptive and predictive data-mining models has become widely accepted in engineering

applications, promising powerful new tools for practicing engineers [15].

Kohestani et al.[16] presented an RF-based model for prediction of seismic liquefaction potential of soil based on the cone penetration test data, and the proposed RF models provide more accurate results than the artificial neural network (ANN) and the support vector machine (SVM) models. Zhou et al.[17] employed eleven algorithms to predict the rockburst classification and found the RF achieved the best result. The above-mentioned applications of the RF-based classification model are all static models. They employed well-trained RF models for all test data sets, which is not suitable for the shield tunneling process. As a real-time early warning model for clogging, the proposed model should be trained via as few rings as possible and is supposed to realize self-updating as new rings have been finished. As a result, a training data set updating mechanism will be designed in this paper to realize the real-time early warning of clogging risk in SPB shield tunneling.

The remainder of the paper is organized as follows: In section 2, we introduce the real-time early warning model based on RF. In section 3, the case study in Nanning metro will be presented. The impact of training data set updating mechanism will be discussed in section 4 before presenting the conclusions.

## 2 Real-time early warning of clogging risks based on RF

Figure 1 shows the process for real-time early warning of clogging risks with training data updating during shield tunneling construction. Suppose there are  $L$  rings in total for a tunnel section, and the shield machine is going to tunneling at the ring  $\#i$  ( $i \leq L$ ). Firstly, we select  $N$  minutes tunneling data at the beginning of each  $i-1$  rings as the early warning input feature ( $N = 0.5, 1, 2, 3, \dots$ ) and take the  $i-1$  rings clogging state as output, which will be composed as the training data set. Secondly, the  $i$ th clogging prediction model is trained with the above training data set, whose hyper-parameters are determined by the 5-fold cross-validation and randomized search[18]. Then, the proposed prediction model is used to predict the clogging state of ring  $\#i$  via the  $N$  minutes tunneling data at the beginning of the  $i$ th ring. Finally, when the ring  $\#i$  has been finished, we will determine the real clogging state for the  $i$ th ring based on the above mentioned three criteria. Moreover, the training data set will be updated by adding the input feature and clogging state of ring  $\#i$ . This process will continue until the tunneling section has been finished.

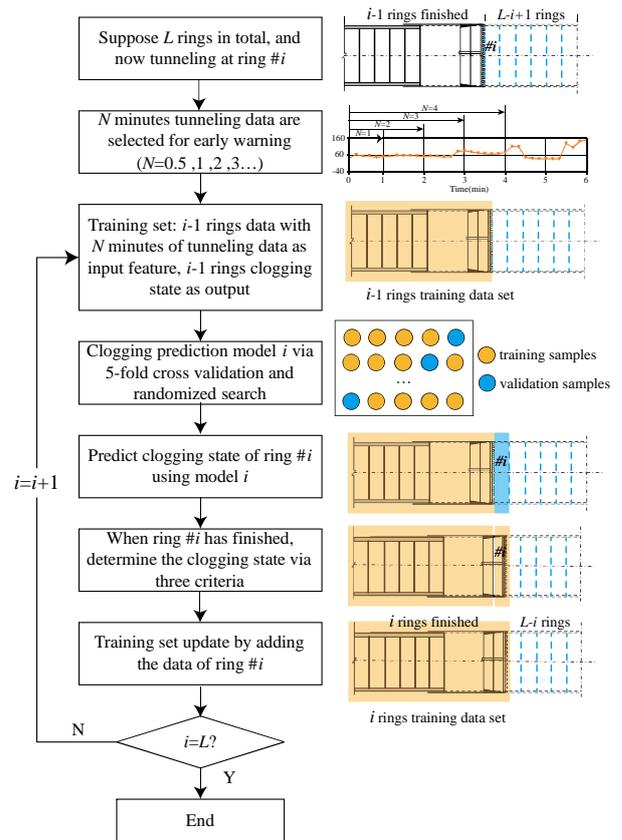


Figure 1. Flow chart for real-time early warning of clogging with training data updating

