Engineering Applications of Artificial Intelligence 114 (2022) 105020

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Prediction of ground surface settlement by shield tunneling using XGBoost and Bayesian Optimization

Jie Su [a,*], Yuzhe Wang [a], Xiaokai Niu [b], Shan Sha [a], Junyu Yu [a]

[a] Key Laboratory of Urban Underground Engineering of Ministry of Education, Beijing Jiaotong University, Beijing 100044, China
[b] Beijing Municipal Engineering Research Institute, Beijing 100037, China

ARTICLE INFO

Keywords:
Ground surface settlement prediction
Shield tunneling
XGBoost model
Bayesian optimization
Model interpretability

ABSTRACT

Ground surface settlement caused by shield tunneling is a complex problem caused by multiple factors. Machine learning models can help in the nonlinear intelligent prediction of ground surface settlement caused by shield tunneling. At present, the Artificial Neural Network model and the Support Vector Machine model are the most widely used models for settlement prediction. Due to the black-box characteristics of these two models, they are inherently deficient in interpretability, which means it is difficult to provide guidance for engineering. To solve the problem of poor interpretability in the prediction of ground surface settlement using these two models, an ensemble learning algorithm called the XGBoost model is introduced. In order to select hyperparameters in XGBoost more efficiently, the Bayesian optimization is used for parameter search. In this study, 533 cases of ground surface settlement monitoring data from a shield tunnel construction project in a city were used. Compared with the prediction results of the ANN model and the SVM model, the XGBoost model has the advantages of prediction accuracy and interpretability, especially for the prediction of out-of-limit settlement points.

1. Introduction

The shield method is a commonly used method in the process of tunnel excavation. The shield machine will inevitably disturb the surrounding soil during the excavation process of the ground. The disturbed soil is subjected to squeezing, shearing, and twisting, resulting in elastic–plastic deformation or even destruction, which eventually causes ground surface settlement.

The main influence factors of ground settlement include tunnel parameters, geotechnical properties, and construction effects. There is a complex nonlinear relationship between ground settlement and these factors. Tunnel parameters mainly include the buried depth, shape and excavation area of the tunnel. In generally, the ground settlement will become greater with a shallower buried depth and larger excavation section (Mair et al., 1993). Geotechnical properties refer to the physical properties of the strata, which are the decisive factors of the ground settlement. If the strata are soft, it will be very difficult to control the ground settlement. The construction effects refer to the human factors during the tunnel construction. Taking the shield tunnel as an example, the ground settlement can be controlled by adjusting the face pressure, the excavation speed or the amount of grouting during the tunnel construction process.

Accurate prediction of the ground surface settlement is of great significance for engineering safety. The traditional prediction methods in the engineering field mainly include empirical formula method, analytical method and numerical simulation method.

The empirical formula is an approximate expression obtained by fitting the field measured data. The Peck formula is the most classic empirical formula to describe the ground surface settlement, and the Gaussian function is used to fit the settlement curve (Peck, 1969). In the Peck formula, the maximum ground surface settlement is related to the ground loss ratio, the internal friction angle of the soil, and the buried depth of the tunnel. After that, some researchers further modified the Peck formula in different situations (O'Reilly and New, 1982; Mair et al., 1993; Wang, 2014). The strata parameter selection of the empirical formula varies greatly from region to region. And the construction effects cannot be considered in empirical formula.

Based on the mechanics principle and certain assumptions, the analytic method could analyze the ground settlement caused by tunnel excavation. The researchers used multiple analytic methods such as the mirror image method, stress function method, complex variable function method, random medium theory and energy conservation method to theoretically analyze the ground settlement caused by tunnel excavation (Loganathan, 1998; Bobet, 2001; Cao, 2020). The analytic method is relatively simple in form and limited in influencing factors, so it is difficult to describe the complex nonlinear relationship between the tunnel and the ground settlement.

* Corresponding author.
E-mail address: sujie@bjtu.edu.cn (J. Su).

https://doi.org/10.1016/j.engappai.2022.105020
Received 7 April 2022; Received in revised form 22 May 2022; Accepted 1 June 2022
Available online 21 June 2022

The numerical simulation method uses the finite element software to discretize the strata and analyze it. Compared with other methods, the numerical simulation method can consider construction process and the nonlinear characteristics of the soil, so researchers have widely used it in the prediction of ground surface settlement (Migliazza et al., 2009; Hasanpour et al., 2014; Shivaei et al., 2020). However, the constitutive model of the soil in the numerical simulation is quite different from the actual soil and the calculation results of the numerical simulation are easily affected by the grid division, which leads to poor prediction accuracy.

In short, the ground settlement caused by tunnel excavation is a complex multi-factor nonlinear problem. There are many deficiencies in traditional prediction methods, which inevitably leads to poor prediction accuracy. Machine learning is a method using computers to achieve intelligent data analysis. Due to the stronger high-dimensional nonlinear fitting ability of the Machine learning method, it has more advantages in settlement prediction compared with other traditional methods.

The neural network algorithm is the machine learning algorithm that is most widely applied to predict ground surface settlement. Back-propagation neural network (BPNN) model was the first neural network algorithm introduced into predict ground surface settlement (Sun and Yuan, 2001; Suwansawat and Einstein, 2006). Some researchers optimized BPNN model and improved the prediction effect of the model (Pourtaghi and Lotfollahi-Yaghin, 2012; Hasanipanah et al., 2016; Qiao et al., 2012; Tian et al., 2017). Since BPNN has the problem of slow learning speed and being easy to fall into the local minimum value, the Adaptive Neuro-Fuzzy Inference System (ANFIS) model was proposed to replace BPNN for settlement prediction (Hou and Zhang, 2009; Bouayad and Emeriault, 2017). Zhou (2011), has adopted another neural network algorithm, RBFNN, to achieve the ground surface settlement.

The support vector machine (SVM) model is another algorithm widely applied to predict ground surface settlement. Compared with the neural network, SVM model, which takes structural risk minimization criteria as the optimization objective, improves the model's generalization ability and achieves good prediction results on small samples. Bai et al. (2013), has applied the SVM model to predict the ground surface settlement caused by shield tunneling, by taking geological parameters, shield buried depth and ground loss as input parameters. Ocak and Seker (2013), has compared the prediction effect of SVM model with ANN model in settlement prediction and found that the prediction effect of SVM model is better. In order to further improve the prediction effect of ground surface settlement, some optimization methods such as wavelet analysis and rough set theory were combined with SVM model (Zhang et al., 2017; Lin et al., 2018).

There are two main problems in the existing studies. First, the models lack interpretability. Both the ANN model and the SVM model belong to the "black box" model type, and there is a lack of correlation between the models' prediction results and the influencing factors, making it difficult to guide practical engineering. Second, the models are prone to suffer "over-fitting", leading to poor prediction effects. In existing studies, one of the reasons why the models easily suffer "over-fitting" is the insufficient sample data (usually less than 50), and it is difficult to train a model with a stronger generalization ability by using a small number of samples. Another reason is the lack of effective optimization of the algorithm, which further affects the prediction accuracy of the model.

