



Swarm-based analysis through social behavior of grey wolf optimization and genetic programming to predict friction capacity of driven piles

Hossein Moayedi^{1,2} · Mohammed Abdullahi Mu'azu³ · Loke Kok Foong⁴

Received: 16 October 2019 / Accepted: 30 October 2019 / Published online: 9 November 2019
© Springer-Verlag London Ltd., part of Springer Nature 2019

Abstract

The advantage of new data mining-based solutions, and more recently, optimization algorithms (i.e., basically swarm-based solutions) have enhanced traditional models of engineering structural analysis. This paper investigates social behavior of Grey Wolf Optimization (GWO) in improving the neural assessment of friction capacity (f_s) of concrete driven pile systems. Besides, the genetic programming (GP) algorithm was also proposed to have comparison with the proposed GWO prediction outputs. To achieve this goal, four f_s influential factors of pile length (m), pile diameter (cm), effective vertical stress (S_v), and undrained shear strength (S_u) are considered for preparing the required dataset. A swarm size-based sensitivity analysis is then carried out to use the best-fitted structures (i.e., more convergency in the final output) of each ensemble. The results of the best prediction network from both above-mentioned sensitivity analyses were compared. The results show that both GWO and GP models presented excellent performance. The findings of neural networks varied based on the number of neurons in a single hidden layer and of course the level of its complexity. Based on R^2 and RMSE, values of (0.9537 and 9.372) and (0.8963 and 7.045) are determined, for the training and testing datasets of MLP-based solution, respectively. On the contrary, for the GP and GWO-MLP proposed predictive models, the R^2 of (0.9783 and 0.982) and (0.913 and 0.892) were found for the training and testing datasets. This proves the better performance of GWO when combined with MLP in predicting engineering solutions comparing to conventional MLP or GP-based combinations.

Keywords Hybrid model · Concrete driven piles · Genetic programming · Multilayer perceptron

1 Introduction

Structural members such as pile foundations can transfer heavy loads of superstructure safely through compressible or weak upper soil layer to underlying stronger bedrock or competent soils. Concrete driven piles can be installed into cohesive soils using impact hammers. Different parameters were affecting the final results of these pile's bearing capacity [1]. Piles are of different types that can be classified based on their installation method. One of the commonly used forms of piles is driven piles that consist of different types such as driven piles of precast concrete, steel or timber. The installation mechanism of this kind of piles is usually by hammering a pile to the ground using impact of a falling weight [2]. The load-carrying capacity of a pile consists of two factors: the factor of friction, named shaft friction or skin friction; and the other is capable of end bearing at the base of the pile toe [3]. Numerous theoretical, experimental, and numerical methods have been proposed for estimating

✉ Hossein Moayedi
hossein.moayedi@tdtu.edu.vn

✉ Loke Kok Foong
lokekokfoong@duytan.edu.vn

¹ Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam

² Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam

³ Civil Engineering Department, College of Engineering, University of Hafr Al-Batin, Al Jamiah District, Eastern Province 39524, Kingdom of Saudi Arabia

⁴ Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

friction capacity of piles under axial compressive load but due to complexity of soil–pile interaction, there is uncertainty about their reliability and accuracy [4]. Therefore, proposing a method able to predict friction capacity of piles is of great importance in geotechnical engineering. In the literature, there are different methods for specification of frictional capacity as effective stress-based approach, total stress-based approach, and combination of two mentioned approaches as a more effective model [5–7]. Nevertheless, developing a theoretical along with the statistical model is difficult and complicated task due to intricate behavior of pile under axial loading and nonlinear interaction between soil and pile in various geotechnical conditions and soil properties as well as rate of loading and effect of ground-level fluctuations during time [8].

Artificial neural network (ANN) is known as one of the most widely utilized AI approaches and was introduced first in the early 1960s. ANN method has been used for a wide variety of civil engineering issues. In comparison with common empirical methods, the AI techniques are known to be proper in estimation. One of the most complex problems in the field of geotechnical engineering is consistency and accurate estimation of pile bearing capacity, especially piles driven into cohesive soils [9–12]. There have been many investigations on the good use of machine learning or other artificial computation techniques in estimating of engineering problems and more particularly friction capacity of the driven piles (e.g., García-Balboa et al. [13]; Samui [14]; Schneider et al. [15]; Schneider et al. [15]). Scholars have done various investigations for the skin friction for piles. For estimating this parameter, Goh (1995) has utilized backpropagation neural networks (BPNN) and showed that in comparison to Engineering News formula, Janbu formula, and Hiley formula, the final load capacity of piles for ANN has more fitting efficiency. Also, in order to predict bearing capacity of piles, scholars have used a hybrid genetic algorithm (GA)-based ANN and proved that GA-based ANN has better performance than the usual ANN models. Samui (2008) utilized support vector machines (SVM) upon the same database and demonstrated the better performance of this approach compared to the ANN model. ANN algorithm has lower generalization because of the gaining local minima in training and requires iterative learning stages for achieving better learning efficiency while SVM has more appropriate generalization in comparison with the ANN algorithm.

Samui [14] investigated the friction capacity of the driven piles by utilizing a well-known classification method of support vector machine. Both, methods, the ANN and SVM, have shown good performances and provided excellent achievements. Similar outputs

were obtained by Schneider et al. [15], but on the sandy soils. Several techniques of calculating ultimate bearing capacity (i.e., considering both of the skin friction and end bearing capacity) of driven piles were mentioned by Xia and Wang [16]. In this regard, machine learning was selected as one of the most accurate methods for calculating the bearing capacity of the driven piles. Tran et al. [17] investigated finding a reliable prediction of skin friction in driven piles. They have installed extensometer instrumentation at top and bottom of the pile where the stress differences along the pile could be measured during the pile load test. In a more advanced solution, Samui [18] applied a technique called MARS (i.e., abbreviated from multivariate adaptive regression spline) to evaluate the same design parameter. A year later, Alkroosh and Nikraz [4] proposed a new artificial intelligent computing technique that could approximate the axial bearing capacity of the driven piles in clayey soils. Mosallanezhad and Moayedi [19] performed a similar study on cohesionless soils. It was found that the results of gene expression programming technique (GEP) are reliable and accurate to be used in practical engineering projects. Samui [20] then presented a model based on the MARS approach to estimate the ultimate bearing capacity of driven piles installed in sands. The results provided an excellent result showing that MARS is very reliable and can be used in geotechnical engineering projects. Moayedi and Hayati [21] provided a series of design solution charts for the driven piles installed in cohesive soils. They have selected several artificial intelligence-based techniques to estimate the ultimate bearing capacity of the driven piles. The best model was selected according to the well-known statistical indexes. The result of the best predictive networks was presented as a series of design solution charts.

After reviewing almost all published works, it should be noted that the algorithm of artificial neural network (ANN) combined with genetic programming (GP) model that is presented in this study has not been proposed for this research area. Novelty of this approach is due to lack of investigation in the use of hybrid GP-based learning systems for estimating the friction capacity of the driven piles and its key parameters and comparison of results with conventional ANN method. Therefore, the main objective of this study was to predict the frictional component of bearing capacity of driven concrete piles by using a novel method. This research is not only for determining friction capacity of driven piles as a particular issue, but it also can help improve AI knowledge for this kind of problem. Also, proposed method can be used for piles with negligible end bearing capacity or floating piles (i.e., pile embedded in soil with no pronounced bearing stratum at the tip). In this work, based on the information collected from field investigation analysis

method, 36 different GP models were investigated to find their best performances. In the following, the results of the GP models were compared with the results of an optimized ANN feedforward learning system.

2 Methodology and established database

In this section, the evaluation of the obtained results in estimating the friction capacity of the driven piles is discussed. Two of the well-known statistical indices that are used in this study are coefficient of determination (R^2) and root-mean-square error (RMSE) as suggested by previous studies (e.g., Mosallanezhad and Moayedi [19], Moayedi and Armaghani [2], and Mosallanezhad and Moayedi [19]).

2.1 Multi-layer perceptron

Artificial neural network (ANN) is the main branch of artificial intelligence (AI) and machine learning. This neural structure can find patterns of complex relationships between input and output layers. One of the most acknowledged advantages of ANN is that they can learn from observing datasets. ANNs are newly developed information processing tools that are based on the present understanding of the human neural network [22–27]. ANN works as a computational system consisting of input and output as well as one or more hidden layers having units that can transform the input into something that the output layer can use. They have been widely used for modeling problems such as medicine, finance, and complex engineering issues. The main advantage of ANNs is their ability to approximate the untrained data. In an ANN architecture, processing elements are linked together by variable weights. Figure 1 shows the overall

structure of a standard paradigm of ANNs, namely Multi-Layer Perceptron (MLP). An MLP has a minimum of three layers, such as one input, one or more hidden, and a single output layer(s). Every layer contains several computational elements, called neurons. The interconnection weights are the equivalents of synapses in a biological neural system. During the ANN performance, the weights are estimated in the initial iteration, and as this process repeats, they are adjusted to minimize the error. More specifically, the performance error is indicated by comparing the system outputs and target data. Then, the system back-propagates the calculated error in an adverse direction (from right to left in Fig. 1) to estimate more appropriate weights.