Random forest (RF) is an ensemble learning method for classification that operates by constructing a large number of decision trees at training time and outputting the class that is the mode of the classes of the individual trees[19]. As illustrated in Figure 2, samples and features are randomly selected from the data set using the bootstrap aggregating method. Therefore, many sub-samples are created by choosing random features with replacement. Then each decision tree will be trained on the sub-sample and the final class (clogging or normal) will be determined by averaging the probabilistic prediction of all the trees. Several hyper-parameters in the RF will be determined by cross-validation, such as the number of trees in the forest ( $n\_estimators$ ), the maximum depth of the tree ( $max\_depth$ ), etc. To improve the training efficiency during shield tunneling construction, the randomized search strategy is employed instead of the traditional grid search method. This strategy implements a randomized search over the hyper-parameters, where each setting is sampled from a distribution over possible hyper-parameter values.

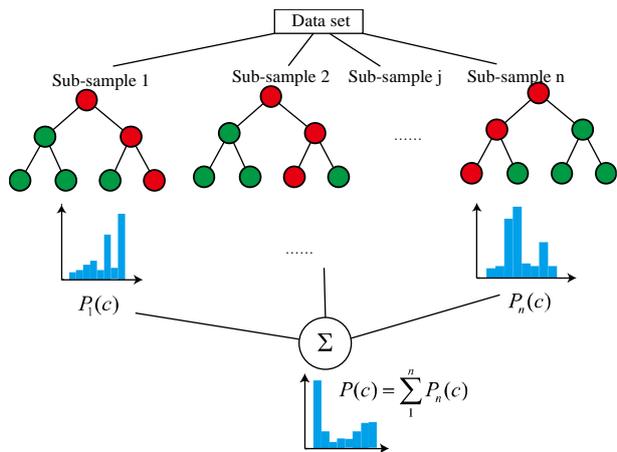


Figure 2. Schematic diagram of random forest classifier

### 3 Case study in Nanning metro

#### 3.1 Project Overview

Nanning metro line 1 is the first metro line with twin tunnels in Guangxi province. The section of left line in Nanning Metro Line 1 between Bai Cang Ling Station and Railway Station (BR section) is 1209 m (806 rings in total with segment width of 1.5 m), which is encountered with the mixed ground condition of mudstone and round gravel and is thus selected as a case study. As a massive of sensitive adjacent buildings along the tunnel section,

shown in Figure 3 (a), the BR section was excavated with a Herrenknecht SPB shield machine with a diameter of 6.28 m.

The tunnel in the BR section is surrounded by complex geological and hydrological conditions. The tunnel passing through geological profiles (illustrated in Figure 3 (b)) are mainly round gravel, mudstone, and sand, which are shown in different colors in Figure 3 (b). The mixed ground containing mudstone locates around ring # 120 to ring # 220, and ring # 283 to ring # 470. A total of 36 boreholes have been drilled to obtain the geotechnical characteristics of the soils found along the BR section (refer to Table 1). The distance between the ground to the tunnel crown ranges from 14 m to 22 m, while the distance between the underground water table to the tunnel crown is about 1.5m to 9m. The round gravel is saturated in the BR section, which mainly consists of gravel with a small number of pebbles (Figure 4 (a)). The average content of particles with a grain size of 2-20mm is 52.8%, and the average content of particles with a grain size greater than 20mm is 25.3%. The inter-particle filling is mainly medium and coarse sand. The round gravel has a very strong permeability coefficient of 90 m/d. In contrast, the mudstone (Figure 4 (b)) has a small permeability coefficient of 0.01 m/d, as listed in Table 1. Figure 4 (c) illustrates the grain size distribution of mudstone samples obtained in the BR section. It can be found that the grain size of mudstone is almost smaller than 0.1 mm and there are about 64% of particles with a grain size smaller than 0.005 mm.

