Ocean Engineering 146 (2017) 1–8

Contents lists available at ScienceDirect

Ocean Engineering

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

Estimating uplift capacity of suction caissons in soft clay: A hybrid computational approach based on model tree and GP

Ali Derakhshani *

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

ARTICLE INFO

Keywords:
Suction caisson
Uplift capacity
Formulation
Hybrid intelligent approach
M5-GP method

ABSTRACT

Stability of suction caissons used as foundations or anchors of offshore structures is a critical challenge in marine structures engineering. To this end, many studies have been conducted including those concentrate on implementing computational intelligence methods to model the response of suction caissons under loading. In this regard, this paper aims at formulating uplift capacity of suction caissons using a hybrid artificial intelligence computational tool based on model tree (M5) and genetic programming (GP), called M5-GP. The formulae are developed in terms of several governing parameters using a reliable experimental database from the literature. The results show that the M5-GP based relationships are able to predict the uplift capacity of suction caissons precisely. Furthermore, to consider the safety in the design process, probabilistic equations are also given for various risk levels. The new formulas compare favorably with the existing relationships in the literature regarding prediction performance. In addition, the simplified formulation is compact, easy to use and physically sound. Therefore, it is especially appropriate to be used in design practice.

1. Introduction

Suction caisson is one of the most widely used anchoring facilities in offshore engineering applications. These tubular steel fabrications were firstly introduced by (Senpere and Auvergne, 1982) as mooring anchors to be used in an offshore project. Also called as suction piles, anchors and buckets, suction caissons are an essential part of anchorage system utilized in offshore drilling equipment (Chen and Randolph, 2007; Ehlers et al., 2004; Zdravkovic et al., 2001).

Suction caissons have several advantages over other types of offshore anchorage systems such as gravity anchors, drag anchors and conventional driven anchor piles (Schneider and Senders, 2010). Design and fabrication is quite easy as they are simply steel tubes closed at the top (Fig. 1). The steel needed to build a suction caisson can be less than that of an equivalent deep pile foundation. The installation process is not complicated since the suction caisson is penetrated into the seafloor due to its self-weight and then the suction is applied by pumping out the trapped water between the seabed and the closed caisson top side. This mechanism considerably facilitates the installation of suction caissons in comparison with pile foundation derivation where several costly ancillary equipment are required (Alavi et al., 2011; Colliat and Dendani, 2002). On the other hand, this method of installation is relatively faster than other types and due to difficult condition of marine construction, the shorter the construction time, the greater the cost reduction (Cheng et al., 2014). Additionally, suction caissons are usually able to carry greater lateral loads than alternatives.

However, in general, the suction caisson systems, may not be as reliable as offshore jackets (Clukey et al., 2000). A suction caisson failure can lead to the collapse of an offshore structure followed by severe negative consequences such as loss of lives, environmental pollution and substantial financial costs. Therefore, precise estimation of suction caisson uplift capacity is an extremely important task to meet the reliable design requirements. To this end, the intricate response of suction caissons under loading should be evaluated carefully.

The main design considerations of a suction caisson are its submerged weight, friction resistance of its wall surface, negative pressure generated under tension and soil strength mobilized at the base (Alavi et al., 2010, 2011; Gandomi et al., 2011).

Artificial intelligence (AI)-based modelling of various engineering problems is considered as a suitable alternative to traditional methods like Finite Element Method (FEM). Recently, soft computing approaches such as Artificial Neural Networks (ANNs), Evolutionary Algorithms, etc.

* Corresponding author. Tel.: +98 2151212084, +98 9125259707
E-mail addresses: adera@shahed.ac.ir, derakhshani85@gmail.com.

https://doi.org/10.1016/j.oceaneng.2017.09.025
Received 15 April 2017; Received in revised form 7 September 2017; Accepted 16 September 2017

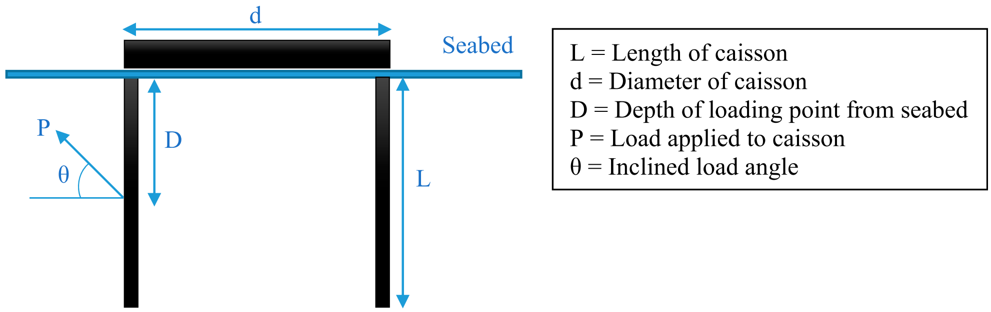

Fig. 1. Cross section sketch of suction caisson.

are implemented in many of civil engineering applications including estimation of the uplift capacity of suction caissons.

M5 model tree is a powerful soft computing tool that is able to classify the data and present simple relationships. It has some advantages over other soft computing approaches like ANNs, due to the fact that it is not opaque and internal parameters do not require optimization. M5 model tree has been successfully used for prediction purposes in several civil engineering problems such as breakwaters stability, scour around pile groups and scour under submarine pipelines (Etemad-Shahidi and Bali, 2012; Etemad-Shahidi and Ghaemi, 2011; Etemad-Shahidi et al., 2011).

Genetic Programming (GP) is a type of evolutionary algorithms in which a supervised machine learning method explores program space rather than data space (Banzhaf et al., 1998). Where a robust theoretical model is not available, GP yields a computer program i.e. an interpretable equation. This has made the GP a strong tool for solving complicated engineering problems. Recently, GP has been used for the estimation of shallow foundations settlement (Shahnazari et al., 2014), scour under submerged pipelines (Azamathulla et al., 2011) and scour around piles (Guven et al., 2009).