Finally, each neuron calculates its output by the following equation:

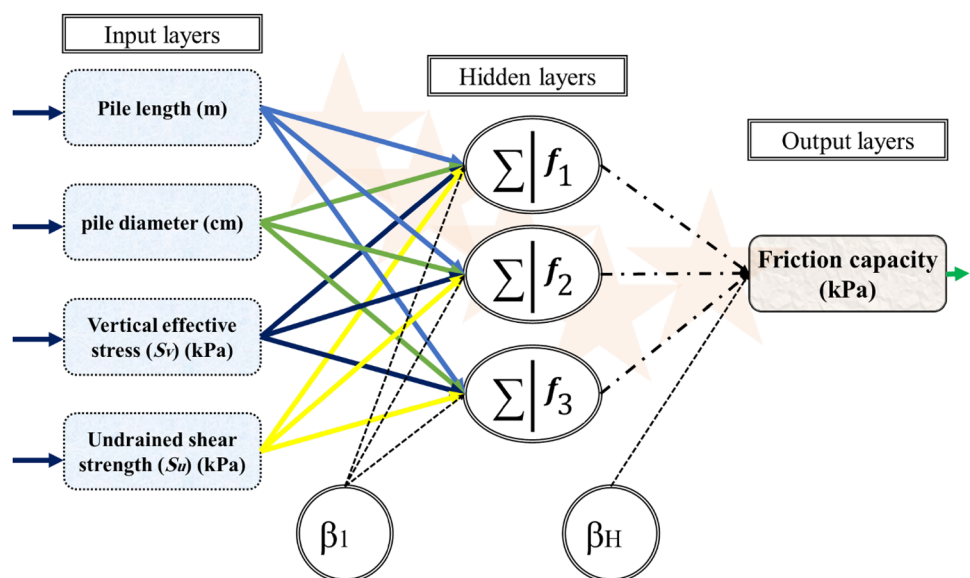
$$y = f(IW + b), \tag{1}$$

where terms I , W and b stand for the input, weight and the bias of the network, respectively. Also, f indicates the activation function, which is considered to be Tansig in this study. Equation 2 defines the Tansig function:

$$\text{Tansig}(x) = \frac{2}{1 + e^{-2x}} - 1, \tag{2}$$

In the present paper, for the friction capacity approximation, we have used a method of feed-forward back-propagation (FFBP) as well as the training algorithm of Levenberg–Marquardt (LM). As was explained before, in the model of FFBP, after computing the system efficiency, initial biases and also weights are adjusted. It should be noted that the Levenberg–Marquardt training algorithm is employed because of its more appropriate performance compared to the usual gradient descent (GD) techniques.

Fig. 1 Typical flowchart for the MLP model to predict f_s



2.2 Grey wolf optimization

To perform an efficient optimization, grey wolf algorithm (GWO) was introduced as a novel meta-heuristic nature-inspired method [28]. The GWO has been first expanded via mimicking the exploring conduct of grey wolves accomplishing in teams of 5–12 singles, which are at the upside of the food chain [29]. In nature, grey wolves commonly pursue a collective hierarchy strictly. In the leader group of wolves, a couple of females and males exist called alpha (α). They make the major decisions during hunting and other behaviors. Beta (β) wolves are at the subsequent level, assisting alpha wolves in making decisions.

Meanwhile, they have to obey the alpha. The β wolves may be female, and their role is to adjust the flock. To substitute the alpha (while they get older or even die), they are the most appropriate candidates. The subsequent level of the flock is named delta (δ). They do the hunting and have the role of scouts, sentinels, etc. The last group of individuals is named omega (ω), which is also known as the weakest level. They have the role of babysitters. Without omega wolves, some fights can be seen in the flock. The grey wolves do hunt, and this is their major social conduct.

Muro et al. [30] introduced three stages of GW hunting manner, which are (1) recognizing, following and nearing the prey; (2) circling the prey, and (3) rushing the prey. These two different social conducts have been considered in the algorithm of GWO [29]. The mathematical modeling of encompassing can be shown as [28] follows:

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)|, \quad (3)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D}, \quad (4)$$

in which \vec{A} and \vec{C} are coefficient vectors. Also, \vec{X}_p shows the prey location and \vec{X} stands for the location of wolves. \vec{D} is a vector and stands to specify a novel location of GWs. The term t stands for the time of iteration. Moreover, we can model \vec{A} and \vec{C} as follows [28]:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a}, \quad (5)$$

$$\vec{C} = 2 \cdot \vec{r}_2; \quad (6)$$

in the above relations, \vec{a} shows the vector set on the iteration, which varies in the range of [0, 2], and \vec{r}_1 and \vec{r}_2 show random vectors between [0, 1].

The α conduct the hunting tactic, β as well as δ attend to this action sporadically. It is considered that α , β and δ have better knowledge about the placement of prey during the model of hunting task. Therefore, the optimized solutions in the case of the three locations may be updated

and the rest of the herd register and follow their locations accordingly.

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}| \quad (7)$$

$$\vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}| \quad (8)$$

$$\vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}| \quad (9)$$

$$\vec{X}_1 = \cdot \vec{X}_\alpha - A_1 \cdot \vec{D}_\alpha \quad (10)$$

$$\vec{X}_2 = \cdot \vec{X}_\beta - A_2 \cdot \vec{D}_\beta \quad (11)$$

$$\vec{X}_3 = \cdot \vec{X}_\delta - A_3 \cdot \vec{D}_\delta \quad (12)$$

$$\vec{X}_{(t+1)} = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (13)$$

The GWs always begin the attack when the prey stops. The vector of A stands for a random amount in the range of $|-2a \cdot 2a|$. The $|A| < 1$ means the GWs attack and $|A| > 1$ means forcing the grey wolves to seek a more appropriate solution [31]. The flowchart of the GWO can be observed in Fig. 2.

2.3 Genetic programming

In 1992, Koza [32] introduced an extension of genetic algorithm (GA), called genetic programming (GP), as another type of evolutionary algorithms (EAs). Evolutionary algorithms are searching for approaches to find solution to multi-variable functions based on biological evolution in nature. Both GA and GP are inspired and formed by Darwin's theory of evolution by natural selection except that GA and GP deal with evolution of string of numbers or programs instead of living creatures. Program in GP is a tree structure consisting of a root node, functional nodes and terminal nodes which are connected by links. Figure 3 shows an example of a tree-like expression of a program $\left(\sqrt{\frac{x_1}{x_2}} + 5\right)$ [33, 34].

Genetic programming procedure starts with the creation of first generation by random selection of functions and terminals from sets that represents the nature of problem. Then specified fitness criteria will be checked for each tree and value of fitness will be assigned to them. Afterward, the next generation of population, which gives better solution, will be created by breeding programs that have higher values of fitness together by crossover and mutation operators. Besides, the third operator (reproduction) copies a portion of population to next-generation based on Roulette Wheel selection without any change. This method of selection causes easier selection of those trees having higher values of fitness. In the crossover operation, after selection of two trees, one random

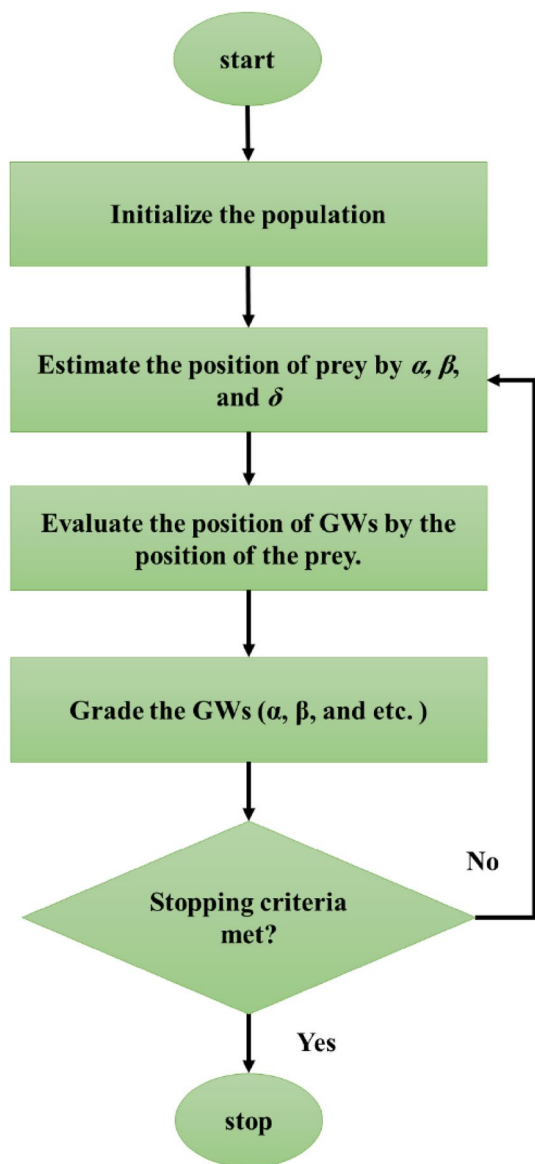


Fig. 2 The flowchart of the GWO algorithm (after Mirjalili et al. [28])

node will be selected from each tree. Then all the links and nodes under that node will be changed to create new trees for next generation. Figure 4 shows an example of crossover procedure [33].

Mutation of a tree is a procedure that a random node of a randomly selected tree will be changed with any random node that has the same number of function nodes, terminal nodes, and output links (arguments). A function node will be changed with a function node, and a terminal node is changeable with another terminal node. A typical illustration of mutation operation is depicted in Fig. 5 [33].

Finally, after the creation of new generation using three operators (reproduction, crossover, and mutation) this

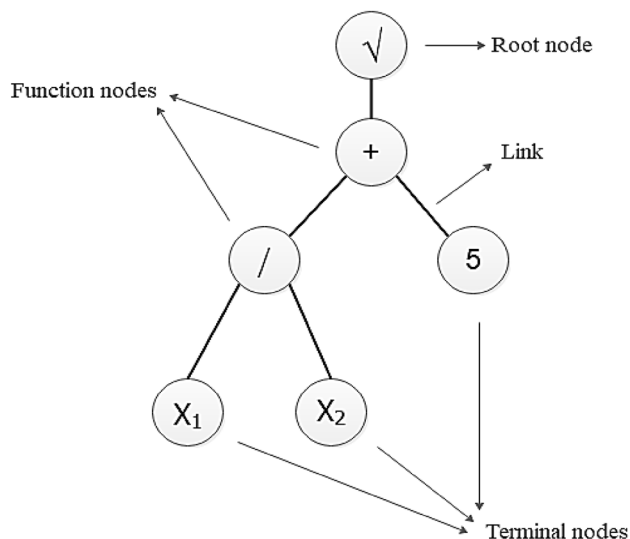


Fig. 3 Tree expression of a genetic program (after [35]) $\left(\sqrt{\left(\frac{X_1}{X_2}\right)} + 5\right)$

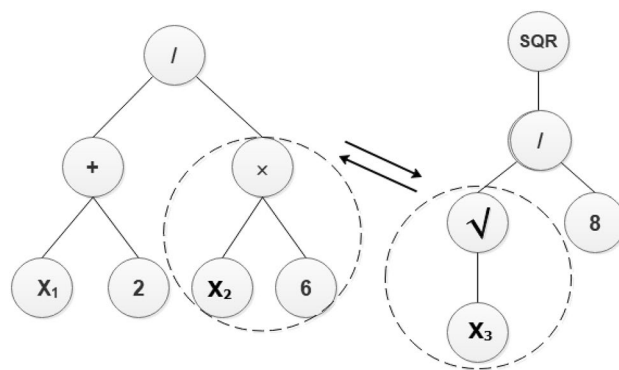


Fig. 4 Typical crossover operation of two trees making new generation by changing random nodes (After [35])