To overcome the above problems, in this study, an ensemble learning algorithm called the XGBoost model is used to predict ground surface settlement. The settlement prediction model is established based on the measured data from the Chengdu metro project. Compared with the prediction results of existing models, this new model has obvious advantages in terms of prediction accuracy and model interpretability.

2. XGBoost model and Bayesian Optimization

The settlement prediction model should clearly reflect the nonlinear relationship between the influence factors and settlement, which means that the prediction model should have strong interpretability. Therefore, in this paper, an ensemble learning algorithm called the Extreme Gradient Boosting model (XGBoost) is proposed as the learner to establish the prediction model for ground surface settlement caused by shield tunneling. Furthermore, we adopt Bayesian optimization to overcome the complex hyperparameter selection problem in XGBoost.

2.1. Algorithm principle for the XGBoost model

Ensemble learning is a method of combining multiple weakly supervised models to get a strong supervised model. XGBoost (Extreme Gradient Boosting) is an ensemble learning algorithm based on boosting strategy (Chen and Guestrin, 2016). The boosting learning strategy of the XGBoost model is shown in Fig. 1. The base learner (the CART model)is a weak learner, and a strong learner is finally formed by analyzing the learning error (e_i) of the weak learner and updating the sample weight during each iteration.

The objective function is the optimization solution object in the model. In the XGBoost model, the objective function is defined as:

$$obj = \sum_{i=1}^{m} L\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{n} \Omega\left(f_k\right) \tag{1}$$

It can be seen from Eq. (1) that the objective function is composed of two parts. The first item $\sum_{i=1}^{m} L\left(y_i, \hat{y}_i\right)$ is the loss function, which measures the difference between the true value and the predicted value and represents the prediction error. The settlement prediction problem considered in this paper is a regression problem, so the loss function is denoted by the mean square error. The second term $\sum_{k=1}^{n} \Omega\left(f_k\right)$ is the regularization penalty term, which is used to optimize the model's complexity. The higher the model's complexity, the more specific the learning will be, which means that the prediction ability for different data sets will be worse. The model's generalization ability can be significantly improved using the regularization penalty term. The XGBoost model introduces the regularization term to the objective function, therefore, it has a better generalization ability than other models.

2.2. XGBoost model hyperparameters

The prediction model's accuracy is closely related to the model's hyperparameters. The XGBoost model is a tree model and its hyperparameters are divided into structure parameters and regularization parameters. Structure parameters include max_depth (the maximum depth of the tree model), minimal_child_weight (the minimum number of leaf node samples required for branch splitting), and n_estimators (the iteration times of the tree model). The structure parameters can control the complexity of the model. The larger the three structure parameters of the tree model are, the more complex the model is, and the better its data fitting ability is. However, if the three parameters are too large, it leads to over-fitting of the model.

In the XGBoost model, the regularization term is used to adjust the model's complexity. Through the regularization term, the model can be effectively optimized, and the generalization ability of the model can be improved. The expression of the regularization term is shown in Eq. (2):

$$\Omega\left(f_t\right) = \gamma T + 1/2\alpha \sum_{j=1}^{T} \left|w_j\right| + 1/2\lambda \sum_{j=1}^{T} w_j^2 \tag{2}$$

Where γT controls the tree depth, T represents the leaf nodes number in the tree structure. γ is the parameter for controlling the leaf node. $1/2\alpha \sum_{j=1}^{T} \left|w_j\right|$ and $1/2\lambda \sum_{j=1}^{T} w_j^2$ denote L1 regularization and L2 regularization terms, respectively, with α and λ representing regularization coefficients. It can be seen from Eq. (2) that the three regularization parameters are gamma (leaf node branch parameter), reg_alpha (L1 regular term penalty coefficient), and reg_lambda (L2 regular term penalty coefficient), respectively.

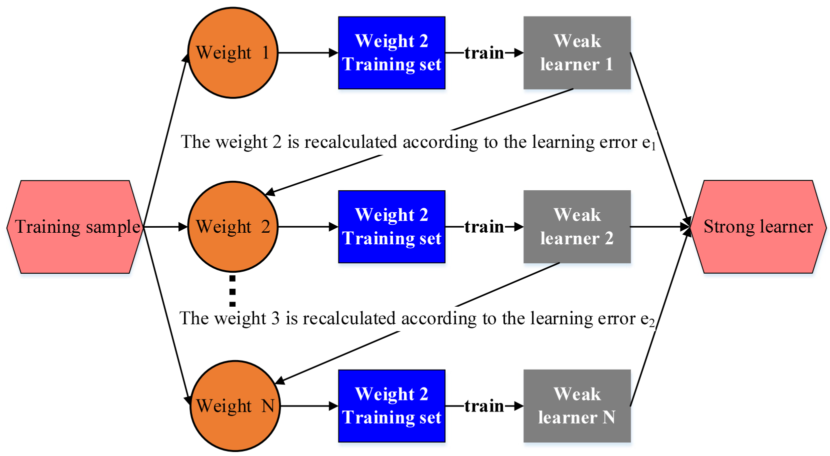


Fig. 1. Boosting learning strategy.

2.3. Bayesian optimization of the model's hyperparameters

In machine learning, the selection of hyperparameters is closely related to the model effect. It is a very important topic to select the most appropriate hyperparameters efficiently and accurately based on known data. To address this problem, Bayesian optimization gives a powerful solution (Shahriari et al., 2015). The Bayesian optimization method obtains prior information from the existing parameter selection results and continuously updates the probability distribution shown by the given hyperparameters in the objective function to guide the selection of the next parameter. In this paper, Bayesian optimization based on Gaussian process is applied to the XGBoost method in order to accomplish the selection of hyperparameters more efficiently.

The optimization process of the model hyperparameters using the Bayesian algorithm is shown in Eqs. (3) and (4):

$$p\left(f \middle| R_{1:t}\right) = \frac{p\left(R_{1:t} \middle| f\right) p(f)}{p\left(R_{1:t}\right)} \tag{3}$$

$$R_{1:t} = \left\{\left(x_1, y_1\right), \left(x_2, y_2\right) \cdots \left(x_t, y_t\right)\right\}$$
$$y_t = f\left(x_t\right) + \varepsilon_t \tag{4}$$

Where f represents the objective function; y_t represents the observed value for the t-th step; x_t represents the hyperparameter of the t-th step; ε_t represents the observation error; $R_{1:t}$ represents the aggregation of observations for the previous t steps; $p\left(R_{1:t} \middle| f\right)$ represents the likelihood distribution of y_t; $p(f)$ represents the prior distribution of f. In the hyperparameter optimization process, $p(f)$ can be regarded as the state assumption for the objective function. $p\left(f \middle| R_{1:t}\right)$ represents the posterior distribution of the objective function.