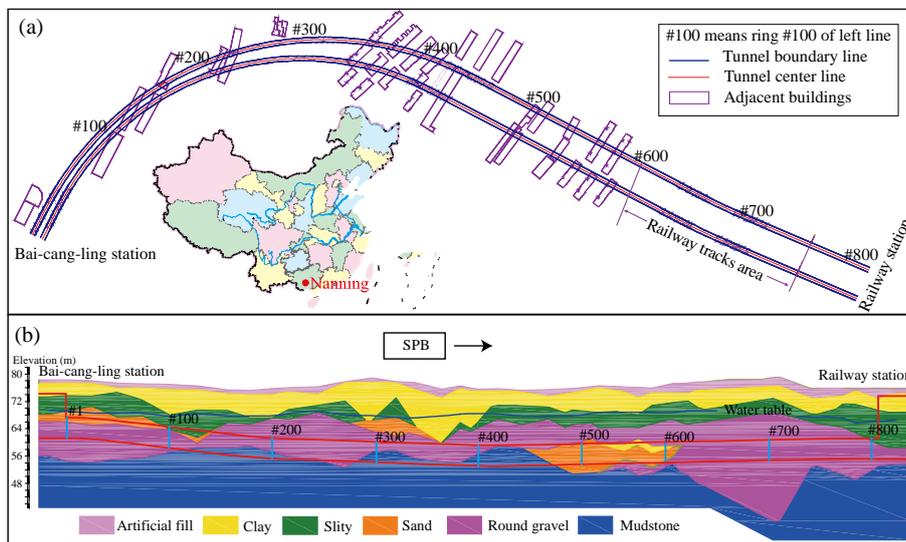


Figure 3. Schematic illustration of BR section of Nanning metro line 1, (a) plan view and parts of adjacent buildings, and (b) cross-section view. # 100 means ring # 100 in the left line

Table 1. Geotechnical characteristics of round gravel and mudstone in the BR section

Soil	$\rho$	$\nu$	$c$	$\phi$	$K$
	g/cm <sup>3</sup>		kPa	°	m/d
Round gravel	2.05	0.27	0.0	35.0	90
Mudstone	2.15	0.20	90.0	21.0	0.01

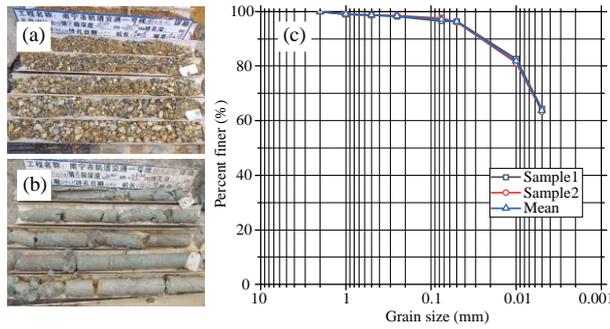


Figure 4. Borehole results of (a) round gravel and (b) mudstone, and (c) grain size distribution curve of mudstone in the BR section

According to the classification diagram raised by F. S. Hollmann and Thewes[1], the mudstone in the BR section is in the form of lumps, this may cause massive clogging because large lumps of this clay could form blockages in critical flow paths of an SPB shield machine. What's worse, to keep tunnel face stability of the water-rich round gravel, the marsh funnel viscosity of the slurry used in the SPB ranges from 24s to 32s, which has made the excavated material more easily clog at the opening of the submerged wall between the excavation chamber and the working chamber, as demonstrated in Figure 5.

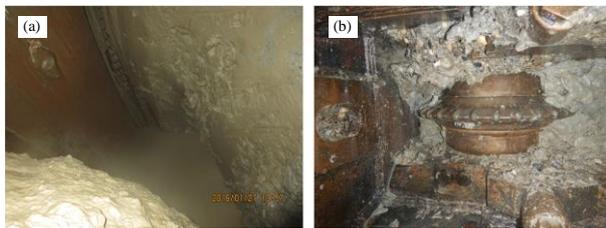


Figure 5. (a) clogging situation observed by opening the excavation chamber and (b) high wear of cutters