This study aims at developing new generic uplift capacity formulae using an AI-based model, a compound of M5 model tree (M5) and genetic programming (GP) called “M5-GP”. This paper is organized as follows: The second Section explains the relevant technical literature. The third addresses the details of the hybrid M5-GP method. In the fourth section, the application of the hybrid M5-GP model for the development of predictive relationships are described. The fifth section analyzes the performance of M5-GP and compares the results with those of available models in the literature. Finally, the conclusions are drawn in section six.

2. Background

Many laboratory tests (Datta and Kumar, 1996; El-Gharbawy and Olson, 2000; Gao et al., 2013; Guo et al., 2012; Hogervorst, 1980; Li et al., 2014, 2015; Randolph et al., 1998; Rao et al., 2006; Tjelta, 1995; Whittle and Germaine, 1998) and field experimental studies (Cho et al., 2002; Dyvik et al., 1993) have been conducted to assess the behavior of suction caissons including installation characteristics, vertical and horizontal load capacities and stress conditions. But, such methods are expensive and exposed to number of restrictions.

In some research programs the suction caissons were modelled numerically to estimate the uplift capacity (Ahn et al., 2015; Aubeny and Murff, 2005; Deng and Carter, 1999, 2002; Sukumaran et al., 1999; Zdravkovic et al., 2001; Zeinoddini et al., 2011). Although numerical models such as those based on FEM are not exposed to limitations of experimental studies, they are restricted to special properties of each model (Pai, 2005). Hence, it is necessary to develop comprehensive mathematical models for prediction of the uplift capacity in a wide range of conditions.

Moreover, many attempts have been made to implement AI approaches for estimation of the uplift capacity of suction caissons. The methods used by different researchers are listed as follows: artificial neural network (ANN) by Rahman et al. (2001), neuro-genetic network (NGN) by Pai (2005), GP and simulated annealing (GP-SA) by Alavi et al. (2010), tree-based genetic programming (TGP), linear genetic programming (LGP) and gene expression programming (GEP) by Alavi et al. (2011), multi expression programming (MEP) by Gandomi et al. (2011), multivariate adaptive regression spline (MARS) by Samui et al. (2011), support vector machine and ANN (SVM-ANN) by Muduli et al. (2013), intelligent fuzzy radial basis function neural network inference method (IFRIM) by Cheng et al. (2014) and Group method of data handling and harmony search (GMDH-HS) by Shahr-Babak et al. (2016).

3. Methodology of data analysis

Conventional regression analysis for derivation of empirical correlations relies on pre-defined relationships between inputs and output. Hence, the main task is finding empirical coefficients of the functional structure. However, a predefined function may not match the data for complicated phenomena and leads to an inexact model.

Modern data mining methods, are used to discover relationships hidden in datasets utilizing different optimization algorithms. Soft computing methods can be used for complicated predictions, building nonlinear relationships, categorizing data and deriving rule-based models (Solomatine and Ostfeld, 2008).

In this study, major criteria to select the most suitable modelling approach are accuracy and simplicity of the prediction formulae that are constructed by implementing an AI-based model. It was inferred that a combination of M5 and GP called hybrid M5-GP model is a suitable option among various possibilities.

3.1. M5 model tree

The M5 model tree is one of the most robust data mining tools that can be implemented for prediction in engineering applications. This approach which was presented by Quinlan (1992) separates complicated problems into smaller parts and deals with a number of simpler problems (Bhattacharya et al., 2007). M5 resembles an inverse tree with a root at the top and leaves at the bottom. Three stages of M5 modeling procedure are: building, pruning and smoothing.

Building: A tree is made of splitting the instance (dataset) space. The intra-subset variability in the values is minimized by applying the classification condition down from the root to the node (via the branch). The standard deviation of the values which reach the node indicates the expected error reduction in variability by testing each attribute at the node. Hence, the attribute (input) causing the minimum expected error reduction is obtained. This process goes on until a limited number of instances remain or the instances arrived at the node (leaf) are quite similar. The standard deviation reduction (SDR) is used in the mentioned process and defined as:

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \times sd(T_i) \tag{1}$$

where T is the set of instances arriving at the node, Ti are the sets that result from classification based on the selected attribute and sd is the standard deviation (Wang and Witten, 1996). After raising the tree, the regression models are developed by the data of the leaves.

Pruning: In this stage, sub-trees are evaluated whether the error of the linear model at the root is lower than /equal to the anticipated error. In such manner, the sub-trees that are not capable of enhancing the model accuracy are pruned. Then, discontinuities possibly occurred among adjacent leaves will be treated in the next step.

Smoothing: In this step, the estimation of the leaf model is filtered through the path toward the root. That value is merged with the value presented by the linear model for that node. Therefore, the prediction passed to the higher node i.e. P' equals to $(np + kq)/(n + k)$ where p is the prediction passed to this node from the lower node, q is the predicted value by the model at this node, n is the number of training data points which arrives at the lower node, and k is a constant (Wang and Witten, 1996).

3.2. Genetic programming (GP)

Genetic programming (GP) is an evolutionary symbolic regression approach that uses the principle of Darwinian natural selection. GP was introduced by (Koza, 1992) for solving problems with different degrees of complexity and is similar to genetic algorithm (GA). However, GP is different from GA such that it represents the solution in the form of an equation (computer program) instead of a set of numbers.

The procedure of GP is explained as follows:

(i) A population of individuals (programs) with a specific size is made by random selection of functions containing mathematical operators like addition, subtraction, multiplication, etc., and the terminals including constants and input parameters. The function and terminal sets make the main structure of GP. Depending on the complexity level of the problem, suitable operators should be selected. Inappropriate selection may lead to derivation of too complicated relationships.

(ii) The fitness quality of individuals in a population is assessed using an appropriate statistical measure. For example, in case the root mean squared error (RMSE) used as the criterion, its smaller value proves the higher quality of fitness.

(iii) Considering a tree-based representation, the genotype is configured in a way that top and middle of the tree is made of members of the function set and the leaves involve members of the terminal set. Then, new sets of models are created implementing genetic operators such as crossover, mutation and reproduction (Koza, 1992). These new models made at each generation called offspring and provide the base of the following generation.