Here the Gaussian process is used as a surrogate model, which is mainly determined by the mean and covariance functions.

$$f(x) \sim \mathrm{GP}\left(m(x), k\left(x, x'\right)\right) \tag{5}$$

$m(x)$ is the mean value function, $k\left(x, x'\right)$ is the covariance function.

$$m(x) = \mathrm{E}[f(x)] \tag{6}$$

$$k\left(x, x'\right) = \mathrm{E}\left[(f(x) - m(x))\left(f\left(x'\right) - m\left(x'\right)\right)\right] \tag{7}$$

3. Database construction based on practical engineering

Generally speaking, the larger the data amount and the more significant the data differences in the database for the model construction, the better the model learning effect becomes. There is a general problem of insufficient data in existing studies (Pourtaghi and Lotfollahi-Yaghin, 2012, 49 data sets; Hasanipanah et al., 2016, 143 data sets; Tian et al., 2017, 53 data sets). Based on the measured data from the Chengdu metro project, in this study, a database containing the monitoring data of 533 cases of ground surface settlement monitoring points was established.

3.1. Engineering background

The data in this paper are from the shield tunnel of the Chengdu metro project. This tunnel is a double-track tunnel, with a total length of about 1.96 km. The shield machine diameter is 6.4 m, and the burial depth is about 9–22 m. The tunnel mainly passes through strata including moderately weathered sandstone, moderately weathered mudstone, plain fill and sandy pebble strata etc. Some tunnel sections pass under roads, bridges, and rivers, making it difficult for engineering construction. The tunnel consists of three sections, and the geological distributions are shown in Fig. 2.

3.2. Data collection

The layout of the points measuring the ground surface settlement is shown in Fig. 3. Along the tunnel direction, one measuring point was laid out every 5 m on the central tunnel axis and one monitoring section was arranged every 30 m. 12 measuring points were arranged on each monitoring section, and the distance between measuring points in each monitoring section is 3 m. The database for the prediction model in this paper was established by collecting settlement data from the measuring points and the data of the factors influencing the settlement. To ensure data consistency, only the measuring points on the tunnel's central axis were selected as the sample points for the settlement data, and a total of 533 cases of settlement monitoring data were collected.

3.3. Feature selection for the prediction model

The surface settlement caused by shield tunneling is mainly affected by factors such as the geotechnical properties, tunnel parameters, and construction effects, so these factors should be considered in the input parameters of the prediction model. According to the monitoring data collection, a total of 11 parameters are selected as input features of the model, including tunnel cover depth (CD), face pressure (FP), cutterhead torque (CT), shield thrust (ST), tunneling velocity (TV), earth volume (EV), grouting amount (GA), grouting pressure (GP), risk sources (RS), cohesion (C) and internal friction angle (φ) of the strata. The description of each variable definition, range and category are shown in Table 1 and part of the data are shown in Table A.1 at the end of the text.

Among them, the construction effect parameters (face pressure, cutterhead torque, shield thrust, tunneling velocity, earth volume, grouting amount and grouting pressure) are obtained through the monitoring

(a) Section 1

(b) Section 2

(c) Section 3

Fig. 2. Geological distributions of the project.

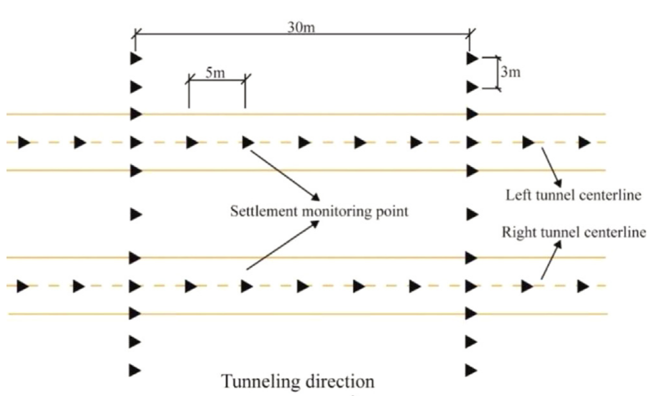


Fig. 3. Layout of the monitoring points.

Table 1
Description of variable definitions, ranges, and categories.

Parametric properties	Symbol	Unit	Parameter description (min–max)	Mean	Category
Tunnel parameter	CD	m	9~23.08	13.67	Input
Construction parameters	FP	Mpa	0.46~1.44	0.90	Input
	ST	MN	5.62~19.8	10.89	Input
	CT	kN m	917~4727	2828.77	Input
	TV	mm/min	20~60	39.66	Input
	EV	m^3	45~65	56.71	Input
	GA	m^3	4.5~10.8	6.31	Input
	GP	Mpa	1~5.6	2.32	Input
Geotechnical parameters	RS	–	0~1	0.06	Input
	C	kPa	5~1230	466.50	Input
	φ	°	8~45	33.46	Input
	S	mm	−52.54~5.39	−4.23	Output

data. Considering the special circumstances of crossing roads and rivers in the project construction process, the risk source near the tunnel is taken as an input feature in the form of a dummy variable. Geotechnical parameters are generally considered to be the most important parameters in affecting surface deformation. Based on existing studies (Qiao et al., 2012) and the geological prospecting reports, the geotechnical properties of the strata where the tunnel is located are represented by

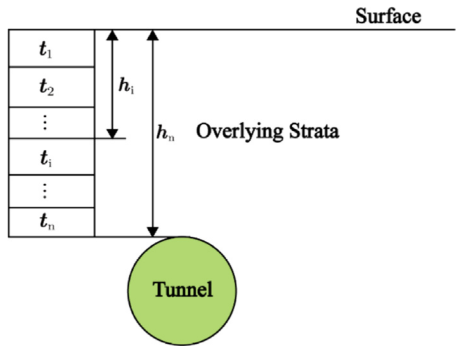


Fig. 4. The schematic view of geological parameter calculation.

the strata cohesion (C) and internal friction angle (φ), and the two parameters are weighted by considering the location and thickness of the overlying strata. The calculation method of geotechnical parameters is as follows.

For the overlying strata, defining the modified parameters of the ith strata:

$$\alpha_i' = t_i/h_n \cdot h_i/h_n \cdot \alpha_i \tag{8}$$

In the equation, α_i is the parameter of the strata in the geological prospecting report. t_i is the thickness of the ith strata. h_i is the depth of the ith strata. h_n is the buried depth of the tunnel. t_i/h_n is thickness coefficient. h_i/h_n is the distance coefficient. The thicker the ith strata is and the closer it is to the tunnel, the greater the coefficients will be. Fig. 4, shows the schematic view of geological parameter calculation.

The weighted average of the strata parameters is used as the input strata parameter of the model, as shown in Eq. (9).

$$\overline{\alpha} = \sum_{i=1}^{n} \alpha_i' \tag{9}$$

4. Prediction performance analysis

4.1. Establishment of prediction model

The prediction model in this paper is based on the Scikit-Learn library, which is the most widely used machine learning library in the Python language, in the Python 3.7 platform. In this study, the XGBoost algorithm was applied to establish a prediction model of ground surface settlement caused by shield tunneling. This model is compared with the two classical prediction models, the ANN model and the SVM model.