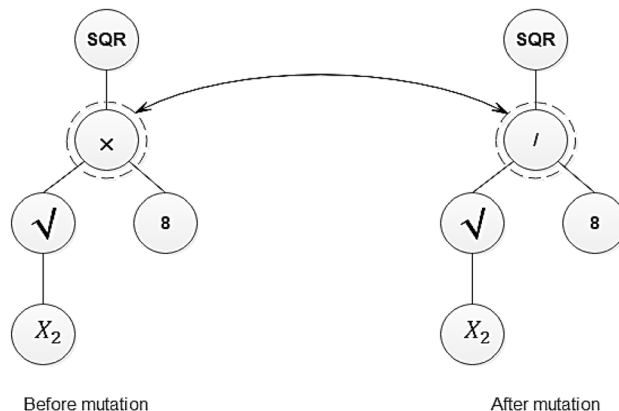


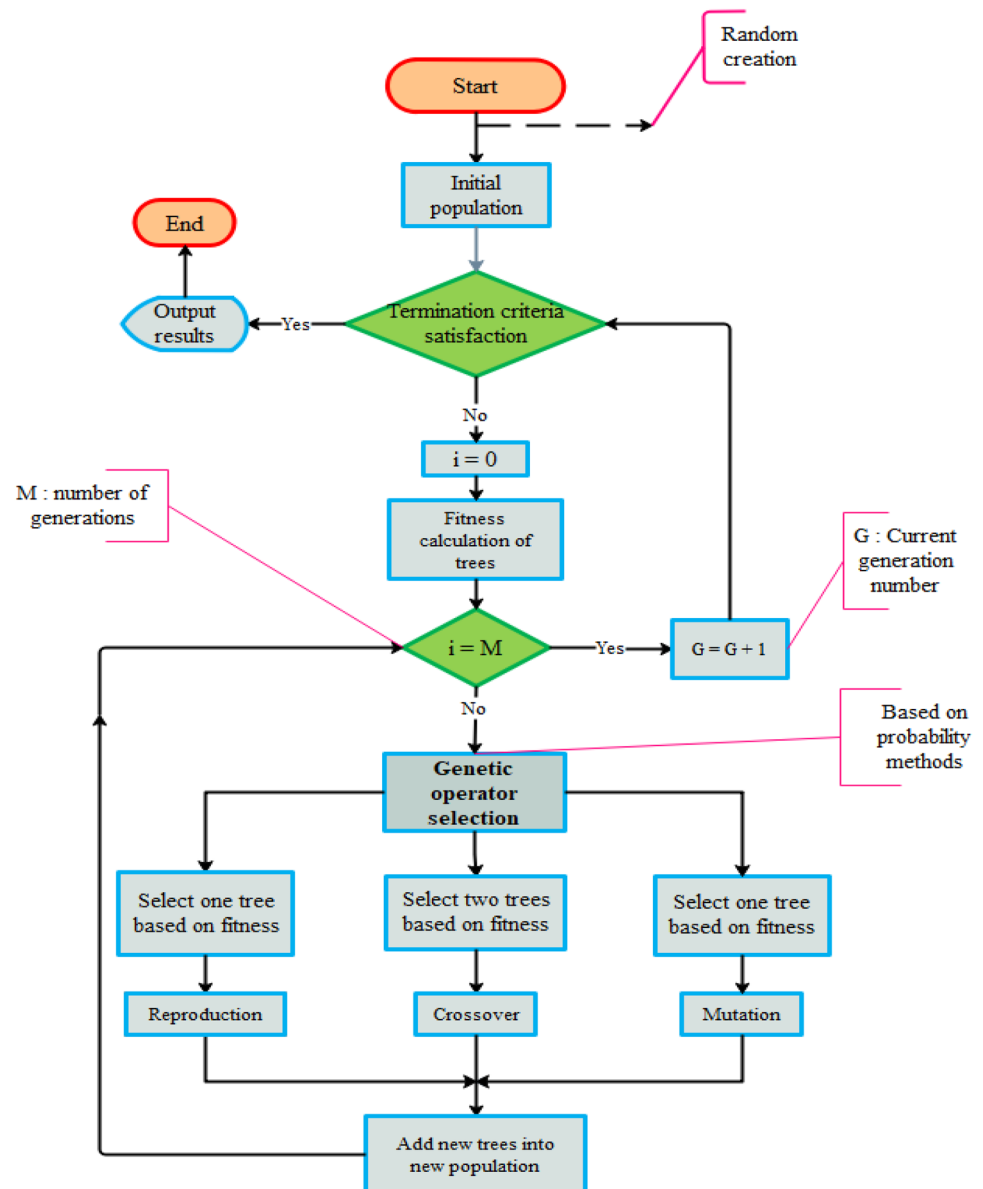
Fig. 5 Mutation procedure of two trees (Chromosomes) [35]

procedure will repeat until reaching number of generations to that of defined by user. The typical procedure of genetic programming is shown in the flowchart Fig. 6.

Genetic programming (GP) is a type of programming that uses the concepts and some of the terminology of biological evolution to help to solve complex problems. In this method the most effective and applicable programs of total possible programs will survive and compete or cross-breed with other programs to approach a suitable and correct solution for problem. Genetic programming is an approach that is suitable for problems having various fluctuating variables like those related to AI. Genetic programming can be expressed as an extension of the genetic algorithm able to test and select the best choice among a set of results. Genetic programming is more advanced and makes the program or

function the testing unit. Generally, selection methods of the most effective programs are cross-breeding and the tournament or competition approach. This model was first introduced by Cramer [36] and then extensively improved by Koza [32]. As stated in various research, the GP has a high capability to establish non-linear functions between the inputs and outputs of a system (e.g., Baziar and Azizkandi [37], Pandey et al. [38], and Fatehnia and Amirinia [34]). This method can optimize an objective function using synthesizing Darwin's theory of natural selection and the survival of the fittest [39]. In a GP process, computer programs that are formed by mathematical functions and input variables or constant values (terminal set) are applied to approximate and solve the problem. Note that the main superiority of GP over the ANN is that it can present a simple

Fig. 6 Flowchart diagram of typical genetic programming process (after [34])



mathematical form of a non-linear relationship between inputs and the respective outputs. When the GP approximation process commences, chromosomes are constructed like the initial population of the model, which contains a collection of user-defined terminals and functions. However, the upper mentioned functions and terminals have been chosen, selected randomly, and arranged in the form of a tree to build the computer model. The overall procedure of GP performance is shown in Fig. 7. According to this figure, there are three types of nodes, which are characterized by one root node and functional and terminal nodes. The functions that indicate the essence of the issue or data are first specified in the GP analysis process. A fitness index is defined for every individual, based on its status relative to the rest of the population’s performance. In the following part of the GP mechanism, existing populations can be replaced with new ones, giving rise to an evolutionary process. This procedure is maintained until a tolerable error, and a maximum number of generations are achieved.

2.3.1 Established database and statistical analysis

A total of 65 in situ tests were collected as for the datasets. The structure of ANN and genetic programming, and GWO-MLP are discussed. The data, used to develop the GP, GWO and ANN models in this study are obtained from Goh [40]. This database, shown graphically in Fig. 8, comprises an extensive real-scale driven pile load experiment. In addition to the result of pile load tests, the properties of the soils adjacent to the piles are also collected. Therefore, the training and testing datasets are collected from extensive in situ experiments. To train both models, a dataset including 52 field experiments is utilized for learning process, and 13 tests are nominated for the testing dataset. According to the developed algorithm, the selection of the test dataset was conducted randomly from 20% of the total number of records in the dataset.

Histogram of the input data layers and output are illustrated in Figs. 9 and 10, respectively.

3 Results and discussion

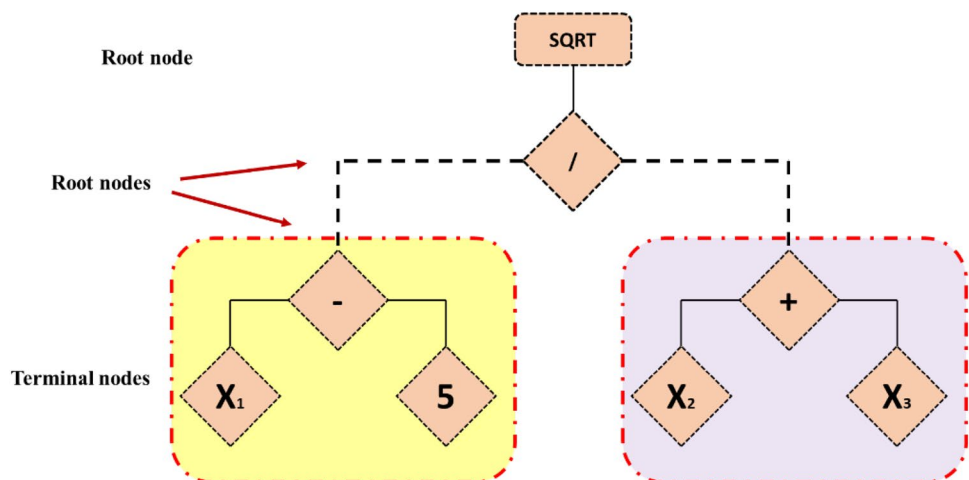
The potential of using artificial computational intelligence and artificial neural network (ANN) models in estimating the engineering properties of structures are well established [41–43]. The present study intends to evaluate the friction capacity of the driven piles installed in cohesive soils by using two intelligent techniques. An MLP neural network and a GP mode were applied to approximate the friction capacity. To do so, the critical parameters were considered (e.g., four practical factors were considered). Similar to previous research, the stock dataset was divided into two parts, randomly, to train and test the networks that have a ratio of 70% and 30%, respectively (e.g., Moayedi and Hayati [21, 44]). Also, for each model, the performance was measured by the index of the root mean square error (RMSE), as well as the accommodation was calculated by the determination coefficient (R^2). These indices were used to measure the error of the performance, as well as for measuring the correlation between the predicted and observed f_s . Equations 1–2 define the formulation of the R^2 and RMSE indices:

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_{i\text{ predicted}} - Y_{i\text{ observed}})^2}{\sum_{i=1}^N (Y_{i\text{ observed}} - \bar{Y}_{\text{observed}})^2}, \tag{14}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N [(Y_{i\text{ observed}} - Y_{i\text{ predicted}})]^2}, \tag{15}$$

where $Y_{i\text{ predicted}}$ and $Y_{i\text{ observed}}$ represent the predicted and actual f_s , and the number of instances is shown by N . Also, $\bar{Y}_{\text{observed}}$ denotes the average of the actual f_s .

