

## Research Paper

## Estimation of diaphragm wall deflections for deep braced excavation in anisotropic clays using ensemble learning

Runhong Zhang<sup>a</sup>, Chongzhi Wu<sup>a</sup>, Anthony T.C. Goh<sup>b</sup>, Thomas Böhlke<sup>c</sup>, Wengang Zhang<sup>a,d,\*</sup><sup>a</sup> School of Civil Engineering, Chongqing University, Chongqing, 400045, China<sup>b</sup> School of Civil and Environmental Engineering, Nanyang Technological University, 639798, Singapore<sup>c</sup> Institute of Engineering Mechanics, Karlsruhe Institute of Technology (KIT), Kaiserstraße 10, 76131, Karlsruhe, Germany<sup>d</sup> Key Laboratory of New Technology for Construction of Cities in Mountain Area, Chongqing University, Ministry of Education, Chongqing, 400045, China

## ARTICLE INFO

## Keywords:

Anisotropic clay

NGI-ADP

Wall deflection

Ensemble learning

eXtreme gradient boosting

Random forest regression

## ABSTRACT

This paper adopts the NGI-ADP soil model to carry out finite element analysis, based on which the effects of soft clay anisotropy on the diaphragm wall deflections in the braced excavation were evaluated. More than one thousand finite element cases were numerically analyzed, followed by extensive parametric studies. Surrogate models were developed via ensemble learning methods (ELMs), including the eXtreme Gradient Boosting (XGBoost), and Random Forest Regression (RFR) to predict the maximum lateral wall deformation ( $\delta_{hmax}$ ). Then the results of ELMs were compared with conventional soft computing methods such as Decision Tree Regression (DTR), Multilayer Perceptron Regression (MLPR), and Multivariate Adaptive Regression Splines (MARS). This study presents a cutting-edge application of ensemble learning in geotechnical engineering and a reasonable methodology that allows engineers to determine the wall deflection in a fast, alternative way.

## 1. Introduction

Nowadays, the design of deep excavations in soft soils remains a challenge for geotechnical engineers. Poor design could lead to the collapse of the excavation system. When fine-grained soils (i.e., clays, fine silts) are loaded quickly by external loading, the pore water in the soil does not dissipate immediately due to the low hydraulic conductivity. Consequently, pore water pressure will increase as a result of this “undrained” condition. Undrained loading is commonly regarded as the vital factor in the underground design for stability or bearing capacity. In these conditions, the shear strength of the soil was referred to as the “undrained shear strength ( $s_u$ )”. This  $s_u$  is not constant, and it depends on the direction of the loading (anisotropy), the rate and duration of the loading, and the sample disturbance. The anisotropy of soft clay proves to be important in excavation analyses, such as Hanson and Clough (1981), Hsieh et al. (2008) and Kong et al. (2012). Teng et al. (2014) studied the anisotropy of clays to obtain more accurate analytical results for geotechnical problems by conducting a series of  $K_0$ -consolidated undrained triaxial compression (CK<sub>0</sub>UC) tests on tube samples of natural Taipei silty clay with multidirectional bender elements. D’Ignazio et al. (2017) conducted a full-scale failure test and finite element analysis by

adopting the NGI-ADP Soft model (Grimstad et al., 2012) to simulate the strain-softening behavior of the clay. However, limited study has been systematically conducted to analyze the anisotropic behavior of soils in the braced excavation. In order to model anisotropy and strain-softening behavior of Singapore marine clay, this paper adopted the elastoplastic constitutive model NGI-ADP, which was developed by the Norwegian Geotechnical Institute (NGI) based on the ADP (Active-Direct shear-Passive) concept (Bjerrum, 1973) and implemented in PLAXIS finite element (FE), to carry out a series of FE analyses, taking soil anisotropic strength and strains as input parameters.

On the other hand, the use of supervised learning (SL) algorithms for the development of descriptive and predictive data-mining models has been widely accepted in geotechnical applications, which provided powerful new tools for practicing engineers. Goh et al. (1995) utilized the artificial neural network (ANN) for estimating lateral wall movements in braced excavations. Kung et al. (2007) estimated the deflection of diaphragm walls caused by excavation in clays using ANN. Chern et al. (2009) predicted lateral wall deflection in top-down excavation by using a Back-Propagation Neural Network (BPNN) model. Zhou et al. (2017) used the Random Forest (RF) to predict the ground settlements induced by the construction of a shield-driven tunnel. Zhang et al. (2017) adopted

\* Corresponding author. School of Civil Engineering, Chongqing University, Chongqing, 400045, China.

E-mail address: [zhangwg@cqu.edu.cn](mailto:zhangwg@cqu.edu.cn) (W. Zhang).

Peer-review under responsibility of China University of Geosciences (Beijing).

**Table 1**  
Parameters of the NGI-ADP model.

Type	Parameters	Description	Unit
Soil stiffness	$G_{ur}/s_u^A$	Ratio unloading/reloading shear modulus over (plane strain) active shear strength	[-]
	$\gamma_f^C$	Shear strain at failure in triaxial compression	[%]
	$\gamma_f^E$	Shear strain at failure in triaxial extension	[%]
	$\gamma_f^{DSS}$	Shear strain at failure in direct simple shear	[%]
Soil strength	$s_{u,ref}^A$	Reference (plane strain) active shear strength	[kN/m <sup>2</sup> /m]
	$s_u^{C,TX}/s_u^A$	Ratio triaxial compressive shear strength over (plane strain) active shear strength (default = 0.99)	[-]
	$y_{ref}$	Reference depth	[m]
	$s_{u,inc}$	Increase of shear strength with depth	[kN/m <sup>2</sup> /m]
	$s_u^P/s_u^A$	Ratio of (plane strain) passive shear strength over (plane strain) active shear strength	[-]
	$\tau_0/s_u^A$	Initial mobilization (default = 0.7)	[-]
	$s_u^{DSS}/s_u^A$	Ratio of direct simple shear strength over (plane strain) active shear strength	[-]
Others	$\nu_u$	Poisson's ratio	[-]

Multivariate adaptive regression splines (MARS) for inverse analysis of soil and wall properties in braced excavation. Goh et al. (2018) utilized the MARS model for the determination of EPB tunnel-related maximum surface settlement. Xie and Peng (2019) utilized Random Forest (RF) modeling and evaluated its predictive capability of for estimating the tunnel Excavation Damaged Zones (EDZs). Although SL techniques have been widely adopted in geotechnical engineering, they were rarely used for lateral wall deformation prediction in deep braced excavations with consideration of the anisotropic shear strength. Based on these considerations, the main objective of this study is to illustrate and compare the accuracy of different SL algorithms including XGBoost, DT, RFR, MLPR, and MARS for prediction of the maximum lateral deflection  $\delta_{hmax}$  in braced excavations.

This paper analyzed the lateral deformation of the diaphragm wall  $\delta_{hmax}$  for braced excavation in anisotropic clays with considering the crucial influential factors including the excavation width, wall stiffness, wall penetration, soil parameters of the ratio of (plane strain) passive shear strength over (plane strain) active shear strength ( $s_u^P/s_u^A$ ), ratio unloading/reloading shear modulus over (plane strain) active shear strength ( $G_{ur}/s_u^A$ ), reference (plane strain) active shear strength ( $s_{u,ref}^A$ ), and unit weight  $\gamma$ . Based on the 1778 FEA results, the different ensemble learning methods were adopted to predict the wall deformation by taking all the seven critical parameters into account. Performance comparisons of these methods were made and some conclusions about the ELMs use were arrived at.