(iv) The program is terminated after the creation of selected number of generations has been completed. Finally, the relationship with the best fitness is given as the best solution.

3.3. Hybrid M5-GP method

In order to employ the advantages of M5 and GP simultaneously, they were coupled for modelling the uplift capacity (Bonakdar et al., 2015). The database was categorized by M5 model tree algorithm to different categories. After the classification of the data, GP is used as a non-linear approach to estimate the uplift capacity of the suction caissons based on the categorized data. The procedures of hybrid M5-GP modelling is presented in Fig. 2 via a simple example.

Window 1 of Fig. 2 displays the application of M5 model tree to the database. This leads to classification of the data into several subsets as demonstrated in window 2. The subsets are actually the nodes at the end of the inverse tree and referred to as leaves. M5 model tree is also capable of developing linear models by the data of each leaf, however, the functional relationship is not necessarily linear. Hence, Genetic Programming is useful in the next stage.

As illustrated in window 3 of Fig. 2, the GP model is applied to the existing data at each subset created by M5. GP optimizes the model and gives an equation for each leaf as shown in the window 4. Therefore, the solution based on M5-GP is the output of GP applied to sub-sets categorized by M5.

4. Modelling of uplift capacity of suction caissons

As illustrated in Fig. 1, the most significant parameters affecting the uplift capacity are as follows (Pai, 2005; Rahman et al., 2001):

$$Q = f\left(\frac{L}{d}, \frac{D}{L}, \theta, S_u, T_k\right) \tag{2}$$

where, Q: uplift capacity of suction caisson; L/d: ratio of embedded length of the caisson (L) to its diameter (d); D/L: ratio of the depth of load application (D) to embedded length of the caisson (L); θ: angle that the inclined load makes with the horizontal; S_u: undrained shear strength of the soil at the caisson tip; $T_k = k/v$: ratio of the permeability of the soil (k) and the steady velocity at which the caisson is pulled from the ground (v);

Considering the mentioned important predictors and using the hybrid M5-GP method, the best equations are selected based on a multi-objective strategy. The criteria consist of choosing the simplest models, including maximum number of inputs and presenting the best fitness on the training and testing datasets. For the assessment of the proposed models, parameters such as correlation coefficient (CC), root mean squared error (RMSE) and mean absolute error (MAE) were calculated. Their relationships are as below:

$$CC = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\sum(x_i - y_i)^2} \tag{4}$$

$$MAE = \frac{\sum_{i=1}^{N}(|x_i - y_i|)}{N} \tag{5}$$

where, x_i and y_i are measured and predicted values and N is the number of samples.

4.1. Database

The database made of the results of 12 experiments were collected by (Rahman et al., 2001). This database that includes 62 cases has been considered in various studies in which the prediction models were developed for the uplift capacity of suction caissons (Alavi et al., 2011; Cheng et al., 2014; Gandomi et al., 2011; Shahr-Babak et al., 2016). It is

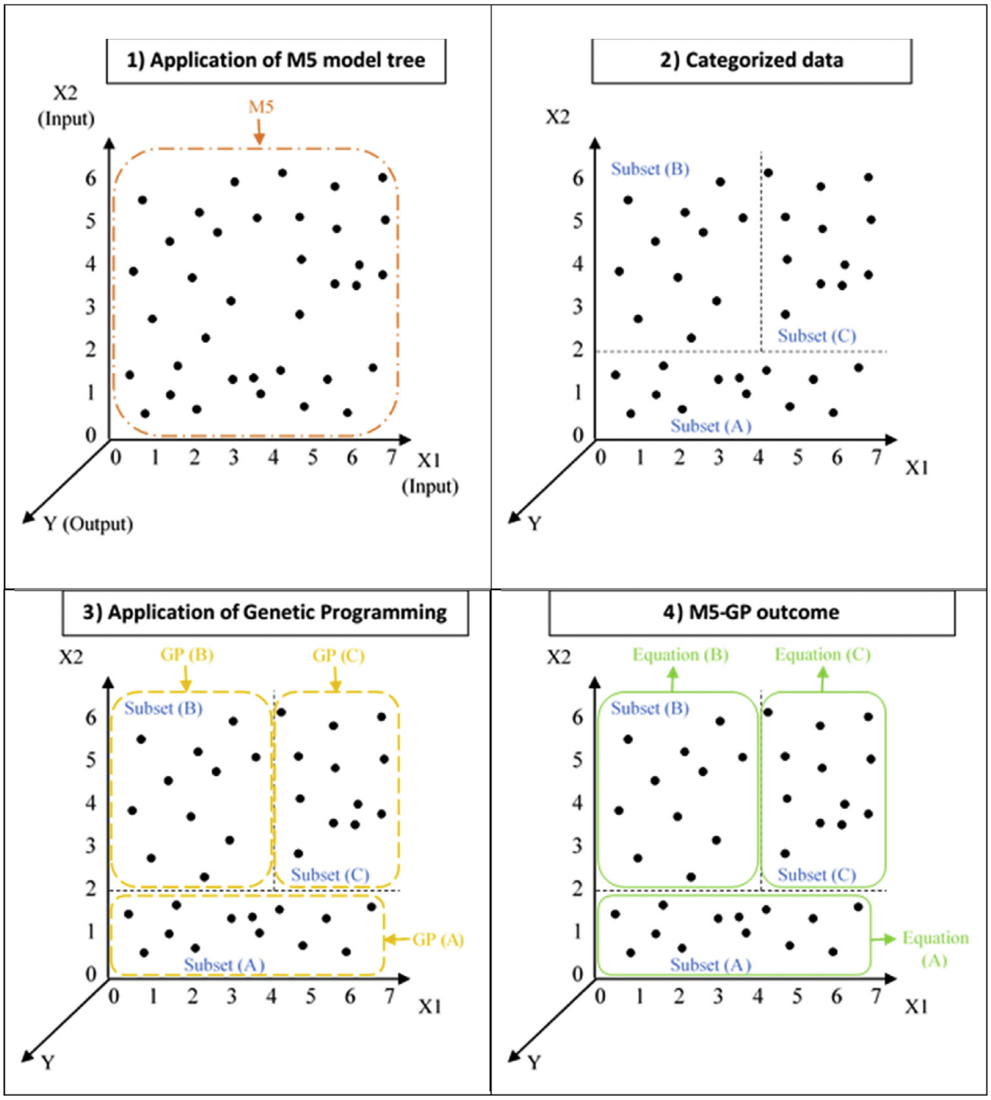

Fig. 2. Procedures of hybrid M5-GP method.

noteworthy that such analyses are just valid for the range of different parameters in the selected database.