Fig. 7 Common GP tree presented for the function $[(X_1 - 5) / (X_2 + X_3)]^2$



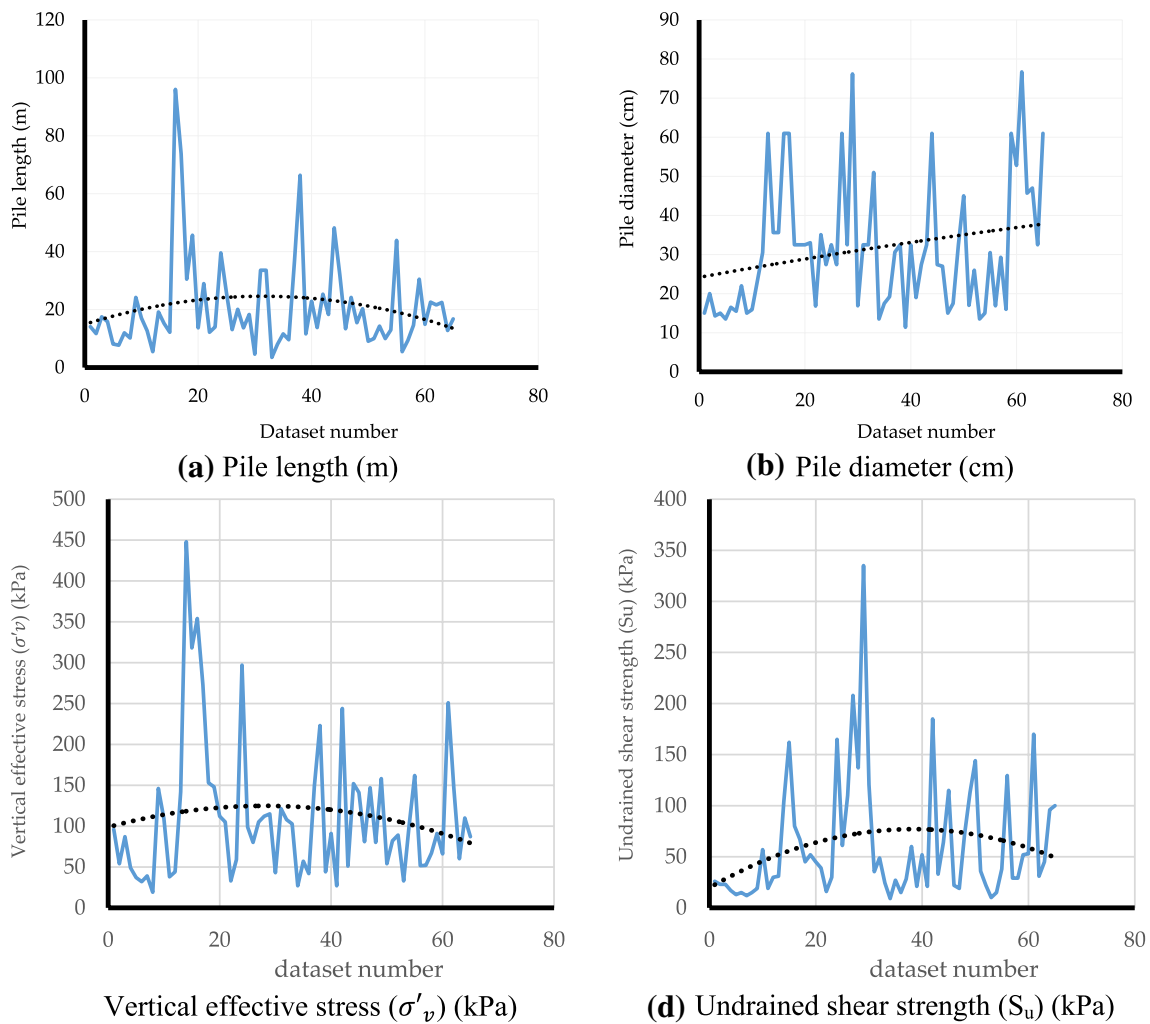


Fig. 8 Graphical view of the large variation of the input and output datasets

3.1 Optimizing multilayer perceptron network

Before GWO and GP models, we need to find the most proper MLP structure through a series of trial and error process. In this regard, an MLP network was tested with different numbers of neurons in its single hidden layer. The network which reported the lowest RMSE was selected as the optimum model. This method is well used in other studies (e.g., [26, 45–47]). The results of these trial and error processes are depicted in Fig. 11. Note that every structure was calculated for six iterations, and, as a general deduction, the MLP network with at least five neurons in its hidden layer can indicate an accruable efficiency for modeling the problem. Considering the reported RMSE and R^2 values, five neurons had the best prediction, and after that, there was a steady trend for both the RMSE and R^2 indices graphs, which means that this is selected as the appropriate neurons number in the hidden layer.

3.2 The hybrid model predicting friction capacity

Finding the proper architecture of the MLP network is the first step before the utilization of hybrid intelligence solutions. In the second step, both of the GWO and GP structure need to be optimized. For the subject of GP optimization, a wide trial and error approach, including 36 various GP models, was performed to find an appropriate GP structure for estimating the friction capacity of the driven piles. The efficiency of the GP method is recognized for distinct amounts of population sizes, numbers of productions and selection tournament sizes. For each one of the three mentioned parameters, 12 different values were considered. For instance, the values of 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 750 and 1000 were considered for understanding the effect of population sizes, sizes of generations, and tournament numbers. The performance of the above-mentioned trial and error process was evaluated using the

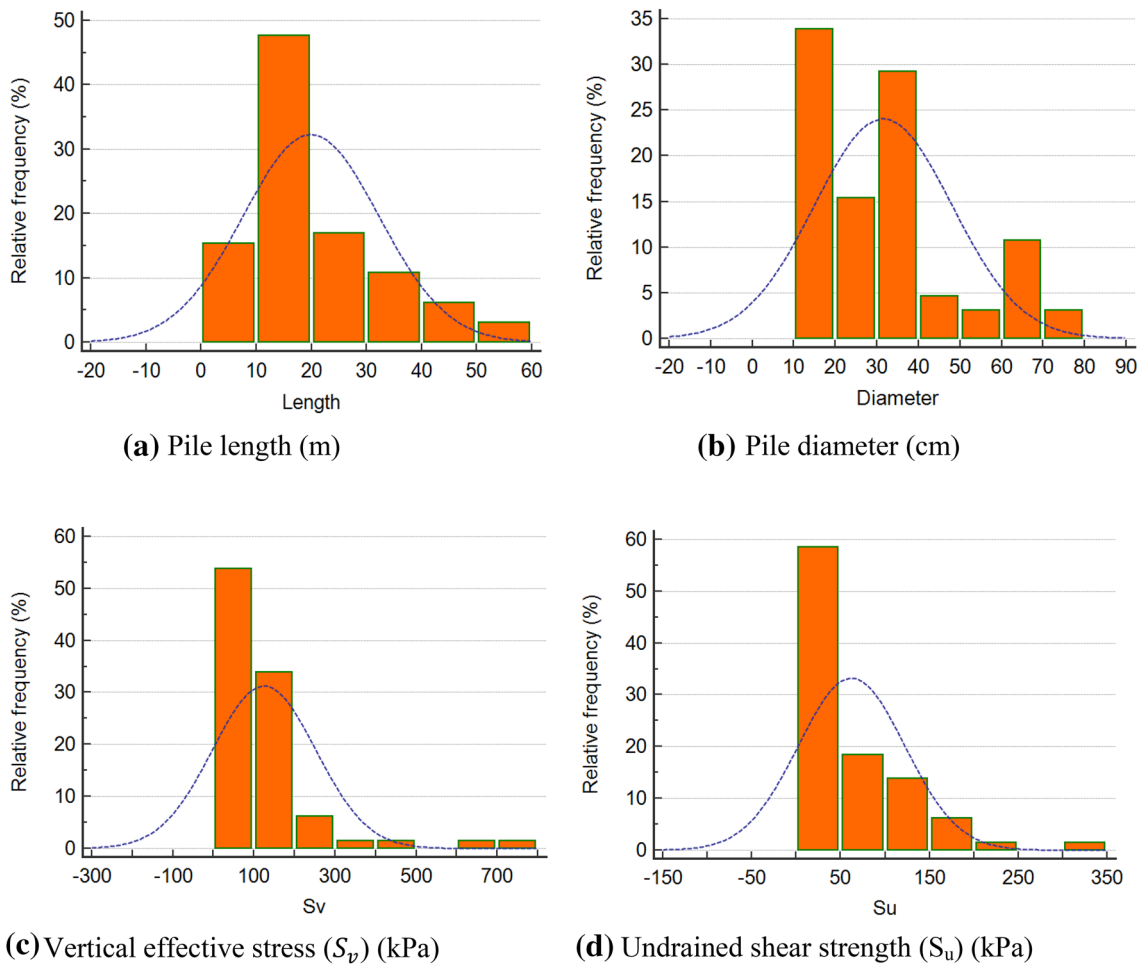


Fig. 9 Histogram of the input data layers, **a** pile length (m), **b** pile diameter, **c** vertical effective stress (s_v), **d** undrained shear strength (s_u)

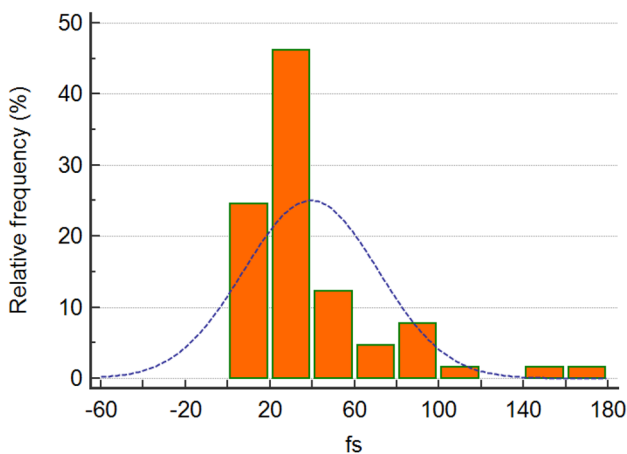


Fig. 10 Histogram of the friction capacity as the main output

RMSE reduction procedure. Considering the trial and error processes provided in Figs. 12, 13 and 14, the model of GP with the amounts of 750, 300 and 250, respectively, in the

case of the population size, some generations and selection tournament size showed the best efficiency, as indicated by its lower RMSE value. Therefore, this structure was introduced as the optimal architecture of the GP for any further estimation of the friction capacity of driven piles in cohesive soil. Noteworthy, the RMSE values obtained for the GP method are calculated based on a non-normalized input and output training and testing datasets.

