Rouhollah Khorrami<sup>a</sup> and Ali Derakhshani<sup>\*</sup>

Department of Civil Engineering, Faculty of Engineering, Shahed University, Tehran, Iran

(Received February 25, 2018, Revised June 21, 2019, Accepted September 26, 2019)

**Abstract.** Available methods to determine the ultimate bearing capacity of shallow foundations may not be accurate enough owing to the complicated failure mechanism and diversity of the underlying soils. Accordingly, applying new methods of artificial intelligence can improve the prediction of the ultimate bearing capacity. The M5' model tree and the genetic programming are two robust artificial intelligence methods used for prediction purposes. The model tree is able to categorize the data and present linear models while genetic programming can give nonlinear models. In this study, a combination of these methods, called the M5'-GP approach, is employed to predict the ultimate bearing capacity of the shallow foundations, so that the advantages of both methods are exploited, simultaneously. Factors governing the bearing capacity of the foundation (L), effective unit weight of the soil ( $\gamma$ ) and internal friction angle of the soil ( $\varphi$ ) are considered for modeling. To develop the new model, experimental data of large and small-scale tests were collected from the literature. Evaluation of the new model by statistical indices reveals its better performance in contrast to both traditional and recent approaches. Moreover, sensitivity analysis of the proposed model indicates the significance of various predictors. Additionally, it is inferred that the new model compares favorably with different models presented by various researchers based on a comprehensive ranking system.

Keywords: artificial intelligence; M5'-GP hybrid model; ultimate bearing capacity; shallow foundation; granular soil

# 1. Introduction

One of the most important issues that civil engineers usually encounter in various projects is the design of the foundations of the structures, which is affected by two factors including ultimate bearing capacity and settlement. The maximum stress from the foundation, which the soil withstands without shear failure is defined as the ultimate bearing capacity. Foundations are normally categorized into two groups including shallow and deep foundations, according to their depth and application. In shallow foundations, the ratio of depth to width is smaller than or equal to four (Das, 2015). Different researchers have investigated the problem of the estimation of the ultimate bearing capacity of shallow foundations. The most wellknown studies are those conducted by Terzaghi (1943), Meyerhof (1963), Hansen (1970) and Vesic (1973). These researches are based on classic methods which tried to simplify the issue by considering some assumptions. Using the results of laboratory model tests on the shallow foundations is very useful for improving the weaknesses of the prediction of ultimate bearing capacity by traditional formulae derived considering some simplifying assumptions.

Model tests of different scales have been conducted to

investigate the behavior of the shallow foundations. Because performing large scale experiments is costly and time consuming, small-scale experiments were also done in addition to large-scale ones.

In addition to laboratory tests, many other researches focused on the numerical modeling of shallow foundations through various methods such as finite element, finite difference, etc. In recent years, soft computational approaches which are able to find hidden complex relations in various phenomena are used to model the ultimate bearing capacity of the foundation. Some investigations conducted on predicting the foundation bearing capacity by these methods include the studies by Padmini et al. (2008) using the neuro-fuzzy model, Kalinli et al. (2011) using the artificial neural network (ANN), Shahnazari and Tutunchian (2012) using multi-gene genetic programming (GP) and multiple linear regression (MLR), Tsai et al. (2013) using the three types of genetic programming (GP), Sadrossadat et al. (2013) using the linear genetic programming (LGP), etc. The equations derived from the application of these methods as well as the classic relationships are listed in Table 1.

Parameters including the width of the foundation (B), the embedment depth of the foundation (D), the length of the foundation (L), the effective unit weight of the soil ( $\gamma$ ) and the internal friction angle of the soil ( $\varphi$ ) are effective in determining the ultimate bearing capacity of the shallow foundations on cohesionless soils (Foye *et al.* 2006). Thus, these parameters are used as the inputs to create models by means of artificial intelligence.

In this research, a new hybrid approach called "M5'-

<sup>\*</sup>Corresponding author, Assistant Professor

E-mail: adera@shahed.ac.ir

<sup>&</sup>lt;sup>a</sup>Graduate Student

Classic approaches	
Reference	Relationship
Terzaghi (1943)	$qu = qN_q + \frac{1}{2}\gamma BN_{\gamma}$
	$N_q = K_q \tan \varphi', N_\gamma = \frac{1}{2} \tan \varphi' (K_\gamma \tan \varphi' - 1)$
Meyerhof (1963)	$qu = qN_qF_{qs}F_{qd}F_{qi} + \frac{1}{2}\gamma BN_\gamma F_{\gamma s}F_{\gamma d}F_{\gamma i}$
	$N_q = tan \left(45 + \frac{\varphi}{2}\right)^2 e^{\pi tan \varphi}, N_\gamma = \left(N_q - 1\right) \tan 1.4\varphi$
Hansen (1970)	$qu = qN_qF_{qs}F_{qd}F_{qi} + \frac{1}{2}\gamma BN_{\gamma}F_{\gamma s}F_{\gamma d}F_{\gamma i}$
	$N_q = tan \left(45 + \frac{\varphi}{2}\right)^2 e^{\pi tan \varphi}, N_{\gamma} = 1.5 \left(N_q - 1\right) \tan \varphi$
Vesic (1973)	$qu = qN_qF_{qs}F_{qd}F_{qi} + \frac{1}{2}\gamma BN_{\gamma}F_{\gamma s}F_{\gamma d}F_{\gamma i}$
	$N_q = tan \left(45 + \frac{\varphi}{2}\right)^2 e^{\pi tan \varphi}, N_{\gamma} = 2(N_q - 1) \tan \varphi$
	Artificial intelligence approaches
Reference	Relationship
Shahnazari-MLR (2012)	$qu = 10^{3}(-2.6264 + 0.0592B + 1.9714D) - 0.0135\frac{L}{B} - 0.0305\gamma + 0.0827\varphi)$
Shahnazari- Multigene GP (2012)	$qu = 2 \times 10^{-12} \varphi^{7} (B + D) \left(\frac{L}{B} + \varphi\right)^{2} + \frac{10^{-8} D^{2} \varphi^{6} (B - \gamma)^{2}}{\frac{L}{B}}$
Tsai-GP (2013)	$qu = e^{0.134\varphi} + \left(\frac{2.88 - \varphi}{D\gamma}\right) + (\log D + B\varphi)^{\log 6.77 + \gamma}$
Sadrossadat-LGP (2013)	$qu = \varphi((\gamma + \varphi + \frac{3.95(\varphi - 35)^2}{\frac{L}{B}})D^2 + 2.5(B(\varphi - 35) + 1))$