**2. Finite element analyses**

**2.1. Soil constitutive model — NGI-ADP**

NGI-ADP model (Grimstad et al., 2012) is an anisotropic shear strength model for clay using non-linear stress path-dependent hardening relationship, defined from direct input of failure strains in the three directions of shearing represented by triaxial compression, simple direct shear, and triaxial extension, and the  $s_u$  profiles for active (A), simple direct shear (D) and passive (P) loading (stress paths) are given as input data. Table 1 shows the soil parameters of the NGI-ADP model. The most critical parameter is  $s_u^P/s_u^A$ . It defines the anisotropy degree of the clay, and it is ranging from 0 to 1. When it equals 1, it indicated that the clay is isotropic.

The NGI-ADP model is formulated for a general stress state, matching both undrained failure shear strengths and strains to the selected design profiles, referred to Andresen and Jostad (1999), Andresen (2002). In the NGI-ADP model, the Tresca approximation after Billington (1988) together with a modified von Mises plastic potential function (Von Mises, 1913) is used to circumvent the possible corner problems. The yield and plastic potential function is independent of the mean stress. Hence zero plastic volume strain develops.

The yield criterion for the NGI-ADP model in plane strain can be expressed by:

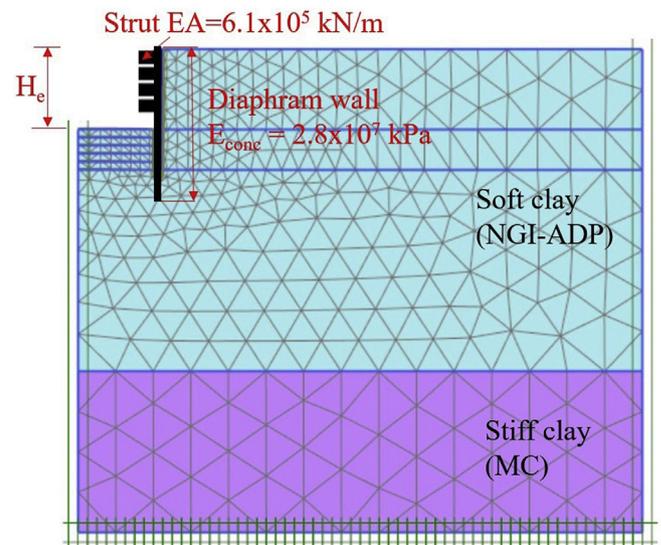


Fig. 1. Cross-sectional profile and typical mesh.

$$f = \sqrt{\left(\frac{\sigma_{yy} - \sigma_{xx}}{2} - (1-k)\tau_0 - k\frac{s_u^A - s_u^P}{2}\right)^2 + \left(\tau_{xy} \frac{s_u^A + s_u^P}{2s_u^{DSS}}\right)^2} - k\frac{s_u^A + s_u^P}{2} = 0 \tag{1}$$

where  $k = 2\frac{\sqrt{\gamma^P/\gamma_f^P}}{1+\gamma^P/\gamma_f^P}$  when  $\gamma^P < \gamma_f^P$ ; else  $k = 1$ ;  $\gamma^P, \gamma_f^P$  are the plastic shear strain and the failure plastic shear strain; and.  $\tau_0 = 0.5(\sigma'_{v0} - \sigma'_{h0}) = 0.5\sigma'_{v0}(1 - K_0)$ .

The NGI-ADP model uses elliptical interpolation between failure strain in passive stress state, direct simple shear, and active stress state. In the implementation of the NGI-ADP model, the yield surface is ensured to remain convex by restricting the input. More details are referred to Brinkgreve et al. (2017).

**2.2. Finite element modeling**

The braced excavation system is analyzed by the finite element method using Plaxis2D (Brinkgreve et al., 2017). The numerical model comprises of the diaphragm wall, four levels of struts, one soft-clay layer, and one stiff-clay layer. The soil bodies are simulated via 15-noded triangular elements. The structural elements of the diaphragm wall are assumed to be linearly elastic and modeled by 5-noded beam elements. The 3-noded bar elements represented the four levels of struts. The nodes on the left and right sides of the mesh are constrained from horizontal displacement, while the nodes at the bottom are constrained from moving both horizontally and vertically. In order to minimize the effects of the boundary restraints effect, the right vertical boundary is extended

**Table 2**  
NGI-ADP soil model properties of soft clay (undrained C).

Parameters	Value	Unit
$s_u^{C,TX}/s_u^A$	0.99	(–)
$\gamma_f^C$	0.75	(%)
$\gamma_f^E$	3.5	(%)
$\gamma_f^{DSS}$	1.735	(%)
$R_{inter}$	1.0	(–)
$s_u^{DSS}/s_u^A$	$(1+s_u^P/s_u^A)/2$	(–)
$\tau_0/s_u^A$	0.7	(–)
$\nu_u$	0.495	(–)
$y_{ref}$	60	(m)
$s_{u,inc}$	0	(kPa/m)

**Table 3**  
Stiff clay properties (undrained C).

Parameters and units	Value
Soil unit weight $\gamma$ (kN/m <sup>3</sup> )	16
Soil undrained shear strength $c_u$ (kPa)	100
Soil modulus ratio $E/c_u$	300
$\phi_u$ (°)	0
$K_o = 1 - \sin\phi$	1.0
Poisson's ratio $\nu_u$	0.495
Interface $R_{inter}$	1.0

**Table 4**  
Parameters and the ranges considered for soft clay.

Parameter	Ranges	Unit
Soil unit weight $\gamma$	15, 16, 18	(kN/m <sup>3</sup> )
$s_u^P/s_u^A$	0.4, 0.5, 0.6, 0.8, 1.0	(–)
$s_{u,ref}^A$	40, 50, 60	(kPa)
$G_{ur}/s_u^A$	300, 600, 900	(–)
Wall width $b$	0.6, 0.8, 1.2	(m)
Excavation width $B$	20, 30	(m)
Wall penetration $D$	3, 5, 10	(m)

**Table 5**  
Construction sequences.

Phase	Construction activities
1	Install wall (ground surface at $y = 60$ m)
2	Excavate to $y = 58$ m; reset displacement to zero
3	Install strut 1 at $y = 59$ m
4	Excavate to $y = 56$ m
5	Install strut 2 at $y = 57$ m
6	Excavate to $y = 54$ m
7	Install strut 3 at $y = 55$ m
8	Excavate to $y = 52$ m
9	Install strut 4 at $y = 53$ m
10	Excavate to $y = 50$ m
11	Set calculation type as safety

far enough from the excavation area.

A parametric study was carried out by adopting the NGI-ADP model for the soft clay with emphasizing on the lateral wall deformation. Fig. 1 shows the schematical cross-section of the excavation system with the final excavation depth  $H_e$ . It is a simplified typical soil stratum in many coastal areas such as Singapore, which comprise of a thick normally consolidated soft clay deposit layer overlying a stiff-clay layer. The Mohr-Coulomb (Undrained C) constitutive model is used for the stiff clay. The struts are used at the depths of 1 m, 3 m, 5 m, and 7 m below the original ground surface, and the horizontal strut spacing ( $L_{spacing}$ ) is 4 m. The strut stiffness per meter EA is assumed to be constant at  $6.1 \times 10^5$  kN/m, and the elastic modulus of the diaphragm wall is  $E_{conc} = 2.8 \times 10^7$  kPa.