The measurements of effective parameters such as L/d, D/L, θ, S_u, T_k, and Q are available in the database. For the estimation of the uplift capacity, 51 data were picked up for the training of the algorithm and the remainder (11 data) were used to test the ability of the models for generalization. To this end, the testing set was never employed in the building of the model. The statistics of different inputs and output variables included in the model derivation are displayed in Table 1.

Table 1
Statistics of the model parameters.

	Minimum	Maximum	Standard deviation	Skewness	Average
Inputs					
L/d	0.23	4	0.77	1.09	1.59
D/L	0	0.69	0.17	2.89	0.06
θ (Rad)	0	$\pi/2$	0.6	−1.58	0.4π
S_u (kPa)	1.8	38	10	1.35	11.75
T_k	1.00E-05	0.04	0.01	4.25	2.20E-03
Output					
Q (kPa)	10.1	387.2	81.67	1.74	90.06

4.2. Classification of the data and derivation of the formulae

As explained above, at first, M5 is applied to the data and categorizes it, so the homogeneous subsets are created. The splitting parameter is the undrained shear strength of the soil at the caisson tip (S_u) and the splitting value is 12.28 kPa. Once the subsets are classified by the application of the M5 model, the data at each leaf is used for the estimation using the GP model. The significant input parameters are the terminal set of the GP model and the function set involves various mathematical operators to be tested for optimization of the GP model. As described before, several important criteria were considered for choosing the best formula among many possible solutions.

In this study, two models were proposed using the hybrid M5-GP approach. The main objective of developing the first model, called "M5-GP-1", was to present a simple formula which is both physically sound and easy to use. So, the terminal and function sets of the GP were chosen such that the final solution is not complex while it has a suitable accuracy. The equations of "M5-GP-1" model are as follows:

$$for\ \ S_u \leq 12.28 \qquad Q = 1.105\ \frac{\left(\frac{D}{L}+1\right)\left(\frac{\pi}{2}+\theta\right)^{0.8} S_u^{0.9}}{T_k^{0.1}} \tag{6-a}$$

Table 2
M values for different levels of risk.

Allowable risk (%)	M
2	2.05
5	1.65
10	1.28
33	0.44
50	0.00

$$for\ \ S_u > 12.28 \qquad Q = 0.083\ \frac{\left(\frac{D}{L}+1\right)^{3.7}\left(\frac{\pi}{2}+\theta\right)^{2.8}S_u^{1.6}}{\left(\frac{L}{d}\right)^{0.2}\left(10^5T_k\right)^{13.1}} \tag{6-b}$$

The second model presented in this research, called "M5-GP-2", mainly accounts for the accuracy. Hence, the function set of the GP was arranged in a way that the prediction errors are minimized considerably, however, the configuration of the given formula is not so simple. The relationships of "M5-GP-2" are as below:

$$for\ \ S_u \le 12.28 \tag{7-a}$$

$$Q = 0.272\frac{\left(\frac{\pi}{2}+\theta\right)^{1.2}S_u}{T_k^{0.1}}\ \exp\left[1.2\left(\frac{D}{L}+1\right)+2.23T_k-0.0004S_u^2\exp\left(\frac{L}{d}\right)\right]$$

$$for\ \ S_u > 12.28 \tag{7-b}$$

$$Q = 11.964\ \frac{\left(\frac{D}{L}+1\right)^{0.2}\left(\frac{\pi}{2}+\theta\right)^{4.1}\left(10^5T_k\right)^{1.4}}{\left(\frac{L}{d}\right)^{1.6}S_u^{2.8}}\ \exp\left[0.36\left(\frac{L}{d}\right)+0.18S_u + 1.69\exp\left(\left(\frac{L}{d}\right)\Big/\left(\frac{\pi}{2}+\theta\right)\right)\right]$$

As can be seen, all the effective parameters are involved in the developed relationships. It is noteworthy that the simple configuration of the equations of the "M5-GP-1" model significantly facilitates the design practice. In addition, a parametric study can be easily performed for further verification of the model. According to the proposed equations, trend of the prediction due to variations of L/d, D/L, θ, S_u and T_k can be evaluated. It can be observed that the uplift capacity increases with the increase of D/L and S_u. When the load direction i.e. θ changes from horizontal to vertical condition, the uplift capacity enhances. Increase in L/d and T_k leads to the decrease of uplift capacity. The findings of parametric study are in good agreement with the results presented by Deng and Carter (2000); Rahman et al. (2001) and Wang et al. (2008).

The equations of the "M5-GP-2" model include an additional exponential term in comparison with the structure of the "M5-GP-1" model. In this way, contributions of input parameters are further considered and the simple initial form is modified. This improves the prediction ability of the relationships, however, the configuration becomes more complicated.