GP" was employed to predict the ultimate bearing capacity of the shallow foundation on granular soil. The basis of the method is classification of the input data using M5' model tree followed by creating the relationships using genetic programming. Certain ranges of parameters involved in the shallow foundation bearing capacity problem were considered and it was assumed that the soil is granular, the ground is horizontal and the load is vertically applied to the center of the foundation. The rest of this paper is outlined as follows. The second section gives an explanation about model tree, genetic programming, and the hybrid model. Details of the collected data are shown in the third section. The fourth section explains about building the model, obtained results, and discussion. The model sensitivity relative to input parameters is discussed in the fifth section. A comparison based on a ranking system is given in the sixth section and finally, the conclusions are provided in the seventh section.

# 2. Hybrid modeling approach

# 2.1 Model tree

The model tree is a strong method highly capable of analyzing the data and presenting predictive models. This capability is based upon the formation of the various relationships in the appropriate branches of the data created by the algorithm according to the output values. Model tree is a special form of regression trees. Also, regression tree is a specific type of the decision tree. A decision tree which predicts the numerical variables is called the regression tree. Moreover, a regression tree, which presents linear regression equations on its leaves, is defined as the model tree. The M5' is a kind of model tree algorithms primarily introduced by (Wang 1997). In order to build a model tree, a tree is first created by the M5' algorithm. The data branching is performed based on the expected standard deviation reduction (SDR) for each attribute (input data). The process of dividing the data continues until the sum of the square deviations from the mean of the data reaches almost zero or a few number of samples remain (Quinlan 1992). After creating the tree, a linear regression relationship will be formed for each leaf of the tree. If the model's standard deviation reduction in the first node (root) of sub-tree is less than or equal to the expected error of subtree, the sub-tree will be pruned (Avval and Derakhshani 2018, Etemad-Shahidi and Ghaemi 2011, Jafariavval and Derakhshani 2019). The optimal tree is selected based on minimizing the prediction error. Finally, in order to compensate the discontinuity which is unavoidably appeared between the linear models in leaves of the pruned tree, smoothing will be performed (Bhattacharya and Solomatine 2005). In this operation, by integrating the available model in each leaf into available models in the root-to-leaf path, the ultimate model is given in each leaf (Kaveh 2018, Wang, 1997).

#### 2.2 Genetic programming

Genetic programming is a robust method developed for automatic programming. Genetic programming is a special form of the genetic algorithm that uses its operators such as crossover, mutation and reproduction, with the chromosome sizes changing by the modified genetic operators. This symbolic optimization technique based on Darwin's theory was introduced by Koza (1992) and Banzhaf *et al.* (1998). The equation trees generated by this method will be used as a model or a computer program. Each of these models or computer programs is a member of the population. New generation will be generated by applying genetic operators such as the crossover, mutation and reproduction to the previous population. In the crossover, a random gene is

Table 1 Classic and artificial intelligence-based equations



Fig. 1 Outline of the M5'-GP hybrid method

selected from both parents that exchange their genes to produce new children. In the mutation, a random gene is selected from the parent. Then, the gene is either deleted or replaced by a new random gene. In reproduction, suitable individuals are copied to new population. Chromosomes representation is in the form of tree and graph. The individuals of the population in genetic programming, are the computer-generated trees (graphs) which are the hierarchical constructs having different sizes and shapes. These are made of sets of functions and terminals which are determined by the user. Functions are mathematical operators and terminals or leaves are the variables or constants. Afterward, the genetic programming will perform a revolutionary search in an enormous space of equations that can be expressed by these initial values. Genetic programming includes the following steps:

1. Generating a primary population from the random combination of functions and terminals which produces N programs (solutions) with different size and shape.

2. Running each program of the population and evaluating it via obtaining the solution fitness.

3. Generating a new population of the individuals (programs) based on the selection of the genetic operators randomly.

- If the selected operator is reproduction, the best individuals of the current population will be selected and copied to the new one.

- If the selected operator is crossover, two parent individuals will be selected. Then, genes of individuals are selected and they are replaced by each other. The two generated children are put in the new population. Crossover plays a vital role in the evolutionary process.

- If the selected operator is mutation, an individual will be chosen from the current population. Then, a gene of this individual is selected and mutation process is performed by either deleting or replacing the gene. Mutation helps keeping diversity.

4. The third stage continues until the new population reaches the N members. Performing stages 2 to 4 continue until a proper solution is obtained and if it does not get any proper solution, it will end after a certain number of iterations. Proper solution assessment is accomplished by an index that is the sum of the absolute difference values between the obtained results and observed results which is originally called Euclidean distance. The lower index is an indication of the superiority of the individual. It is worth



Fig. 2 An example of the M5'-GP method application

noting that different variations of genetic programming have been successfully used to solve prediction problems in geotechnical engineering e.g., (Alavi *et al.* 2010, Javadi and Rezania 2009, Li *et al.* 2016, Talebi and Derakhshani 2019).

#### 2.3. M5'-GP hybrid approach

M5'-GP hybrid method is based on using the capabilities and strong points of the M5' and GP methods for solving the prediction problems. The hybrid use of these two methods overcomes the incapability of M5' model tree to present the non-linear prediction models by using the GP. On the other hand, the capability of M5' in categorizing the data space in conjunction with the ability of the GP for non-

Table 2 Parameters used to develop the prediction model



T 11 A	a		C .1	•	1 .
Table 4	Statistical	decomption	of the	inniif	data
Table 5	Statistical	ucscribtion	or the	mout	uata
-					

Parameters		B (m)	D (m)	L (m)	$\gamma$ (kN/m <sup>3</sup> )	φ (rad)
м. :	Train	3.02	0.89	3.02	20.80	0.80
Maximum	Test	3.00	0.76	3.00	20.80	0.80
	Train	0.03	0.00	0.03	9.85	0.56
Minimum	Test	0.04	0.00	0.04	10.2	0.56
М	Train	0.52	0.12	0.78	15.58	0.68
Mean	Test	0.59	0.10	0.87	15.88	0.68
Ctow dowed above the se	Train	0.52	0.21	0.65	2.94	0.06
Standard deviation –	Test	0.71	0.20	0.87	2.93	0.06

linear modelling can provide better prediction results (Bonakdar *et al.* 2015, Derakhshani 2017, 2018). In the hybrid method, input data is first divided into different groups with the help of branching feature of the model tree algorithm. Then, the GP is applied to each of the branches so that a predictive model with higher degree of accuracy is obtained. The procedures of the hybrid M5'-GP method are

displayed in Fig. 1 and described in details as follows.