In the case of the GWO-MLP method, a population size (N_{pop}) starting from 25 to 500 was considered. This can be called a population-based sensitivity analysis. Having 1000 iterations, 11 cases of GWO-MLP ensembles were tested. Different N_{pop} were considered, including 25, 50, 100, 150, 200, 250, 300, 400 and 500. The performances of GWO predictive networks were evaluated using Mean Square Error (MSE). It was expected that the best predictive, among all proposed models, shows the minimum MSE error. This can be found by looking at illustrated population-based sensitivity analysis (see Figs. 15, 16) or tabulated training and testing R^2 and RMSE outputs of each model in Table 1. The MSE performance of

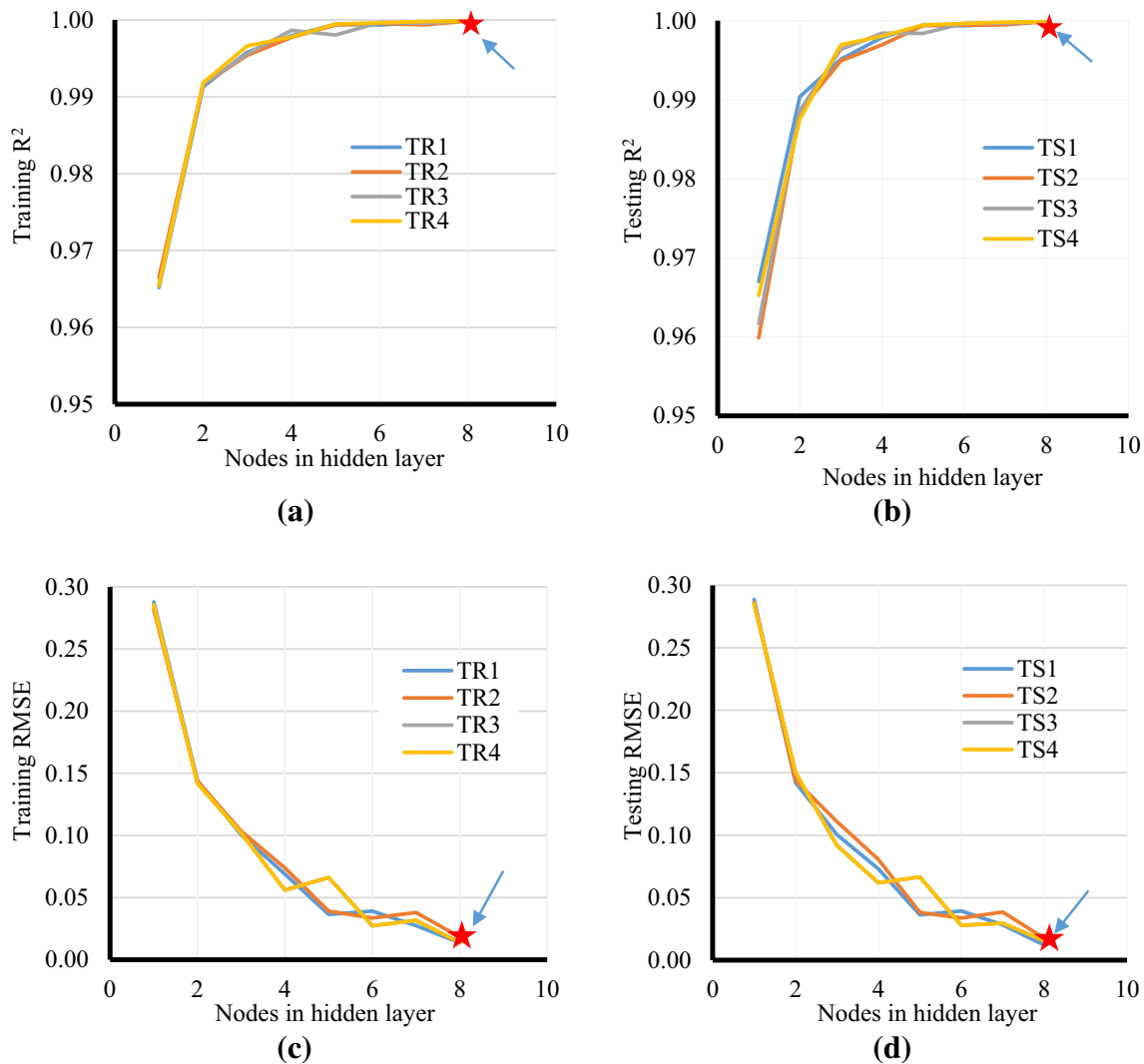


Fig. 11 Selection of appropriate nodes in the hidden layer using R^2 and RMSE for training and testing databases

GWO-MLP after 1000 iterations is also shown in Fig. 17. The results show that the population size equal to 500 will provide the most accurate predictive GWO network.

3.3 Model development findings

A reliable prediction process that is implemented using hybrid models of ANN must be constructed from several stages like (1) data processing and normalization, (2) selecting a suitable hybrid approach and finally (3) finding a proper hybrid structure for the proposed method, which may be obtained via a trial and error procedure. Both of the proposed ANN and GP models could provide excellent prediction results. Table 1 illustrates the result of the comparison of the proposed models in their optimal structures. In this regard, the CRS and TRM ranking systems were also used. Table 1 shows the advantage of using GP model instead of

conventional ANN model in this study. The result of the network performance of training and testing databases to the suggested ANN and GP model in estimating friction capacity of the driven piles is presented in Figs. 18 and 19, respectively. The R^2 of 0.8963 in Fig. 18b indicates approximately fair agreement between measured and predicted for testing dataset using ANN model.

On the other hand, GP model gives better coefficient of determination of $R^2=0.913$ in Fig. 19b. GP model has more accuracy in predicting pile skin friction capacity in addition to faster training procedures. The GWO-MLP technique showed better performance as the R^2 of both training and testing datasets were 0.982, 0.982, respectively. The results obtained from the GWO was superior compared to conventional MLP or GP techniques. The Eq. (5) is also provided for the estimating pile capacity values, but it is from the GP predictive network algorithm (Fig. 20):

Fig. 12 Selection of best population size for GP network using RMSE

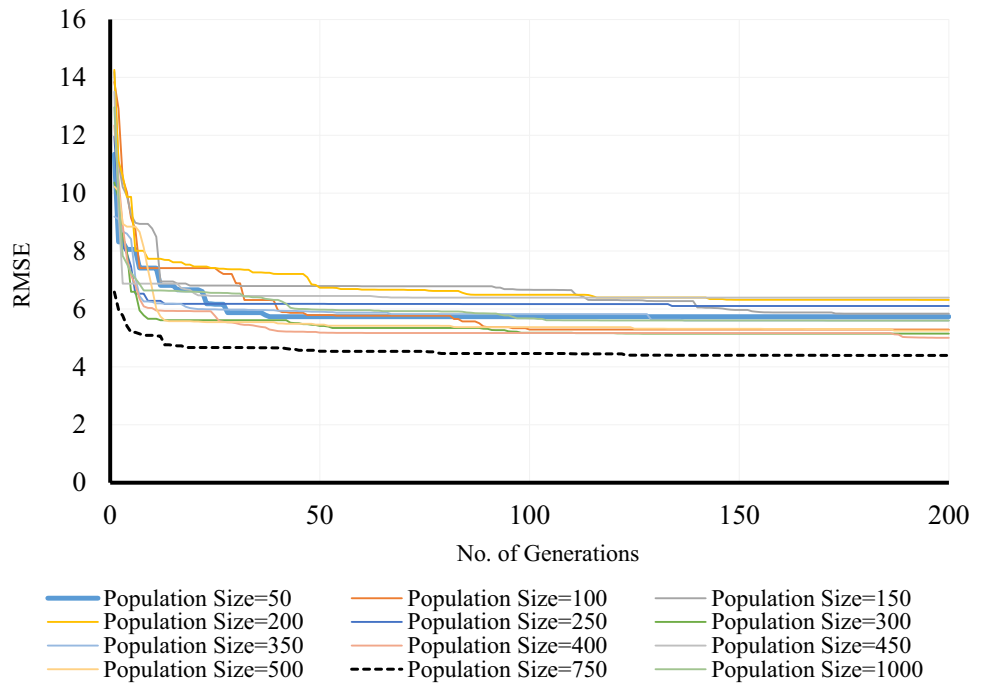
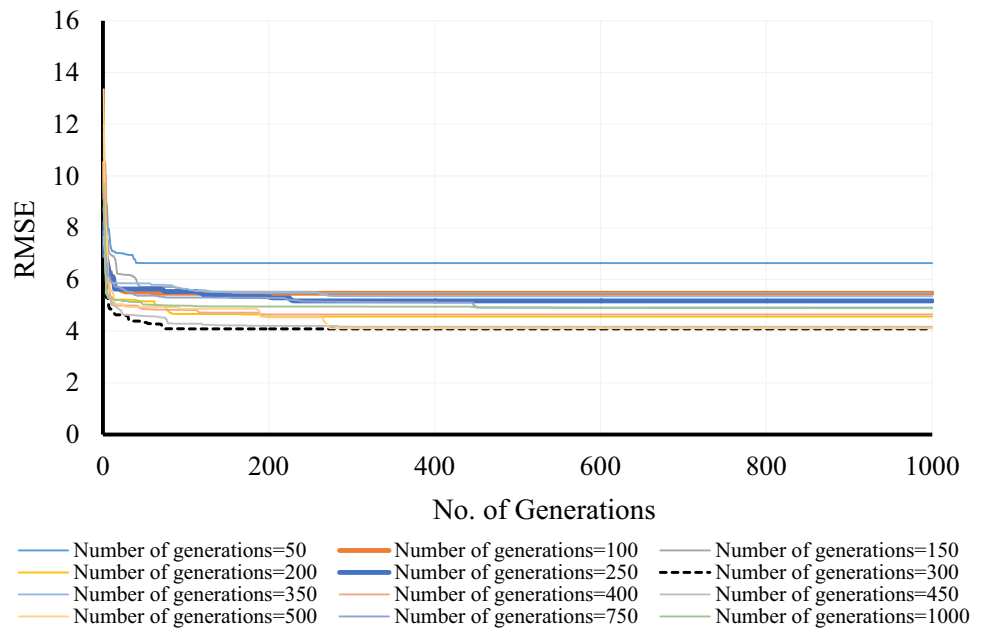