The model building process involves data preprocessing, division of training sets and test sets, training model creation, the output of results, and model evaluation. The data needs to undergo normalized pretreatment before input into the model for eliminating the dimensional influence and enabling data comparability. In this study, the StandardScaler module was used to normalize the data through the standard deviation standardization method, and the test set and training set were divided through the hold-out method. 70% of the data were randomly selected as the training set and the remaining 30% as the test set. We used the same dataset to build ANN model, SVM model, XGBoost model and optimized the hyperparameters of each model.

The considered hyperparameters of the ANN model are the number of hidden layers, the number of nodes, and the type of activation function. In the existing research, the activation function of the model is usually specified as sigmoid function (Ocak and Seker, 2013), the number of hidden layers and nodes is determined by verifying the performance of different models. Suwansawat and Einstein (2006), has established 18 models according to different number of hidden layers and nodes. By comparing the errors of different models, the optimal network with 1 hidden layer and 20 nodes were obtained. Hasanipanah et al. (2016), has established a single-layer neural network and compared the performance of the models under different number of nodes. In this study, these three hyperparameters are considered and obtained by grid search method. Table 2 shows the search scope of hyperparameters for ANN model. According to the grid search results, the hyperparameters of the ANN model were finally determined as: hidden layers = 1; nodes = 5; activation = "logistic".

Table 2
Search scope of hyperparameters for ANN model.

Hyperparameters	Illustration	Search scope of parameters
hidden_layer_sizes	The number of hidden layers and nodes	hidden layers (1,2,3) nodes (1,2,3,...,10)
activation	Activation function for the hidden layer	"identity", "logistic", "tanh", "relu"

Table 3
Search scope of hyperparameters for SVM model.

Hyperparameters	Illustration	Search scope of parameters
kernel	Specifies the kernel type to be used in the algorithm	"linear", "poly", "rbf", "sigmoid"
gamma	Kernel coefficient for 'rbf', 'poly' and 'sigmoid'	(0.1,0.2,0.3,...,1)
C	Regularization parameter	(1,2,3,...,100)

For the SVM model, the influence of the three hyperparameters "the kernel type", "the kernel function coefficient (gamma) ", and "the penalty coefficient (C) " are usually considered in the existing research. Bai et al. (2013), has used the RBF function as the kernel function to establish the SVM model, and determined the values of the hyperparameters "C" and "gamma" by the grid search method. Zhang et al. (2017) has optimized the SVM model hyperparameters "C" and "gamma" by using the genetic algorithm. In this study, these three hyperparameters are considered and obtained by grid search method. Table 3 shows the search scope of hyperparameters for SVM model. According to the grid search results, the hyperparameters of the SVM model were finally determined as: kernel = "rbf"; gamma = 0.1; C = 35.

Since more hyperparameters are considered in the XGBoost model, it is computationally difficult to optimize the hyperparameters by using the grid search method. In this study, the Bayesian optimization method was used to optimize the hyperparameters of XGBoost model. Compared with grid search method, Bayesian optimization method greatly improves the search efficiency of parameters.

The Bayesian optimization is implemented through the Skopt package. According to the introduction of XGBoost model hyperparameters in Section 2.2, we have optimized the six hyperparameters of the model: "n_estimators", "max_depth", "min_child_weight", "gamma", "reg_alpha" and "reg_lambda". Table 4 shows the search scope of hyperparameters for XGBoost model.

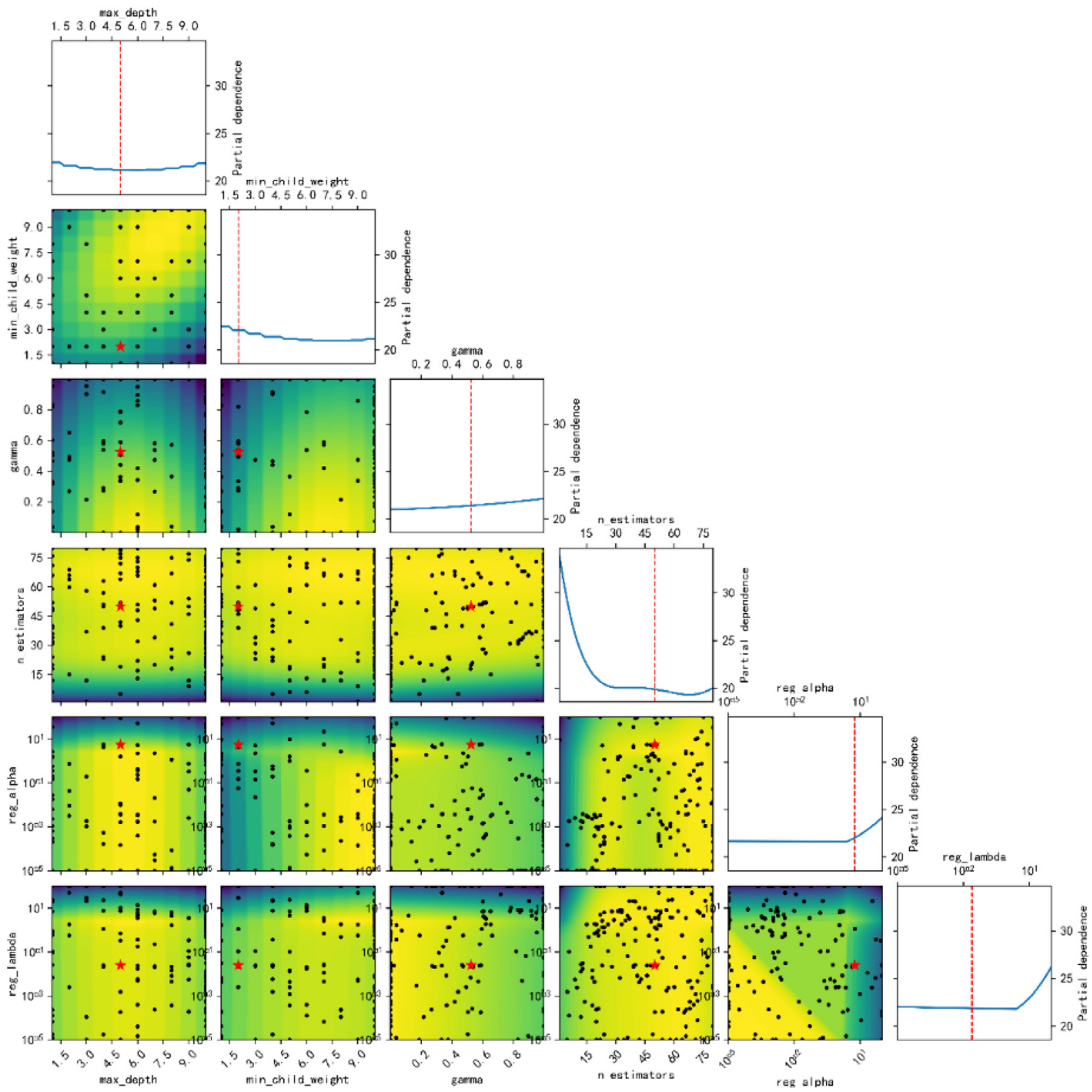


Fig. 5. Bayesian search for parameters of XGBoost.