#### 2.3.1 Categorizing the data by M5' model tree

In this stage, by introducing the input data to M5' algorithm, data is divided into various branches based on the criteria proposed by the model tree algorithm. The data is divided into branches based on the independent variables.

Hence, based on the defined condition in each node, data is divided into two groups in that node. If suitable subsets are not obtained, branching will be continued in each of the branches, otherwise branching will be stopped.

#### 2.3.2 Deriving prediction equations by GP

In the second stage, the input data of each branch is separately introduced to the model.

These groups of data are divided into training and testing datasets. After introducing the data to the model, the required settings of the genetic programming are also adjusted. Then, the genetic programming is applied to different subsets and their prediction equations will be obtained.

#### 2.3.3 Evaluating the developed model

To assess the model performance, the predicted values on each branch using the corresponding equations and measured values are taken into consideration and their error values are calculated. An example of the application of the M5'-GP hybrid method is shown in Fig. 2.

#### 3. The database for model development

To develop the M5'-GP hybrid model, 169 data were collected from the 12 publications by Golder et al. (1941), Eastwood (1951), Subrahmanyam (1967), Muhs et al. (1969), Weiß (1970), Muhs and Weiß (1971), Muhs and Weiß (1973), Briaud and Gibbens (1997) Briaud and Gibbens (1999), Gandhi (2003), Cerato and Lutenegger (2007), and Akbas and Kulhawy (2009). These data are the observed values of ultimate bearing capacity of the shallow foundations on the granular soils corresponding to different values of input parameters. Such experimental results can reflect the physical behavior of the system well, although the model tests have their own limitations. Input and output parameters introduced to the model, are presented in Table 2. The reference experiments were performed with different scales, where, 65 and 104 data out of 169 data are related to the small and large scale experiments, respectively. Hence, suitable ranges of predictors were considered for modeling.

Fig. 3 depicts the distribution of the values of various parameters in the database. It can be seen that the distributions of different variables are not uniform and the densities of the samples differ for various intervals. For example, Fig. 3(b) and 3(e) regarding D and  $\gamma$ , illustrate that the samples are mainly weighted in the intervals of [0, 0.1] m and [14, 17] kN/m<sup>3</sup>, respectively. Better prediction performance is probably achieved in case of greater densities of variables.

The statistical specifications of the data are given in Table 3. In order to create the model and assess its performance, the data is divided into two parts including training (80%) and evaluation (20%). These classified parts of the data should have close statistical indices (Shahin *et al.*, 2004)

# 4. Creating the M5'-GP model and discussion on the results

#### 4.1 Building the model

In order to construct the prediction model of the ultimate

Table 4 Statistical performance of M5'-GP

Data		Error measure						
	CC (%)	RMSE	MAE					
Train	93.61	168.23	99.87					
Test	97.1	111.61	67.83					
All	94.2	158.48	93.42					

bearing capacity, the data of different input parameters and the output were introduced to the M5' algorithm. By using the model tree method in WEKA (Witten *et al.*, 2016), the data was divided into two branches by selecting the foundation width as the split criterion. This means that objects in the same group are more similar to each other than to those in other group. The splitting value was selected to be (B =) 0.275m (determined using the classification procedure in model tree) upon which the two branches were formed and separately considered for modeling by genetic programming (Searson, 2015). The initial settings that should be adjusted for genetic programming was changed many times for both subsets.

Then, the best cases were selected among the results achieved based on the various settings used for the implementation of genetic programming. Finally, the prediction model of the ultimate bearing capacity was yielded by the following equations:

For 
$$B \leq 0.275$$

$$qu = 4.83\varphi^{2}(B+D) + 298(B+D)(\gamma - \varphi)$$
(1-a)  
$$-\frac{1.08\gamma(B-D)}{L} + 63.35$$

For 
$$B > 0.275$$

$$qu = 58.2 B - 447\varphi + 505 D(1 - \varphi) - 752D(L + \gamma) + 6.35\varphi(B + \varphi) + 11.9D\varphi(\varphi - \gamma) + 7666$$
(1-b)

The new M5'-GP model is a structured representation (with explicit form) of the phenomena being investigated. This transparent configuration is the important advantage of the proposed model over black-box approaches like ANN, DNN, etc. used for various prediction purposes in civil engineering (Derakhshani and Foruzan 2019).

#### 4.2 Assessment of the model performance

In order to evaluate the accuracy of the proposed "M5'-GP" model, three statistical indicators, namely, correlation coefficients (CC), root mean square error (RMSE) and mean absolute error (MAE) were used. The equations of the three evaluation indices are as follows

$$CC = \frac{\sum (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$
(2)

$$RMSE = \sqrt{\frac{\sum (x_i - y_i)^2}{N}}$$
(3)

$$MAE = \frac{\sum(x_i - y_i)}{N} \tag{4}$$



Table 5 Comparison of the accuracy of the results by the M5'- GP and other models

Fig. 4 Observed and estimated ultimate bearing capacity by M5'-GP: (a) train (b) test (c) all

In equations, xi is the measured value, yi is the predicted value, N is the number of measured values and  $\bar{x}$  and  $\bar{y}$  are the measured and predicted mean values. The correlation coefficient greater than 0.8 (CC> 0.8) shows a strong correlation between the model and experimental data (Smith, 1986). Also, the lower RMSE and MAE errors indicate better performance of a prediction model. The results of the evaluation of the new model by the statistical indices considering the train, test and all data are presented in Table 4.

The results of Table 4 illustrate the ability of the proposed hybrid computational approach to predict the

ultimate bearing capacity of shallow foundations. Various measures obtained for different sets of data show that the new approach is capable of presenting accurate estimates of the ultimate bearing capacity.