Reliability of the prediction of the uplift capacity is extremely important for design of suction caissons in practice. However, it is not possible to evaluate the uncertainty incorporated in the formulas and the failure risk of the suction caissons is not considered. In order to conquer this limitation, the proposed formulas were modified for probabilistic design utilizing the standard deviation of data. The "M5-GP-1" model was revised as:

$$for\ \ S_u \le 12.28 \qquad Q = (1.105+0.223M)\ \frac{\left(\frac{D}{L}+1\right)\left(\frac{\pi}{2}+\theta\right)^{0.8}S_u^{0.9}}{T_k^{0.1}} \tag{8-a}$$

$$for\ \ S_u > 12.28 \qquad Q = (0.083+0.026M)\ \frac{\left(\frac{D}{L}+1\right)^{3.7}\left(\frac{\pi}{2}+\theta\right)^{2.8}S_u^{1.6}}{\left(\frac{L}{d}\right)^{0.2}\left(10^5T_k\right)^{13.1}} \tag{8-b}$$

and the "M5-GP-2" model was modified as:

$$for\ \ S_u \le 12.28 \tag{9-a}$$

$$Q = (0.272+0.024M)\frac{\left(\frac{\pi}{2}+\theta\right)^{1.2}S_u}{T_k^{0.1}}\ \exp\left[1.2\left(\frac{D}{L}+1\right)+2.23T_k - 0.0004S_u^2\exp\left(\frac{L}{d}\right)\right]$$

$$for\ \ S_u > 12.28 \tag{9-b}$$

$$Q = (11.964+1.433M)\ \frac{\left(\frac{D}{L}+1\right)^{0.2}\left(\frac{\pi}{2}+\theta\right)^{4.1}\left(10^5T_k\right)^{1.4}}{\left(\frac{L}{d}\right)^{1.6}S_u^{2.8}}\ \exp\left[0.36\left(\frac{L}{d}\right) + 0.18S_u + 1.69\exp\left(\left(\frac{L}{d}\right)\Big/\left(\frac{\pi}{2}+\theta\right)\right)\right]$$

These equations can be used for various levels of allowable risk by substituting suitable values for M. Table 2 shows the M values obtained from the normal distribution curve for the different desired levels of risk. For instance, for the risk of 10%, M equals to 1.28.

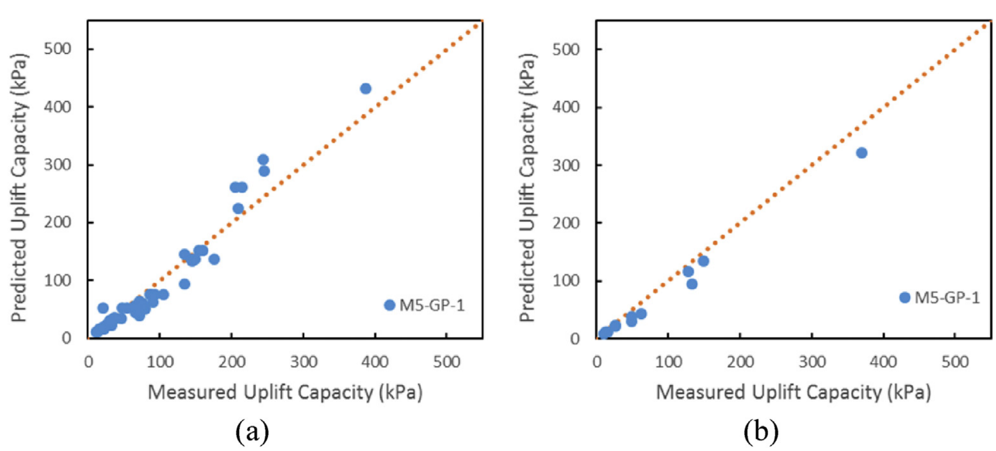

Fig. 3. Estimated versus observed uplift capacity by M5-GP-1: a) Training set, b) Testing set.

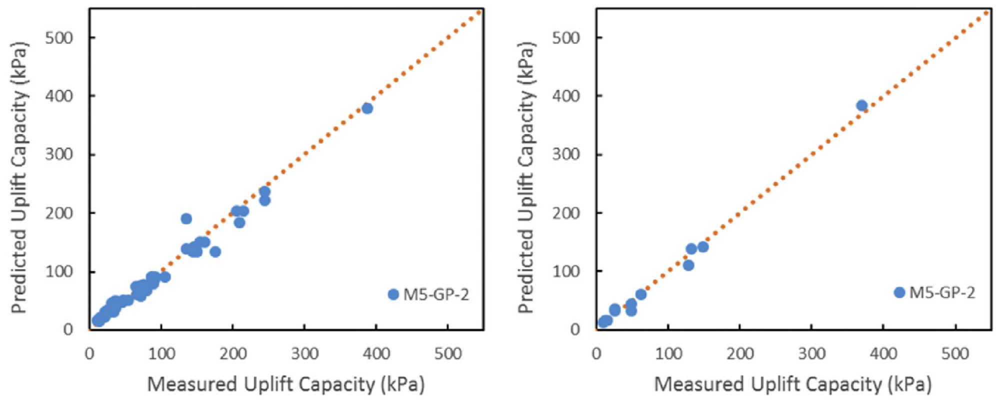

Fig. 4. Estimated versus observed uplift capacity by M5-GP-2: a) Training set, b) Testing set.

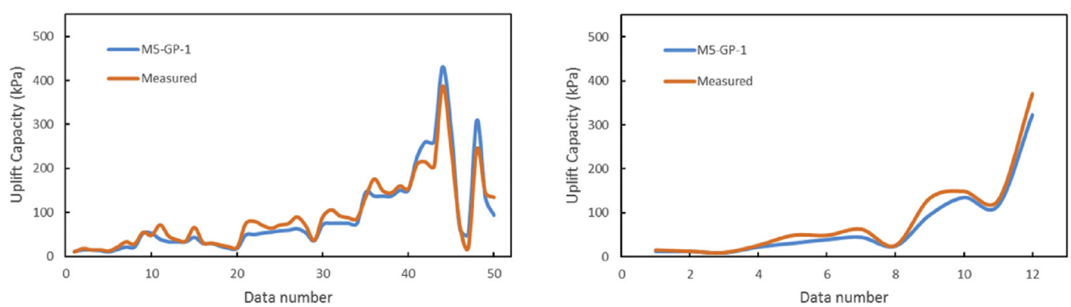

Fig. 5. Comparison of measured and predicted uplift capacity by M5-GP-1: a) Training set, b) Testing set.

5. Performance analyses

Two models were developed for the prediction of uplift capacity of suction caissons. Figs. 3 and 4 display estimated versus observed uplift capacity using M5-GP-1 and M5-GP-2 for training and testing data sets. The M5-GP-1 model gives correlation coefficient (CC) of 0.976 and 0.996 for the training and testing datasets, respectively. The correlation coefficient (CC) given by M5-GP-2 model for the training and testing datasets are 0.986 and 0.996.