The properties of soft and stiff clay are listed in Tables 2 and 3, respectively.

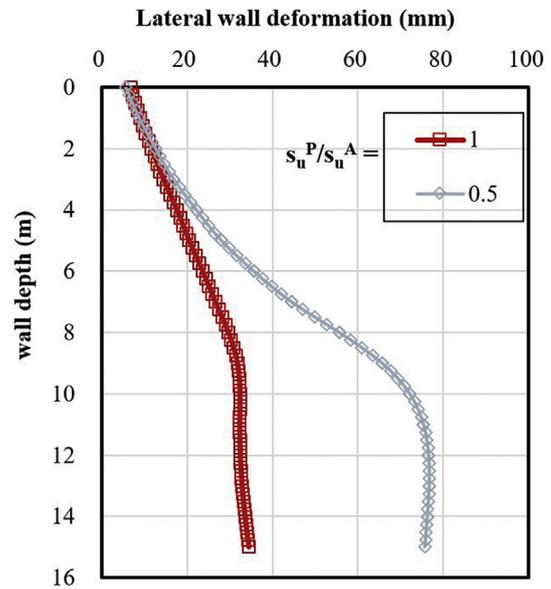


Fig. 2. Comparison of  $\delta_{hmax}$  for isotropic and anisotropic clays.

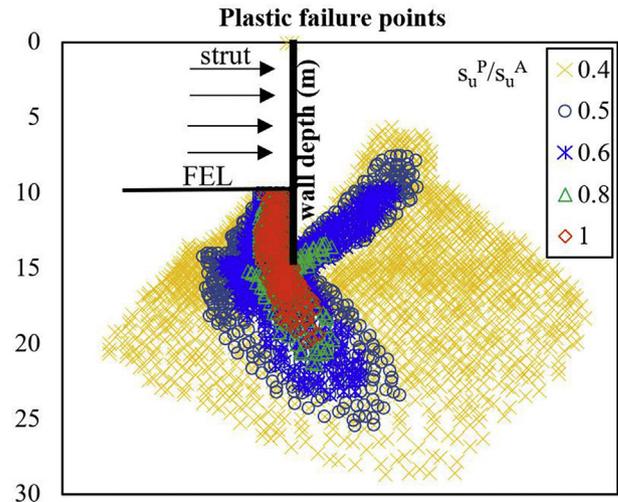


Fig. 3. Plastic failure points for  $s_u^P/s_u^A = \{0.4, 0.5, 0.6, 0.8, 1.0\}$ .

The parameters considered and ranges of properties for soft clay are shown in Table 4. The wall width is  $b$ , the excavation width is  $B$ , the penetration depth of the wall into the stiff layer is  $D$ . A total of 1778 hypothetical cases were analyzed. The details for the construction sequence are listed in Table 5.

### 2.3. Results and analyses

Fig. 2 shows the comparison of  $\delta_{hmax}$  for the cases with  $s_u^P/s_u^A$  equals 0.5 and 1. The case with  $s_u^P/s_u^A = 1$  implies that the soft clay is isotropic while the case with  $s_u^P/s_u^A = 0.5$  means that the soft clay is anisotropic. It is shown that the influence of  $s_u^P/s_u^A$  on the  $\delta_{hmax}$  is significant. The  $\delta_{hmax}$  of the case with  $s_u^P/s_u^A = 1$  is less than half of  $\delta_{hmax}$  for the case with  $s_u^P/s_u^A = 0.5$ .

Fig. 3 shows the plastic failure points for  $s_u^P/s_u^A = \{0.4, 0.5, 0.6, 0.8, 1.0\}$ . It is shown that as the  $s_u^P/s_u^A$  decreases, not only the plastic failure point number increase but also the plastic failure area shape and coverage changed. Moreover, there is an evident sliding surface that develops behind the wall only for the cases with  $s_u^P/s_u^A$  smaller than 1.

As shown in Fig. 4, the excavation width, wall penetration, and wall

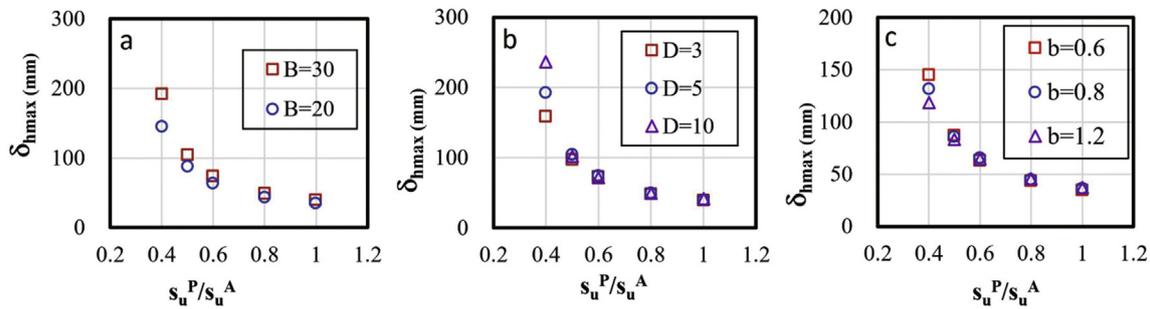


Fig. 4. Influence of (a)  $B$ , (b)  $D$ , and (c)  $b$  on  $\delta_{hmax}$  for  $s_{u,ref}^A = \{0.4, 0.5, 0.6, 0.8, 1.0\}$ .

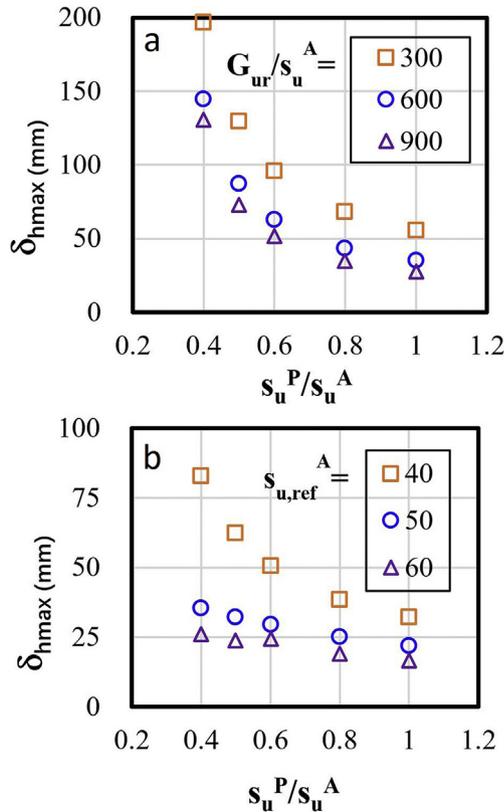


Fig. 5. Influence of (a)  $G_{ur}/s_u^A$  and (b)  $s_{u,ref}^A$  on  $\delta_{hmax}$  for  $s_u^P/s_u^A = \{0.4, 0.5, 0.6, 0.8, 1.0\}$ .

width have little influence on the wall deformation, except for the case with the  $s_u^P/s_u^A = 0.4$ . Since the wall movement is rigid for the “stiff” walls, the soil of the passive side has not failed for the higher  $s_u^P/s_u^A$ . Thus, the influence on the wall deformation is less significant. It requires further study in the future involving flexible walls, for which it is expected that the influence of  $B$ ,  $D$ , and  $b$  will be more significant.

