ORIGINAL ARTICLE

A novel whale optimization algorithm optimized XGBoost regression for estimating bearing capacity of concrete piles

Hieu Nguyen[1] · Minh-Tu Cao[2] · Xuan-Linh Tran[3,4] · Thu-Hien Tran[3,4] · Nhat-Duc Hoang[3,4]

Abstract

This paper presents a hybrid model combining the extreme gradient boosting machine (XGBoost) and the whale optimization algorithm (WOA) to predict the bearing capacity of concrete piles. The XGBoost provides the ultimate prediction from a set of explanatory experiment variables. The WOA, which is configured to search for an optimal set of XGBoost parameters, helps increase the model's accuracy and robustness. The hybrid method is constructed by a dataset of 472 samples collected from static load tests in Vietnam. The results indicate that the hybrid model consistently outperforms the default XGBoost model and deep neural network (DNN) regression. In an experiment of 20 runs, the proposed model has gained roughly 12, 11.7, 9, and 12% reductions in root mean square error compared to the DNN with 2, 3, 4, and 5 hidden layers, respectively. The Wilcoxon signed-rank tests confirm that the proposed model is highly suitable for concrete pile capacity prediction.

✉ Xuan-Linh Tran
tranxuanlinh@duytan.edu.vn

Hieu Nguyen
trunghieu.nguyen@fulbright.edu.vn

Minh-Tu Cao
mtcao@nycu.edu.tw

Thu-Hien Tran
tranthuhien4@dtu.edu.vn

Nhat-Duc Hoang
hoangnhatduc@duytan.edu.vn

1   Fulbright University Vietnam, 105 Ton Dat Tien, Tan Phu Ward, District 7, Ho Chi Minh City, Vietnam

2   Department of Civil Engineering, National Yang Ming Chiao Tung University, No. 1001, Daxue Rd., East Dist., Hsinchu City 300093, Taiwan

3   Institute of Research and Development, Duy Tan University, P809 - 03 Quang Trung, Danang 550000, Vietnam

4   Faculty of Civil Engineering, Duy Tan University, P809 - 03 Quang Trung, Danang 550000, Vietnam

1 Research background and motivation

A deep foundation is a common and obligatory type of foundation to support superstructure having heavy loads or laying on weak ground. Beside drilled shafts, driven piles made of timber, steel, precast concrete, and composite are also an effective solution in terms of cost and quality. In order to design pile foundation, the axial pile bearing capacity is regarded as the most important parameter. Therefore, estimating this parameter has been the subject of numerous theoretical and experimental studies in geotechnics.

Overall, there are five main methods to evaluate the pile bearing capacity, namely the static analysis, dynamic analysis, dynamic testing, pile load test, and in-situ testing [49, 50, 57]. Design guidelines based on static analysis often recommend using the critical depth concept. However, the critical depth is an idealization that has neither theoretical nor reliable experimental support, and it contradicts physical laws.

Dynamic analysis methods are based on wave mechanics for the hammer-pile-soil system. The ambiguity in the hammer impact effect, as well as changes in soil strength from the conditions at the time of pile driving, and at the

time of loading, causes uncertainties in bearing capacity determination. Dynamic testing methods are based on monitoring acceleration and strain near the pile head during driving. However, the measurements can only be analyzed by an experienced person. In addition, another considerable limitation is that the capacity estimation is not available until the pile is driven [53]. The pile load test, a field measurement of full-scale pile settlement subjected to static load, is believed to provide the most accurate results. However, this method is time-consuming and costly [59]. Therefore, developing a simple, economical, and accurate method is highly desired.

The measurements of soil properties by in-situ test methods have developed rapidly since 1970's. Concurrent with this development is the increasing use of in-situ test data in prediction of pile bearing capacity. The common tests include: standard penetration test (SPT), cone penetration test (CPT), flat dilatometer test (DMT), pressuremeter test (PMT), plate loading test (PLT), dynamic probing test (DP), press-in and screw-on probe test (SS) and field vane test (FVT). Each test applies different loading schemes to measure the corresponding soil response in an attempt to evaluate material characteristics. Among these in-situ test data, the SPT is commonly used to predict the bearing capacity of piles [5, 7].

Different SPT data based methods for determining the bearing capacity of piles have been proposed in the literature. They can be categorized into two main approaches, direct and indirect methods. Among the two, the direct methods are more widely accepted among field engineers due to the ease of computation. For example, [2, 4, 8, 18, 34, 58] proposed SPT direct methods for sandy or clayed soil proposed. For a case study in Iran [59], the authors analyzed the pile by the finite element method and compared it with four different SPT direct methods to find a reasonable prediction for its bearing capacity. However, according to [57], all of these empirical formulations have some inadequacies. Therefore, researchers have been exploring other ways to utilize SPT data to predict pile bearing capacity. Previous studies show that using machine learning algorithms is a viable option [24].

Machine learning (ML), a branch of artificial intelligence, that mimics the operation of the human brain, can nonlinearly infer new facts from adaptively learning historical data [36, 43, 52]. Moreover, the performance of machine learning (ML)-based models can be improved gradually along with the increase of learning data, so they can be kept up-to-date with the high requirements of accuracy for complex engineering problems. Many contributions have demonstrated the effectiveness and efficiency of ML-based models to deal with civil engineering-related problems, for example, predicting mechanical properties (compressive/tensile strength/shear) of hardened

concrete [10, 23, 28], the ultimate bond strength of corroded reinforcement and surrounding concrete [25], the bearing capacity of piles [11, 13], the pulling capacity of ground anchor [14, 56], etc.

ML-based models, especially Artificial Neural Network (ANN), have been extensively used to predict pile bearing capacity. Early works in this direction include [30, 61], where ANN with error back propagation is utilized. In [39], a combination of ANN and Genetic Algorithm (GA), where the weights of ANN is optimized by GA, is trained using data from 50 dynamic load tests conducted on precast concrete piles. A similar approach is proposed in [40], where in addition to GA, Particle Swarm Optimization (PSO) is utilized to optimize ANN connection weights. GA is also used to select the most important features in the raw dataset in the application of ANN to predict the bearing capacity of piles [50]. ML-based techniques other than ANN have also been considered, for instance, Samui [55] uses Support Vector Machine (SVM), Pham et al. [49] investigates Random Forest and Chen et al. [13] studied the neuro-imperialism and neuro-genetic methods.

Based on this literature review, it is clear that ANN is the current state-of-the-art method due to its black-box nature and ease of use. However, the prediction accuracy provided by ANN can be improved, and the prediction robustness with respect to different data modeling methods should be studied more thoroughly. There is also a need to investigate more advanced machine learning algorithms. In this paper, we propose to use the Extreme Gradient Boosting (XGBoost) [12], an ensemble tree model, to predict axial pile bearing capacity from a set of influential variables and a large-scale dataset. The XGBoost is the winning algorithm in multiple machine learning competitions. It has also become a very popular algorithm in science and engineering.

Le et al. [29] recently utilized an XGBoost-based ensemble model for predicting the heating load of buildings for smart city planning and concluded that the proposed ensemble model is the most robust in comparing with other machine learning models, including a standard XGBoost model, SVM, Random Forest (RF), Gaussian Process (GP), and Classification and Regression Trees (CART). Nguyen et al. [42] recently demonstrates that the XGBoost is a promising tool to assist civil engineers in forecasting deflections of reinforced concrete members. In addition, the outstanding performance of XGBoost-based models has been further convincingly demonstrated in a variety of practical problems [19, 20, 31, 60, 67, 68].

Motivated by the successes of XGBoost-based ensemble models, this study aimed to investigate an XGBoost-based model to predict the bearing capacity of reinforced concrete piles and compared its performance with that of the deep ANN, which is a popular machine learning model

used for regression analysis. For the XGBoost, as well as other machine learning models, determining an optimal set of hyper-parameters to acquire the ultimate generalization of models is a crucial task [15, 66]. Nevertheless, most of the previous studies, which are related to machine learning-based pile bearing capacity estimation, relied on manual or simple grid search method for hyper-parameter setting [1, 3, 17]. Employing the default hyper-parameters, performing model selection on the basis of experience, and selecting hyper-parameters based on trial-and-error processes usually lead to suboptimal performances. Systematic and automated approaches for model selection should be used to effectively construct machine learning-based pile bearing capacity prediction models.

Accordingly, using metaheuristic approaches (i.e. heuristic optimization algorithms), such as, the PSO, GA, bat algorithm, history-based adaptive differential evolution, differential flower pollination, etc., to find an optimal set of parameters for machine learning models has become popular [25, 42, 44–46, 63]. In this research, we propose to use the Whale Optimization Algorithm (WOA) to optimize the search for the best parameter configuration of XGBoost. WOA is modeled after humpback whales' prey searching and feeding behaviors [35]. The motivation for selecting the WOA is that it is one of the current state-of-the-art metaheuristics and has been successfully used in my applications in engineering [22].

Our proposed hybrid model WOA-XGBoost has XGBoost as the main machine learning-based prediction method, which can estimate the axial pile bearing capacity from a set of explanatory variables. The WOA is configured to search for an optimal set of XGBoost parameters that yields the smallest root mean squared error. The proposed model is tested using a dataset of 472 samples of static load pile tests. The hybrid model demonstrates superior performances over those of default XGBoost, and especially deep ANN, a popular model in previous works. In summary, the main contributions of the current paper can be stated as follows:

(i)   Machine learning-based models for estimation of pile bearing capacity are constructed and verified using a large-scale dataset consisting of 472 pile test experiments. It is noted that most of the previous studies only relied on small-scale datasets [16, 38, 39].

(ii) Although various models have been proposed [3], few studies have been dedicated to the investigation of XGBoost, a currently state-of-the-art regressor, for estimating pile bearing capacity.

(iii) Since the problem of model selection (i.e. the determination of a suitable set of hyper-parameters) is crucial [6], a novel hybridization of WOA and XGBoost-based regression machine is proposed and verified.

The rest of the paper is organized as follows. In Sect. 2, research methodology, the dataset, experiment variables as well as the main ideas of WOA and XGBoost are presented. The hybrid model WOA-XGBoost is described in Sect. 3. In Sect. 4, experiments are presented and the performances are compared and discussed. The paper ends with some conclusions in Sect. 5.

2 Research methodology

2.1 The collected dataset of static load tests

To train and validate the proposed machine learning method, this study relies on a dataset of static load test of driven reinforced concrete piles. This dataset includes 472 compiled in the previous work of [50]. This is a fairly large dataset and highly appropriate for constructing and verifying sophisticated machine learning models. It is noted that precast piles with closed tips are driven into soil layers with the employment of hydraulic pile driven machine to record capacity of piles. Figure 1 demonstrates the experimental set-up used for data measurement. Figure 2 provides illustrations for the pile structure, its geometrical variables, and soil stratigraphy. Table 1 summarizes the ten conditioning factors employed to predict the dependent variable Y which is the axial pile bearing capacity. In addition, this table also reports statistical descriptions of the predictor and dependent variables.

2.2 Whale optimization algorithm

The Whale Optimization Algorithm was first introduced by Mirjalili and Lewis [35]. It is a swarm-based metaheuristics algorithm whose exploration and exploitation phases are modeled after humpback whales' prey searching and feeding behaviors. In nature, after identifying the prey (usually a school of small fishes or plankton) humpback whales dive down nearly 12 m and start to swim upward to the surface in a spiral trajectory while creating bubbles to form a virtual net to herd and corral the prey. This is referred to as the bubble-net feeding method [9, 51].

More specifically, the algorithm considers a pod of n whales hunting for food. The coordinates of their positions $\vec{X_i}$ are candidates for the parameters need to be optimized. In the exploitation (feeding) phase, the current best position, $\overrightarrow{X^*}$, is treated as the target prey. To mimic the prey shrinking encircling behavior, the position of the ith whale in the $(t + 1)$th iteration is updated by:
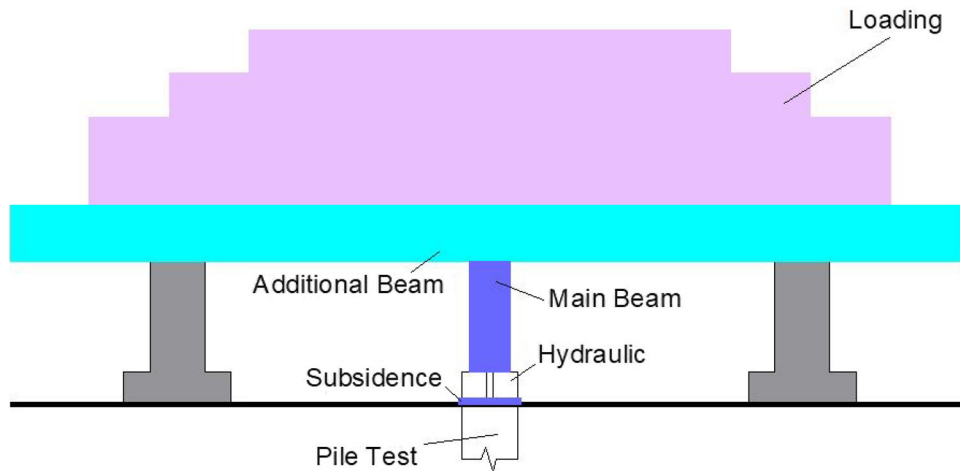
Fig. 1 Static load test experiment for measuring pile bearing dcapacity

$$\overrightarrow{X_i}(t+1) = \overrightarrow{X^*}(t) - A\vec{D} \tag{1}$$

$$\vec{D} = |C\,\overrightarrow{X^*}(t) - \overrightarrow{X_i}(t) \tag{2}$$

where A and C are coefficients determined by:

$$A = 2a \cdot r - a \text{ and } C = 2r \tag{3}$$

where r is a random number in $[0, 1]$ and a linearly decreases from 2 to 0.

In addition, the following equation is utilized to imitate the position spiral updating mechanism:

$$\overrightarrow{X_i}(t+1) = \vec{D}'e^{bl}\cos 2\pi l + \overrightarrow{X^*}(t) \tag{4}$$

where $\overrightarrow{D}' = |\overrightarrow{X^*}(t) - \overrightarrow{X_i}(t)|$, b is a predefined parameter and l is a random number in $[-1, 1]$.

As whales swim in shrinking circles and a spiral trajectory simultaneously, a random number $p \in [0, 1]$ can be used to determine the particular behavior a whale exhibits in a given iteration.

In the exploitation phase, the algorithm has assumed that the current best position is the location of the prey. However, this is not always true and whales often search randomly according to the position of other whales in the pod. This exploration is modeled using the following equations:

$$\overrightarrow{X_i}(t+1) = \vec{X}_{\textbf{rand}}(t) - A\vec{D} \tag{5}$$

$$\vec{D} = \left| C\vec{X}_{\textbf{rand}}(t) - \overrightarrow{X_i}(t) \right| \tag{6}$$

where $\overrightarrow{X}_{\text{rand}}$ is the position of a random whale in the population other than the ith.

The decision whether to search locally (exploitation) or globally (exploration) is made upon checking if $|A| \leq 1$. The pseudo code of the algorithm is presented in Fig. 3.

For more information, new variants and applications of WOA, we refer the reader to [22, 32, 35] and the references therein.