Table 4
Search scope of hyperparameters for XGBoost model.

Hyperparameters	Illustration	Search scope of parameters
n_estimators	Iteration boosting number of decision tree	(1,2,3,...,100)
max_depth	Maximum depth of decision tree	(1,2,3,...,10)
min_child_weight	The minimum number of leaf node samples	(1,2,3,...,10)
gamma	Leaf node branch parameter	(0.1,0.2,0.3,...,1)
reg_alpha	L1 regular term penalty coefficient	(1e−5,1e−4,1e−3,...,100)
reg_lambda	L2 regular term penalty coefficient	(1e−5,1e−4,1e−3,...,100)

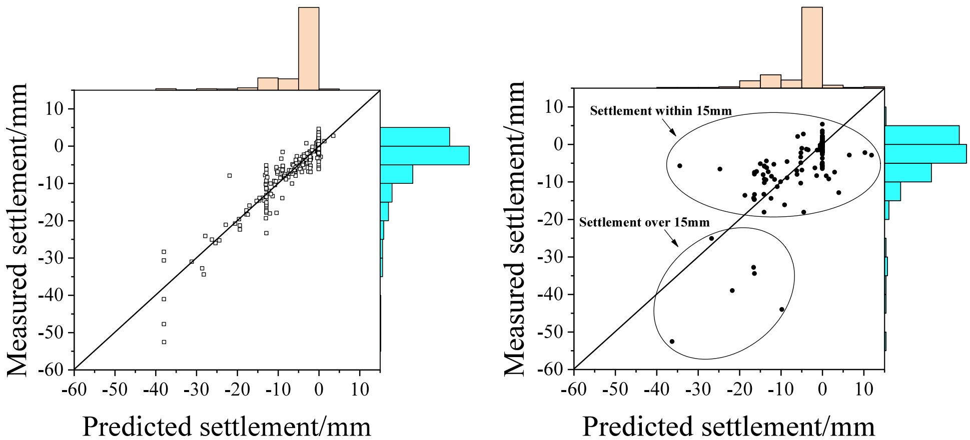
The Bayesian search process of the model's parameters is shown in Fig. 5. The parameter with the minimum mean square error was selected as the hyperparameter of the model. The coordinate axes in the figure represent the search range of different parameters. The black dots represent the sampling points of the Bayesian optimization search process, and the red star represents the optimal parameter points in the search results.

According to the Bayesian search results of the hyperparameters, the model's hyperparameters were finally determined as: n_estimators = 50; max_depth = 5; min_child_weight = 2; gamma = 0.5; reg_lambda = 1e-1; reg_alpha = 1e-0. The XGBoost model was established based on these parameters.

4.2. Error evaluation of the prediction results on ground surface settlement

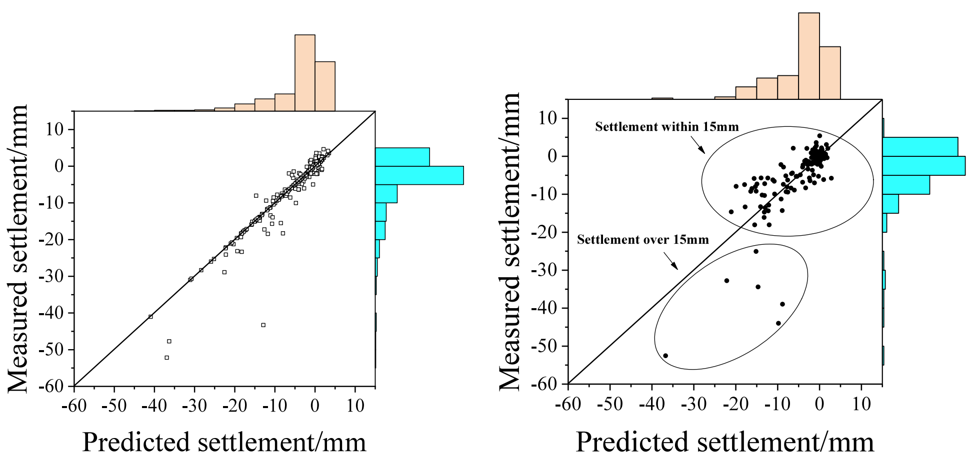
The measured data collected in this study are settlement values of the central tunnel axis directly above the shield tunneling cross-section. Therefore, the prediction model obtained in this study was only used to predict the ground surface settlement at the center of the tunnel section. According to the method mentioned above, the ANN model, SVM model and XGBoost model were established respectively, with 373 cases of data for the training set and 160 cases of data for the test set.

Figs. 6–8 show the prediction results distribution of settlement value range by the ANN model, the SVM model, and the XGBoost model, respectively. The closer the scatter in the figure is to the oblique line, the closer the predicted value is to the true value. The prediction results on the training set represent the model's ability to fit the in-sample data. From the prediction results on the training sets by the three models shown in Fig. 6(a), 7(a) and 8(a), it can be seen that the prediction results on the training sets by the three models are all



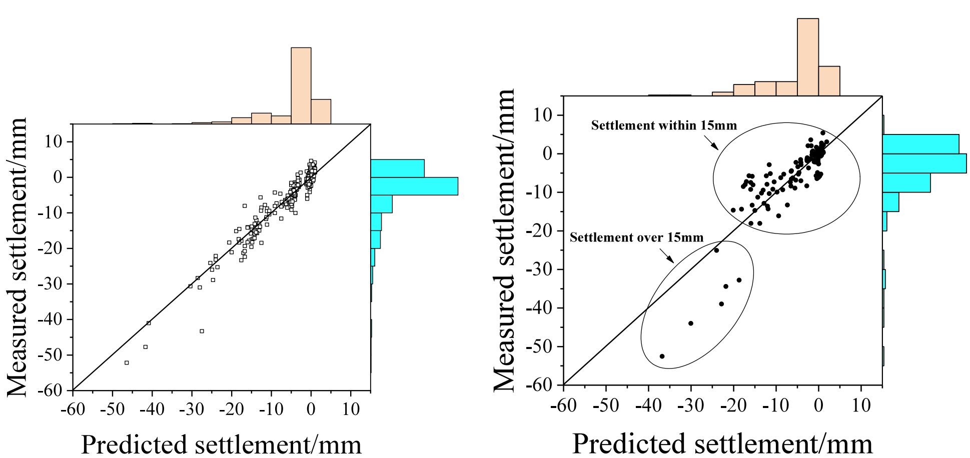
(a) Results on the training set (b) Results on the test set

Fig. 6. Prediction results by ANN model.