However, compared to the three parameters shown in Fig. 5, the soil strength parameters such as  $s_u^P/s_u^A$ ,  $G_{ur}/s_u^A$ , and  $s_{u,ref}^A$  have more significant influences on the  $\delta_{hmax}$  than  $G_{ur}/s_u^A$ . Fig. 5a shows that the influence of  $G_{ur}/s_u^A$  on the  $\delta_{hmax}$  is less significant especially when  $s_u^P/s_u^A$  and  $G_{ur}/s_u^A$  are greater. Fig. 5b indicates that with the  $s_u^P/s_u^A$  increases, the  $\delta_{hmax}$  decreases more when the  $s_{u,ref}^A$  is 40, compared with the case of  $s_{u,ref}^A = 60$ , indicating that  $s_u^P/s_u^A$  has more considerable influence on the wall deformation when  $s_{u,ref}^A$  is smaller. It also can be concluded from Fig. 5 that the higher the  $s_u^P/s_u^A$ , the marginal the influence of  $s_{u,ref}^A$  on the wall deformation.

### 3. Estimation models of $\delta_{hmax}$

This section presents the development of the surrogate models, as well as the interpretations of the built models. Some new algorithms are in detail while the conventional methods are in brief.

#### 3.1. XGBoost

XGBoost is an algorithm proposed by Chen and Guestrin (2016), and it improves the accuracy of a model based on a decision tree model through the idea of integration. The procedures for XGBoost mainly are: firstly, a Classification and Regression Tree, (i.e., CART, Breiman et al., 1984) is trained with a part of the randomly taken training data. Second, the samples which incorrectly predicted are fed back to the training model. Then, the weights of these incorrect samples are increased, and a tree based on the old tree is trained. The loop is repeated many times till the pre-defined tree number is reached. The final result is the weighted average of all of the trees developed. More detailed explanations of the XGBoost algorithm are referred to Chen and Guestrin (2016), Zhang et al. (2020a). The boosting process is shown in Fig. 6.

Since CART is adopted in XGBoost, the complexity of the built model is directly related to the depth of the tree, the number of leaf nodes, as well as the output value of the leaf node (i.e., the leaf node weight in XGBoost). Assume the number of trees is  $k$ , the number of samples is  $i$ , the predicted score on the regression tree for the feature  $x_i$  is  $f_k$ . The output of the predictive model  $\hat{y}_i$  is expressed as follows:

$$\hat{y}_i = \sum_{k=1}^k f_k(x_i) \quad (2)$$

For machine learning, objective functions are the most basic expressions. The objective function generally consists of a loss function and a regularization term:

$$Obj(\Theta) = L(\Theta) + \Omega(\Theta) \quad (3)$$

in which  $L(\Theta)$  is the loss function to describe the degree to which the model fits the target data;  $\Omega(\Theta)$  is the regularization term describing the model complexity. Assume  $L(\Theta) = \sum_{i=1}^n l(y_i, \hat{y}_i)$  and  $\Omega(\Theta) = \sum_{k=1}^K \Omega(f_k)$ , Eq. (3) can be transformed to:

$$Obj(\Theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (4)$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (5)$$

in which  $\gamma$  is the complexity cost by introducing the additional leaf,  $T$  is the number of leaves,  $\lambda$  is the hyperparameter, term  $\frac{1}{2} \lambda \sum_{j=1}^T w_j^2$  is the

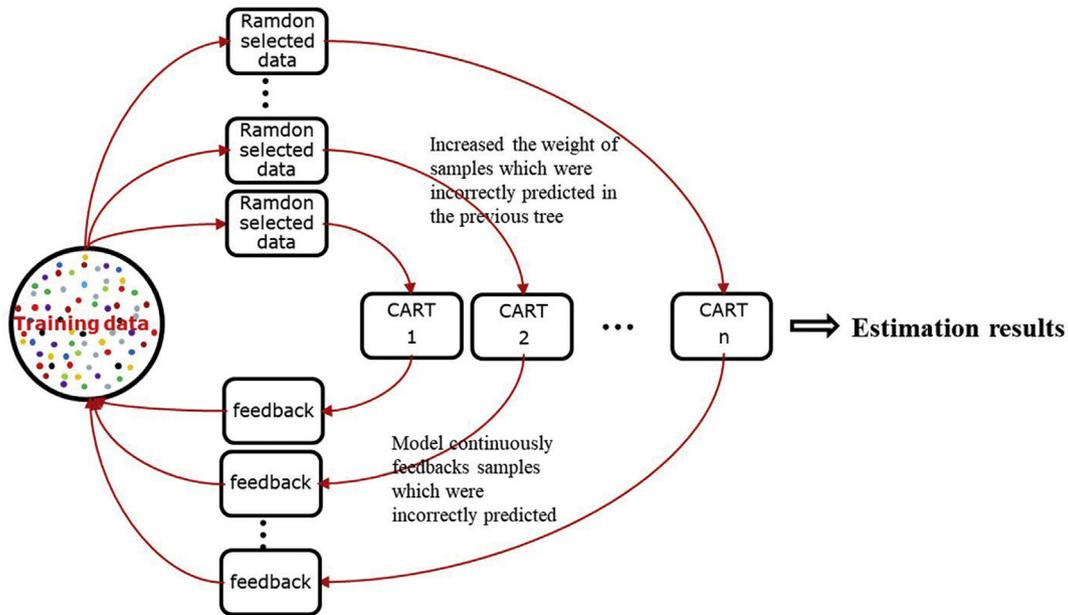


Fig. 6. Flow chart of boosting.

regularization of the weight of leaf nodes.

Let  $\hat{y}_i^{(k)}$  be the prediction of the  $i$ th instance at the  $k$ -th iteration,

$$\hat{y}_i^{(k)} = \hat{y}_i^{(k-1)} + f_k(x_i) \tag{6}$$

in which  $f_k(x_i)$  is determined via finding a tree to that minimizes the following approximated objective function using the second-order Taylor expansion:

$$Obj^{(k)} \simeq \sum_{i=1}^n \left[ g f_k(x_i) + \frac{1}{2} h_i f_k^2(x_i) \right] + \Omega(f_k) \tag{7}$$

The weight (output value) function of leaf nodes follows:

$$f_k(x_i) = w_j \tag{8}$$

in which vector  $w$  represents the score of each region; let  $I_j$  represents the sample set of leaf  $j$ . Eq. (7) can be transformed to:

$$\begin{aligned} Obj^{(k)} &= \sum_{i=1}^n \left[ g f_k(x_i) + \frac{1}{2} h_i f_k^2(x_i) \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \\ &= \sum_{j=1}^k \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left( \sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \end{aligned} \tag{9}$$

Finally, the last step is to find the  $w_j$  to minimize the  $Obj^{(k)}$  in Eq. (7). This problem involves simple quadratic programming.

The Gain in the loss reduction after the split can be described as follows:

$$Gain = \frac{1}{2} \left[ \frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \tag{10}$$

$\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda}$  is the score of the left leaf node after division;

in which  $\frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda}$  is the score of the right leaf node after division;

$\frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda}$  is the score of the present node before division;

$\gamma$  is the complexity cost of dividing the node into leaf nodes;

Eq. (10) compares the scores before and after division,  $\lambda$  is the threshold. If the result difference after division is greater than  $\lambda$ , then division; vice versa.

Generally, the input sample matrix  $X$  is sparse due to the missing data as well as the rounding errors (zero or very small values). When dealing with the sparse matrix, XGBoost suggests two default directions for each sample at each tree node. Then the node learns from the data to decide a better direction for samples, and subsequently this node is regarded as the next most suitable node.

