

# Prediction of Settlement of Shallow Footings on Granular Soils using Genetic Algorithm

Prasanth S<sup>1</sup> and N. Sankar<sup>2</sup>

<sup>1</sup>Post Graduate student in Civil Engineering  
National Institute of Technology Calicut  
*prasanththejus@gmail.com*

<sup>2</sup>Professor in Civil Engineering  
National Institute of Technology Calicut  
*sankar@nitc.ac.in*

---

**ABSTRACT**— *In this paper, a genetic algorithm approach is used for predicting the settlement of shallow foundations on granular soils. The development and verification of the genetic model was done using a large database containing about 198 case histories from various published literatures. The results of the model obtained were compared with various empirical equations available for measuring settlement. The correlation of predicted data with actual field measurements was determined and it was found out that the genetic algorithm approach can be used to fit a function to set of experimental data with a high degree of accuracy. The equation for settlement thus obtained connecting settlement and its contributing factors can be used in predicting settlements for new cases that were not used for the development of the genetic model.*

**Keywords**— Foundation Settlement, Genetic Algorithm, Geotechnical Model, Granular Soils

---

## 1. INTRODUCTION

Settlement is an important parameter that controls the design of shallow foundations, especially when the footing width exceeds 1m. Excessive settlements can lead to serviceability problems which makes settlement a more important parameter when compared to the bearing capacity of soils in the design of foundations.

The various theoretical relations connecting settlement and its related factors have been derived based on theory of elasticity concept, assuming soil as homogeneous, elastic etc. But in actual case, soil stands much far away from this assumption, hence the settlements calculated based on these relationships cannot be used in the design of shallow foundations as there will be always a question on the safety of the structure to be constructed above the particular granular soil deposit. Determination of settlement of shallow foundations on sand is a highly complex problem due to the various uncertainties associated with the stress-strain history, compressibility and applied stress distribution of soil (Shahin et al. 2002). Moreover there are a lot of difficulties involved in sampling and determination of in-situ properties of sandy soil deposits, which makes the problem even more complex.

Taking into consideration the various complexities involved in the determination of settlement, researchers have started to model the settlement phenomenon using various approaches. A number of empirical, semi-empirical and numerical methods have been developed which focuses on predicting the settlement of shallow footings based on various case histories. It has been reported that more than 40 different methods are available for predicting foundation settlements in granular soils (Douglas. 1986). In situations where the theoretical and other methods fail to give satisfactory results, recently developed methods like neural networks and Genetic Algorithm have the potential to make reliable predictions of settlement based on adaptive learning. Genetic Algorithm proves to be a powerful tool in this context which can model the complex soil behavior and generate the output in the form of an equation which best fits the set of experimental data used. The equation connecting settlement and its related parameters can be used for making reliable predictions of settlement for all cases.

## 2. SETTLEMENT PREDICTION METHODS

Traditional methods available for predicting settlement considers the net applied pressure, foundation width and soil stiffness as important parameters governing settlement. The soil stiffness is indirectly measured using standard penetration test data. The other factors governing settlement are shape of foundation, depth of embedment of foundation as well as the depth of water table below the footing level. The effect of water table may be reflected in the SPT blow count value, but it is always better to consider the water table depth in formulating the equation for settlement.

Developments in computational software have led to the emergence of numerous data classification and pattern recognition techniques like fuzzy logic, artificial neural network, genetic algorithm etc. These techniques operate on large sets of data and learn complex model functions. The genetic algorithm model is found to be much more accurate than the various traditional methods available and is much easier to use (Rezania et al. 2007). Genetic algorithm has been applied to many geotechnical problems like liquefaction prediction (Bagheripour et al. 2012), slope stability analysis (Mc Combie et al. 2005), soil parameter identification (Levassuer et al. 2007), optimization of pile groups (Chan et al. 2009), prediction of pile bearing capacities (Momeni et al. 2014) etc with a remarkable degree of success. This paper presents a genetic algorithm approach for predicting foundation settlements on granular soils using a database of 198 case histories.

### 3. GENETIC ALGORITHM APPROACH

Genetic algorithm is an evolutionary computational method which makes use of the evolutionary mechanisms such as selection, crossover and mutation to search for a function that best fits a set of experimental data. A flow chart for the genetic algorithm procedure is shown in Figure. 1.