(a) Results on the training set (b) Results on the test set

Fig. 7. Prediction results by SVM model.



(a) Results on the training set (b) Results on the test set

Fig. 8. Prediction results by the XGBoost model.

satisfied, reflecting that the models all have strong fitting ability. The prediction results on the test sets represent the models' ability to predict out-sample data, which can better reflect the models' prediction effects. As can be seen from the prediction results on the test sets by the models

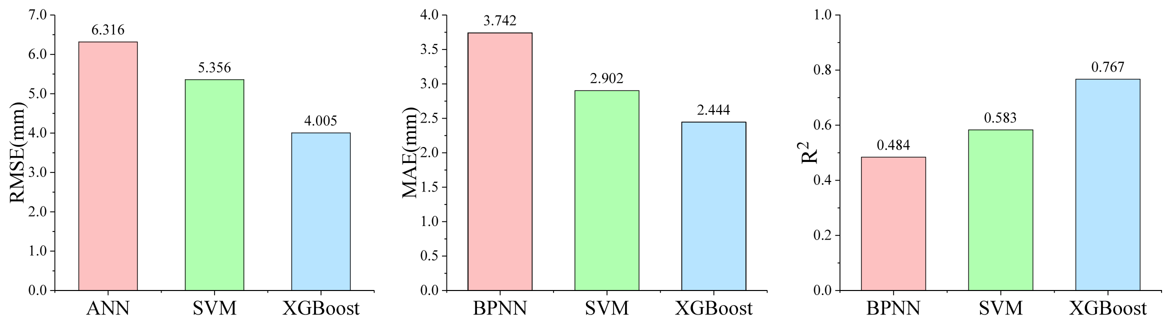


Fig. 9. Model evaluation indexes.

in Fig. 6(b), 7(b) and 8(b), The prediction result of the XGBoost model is closest to the measured values, while the prediction results of ANN model have the largest gap with the measured values.

The settlement data are divided into two parts, which include the settlement over 15 mm and the settlement within 15 mm, as shown in the circle in Fig. 6(b), 7(b) and 8(b). It can be seen that the prediction performance of the ANN model in the two parts of the data are both poor. The SVM model has a larger error in the measurement points with a settlement value greater than 15 mm, and the XGBoost model has better prediction results in the two parts of the data. The reason for the different performance of SVM model in the prediction results of two parts of data is that the data with the settlement within 15 mm account for the vast majority, which leads to the learning weight of this part of data are increased. Therefore, the prediction results based on these data are better, while the prediction effect for some outliers (settlement over 15 mm) is poor. Due to its ensemble learning characteristic, the XGBoost model shows the reinforcement learning ability for a few outliers, so it still has a relatively reliable prediction effect on this part of the data. In engineering applications, the out-of-limit settlement points are always the focus of engineers, so the XGBoost model has more significant engineering application value.

In order to effectively evaluate the prediction effect of the model, the root mean square error (RMSE), mean absolute error (MAE), and goodness of fit (R^2) are used as evaluation indexes for prediction models. RMSE and MAE represent the error between the predicted value and the monitoring value, whose unit is mm. R^2 represents the correlation between the features of the model and the prediction target. The calculation equations of the evaluation indexes are presented as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{10}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{11}$$

$$\mathrm{R}^2 = \frac{\sum\left(\hat{y}_i - \overline{y}\right)^2}{\sum\left(y_i - \overline{y}\right)^2} \tag{12}$$

where y_i represents the observed value, and $\hat{y}_i$ represents the predicted value of the model. $\overline{y}$ represents the average of the observed values, and n denotes the number of samples in the training or testing stages.

The three indexes for the prediction results on the test sets of the three models are summarized and shown in Fig. 9. The RMSE and MAE reflect the prediction error of models and their values for the three models are: XGBoost < SVM < BPNN. This indicates that the XGBoost model has the smallest prediction error (RMSE = 4.005; MAE = 2.444). The prediction error of the SVM model is smaller than that of the ANN model, which is consistent with existing research conclusions (Ocak and Seker, 2013). In terms of the goodness of fit(R^2), which reflects the fitting degree of the models to independent variables, the performance of the models is: XGBoost > SVM > BPNN. This indicates that the XGBoost model has the best fitting effect (R^2 = 0.767). To sum up, the prediction indexes above reflect that the XGBoost model has the best accuracy for ground settlement prediction caused by shield tunneling.

4.3. The interpretability of prediction model

In geotechnical engineering, it is important to obtain the complex nonlinear relationship between strata deformation and its influencing factors by constructing prediction models. Therefore, the interpretability of the settlement prediction model is particularly important, which is often overlooked in existing studies (Suwansawat and Einstein, 2006; Pourtaghi and Lotfollahi-Yaghin, 2012; Hasanipanah et al., 2016). Obviously, it is difficult for the ANN model and the SVM model, as black box models, to satisfy requirements. XGBoost is a model established with clearly tree model structures. One of the tree structures generated during model iteration is shown in Fig. 10. It can be seen that the model structure has six layers, including five layers of tree structure and one layer of leaf structure, and nodes of each layer are divided by the value of input features.

By extracting the model's iterative process, the role of each influencing factor in settlement prediction can be analyzed, and the important evaluation index for each influencing factor can be fetched by using the feature importance module in the software. The feature importance scores in the settlement prediction models established in this study are listed in Fig. 11.

According to the feature importance evaluation, the strata cohesion in the shield tunneling process is the index with the highest score. Then, in order of importance scores, they are face pressure, tunnel cover depth, tunneling speed, whether there is a risk source, cutter head torque, earth volume, and grouting pressure. From the model's feature evaluation, it can be seen that for the settlement prediction model, geotechnical property is the most important factor determining the surface settlement, while subjective factors (shield tunneling parameters) will have a certain influence on settlement. This is consistent with the general understanding of settlement analysis currently, thus further verifying the reliability of the model.

4.4. Summary of model effects

Based on the above analysis, we can obtain the following conclusions:

(1) In terms of both the prediction indexes and the overall distribution of prediction results, the XGBoost model shows the best prediction effect, the ANN model shows the worst prediction effect, and the SVM model is between the two.