The scalability of XGBoost is attributed to several important system and algorithm optimizations, including:

- (i) A novel tree learning algorithm for processing sparse data;
- (ii) the theoretically reasonable weighted quantile sketch process enables instance weights to process in approximate tree learning;
- (iii) introduced a novel sparsity-aware algorithm for parallel tree learning, parallel and distributed computing enable a faster learning and model exploration;
- (iv) proposed an efficient cache perceptual block structure for extra-nuclear tree learning.

### 3.2. RFR

Random forest refers to a classifier that uses multiple trees to train and predict samples. The classifier was first proposed by Breiman (2001), inspired by earlier work by Amit and Geman (1997) and then further developed by Cutler et al. (2011) and registered as a trademark. Random forest is composed up of multiple CARTs. For each CART, a smaller training set is back-sampled from the training set. In other words, some samples in the training set may be used multiple times in the training process (Zhang et al., 2020b). During the node training process of each tree, the features are also randomly extracted from all of the features according to a proportion of  $\sqrt{M}$ ,  $\frac{1}{2} \sqrt{M}$  or  $2\sqrt{M}$ , in which  $M$  is the total number of the features (Breiman, 2001).

The procedures for Random Forest are: firstly, the number of CART is determined as  $k$ , and the  $k$  is usually more than one thousand; then a part of training data is randomly back-sampled from the training set to train  $k$  trees simultaneously. All of these trees form a forest. The final prediction result employs the average values of all the independent trees, which

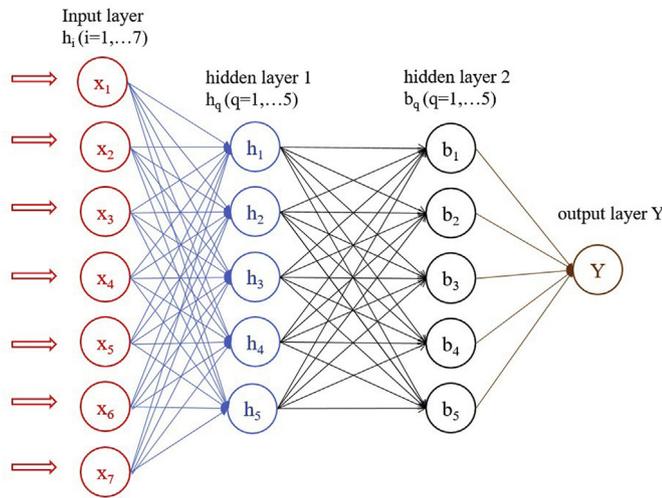


Fig. 7. MLPR architecture used in this study.

reduces the risk of overfitting significantly.

In summary, although the basic models for both XGBoost and random forest are the same decision trees, they are different in the training process. XGBoost relies on a serial combination decision tree, i.e., every single tree is trained based on the previous tree, except for the first tree. Random Forest is a parallel combination that all the trees can be calculated in a parallel procedure simultaneously.

### 3.3. DTR

Decision Tree (DT) is an SL method that can summarise decision rules from data sets with features and labels to solve classification and regression problems. Similar to k-nearest neighbors, the decision tree algorithm and its variants are algorithms that divide the input space into regions, and each region has separate parameters (Goodfellow et al., 2016).

The structure of DT is a flowchart-like tree that acts as a decision support system. A tree structure consists of a root node, internal nodes, and leaf nodes. Each path from the root to the leaf indicates a decision rule, which can be presented briefly as if-then rule denotes the relationship between input and output variables. When the dependent variable of the data set is a continuous value, the structure can be used as a regression tree, and a predicted value can be obtained by decision tree regression (DTR), which equals the mean value of the leaf nodes.

### 3.4. MLPR

A four-layer neural network is adopted in this study composed of one input layer, two hidden layers and one output layer. As shown in Fig. 7, the Multilayer Perceptron (MLP) algorithm is a forward-structured artificial neural network that maps a set of input vectors to a set of output vectors. An MLP can be regarded as a directed graph consisting of multiple node layers, and each node is connected to the next layer. Each node is a neuron with a nonlinear activation function except for the input nodes. A SL method back-propagation algorithm is used to generalize the perceptron, which overcomes the deficiency of perceptron MLPR cannot identify linearly indivisible data. MLP can approximate the nonlinear input function in regression by implementing nonlinear discriminants. There have been extensive researches focusing on underground excavation analyses using MLPR. Kung et al. (2007) and Chern et al. (2009) estimated the deflection of diaphragm wall caused by excavation in clays using MLPR. Huang and Wang (2007) studied the Neural network forecast model in deep excavation. Jan et al. (2002) adopted MLPR-based method to analyze the reliability of deep excavation. Sou-Sen and Chuang (2004), Yu et al. (2009) and Tsekouras et al. (2010) investigated

Table 6

Expression of the MARS model for lateral wall deflection.

Basis Function	Coefficient $\beta_n$	Basis Function	Coefficient $\beta_n$
Intercept $\beta_0$	19385.9	BF18 = $X_1 \times X_5 \times X_7 \times X_4$	-0.0261989
BF1 = $X_4$	-375.784	BF19 = $X_6 \times X_6 \times X_5$	4.81E-06
BF2 = $X_7$	14278.9	BF20 = $X_2 \times X_4$	-0.0944681
BF3 = $X_5$	-4119.83	BF21 = $X_7 \times X_4 \times X_7 \times X_4$	3.86752
BF4 = $X_7 \times X_4$	-500.085	BF22 = $X_7 \times X_5 \times X_7 \times X_4$	-27.1847
BF5 = $X_5 \times X_4$	97.5886	BF23 = $X_7 \times X_7$	-13427
BF6 = $X_4 \times X_5 \times X_4$	-0.386583	BF24 = $X_7 \times X_7 \times X_5 \times X_7$	-276.179
BF7 = $X_6 \times X_5$	-0.0112166	BF25 = $X_7 \times X_7 \times X_7 \times X_4$	78.9361
BF8 = $X_5 \times X_7 \times X_4$	37.971	BF26 = $X_5 \times X_5 \times X_7$	-152.455
BF9 = $X_4 \times X_7 \times X_4$	6.52654	BF27 = $X_5 \times X_5 \times X_7 \times X_4$	2.72672
BF10 = $X_7 \times X_5 \times X_7$	2132.57	BF28 = $X_1 \times X_7 \times X_7 \times X_4$	0.691492
BF11 = $X_7 \times X_7 \times X_4$	-154.474	BF29 = $X_1 \times X_7 \times X_7$	-22.146
BF12 = $X_7 \times X_4 \times X_5 \times X_4$	-0.808477	BF30 = $X_1 \times X_5 \times X_4$	-0.0895793
BF13 = $X_6 \times X_4 \times X_7 \times X_4$	2.03E-05	BF31 = $X_1 \times X_4 \times X_4$	0.0128747
BF14 = $X_2 \times X_5$	0.41908	BF32 = $X_5 \times X_4 \times X_5 \times X_4$	0.0339091
BF15 = $X_5 \times X_5$	214.956	BF33 = $X_2 \times X_5 \times X_7$	-0.108153
BF16 = $X_5 \times X_5 \times X_4$	-5.77068	BF34 = $X_2 \times X_5 \times X_7$	-0.108153
BF17 = $X_1 \times X_5 \times X_5$	0.183979		

MARS Expression:  $y = \beta_0 + \sum BF(X)\beta_n$ .