2.3 Extreme gradient boosting machine (XGBoost)

The XGBoost is an open-source library that provides machine learning algorithms, both regression and classification, in the *gradient boosting framework* [21, 33]. It originated from an academic research project but has become a widely used library in both academia and industry [65]. The library is highly efficient, flexible, and portable. It supports multiple languages, including C+ +, python, R, etc. The library also supports distributed training on clusters on cloud computing platforms, such as Amazon Web Services, Google Cloud Platform and Microsoft Azure.

Under the hood, the XGBoost algorithm builds a series of weak learners, which are *classification or regression trees* (CART) [54, 64]. These weak learners are then combined to form the final prediction model. Like other boosting methods, XGBoost do not build all the regression trees at the same time but step by step. The tree in the current step is constructed in such a way that it minimizes the average value of the loss function of all the steps on the training set.

More specifically, let the training data be $D = \{x_i, y_i\}_{i=1}^{n}$, where $x_i \in R^m$ is an input vector with m features, and $y_i \in R$ is the corresponding output. Assume $\widehat{y}_i^{(t-1)}$ is the prediction output at step $t-1$. Then, at step t, the XGBoost builds the tree that minimizes the following objective function:
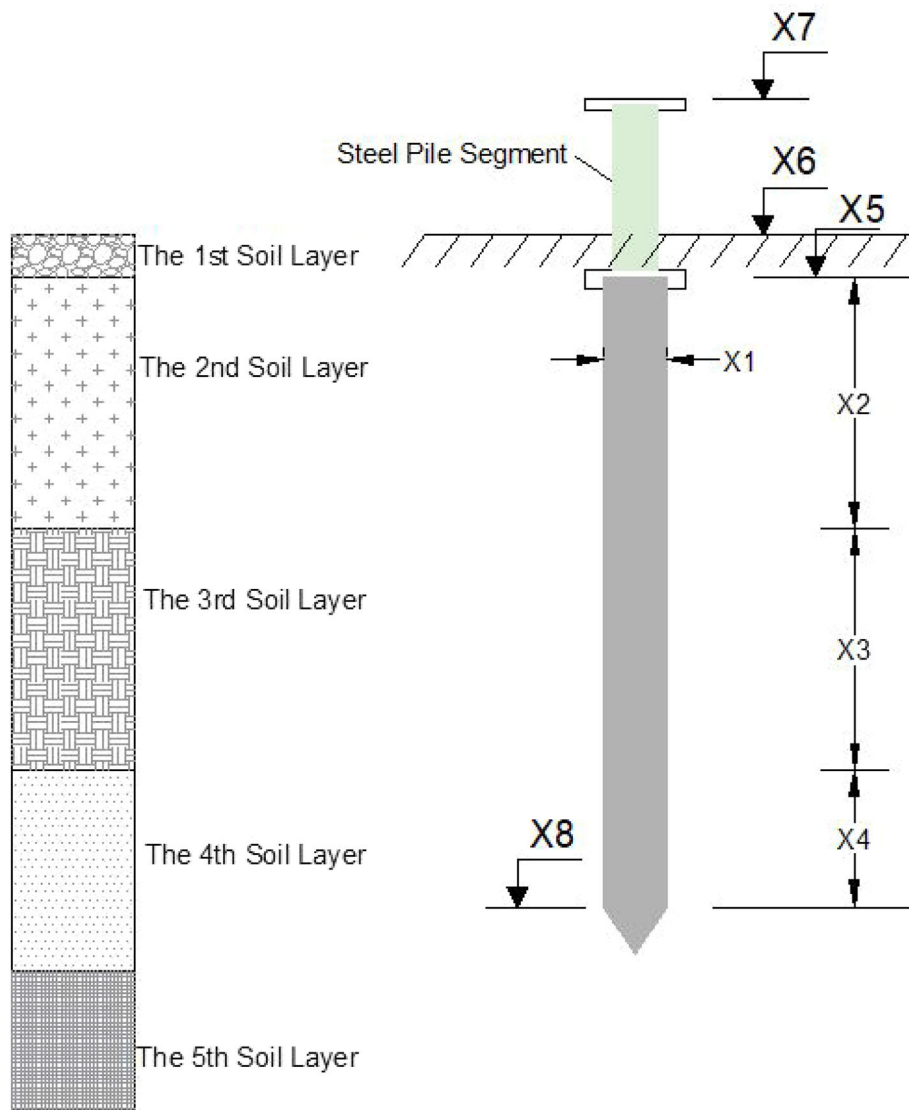
Fig. 2 Demonstration of the pile structure and soil stratigraphy

$$L^t = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t)}\right) + \Omega(f_t)$$

$$= \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (7)$$

where l is any convex and differentiable loss function measuring the difference of the prediction and the provided output, $f_t : R^m \to R, f_t(x) = w_{q(x)}$ is the prediction function of the tree, where $w \in R^T$ is the vector of scores on leaves, $q : R^m \to \{1, 2, \ldots, T\}$ is a function assigning each data point to the corresponding leaf, and T is the number of leaves. The last term, Ω, is the regularization term. Its purpose is to reduce overfitting, a common issue in machine learning. This term penalizes complex trees with many leaves and gives priority to more simple and predictive trees, more specifically

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} w_j^2 \quad (8)$$

where γ and λ are parameters.

Approximating the right hand side of Eq. (6) by using the second-order Taylor expansion of l w.r.t to the second variable we have:

$$L^t \approx \sum_{i=1}^{n} \left[l\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad (9)$$

Table 1 Statistical descriptions of the employed variables

Variables	Notation	Min	Average	Std	Max
Diameter of pile (mm)	X1	300.00	363.77	48.12	400.00
The thickness of the first soil Layer that pile embedded (m)	X2	3.40	3.83	0.48	5.72
The thickness of the second Soil layer that pile embedded (m)	X3	1.50	6.58	1.64	8.00
The thickness of the third Soil layer that pile embedded (m)	X4	0.00	0.33	0.46	1.69
Pile top elevation (m)	X5	0.68	2.80	0.62	3.40
Natural ground elevation (m)	X6	3.04	3.50	0.08	4.12
The elevation of extra segment pile top (m)	X7	1.03	2.92	0.60	4.35
The depth of pile tip (m)	X8	8.30	13.54	1.80	16.09
The average SPT blow count along the pile shaft	X9	5.60	10.74	2.26	15.41
The average SPT blow count at the pile tip	X10	4.38	7.06	0.66	7.75
The axial bearing capacity load of pile (kN)	Y	407.20	984.20	353.21	1551.00

The Whale Optimization Algorithm

Initialize the whale population $\vec{X}_i$ ($i = 1, 2, \dots, n$).
Calculate the fitness of each whale
X*=the current best position
For each iteration t
    For each whale
        Update a, A, C, l and p
        If ($p < 0.5$) then
            If ($|A| < 1$) then
                Update the position of the current whale using **Eq. (1)**
            Elseif ($|A| \geq 1$) then
                Randomly choose a whale (X_{rand})
                Update the position of the current whale using **Eq. (4)**

            Endif
        Elseif ($p \geq 0.5$)

            Update the position of the current whale using **Eq. (3)**
        Endif
    Endfor
    Fix the positions of any whales go beyond the search space
    Recalculate the fitness of each whale
    Update the current best position X^*
     t = t+1
Endfor
Return the current best position X^*

Fig. 3 The Whale Optimization algorithm

The XGBoost algorithm

Specify the algorithm parameters
Sort the input set by feature values
For each iteration t
    Calculate g_i and h_i
        Examine the current tree and decide the best split using **Eq. (12)**
        The weights of the new tree $f_t(x)$ is computing using **Eq. (11)**
        The new prediction is computed using **Eq. (13)**
Return the trained XGBoost model

Fig. 4 The XGBoost algorithm

where $g_i = \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ and $h_i = \partial^2_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$.

Removing the term $\sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)}\right)$, which does not depend on the choice of the decision tree in the current step, from Eq. (8) and collecting the terms associated with the same score (the data points on the same leaf get the same score), we are left with a simpler objective function that needs to be minimized:

$$
\tilde{L}^t = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t)
$$
$$
= \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T
$$
(10)

where I_j is the subset of the input set associated with leaf j, i.e., $I_j = \{i : q(x_i) = j\}$.

It can be noted that the first term of Eq. (10) is a sum of independent quadratic functions in w_j. Therefore, the optimal w_j^* and the minimal objective $\tilde{L}^{*t}$ are given by:

$$
w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}, \quad \tilde{L}^{*t}(q) = -\frac{1}{2} \sum_{j}^{T} \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T
$$
(11)

This equation can be used to measure how good a tree is as a candidate for the current step. However, these optimal values can only be calculated when the structure of the tree in the current step has already determined. As it is not feasible to consider all the possible tree structures, the XGBoost built trees iteratively.

At the beginning, the XGBoost sorts the input data set according to feature values to form a tree with zero depth. Then in each step, a new tree is created by an optimal branch splitting. According to Eq. (11), this splitting, which maximizes the lost reduction, is calculated by

$$
\tilde{L}_{\text{split}} = -\frac{1}{2} \left[\frac{\left(\sum_{i \in I_L} g_i \right)^2}{\sum_{i \in I_R} h_i + \lambda} + \frac{\left(\sum_{i \in I_L} g_i \right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i \right)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma
$$
(12)

where I_L is the subset of input indices on the left of the split and I_R is the subset of indices on the right of the split.

The XGBoost algorithm is summarized in Fig. 4. For more detailed information, we refer the reader to [41, 42, 64]. In addition to using the regularization term Ω in Eq. (7), the XGBoost let users specify two parameters, max_depth and learning rate η, to combat overfitting. The parameter max_depth, as suggested by the name, limits the maximal tree depth that is allowed by XGBoost. The possible range for this parameter is $[0, \infty]$ and the default value is max_depth = 6. The learning rate, also called shrinkage, scales the prediction of newly built tree by a factor $0 < \eta < 1$ to reduce the influence of each individual tree and allow trees in the later steps chances to improve the model. More specifically:

$$
\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(x_i)
$$
(13)

The default value of η is 0.3.

3 The proposed WOA-XGBoost model for intelligent estimation of pile bearing capacity

Our study proposes to combine XGBoost and WOA into a hybrid model of WOA-XGBoost to predict the bearing capacity of concrete piles. In this hybrid model, XGBoost is served as the main prediction machinery which establishes a function that can derive the axial pile bearing capacity from a set of explanatory variables, including pile diameter (X1), thickness of the first soil layer (X2),
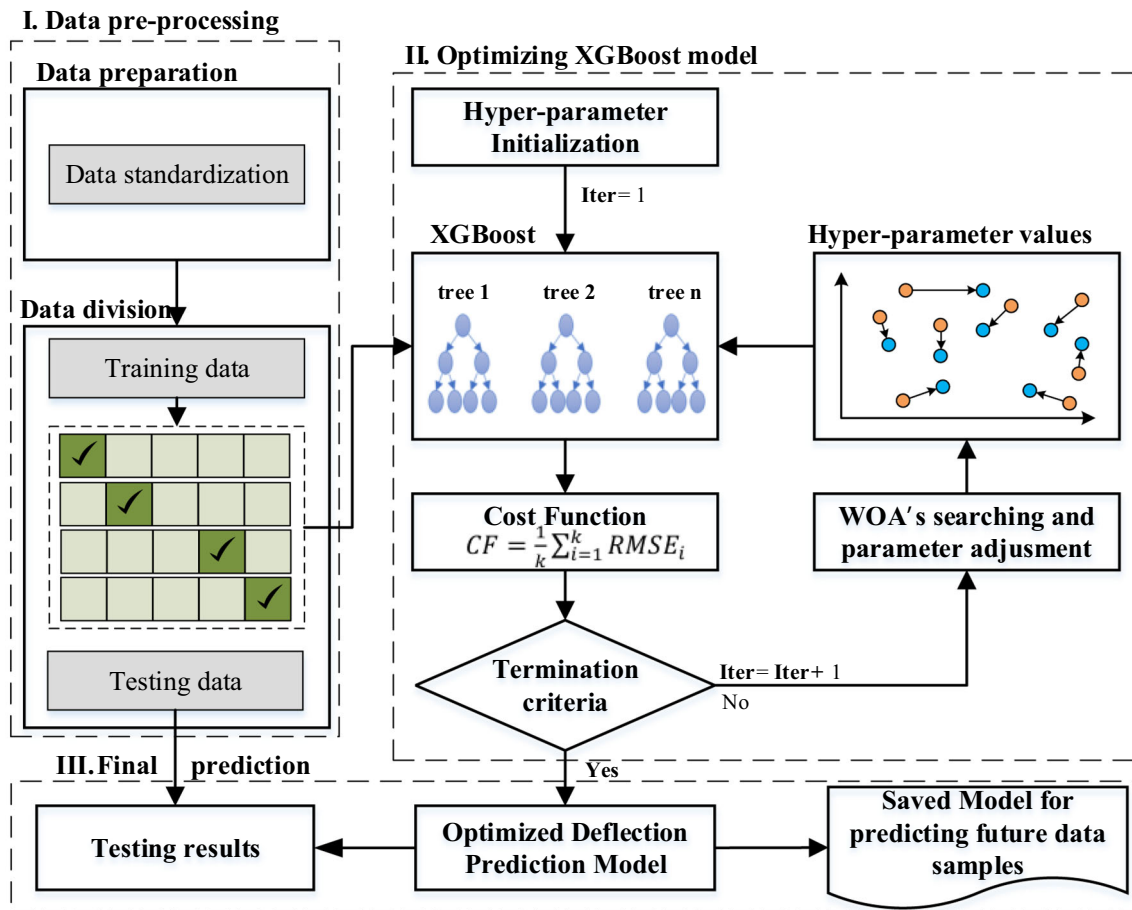
Fig. 5 Operation procedure of the WOA-XGBoost model

thickness of the second soil layer (X3), thickness of the third soil layer (X4), elevation of the natural ground (X5), top of pile elevation (X6), elevation of the extra segment of pile top (X7), depth of pile tip (X8), mean value of SPT blow count along pile shaft (X9), and mean value of SPT blow count at pile tip (X10). However, even though the XGBoost algorithm is quite robust, its performance does depend on the choice of its parameters. Therefore, the WOA is utilized to determine an optimal set of the XGBoost parameters. Once these optimal parameters have been determined, the corresponding XGBoost model is then used to train and provide the final prediction.

The complete work flow of the WOA-XGBoost is illustrated in Fig. 5. It consists of three phases: *preprocessing*, *parameter optimization* and *final prediction*. In the preprocessing phase, the input data is normalized to have zero mean and roughly similar magnitudes using the following Z-score formula:

$$X_N = \frac{X_O - m_X}{s_X} \tag{14}$$

where X_N and X_O are the normalized and original feature variables, m_X and s_X are the mean and standard deviation of the considered feature in the whole input data. Finally, the whole data set is spit into a training data set and a testing data set. The training data set is used for training the model and optimizing parameters. The testing set is only used to benchmark the performance of trained models.