### 3.2 Data Description

The tunneling parameters in the BR section are recorded via the programmable logic controller (PLC)

every 10 seconds. There is a total of 666 complete ring data that will be analyzed in this study. Four kinds of tunneling parameters are selected for clogging prediction, which are the difference ( $\Delta SP$ ) between slurry pressure in the excavation chamber (SPE) and slurry pressure in the working chamber (SPW), the cutterhead torque (TOR), total thrust (THR), penetration rate (PR). The SPE and SPW are measured by pressure sensors locating at the spring line of both chambers. To elevate the influence of mudstone, a mixed ratio  $\lambda$  [4] is defined as Eq. (2).

$$\Delta SP = SPW - SPE \quad (1)$$

$$\lambda = H_m / D \quad (2)$$

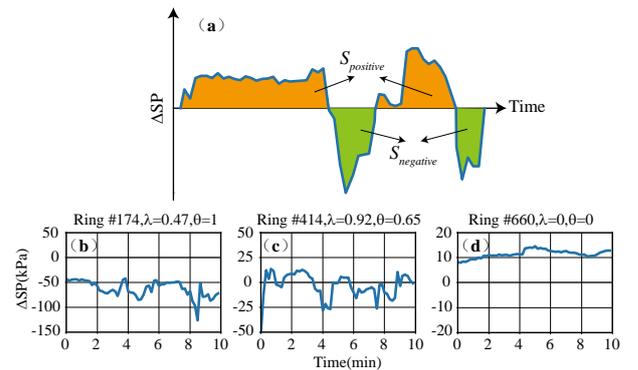
where  $H_m$  is the mudstone thickness in the tunnel excavation face and is determined by the geological report.  $D$  is the SPB cutterhead diameter.

Here we pay more attention to the instantaneous values of  $\Delta SP$ , so we define a slurry pressure fluctuation index (SPFI)  $\theta$ , as illustrated in Figure 6(a) and in E.q.(3). In normal conditions, the values of  $\Delta SP$  are in the range of 0 to 20 kPa (Figure 6(d)). However, when clogging occurs, the values of  $\Delta SP$  may range from -50 kPa to -150 kPa (Figure 6(b)). We plot 10 minutes of instantaneous values of  $\Delta SP$  and their SPFIs in Figure 6 considering different ground conditions.

$$S_{positive} = \frac{1}{2} \left( \int \Delta SP dt + \int |\Delta SP| dt \right)$$

$$S_{negative} = \frac{1}{2} \left( \int \Delta SP dt - \int |\Delta SP| dt \right) \quad (3)$$

$$\theta = \frac{abs(S_{negative})}{S_{positive} + abs(S_{negative})}$$


 Figure 6. Definition of slurry pressure fluctuation index and typical examples in different ground conditions (10 minutes data), (a) schematic diagram of  $\theta$ , (b) ring #174,  $\theta=1$  in half mudstone and half round gravel condition, (c) ring

#414,  $\theta=0.65$  in mudstone condition, and (d) ring #660,  $\theta=0$  in round gravel condition.

According to field observation, we plot the clogging rings distribution, as well as the SPFI  $\theta$  and mixed ratio  $\lambda$ , as shown in Figure 7. It can be found that there are dramatic fluctuations of SPE in most clogging rings as the values of  $\theta$  are all near one when clogging occurs. Also, the distribution of the mixed ratio is not always consistent with the clogging state, which may be caused by the uncertainty of the mudstone distribution in the geological survey as well as the mudstone that stuck to the cutterhead.