Comparisons of the measured and predicted uplift capacity of suction caissons by the two M5-GP models are illustrated for the training and testing sets in Figs. 5 and 6. The models perform quite better on the testing data in contrast to the training data.

Explicit formulations proposed by different researchers for the prediction of upload capacity of suction caissons are summarized in Table 3 (Alavi et al., 2011; Gandomi et al., 2011; Muduli et al., 2013). As can be seen, various AI-based methods have been employed and in several cases the normalized values of the input parameters have been used.

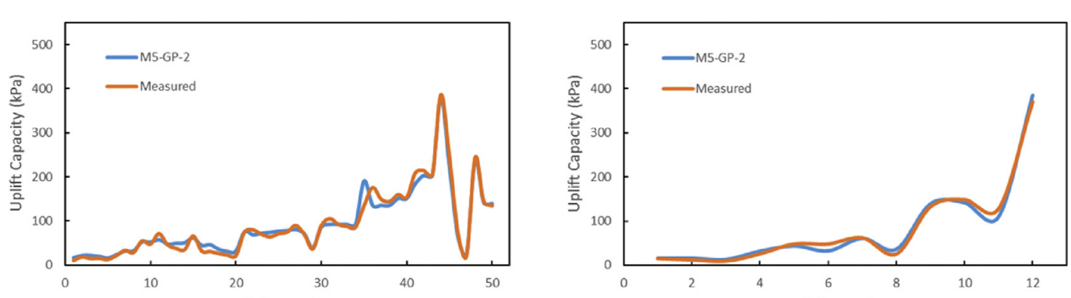

Fig. 6. Comparison of measured and predicted uplift capacity by M5-GP-2: a) Training set, b) Testing set.

Table 3
Recommended models for uplift capacity prediction.

Author	Method	Formula
Gandomi et al. (2011)	MEP	$Q_{MEP} = S_u\left(24\left(\frac{D}{L}\right)^2 + \frac{40}{L/d+10} - 160T_k + 4\theta\right)$
Alavi et al. (2011)	TGP	$Q_{TGP} = 406.56S_{u,n}\left(\left(\frac{D}{L}\right)_n + \theta_n - \left(T_{k,n} + S_{u,n}\left(S_{u,n} + \theta_n - \left(\frac{L}{d}\right)_n - 1\right)\right)\right)$
	LGP	$Q_{LGP} = 406.56S_{u,n}\left(\left(\left(\frac{D}{L}\right)_n^2 - 0.5\left(\frac{L}{d}\right)_n\right)^2 + \left(\frac{D}{L}\right)_n - 0.25\left(\frac{L}{d}\right)_n + 0.25S_{u,n} + 0.75\theta_n - 0.25T_{k,n} + 0.125\right)$
	GEP	$Q_{GEP} = 406.56Log(e^{S_{u,n}})\left(e^{\left(\frac{D}{L}\right)_n} + S_{u,n}\left(\theta_n - e^{T_{k,n}+\left(\frac{L}{d}\right)_n} + e^{\left(\frac{D}{L}\right)_n}\right) + \theta_n\right)$
	MLSR	$Q_{MSLR} = -4.56\left(\frac{L}{d}\right) + 8.83S_u + 749.63T_k + 1.77\theta + 304.36\left(\frac{D}{L}\right) - 148.92$
Muduli et al. (2013)	GP	$Q_{GP} = 8.739S_u\left(\frac{D}{L}\right) - 0.062(S_u + 11.850)\left(\left(\frac{L}{d}\right) - S_u\right) - (8.739 + 1.595S_uLog(T_k))\tanh(\theta) + 11.46$

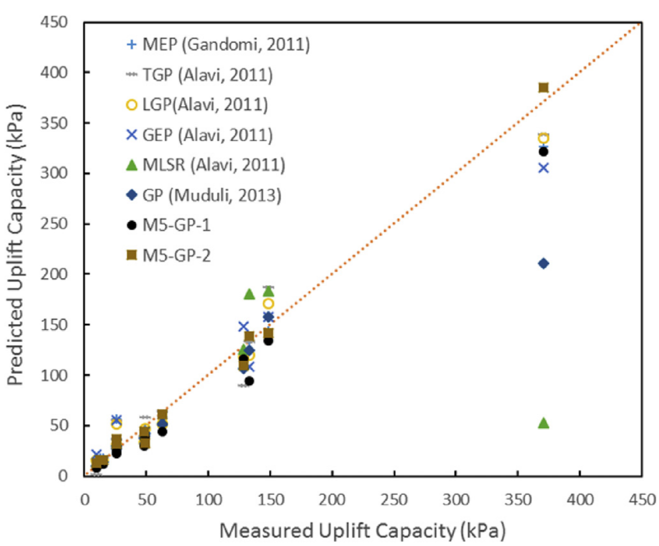

Fig. 7. Comparison of the uplift capacity obtained by different approaches.

Predictions made by MEP, TGP, LGP, GEP, MLSR and GP for the testing dataset are demonstrated in Fig. 7. The prediction capabilities of different models on the testing dataset are provided in Table 4. As shown in Fig. 7 and Table 4, the "M5-GP-1" and "M5-GP-2" models are able to predict the desired values accurately.

It can be seen that the M5-GP based solutions yield the best results among different methods. The two M5-GP models have the similar values of correlation coefficient (CC), however, "M5-GP-2" performs better than "M5-GP-1" in terms of errors i.e. the RMSE and MAE.

The CC values in Table 4 show that the generalization capability of M5-GP models is better than all other available models. The next models exhibiting well performance are MEP, LGP, GEP, TGP, GP and MLSR respectively. Among these, the MLSR is weaker due to significant drawbacks of regression techniques. Regarding the RMSE and MAE values, the "M5-GP-2" performs well followed by MEP, LGP and "M5-GP-1".