In the parameter optimization phase, in order for WOA to be able to find optimal XGBoost parameters, we need to provide it with a way to evaluate how good a given set of parameters is. This is done via formulating an objective/cost function based on Root Mean Square Error (RMSE) and k-fold cross-validation. The training data set is further split into k folds (subsets of roughly equal size). Given a set of parameters, the associated XGBoost model is trained and validated k times using different training and validating sets. For each fold, the model is trained with data

Convergence History
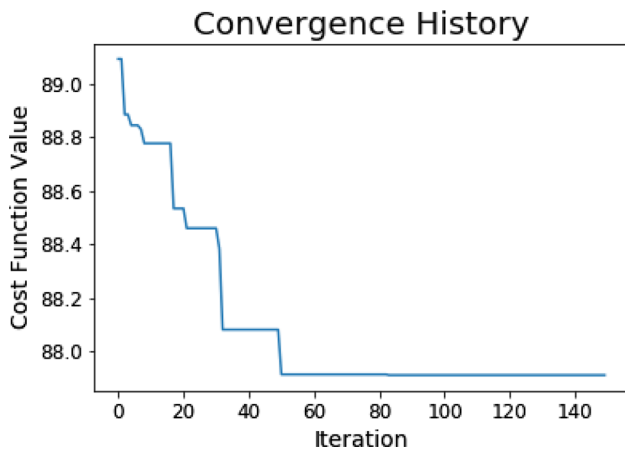
Fig. 6 Convergence curve of the cost function in the parameter optimization phase

from the other $k-1$ folds and is validated against the data in the current fold. The cost function is defined as the average of the validating RMSE, more specifically:

$$\mathbf{CF} = \frac{1}{k} \sum_{i=1}^{k} \mathbf{RMSE}_i, \tag{15}$$

where $\mathbf{RMSE}_i$ is the validating RMSE associated with the ith fold (when the training set is the union of other folds):

$$\mathbf{RMSE}_i = \sqrt{\frac{\sum_{j \in S_i} \left(Y_{A,j} - Y_{P,j}\right)^2}{|S_i|}} \tag{16}$$

where S_i is the index set of the ith fold, $|.|$ denotes the cardinal of a set, $Y_{A,j}$ is the actual output for the jth data set and $Y_{P,j}$ is the corresponding predicted output.

The WOA search stops when the cost function does not improve (reduce) after a certain number of iterations or when the prescribed maximal number of iterations has been reached. When an optimal set of parameters has been determined, the associated XGBoost model is used in the final prediction phase. Then, the predicted values of pile bearing capacity can be obtained and documented.

4 Experimental results and discussion

4.1 Experiment Setup

In the experiments, the integrated WOA-XGBoost model was developed in Python with the uses of the following packages: (i) XGBoost Python package version 0.90, the official implementation of XGBoost in Python (ii) mealpy version 1.1.0, a Python module of cutting-edge nature-

Table 2 An optimal set of tuned parameters

Model	Parameters	Value
XGBoost	Learning rate (η)	0.11769209
	Max depth	3.52531112

inspired meta-heuristic algorithms, including WOA [62] (iii) scikit-learn version 0.23.2, a machine learning library for Python language. All experiments of the WOA-XGBoost model were performed on a laptop computer equipped with Intel® Core™ i5-3437U CPU @ 1.90 GHz × 4, 8 GB DDR3 RAM 1600 MHz and the operating system Ubuntu 20.04.4 LTS.For WOA, we set the whale population (pop_size) to be 50 and the maximal number of iterations (epoch) to be 150. It is configured to optimize two XGBoost parameters, namely the learning rate (η) and the maximal tree depth (maxdepth). The search ranges of η and maxdepth are $[0.05, 0.3]$ and $[3,10]$, respectively.

For XGBoost, the objective is set to be reg: squarederror (regression with square error) and the number of boost rounds (number of iterations) is chosen to be 100. The data consisting of 472 samples is randomly split into two sets: a training set of 424 samples (90%) and a testing set of 48 samples (10%). During the parameter optimization phase, the training set is further partitioned into 5 folds of roughly equal size to evaluate the cost function according to Eqs. (15) and (16).

In Fig. 6, the convergence history of the cost function in the parameter optimization phase of a typical run of WOA-XGBoost is illustrated. It can be seen that, in this particular run, the cost function decreases very quickly at the beginning and stall after roughly 50 iterations. Currently, early stopping has not been supported in the implementation of the WOA algorithm. Hence, the code continues to run until the maximum iteration of 150 is achieved. However, it is necessary to keep the maximal number of iterations of WOA at 150 to maintain the robust performance of the optimization process. It is because the number of iterations required for the WOA convergence may vary. In Table 2, the tuned learning rate and max depth parameters of XGBoost obtained by WOA in this particular run are given. It should be noted that as tree depth is an integer parameter, the effective value of the maxdepth found is actually 3. In addition, when the cost function is computed, the XGBoost model is trained and tested five times to perform a fivefold cross-validation. On average, each training and testing phase last 0.082511 and 0.002290 s, respectively. In total, the WOA-XGBoost

Table 3 Performance comparison of XGBoost and WOA-XGBoost

Phase	Metrics	XGBoost	WOA-XGBoost
Training	RMSE	39.64	62.74
	MAE	21.02	47.92
	MAPE (%)	2.51	5.21
	R^2	0.99	0.97
Testing	RMSE	101.30	87.72
	MAE	73.77	64.55
	MAPE (%)	7.78	7.34
	R^2	0.92	0.94

model requires nearly 1 h and 15 min (4454 s) to find a set of optimal XGBoost parameters.

In order to accurately assess the predictive capability of different models, the following performance metrics are considered: RMSE, the mean absolute percentage error (MAPE), the mean absolute error (MAE) and the coefficient of determination (R^2). RMSE has been introduced in the previous section. MAPE, MAE, and R^2 are given by:

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^{N} \frac{\left|Y_{A,i} - Y_{P,i}\right|}{Y_{A,i}} \tag{17}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \frac{\left|Y_{A,i} - Y_{P,i}\right|}{Y_{A,i}} \tag{18}$$

$$R^2 = 1 - \sum_{i=1}^{N} \left(Y_{A,i} - Y_{P,i}\right)^2 \Big/ \sum_{i=1}^{N} \left(Y_{A,i} - \overline{Y}\right)^2 \tag{19}$$

where $Y_{A,i}$ and $Y_{P,i}$ are the actual and the predicted bearing capacity, respectively; N is the number of data instances; $\overline{Y}$ is the average of actual values. It should be noted that smaller RMSE, MAE or MAPE is better, while higher R^2 is better.

4.2 Comparison between the tuned and the default model

Table 3 presents the comparison in performance of XGBoost with default parameters recommended by the XGBoost toolbox and WOA-XGBoost with parameters

optimized by the metaheuristic algorithm. We can see that in the training phase, default XGBoost is better than WOA-XGBoost in all metrics. The RMSE and MAE of XGBoost are almost 37 and 56%, respectively, smaller than those of WOA-XGBoost. However, in the testing phase the opposite holds. WOA-XGBoost outperformed XGBoost in all the metrics. The RMSE and MAE of WOA-XGBoost are 13.4 and 12.4%, respectively, smaller than those of XGBoost. In the testing phase, the MAPE and R^2 of WOA-XGBoost are also noticeably better than those of XGBoost. This result indicates that tuning parameters using WOA does help WOA-XGBoost to reduced overfitting and provide more accurate prediction.

All of the above results are exclusive to the considered partition of the data. They could be biased toward this specific partition. Therefore, we further compare the performance of XGBoost (with default parameters) and WOA-XGBoost (with the optimal parameters found by WOA and listed in Table 2) on other 20 randomly partitions of the data with the same splitting ratio (9:1). In Fig. 7, the distributions of the residuals of both models are presented. It can be seen that the distribution is more symmetric and closer to a normal distribution in the case of the tuned model. This is the first indication that WOA-XGBoost is more robust than the default XGBoost. In Fig. 8, model prediction capability is compared. In the training phase, the default XGBoost model fits the data better (the predicted output is closer to the line of best fit). However, in the testing phase, the tuned model (WOA-XGBoost) provides a more accurate prediction. Thus, it is apparent that tuned hyper-parameters have helped XGBoost be less prone to overfitting.

The comparison of performance in the testing phase is presented in Fig. 9. In terms of RMSE, the error incorporated in the objective/cost function, the tuned model is clearly the winner. In 20 experiments, there are only two cases where the RMSE of the tuned model is larger than that of the default model. In term of MAE and MAPE, the tuned model is no longer the clear winner, but it is still better in more than half of the occasions. In terms of R^2, an important benchmark on how well a model predicts the observed data, the tuned model again beats the default model. There is only one instance where the R^2 of the tuned

Table 4 Wilcoxon signed-rank test the claim that WOA-XGBoost is better than XGBoost in different benchmarks

Benchmark	p value	p value < 0.05	Test outcome
RMSE	0.00020	Yes	Accept claim (H_1)
MAE	0.01812	Yes	Accept claim (H_1)
MAPE	0.26084	No	Reject claim (H_1)
R^2	0.00016	Yes	Accept claim (H_1)

model is smaller (worse) than that of the default model. However, even in that case, the R^2 of the tuned model is already 0.9, which indicates good prediction capability.

In order to have a more reliable conclusion, we perform Wilcoxon signed-rank tests on the claim that WOA-XGBoost is better than default XGBoost in a specific metric. Here, "better" means smaller RMSE, MAE, and MAPE, but also means greater R^2. The hypotheses are as follows:

- H_0: The (pseudo) median of the benchmark of WOA-XGBoost in the 20 experiments is the same as that of default XGBoost.

- H1: The (pseudo) median of the benchmark of WOA-XGBoost in the 20 experiments is better than that of default XGBoost (claim).

The results of the test are summarized in Table 4. At $\alpha = 0.05$, there is not enough evidence to accept the claim for MAPE, but there are strong evidences to support the claim for RMSE, MAE and R^2. Especially, the decision is very conclusive for RMSE and R^2 because the p-values are very small. In summary, the tuned model is more accurate and reliable than the default model because it is much better in the objective benchmark, RMSE, and in important benchmarks, namely MAE and R^2.
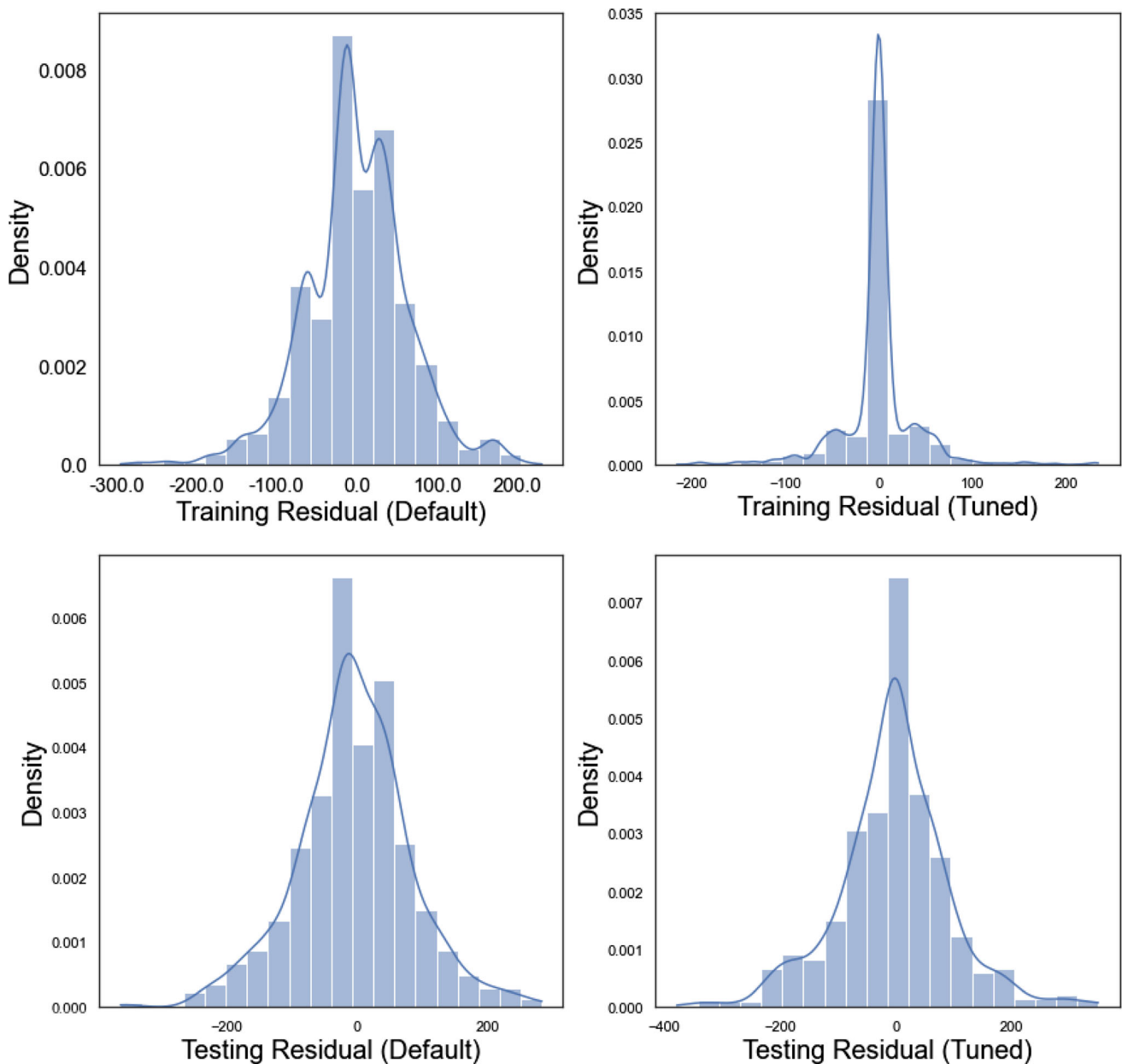
Fig. 7 Residual comparison
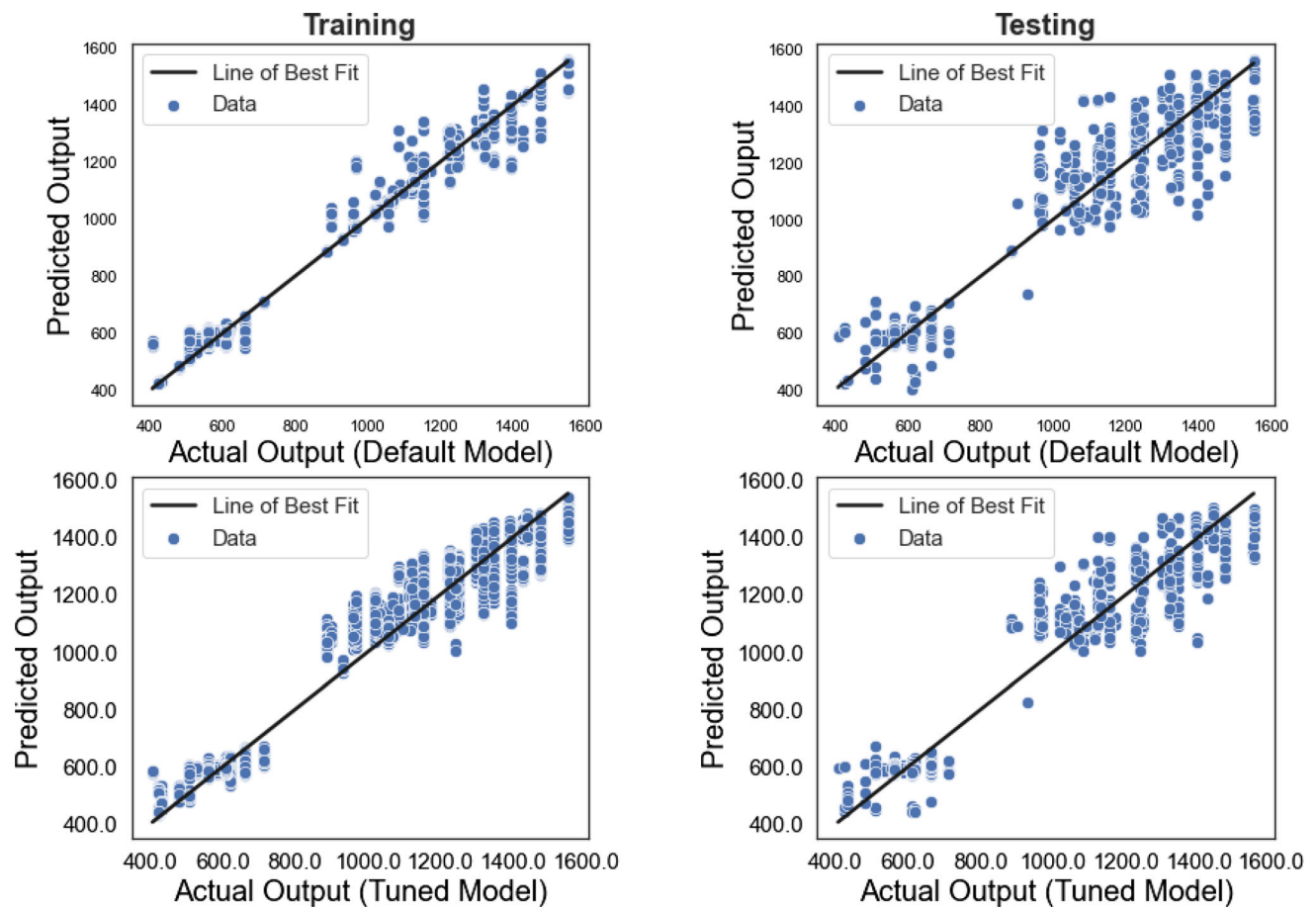
Fig. 8 Model prediction comparison

Thus, it can be observed that even though the XGBoost algorithm is robust, its performance significantly depends on the selection of its hyper-parameters. Hence, a meta-heuristic is definitely needed to optimize the process of finding these hyper-parameters. The WOA has proven to be highly appropriate in assisting the learning phase of XGBoost. This metaheuristic shows good convergence property and helps locate a good solution of hyper-parameters of the XGBoost algorithm. Once a set of optimal parameters has been determined, the corresponding XGBoost model is then used to train and provide the final prediction.