#### 4.3 Comparison with other approaches

The new model for prediction of the ultimate bearing capacity of the shallow foundation on cohesionless soil was assessed using the statistical measures. These indices can be also considered to compare the new hybrid model with the classical methods suggested by Terzaghi (1943), Meyerhof



(a) Classic

(b) Artificial Intelligence





Fig. 6 DR values of various models for prediction of the ultimate bearing capacity

(1963), Hansen (1970) and Vesic (1973), as well as the methods based on artificial intelligence presented by Shahnazari and Tutunchian (2012) (MLR), Shahnazari and Tutunchian (2012) (MGP), Tsai *et al.* (2013) and Sadrossadat *et al.* (2013). Table 5 compares the evaluation indices of new hybrid model with those of various classical methods and artificial intelligence-based methods available for predicting the ultimate bearing capacity.

As shown in Table 5 the performance of the proposed model is better and it is more accurate than the classical and artificial intelligence-based methods for predicting the ultimate bearing capacity. Fig. 4 depicts the correlation between the measured and predicted values of the hybrid model regarding the training, test, and total data. In Fig. 4, the closeness of points to the ideal line (predicted = measured) indicates the ability of the model to predict the ultimate bearing capacity. According to the plots of Fig. 4, it can be found that the model's accuracy at lower values of ultimate bearing capacity is higher than that of greater values. In Fig. 5, the results of the proposed prediction model are compared with classical and artificial intelligence-based methods in separate graphs regarding the test data. Fig. 5 displays that the predicted values by the hybrid model are closer to those observed in the experiments comparing with classical and computational intelligence-based methods. The conformity between the predicted and the measured values can be determined via calculation of DR<sub>i</sub> as follows

$$DR_i = \frac{y_i}{x_i} \tag{5}$$

In Eq. (5),  $x_i$  is the measured value and  $y_i$  is the predicted value. The closer the average  $DR_i$  of the model results to one, the more accurate is the model in predicting the bearing capacity. Fig. 6, illustrates the histograms of the  $DR_i$  values of different methods. The models that have greater number of predicted-to-measured ratios in the range of 0.75 to 1.25 are more reliable to predict the ultimate bearing capacity. As shown in fig. 6, for the hybrid model, about 80% of the data is in this range, which is the best value among different models.

#### 4.4 Parametric study

The response of the prediction models to their independent variables can be investigated via conducting a parametric analysis (Rostami *et al.* 2018, Sadrossadat *et al.* 2018). This investigation is performed by evaluating the response of the predictive model to the variations of a desired predictor while other predictors are kept constant at their mean values in the whole dataset. The values of every input parameter being studied were introduced to the model equations and the qu was computed. Fig. 7 illustrates the tendency of the qu estimations to the changes of different



Fig. 7 Parametric study on the developed M5'-GP model

governing parameters. These charts demonstrate that qu increases with increasing B, D,  $\gamma$ , and f. Also, it is shown that qu decreases with increase in L/B. In general, it is found that the trends are consistent with those of experimental data in accordance with the results reported by Sadrossadat *et al.* (2013).

# 5. Sensitivity analysis

To investigate the effects of different predictors on the prediction of the ultimate bearing capacity of the shallow foundation provided by the hybrid method, a set of sensitivity analyses are performed. The significance of an input parameter on the prediction results can be determined by changing its value and keeping other parameters constant. Initially, the average and the standard deviation values of the data for all parameters were calculated. Then, the influences of the addition and subtraction of the standard deviation of a certain parameter with respect to its average value were considered. To this aim, the values of SA<sub>1</sub><sup>1</sup> and SA<sub>2</sub><sup>2</sup> were computed as follows

Table 6 Sensitivity analysis of the input parameters

		-		-	
Parameter	В	D	L	φ	γ
Equation (1-a)	58	64	72	95	74
Ra	5	4	3	1	2
Equation (1-b)	330	1271	155	726	324
Rb	3	1	5	2	4
Ra+Rb	8	5	8	3	6
R	4	2	4	1	3

$$SA_i^1 = Qu(M + \sigma_i) - Qu(M) \tag{6}$$

$$SA_i^2 = Qu(M - \sigma_i) - Qu(M) \tag{7}$$

where Qu is the predictive model for determining the ultimate bearing capacity, M denotes the mean values of all data for each parameter and  $\sigma_i$  is the standard deviation of the ith input which its significance is evaluated. Accordingly, the SA value for each parameter is calculated by the following equation

$$SA_i = |SA_i^1 - SA_i^2| \tag{8}$$

Method	Correlation coefficient		Arithmetic calculations of qup /qum			Cumulative probability			Overall ranking	
	CC (%)	$R_1$	μ	σ	$R_2$	P50	P90	<b>R</b> <sub>3</sub>	RI	R
Terzaghi (1943)	84.90	7	0.99	0.41	2	0.86	1.78	4	13	4
Meyerhof (1963)	94.50	4	0.89	0.31	3	0.90	1.36	2	9	2
Hansen (1970)	94.40	5	0.73	0.30	7	0.71	1.08	9	21	7
Vesic (1973)	94.51	3	0.88	0.31	5	0.87	1.28	3	11	3
Shahnazari-MLR (2012)	72	9	0.80	1.61	6	0.76	1.84	8	23	8
Shahnazari-MGP (2012)	93.7	6	0.84	0.39	4	0.80	1.28	5	15	5
Tsai-GP (2013)	73.27	8	0.75	0.27	9	0.82	1.66	6	24	9
Sadrossadat-LGP (2013)	95.9	2	0.73	0.37	8	0.78	1.08	7	17	6
M5′ - GP	97.1	1	1.02	0.21	1	1.01	1.28	1	3	1

Table 7 Ranking system of the prediction models for ultimate bearing capacity

The greater SA<sub>i</sub> value for a parameter indicates the greater effect of that parameter on the estimation of the ultimate bearing capacity. The value of SAi was calculated for each parameter in each branch of the model separately and the importance of the parameters on the predictions of each branch were obtained. Table 6 summarizes the values of SA<sub>i</sub> for different governing parameters. As can be seen, the effectiveness of the parameters in the first branch is ranked by the parameter Ra and that of the second branch is ranked by the parameter Rb. Then, the significance of the parameters on the estimations of both branches is obtained by the summation of Ra and Rb. The lower the value of this summation for a certain parameter, the greater the effect of that parameter. The final ranking of the importance of the parameters is specified by the parameter R. As shown in Table6, the most significant parameter is the internal friction angle of the soil ( $\varphi$ ) followed by the embedment depth of the foundation (D). The next important input parameters are respectively the effective unit weight of the soil ( $\gamma$ ) and, width and length of the foundation (B and L).