(2) There is the unbalanced settlement distribution in the samples used for training, the measuring points with larger settlement values account for a smaller proportion of the samples. The ANN model and the SVM model have poor prediction effects on these data, while the

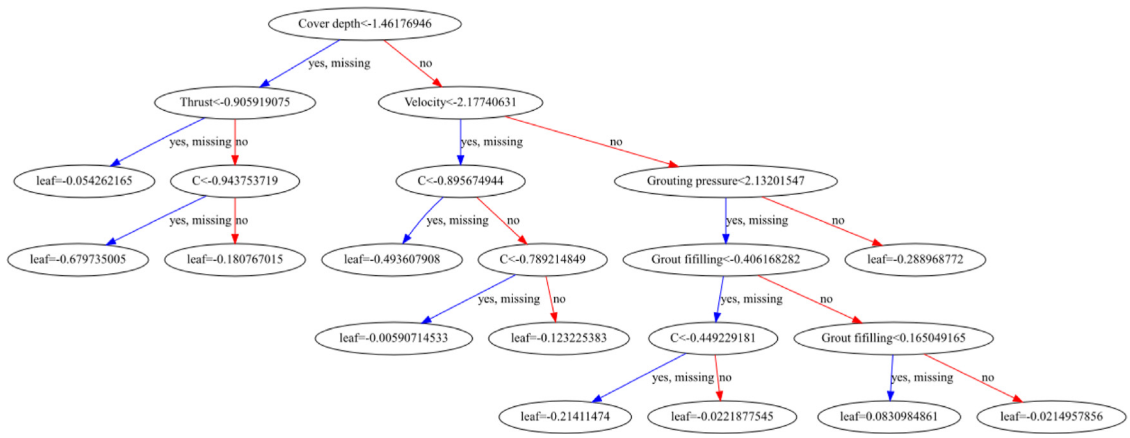


Fig. 10. The tree structure of XGBoost model.

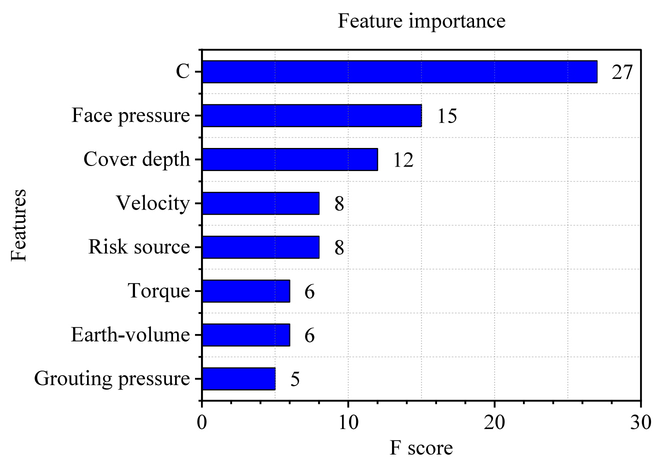


Fig. 11. Feature importance evaluation.

XGBoost model still presents a relatively reliable prediction effect on these measuring points. This is also the main reason why the XGBoost model is superior to the other two models in predicting evaluation indexes.

(3) The XGBoost model has obvious advantages in model interpretability. It can reflect the contribution degrees of different factors to settlement development by extracting feature importance scores of the model, so as to realize the close connection between the prediction model and practical engineering.

(4) The evaluation results on the feature importance of the XGBoost model show that the geotechnical properties are the most important factors determining the development of ground surface settlement, while shield tunneling parameters have a certain influence on settlement control.

4.5. Discussion and prospect of the XGBoost model

Although XGBoost model is better than SVM model and ANN model in settlement prediction effect through our research, it still has some limitations. Firstly, the model prediction effect depends on the data size. Fig. 12 shows the model prediction results under all data and the prediction results under 100 sets of data. It can be seen that when the sample size is insufficient, the prediction performance of the XGBoost model will decrease significantly. Through our tests, the model can only have stable performance when the number of samples is more than 200. Secondly, there are still some errors between the model prediction value and the monitoring value, which may be caused by insufficient model optimization or insufficient data features. These limitations deserve further study in the future.

5. Conclusion

In this paper, the XGBoost model and Bayesian optimization are introduced to predict the ground surface settlement caused by shield tunneling.

(1) By using XGBoost algorithm and Bayesian optimization, a prediction model for the settlement induced by the shield tunneling was established based on 533 monitoring data, which can achieve reliable prediction of ground surface settlement caused by shield tunneling.

(2) Compared with the prediction results of the ANN model and the SVM model, the XGBoost model has advantages in prediction accuracy, especially for the prediction of out-of-limit settlement points.

(3) Compared with the ANN model and the SVM model, the XGBoost model has better interpretability, and can evaluate the importance of features. According to the results feature importance evaluation, geotechnical properties are the determining factor for predicting the settlement.

CRediT authorship contribution statement

Jie Su: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition. **Yuzhe Wang:** Writing – original draft, Formal analysis, Supervision. **Xiaokai Niu:** Project administration, Resources. **Shan Sha:** Data curation. **Junyu Yu:** Validation.

Declaration of competing interest

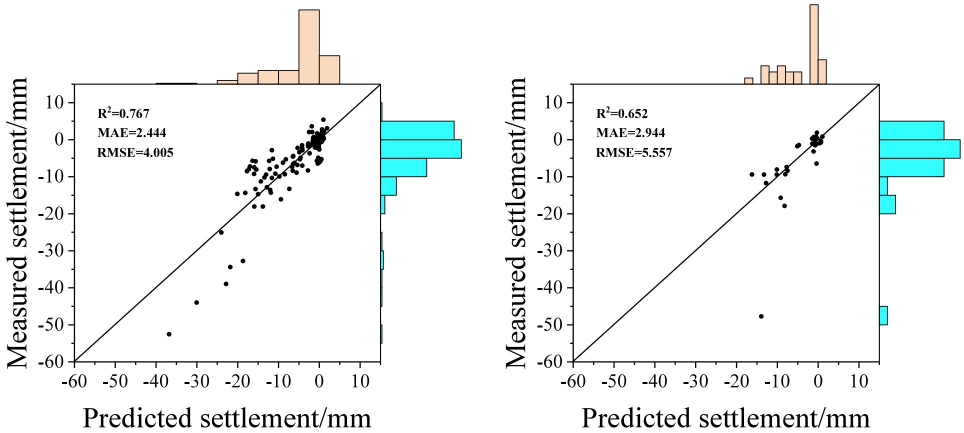
The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jie Su reports financial support was provided by National Natural Science Foundation of China.

Acknowledgments

The authors gratefully acknowledge the financial support by the National Natural Science Foundation of China under Grant 52178379, and the National Key R&D Program of China under Grant 2017YFC0805401.

Appendix

See Table A.1.



(a) Prediction results under 533 sets of data (b) Prediction results under 100 sets of data

Fig. 12. Prediction results of XGBoost model under different data amounts.

Table A.1
Part of the monitoring data.