4.3 Comparison with the benchmark approaches

In this section, to confirm the predictive capability of the newly developed hybrid model WOA-XGBoost used for pile bearing capacity prediction, its performance is compared to that of the capable machine-learning-based models based on Deep Neural Network (DNN) for regression. The DNN has been trained with the state-of-the-art Adam

optimizer [27] and implemented via the scikit-learn Python library [48]. The DNN has been constructed with 2, 3, 4, and 5 hidden layers. The DNN is selected as the benchmark models in this section because neural networks have been extensively and successfully employed in data-driven pile capacity estimation [26, 37, 39, 47]. In DNN, ReLU (Rectified Linear Unit) is employed and the number of training epochs is set to be 1000; the number of neurons in the hidden layers is selected via a fivefold cross-validation.

In Tables 5 and 6, the means and standard deviations of all the metrics are provided for all considered models in both training and testing phases. Variants of DNN with different numbers of hidden layers (2, 3, 4 and 5) appear to have similar performances, especially in the testing phase. Among them, the variant with 4 layers is the better one. However, even this variant lags behind WOA-XGBoost in both training and testing phase. The reductions in the average of testing RMSE of WOA-XGBoost compared to DNN with 2, 3, 4 and 5 layers are roughly 12, 11.7, 9 and 12%, respectively. The box plots in Fig. 10, further confirm this for the testing phase. Not only WOA-XGBoost has a
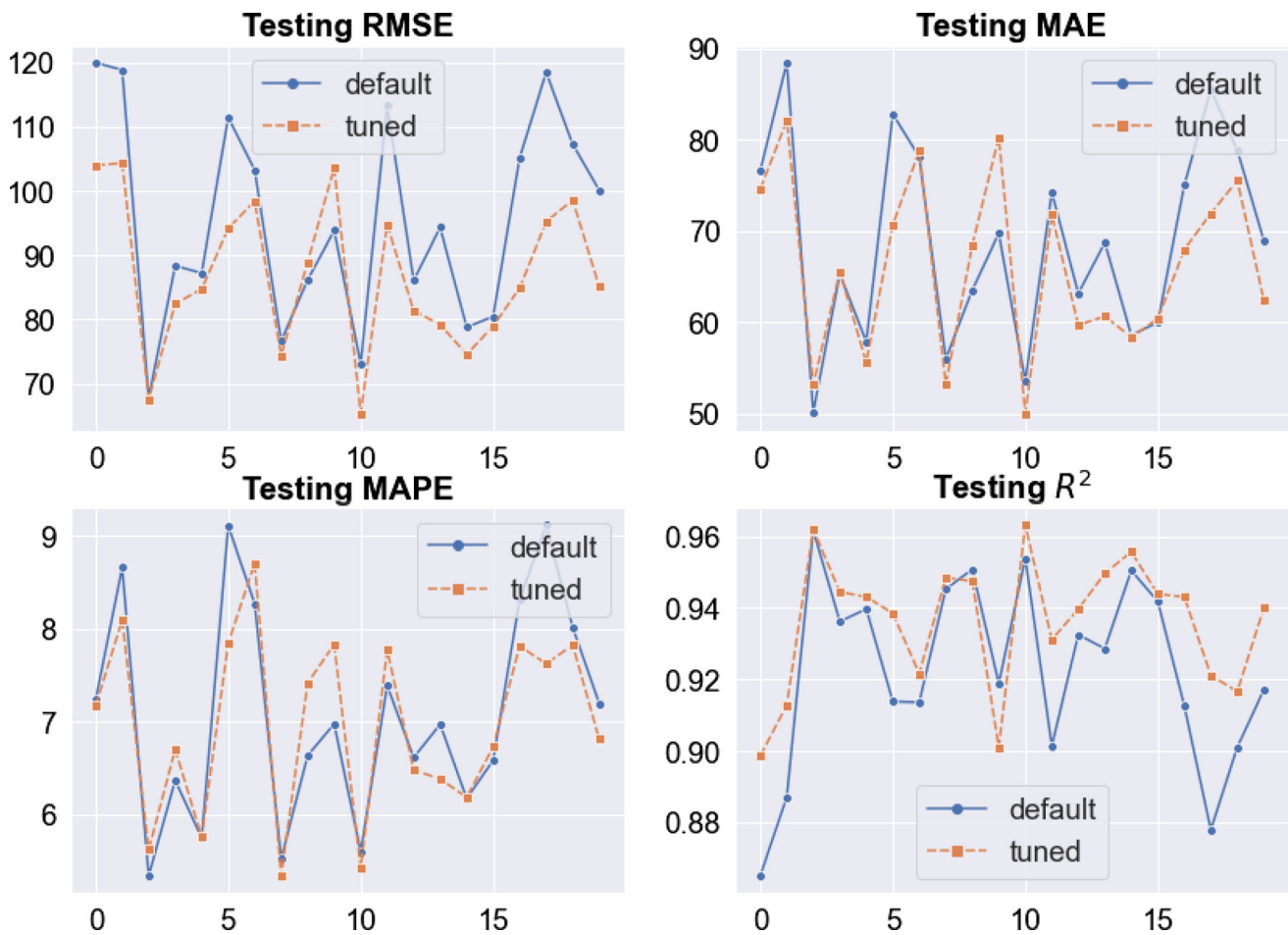
Fig. 9 Performance comparison in the testing phase

Table 5 Performance statistic in the training phase

Indices	DNN 2 layers		DNN 3 layers		DNN 4 layers		DNN 5 layers		WOA-XGB	
	Mean	Std.	Mean	Std	Mean	Std.	Mean	Std.	Mean	Std.
RMSE	93.85	3.22	89.37	4.25	90.71	1.29	94.97	3.57	64	1.28
MAE	72.47	2.11	68.94	3.4	70.19	0.73	74.35	3.59	48.44	1.06
MAPE	7.63	0.21	7.24	0.34	7.33	0.06	7.8	0.42	5.26	0.1
R^2	0.93	0.005	0.94	0.006	0.93	0.03	0.93	0.006	0.97	0.001

better mean but it also has a better median and interquartile range in all metrics. The performances of WOA-XGBoost in all metrics are also very robust. There is no outliner in the negative direction in the WOA-XGBoost, while there are relatively numerous ones in the DNN models, especially in the R^2 index.

Similar to Sect. 4.2, we perform Wilcoxon signed-rank tests to compare WOA-XGBoost with each and every variant of DNN in all considered benchmarks. The general hypotheses can be written as:

- H_0: The (pseudo) median of the benchmark of WOA-XGBoost in the 20 experiments is the same as that of DNN with n layers
- H1: The (pseudo) median of the benchmark of WOA-XGBoost in the 20 experiments is better than that of DNN with n layers (claim)

where $n = 2, 3, 4$ and 5.

The test results are summarized in Tables 7, 8, 9 and 10. It can be seen that, there is sufficient evidence to support the claim in all comparisons in all benchmarks, except for

Table 6 Performance statistic in the testing phase

Indices	DNN 2 layers		DNN 3 layers		DNN 4 layers		DNN 5 layers		WOA-XGB	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
RMSE	98.85	11.58	98.59	15.13	95.49	18.9	98.9	10.22	87.03	11.8
MAE	75.58	9.61	74.54	10.23	73.53	10.6	76.85	8.35	66.07	9.61
MAPE	8.06	0.9	7.82	0.95	7.7	0.97	8.07	0.84	6.98	0.98
R^2	0.92	0.02	0.92	0.03	0.93	0.03	0.93	0.02	0.94	0.02
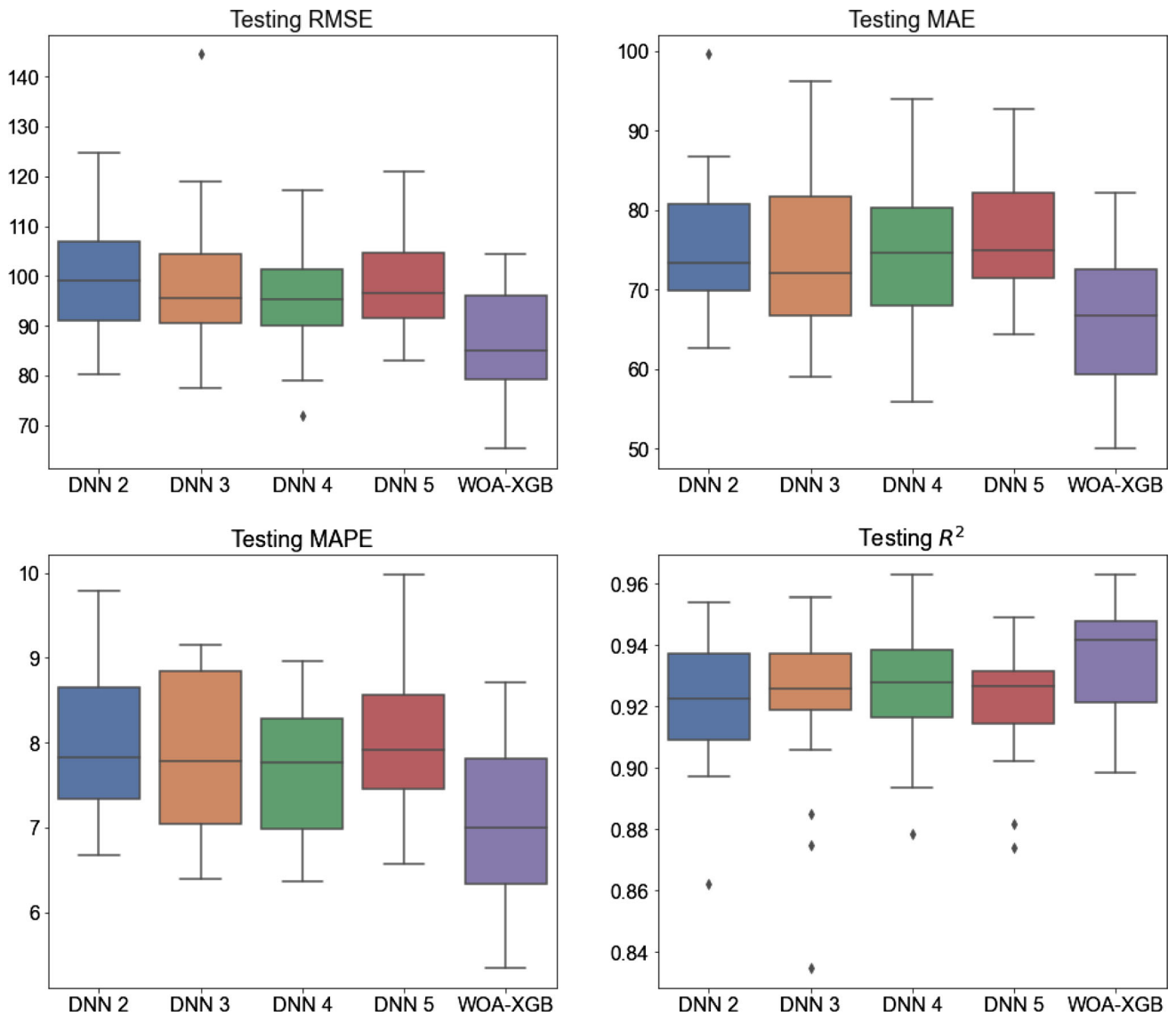
Fig. 10 Performance comparison of different models in the testing phase

Table 7 Wilcoxon signed-rank pairwise comparison in RMSE

Pairwise comparison	p value	p value < 0.05	Test outcome
WOA-XGBoost versus DNN with 2 layers	0.00279	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 3 layers	0.01332	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 4 layers	0.02422	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 5 layers	0.00099	Yes	Accept claim (H_1)

Table 8 Wilcoxon signed-rank pairwise comparison in MAE

Pairwise comparison	p value	p value < 0.05	Test outcome
WOA-XGBoost versus DNN with 2 layers	0.00472	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 3 layers	0.01074	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 4 layers	0.03186	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 5 layers	0.00158	Yes	Accept claim (H_1)

Table 9 Wilcoxon signed-rank pairwise comparison in MAPE

Pairwise comparison	p value	p value < 0.05	Test outcome
WOA-XGBoost versus DNN with 2 layers	0.01479	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 3 layers	0.02422	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 4 layers	0.03186	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 5 layers	0.00211	Yes	Accept claim (H_1)

Table 10 Wilcoxon signed-rank pairwise comparison in R^2

Pairwise comparison	p value	p value < 0.05	Test outcome
WOA-XGBoost versus DNN with 2 layers	0.01479	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 3 layers	0.02913	Yes	Accept claim (H_1)
WOA-XGBoost versus DNN with 4 layers	0.08842	No	Reject claim (H_1)
WOA-XGBoost versus DNN with 5 layers	0.00859	Yes	Accept claim (H_1)

only one inconclusive case, where R^2 of WOA-XGBoost is compared to that of DNN with 4 layers. With all of these evidences, it can be concluded that WOA-XGBoost is the best model in these experiments.

5 Concluding remarks

In this paper, we have formulated and tested the hybrid model WOA-XGBoost in predicting the bearing capacity of concrete piles. XGBoost is the crucial part of the model providing the prediction from a set of ten feature variables, including thickness of the first, second and third soil layer, pile diameter, elevation of the natural ground, of top of pile and of the extra segment of pile top, depth of pile tip, mean

value of SPT blow count along pile shaft and of SPT blow count at pile tip.