Fig. 13 Selection of the best number of generations for GP network using RMSE



$$\begin{aligned}
 Q_f &= 0.5573 \times Y_1 - 0.7729 \times Y_2 + 0.0055 \times Y_3 + 0.3298 \times Y_4 - 0.7695 \times Y_5 - 0.3514 \times Y_6 - 0.0923 \times Y_7 + 0.1238 \times Y_8 - 0.3556 \\
 Y_1 &= \text{Tansig}(0.9623 \times A + 0.9304 \times B - 1.6039 \times C + 1.5458 \times D - 2.4960) \\
 Y_2 &= \text{Tansig}(1.3168 \times A - 0.8490 \times B - 0.8438 \times C + 1.4584 \times D - 1.7208) \\
 Y_3 &= \text{Tansig}(-0.9937 \times A + 2.5840 \times B + 1.2616 \times C + 0.6396 \times D + 0.4859) \\
 Y_4 &= \text{Tansig}(0.7810 \times A - 1.7112 \times B + 1.1512 \times C - 2.4210 \times D + 0.1313) \\
 Y_5 &= \text{Tansig}(0.5206 \times A - 0.5597 \times B + 0.1139 \times C - 1.9060 \times D - 0.3959) \\
 Y_6 &= \text{Tansig}(1.6258 \times A + 0.4894 \times B - 2.0704 \times C + 0.1087 \times D + 0.7643) \\
 Y_7 &= \text{Tansig}(-1.8600 \times A - 0.9789 \times B + 0.6155 \times C - 0.7022 \times D - 0.7961) \\
 Y_8 &= \text{Tansig}(2.1393 \times A + 0.8077 \times B + 0.6632 \times C + 2.1325 \times D + 3.6012),
 \end{aligned}
 \tag{16}$$

Fig. 14 Selection of best tournament size for GP network using RMSE

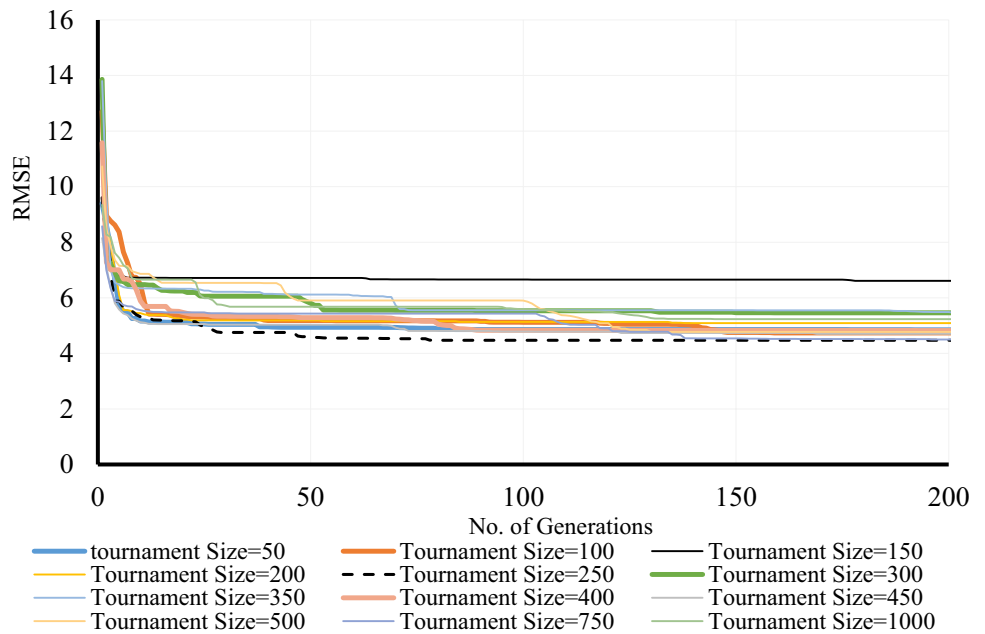
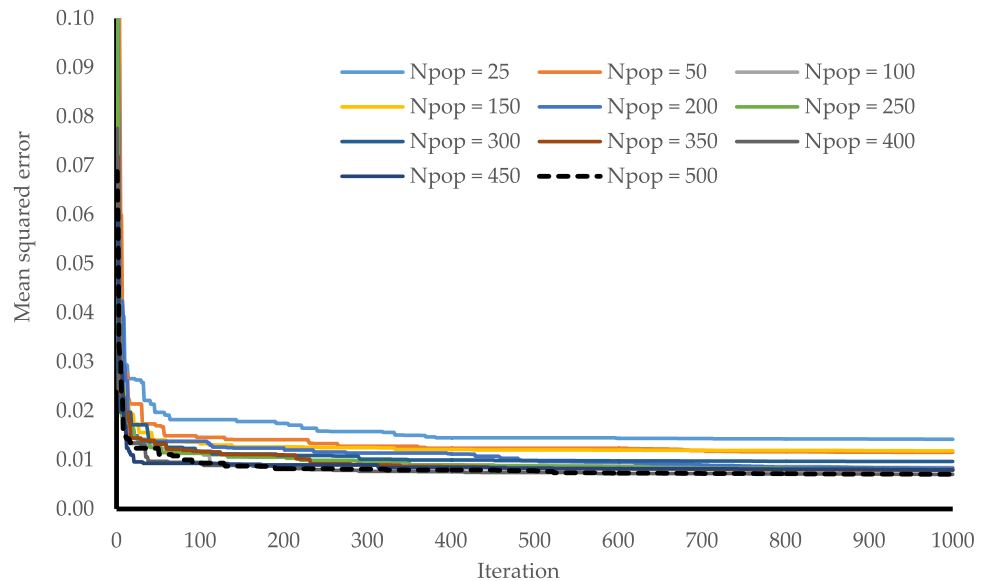


Fig. 15 Variation of the MSE with N_{pop}



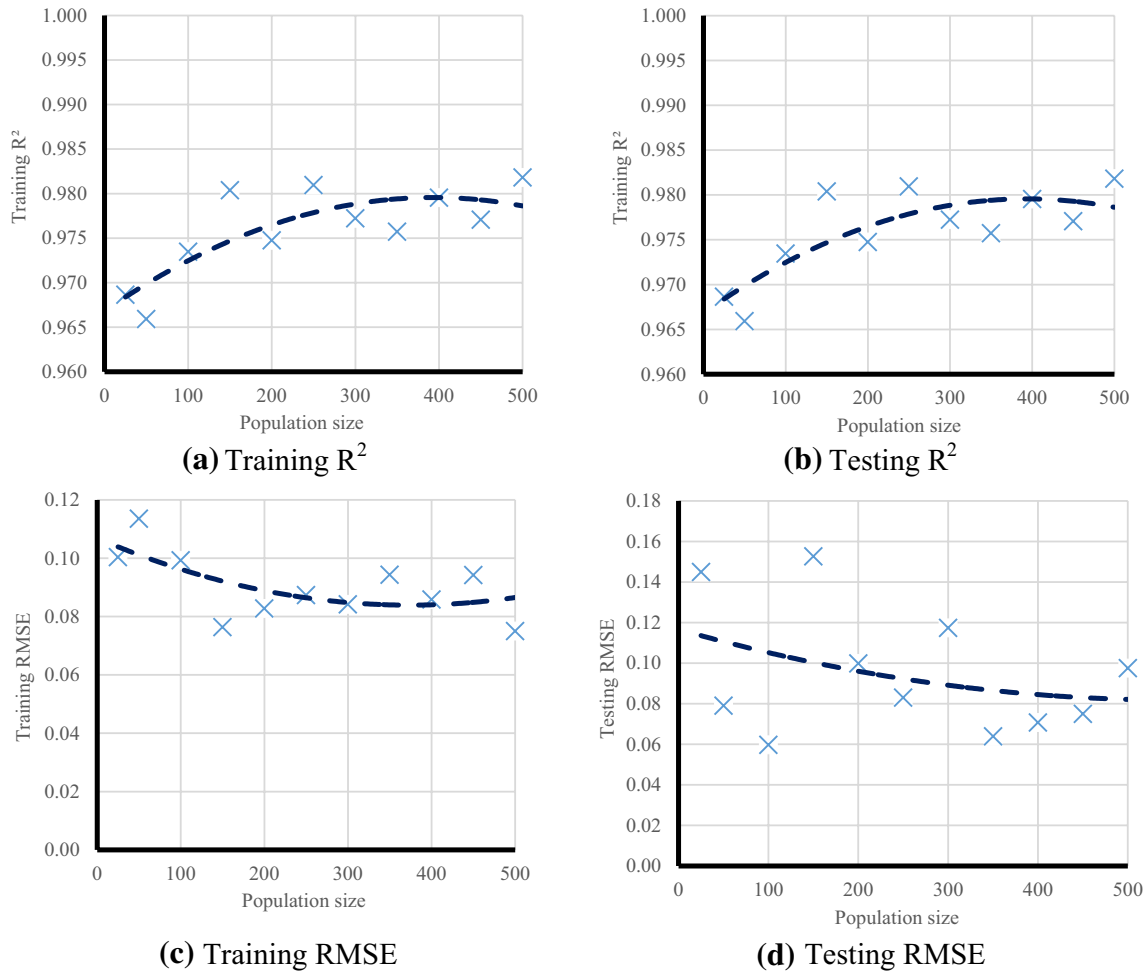


Fig. 16 Sensitivity analysis of population size in GWO-MLP for both of the training and testing datasets