## 6. Ranking

In order to compare different methods of predicting the bearing capacity of the shallow foundations on granular soils comprehensively, a ranking system was created (Abu-Farsakh and Titi 2004). The ranking presented in Table 7, is conducted for the classical and the computational intelligence-based methods in addition to the new hybrid model. Three scoring indices of this system are:

R1 is the scoring based on the correlation rate of the predicted and measured values using the correlation coefficient (CC). As shown in Table 7, the correlation coefficient of the hybrid model is better than other models, hence the value of R1 for the hybrid model is considered to be 1. The other models are also scored based on their correlation coefficient.

R2 is the scoring based on the mean ( $\mu$ ) and the standard deviation ( $\sigma$ ) of the predicted to the measured values,  $\frac{qu_p}{qu_m}$ . The ranking of models were determined based on the

closeness of the mean values of  $\frac{qu_p}{qu_m}$  to one and their standard deviation to zero. According to the R2 index, the hybrid model ranks first due to its proper performance regarding both criteria.

R3 is the scoring based on P50 and P90 that are the cumulative probabilities of 50% and 90% of the  $\frac{qu_p}{qu_m}$  values. In other words, 50% of the ratios of the predicted to measured values are smaller than P50; and P90 is the upper bound value of 90% of the ratios of the predictions to observations. The less the difference between the two values of P50 and P90, the better the performance of the model.

The final ranking is determined by the index RI which equals to summation of R1, R2 and R3. The lower RI exhibits the greater ability of a method in predicting the ultimate bearing capacity. Table 7 shows that the proposed model has the best performance with regard to all indices as it obtained the best rank compared to other methods. This reveals the superiority of the hybrid model from different viewpoints, in contrast to the other approaches considered.

# 7. Conclusions

In this research, a new hybrid model was developed for predicting the ultimate bearing capacity of shallow foundations on granular soils. The model was built based on M5' model tree and genetic programming methods with consideration of strong points and ignoring the weak points of each method. To create this model, 169 laboratory data was collected from the experimental studies of the previous researchers. The data is initially split by the model tree. Then, the GP model was applied on each branch of data and the prediction equations for the ultimate bearing capacity were presented. The results of the novel hybrid method were evaluated by several statistical indices. The model was also compared with classic and artificial intelligence techniques, previously suggested to determine the ultimate bearing capacity of shallow foundations. The following conclusions can be drawn from the present study:

1. A robust model consisting of two equations was

derived for predicting the ultimate bearing capacity of shallow foundations on granular soil using a hybrid approach. The evaluation of this model indicated that the ability of the M5' algorithm for branching the data associated with the ability of genetic programming for nonlinear modeling, can well simulate the experimental values of the ultimate bearing capacity by suggesting a relationship with high correlation coefficient.

2. The M5'-GP hybrid model was compared with the classic methods including Terzaghi, Meyerhof, Hansen and Vesic, and the artificial intelligence-based models such as MLR, Multi-gene GP, GP, and LGP. It was concluded that the new model proposed to determine the ultimate bearing capacity of the shallow foundation performs better than the previous methods regarding all the three indices of MAE, RMSE, and CC.

3. According to the sensitivity analysis conducted on the hybrid model, the internal friction angle of the soil ( $\varphi$ ) had the greatest impact on the model. The embedment depth of the foundation (D) also had a significant effect on the new model. The effective unit weight of the soil ( $\gamma$ ), width of the foundation (B), and length of the foundation (L) are the other important parameters in determining the ultimate bearing capacity by the proposed model, respectively.

4. A comprehensive ranking system was created based on four important indices: the correlation coefficient of predicted and measured values, mean and standard deviation of measured values to predicted ones, cumulative probability of 50% and 90% of observation to prediction ratios. The results of this ranking indicated the superiority of the M5'-GP hybrid approach in predicting the ultimate bearing capacity of shallow foundations from different aspects, in contrast to other methods.

#### References

Abu-Farsakh, M.Y. and Titi, H.H. (2004), "Assessment of direct cone penetration test methods for predicting the ultimate capacity of friction driven piles", J. Geotech. Geoenviron. Eng., 130(9), 935-944.

https://doi.org/10.1061/(ASCE)1090-0241(2004)130:9(935).

- Akbas, S.O. and Kulhawy, F.H. (2009), "Axial compression of footings in cohesionless soils. II: Bearing capacity", J. Geotech. Geoenviron. Eng., 135(11), 1575-1582.
- https://doi.org/10.1061/(ASCE)GT.1943-5606.0000136. Alavi, A.H., Gandomi, A.H., Mousavi, M. and Mollahasani, A.
- (2010), "High-precision modeling of uplift capacity of suction caissons using a hybrid computational method", *Geomech. Eng.*, 2(4), 253-280. https://doi.org/10.12989/gae.2010.2.4.253.
- Avval, Y.J. and Derakhshani, A. (2018), "New formulas for predicting liquefaction-induced lateral spreading: Model tree approach", *Bull. Eng. Geol. Environ.*, 1-13. https://doi.org/10.1007/s10064-018-1319-1.
- Banzhaf, W., Nordin, P., Keller, R.E. and Francone, F.D. (1998), Genetic Programming: An Introduction, Morgan Kaufmann, San Francisco, California, U.S.A.
- Bhattacharya, B. and Solomatine, D.P. (2005), "Neural networks and M5 model trees in modelling water level–discharge relationship", *Neurocomputing*, **63**, 381-396. https://doi.org/10.1016/j.neucom.2004.04.016.
- Bonakdar, L., Oumeraci, H. and Etemad-Shahidi, A. (2015), "Wave load formulae for prediction of wave-induced forces on a

slender pile within pile groups", *Coastal Eng.*, **102**, 49-68. https://doi.org/10.1016/j.coastaleng.2015.05.003.