Note:  $X_1 = D$ ;  $X_2 = B$ ;  $X_3 = b$ ;  $X_4 = s_{u,ref}^A$ ;  $X_5 = \gamma$ ;  $X_6 = G_{ur}/s_u^A$ ;  $X_7 = s_u^P/s_u^A$

the ground surface settlement induced by excavations by employing MLPR algorithms.

In this study, there are seven nodes of the input layer represent  $\gamma$ ,  $s_u^P/s_u^A$ ,  $s_{u,ref}^A$ ,  $G_{ur}/s_u^A$ ,  $b$ ,  $B$ ,  $D$ , respectively. Five nodes of each two hidden layers represent five perceptrons, and one node of the output layer represents the wall deflection. In order to solve the optimization problem, the gradient descent method (SGD) is utilized: firstly, randomly initialize all parameters, then iteratively train the data and continuously calculate the gradient and update the parameters until the error is small enough and the number of iterations is sufficient.

### 3.5. MARS

MARS is an implementation of techniques popularized by Friedman (1991) for solving regression-type problems. It is a nonlinear and nonparametric regression method based on a divide and conquers strategy in which the training data sets are partitioned into separate piecewise linear segments (splines) of differing gradients (slope). No specific assumption about the underlying functional relationship between the input variables and the output is required. The endpoints of the segments are called knots. A knot marks the end of one region of data and the beginning of another. The resulting piecewise curves, which known as the basis functions (BF), give greater flexibility to the model, allowing for bends, thresholds, and other departures from linear functions. MARS has been widely applied in geotechnical engineering, previous applications of MARS algorithm in civil engineering can be found in Samui and Karup (2011), Lashkari (2012), Zhang and Goh (2013, 2016), Goh and Zhang (2014), Zhang et al. (2015, 2016, 2017, 2019, 2020a), Goh et al. (2017, 2018).

MARS modeling is a data-driven process constructed in a two-phase procedure; the forward phase generates BFs by searching in a stepwise manner and finds potential knots locations using the adaptive regression algorithm to improve the performance, preliminary resulting in an overfitting model; the backward is to remove extraneous variables and then find a close to the optimal model. Open source code on MARS from

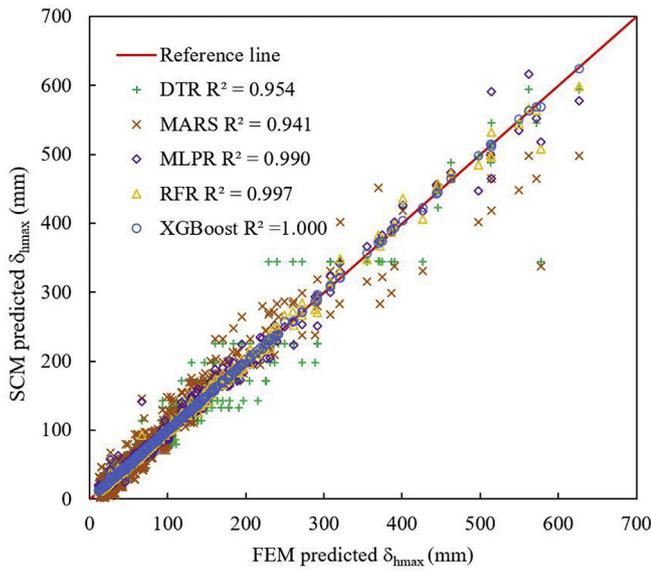


Fig. 8. Comparison between the training results of SCM and FEM.

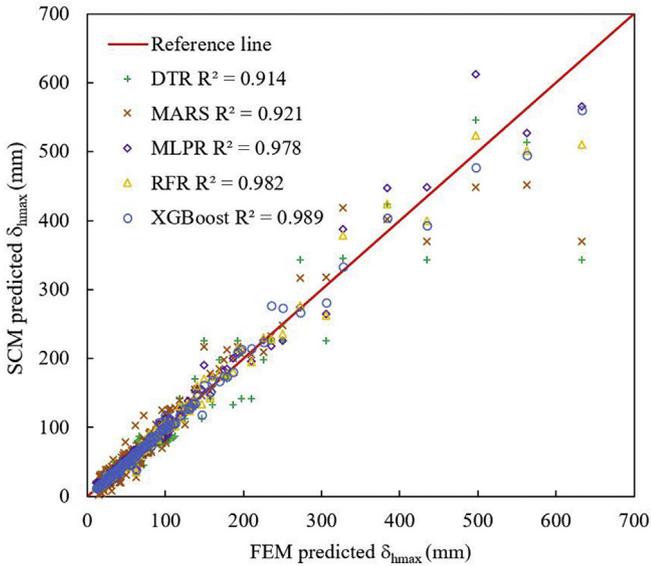


Fig. 9. Comparison between the testing results of SCM and FEM.

Jekabsons (2010) was used to carry out the analyses presented in this paper.

The optimal MARS model is determined by grouping together all the BFs that involve one variable and another cluster of BFs that involve pairwise interactions (and even higher-level interactions when applicable). The results of the BFs equation of this study are listed in Table 6.

### 3.6. Comparison of the four methods

Figs. 8 and 9 show the comparison of the results of training and testing results of the prediction of the  $\delta_{hmax}$  by XGBoost, MARS, MLPR, and RFR, respectively. The  $R^2$  (i.e., coefficient of determination) of the four methods are also shown in Figs. 8 and 9. The order of accuracy of the five methods can be: XGBoost > RFR > MLPR > MARS > DTR. Especially for the wall deflection is smaller than 200 mm, the points in Figs. 8 and 9 for XGBoost is closer to the reference line, indicating that the XGBoost outperforms other methods. As shown in Fig. 8, the data is non-uniform distributed, the data density for  $\delta_{hmax}$  smaller than 200 mm is much

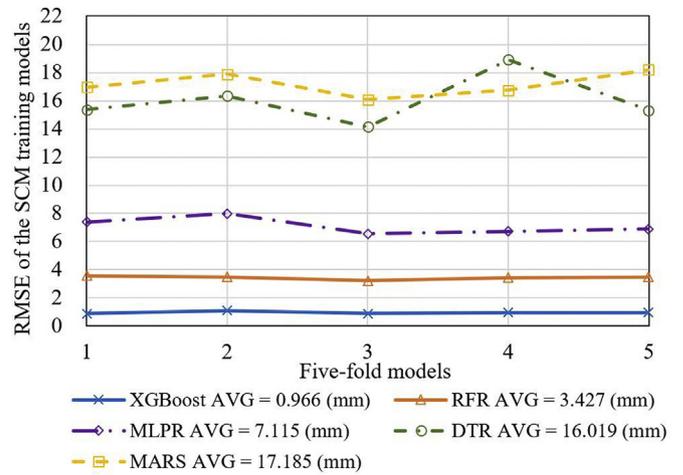


Fig. 10. RMSE change curve of training models under 5-fold cross-validations.