Table 1 Population-based sensitivity analysis in GWO-MLP algorithm

Population size	Network result				Ranking				Total rank	RANK
	Train		Test		Train		Test			
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE		
25	0.969	0.100	0.955	0.145	2	2	3	2	9	11
50	0.966	0.114	0.954	0.079	1	1	2	7	11	10
100	0.973	0.099	0.986	0.060	3	3	11	11	28	3
150	0.980	0.076	0.948	0.153	9	10	1	1	21	9
200	0.975	0.083	0.979	0.100	4	9	7	4	24	7
250	0.981	0.087	0.956	0.083	10	6	4	6	26	5
300	0.977	0.084	0.970	0.117	7	8	5	3	23	8
350	0.976	0.094	0.982	0.064	5	4	9	10	28	3
400	0.980	0.086	0.980	0.071	8	7	8	9	32	2
450	0.977	0.094	0.972	0.075	6	5	6	8	25	6
500	0.982	0.075	0.982	0.097	11	11	10	5	37	1

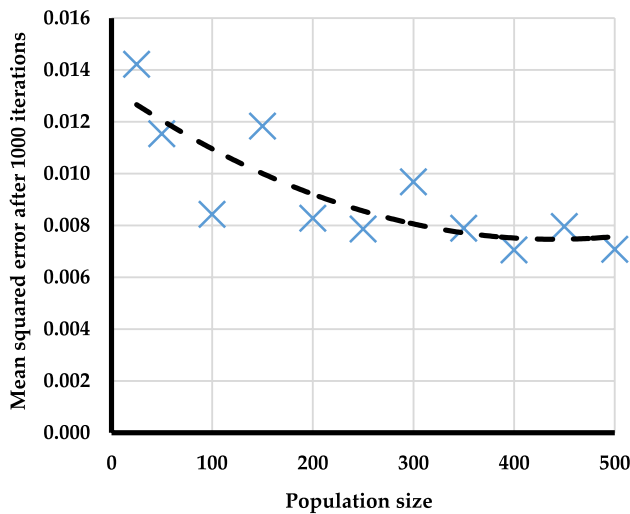


Fig. 17 The MSE performance of GWo-MLP after 1000 iterations

where Q_f is the friction capacity of the driven piles (kPa), the term A is the pile’s length (in meters), B is the diameter of pile (should be in centimeters), the term C is vertical stress σ'_v (kPa) in pile toe, and D which normally denotes as S_u , represents as the values of undrained soil shear strength (kPa).

$$\begin{aligned}
 Q_f = & \frac{2885452003929751 \times (B)}{36028797018963968} + \frac{27981927016846971 \times (C)}{72057594037927936} + \frac{735861421529235 \times (D)}{1125899906842624} \\
 & + \frac{1226640469248279 \times \text{COS}(\exp(A))}{562949953421312} + \frac{1226640469248279 \times \text{COS}(C)}{562949953421312} - \frac{2885452003929751 \times (A \times C \times D)^{0.5}}{72057594037927936} \\
 & - \frac{8193276386513187 \times (B) \times (C)}{9223372036854775808} - \frac{8193276386513187 \times (D) \times ((2 \times C) + D)}{18446744073709551616} - \frac{1226640469248279 \times (C^{0.5})}{562949953421312} + 2.3493.
 \end{aligned}
 \tag{17}$$

4 Conclusions

The importance of estimating friction capacity of the driven piles (i.e., embedded in cohesive soils) reliably and its complexity in engineering projects is well established. Therefore, the main aim of this work was to assess the capability of two artificial intelligence techniques in the assessment of friction capacity of the driven piles installed in the cohesive soil environment. To do this, optimized GP and GWO were implemented to optimize conventional MLP methods. First, results of 65 in situ tests conducted on concrete driven piles were collected. Then, four effective factors that affect the friction capacity of the driven piles, namely length, diameter, effective stress and undrained cohesion strength of the adjacent soils were considered as input data. The optimization activity was done using a long trial and error procedure to find the networks with the best network performance. The obtained architecture for the ANN was $4 \times 8 \times 1$, which indicates an MLP network to 8 neurons in its single hidden layer as chosen. Also, the model of GP with the amounts of 750, 300, and 250, respectively, in the case of population size, number of generations and selection tournament size provided the most appropriate prediction and was identified as the optimum GP model. According to the results, both the

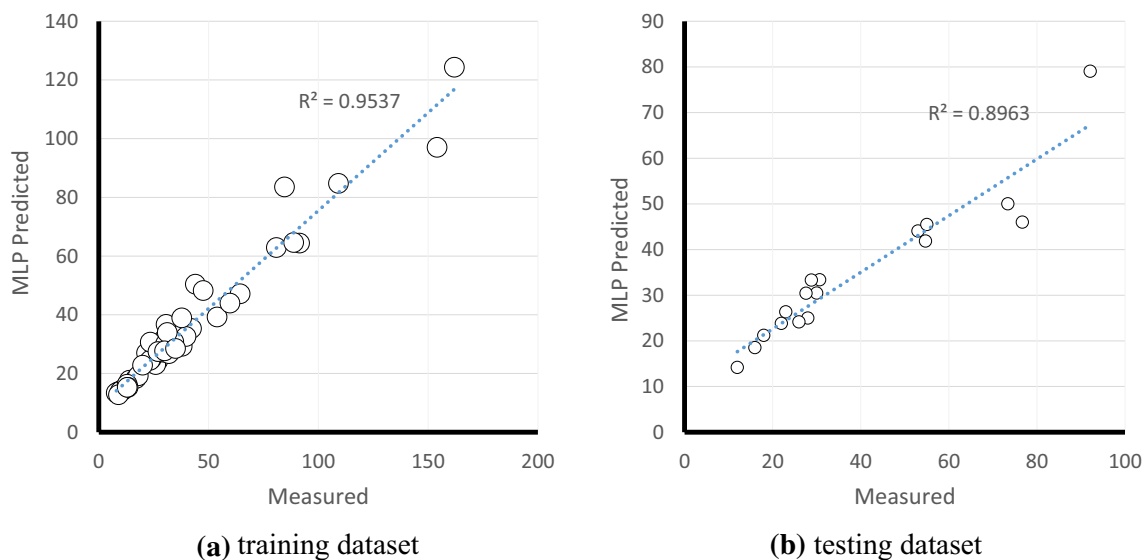


Fig. 18 Predicted and measured pile friction capacity using non-normalized MLP model **a** training and **b** testing dataset

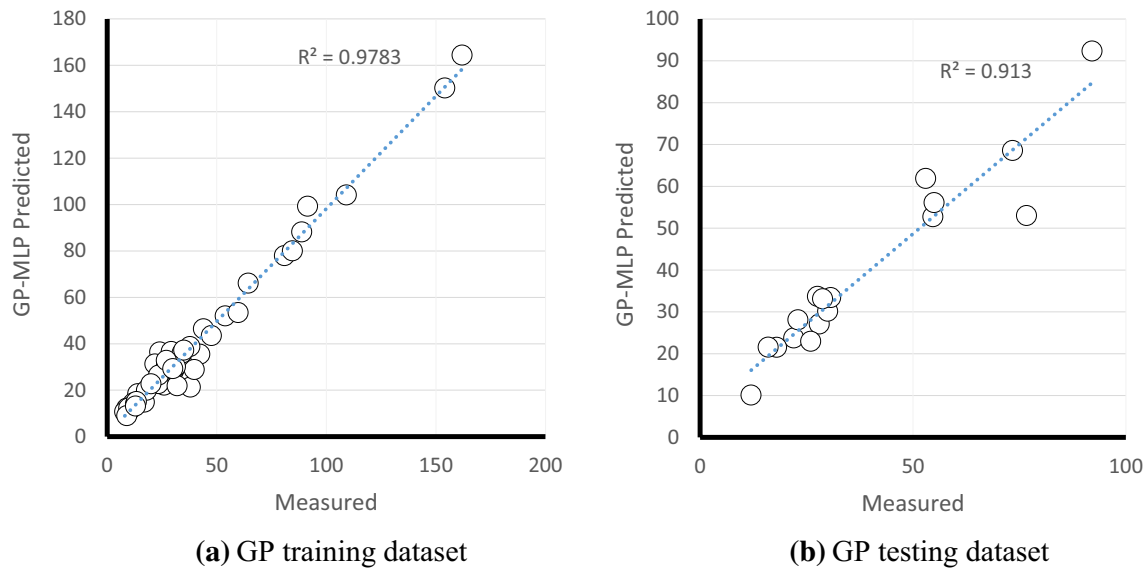


Fig. 19 Predicted and measured pile friction capacity using non-normalized GP model (training and testing dataset)

