Evaluation of Pile Lateral Capacity in Clay Applying Evolutionary Approach

Alkroosh I. and Nikraz H.
Curtin University, Perth, Australia

ABSTRACT

This paper presents the development of a new model to predict the lateral capacity of piles inserted into clayey soils and subjected to lateral loads. Gene Expression Programming (GEP) has been utilized for this purpose. The data used for development of the GEP model is collected from the literature and comprise 38 data points. The data are divided into two subsets: Training set for model calibration and independent validation set for model verification. Predictions from the GEP model are compared with the results of experimental data. The model has achieved a coefficient of correlation, r, of 0.95 for training and validation sets and average prediction ratio (APR) of 0.97 and 1.04 for training and validation sets respectively. The results indicate that the GEP model performs very well and able to predict the pile lateral capacity accurately.

Keywords: Pile; Capacity; GEP; Training and validation

1. INTRODUCTION

Geotechnical engineers often recommend piles as foundations to support the proposed superstructure subjected to lateral loads. Hence, pile lateral carrying capacity is required to be evaluated.

Several researchers have attempted evaluation of pile lateral capacity based on analytical solutions [e.g. 1], semiempirical solutions [e.g. 2, 3] and finite element solutions [e.g. 4]. However, because of the nonlinearity of the soil behaviour and the variability of soil properties, the proposed methods have achieved limited success in terms of giving accurate prediction of pile lateral capacity.

In this respect, artificial intelligence techniques may be more efficient. Recently, several attempts have been made to use artificial neural networks (ANNs) for modeling the axial capacity of pile foundations [e.g. 5, 6, 7] and lateral capacity [e.g. 8]. The modelling advantage of ANNs over traditional methods is the ability of ANNs to capture the nonlinear and complex relationship between the bearing capacity and the factors affecting it without having to assume a priori formula of what could be this relationship. However, the main shortcoming of ANNs is the large complexities of the network structure, as it represent the knowledge in terms of weight matrices together with biases that are not accessible to the users [9].

In this regard, the genetic programming (GP) may represent better alternative. The main advantage of the GP over the ANNs is the ability to provide the relationship between a set of inputs and the corresponding outputs in a simple mathematical form accessible to the users. Recently, the GP has been found successful in solving several problems in the field of engineering [e.g. 10, 11].

In this paper, the lateral capacity of piles in clayey soils has been correlated with undrained shear strength and load eccentricity using a developed version of genetic programming that is gene expression programming (GEP). Recently, GEP has been applied with success in solving engineering problems [e.g. 12, 13, 14, 15, 16a, 16b]

The objectives of this paper:
1. Applying the GEP technique for modelling the lateral load capacity of pile foundations embedded in clayey soils.
2. Evaluating the performance of the GEP model by comparing its predictions with experimental data.
3. Measuring the accuracy of the GEP model via statistical analysis.

2. OVERVIEW OF GENE EXPRESSION PROGRAMMING

Gene Expression Programming is an instance of an Evolutionary Algorithm from the field of Evolutionary Computation, invented by Ferreira [17] as a global optimization algorithm. It has similarities to other Evolutionary Algorithms such as the Genetic Algorithm as well as other Evolutionary Automatic Programming techniques such as Genetic Programming. Similar to the GAs, the GEP utilizes evolution of computer programs (individuals or chromosomes) that are encoded linearly in chromosomes of fixed length and likewise the GP the evolved programs are expressed nonlinearly in a form of expression trees (ETs) of different sizes and shapes. However, the GEP implements different evolutionary computational method.

The GEP distinct itself from GAs in that the evolved solutions are expressed in forms of parse trees of different sizes and structures and unlike GP genetic variations are performed on chromosomes before they are translated into ETs. The GEP chromosomes are composed of multiple genes, each gene is encoded a smaller sub-program. Every gene has a constant length and includes a head that contains functions and

terminals, and a tail that composes of terminals only. The genetic code represents a one-to-one relationship between the symbols of the chromosome; the functions or terminals. The process of information decoding from chromosomes to expression trees is called translation which is based on sets of rules that determine the spatial organization of the functions and terminals in the ETs and the type of interaction (link) between the sub-ETs [17].