Along with suitable performance, M5-GP based formulae have some other important advantages. The equations are relatively simple, physically sound and easy to use in engineering practice. However, ANN (Rahman et al., 2001), NGN (Pai, 2005), IFRIM (Cheng et al., 2014) and GMDH-HS (Shahr-Babak et al., 2016) are black-box models that do not present a transparent relationship between the output and inputs. In addition, these methods may need prior adjustment of settings. The M5-GP equations can uniformly handle data of any conditions while the FEM models are sensitive to the individual cases (Alavi et al., 2010).

6. Conclusions

A combination of model tree (M5) and genetic programming (GP) called hybrid "M5-GP" model was used to predict the uplift capacity of suction caissons. A well-known database of the uplift capacity experimental results was employed to develop the model. Uplift capacity, is a function of the most effective corresponding geotechnical and structural parameters.

The findings regarding the employment of M5-GP and proposed relationships are concluded as follows:

(i) The new equations developed by the hybrid model are easy to use due to simple configuration and they are transparent because the influence of input parameters on the output is physically interpretable. Hence, the proposed formulae are suitable for practical applications. Additionally, the equations were modified to account for different levels of risk to be considered in engineering practice.

(ii) Different from FEM analysis, M5-GP model can systematically evaluate the uplift capacity in all various conditions of the governing parameters. Effects of the predictors on the uplift capacity were assessed and it was found that the new model can reasonably reflect the influence of the variations of the input parameters. It is noteworthy that the M5-GP is capable of giving more than one model for a complicated system using different terminal and function sets included in genetic programming process of its algorithm.

(iii) The high accuracy of the proposed relationships by the M5-GP model derived from the analysis of the data, is confirmed by

Table 4
Performance indices of different models for uplift capacity prediction.

	TGP (Alavi et al., 2011)	LGP (Alavi et al., 2011)	GEP (Alavi et al., 2011)	MLSR (Alavi et al., 2011)	MEP (Gandomi et al., 2011)	GP (Muduli et al., 2013)	M5-GP-1 (Current study)	M5-GP-2 (Current study)
CC	0.979	0.990	0.983	0.631	0.994	0.946	0.996	0.996
RMSE	20.79	16.03	23.49	152.45	16.14	46.89	20.57	9.51
MAE	15.56	12.21	16.18	131.06	9.86	19.59	14.58	7.64

obtaining the appropriate values of correlation coefficient (CC), root mean squared error (RMSE) and mean absolute error (MAE).

(iv) The estimations made by derived formulas were compared with the results calculated by different available models representing explicit formulations such as TGP, LGP, GEP, MEP and GP. It was inferred that the new equations proposed by M5-GP model provides reliable predictions of the uplift capacity of suction caissons. The new predictive models perform better than or similar to the previously presented models.

References

Ahn, J., Lee, H., Kim, Y.-T., 2015. Finite element analysis of the holding capacity of shallow suction caisson anchors. Mar. Georesources Geotechnol. 33 (1), 33–44.

Alavi, A.H., Aminian, P., Gandomi, A.H., Esmaeili, M.A., 2011. Genetic-based modeling of uplift capacity of suction caissons. Expert Syst. Appl. 38 (10), 12608–12618.

Alavi, A.H., Gandomi, A.H., Mousavi, M., Mollahasani, A., 2010. High-precision modeling of uplift capacity of suction caissons using a hybrid computational method. Geomechanics Eng. 2 (4), 253–280.

Aubeny, C., Murff, J.D., 2005. Simplified limit solutions for the capacity of suction anchors under undrained conditions. Ocean. Eng. 32 (7), 864–877.

Azamathulla, H.M., Guven, A., Demir, Y.K., 2011. Linear genetic programming to scour below submerged pipeline. Ocean. Eng. 38 (8–9), 995–1000.

Banzhaf, W., Nordin, P., Keller, R.E., Francone, F.D., 1998. Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and Its Applications. dpunkt–Verlag fur digitale. Technologie GbmH and Morgan Kaufmann Publishers. Inc., Heidelberg and San Francisco CA, resp.

Bhattacharya, B., Price, R.K., Solomatine, D.P., 2007. Machine learning approach to modeling sediment transport. J. Hydraulic Eng. 133 (4), 440–450.

Bonakdar, L., Oumeraci, H., Etemad-Shahidi, A., 2015. Wave load formulae for prediction of wave-induced forces on a slender pile within pile groups. Coast. Eng. 102, 49–68.

Chen, W., Randolph, M.F., 2007. External radial stress changes and axial capacity for suction caissons in soft clay. Géotechnique 57, 499.

Cheng, M.-Y., Cao, M.-T., Tran, D.-H., 2014. A hybrid fuzzy inference model based on RBFNN and artificial bee colony for predicting the uplift capacity of suction caissons. Automation Constr. 41, 60–69.

Cho, Y., Lee, T., Park, J., Kwag, D., Chung, E., Bang, S., 2002. Field Tests on Suction Pile Installation in Sand, ASME 2002 21st International Conference on Offshore Mechanics and Arctic Engineering. American Society of Mechanical Engineers, pp. 765–771.

Clukey, E., Banon, H., Kulhawy, F., 2000. Reliability assessment of deepwater suction caissons, Offshore Technology Conference. In: Offshore Technology Conference.

Colliat, J., Dendani, H., 2002. Girassol: geotechnical design analyses and installation of the suction anchors, offshore site investigation and geotechnics' diversity and sustainability. In: Proceedings of an International Conference. Society of Underwater Technology.

Datta, M., Kumar, P., 1996. Suction Beneath Cylindrical Anchors in Soft Clay, the Sixth International Offshore and Polar Engineering Conference. International Society of Offshore and Polar Engineers.

Deng, W., Carter, J., 1999. Vertical Pullout Behaviour of Suction Caissons. Research Re~Ort, Centre for Geotechnical Research. The University of Sydney.

Deng, W., Carter, J., 2000. Uplift capacity of suction anchors in uniform soils. Proc. GeoEng.

Deng, W., Carter, J., 2002. A theoretical study of the vertical uplift capacity of suction caissons. Int. J. Offshore Polar Eng. 12 (02).