- Briaud, J.L. and Gibbens, R. (1997), "Large scale load tests and data base of spread footings on sand", https://rosap.ntl.bts.gov/view/dot/4563.
- Briaud, J.L. and Gibbens, R. (1999), "Behavior of five large spread footings in sand", J. Geotech. Geoenviron. Eng., 125(9), 787-796.

https://doi.org/10.1061/(ASCE)1090-0241(1999)125:9(787)

- Cerato, A.B. and Lutenegger, A.J. (2007), "Scale effects of shallow foundation bearing capacity on granular material", J. Geotech. Geoenviron. Eng., 133(10), 1192-1202. https://doi.org/10.1061/(ASCE)1090-0241(2007)133:10(1192).
- Das, B.M. (2015), Principles of Foundation Engineering, Cengage Learning
- Derakhshani, A. (2017), "Estimating uplift capacity of suction caissons in soft clay: A hybrid computational approach based on model tree and GP", Ocean Eng., 146, 1-8. https://doi.org/10.1016/j.oceaneng.2017.09.025.
- Derakhshani, A. (2018), "On the uncertainty analysis of uplift capacity of suction caissons in clay based on the fuzzy sets theory", *Ocean Eng.*, **170**, 416-425. https://doi.org/10.1016/j.oceaneng.2018.10.045.
- Derakhshani, A. and Foruzan, A.H. (2019), "Predicting the principal strong ground motion parameters: A deep learning approach", *Appl. Soft Comput.*,
- https://doi.org/10.1016/j.asoc.2019.03.029
- Eastwood, W. (1951), "A comparison of the bearing power of footings on dry and inundated sand", *Struct. Engineer*, **29**(1), 1-11.
- Etemad-Shahidi, A. and Ghaemi, N. (2011), "Model tree approach for prediction of pile groups scour due to waves", *Ocean Eng.*, 38(13), 1522-1527.

https://doi.org/10.1016/j.oceaneng.2011.07.012

- Foye, K., Salgado, R. and Scott, B. (2006), "Assessment of variable uncertainties for reliability-based design of foundations", J. Geotech. Geoenviron. Eng., 132(9), 1197-1207. DOI: http://10.1061/(ASCE)1090-0241(2006)132:9(1197).
- Gandhi, G. (2003), "Study of bearing capacity factors developed from lab. Experiments on shallow footings on cohesionless soils", Ph.D. Thesis, Shri GS Institute of Technology and Science, Indore, India.
- Golder, H., Fellenius, W., Kogler, F., Meischeider, H., Krey, H. and Prandtl, L. (1941), "The ultimate bearing pressure of rectangular footings", *J. Inst. Civ. Eng.*, **17**(2), 161-174. https://doi.org/10.1680/ijoti.1941.13728.
- Hansen, J.B. (1970), "A revised and extended formula for bearing capacity", https://trid.trb.org/view/125129.
- Jafariavval, Y. and Derakhshani, A. (2019), "New formulae for capacity energy-based assessment of liquefaction triggering", *Mar. Georesour. Geotechnol.*, 1-9.

https://doi.org/10.1080/1064119X.2019.1566297.

Javadi, A. and Rezania, M. (2009), "Applications of artificial intelligence and data mining techniques in soil modeling", *Geomech Eng.* 1(1), 53-74.

https://doi.org/10.12989/gae.2009.1.1.053.

- Kalinli, A., Acar, M.C. and Gündüz, Z. (2011), "New approaches to determine the ultimate bearing capacity of shallow foundations based on artificial neural networks and ant colony optimization", *Eng. Geol.*, **117**(1), 29-38. https://doi.org/10.1016/j.enggeo.2010.10.002.
- Kaveh, A., Hamze-Ziabari, S.M. and Bakhshpoori, T. (2018), "Soft computing-based slope stability assessment: A comparative study", *Geomech. Eng.*, 14(3), 257-269. https://doi.org/10.12989/gae.2018.14.3.257.
- Koza, J.R. (1992), Genetic Programming: On the Programming of Computers by means of Natural Selection, MIT Press.

- Li, S., Yu, S., Shangguan, Z. and Wang, Z. (2016), "Estimating model parameters of rockfill materials based on genetic algorithm and strain measurements", *Geomech. Eng.*, **10**(1), 37-48. http://dx.doi.org/10.12989/gae.2016.10.1.037.
- Meyerhof, G.G. (1963), "Some recent research on the bearing capacity of foundations", *Can. Geotech. J.*, **1**(1), 16-26. https://doi.org/10.1139/cgj-2018-0433.
- Muhs, H. and Weiß, K. (1971), Untersuchung von Grenztragfähigkeit und Setzungsverhalten flachgegründeter Einzelfundamente in ungleichförmigen nichtbindigen Boden, Abschlußbericht des Forschungsauftrags Untersuchungen des Setzungsverhaltens und der Grenztragfähigkeit von flachgegründeten Fundamenten in grobkörnigen Sanden des Ministers für Wohnungsbau und öffentliche Arbeiten des Landes Nordrhein-Westfalen 1970, Ernst
- Muhs, H. and Weiß, K. (1973), "Inclined load tests on shallow strip footings", Proceedings of the 8th International Conference on Soil Mechanism and Foundation Engineering, Moscow, Russia, August.
- Muhs, H., Elmiger, R. and Weiß, K. (1969), Sohlreibung und Grenztragfähigkeit unter lotrecht und schräg belasteten Einzelfundamenten; mit 128 Bildern und 13 Zahlentafeln, Ernst.
- Padmini, D., Ilamparuthi, K. and Sudheer, K. (2008), "Ultimate bearing capacity prediction of shallow foundations on cohesionless soils using neurofuzzy models", *Comput. Geotech.*, 35(1), 33-46.
- https://doi.org/10.1016/j.compgeo.2007.03.001.
- Quinlan, J.R. (1992), "Learning with continuous classes", Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Hobart, Tamansia, November.
- Rostami, M.F., Sadrossadat, E., Ghorbani, B. and Kazemi, S.M. (2018), "New empirical formulations for indirect estimation of peak-confined compressive strength and strain of circular RC columns using LGP method", *Eng. Comput.*, **34**(4), 865-880. https://doi.org/10.1007/s00366-018-0577-7.
- Sadrossadat, E., Ghorbani, B., Hamooni, M. and Moradpoor Sheikhkanloo, M.H. (2018), "Numerical formulation of confined compressive strength and strain of circular reinforced concrete columns using gene expression programming approach", *Struct. Concrete*, **19**(3), 783-794. https://doi.org/10.1002/suco.201700131.
- Sadrossadat, E., Soltani, F., Mousavi, S.M., Marandi, S.M. and Alavi, A.H. (2013), "A new design equation for prediction of ultimate bearing capacity of shallow foundation on granular soils", J. Civ. Eng. Manage., 19(sup1), S78-S90. https://www.tandfonline.com/doi/abs/10.3846/13923730.2013.8 01902.
- Searson, D.P. (2015), GPTIPS 2: An Open-Source Software Platform for Symbolic Data Mining, in Handbook of Genetic Programming Applications, Springer International Publishing, Cham, Germany.
- Shahin, M.A., Maier, H.R. and Jaksa, M.B. (2004), "Data division for developing neural networks applied to geotechnical engineering", J. Comput. Civ. Eng., 18(2), 105-114. https://doi.org/10.1061/(ASCE)0887-3801(2004)18:2(105).
- Shahnazari, H. and Tutunchian, M.A. (2012), "Prediction of ultimate bearing capacity of shallow foundations on cohesionless soils: An evolutionary approach", *KSCE J. Civ. Eng.*, **16**(6), 950-957. https://doi.org/10.1007/s12205-012-1651-0.
- Smith, G.N. (1986), *Probability and Statistics in Civil Engineering*, Collins Professional and Technical Books.
- Subrahmanyam, G. (1967), "The effect of roughness of footings on bearing capacity", J. Int. Soc. Soil Mech. Found. Eng., 6, 33-45.
- Talebi, A. and Derakhshani, A. (2019), "Estimation of Pmultipliers for laterally loaded pile groups in clay and sand",