The modelling process of GEP begins with random generation of chromosomes of initial population. Each individual chromosome is expressed and its fitness is evaluated through the fitness function which measures how good the individual is at competition with the rest of the population. The best individuals are kept for modifications which are performed by the genetic operators such as mutation and recombination. New offspring of chromosomes with new traits are generated and used to replace the existing population. The individuals of the new generation are then subjected to the same developmental process which is repeated until stopping criteria are satisfied

3. DEVELOPMENT OF THE GEP MODEL

In this work, the GEP model is developed using the commercial available software package *GeneXproTools* 4.0 [18]. The data used for the model development are collected from the literature and comprise 38 data points of piles inserted in clayey soil reported by Rao and Kumar [19] and found in Das and Basudhar [8]. The piles have different sizes with diameters ranging from 6.35 mm to 25.4 mm and lengths from 130 mm to 300 mm. In order to accurately predict the pile lateral capacity, the significant factors that influence the capacity need to be identified and presented to the GEP as input variables. These include the pile geometry, load leaver arm and soil properties. The pile geometry is represented by the pile diameter, D, and pile embedment length, L. The load leaver arm is represented by the eccentricity, e. The soil properties are represented by the undrained shear strength, $S_{u.}$. The lateral pile capacity, Q_u, is the single output.

3.1 Data Division

The next step in development of the GEP model is the data division. In this work, the data are randomly divided into two statistically consistent sets, as recommended by Masters [20] and detailed by Shahin et al. [21]. This includes a training set for model calibration and an independent validation set for model verification. In total, 29 data points (75%) of the available 38 data points were used for training and 9 data points (25%) for validation. The statistics of the data used for the training and validation sets are presented in Table 1, which includes the mean, standard deviation, maximum, minimum and range. It should be noted that, like all empirical models, GEP performs best in interpolation rather than extrapolation, thus, the extreme values of the data used were included in the training set.

Table 1 GEP model input and output statistics

Model variable and data sets	Statistical parameters				
	Mean	SD*	Max.*	Min.*	Range
Pile diameter, *D* (mm)					
Training	18	7	33	6	27
Validation	17	4	25	12	13
Pile embedment length, *L* (mm)					
Training	282	50	300	130	170
Validation	269	63	300	132	168
Load eccentricity, *e* (mm)					
Training	45	14	50	0	50
Validation	44	17	50	0	50
Undrained shear strength, *Su* (kPa)					
Training	10	10	39	3	35
Validation	10	12	39	3	35
Pile lateral capacity, *Qu* (N)					
Training	77	40	225	30	196
Validation	62	29	128	35	93

* SD: Standard deviation; Max: Maximum; Min: Minimum

3.2 Modelling Attempts

The success of the modelling process using GEP technique depends significantly on the design of the model structure. In this, the optimal model parameters are determined to ensure that the best performing model is achieved. In the search for a model using the GEP, the number of chromosomes, chromosome structure, functional set, fitness function, linking function and rates of genetic operators play important role during modelling process and choosing suitable rates of these parameters can reduce modelling time and effort and produce a robust solution.

In this work, the trial-and-error approach was used to determine the values of setting parameters. This approach involved using different settings and conducting runs in steps. During each step, runs were carried out and the values of one of the above mentioned parameters (with its optimal value being searched) were varied, whereas the values of the other parameters were set constant (i.e. number of chromosomes = 30, number of genes = 3, gene's head size = 8, functions set = +, -, ×, and /, fitness function = mean squared error (MSE), linking function = +, mutation rate = 0.04, and gene recombination rate = 0.1). The runs were stopped after fifteen thousand generations, which were found sufficient to evaluate the fitness of the output. At the end of each run, the MSE for both training and validation sets were recorded in order to identify the values that give the least MSE. The search attempts for optimal parameters values are presented in Table 2.

3.3 Model Formulation

As mentioned earlier, one of the advantages of the GEP is that it presents the relationship between the input and output in a form of expression trees as shown in Fig. 1.