	Face pressure	Thrust	Torque	Earth-volume	Velocity	Grouting pressure	Grout filling	c	φ	Risk source	Cover depth	Maximum settlement
	Mpa	MN	kN m	m^3	mm/min	Mpa	m^3	MPa	°		m	mm
1	0.756	8.40	1677	61	33	0.20	6.1	240	32	0	11.967	−1.7
2	1	15.60	2800	55	40	0.33	6.5	1000	45	0	19.439	−2.03
3	0.876	11.00	3000	62	45	0.20	5.7	150	30	1	13.888	−15.71
4	0.75	10.60	3385	55	50	0.27	7.0	150	20	0	17.693	−1.57
5	0.675	8.24	1425	56	38	0.20	6.5	10	8	0	10.620	−13.41
6	0.824	12.50	3100	63	42	0.21	6.2	200	30	0	13.791	−1.44
7	0.787	10.78	2636	60	33	0.22	6.7	1000	40	0	11.852	0.87
8	1.03	11.70	3607	56	43	0.23	5.8	500	38	0	14.780	−1.06
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
260	0.87	9.88	2266	56	51	0.27	6.2	150	20	0	13.394	−3.68
261	0.873	11.20	2600	63	41	0.22	5.1	300	30	1	13.871	−11.21
262	1.34	11.61	3440	55	58	0.19	5.8	12	25	0	13.853	−8.06
263	0.686	8.03	1215	56	41	0.21	6.6	10	8	0	10.718	−8.53
264	0.78	10.52	2324	55	41	0.26	6.2	1000	45	0	14.524	−0.24
265	0.807	8.21	2450	57	38	0.23	6.5	80	23	0	10.939	0.76
266	0.92	11.76	3540	55	40	0.24	6.0	500	38	0	14.525	−0.65
267	0.8	14.20	3027	58	32	0.31	6.7	1000	45	0	19.571	−0.14
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
529	0.92	11.33	3550	55	33	0.29	6.3	500	38	0	13.527	−2.1
530	0.97	9.65	2248	56	42	0.26	6.4	1000	45	0	15.507	0.73
531	0.92	8.38	1562	54	31	0.20	6.8	500	36	0	11.596	0.8
532	0.84	11.40	2950	56	40	0.19	6.5	60	36	0	11.212	−2.18
533	0.8	13.10	3020	55	35	0.30	5.5	12	25	0	11.622	−25.07

References

Bai, Y.X., Qi, T.Y., Li, T.D., et al., 2013. Prediction for surface collapse deformation of shield construction based on LSSVM. Chin. J. Rock Mech. Eng. 32 (S2), 3666–3674.

Bobet, A., 2001. Analytical solutions for shallow tunnels in saturated ground. J. Eng. Mech. 127 (12), 1258–1266.

Bouayad, D., Emeriault, F., 2017. Modeling the relationship between ground surface settlements induced by shield tunneling and the operational and geological parameters based on the hybrid PCA/ANFIS method. Tunn. Undergr. Space Technol. 68, 142–152.

Cao, L.Q., 2020. Mechanical Responses and Safety Control of Composite Muiti-Layered Ground and Built Environment Due to Shield Tunneling in Urban Areas (Ph.D. thesis). Beijing Jiaotong University.

Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system. In: The 22nd ACM SIGKDD International Conference. ACM.

Hasanipanah, M., Noorian-Bidgoli, M., Jahed Armaghani, D., et al., 2016. Feasibility of PSO-ANN model for predicting surface settlement caused by tunneling. Eng. Comput. 32 (4), 705–715.

Hasanpour, R., Rostami, J., Ünver, B., 2014. 3D finite difference model for simulation of double shield TBM tunneling in squeezing grounds. Tunn. Undergr. Space Technol. 40, 109–126.

Hou, J., Zhang, M.X., 2009. Prediction of surface settlements induced by shield tunneling: An ANFIS model. In: Proceeding of the 6th International Symposium on Geotechnical Aspects of Underground Construction in Soft Ground.

Lin, R.A., Sun, Y.F., Dai, Z.H., et al., 2018. Predicting for ground surface settlement induced by shield tunneling in upper-soft and lower-hard ground based on RS-SVR. China J. Highw. Transp. 31 (11), 130–137.

Loganathan, N., 1998. Analytical prediction for tunneling-induced ground movements in clays. J. Geotech. Geoenviron. Eng. 124 (9).

Mair, R.J., Taylor, R.N., Bracegirdle, A., 1993. Subsurface settlement profiles above tunnels in clays. Géotechnique 43 (2).

Migliazza, M., Chiorboli, M., Giani, G.P., 2009. Comparison of analytical method, 3D finite element model with experimental subsidence measurements resulting from the extension of the Milan underground. Comput. Geotech. 36 (1–2), 113–124.

Ocak, I., Seker, S.E., 2013. Calculation of surface settlements caused by EPBM tunneling using artificial neural network, SVM, and Gaussian processes. Environ. Earth Sci. 70 (3), 1263–1276.

O'Reilly, M.P., New, B.M., 1982. Settlement above tunnels in the United Kingdom-their magnitude and prediction. In: 3rd International Symposium, Brighton, London. pp. 173–181.

Peck, R.B., 1969. Deep excavations and tunneling in soft ground. In: Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering. Sociedad Maxicana de Mecanica de Suelos, Mexico City.

Pourtaghi, A., Lotfollahi-Yaghin, M.A., 2012. Wavenet ability assessment in comparison to ANN for predicting the maximum surface settlement caused by tunneling. Tunnell. Undergr. Space Technol. Inc. Trench. Technol. Res. 28, 257–271.

Qiao, J.L., Fan, Y.L., Liu, B., et al., 2012. Predicting the surface settlement by shield tunneling based on modified BP network. Chin. J. Undergr. Space Eng. 8 (02), 352–357.

Shahriari, B., Swersky, K., Wang, Z., et al., 2015. Taking the human out of the loop: A review of Bayesian optimization. Proc. IEEE 104 (1), 148–175.

Shivaei, S., Hataf, N., Pirastehfar, K., 2020. 3D numerical investigation of the coupled interaction behavior between mechanized twin tunnels and groundwater – A case study: Shiraz metro line 2. Tunn. Undergr. Space Technol. 103, 103458.

Sun, J., Yuan, J.R., 2001. Soil disturbance and ground movement under shield tunneling and its intelligent prediction by using ANN technology. Chin. J. Geotech. Eng. (03), 261–267.

Suwansawat, S., Einstein, H.H., 2006. Artificial neural networks for predicting the maximum surface settlement caused by EPB shield tunneling. Tunn. Undergr. Space Technol. 21 (2), 133–150.

Tian, M.J., Mou, J.H., Qiu, W.G., 2017. Research of the model comprehensive judgement for tunnel stability based on BP neural network. China Civ. Eng. J. 50 (S2), 260–266.

Wang, J.C., 2014. Mechanical Response and its Control of Overlying Existing Underground Structure with Shallow Tunnel Construction (Ph.D. thesis). Beijing Jiaotong University.

Zhang, L., Wu, X., Ji, W., et al., 2017. Intelligent approach to estimation of tunnel-induced ground settlement using wavelet packet and support vector machines. J. Comput. Civ. Eng. 31 (2), 4016053.

Zhou, C., 2011. Research on Time-Space Evolution Law and Early-Warning of the Ground Surface Settlement in Metro Shield Tunneling (Ph.D. thesis). Huazhong University of Science and Technology.