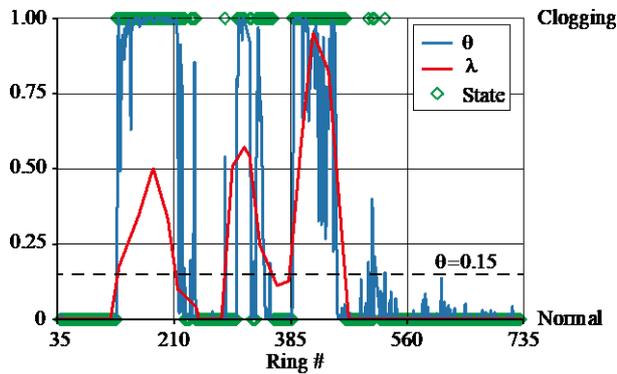


Figure 7. Distribution of mixed ratio, SPFI, and clogging state in the BR section

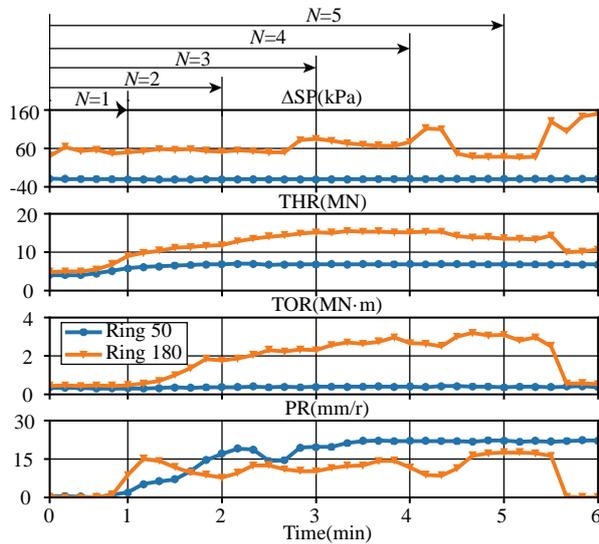


Figure 8. Demonstration of six minutes of tunneling data of clogging ring (ring #180) and normal ring (ring #50) at the beginning of each ring

Figure 8 demonstrates an example of the six minutes tunneling data that will be used as the input feature for

clogging prediction. We employ 17 kinds of features as the input, including the statistical indexes (mean values, standard deviations, maximum values, and range values) of the  $\Delta SP$ , THR, TOR, PR, as well as the SPFI ( $\theta$ ). We will explore the influence of parameter length  $N$  because the smaller of  $N$  value, the better for real-time early warning.

### 3.3 Clogging Prediction Results in the BR Section

We define the confusion matrix for clogging prediction in Table 2, where the P=Positive, that means the clogging state; the N=Negative, that means the normal state; TP=True Positive, that means the actual state of one ring is clogging and the predicted state is also clogging; FP=False Positive, that means the actual state of one ring is normal but the predicted state is clogging; TN=True Negative, that means the actual state of one ring is clogging but the predicted state is normal; FN=False Negative, that means the actual state of one ring is normal and the predicted state is also normal. We conduct the error rate, precision, recall, and F1 to evaluate the prediction model performance, as shown in E.q. (4). The high precision relates to a low false positive (FP) rate, and high recall relates to a low false negative (FN) rate. High scores for both show that the classifier is returning accurate results (high precision), as well as returning a majority of all positive (clogging) results (high recall).

$$\begin{aligned}
 \text{error rate} &= \frac{FP + FN}{TP + TN + FP + FN} \times 100\% \\
 \text{precision} &= \frac{TP}{TP + FP} \times 100\% \\
 \text{recall} &= \frac{TP}{TP + FN} \times 100\% \\
 F1 &= \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}
 \end{aligned} \tag{4}$$

Table 2. Confusion matrix for clogging prediction

		Actual class	
		P: Clogging	N: Normal
Predicted class	P: Clogging	TP	FP
	N: Normal	FN	TN