Table 2 Input parameters used for developing GEP model

Parameter	Used input
Number of chromosomes	15, 16, 17, …30
Number of genes	1, 2, 3
Head size	7, 8, 9, …12
Function set	$+, -, \times, \div, \sqrt{\ }, \sqrt[3]{\ }, \sqrt[4]{\ }$, Power
Fitness function	MSE (Mean Squared Error)
Linking function	+, ×
Mutation rate	0.045, 0.05, 0.055, …, 0.08
Recombination rate	0.1, 0.2, 0.3, …, 0.7

As can be seen, the figure illustrates the mathematical operations and interactions between the components of the solution. This can give insight to the nature of the relationship between the input and the output. The trees can be easily translated and arranged into mathematical expression as follows:

$$Q_p = 2D + e + \frac{L}{\sqrt[2]{e} + \sqrt[3]{\frac{s_u^2}{L}}} - \frac{e^2}{D(D+9.8)} - \frac{L}{S_u} \qquad (1)$$

where;

Q_p : predicted pile lateral capacity; D: pile diameter; L: embedment depth; e: eccentricity; S_u: undrained shear strength.
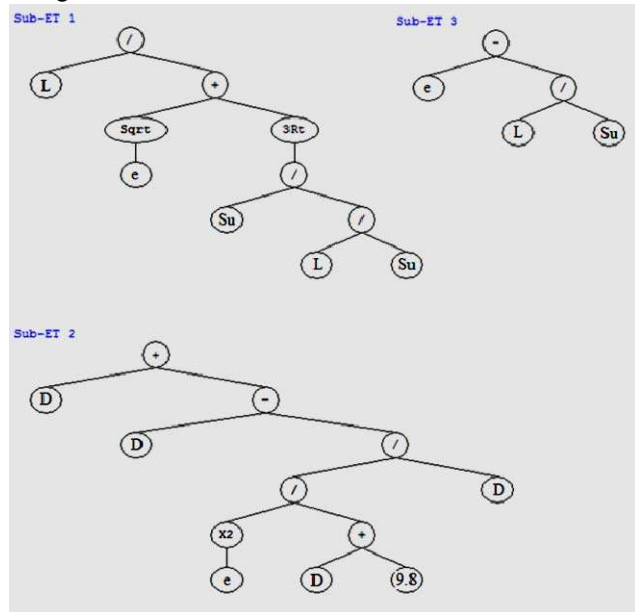

Fig. 1 Expression trees of the developed GEP model; Sqrt = square root; 3Rt = cubic root; X^2 = to power 2

4. RESULTS AND MODEL VALIDATION

The performance of the optimum GEP model is shown numerically in Table 3 and is depicted graphically in Fig. 2. It can be seen from Table 3 that the model performs well with high coefficients of correlation, r, of 0.95 for the training and validation sets. It can also be seen that the model has good average prediction ratios, APR, of 0.97 and 1.04 for the training and validation sets, respectively. The APR is calculated from

$$APR = \sum_{i=1}^{n} \left(\frac{Q_p}{Q_m} \right) / n \qquad (2)$$

where;

Q_p: predicted capacity; Q_m: measured capacity and n: the number of case records. Fig. 2 also indicates that the model has minimum scatter around the line of equality between the measured and predicted pile capacities for the training and validation sets. The results demonstrate that the developed GEP model performs well and provides accurate predictions.
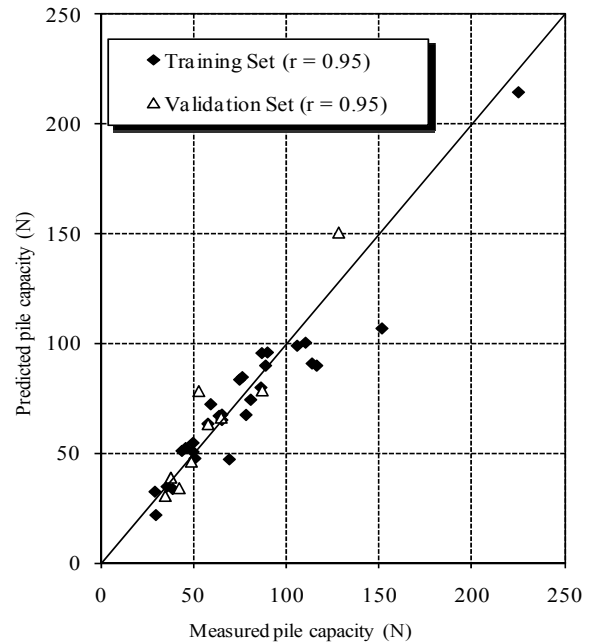

Fig. 2 GEP model performance in training and validation sets

Table 3 Numerical evaluation of the GEP model performance

Performance measure	Data set	
	Training	Validation
Correlation coefficient, r	0.95	0.95
Average prediction ratio, APR	0.97	1.04

5. CONCLUSION

The results of this study indicate that the GEP model possess a good capability in predicting the lateral capacity of piles embedded into clayey soils; the model has achieved high coefficients of correlation, r, of 0.95 for the data used in model calibration and validation. The model has also low average prediction ratio, APR, values of 0.97 and 1.04 for the data used in model calibration and validation, respectively; these values indicate that the model may tend to under-predict the pile lateral capacity. The results also demonstrate that GEP model performs well in comparison with the experimental data. Overall, the output of this study has demonstrated that resulting model correlates pile lateral capacity and undrained shear strength of soil accurately.