Dyvik, R., Andersen, K.H., Hansen, S.B., Christophersen, H.P., 1993. Field tests of anchors in clay. I: description. J. Geotechnical Eng. 119 (10), 1515–1531.

Ehlers, C.J., Young, A.G., Chen, J.-H., 2004. Technology assessment of deepwater anchors, Offshore technology conference. In: Offshore Technology Conference.

El-Gharbawy, S., Olson, R., 2000. Modeling of suction caisson foundations, the tenth international offshore and polar engineering conference. Int. Soc. Offshore Polar Eng.

Etemad-Shahidi, A., Bali, M., 2012. Stability of rubble-mound breakwater using H50 wave height parameter. Coast. Eng. 59 (1), 38–45.

Etemad-Shahidi, A., Ghaemi, N., 2011. Model tree approach for prediction of pile groups scour due to waves. Ocean. Eng. 38 (13), 1522–1527.

Etemad-Shahidi, A., Yasa, R., Kazeminezhad, M.H., 2011. Prediction of wave-induced scour depth under submarine pipelines using machine learning approach. Appl. Ocean Res. 33 (1), 54–59.

Gandomi, A.H., Alavi, A.H., Yun, G.J., 2011. Formulation of uplift capacity of suction caissons using multi expression programming. KSCE J. Civ. Eng. 15 (2), 363–373.

Gao, Y., Qiu, Y., Li, B., Li, D., Sha, C., Zheng, X., 2013. Experimental studies on the anti-uplift behavior of the suction caissons in sand. Appl. Ocean Res. 43, 37–45.

Guo, Z., Wang, L., Yuan, F., Li, L., 2012. Model tests on installation techniques of suction caissons in a soft clay seabed. Appl. Ocean Res. 34, 116–125.

Guven, A., Azamathulla, H.M., Zakaria, N.A., 2009. Linear genetic programming for prediction of circular pile scour. Ocean. Eng. 36 (12–13), 985–991.

Hogervorst, J., 1980. Field trails with large diameter suction piles, offshore technology conference. In: Offshore Technology Conference.

Koza, J.R., 1992. Genetic Programming: on the Programming of Computers by Means of Natural Selection. MIT press.

Li, D., Feng, L., Zhang, Y., 2014. Model tests of modified suction caissons in marine sand under monotonic lateral combined loading. Appl. Ocean Res. 48, 137–147.

Li, D., Zhang, Y., Feng, L., Gao, Y., 2015. Capacity of modified suction caissons in marine sand under static horizontal loading. Ocean. Eng. 102, 1–16.

Muduli, P.K., Das, M.R., Samui, P., Kumar Das, S., 2013. Uplift capacity of suction caisson in clay using artificial intelligence techniques. Mar. Georesources Geotechnol. 31 (4), 375–390.

Pai, G.A.V., 2005. Prediction of uplift capacity of suction caissons using a neuro-genetic network. Eng. Comput. 21 (2), 129–139.

Quinlan, J.R., 1992. Learning with Continuous Classes, 5th Australian Joint Conference on Artificial Intelligence. Singapore, pp. 343–348.

Rahman, M.S., Wang, J., Deng, W., Carter, J.P., 2001. A neural network model for the uplift capacity of suction caissons. Comput. Geotechnics 28 (4), 269–287.

Randolph, M., O'neill, M., Stewart, D., Erbrich, C., 1998. Performance of suction anchors in fine-grained calcareous soils, Offshore technology conference. In: Offshore Technology Conference.

Rao, S.N., Latha, K.H., Pallavi, B., Surendran, S., 2006. Studies on pullout capacity of anchors in marine clays for mooring systems. Appl. Ocean Res. 28 (2), 103–111.

Samui, P., Das, S., Kim, D., 2011. Uplift capacity of suction caisson in clay using multivariate adaptive regression spline. Ocean. Eng. 38 (17–18), 2123–2127.

Schneider, J.A., Senders, M., 2010. Foundation design: a comparison of oil and gas platforms with offshore wind turbines. Mar. Technol. Soc. J. 44 (1), 32–51.

Senpere, D., Auvergne, G.A., 1982. Suction anchor piles-a proven alternative to driving or drilling, Offshore Technology Conference. In: Offshore Technology Conference.

Shahnazari, H., Shahin, M.A., Tutunchian, M.A., 2014. Evolutionary-based approaches for settlement prediction of shallow foundations on cohesionless soils. Int. J. Civ. Eng. 12 (1), 55–64.

Shahr-Babak, M.M., Khanjani, M.J., Qaderi, K., 2016. Uplift capacity prediction of suction caisson in clay using a hybrid intelligence method (GMDH-HS). Appl. Ocean Res. 59, 408–416.

Solomatine, D.P., Ostfeld, A., 2008. Data-driven modelling: some past experiences and new approaches. J. Hydroinformatics 10 (1), 3–22.

Sukumaran, B., McCarron, W.O., Jeanjean, P., Abouseeda, H., 1999. Efficient finite element techniques for limit analysis of suction caissons under lateral loads. Comput. Geotechnics 24 (2), 89–107.

Tjelta, T., 1995. Geotechnical experience from the installation of the europipe jacket with bucket foundations, offshore technology conference. In: Offshore Technology Conference.

Wang, D., Hu, Y., Randolph, M.F., 2008. Effect of loading rate on the uplift capacity of plate anchors, the eighteenth international offshore and polar engineering conference. Int. Soc. Offshore Polar Eng.

Wang, Y., Witten, I.H., 1996. Induction of Model Trees for Predicting Continuous Classes.

Whittle, A.J., Germaine, J.T., 1998. Behavior of miniature suction casissons in clay, offshore site investigation and foundation behaviour'new frontiers. In: Proceedings of an International Conference. Society of Underwater Technology.

Zdravkovic, L., Potts, D.M., Jardine, R.J., 2001. A parametric study of the pull-out capacity of bucket foundations in soft clay. Géotechnique 51 (1), 55–67.

Zeinoddini, M., Keyvani, J., Nabipour, M., 2011. Wall slope effects on the vertical pull-out capacity of tapered suction caissons. Sci. Iran. 18 (3), 313–325.