Although the XGBoost algorithm is the state-of-the-art machine learning method and is highly effective for complex function approximation. This study has proven that its performance can still be meliorated with the use of advanced metaheuristic algorithms. Accordingly, the WOA is used to find an optimal set of values of the learning rate and the maximal depth, two important hyper-parameters of XGBoost. The hybrid model is set up so that the selected set of parameters minimizes the average RMSE error in a fivefold cross-validation. The model is trained, validated, and compared on subsets of a dataset consisting of 472 samples.

The experimental results, supported by the statistical hypothesis tests, confirm that the hybridization of XGBoost

and WOA significantly outperforms the individual XGBoost model as well as the DNN-based regression models. Therefore, it is highly recommended to use in pile bearing capacity prediction. Incorporating advanced feature selection and utilizing other state-of-the-art metaheuristics are two lines of research that can be considered to advance the current study. Since the proposed method combines machine learning and metaheuristic, the model construction phase is completely data-dependent. Hence, the hybrid WOA-XGBoost can be trained and implemented autonomously without domain knowledge in machine learning. Therefore, the proposed method has a high potential to be used for modeling other sophisticated problems in the field of civil engineering.

Appendix

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
1	400.00	4.35	8.00	1.00	2.05	3.48	2.08	15.40	13.35	7.63	1395.00
2	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	559.80
3	300.00	3.40	5.30	0.00	3.40	3.52	3.42	12.10	8.70	6.76	508.90
4	400.00	4.25	8.00	0.90	2.15	3.56	2.26	15.30	13.15	7.61	1395.00
5	400.00	3.40	7.30	0.00	3.40	3.49	3.39	14.10	10.70	7.28	1068.80
6	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	661.60
7	400.00	4.35	8.00	1.06	2.05	3.55	2.09	15.46	13.41	7.66	1321.00
8	400.00	3.85	7.55	0.00	2.95	3.63	3.28	14.35	11.40	7.14	1440.00
9	400.00	4.65	7.35	0.00	2.15	3.55	3.40	14.15	12.00	6.79	1392.00
10	400.00	4.35	8.00	1.06	2.05	3.56	2.10	15.46	13.41	7.66	1321.00
11	400.00	3.85	7.30	0.00	2.95	3.70	3.60	14.10	11.15	7.08	1440.00
12	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	559.80
13	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	585.40
14	400.00	3.40	7.22	0.00	3.40	3.45	3.43	14.02	10.62	7.26	1240.00
15	300.00	3.40	5.27	0.00	3.40	3.51	3.44	12.07	8.67	6.75	661.60
16	400.00	4.10	2.08	0.00	2.70	3.63	2.75	8.88	6.18	4.86	432.00
17	400.00	3.45	8.00	0.30	2.95	3.65	2.95	14.70	11.75	7.59	1152.00
18	400.00	4.75	7.40	0.00	2.05	3.55	3.35	14.20	12.15	6.76	1440.00
19	400.00	4.10	1.71	0.00	2.70	3.26	2.75	8.51	5.81	4.56	423.90
20	400.00	4.65	7.50	0.00	2.15	3.59	3.29	14.30	12.15	6.82	1551.00
21	400.00	3.40	7.28	0.00	3.40	3.48	3.40	14.08	10.68	7.27	1318.00
22	300.00	3.40	5.22	0.00	3.40	3.45	3.43	12.02	8.62	6.74	559.80
23	300.00	3.40	5.20	0.00	3.40	3.40	3.40	12.00	8.60	6.73	559.00
24	400.00	3.45	8.00	0.19	2.95	3.56	2.97	14.59	11.64	7.52	1318.00
25	400.00	5.40	6.30	0.00	2.15	3.52	1.06	13.10	14.70	5.50	1056.00
26	400.00	4.45	8.00	1.04	1.95	3.44	2.00	15.44	13.49	7.61	1128.60
27	400.00	4.35	8.00	0.10	2.05	3.48	2.98	14.50	12.45	7.10	1392.00
28	400.00	4.25	8.00	0.40	2.15	3.59	2.79	14.80	12.65	7.32	1551.00
29	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	661.60
30	400.00	3.45	8.00	0.07	2.95	3.42	2.95	14.47	11.52	7.44	1240.00
31	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	610.70
32	300.00	3.40	5.20	0.00	3.40	3.40	3.40	12.00	8.60	6.73	559.80
33	400.00	4.25	8.00	0.91	2.15	3.56	2.25	15.31	13.16	7.62	1473.00
34	400.00	4.35	8.00	0.60	2.05	3.50	2.50	15.00	12.95	7.40	1297.80
35	400.00	3.85	7.20	0.00	2.95	3.57	3.57	14.00	11.05	7.06	1440.00

| Sample | Influencing factors | | | | | | | | | | Pile bearing capacity |
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
36	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	610.70
37	300.00	3.40	5.28	0.00	3.40	3.50	3.42	12.08	8.68	6.76	661.60
38	400.00	4.35	8.00	1.09	2.05	3.58	2.09	15.49	13.44	7.68	1224.80
39	400.00	4.35	8.00	0.95	2.05	3.44	2.09	15.35	13.30	7.60	1152.00
40	400.00	4.35	8.00	1.02	2.05	3.48	2.06	15.42	13.37	7.64	1248.00
41	400.00	3.40	7.30	0.00	3.40	3.54	3.44	14.10	10.70	7.28	967.00
42	300.00	3.40	5.30	0.00	3.40	3.52	3.42	12.10	8.70	6.76	610.70
43	400.00	3.40	7.30	0.00	3.40	3.51	3.41	14.10	10.70	7.28	1068.80
44	400.00	3.45	8.00	0.13	2.95	3.48	2.95	14.53	11.58	7.48	1240.00
45	300.00	3.40	5.30	0.00	3.40	3.51	3.41	12.10	8.70	6.76	661.60
46	400.00	3.45	8.00	0.25	2.95	3.64	2.99	14.65	11.70	7.55	1119.70
47	300.00	3.40	5.32	0.00	3.40	3.55	3.43	12.12	8.72	6.77	661.60
48	400.00	4.20	8.00	0.80	2.20	3.49	2.29	15.20	13.00	7.58	1395.00
49	400.00	4.35	8.00	0.96	2.05	3.43	2.07	15.36	13.31	7.61	1119.70
50	400.00	4.35	8.00	0.75	2.05	3.45	2.30	15.15	13.10	7.49	1323.20
51	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	559.80
52	400.00	4.35	8.00	1.05	2.05	3.50	2.05	15.45	13.40	7.66	1344.00
53	400.00	4.35	8.00	0.90	2.05	3.42	2.12	15.30	13.25	7.58	1395.00
54	400.00	4.75	7.25	0.00	2.05	3.62	3.57	14.05	12.00	6.73	1425.00
55	400.00	4.35	8.00	0.99	2.05	3.49	2.10	15.39	13.34	7.63	1224.80
56	400.00	3.40	7.30	0.00	3.40	3.50	3.40	14.10	10.70	7.28	1056.00
57	400.00	3.40	7.26	0.00	3.40	3.46	3.40	14.06	10.66	7.27	1152.00
58	300.00	3.40	5.26	0.00	3.40	3.47	3.41	12.06	8.66	6.75	559.80
59	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	661.60
60	400.00	4.35	8.00	0.95	2.05	3.41	2.06	15.35	13.30	7.60	1323.20
61	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	559.80
62	400.00	4.35	8.00	1.08	2.05	3.56	2.08	15.48	13.43	7.67	1344.00
63	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	661.60
64	400.00	4.35	8.00	1.08	2.05	3.53	2.05	15.48	13.43	7.67	1248.00
65	400.00	4.35	8.00	1.20	2.05	3.62	2.02	15.60	13.55	7.74	1119.70
66	400.00	4.45	8.00	1.10	1.95	3.50	2.00	15.50	13.55	7.65	1128.60
67	300.00	3.40	5.23	0.00	3.40	3.47	3.44	12.03	8.63	6.74	559.80
68	400.00	4.35	8.00	1.00	2.05	3.55	2.15	15.40	13.35	7.63	1344.00
69	400.00	4.25	8.00	1.02	2.15	3.58	2.16	15.42	13.27	7.68	1248.00
70	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	559.80
71	400.00	3.45	8.00	0.14	2.95	3.52	2.98	14.54	11.59	7.48	885.00
72	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	610.70
73	400.00	3.40	7.32	0.00	3.40	3.57	3.45	14.12	10.72	7.28	1032.40
74	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	585.35
75	400.00	4.35	8.00	1.22	2.05	3.67	2.05	15.62	13.57	7.75	1248.00
76	400.00	4.10	1.90	0.00	2.70	3.43	2.73	8.70	6.00	4.72	620.00
77	400.00	4.25	8.00	1.00	2.15	3.55	2.15	15.40	13.25	7.67	1248.00
78	300.00	3.40	5.23	0.00	3.40	3.45	3.42	12.03	8.63	6.74	559.80
79	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	610.70
80	300.00	3.40	5.20	0.00	3.40	3.40	3.40	12.00	8.60	6.73	610.70
81	400.00	4.10	2.00	0.00	2.70	3.54	2.74	8.80	6.10	4.80	712.50
82	400.00	3.85	7.20	0.00	2.95	3.56	3.56	14.00	11.05	7.06	1440.00
83	400.00	3.45	8.00	0.11	2.95	3.46	2.95	14.51	11.56	7.46	1240.00

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
84	400.00	4.35	8.00	0.95	2.05	3.48	2.13	15.35	13.30	7.60	1392.00
85	400.00	4.05	8.00	0.80	2.35	3.56	2.36	15.20	12.85	7.64	1318.00
86	400.00	4.75	7.20	0.00	2.05	3.68	3.68	14.00	11.95	6.72	1344.00
87	400.00	4.35	8.00	1.01	2.05	3.46	2.05	15.41	13.36	7.64	1473.00
88	300.00	3.40	5.30	0.00	3.40	3.52	3.42	12.10	8.70	6.76	508.90
89	300.00	3.40	5.20	0.00	3.40	3.41	3.41	12.00	8.60	6.73	610.70
90	400.00	4.35	8.00	1.08	2.05	3.58	2.10	15.48	13.43	7.67	1224.80
91	400.00	4.35	8.00	0.90	2.05	3.40	2.10	15.30	13.25	7.58	1395.00
92	400.00	4.35	8.00	0.97	2.05	3.47	2.10	15.37	13.32	7.61	1395.00
93	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	610.70
94	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	559.80
95	300.00	3.40	5.20	0.00	3.40	3.48	3.48	12.00	8.60	6.73	611.60
96	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	610.70
97	400.00	4.35	8.00	0.60	2.05	3.40	2.40	15.00	12.95	7.40	1473.00
98	400.00	4.35	8.00	1.02	2.05	3.47	2.05	15.42	13.37	7.64	1473.00
99	400.00	4.35	8.00	0.30	2.05	3.45	2.75	14.70	12.65	7.22	1473.00
100	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	610.70
101	400.00	4.35	8.00	0.90	2.05	3.46	2.16	15.30	13.25	7.58	1395.00
102	400.00	4.35	8.00	1.05	2.05	3.53	2.08	15.45	13.40	7.66	1473.00
103	400.00	3.45	8.00	0.15	2.95	3.50	2.95	14.55	11.60	7.49	1240.00
104	400.00	3.50	8.00	0.17	2.90	3.47	2.90	14.57	11.67	7.48	960.00
105	400.00	4.35	8.00	1.00	2.05	3.46	2.06	15.40	13.35	7.63	1344.00
106	400.00	4.75	7.45	0.00	2.05	3.67	3.42	14.25	12.20	6.77	1440.00
107	400.00	4.35	8.00	1.06	2.05	3.56	2.10	15.46	13.41	7.66	1224.80
108	400.00	3.45	8.00	0.20	2.95	3.52	2.92	14.60	11.65	7.52	967.00
109	400.00	4.35	8.00	0.94	2.05	3.49	2.15	15.34	13.29	7.60	1395.00
110	300.00	3.40	5.20	0.00	3.40	3.40	3.40	12.00	8.60	6.73	559.80
111	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	661.60
112	400.00	4.10	2.00	0.00	2.70	3.56	2.76	8.80	6.10	4.80	620.00
113	400.00	4.05	8.00	0.66	2.35	3.46	2.40	15.06	12.71	7.56	1318.00
114	400.00	3.40	7.30	0.00	3.40	3.50	3.40	14.10	10.70	7.28	960.00
115	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	660.60
116	400.00	4.10	1.95	0.00	2.70	3.49	2.74	8.75	6.05	4.76	712.50
117	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	508.90
118	400.00	4.10	2.20	0.00	2.70	3.72	2.72	9.00	6.30	4.94	610.70
119	300.00	3.40	5.20	0.00	3.40	3.45	3.45	12.00	8.60	6.73	610.70
120	400.00	3.45	8.00	0.09	2.95	3.44	2.95	14.49	11.54	7.45	1318.00
121	400.00	4.35	8.00	1.00	2.05	3.48	2.08	15.40	13.35	7.63	1395.00
122	300.00	3.40	5.27	0.00	3.40	3.50	3.43	12.07	8.67	6.75	610.70
123	400.00	4.10	2.00	0.00	2.70	3.52	2.72	8.80	6.10	4.80	610.70
124	300.00	3.40	5.25	0.00	3.40	3.51	3.46	12.05	8.65	6.75	559.80
125	400.00	3.45	8.00	0.20	2.95	3.58	2.98	14.60	11.65	7.52	1068.80
126	400.00	3.40	7.35	0.00	3.40	3.56	3.41	14.15	10.75	7.29	1052.40
127	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	559.80
128	400.00	4.05	8.00	0.70	2.35	3.47	2.37	15.10	12.75	7.58	1318.00
129	400.00	4.35	8.00	0.05	2.05	3.58	3.13	14.45	12.40	7.07	1344.00
130	300.00	3.40	5.26	0.00	3.40	3.51	3.45	12.06	8.66	6.75	610.70
131	400.00	4.05	8.00	0.70	2.35	3.48	2.38	15.10	12.75	7.58	1240.00