2 REFERENCES

[1] Poulos HG, and Davis EH, Pile Foundation Analysis & Design. New York: Wiley, 1980.

[2] Hansen B, "The ultimate resistance of rigid piles against transversal force," Danish Geotechnical Institute, Bulletin No 12. 1961, pp. 5-9.

[3] Brom BB, "Lateral resistance of piles in cohesive soils," J. Soil Mech Found Eng., ASCE, 90 (SM2), 1964, pp. 27-63.

[4] Portugal JC, and Sceo Pinto PS, "Analysis and design of pile under lateral loads", in Proc. 11th Int. Conf. on geotechnical seminar on deep foundation on bored and auger piles, Belgium, 1993, pp. 309-13.

[5] Abu-Kiefa M, "General regression neural networks for driven piles in cohesionless soils," J. of Geotechnical & Geoenvironmental Engineering, vol. 124 (12), 1998, pp. 1177-1185.

[6] Chan WT, Chow YK, Liu LF, "Neural network: an alternative to pile driving formulas," J. Computer & Geotechnics, vol. 17, 1995, pp. 135-56.

[7] Shahin M, 2010. "Intelligent computing for modeling axial capacity of pile Foundations," Canadian Geotechnical Journal, vol. 47 (2), 2010, pp. 230-243.

[8] Das SK, and Basudhar PK, "Undrained Lateral Load Capacity of Piles in Clay Using Artificial Neural Networks," Computers and Geotechnics, vol. 33 (8), 2006, pp. 454-459.

[9] Rezania M, and Javadi A, "A new genetic programming model for predicting settlement of shallow foundations," Canadian Geotechnical J., vol. 44 (12), 2007, 1462-1473.

[10] Javadi AA, Rezania M, Nezhad MM, "Evaluation of liquefaction induced lateral displacements using genetic programming," Computers and Geotechnics, vol. 33 (4-5), 2006, pp. 222-233.

[11] Bayksoglu A, Gullu H, Canakci H, Ozbakir L, "Prediction of compressive and tensile strength of limestone via genetic programming," Expert Syst. Appl. vol. 35 (1-2), 2008, pp. 111–123.

[12] Cevic A, and Cabalar AF, "Modelling damping ratio and shear modulus of sandmica mixtures using genetic programming," Expert Syst. Appl., vol. 36 (4), 2009, pp. 7749-7757.

[13] Alkroosh I, Shahin M, Nikraz H, "Modelling axial capacity of bored piles using genetic programming technique," in Proc. GEO-CHIANGMIA Int. Conf., Thailand, 2008.

[14] Alkroosh I, Shahin M, Nikraz H, "Genetic programming for predicting axial capacity of driven piles," in Proc. 1st Int. Symp. on Computational Geomechanics. Cote d'Azur, France, 2009.

[15] Alkroosh I, and Nikraz H, "Correlation of pile axial capacity and CPT data using gene expression programming," Geotechnical and Geological J., 2011a. DOI: 10:1007/s10706-011-9413-1).

[16] Alkroosh I, and Nikraz H, "Predicting axial capacity of driven piles in cohesive soils using intelligent computing," Engineering Applications of Artificial Intelligence, vol. (25), 2011b, pp. 618-627

[17] Ferreira C. Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence. Portugal, Angra do Heroismo, 2002.

[18] Gepsoft, Gene Expression Programming Tool, 2002, http://www.gepsoft.com.

[19] Rao KM, Suresh Kumar V, "Measured and predicted response of laterally loaded piles," in Proc. 6th Int. Conf. and exhibition on piling and deep foundation. India, 1996, 1.6.1-1.6.7

[20] Master T, Practical neural network recipes in C++. In Academic Press. San Diego, California, 1993.

[21] Shahin M, Maier H, Jaska M, 2004. "Data division for developing neural networks applied to geotechnical engineering," J. Computing in Civil Eng., vol. 18 (2), 2004, pp. 105-114.

Corresponding Author: **Alkroosh I.**