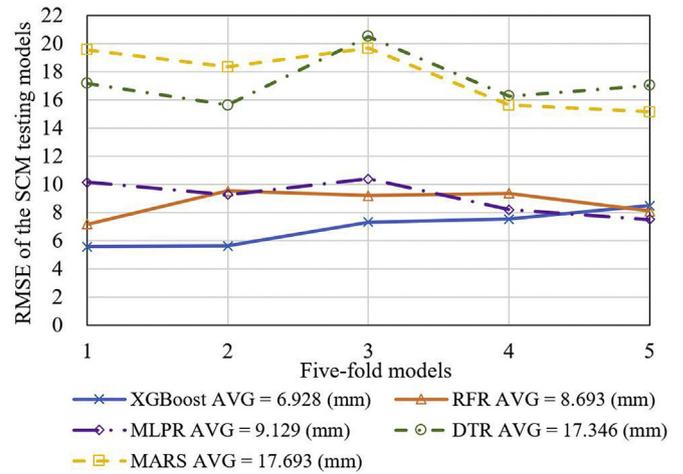


Fig. 11. RMSE change curve of testing models under 5-fold cross-validations.

intensive than the  $\delta_{hmax}$  larger than 200 mm. The MARS model better fits the lower values of wall deformation but shows a higher scatter for higher values. Not so significantly, the RFR, MLPR, and DTR models present the same trend. However, the XGBoost model still perfectly fits the higher values of wall deformation. It reveals the outstanding ability of XGBoost in processing sparse data.

Figs. 10 and 11 show the RMSE for training and testing models under 5-fold cross-validations, respectively. As a smaller RMSE value indicate high confidence in the model predicted values, the XGBoost performs the best among the five methods for both training model and testing model. Moreover, as can be seen in Figs. 10 and 11, the fluctuation for XGBoost and RFR results are evidently lower than DTR, MLPR, and MARS, implying that XGBoost and RFR are more stable and robust.

It can be concluded that the overall performance has been slightly improved by ensemble learning XGBoost and RFR method in this study, compared with the more conventional MARS and MLPR. The slight improvement is insignificant for this problem in the manuscript since the data patterns are from numerical synthetic cases, i.e., with less noise. As shown in Figs. 10 and 11, the fluctuation of RMSE for XGBoost and RFR is evidently lower than DTR, MLPR, and MARS, which indicates that XGBoost and RFR are more stable and robust than the rest. For other applications with instrumented results, the use of ensemble learning will outperform ordinary learning. As a robust tree-based tool, XGBoost and RFR method can balance the relationship between the predictive accuracy and requirements of intelligibility.

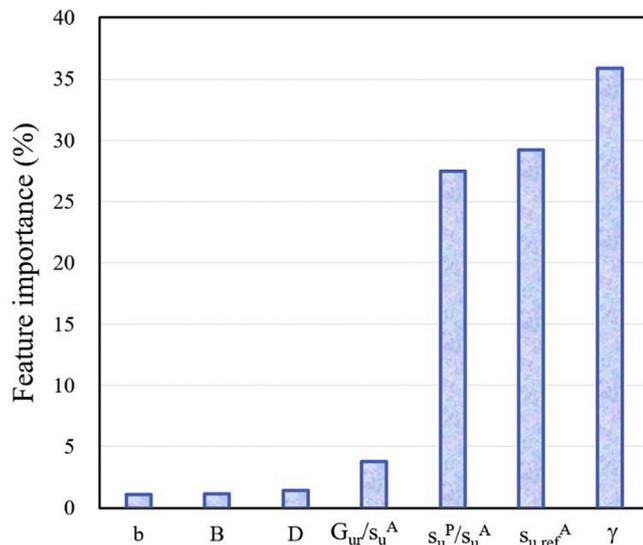


Fig. 12. Feature importance of the XGBoost.

### 3.7. Sensitivity analysis

Fig. 12 shows the sensitivity analysis results of the XGBoost. The result is consistent with the result of the FEM method. The  $\gamma$  and the anisotropic parameter of  $s_u^P/s_u^A$  and  $s_{u,ref}^A$  show significant influence in determining the  $\delta_{hmax}$ . However, the  $G_{ur}/s_u^A$  shows lower importance than  $s_u^P/s_u^A$  and  $s_{u,ref}^A$ , and the  $b, B, D$  have marginal influence.

## 4. Summary and conclusions

This paper firstly carried out finite element simulations with adopting the NGI-ADP soil model to assess the lateral wall deflection for excavations in soft clays. The effects of seven key parameters on diaphragm wall deflections are evaluated, including excavation width, wall stiffness, wall penetration, soil parameters of  $s_u^P/s_u^A$ ,  $G_{ur}/s_u^A$ ,  $s_{u,ref}^A$ , and  $\gamma$ . The results show that the  $b, B, D$  have a marginal influence on the  $\delta_{hmax}$  when the wall is stiff, except for the case with  $s_u^P/s_u^A = 0.4$ . The soil properties of the  $s_{u,ref}^A$ ,  $G_{ur}/s_u^A$ , and  $s_u^P/s_u^A$  have significant influences on the  $\delta_{hmax}$ . Moreover, the results are consistent with the sensitivity analysis results of the XGBoost. The anisotropic characteristic of soil parameters is essential in determining the lateral wall deflections induced by braced excavation supported by the diaphragm wall.

The results of a total of 1778 synthetic cases with different excavation geometries and soil properties were utilized by the ELMs, including XGBoost and RFR to predict the  $\delta_{hmax}$ . The ELMs results were then compared with the predictions obtained from conventional soft computing methods such as DTR, MLPR, and MARS. The result shows that the XGBoost and RFR outperform DTR, MLPR, and MARS in the predictions of  $\delta_{hmax}$ , especially for a higher  $\delta_{hmax}$  which in a sparse distribution out of the data set, and also shows a characteristic of stabilization.

It should be mentioned that it is the data and features that determine the upper limit of machine learning, while models or algorithms only approach this upper limit from different perspectives.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgment

This work was supported by the High-end Foreign Expert

Introduction program (No. G20190022002), Chongqing Construction Science and Technology Plan Project (2019-0045) and the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJZD-K201900102). The financial support is gratefully acknowledged.