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
132	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	559.80
133	400.00	3.40	7.24	0.00	3.40	3.44	3.40	14.04	10.64	7.26	1395.00
134	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	508.90
135	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	610.70
136	400.00	4.35	8.00	1.05	2.05	3.54	2.09	15.45	13.40	7.66	1224.80
137	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	610.70
138	300.00	3.40	5.30	0.00	3.40	3.51	3.41	12.10	8.70	6.76	559.80
139	400.00	4.45	8.00	1.04	1.95	3.43	1.99	15.44	13.49	7.61	1128.60
140	300.00	3.40	5.20	0.00	3.40	3.38	3.38	12.00	8.60	6.73	610.70
141	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	559.80
142	400.00	4.35	8.00	1.01	2.05	3.46	2.05	15.41	13.36	7.64	1318.00
143	400.00	4.75	7.31	0.00	2.05	3.61	3.50	14.11	12.06	6.74	1440.00
144	300.00	3.40	5.22	0.00	3.40	3.45	3.43	12.02	8.62	6.74	559.80
145	400.00	4.35	8.00	1.00	2.05	3.45	2.05	15.40	13.35	7.63	1248.00
146	400.00	4.45	8.00	1.16	1.95	3.55	1.99	15.56	13.61	7.68	1224.80
147	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	610.70
148	400.00	4.35	8.00	0.98	2.05	3.48	2.10	15.38	13.33	7.62	1224.80
149	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	610.70
150	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	559.80
151	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	555.30
152	300.00	3.40	5.20	0.00	3.40	3.40	3.40	12.00	8.60	6.73	559.80
153	400.00	4.35	8.00	1.05	2.05	3.55	2.10	15.45	13.40	7.66	1224.80
154	300.00	3.40	5.24	0.00	3.40	3.45	3.41	12.04	8.64	6.75	661.60
155	400.00	4.35	8.00	0.30	2.05	3.50	2.80	14.70	12.65	7.22	1297.80
156	400.00	4.25	8.00	1.00	2.15	3.56	2.16	15.40	13.25	7.67	1344.00
157	300.00	3.40	5.20	0.00	3.40	3.41	3.41	12.00	8.60	6.73	610.70
158	400.00	4.10	1.66	0.00	2.70	3.21	2.75	8.46	5.76	4.52	423.90
159	400.00	4.75	7.25	0.00	2.05	3.65	3.60	14.05	12.00	6.73	1425.00
160	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	559.80
161	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	559.80
162	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	610.70
163	400.00	4.75	7.60	0.00	2.05	3.49	3.09	14.40	12.35	6.81	1473.00
164	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	610.70
165	400.00	3.45	8.00	0.13	2.95	3.51	2.98	14.53	11.58	7.48	1344.00
166	400.00	3.45	8.00	0.15	2.95	3.51	2.96	14.55	11.60	7.49	1017.90
167	400.00	4.35	8.00	1.00	2.05	3.50	2.10	15.40	13.35	7.63	1224.80
168	400.00	3.45	8.00	0.13	2.95	3.48	2.95	14.53	11.58	7.48	1318.00
169	400.00	4.35	8.00	1.02	2.05	3.47	4.05	15.42	13.37	7.64	1318.00
170	400.00	4.35	8.00	1.00	2.05	3.55	2.15	15.40	13.35	7.63	1344.00
171	400.00	3.50	8.00	0.20	2.90	3.50	2.90	14.60	11.70	7.50	1056.00
172	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	610.70
173	400.00	4.35	8.00	0.90	2.05	3.38	2.08	15.30	13.25	7.58	1395.00
174	400.00	4.45	8.00	1.10	1.95	3.46	1.96	15.50	13.55	7.65	1248.00
175	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	610.70
176	400.00	3.40	7.31	0.00	3.40	3.56	3.45	14.11	10.71	7.28	1224.80
177	300.00	3.40	5.26	0.00	3.40	3.47	3.41	12.06	8.66	6.75	559.80
178	400.00	4.35	8.00	0.98	2.05	3.53	2.15	15.38	13.33	7.62	1395.00
179	400.00	4.45	8.00	0.99	1.95	3.38	1.99	15.39	13.44	7.59	1128.60

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
180	400.00	3.45	8.00	0.30	2.95	3.65	2.95	14.70	11.75	7.59	1017.90
181	400.00	4.35	8.00	1.11	2.05	3.60	2.09	15.51	13.46	7.69	1128.60
182	400.00	5.72	8.00	1.69	0.68	4.12	1.03	16.09	15.41	7.50	1344.00
183	400.00	4.35	8.00	1.05	2.05	3.55	2.10	15.45	13.40	7.66	1473.00
184	400.00	4.35	8.00	0.95	2.05	3.45	2.10	15.35	13.30	7.60	1395.00
185	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	407.20
186	400.00	3.45	8.00	0.17	2.95	3.54	2.97	14.57	11.62	7.50	1056.00
187	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	661.60
188	400.00	3.45	8.00	0.27	2.95	3.63	2.96	14.67	11.72	7.57	1152.00
189	400.00	4.35	8.00	1.05	2.05	3.51	2.06	15.45	13.40	7.66	1248.00
190	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	610.70
191	400.00	4.35	8.00	1.06	2.05	3.56	2.10	15.46	13.41	7.66	1224.80
192	400.00	4.35	8.00	1.11	2.05	3.60	2.09	15.51	13.46	7.69	1128.60
193	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	610.70
194	300.00	3.40	5.20	0.00	3.40	3.38	3.38	12.00	8.60	6.73	610.70
195	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	610.70
196	400.00	4.75	7.60	0.00	2.05	3.65	3.25	14.40	12.35	6.81	1440.00
197	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	559.80
198	300.00	3.40	5.23	0.00	3.40	3.44	3.41	12.03	8.63	6.74	585.35
199	400.00	4.35	8.00	1.13	2.05	3.63	2.10	15.53	13.48	7.70	1224.80
200	300.00	3.40	5.20	0.00	3.40	3.40	3.40	12.00	8.60	6.73	559.80
201	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	610.70
202	400.00	4.75	7.60	0.00	2.05	3.46	3.06	14.40	12.35	6.81	1473.00
203	400.00	4.35	8.00	1.00	2.05	3.47	2.07	15.40	13.35	7.63	1344.00
204	400.00	3.45	8.00	0.16	2.95	3.53	2.97	14.56	11.61	7.50	1056.00
205	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	661.60
206	400.00	4.35	8.00	1.05	2.05	3.50	4.35	15.45	13.40	7.66	1297.80
207	400.00	3.45	8.00	0.22	2.95	3.57	2.95	14.62	11.67	7.53	1318.00
208	400.00	4.35	8.00	1.09	2.05	3.57	2.08	15.49	13.44	7.68	1395.00
209	400.00	4.75	7.50	0.00	2.05	3.45	3.15	14.30	12.25	6.79	1297.80
210	400.00	4.35	8.00	0.90	2.05	3.38	2.08	15.30	13.25	7.58	1395.00
211	300.00	3.40	5.23	0.00	3.40	3.44	3.41	12.03	8.63	6.74	610.70
212	300.00	3.40	5.35	0.00	3.40	3.57	3.42	12.15	8.75	6.78	661.60
213	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	559.80
214	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	610.70
215	300.00	3.40	5.35	0.00	3.40	3.57	3.42	12.15	8.75	6.78	661.60
216	400.00	4.10	2.00	0.00	2.70	3.55	2.75	8.80	6.10	4.80	610.70
217	300.00	3.40	5.24	0.00	3.40	3.49	3.45	12.04	8.64	6.75	610.70
218	300.00	3.40	5.24	0.00	3.40	3.46	3.42	12.04	8.64	6.75	610.70
219	400.00	4.35	8.00	1.07	2.05	3.56	2.09	15.47	13.42	7.67	1224.80
220	300.00	3.40	5.23	0.00	3.40	3.44	3.41	12.03	8.63	6.74	610.70
221	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	610.70
222	400.00	4.35	8.00	1.10	2.05	3.60	2.10	15.50	13.45	7.69	1224.80
223	400.00	4.05	8.00	0.70	2.35	3.47	2.37	15.10	12.75	7.58	1240.00
224	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	661.60
225	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	559.80
226	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	559.80
227	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	559.80

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
228	300.00	3.40	5.18	0.00	3.40	3.36	3.38	11.98	8.58	6.73	559.80
229	400.00	4.35	8.00	1.05	2.05	3.50	2.05	15.45	13.40	7.66	1128.60
230	300.00	3.40	5.30	0.00	3.40	3.52	3.42	12.10	8.70	6.76	508.90
231	400.00	3.40	7.27	0.00	3.40	3.48	3.41	14.07	10.67	7.27	1152.00
232	400.00	4.35	8.00	1.03	2.05	3.48	2.05	15.43	13.38	7.65	1318.00
233	400.00	4.35	8.00	1.01	2.05	3.46	2.05	15.41	13.36	7.64	1473.00
234	400.00	3.40	7.30	0.00	3.40	3.61	3.51	14.10	10.70	7.28	1115.20
235	400.00	4.35	8.00	1.18	2.05	3.66	2.08	15.58	13.53	7.73	1056.00
236	400.00	4.35	8.00	1.03	2.05	3.51	2.08	15.43	13.38	7.65	1224.80
237	400.00	3.45	8.00	0.22	2.95	3.57	2.95	14.62	11.67	7.53	1344.00
238	400.00	4.35	8.00	1.06	2.05	3.52	2.06	15.46	13.41	7.66	1344.00
239	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	600.70
240	400.00	4.10	1.92	0.00	2.70	3.44	2.72	8.72	6.02	4.73	712.50
241	400.00	3.40	7.35	0.00	3.40	3.55	3.40	14.15	10.75	7.29	1017.90
242	400.00	4.35	8.00	0.70	2.05	3.49	2.39	15.10	13.05	7.46	1392.00
243	400.00	3.85	7.24	0.00	2.95	3.62	3.58	14.04	11.09	7.07	1344.00
244	400.00	3.40	7.35	0.00	3.40	3.57	3.42	14.15	10.75	7.29	1068.80
245	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	610.70
246	400.00	4.10	2.10	0.00	2.70	3.63	2.73	8.90	6.20	4.87	480.00
247	400.00	4.25	8.00	0.20	2.15	3.56	2.96	14.60	12.45	7.20	1392.00
248	400.00	3.40	7.27	0.00	3.40	3.49	3.42	14.07	10.67	7.27	1248.00
249	300.00	3.40	5.22	0.00	3.40	3.46	3.44	12.02	8.62	6.74	661.60
250	300.00	3.40	5.24	0.00	3.40	3.48	3.44	12.04	8.64	6.75	610.70
251	400.00	3.40	7.40	0.00	3.40	3.61	3.41	14.20	10.80	7.30	1088.80
252	400.00	4.35	8.00	0.95	2.05	3.45	2.10	15.35	13.30	7.60	1221.50
253	400.00	3.45	8.00	0.25	2.95	3.62	2.97	14.65	11.70	7.55	1170.60
254	400.00	4.45	8.00	1.18	1.95	3.58	2.00	15.58	13.63	7.69	1032.40
255	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	617.00
256	400.00	3.40	7.28	0.00	3.40	3.53	3.45	14.08	10.68	7.27	1032.40
257	300.00	3.40	5.27	0.00	3.40	3.49	3.42	12.07	8.67	6.75	559.80
258	400.00	4.25	8.00	0.96	2.15	3.53	2.17	15.36	13.21	7.65	1344.00
259	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	610.70
260	400.00	4.35	8.00	1.10	2.05	3.55	2.05	15.50	13.45	7.69	1425.00
261	400.00	3.45	8.00	0.25	2.95	3.60	2.95	14.65	11.70	7.55	960.00
262	400.00	4.35	8.00	1.05	2.05	3.52	2.07	15.45	13.40	7.66	1248.00
263	400.00	4.35	8.00	1.05	2.05	3.55	2.10	15.45	13.40	7.66	1224.80
264	400.00	3.45	8.00	0.27	2.95	3.65	2.98	14.67	11.72	7.57	1248.00
265	400.00	4.35	8.00	1.04	2.05	3.52	2.08	15.44	13.39	7.65	1321.00
266	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	617.00
267	400.00	4.35	8.00	1.10	2.05	3.54	2.04	15.50	13.45	7.69	1323.20
268	300.00	3.40	5.27	0.00	3.40	3.51	3.44	12.07	8.67	6.75	661.60
269	400.00	4.35	8.00	0.10	2.05	3.40	2.90	14.50	12.45	7.10	1473.00
270	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	661.60
271	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	559.80
272	400.00	3.85	7.50	0.00	2.95	3.66	3.36	14.30	11.35	7.13	1425.00
273	400.00	3.85	7.30	0.00	2.95	3.70	3.60	14.10	11.15	7.08	1440.00
274	400.00	3.50	8.00	0.20	2.90	3.50	2.90	14.60	11.70	7.50	1056.00
275	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	610.70

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
276	400.00	4.10	2.00	0.00	2.70	3.52	2.72	8.80	6.10	4.80	610.70
277	400.00	3.40	7.22	0.00	3.40	3.47	3.45	14.02	10.62	7.26	1128.60
278	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	610.70
279	400.00	3.45	8.00	0.07	2.95	3.42	2.95	14.47	11.52	7.44	1240.00
280	400.00	4.10	2.05	0.00	2.70	3.58	2.73	8.85	6.15	4.83	661.60
281	400.00	4.45	7.21	0.00	2.35	3.41	2.40	14.01	11.66	6.83	1318.00
282	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	555.40
283	400.00	4.35	8.00	0.96	2.05	3.42	2.06	15.36	13.31	7.61	1244.00
284	300.00	3.40	5.25	0.00	3.40	3.45	3.40	12.05	8.65	6.75	610.70
285	300.00	3.40	5.27	0.00	3.40	3.50	3.43	12.07	8.67	6.75	610.70
286	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	559.80
287	400.00	4.35	8.00	0.90	2.05	3.40	2.10	15.30	13.25	7.58	1473.00
288	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	610.70
289	400.00	4.45	8.00	0.94	1.95	3.34	2.00	15.34	13.39	7.56	1128.60
290	400.00	3.50	8.00	0.20	2.90	3.51	2.91	14.60	11.70	7.50	960.00
291	300.00	3.40	5.20	0.00	3.40	3.45	3.45	12.00	8.60	6.73	610.70
292	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	508.90
293	400.00	4.35	8.00	0.18	2.05	3.38	2.80	14.58	12.53	7.15	1473.00
294	300.00	3.40	5.25	0.00	3.40	3.45	3.40	12.05	8.65	6.75	610.70
295	400.00	3.40	7.28	0.00	3.40	3.53	3.45	14.08	10.68	7.27	1032.40
296	300.00	3.40	5.21	0.00	3.40	3.44	3.43	12.01	8.61	6.74	508.50
297	400.00	4.35	8.00	0.96	2.05	3.42	2.06	15.36	13.31	7.61	1119.70
298	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	559.80
299	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	559.80
300	400.00	4.35	8.00	0.98	2.05	3.44	2.06	15.38	13.33	7.62	1344.00
301	300.00	3.40	5.30	0.00	3.40	3.52	3.42	12.10	8.70	6.76	559.80
302	400.00	4.35	8.00	1.03	2.05	3.48	2.05	15.43	13.38	7.65	1344.00
303	400.00	4.10	2.01	0.00	2.70	3.53	2.72	8.81	6.11	4.80	528.00
304	400.00	4.65	7.40	0.00	2.15	3.59	3.39	14.20	12.05	6.80	1551.00
305	400.00	4.35	8.00	1.05	2.05	3.55	2.10	15.45	13.40	7.66	1221.50
306	300.00	3.40	5.20	0.00	3.40	3.45	3.45	12.00	8.60	6.73	610.70
307	400.00	3.40	7.30	0.00	3.40	3.50	3.40	14.10	10.70	7.28	1152.00
308	400.00	4.35	8.00	0.60	2.05	3.46	2.46	15.00	12.95	7.40	1395.00
309	400.00	3.50	8.00	0.16	2.90	3.47	2.91	14.56	11.66	7.47	1056.00
310	400.00	4.35	8.00	1.03	2.05	3.53	2.10	15.43	13.38	7.65	1395.00
311	400.00	3.40	7.30	0.00	3.40	3.48	3.38	14.10	10.70	7.28	967.00
312	400.00	3.45	5.24	0.00	3.35	3.44	3.40	12.04	8.69	6.72	1240.00
313	400.00	4.35	8.00	1.03	2.05	3.48	2.05	15.43	13.38	7.65	1395.00
314	300.00	3.40	5.27	0.00	3.40	3.49	3.42	12.07	8.67	6.75	610.70
315	400.00	3.45	8.00	0.20	2.95	3.56	2.96	14.60	11.65	7.52	1152.00
316	400.00	4.35	8.00	1.07	2.05	3.52	2.05	15.47	13.42	7.67	1082.30
317	400.00	3.40	7.30	0.00	3.40	3.50	3.40	14.10	10.70	7.28	900.00
318	300.00	3.40	5.20	0.00	3.40	3.44	3.44	12.00	8.60	6.73	559.80
319	400.00	4.20	8.00	0.86	2.20	3.48	2.22	15.26	13.06	7.61	1344.00
320	400.00	4.10	1.84	0.00	2.70	3.39	2.75	8.64	5.94	4.67	423.90
321	400.00	4.35	8.00	0.98	2.05	3.45	2.07	15.38	13.33	7.62	1395.00
322	400.00	4.35	8.00	0.95	2.05	3.41	2.06	15.35	13.30	7.60	1110.60
323	300.00	3.40	5.29	0.00	3.40	3.54	3.45	12.09	8.69	6.76	661.60