As mentioned before, there are 666 rings for analysis in the BR section, and 229 rings are regarded as clogging state (Figure 7). We choose the first 90 rings as the initial training data set and then examine the different tunneling data length ( $N=0.5, 1, 2, 3, 4, 5, 6, 8$ ). The hyperparameters ( $n\_estimators$  and  $max\_depth$ )

distribution for the RF model is the random number between 0 to 1000. To compare different model performances, we also employ the  $K$ -nearest neighbor (KNN) model[20], the support vector classification (SVC) model[21], and the multi-layer perceptron (MLP) model[22]. The hyper-parameters (  $n\_neighbors$  ) distribution for the KNN model is the random number between 0 to 10. The RBF function is adopted as the kernel function in the SVC model, which has a logarithmically spaced  $C$  values range from 1 to 1000 with a vector length of 20, a logarithmically spaced  $\gamma$  values range from 0.001 to 1000 with a vector length of 7. The MLP model has one hidden layer with different neurons from 5 to 25 with an increase of 5. The learning rate of the MLP model is set as a logarithmically spaced values range from 1 to 1000 with a vector length of 4.

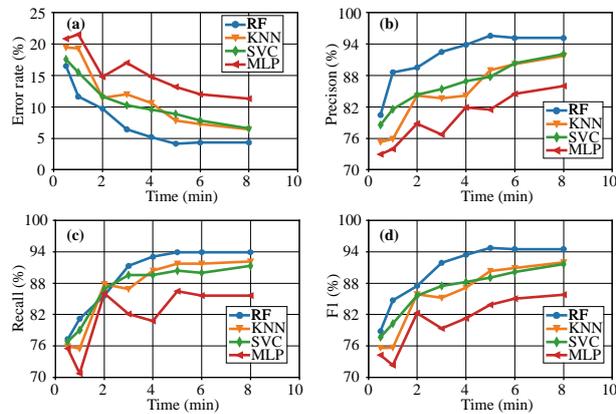


Figure 9. Clogging prediction model performances with different tunneling data length, (a) error rate, (b) precision, (c) recall, and (d) F1

After the simulation of 666 rings in the BR section, Figure 9 illustrates the metrics of the four clogging

prediction models with different tunneling data length. The model performance becomes better as the tunneling data length increases, and the RF model performs best among these four models considering all the metrics. The KNN model performance is similar to the SVC model while the MLP model performs worst with a very high error rate. From Figure 9 (a), it can be found that when we use four minutes of tunneling data to predict the clogging state, the RF model error rate is 5%, which seems well enough for tunneling construction. What's more, the four minutes tunneling data makes the RF model yield a high precision (Figure 9 (b)), recall (Figure 9 (c)) and F1 score (Figure 9 (d)), which are about 94%. With the increase of tunneling data length, the RF model performance increases slightly. Therefore, we believe the data length of four minutes at the beginning of one ring tunneling period is a good choice for clogging state prediction.

### 3.4 Discussion on Model Performance

#### 3.4.1 Model Performance without Training Data Updating

Table 3 lists the metrics of the RF clogging prediction model considering different training strategies. One strategy uses the updating training data set and the other one just employs the initial training data set without updating. We can see that the model with updating the training data set achieves a better prediction accuracy than the model with the static training data set. We can realize the early warning of clogging with the four minutes of tunneling data with an error rate of 5.2% in the proposed model. However, we need five to eight minutes of tunneling data to realize the same goal when the training data remains unchanged. One possible reason for this difference is that the RF model learns more patterns about clogging when the training data set is updated as shield machine advances.

Table 3. Metrics comparison of RF clogging prediction model considering different training strategies

Tunneling data length (minutes)	Error rate (%)		Precision (%)		Recall (%)		F1 (%)	
	Updating	No updating	Updating	No updating	Updating	No updating	Updating	No updating
0.5	16.5	20.7	80.5	80.2	77.3	62.8	78.9	70.4
1	11.6	18.6	88.6	77.5	81.2	75.2	84.7	76.3
2	9.7	24.5	89.5	64.3	85.6	84.5	87.5	73.0
3	6.4	12.5	82.5	80.1	91.3	91.2	86.7	85.3
4	5.2	9.7	93.8	84.6	93.4	92.7	93.6	88.5
5	4.2	5.4	95.6	94.1	93.9	92.8	94.7	93.4
6	4.3	5.5	95.1	95.2	93.9	92.6	94.5	93.9
8	4.3	5.2	95.1	94.2	94.0	92.8	94.6	93.5