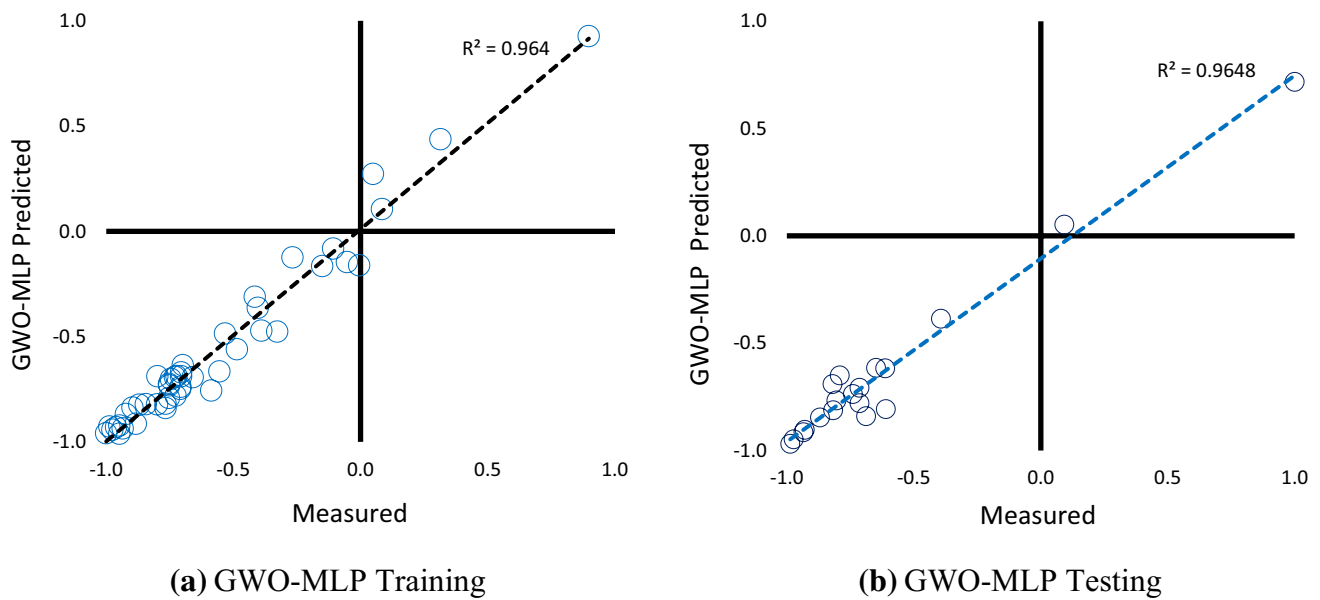


Fig. 20 Predicted and measured pile friction capacity using normalized (−1.0 to +1.0) GWO-MLP model, a training, b testing datasets

RMSE and R^2 values showed a slightly better approximation for the GWO-MLP model in training and testing datasets than the conventional MLP or even optimized GP techniques. It can be noted that due to using only four parameters consisting of pile diameter, pile length, effective stress and undrained shear strength to develop MLP, GWO-MLP and GP employed models as well as inherent complexity of soil properties and layering this study may have some limitations but, when available input parameters are limited it will

provide good assessment of friction capacity of concrete driven piles.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Moayedi H, Raftari M, Sharifi A, Jusoh WAW, Rashid ASA (2019) Optimization of ANFIS with GA and PSO estimating α ratio in driven piles. *Eng Comput* 35:1–12
- Moayedi H, Armaghani DJ (2017) Optimizing an ANN model with ICA for estimating bearing capacity of driven pile in cohesionless soil. *Eng Comput* 34:347–356
- Liu L, Moayedi H, Rashid ASA, Rahman SSA, Nguyen H (2019) Optimizing an ANN model with genetic algorithm (GA) predicting load-settlement behaviours of eco-friendly raft-pile foundation (ERP) system. *Eng Comput* 36:1–13
- Alkroosh I, Nikraz H (2012) Predicting axial capacity of driven piles in cohesive soils using intelligent computing. *Eng Appl Artif Intell* 25:618–627
- Randolph MF (2003) Science and empiricism in pile foundation design. *Géotechnique* 53:847–874
- Coduto R (2001) Foundation design principles and practices. Prentice Hall, Englewood Cliffs
- Baker VA, Thomsen NA, Nardi CR, Talbot MJ (1984) Pile foundation design using pile driving analyzer. *Analysis and Design of Pile Foundations*, San Francisco
- Tomlinson M, Woodward J (2008) Pile design and construction practice, 5th edn. Taylor and Francis Group, New York
- Nazir R, Moayedi H, Subramaniam P, Ghareh S (2017) Ground improvement using SPVD and RPE. *Arab J Geosci* 10:21
- Moayedi H, Huat BBK, Ali TAM, Asadi A, Moayedi F, Mokheri M (2011) Preventing landslides in times of rainfall: case study and FEM analyses. *Disaster Prev Manag* 20:115–124
- Gao W, Karbasi M, Derakhsh AM, Jalili A (2019) Development of a novel soft-computing framework for the simulation aims: a case study. *Eng Comput* 35:315–322
- Gao W, Raftari M, Rashid ASA, Mu'azu MA, Jusoh WAW (2019) A predictive model based on an optimized ANN combined with ICA for predicting the stability of slopes. *Eng Comput* 2019:1–20
- García-Balboa JL, Alba-Fernández MV, Ariza-López FJ, Rodríguez-Avi J (2018) Analysis of thematic similarity using confusion matrices. *ISPRS Int J Geoinform* 7:233
- Samui P (2008) Prediction of friction capacity of driven piles in clay using the support vector machine. *Can Geotech J* 45:288–295
- Schneider JA, Xu XT, Lehane BM (2008) Database assessment of CPT-based design methods for axial capacity of driven piles in siliceous sands. *J Geotech Geoenviron Eng* 134:1227–1244
- Xia T, Wang W (2009) Study on evaluating methods for time-dependent ultimate bearing capacity of single driven pile. *IEEE Computer Soc, Los Alamitos*
- Tran KT, McVay M, Herrera R, Lai P (2011) A new method for estimating driven pile static skin friction with instrumentation at the top and bottom of the pile. *Soil Dyn Earthq Eng* 31:1285–1295
- Samui P (2011) Multivariate adaptive regression spline applied to friction capacity of driven piles in clay. *Geomech Eng* 3:285–290
- Mosallanezhad M, Moayedi H (2017) Developing hybrid artificial neural network model for predicting uplift resistance of screw piles. *Arab J Geosci* 10:10
- Samui P (2012) Determination of ultimate capacity of driven piles in cohesionless soil: a multivariate adaptive regression spline approach. *Int J Numer Anal Methods Geomech* 36:1434–1439
- Moayedi H, Hayati S (2018) Artificial intelligence design charts for predicting friction capacity of driven pile in clay. *Neural Comput Appl* 31 (in press)
- Gao W, Wang W, Dimitrov D, Wang Y (2018) Nano properties analysis via fourth multiplicative ABC indicator calculating. *Arab J Chem* 11:793–801
- Nazir R, Moayedi H, Subramaniam P, Gue SS (2018) Application and design of transition piled embankment with surcharged prefabricated vertical drain intersection over soft ground. *Arab J Sci Eng* 43:1573–1582
- Gao W, Wu H, Siddiqui MK, Baig AQ (2018) Study of biological networks using graph theory. *Saudi J Biol Sci* 25:1212–1219
- Bui DT, Moayedi H, Gör M, Jaafari A, Foong LK (2019) Predicting slope stability failure through machine learning paradigms. *ISPRS Int J Geoinform* 8:395
- Gao W, Guirao JLG, Basavanagoud B, Wu J (2018) Partial multi-dividing ontology learning algorithm. *Inf Sci* 467:35–58
- Moayedi H, Tien Bui D, Gör M, Pradhan B, Jaafari A (2019) The feasibility of three prediction techniques of the artificial neural network, adaptive neuro-fuzzy inference system, and hybrid particle swarm optimization for assessing the safety factor of cohesive slopes. *ISPRS Int J Geoinform* 8:391
- Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. *Adv Eng Softw* 69:46–61
- Bozorg-Haddad O (2018) Advanced optimization by nature-inspired algorithms. Springer, Berlin
- Muro C, Escobedo R, Spector L, Coppinger R (2011) Wolf-pack (*Canis lupus*) hunting strategies emerge from simple rules in computational simulations. *Behav Process* 88:192–197
- Dehghani M, Riahi-Madvar H, Hooshyaripour F, Mosavi A, Shamsirband S, Zavadskas EK, Chau K-w (2019) Prediction of hydropower generation using grey wolf optimization adaptive neuro-fuzzy inference system. *Energies* 12:289
- Koza JR (1994) Genetic programming as a means for programming computers by natural selection. *Stat Comput* 4:87–112
- Rezania M, Javadi AA (2007) A new genetic programming model for predicting settlement of shallow foundations. *Can Geotech J* 44:1462–1473
- Fatehnia M, Amirinia G (2018) A review of genetic programming and artificial neural network applications in pile foundations. *Int J Geoen* 9:20
- Obiedat R, Alkasasbeh M, Faris H, Harfoushi O (2013) Customer churn prediction using a hybrid genetic programming approach. *Sci Res Essays* 8:1289–1295
- Cramer NL (1985) A representation for the adaptive generation of simple sequential programs. In: *Proceedings of the first international conference on genetic algorithms*
- Baziar MH, Azizkandi AS (2013) Evaluation of lateral spreading utilizing artificial neural network and genetic programming. *Int J Civ Eng* 11:100–111
- Pandey DS, Pan I, Das S, Leahy JJ, Kwapinski W (2015) Multi-gene genetic programming based predictive models for municipal solid waste gasification in a fluidized bed gasifier. *Bioresour Technol* 179:524–533
- Zhai S, Mao J, Liu H (2017) Average velocity of debris flow forecast based on genetic programming combined with rough set theory. In: *DEStech transactions on engineering and technology research*
- Goh ATC (1996) Pile driving records reanalyzed using neural networks. *J Geotech Eng ASCE* 122:492–495
- Moayedi H, Mosallanezhad M (2017) Physico-chemical and shrinkage properties of highly organic soil treated with non-traditional additives. *Geotech Geol Eng* 35:1409–1419
- Moayedi H, Armaghani DJ (2018) Optimizing an ANN model with ICA for estimating bearing capacity of driven pile in cohesionless soil. *Eng Comput* 34:347–356
- Vakili AH, Bin Selamat MR, Mohajeri P, Moayedi H (2018) A critical review on filter design criteria for dispersive base soils. *Geotech Geol Eng* 36:1933–1951

44. Moayedi H, Hayati S (2018) Modelling and optimization of ultimate bearing capacity of strip footing near a slope by soft computing methods. *Appl Soft Comput* 66:208–219
45. Asadi A, Moayedi H, Huat BBK, Boroujeni FZ, Parsaie A, Sojoudi S (2011) Prediction of zeta potential for tropical peat in the presence of different cations using artificial neural networks. *Int J Electrochem Sci* 6:1146–1158
46. Asadi A, Moayedi H, Huat BBK, Parsaie A, Taha MR (2011) Artificial neural networks approach for electrochemical resistivity of highly organic soil. *Int J Electrochem Sci* 6:1135–1145
47. Raftari M, Rashid ASA, Kassim KA, Moayedi H (2014) Evaluation of kaolin slurry properties treated with cement. *Measurement* 50:222–228

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.