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
324	400.00	4.25	8.00	0.10	2.15	3.54	3.04	14.50	12.35	7.14	1551.00
325	400.00	4.65	7.50	0.00	2.15	3.55	3.25	14.30	12.15	6.82	1551.00
326	300.00	3.40	5.23	0.00	3.40	3.45	3.42	12.03	8.63	6.74	559.80
327	400.00	3.40	7.26	0.00	3.40	3.48	3.42	14.06	10.66	7.27	1032.40
328	400.00	4.35	8.00	0.96	2.05	3.46	2.10	15.36	13.31	7.61	1395.00
329	400.00	3.40	7.33	0.00	3.40	3.55	3.42	14.13	10.73	7.28	1094.25
330	300.00	3.40	5.24	0.00	3.40	3.45	3.41	12.04	8.64	6.75	610.70
331	400.00	4.35	8.00	1.04	2.05	3.49	2.05	15.44	13.39	7.65	1344.00
332	400.00	4.35	8.00	1.04	2.05	3.49	2.05	15.44	13.39	7.65	1395.00
333	400.00	4.35	8.00	1.05	2.05	3.50	2.05	15.45	13.40	7.66	1152.00
334	400.00	3.45	8.00	0.20	2.95	3.50	2.90	14.60	11.65	7.52	1017.90
335	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	559.80
336	400.00	3.40	7.35	0.00	3.40	3.57	3.42	14.15	10.75	7.29	1119.70
337	400.00	4.75	7.40	0.00	2.05	3.52	3.32	14.20	12.15	6.76	1425.00
338	400.00	4.35	8.00	1.02	2.05	3.47	2.05	15.42	13.37	7.64	1395.00
339	400.00	4.10	1.96	0.00	2.70	3.49	2.73	8.76	6.06	4.76	508.90
340	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	611.70
341	400.00	3.45	8.00	0.14	2.95	3.52	2.98	14.54	11.59	7.48	885.00
342	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	559.80
343	400.00	4.35	8.00	1.05	2.05	3.50	2.05	15.45	13.40	7.66	1152.00
344	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	661.60
345	300.00	3.40	5.24	0.00	3.40	3.46	3.42	12.04	8.64	6.75	559.80
346	400.00	3.85	7.50	0.00	2.95	3.68	3.38	14.30	11.35	7.13	1425.00
347	400.00	4.45	8.00	1.01	1.95	3.40	1.99	15.41	13.46	7.60	1128.60
348	400.00	3.85	7.36	0.00	2.95	3.66	3.50	14.16	11.21	7.10	1344.00
349	400.00	3.45	8.00	0.15	2.95	3.64	3.09	14.55	11.60	7.49	1344.00
350	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	610.70
351	300.00	3.40	5.30	0.00	3.40	3.52	3.42	12.10	8.70	6.76	661.60
352	400.00	4.25	8.00	0.20	2.15	3.55	2.95	14.60	12.45	7.20	1392.00
353	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	559.80
354	400.00	3.40	7.31	0.00	3.40	3.56	3.45	14.11	10.71	7.28	1032.40
355	400.00	4.10	1.72	0.00	2.70	3.27	2.75	8.52	5.82	4.57	423.90
356	300.00	3.40	5.22	0.00	3.40	3.46	3.44	12.02	8.62	6.74	661.60
357	400.00	3.45	8.00	0.10	2.95	3.60	3.10	14.50	11.55	7.46	1344.00
358	400.00	4.35	8.00	0.87	2.05	3.37	2.10	15.27	13.22	7.56	1473.00
359	400.00	4.35	8.00	1.11	2.05	3.56	2.05	15.51	13.46	7.69	1128.60
360	400.00	4.35	8.00	0.80	2.05	3.45	2.25	15.20	13.15	7.52	1392.00
361	400.00	4.35	8.00	1.08	2.05	3.56	2.08	15.48	13.43	7.67	1224.80
362	400.00	3.45	8.00	0.20	2.95	3.53	2.93	14.60	11.65	7.52	1068.80
363	400.00	4.35	8.00	1.02	2.05	3.48	2.06	15.42	13.37	7.64	1248.00
364	400.00	4.45	8.00	1.04	1.95	3.43	1.99	15.44	13.49	7.61	1224.80
365	400.00	3.45	8.00	0.10	2.95	3.54	3.04	14.50	11.55	7.46	1017.90
366	400.00	3.85	7.30	0.00	2.95	3.68	3.58	14.10	11.15	7.08	1440.00
367	400.00	4.35	8.00	1.08	2.05	3.53	2.05	15.48	13.43	7.67	1248.00
368	400.00	4.35	8.00	1.01	2.05	3.46	2.05	15.41	13.36	7.64	1550.00
369	400.00	4.45	8.00	1.03	1.95	3.43	2.00	15.43	13.48	7.61	1128.60
370	400.00	3.40	7.23	0.00	3.40	3.43	3.40	14.03	10.63	7.26	960.00
371	400.00	4.75	7.50	0.00	2.05	3.60	3.30	14.30	12.25	6.79	1425.00

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
372	400.00	3.40	7.24	0.00	3.40	3.44	3.40	14.04	10.64	7.26	1240.00
373	400.00	3.40	7.40	0.00	3.40	3.62	3.42	14.20	10.80	7.30	1017.90
374	400.00	4.25	8.00	1.02	2.15	3.58	2.16	15.42	13.27	7.68	1248.00
375	300.00	3.40	5.22	0.00	3.40	3.45	3.43	12.02	8.62	6.74	559.80
376	400.00	3.45	8.00	0.18	2.95	3.55	2.97	14.58	11.63	7.51	1344.00
377	400.00	4.75	7.60	0.00	2.05	3.44	3.04	14.40	12.35	6.81	1473.00
378	400.00	4.35	8.00	1.03	2.05	3.53	2.10	15.43	13.38	7.65	1224.80
379	400.00	3.45	8.00	0.22	2.95	3.59	2.97	14.62	11.67	7.53	1152.00
380	400.00	3.45	8.00	0.30	2.95	3.66	2.96	14.70	11.75	7.59	960.00
381	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	559.80
382	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	559.80
383	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	532.40
384	300.00	3.40	5.23	0.00	3.40	3.46	3.43	12.03	8.63	6.74	559.80
385	400.00	4.35	8.00	1.10	2.05	3.55	2.05	15.50	13.45	7.69	1119.70
386	400.00	3.40	7.31	0.00	3.40	3.54	3.43	14.11	10.71	7.28	1032.80
387	300.00	3.40	5.18	0.00	3.40	3.38	3.40	11.98	8.58	6.73	559.80
388	400.00	3.85	7.35	0.00	2.95	3.64	3.49	14.15	11.20	7.09	1425.00
389	300.00	3.40	5.30	0.00	3.40	3.54	3.44	12.10	8.70	6.76	661.60
390	300.00	3.40	5.20	0.00	3.40	3.45	3.45	12.00	8.60	6.73	559.80
391	400.00	4.10	1.75	0.00	2.70	3.26	2.71	8.55	5.85	4.59	508.90
392	400.00	4.35	8.00	0.95	2.05	3.42	2.07	15.35	13.30	7.60	1119.70
393	400.00	4.10	2.08	0.00	2.70	3.63	2.75	8.88	6.18	4.86	432.00
394	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	532.40
395	300.00	3.40	5.30	0.00	3.40	3.53	3.43	12.10	8.70	6.76	559.80
396	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	559.80
397	400.00	4.35	8.00	1.01	2.05	3.51	2.10	15.41	13.36	7.64	1473.00
398	400.00	3.40	7.31	0.00	3.40	3.51	3.40	14.11	10.71	7.28	1395.00
399	400.00	3.40	7.35	0.00	3.40	3.57	3.42	14.15	10.75	7.29	1119.70
400	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	661.60
401	400.00	4.35	8.00	0.99	2.05	3.44	2.05	15.39	13.34	7.63	1395.00
402	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	610.70
403	400.00	4.10	2.08	0.00	2.70	3.58	2.70	8.88	6.18	4.86	480.00
404	400.00	4.35	8.00	0.90	2.05	3.43	2.13	15.30	13.25	7.58	1395.00
405	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	661.60
406	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	661.60
407	300.00	3.40	5.20	0.00	3.40	3.42	3.42	12.00	8.60	6.73	661.60
408	400.00	3.50	8.00	0.05	2.90	3.36	2.91	14.45	11.55	7.40	1056.00
409	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	559.80
410	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	559.80
411	300.00	3.40	5.15	0.00	3.40	3.36	3.41	11.95	8.55	6.72	610.70
412	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	610.70
413	400.00	4.35	8.00	1.04	2.05	3.54	2.10	15.44	13.39	7.65	1224.80
414	400.00	4.35	8.00	1.07	2.05	3.55	2.08	15.47	13.42	7.67	1224.80
415	300.00	3.40	5.23	0.00	3.40	3.46	3.43	12.03	8.63	6.74	661.60
416	400.00	4.35	8.00	1.08	2.05	3.54	2.06	15.48	13.43	7.67	1395.00
417	400.00	4.35	8.00	0.90	2.05	3.37	2.07	15.30	13.25	7.58	1395.00
418	400.00	4.25	8.00	1.00	2.15	3.56	2.16	15.40	13.25	7.67	1248.00
419	400.00	4.35	8.00	0.80	2.05	3.46	2.26	15.20	13.15	7.52	1395.00

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
420	400.00	4.35	8.00	1.08	2.05	3.53	2.05	15.48	13.43	7.67	1152.00
421	400.00	4.35	8.00	0.80	2.05	3.45	2.25	15.20	13.15	7.52	1392.00
422	400.00	4.35	8.00	1.00	2.05	3.45	2.05	15.40	13.35	7.63	1119.70
423	300.00	3.40	5.22	0.00	3.40	3.42	3.40	12.02	8.62	6.74	508.90
424	400.00	3.45	6.29	0.00	3.35	3.44	3.35	13.09	9.74	7.02	1240.00
425	400.00	4.35	8.00	1.10	2.05	3.60	2.10	15.50	13.45	7.69	1395.00
426	400.00	4.35	8.00	1.05	2.05	3.51	2.06	15.45	13.40	7.66	1344.00
427	300.00	3.40	5.20	0.00	3.40	3.44	3.44	12.00	8.60	6.73	611.60
428	400.00	3.40	7.40	0.00	3.40	3.61	3.41	14.20	10.80	7.30	1152.00
429	300.00	3.40	5.20	0.00	3.40	3.44	3.44	12.00	8.60	6.73	661.60
430	400.00	4.35	8.00	1.00	2.05	3.50	2.10	15.40	13.35	7.63	1395.00
431	400.00	3.40	7.28	0.00	3.40	3.48	3.40	14.08	10.68	7.27	1318.00
432	400.00	3.55	5.39	0.00	3.25	3.44	3.25	12.19	8.94	6.72	1083.00
433	300.00	3.40	5.30	0.00	3.40	3.50	3.40	12.10	8.70	6.76	610.70
434	400.00	4.10	1.80	0.00	2.70	3.39	2.79	8.60	5.90	4.64	620.00
435	300.00	3.40	5.23	0.00	3.40	3.45	3.42	12.03	8.63	6.74	559.80
436	300.00	3.40	5.30	0.00	3.40	3.52	3.42	12.10	8.70	6.76	661.60
437	400.00	4.35	8.00	1.00	2.05	3.52	2.12	15.40	13.35	7.63	1395.00
438	400.00	4.35	8.00	0.95	2.05	3.45	2.10	15.35	13.30	7.60	1224.80
439	400.00	3.40	7.40	0.00	3.40	3.58	3.38	14.20	10.80	7.30	967.00
440	400.00	4.35	8.00	0.97	2.05	3.42	2.05	15.37	13.32	7.61	1317.00
441	400.00	3.45	8.00	0.12	2.95	3.47	2.95	14.52	11.57	7.47	1318.00
442	300.00	3.40	5.24	0.00	3.40	3.48	3.44	12.04	8.64	6.75	559.80
443	300.00	3.40	5.20	0.00	3.40	3.43	3.43	12.00	8.60	6.73	661.60
444	400.00	4.45	8.00	1.10	1.95	3.49	1.99	15.50	13.55	7.65	1128.60
445	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	559.80
446	300.00	3.40	5.18	0.00	3.40	3.38	3.40	11.98	8.58	6.73	559.80
447	400.00	3.45	8.00	0.20	2.95	3.61	3.01	14.60	11.65	7.52	1344.00
448	300.00	3.40	5.25	0.00	3.40	3.47	3.42	12.05	8.65	6.75	559.80
449	400.00	4.05	8.00	0.60	2.35	3.41	2.41	15.00	12.65	7.52	1240.00
450	400.00	4.45	8.00	1.00	1.95	3.40	2.00	15.40	13.45	7.59	1128.60
451	300.00	3.40	5.24	0.00	3.40	3.48	3.44	12.04	8.64	6.75	610.70
452	300.00	3.40	5.22	0.00	3.40	3.44	3.42	12.02	8.62	6.74	617.00
453	400.00	3.55	5.36	0.00	3.25	3.41	3.25	12.16	8.91	6.71	930.00
454	400.00	3.50	8.00	0.25	2.90	3.57	2.92	14.65	11.75	7.53	1056.00
455	400.00	4.35	8.00	1.02	2.05	3.47	2.05	15.42	13.37	7.64	1152.00
456	300.00	3.40	5.25	0.00	3.40	3.46	3.41	12.05	8.65	6.75	610.70
457	400.00	4.75	7.60	0.00	2.05	3.49	3.09	14.40	12.35	6.81	1473.00
458	400.00	3.40	7.24	0.00	3.40	3.44	3.40	14.04	10.64	7.26	967.00
459	400.00	4.35	8.00	1.10	2.05	3.60	2.10	15.50	13.45	7.69	1297.80
460	400.00	3.45	8.00	0.12	2.95	3.47	2.95	14.52	11.57	7.47	1318.00
461	400.00	3.45	8.00	0.06	2.95	3.41	2.95	14.46	11.51	7.43	1240.00
462	400.00	4.45	8.00	0.94	1.95	3.49	2.15	15.34	13.39	7.56	1318.00
463	300.00	3.40	5.25	0.00	3.40	3.48	3.43	12.05	8.65	6.75	508.90
464	300.00	3.40	5.30	0.00	3.40	3.51	3.41	12.10	8.70	6.76	559.80
465	400.00	4.10	1.50	0.00	2.70	3.04	2.74	8.30	5.60	4.38	508.90
466	400.00	4.25	8.00	0.40	2.15	3.55	2.75	14.80	12.65	7.32	1392.00
467	400.00	3.45	5.25	0.00	3.35	3.45	3.40	12.05	8.70	6.72	1240.00

Sample	Influencing factors										Pile bearing capacity
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y
468	300.00	3.40	5.25	0.00	3.40	3.49	3.44	12.05	8.65	6.75	559.80
469	400.00	4.35	8.00	1.07	2.05	3.52	2.05	15.47	13.42	7.67	1425.00
470	400.00	4.35	8.00	1.07	2.05	3.52	2.05	15.47	13.42	7.67	1425.00
471	300.00	3.40	5.26	0.00	3.40	3.49	3.43	12.06	8.66	6.75	508.90
472	400.00	3.85	7.60	0.00	2.95	3.67	3.27	14.40	11.45	7.15	1425.00

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors confirm that there are no conflicts of interest regarding the publication of the manuscript.