### 3.4.2 Influence of Input Features on Clogging Prediction

In section 3.2, we employ 17 kinds of features as the clogging prediction model input. For better guidance for shield tunneling construction, here we carry out feature importance to investigate which kind of features has a crucial impact on clogging prediction. The permutation feature importance technology [23] is adopted with four minutes of tunneling data in the RF model. Figure 10 shows the feature importance result with the metric of error rate in different states. Firstly, we consider the whole tunnel section, and there are two kinds of input feature that has feature importance larger than 10%, which are the mean and maximum values of the TOR. Also, the maximum and range values of the PR should be considered. Then, if we just consider the normal state rings, the influence of PR seems quite important as the feature importances of maximum values, range values, and standard deviations of PR are larger than 10%. Finally, when we consider the clogging rings, it can be found that the feature importances of the mean and maximum values of the TOR are larger than 15%, which means the variation of TOR should be taken as the priority to make the judgment of clogging.

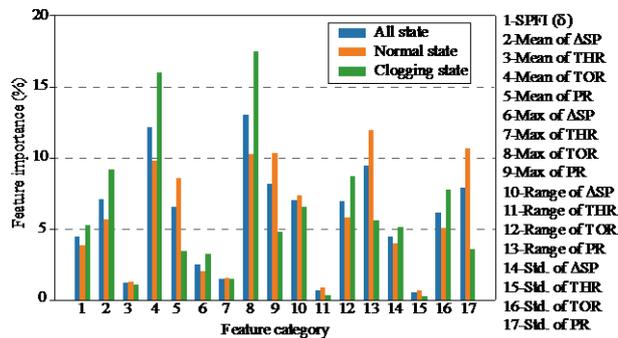


Figure 10. Feature importance of the RF model considering different states

## 4 Conclusion

This paper proposed a real-time early warning strategy of clogging risks based on RF. The clogging early warning model performance was evaluated. Results of conducted analyses show that:

(1) The RF model can realize the real-time early warning of clogging with four minutes of tunneling data at the beginning of one ring, which achieves an error rate of 5%. Compared with the KNN model, the SVC model, and the MLP model, the RF model gives the best performances.

(2) The statistical indexes of TOR are essential for clogging judgment while the PR is crucial for normal state identification based on the importance of the

explanatory features. Without updating the training data set, the RF model needs more tunneling data to make a good prediction of clogging.

Although we have realized some achievements in the analysis and prediction of the clogging risks based on the tunneling parameters, we still pay more attention to the control measures to mitigate the clogging problem. For example, more appropriate slurry properties design for mixed ground containing mudstone is vital, which can keep the tunnel face stability and convey the excavated material more smoothly. Besides, it's significant to carry out the intelligent models that can have good control of tunneling parameters in high clogging risk areas instead of the trial-error process based on field experience.

## References

- [1] F.S. Hollmann, M. Thewes, Assessment method for clay clogging and disintegration of fines in mechanised tunnelling, *Tunn. Undergr. Sp. Technol.* 37 (2013) 96–106.
- [2] D.G.G. de Oliveira, M. Thewes, M.S. Diederichs, Clogging and flow assessment of cohesive soils for EPB tunnelling: Proposed laboratory tests for soil characterisation, *Tunn. Undergr. Sp. Technol.* 94 (2019) 103110.
- [3] N. Tokgöz, I. Serkan Binen, E. Avunduk, An evaluation of fine grained sedimentary materials in terms of geotechnical parameters which define and control excavation performance of EPB TBM's, *Tunn. Undergr. Sp. Technol.* 47 (2015) 211–221.
- [4] X. Xie, Q. Wang, Z. Huang, Y. Qi, Parametric analysis of mixshield tunnelling in mixed ground containing mudstone and protection of adjacent buildings: case study in Nanning metro, *Eur. J. Environ. Civ. Eng.* 22 (2018) s130–s148.
- [5] J. Zhao, Q.M. Gong, Z. Eisensten, Tunnelling through a frequently changing and mixed ground: A case history in Singapore, *Tunn. Undergr. Sp. Technol.* 22 (2007) 388–400.
- [6] L. Bo, B. Qin, Key Techniques for Construction of Sanyang Road Cross-river Tunnel of Wuhan Rail Transit Line 7, *Tunn. Constr.* 39 (2019) 820–831.
- [7] Q. Wang, X. Xie, Z. Huang, Y. Qi, Protection of adjacent buildings due to mixshield tunnelling in mixed ground with round gravel and mudstone, in: *Proc. World Tunn. Congr. 2017 – Surf. Challenges – Undergr. Solut.*, Bergen, Norway, 2017: pp. 14635-1–10.
- [8] X. Ye, S. Wang, J. Yang, D. Sheng, C. Xiao, Soil conditioning for EPB shield tunneling in argillaceous siltstone with high content of clay