*Ships Offshore Struct.*, **14**(3), 229-237. https://doi.org/10.1080/17445302.2018.1495542.

- Terzaghi, K. (1943), *Theoretical Soil Mechanics*, Wiley Online Library
- Tsai, H.C., Tyan, Y.Y., Wu, Y.W. and Lin, Y.H. (2013), "Determining ultimate bearing capacity of shallow foundations using a genetic programming system", *Neural Comput. Appl.*, 23(7-8), 2073-2084. https://doi.org/10.1007/s00521-012-1150-8.
- Vesic, A. (1973), "Analysis of ultimate loads of shallow foundations: closure of discussion of original paper", J. Soil Mech. Found. Div., 99(sm1).
- Wang, Y.W. (1997), "IH: Inducing model trees for predicting continuous classes", *Proceedings of the European Conference* on Machine Learning, Prague, Czech Republic, April.
- Weiß, K. (1970), Der Einfluß der Fundamentform auf die Grenztragfähigkeit flachgegründeter Fundamente, Untersuchungen ausgef.... von Klaus Weiß: mit 14 Zahlentaf, Ernst.
- Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J. (2016), *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann.

CC

# Appendix A: Compiled database for developing the predictive model of the ultimate bearing capacity

T 11 A 1		1 .
Table A1	Training	data

Table A1	Training	data				0.0585	0.029	5.95	17.1	42.5
Foundation width (B)	Foundation depth (D)	Aspect ratio (L/B)	Specific gravity (y)	Internal friction	Ultimate bearing	0.0585	0.058	5.95	17.1	42.5
0.6	0.3	2	9.85	34.9	270	0.094	0.047	6	15.7	34
0.6	0.3	2	10.2	37.7	570	0.094	0.047	6	16.1	37
0.6	0	2	10.85	44.8	860	0.094	0.094	6	16.1	37
0.6	0.3	2	10.85	44.8	1760	0.094	0.047	6	16.5	39.5
0.5	0	1	10.2	37.7	165	0.094	0.094	6	16.5	39.5
0.5	0	2	10.2	37.7	203	0.094	0.094	6	16.8	41.5
0.5	0	2	10.2	37.7	195	0.094	0.047	6	17.1	42.5
0.5	0	3	10.2	37.7	214	0.094	0.094	6	17.1	42.5
0.52	0	3.85	10.2	37.7	186	0.152	0.075	5.95	15.7	34
0.5	0.3	1	10.2	37.7	681	0.152	0.15	5.95	15.7	34
0.5	0.3	2	10.2	37.7	530	0.152	0.15	5.95	16.1	37
0.5	0.3	3	10.2	37.7	402	0.152	0.075	5.95	16.5	39.5
0.52	0.3	3.85	10.2	37.7	413	0.152	0.15	5.95	16.5	39.5
0.5	0	1	11.7	37	111	0.152	0.075	5.95	16.8	41.5
0.5	0	1	11.7	37	132	0.152	0.075	5.95	17.1	42.5
0.5	0	2	11.7	37	143	0.152	0.15	5.95	17.1	42.5
0.5	0.013	1	11.7	37	137	0.094	0.047	1	15.7	34
0.5	0.029	4	11.7	37	109	0.094	0.094	1	15.7	34
0.5	0.127	4	11.7	37	187	0.094	0.094	1	16.1	37
0.5	0.3	1	11.7	37	406	0.094	0.047	1	16.5	39.5
0.5	0.3	1	11.7	37	446	0.094	0.094	1	16.5	39.5
0.5	0.5	2	11.7	37	565	0.094	0.047	1	16.8	41.5
0.5	0.5	4	11.7	37	425	0.094	0.047	1	17.1	42.5
0.5	0	1	12.41	44	782	0.094	0.094	1	17.1	42.5
0.5	0.3	1	12.41	44	1940	0.152	0.075	1	15.7	34
0.5	0.5	2	12.41	44	2847	0.152	0.15	1	15.7	34
0.5	0.5	4	12.41	44	2033	0.152	0.15	1	16.1	37
0.5	0.49	4	12.27	42	1492	0.152	0.075	1	16.5	39.5
0.5	0	2	11.77	37	123	0.152	0.15	1	16.5	39.5
0.5	0	2	11.77	37	134	0.152	0.075	1	16.8	41.5
0.5	0.3	1	11.77	37	370	0.152	0.075	1	17.1	42.5
0.5	0.5	2	11.77	37	464	0.152	0.15	1	17.1	42.5
0.5	0	4	12	40	461	0.08	0	1	17.2	42.8
0.5	0.5	4	12	40	1140	0.15	0	1	17.2	42.8
1	0.2	3	11.97	39	710	0.05	0	1	17.2	42.8
0.991	0.711	1	15.8	32	1773.7	0.1	0	1	17.1	42.8
2.489	0.762	1	15.8	32	1158	0.15	0	1	17.1	42.8
1.492	0.762	1	15.8	32	1540	0.25	0	1	17.1	42.8
3.016	0.889	1	15.8	32	1161.2	0.3	0	1	17.1	42.8
0.0585	0.029	5.95	15.7	34	58.5	0.03	0	1	15.89	42
0.0585	0.029	5.95	16.1	37	82.5	0.05	0	1	15.89	42
0.0585	0.058	5.95	16.1	37	98.9	0.06	0	1	13.2	32
0.0585	0.029	5.95	16.5	39.5	121.5	0.06	0	1	15.4	42
0.0585	0.058	5.95	16.5	39.5	142.9	1	0	1	19.5	38.75
	-	-			<u></u>	1	0	1	19.5	38.75