References

1. Alkroosh IS, Bahadori M, Nikraz H, Bahadori A (2015) Regressive approach for predicting bearing capacity of bored piles from cone penetration test data. J Rock Mech Geotech Eng 7:584–592. https://doi.org/10.1016/j.jrmge.2015.06.011
2. Aoki N, Velloso DDA (1975) An approximate method to estimate the bearing capacity of piles. International Society of Soil Mechanics and Geotechnical Engineering Buenos, pp 367–376
3. Baghbani A, Choudhury T, Costa S, Reiner J (2022) Application of artificial intelligence in geotechnical engineering: a state-of-the-art review. Earth Sci Rev 228:103991. https://doi.org/10.1016/j.earscirev.2022.103991
4. Bazaraa AR, Kurkur MM (1986) N-values used to predict settlements of piles in Egypt. ASCE, pp 462–474
5. Benali A, Nechnech A, Bouafia A (2013) Bored pile capacity by direct SPT methods applied to 40 case histories. Civ Environ Res 5:118–122
6. Bishop CM (2011) Pattern recognition and machine learning (Information Science and Statistics). Springer (**ISBN-10: 0387310738**)
7. Bouafia A, Derbala A (2002) Assessment of SPT-based method of pile bearing capacity–analysis of a database, pp 369–374
8. Briaud JL, Tucker LM (1988) Measured and predicted axial capacity of 98 piles. ASCE J Geotech Eng 114:984–1001
9. Bryngelson SH, Colonius T (2020) Simulation of humpback whale bubble-net feeding models. J Acoust Soc Am 147:1126–1135. https://doi.org/10.1121/10.0000746
10. Bui D-K, Nguyen T, Chou J-S, Nguyen-Xuan H, Ngo TD (2018) A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete. Constr Build Mater 180:320–333. https://doi.org/10.1016/j.conbuildmat.2018.05.201
11. Cao M-T, Nguyen N-M, Wang W-C (2022) Using an evolutionary heterogeneous ensemble of artificial neural network and multivariate adaptive regression splines to predict bearing capacity in axial piles. Eng Struct 268:114769. https://doi.org/10.1016/j.engstruct.2022.114769
12. 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, San Francisco, California, USA, pp 785–794. https://doi.org/10.1145/2939672.2939785
13. Chen W, Sarir P, Bui X-N, Nguyen H, Tahir MM, Jahed Armaghani D (2020) Neuro-genetic, neuro-imperialism and genetic programing models in predicting ultimate bearing capacity of pile. Eng Comput 36:1101–1115. https://doi.org/10.1007/s00366-019-00752-x
14. Cheng M-Y, Cao M-T, Tsai P-K (2020) Predicting load on ground anchor using a metaheuristic optimized least squares support vector regression model: a Taiwan case study. J Comput Des Eng 8:268–282. https://doi.org/10.1093/jcde/qwaa077
15. Cheng M-Y, Prayogo D, Wu Y-W (2018) Prediction of permanent deformation in asphalt pavements using a novel symbiotic organisms search–least squares support vector regression. Neural Comput Appl. https://doi.org/10.1007/s00521-018-3426-0
16. Chow YK, Chan WT, Liu LF, Lee SL (1995) Prediction of pile capacity from stress-wave measurements: a neural network approach. Int J Numer Anal Meth Geomech 19:107–126. https://doi.org/10.1002/nag.1610190204
17. Das SK, Basudhar PK (2006) Undrained lateral load capacity of piles in clay using artificial neural network. Comput Geotech 33:454–459. https://doi.org/10.1016/j.compgeo.2006.08.006
18. Decourt L (1995) Prediction of load settlement relationships for foundations on the basis of the SPT-T. In: Ciclo de conferencias Inter"Leonardo Zeevaert", UNAM, Mexico, pp 85–104
19. Dong W, Huang Y, Lehane B, Ma G (2020) XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring. Autom Constr 114:103155. https://doi.org/10.1016/j.autcon.2020.103155
20. Duan J, Asteris PG, Nguyen H, Bui X-N, Moayedi H (2021) A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model. Eng Comput 37:3329–3346. https://doi.org/10.1007/s00366-020-01003-0
21. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29:1189–1232. https://doi.org/10.1214/aos/1013203451
22. Gharehchopogh FS, Gholizadeh H (2019) A comprehensive survey: whale optimization algorithm and its applications. Swarm Evol Comput 48:1–24. https://doi.org/10.1016/j.swevo.2019.03.004

23. Hoang N-D (2019) Estimating punching shear capacity of steel fibre reinforced concrete slabs using sequential piecewise multiple linear regression and artificial neural network. Measurement 137:58–70. https://doi.org/10.1016/j.measurement.2019.01.035

24. Hoang N-D, Tran X-L, Huynh T-C (2022) Prediction of pile bearing capacity using opposition-based differential flower pollination-optimized least squares support vector regression (ODFP-LSSVR). Adv Civ Eng 2022:7183700. https://doi.org/10.1155/2022/7183700

25. Hoang N-D, Tran X-L, Nguyen H (2020) Predicting ultimate bond strength of corroded reinforcement and surrounding concrete using a metaheuristic optimized least squares support vector regression model. Neural Comput Appl 32:7289–7309. https://doi.org/10.1007/s00521-019-04258-x

26. Jahed Armaghani D, Shoib RS, Faizi K, Rashid AS (2017) Developing a hybrid PSO–ANN model for estimating the ultimate bearing capacity of rock-socketed piles. Neural Comput Appl 28:391–405. https://doi.org/10.1007/s00521-015-2072-z

27. Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. arXiv:14126980. https://doi.org/10.48550/arXiv.1412.6980

28. Kurtoglu AE, Gulsan ME, Abdi HA, Kamil MA, Cevik A (2017) Fiber reinforced concrete corbels: modeling shear strength via symbolic regression. Comput Concr 20:065–075

29. Le LT, Nguyen H, Zhou J, Dou J, Moayedi H (2019) Estimating the heating load of buildings for smart city planning using a novel artificial intelligence technique PSO-XGBoost. Appl Sci 9:2714. https://doi.org/10.3390/app9132714

30. Lee I-M, Lee J-H (1996) Prediction of pile bearing capacity using artificial neural networks. Comput Geotech 18:189–200. https://doi.org/10.1016/0266-352X(95)00027-8

31. Liu W, Liu WD, Gu J (2020) Predictive model for water absorption in sublayers using a joint distribution adaption based XGBoost transfer learning method. J Petrol Sci Eng 188:106937. https://doi.org/10.1016/j.petrol.2020.106937

32. Mafarja MM, Mirjalili S (2017) Hybrid whale optimization algorithm with simulated annealing for feature selection. Neurocomputing 260:302–312. https://doi.org/10.1016/j.neucom.2017.04.053

33. Mason L, Baxter J, Bartlett P, Frean M (1999) Boosting algorithms as gradient descent in function space, pp 512–518

34. Meyerhof GG (1976) Bearing capacity and settlement of pile foundations. J Geotech Eng Div 102:197–228. https://doi.org/10.1061/AJGEB6.0000243

35. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

36. Mirrashid M, Naderpour H (2021) Recent trends in prediction of concrete elements behavior using soft computing (2010–2020). Arch Comput Methods Eng 28:3307–3327. https://doi.org/10.1007/s11831-020-09500-7

37. Moayedi H, Jahed Armaghani D (2018) Optimizing an ANN model with ICA for estimating bearing capacity of driven pile in cohesionless soil. Eng Comput 34:347–356. https://doi.org/10.1007/s00366-017-0545-7

38. Mohanty R, Suman S, Das SK (2018) Prediction of vertical pile capacity of driven pile in cohesionless soil using artificial intelligence techniques. Int J Geotech Eng 12:209–216. https://doi.org/10.1080/19386362.2016.1269043

39. Momeni E, Nazir R, Jahed Armaghani D, Maizir H (2014) Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. Measurement 57:122–131. https://doi.org/10.1016/j.measurement.2014.08.007

40. Murlidhar BR, Sinha RK, Mohamad ET, Sonkar R, Khorami M (2020) The effects of particle swarm optimisation and genetic algorithm on ANN results in predicting pile bearing capacity. Int

J Hydromechatronics 3:69–87. https://doi.org/10.1504/IJHM.2020.105484

41. Nguyen H (2020) PSO-XGBoost hybrid model to predict long-term deflection of reinforced concrete members. 10.5281/zenodo.3932822

42. Nguyen H, Nguyen N-M, Cao M-T, Hoang N-D, Tran X-L (2021) Prediction of long-term deflections of reinforced-concrete members using a novel swarm optimized extreme gradient boosting machine. Eng Comput. https://doi.org/10.1007/s00366-020-01260-z

43. Nguyen H, Vu T, Vo TP, Thai H-T (2021) Efficient machine learning models for prediction of concrete strengths. Constr Build Mater 266:120950. https://doi.org/10.1016/j.conbuildmat.2020.120950

44. Nguyen T-D, Tran T-H, Hoang N-D (2020) Prediction of interface yield stress and plastic viscosity of fresh concrete using a hybrid machine learning approach. Adv Eng Inform 44:101057. https://doi.org/10.1016/j.aei.2020.101057

45. Nguyen T-D, Tran T-H, Nguyen H, Nhat-Duc H (2021) A success history-based adaptive differential evolution optimized support vector regression for estimating plastic viscosity of fresh concrete. Eng Comput 37:1485–1498. https://doi.org/10.1007/s00366-019-00899-7

46. Nhu V-H et al (2020) Effectiveness assessment of Keras based deep learning with different robust optimization algorithms for shallow landslide susceptibility mapping at tropical area. CATENA 188:104458. https://doi.org/10.1016/j.catena.2020.104458

47. Park HI, Cho CW (2010) Neural network model for predicting the resistance of driven piles. Mar Georesour Geotechnol 28:324–344. https://doi.org/10.1080/1064119X.2010.514232

48. Pedregosa F et al (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830

49. Pham TA, Ly H-B, Tran VQ, Giap LV, Vu H-LT, Duong H-AT (2020) Prediction of pile axial bearing capacity using artificial neural network and random forest. Appl Sci 10:1871

50. Pham TA, Tran VQ, Vu H-LT, Ly H-B (2020) Design deep neural network architecture using a genetic algorithm for estimation of pile bearing capacity. PLoS ONE 15:e0243030. https://doi.org/10.1371/journal.pone.0243030

51. Pirotta V, Owen K, Donnelly D, Brasier MJ, Harcourt R (2021) First evidence of bubble-net feeding and the formation of 'super-groups' by the east Australian population of humpback whales during their southward migration. Aquat Conserv Mar Freshw Ecosyst 31:2412–2419. https://doi.org/10.1002/aqc.3621

52. Prayogo D, Cheng M-Y, Wu Y-W, Tran D-H (2020) Combining machine learning models via adaptive ensemble weighting for prediction of shear capacity of reinforced-concrete deep beams. Eng Comput 36:1135–1153. https://doi.org/10.1007/s00366-019-00753-w

53. Rausche F, Goble GG, Likins GE (1985) Dynamic determination of pile capacity. J Geotech Eng 111:367–383. https://doi.org/10.1061/(ASCE)0733-9410(1985)111:3(367)

54. Rokach L, Maimon OZ (2007) Data mining with decision trees: theory and applications. World Scientific

55. Samui P (2011) Prediction of pile bearing capacity using support vector machine. Int J Geotech Eng 5:95–102. https://doi.org/10.3328/IJGE.2011.05.01.95-102

56. Shahin MA, Jaksa MB (2005) Neural network prediction of pullout capacity of marquee ground anchors. Comput Geotech 32:153–163. https://doi.org/10.1016/j.compgeo.2005.02.003

57. Shariatmadari N, Eslami A, Karimpour-Fard M (2008) Bearing capacity of driven piles in sands from SPT-applied to 60 case histories Iranian Journal of Science & Technology. Trans B Eng 32:125–140

58. Shioi Y, Fukui J (1982) Application of N-value to design of foundations in Japan. In: Proceedings of the second European symposium on penetration testing, Amsterdam, 24–27 May 1982. Routledge, pp 159–164

59. Shooshpasha I, Hasanzadeh A, Taghavi A (2013) Prediction of the axial bearing capacity of piles by SPT-based and numerical design methods. Int J GEOMATE 4:560–565

60. Tang J, Zheng L, Han C, Liu F, Cai J (2020) traffic incident clearance time prediction and influencing factor analysis using extreme gradient boosting model. J Adv Transp 2020:e6401082. https://doi.org/10.1155/2020/6401082

61. Teh CI, Wong KS, Goh ATC, Jaritngam S (1997) Prediction of pile capacity using neural networks. J Comput Civ Eng 11:129–138. https://doi.org/10.1061/(ASCE)0887-3801(1997)11:2(129)

62. Thieu NV (2021) A collection of the state-of-the-art meta-heuristics algorithms in PYTHON: mealpy. Zenodo. https://doi.org/10.5281/zenodo.3711948

63. Tien Bui D, Hoang N-D, Nguyen H, Tran X-L (2019) Spatial prediction of shallow landslide using Bat algorithm optimized machine learning approach: a case study in Lang Son Province. Vietnam Adv Eng Inform 42:100978. https://doi.org/10.1016/j.aei.2019.100978

64. XGBoost-Documentation (2021) XGBoost documentation—xgboost 1.5.0-dev documentation. https://xgboost.readthedocs.io/en/latest/index.html

65. XGBoost (2021) XGBoost: scalable and flexible gradient boosting. https://xgboost.ai/

66. Xia D, Zheng Y, Bai Y, Yan X, Hu Y, Li Y, Li H (2022) A parallel grid-search-based SVM optimization algorithm on Spark for passenger hotspot prediction. Multimed Tools Appl 81:27523–27549. https://doi.org/10.1007/s11042-022-12077-x

67. Zhang W, Wu C, Zhong H, Li Y, Wang L (2021) Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. Geosci Front 12:469–477. https://doi.org/10.1016/j.gsf.2020.03.007

68. Zhang X, Nguyen H, Bui X-N, Tran Q-H, Nguyen D-A, Bui DT, Moayedi H (2020) Novel soft computing model for predicting blast-induced ground vibration in open-pit mines based on particle swarm optimization and XGBoost. Nat Resour Res 29:711–721. https://doi.org/10.1007/s11053-019-09492-7