- minerals: Case study, *Int. J. Geomech.* 17 (2017) 1–8.
- [9] F. Min, W. Zhu, C. Lin, X. Guo, Opening the excavation chamber of the large-diameter size slurry shield: A case study in Nanjing Yangtze River Tunnel in China, *Tunn. Undergr. Sp. Technol.* 46 (2015) 18–27..
- [10] R. Zumsteg, A.M. Puzrin, G. Anagnostou, Effects of slurry on stickiness of excavated clays and clogging of equipment in fluid supported excavations, *Tunn. Undergr. Sp. Technol.* 58 (2016) 197–208.
- [11] F.S. Hollmann, M. Thewes, Assessment method for clay clogging and disintegration of fines in mechanised tunnelling, *Tunn. Undergr. Sp. Technol. Inc. Trenchless Technol. Res.* 37 (2013) 96–106.
- [12] N. Tokgöz, I. Serkan Binen, E. Avunduk, An evaluation of fine grained sedimentary materials in terms of geotechnical parameters which define and control excavation performance of EPB TBM's, *Tunn. Undergr. Sp. Technol.* 47 (2015) 211–221.
- [13] J.N. Shirlaw, Pressurised TBM tunnelling in mixed face conditions resulting from tropical weathering of igneous rock, *Tunn. Undergr. Sp. Technol.* 57 (2016) 225–240.
- [14] C. Budach, D. Placzek, E. Kleen, Quantitative Bestimmung des Verklebungspotenzials feinkörniger Böden auf Basis von Adhäsionsspannung: Aktuelle Untersuchungen und neue Erkenntnisse, *Geotechnik.* 42 (2019) 2–10.
- [15] S.E. Seker, I. Oca, Performance prediction of roadheaders using ensemble machine learning techniques, *Neural Comput. Appl.* 31 (2019) 1103–1116.
- [16] V.R. Kohistani, M. Hassanlourad, A. Ardakani, Evaluation of liquefaction potential based on CPT data using random forest, *Nat. Hazards.* 79 (2015) 1079–1089.
- [17] J. Zhou, X. Li, H.S. Mitri, Classification of rockburst in underground projects: Comparison of ten supervised learning methods, *J. Comput. Civ. Eng.* 30 (2016) 1–19.
- [18] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, *J. Mach. Learn. Res.* 13 (2012) 281–305.
- [19] L. Breiman, Random forests, *Mach. Learn.* 45 (2001) 5–32.
- [20] R. Akhavian, A.H. Behzadan, Smartphone-based construction workers' activity recognition and classification, *Autom. Constr.* 71 (2016) 198–209.
- [21] Q. Zhang, Z. Liu, J. Tan, Prediction of geological conditions for a tunnel boring machine using big operational data, *Autom. Constr.* 100 (2019) 73–83.
- [22] C.O. Sakar, S.O. Polat, M. Katircioglu, Y. Kastro, Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks, *Neural Comput. Appl.* 31 (2019) 6893–6908.
- [23] A. Fisher, C. Rudin, F. Dominici, All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously, *J. Mach. Learn. Res.* 20 (2019) 1–81.