## References

- Amit, Y., Geman, D., 1997. Shape quantization and recognition with randomized trees. *Neural Comput.* 9 (7), 1545–1588.
- Andresen, L., 2002. Capacity Analysis of Anisotropic and Strain-Softening Clay. Ph.D. thesis. University of Oslo, Norway.
- Andresen, L., Jostad, H.P., 1999. Application of an anisotropic hardening model for undrained response of saturated clay. *Proc. NUMOG VII* 581–585.
- Billington, E.W., 1988. Generalized isotropic yield criterion of incompressible materials. *Acta Mech.* 72, 1–20.
- Bjerrum, L., 1973. Problems of soil mechanics and construction on soft clays. State-of-the-art report. In: *Proceedings, 8th ICSMFE, Moscow*, vol. 3, pp. 111–159.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32. <https://doi.org/10.1023/a:1010933404324>.
- Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. *Classification and Regression Trees*. CRC Press, Boca Raton, FL, p. 368.
- Brinkgreve, L.B.J., Engin, E., Swolfs, W.M., 2017. *Plaxis Manual*. PLAXIS bv, Netherlands.
- Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.
- Chern, S., Tsai, J.H., Chien, L.K., Huang, C.Y., 2009. Predicting lateral wall deflection in top-down excavation by neural network. *Int. J. Offshore Polar Eng.* 19 (2), 151–157.
- Cutler, A., Cutler, D.R., Stevens, J.R., 2011. Random forests. *Mach. Learn.* 45 (1), 157–176.
- Friedman, J.H., 1991. Multivariate adaptive regression splines. *MLPR Stat* 19, 1–141.
- Goh, A.T.C., Zhang, W.G., 2014. An improvement to MLR model for predicting liquefaction-induced lateral spread using Multivariate Adaptive Regression Splines. *Eng. Geol.* 170, 1–10.
- Goh, A.T.C., Wong, K.S., Broms, B.B., 1995. Estimation of lateral wall movements in braced excavations using neural networks. *Can. Geotech. J.* 32 (6), 1059–1064.
- Goh, A.T.C., Zhang, Y., Zhang, R.H., Zhang, W.G., Xiao, Y., 2017. Evaluating stability of underground entry-type excavations using multivariate adaptive regression splines and logistic regression. *Tunn. Undergr. Space Technol.* 70, 148–154.
- Goh, A.T.C., Zhang, W.G., Zhang, Y.M., Xiao, Y., Xiang, Y.Z., 2018. Determination of earth pressure balance tunnel-related maximum surface settlement: a multivariate adaptive regression splines approach. *Bull. Eng. Geol. Environ.* 77, 489–500.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. The MIT Press.
- Grimstad, G., Andresen, L., Jostad, H.P., 2012. NGI-ADP: anisotropic shear strength model for clay. *Int. J. Numer. Anal. Methods GeoMech.* 36 (4), 483–497.
- Hanson, L.A., Clough, G.W., 1981. The significance of clay anisotropy in finite element analysis of supported excavations. In: *Proc. Symp. Implementation of Computer Procedure of Stress Strain Laws in Geotechnical Engineering, I-II*. Chicago Illinois.
- Hsieh, P.G., Ou, C.Y., Liu, H.T., 2008. Basal heave analysis of excavations with consideration of anisotropic undrained strength of clay. *Can. Geotech. J.* 45 (6), 788–799.
- Huang, F.K., Wang, G.S., 2007. *MLPR-based Reliability Analysis for Deep Excavation*. IEEE, New York.
- D'Ignazio, M., Lansivaara, T., Jostad, H.P., 2017. Failure in anisotropic sensitive clays: a finite element study of the Perni failure test. *Can. Geotech. J.* 54 (7), 1013–1033.
- Jan, J.C., Hung, S.L., Chi, S.Y., Chern, J.C., 2002. Neural network forecast model in deep excavation. *J. Comput. Civ. Eng.* 16 (1), 59–65.
- Jekabsons, G., 2010. *VariReg: A Software Tool for Regression Modeling Using Various Modeling Methods*. Riga Technical University. <http://www.cs.rtu.lv/jekabsons>.
- Kong, D.S., Men, Y.Q., Wang, L.H., Zhang, Q.H., 2012. Basal heave stability analysis of deep foundation pits in anisotropic soft clays. *J. Cent. S. Univ.* 43 (11), 4472–4476.
- Kung, G.T.C., Hsiao, E.C.L., Schuster, M., Juang, C.H., 2007. A neural network approach to estimating deflection of diaphragm walls caused by excavation in clays. *Comput. Geotech.* 34 (5), 385–396. <https://doi.org/10.1016/j.comgeo.2007.05.007>.
- Lashkari, A., 2012. Prediction of the shaft resistance of non-displacement piles in sand. *Int. J. Numer. Anal. Methods GeoMech.* 37, 904–931.
- Samui, P., Karup, P., 2011. Multivariate adaptive regression spline and least square support vector machine for prediction of undrained shear strength of clay. *Int. J. Appl. Metaheuristic Comput. (IJAMC)* 3 (2), 33–42.
- Sen-Sou, L., Chuang Hsien, L., 2004. Neural-network-based regression model of ground surface settlement induced by deep excavation. *Autom. Constr.* 13 (3), 279–289. [https://doi.org/10.1016/s0926-5805\(03\)00018-9](https://doi.org/10.1016/s0926-5805(03)00018-9).
- Teng, F.C., Ou, C.Y., Hsieh, P.G., 2014. Measurements and numerical simulations of inherent stiffness anisotropy in soft Taipei clay. *J. Geotech. Geoenviron. Eng.* 140 (1), 237–250. [https://doi.org/10.1061/\(asce\)gt.1943-5606.0001010](https://doi.org/10.1061/(asce)gt.1943-5606.0001010).
- Tsekouras, G.J., Koukoulis, J., Mastorakis, N.E., 2010. An optimized neural network for predicting settlements during tunneling excavation. *World. Sci. Eng. Acad. Soc.* 9, 1153–1167.
- Von Mises, R., 1913. *Mechanik der festen korper in plastisch deformablem zustand*. Gottinger Nachrichten Math. Phys. Klasse 1, 582–592 (in German).
- Xie, Q., Peng, K., 2019. Space-time distribution laws of tunnel excavation damaged Zones (EDZs) in deep mines and EDZ prediction modeling by random forest regression. *Adv. Civ. Eng.* 1–13. <https://doi.org/10.1155/2019/6505984>.

- Yu, J., Chen, H.M., Yu, J., Chen, H.M., 2009. Artificial neural network's application in intelligent prediction of surface settlement induced by foundation pit excavation. In: *Icicta: 2009 Second International Conference On Intelligent Computation Technology And Automation*, Vol 1, Proceedings. 2009, Los Alamitos. IEEE Computer Society, pp. 303–305.
- Zhang, W.G., Goh, A.T.C., 2013. Multivariate adaptive regression splines for analysis of geotechnical engineering systems. *Comput. Geotech.* 48, 82–95.
- Zhang, W.G., Goh, A.T.C., 2016. Multivariate adaptive regression splines and neural network models for prediction of pile drivability. *Geosci. Front* 7, 45–52.
- Zhang, W.G., Goh, A.T.C., Zhang, Y.M., Chen, Y.M., Xiao, Y., 2015. Assessment of soil liquefaction based on capacity energy concept and multivariate adaptive regression splines. *Eng. Geol.* 188, 29–37.
- Zhang, W.G., Goh, A.T.C., Zhang, Y.M., 2016. Multivariate adaptive regression splines application for multivariate geotechnical problems with big data. *Geotech. Geol. Eng.* 34 (1), 193–204.
- Zhang, W.G., Zhang, Y.M., Goh, A.T.C., 2017. Multivariate adaptive regression splines for inverse analysis of soil and wall properties in braced excavation. *Tunn. Undergr. Space Technol.* 64, 24–33.
- Zhang, W.G., Wu, C.Z., Li, Y.Q., Wang, L., Samui, P., 2019. Assessment of pile drivability using random forest regression and multivariate adaptive regression splines. *Georisk.* <https://doi.org/10.1080/17499518.2019.1674340>.
- Zhang, W.G., Li, H.R., Wu, C.Z., Li, Y.Q., Liu, Z.Q., Liu, H.L., 2020a. Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. *Undergr. Space* in press. <https://doi.org/10.1016/j.undsp.2019.12.003>.
- Zhang, W.G., Zhang, R.H., Wu, C.Z., Goh, A.T.C., Lacasse, S., Liu, Z.Q., Liu, H.L., 2020b. State-of-the-art review of soft computing applications in underground excavations. *Geosci. Front* 11 (4), 1095–1106. <https://doi.org/10.1016/j.gsf.2019.12.003>.
- Zhou, J., Shi, X.Z., Du, K., Qiu, X.Y., Li, X.B., Mitri, H.S., 2017. Feasibility of random-forest approach for prediction of ground settlements induced by the construction of a shield-driven tunnel. *Int. J. GeoMech.* 17 (6), 04016129. [https://doi.org/10.1061/\(asce\)gm.1943-5622.0000817](https://doi.org/10.1061/(asce)gm.1943-5622.0000817).