Table A1 Continued Foundation Foundation

depth (D)

width (B)

Aspect ratio (L/B)

gravity  $(\gamma)$ 

Specific Internal friction Ultimate bearing

capacity (qu)

180.5

211

74.4

104.8

127.5 155.8

185.6

244.6

235.6

279.6

98.2

122.3

176.4

211.2

254.5

285.3

335.3

400.6

67.7

90.5

131.5

147.8

191.6 196.8

228.8

295.6 91.2

124.4

182.4 201.2

264.5

276.3

325.3

423.6

133

246

109

152

214

333

404

52

95

14

106

377

335

angle (ø)

Table A2 Continued

Table A1 Continued

Foundation width (B)	Foundation depth (D)	Aspect ratio (L/B)	Specific gravity (y)	Internal friction angle (\u00fc)	Ultimate bearing capacity (q <sub>u</sub> )	Foundation width (B)	Foundation depth (D)	Aspect ratio (L/B)	Specific gravity (y)	Internal friction angle (\u00f6)	Ultimate bearing capacity (q <sub>11</sub> )
1	0	1	19.5	38.75	305	0.5	0	1	10.2	37.7	154
1	0	1	19.5	38.75	400	0.5	0.3	2	10.2	37.7	542
1	0	1	19.5	38.75	296	0.5	0.3	4	11.7	37	322
1	0	1	19.5	38.75	390	0.5	0	4	12.41	44	797
0.71	0	1	19.5	38.75	438.4	0.5	0.5	2	12.41	44	2266
1	0	1	16.8	40.55	773	1	0	3	11.93	40	630
1	0	1	16.8	40.55	685	3.004	0.762	1	15.8	32	1019.4
1	0	1	16.8	40.55	560	0.0585	0.058	5.95	15.7	34	70.91
1	0	1	16.8	40.55	598	0.0585	0.029	5.95	16.8	41.5	157.5
1	0	1	16.8	40.55	584	0.094	0.094	6	15.7	34	91.5
1	0	1	16.8	40.55	716	0.094	0.047	6	16.8	41.5	206.8
1	0	1	16.8	40.55	922	0.152	0.075	5.95	16.1	37	143.3
1	0	1	16.8	40.55	659	0.152	0.15	5.95	16.8	41.5	342.5
1	0	1	16.8	40.55	640	0.094	0.047	1	16.1	37	98.8
1	0	1	16.8	40.55	626	0.094	0.094	1	16.8	41.5	253.6
1	0	1	16.8	40.55	927	0.152	0.075	1	16.1	37	135.2
0.7	0	1	16.8	40.55	612.2	0.152	0.15	1	16.8	41.5	361.5
0.75	0	1	20.8	44.95	856.9	0.08	0	1	17.1	42.8	130
0.45	0	1	20.8	44.95	953.1	0.2	0	1	17.1	42.8	266
0.45	0	1	20.8	44.95	454.3	0.04	0	1	15.89	42	92
0.3	0	1	20.8	45.7	422.2	0.06	0	1	14.8	42	72
0.3	0	1	20.8	45.7	900	1	0	1	19.5	38.75	368
0.3	0	1	20.8	45.7	1688.9	1	0	1	19.5	38.75	435
0.91	0	1	14.6	31.95	324.8	1	0	1	16.8	40.55	500
0.61	0	1	14.6	31.95	94.1	1	0	1	16.8	40.55	726
0.61	0	1	14.6	31.95	322.5	1	0	1	16.8	40.55	825
0.61	0	1	19	37	258	0.75	0	1	20.8	44.95	1020.4
0.8	0	1	17.1	39.75	348.4	0.3	0	1	20.8	45.7	600
0.63	0	1	17.1	39.75	365.3	0.61	0	1	14.6	31.95	196.2
0.46	0	1	17.1	39.75	104	0.46	0	1	14.6	31.95	259.9
0.31	0	1	15.8	37.9	478.7	1.2	0	1	20.4	41	978.5
1.2	0	1	20.4	41	1129.9	0.3	0	1	20.4	41	522.2
0.3	0	1	20.4	41	1277.8	3	0.76	1	15.5	35.3	1144.4
0.3	0	1	20.4	41	811.1						
0.3	0	1	20.4	41	333.3						
0.3	0	1	20.4	41	233.3						
0.76	0	1	16.2	40.8	744.4						
0.31	0	1	16.2	40.8	260.1						
0.31	0	1	16.2	40.8	468.3						
1	0.71	1	15.5	35.3	1550						

# Table A2 Testing data

0.76

0.76

0.89

1

1

1

1.5

2.5

3

Foundation	Foundation	Aspect	Specific	Internal friction	Ultimate bearing
width (B)	depth (D)	ratio (L/B)	gravity (γ)	angle (ø)	capacity (q <sub>u</sub> )
0.6	0	2	10.2	37.7	200

15.5

15.5

15.5

35.3

35.3

35.3

1355.6

1152

1011.1