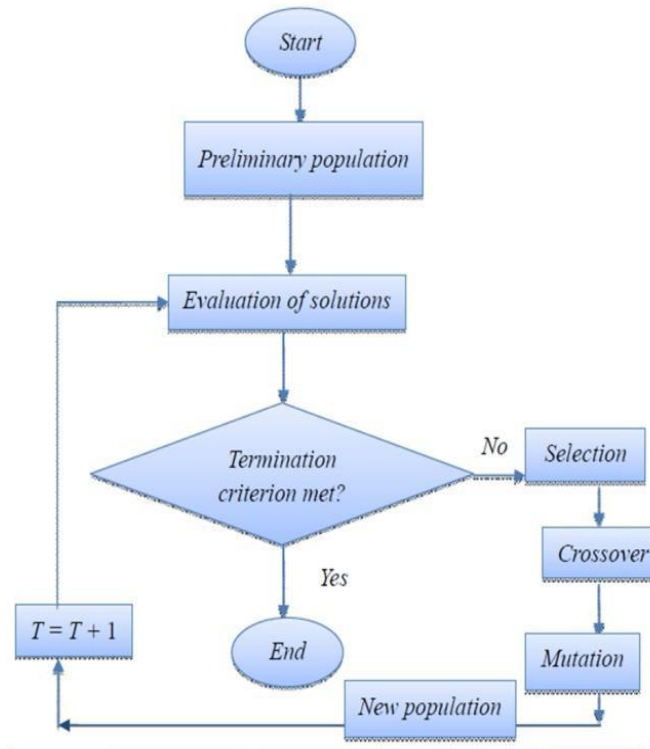


Figure 1: Genetic Algorithm flowchart

Genetic algorithm was applied for the prediction of settlement of shallow foundations on granular soils. The database consisted of 198 field measurement cases from various published literatures as shown in Table 1. The inputs to the model were Length of footing (L), width of footing (B), Depth of embedment of footing ( $D_f$ ), SPT N value (N), Depth of water table below the level of footing ( $H_w$ ) and the contact pressure at the base of the footing (q). These input variables were used by the model to predict settlement of shallow footings on granular soils. The database covers a wide range of measurements of footing dimensions, soil densities and various soil types.

Table 1: Database for Prediction

Reference	No: of cases
Bazaara, 1967	4
Briaud & Gibbens, 1999	4
Burbidge, 1982	22
Burland & Burbidge, 1985	134
Maugeri et al, 1998	2
Picornell & Delmonte, 1988	1
Wahls, 1997	31
Total	198

### 3.1 Data Division and Processing

The main advantage of the genetic model is that the data can be directly provided to the model without any normalization or calibration. Generally the model is trained by adaptive learning over number of cases from the training dataset and the performance of the model so developed is tested using another validation dataset which is completely new to the model.

The data are divided into training and testing datasets by using a statistically consistent approach. The data are randomly divided into training and testing datasets and statistical analysis was carried out. Statistical analysis measured the mean as well as standard deviations of the both the subsets. This random process was continued until a robust combination was achieved in which the mean and standard deviation of both the subsets were almost close to each other. Statistically consistent approach ensures that the statistical parameters of both the datasets are as close as possible and hence represent the same statistical population. Mean and standard deviation of the data were used for statistical analysis and the results of the statistical analysis are presented in table 2. Considering that the data contains events that cannot be replicated everywhere in the dataset, there may be still some minor inconsistencies in the statistical parameters of the two different subsets that are used for training and verification of the developed model. 158 cases (80%) of the data were used for training the model and the remaining 40 cases (20%) were used for testing the performance of the model.

**Table 2:** Statistical parameters of data subsets

Model Input	Statistical Parameter	Training set	Testing set
N-Value	Mean	24.31	21.23
	Standard Deviation	13.09	13.78
Breadth of footing (B)	Mean	9.61	9.03
	Standard Deviation	14.52	8.43
Length of Footing (L)	Mean	18.91	20.08
	Standard Deviation	25.57	36.66
Depth of Embedment ( $D_f$ )	Mean	2.19	2.14
	Standard Deviation	2.46	1.68
Height of water table ( $H_w$ )	Mean	2.86	3.90
	Standard Deviation	3.47	2.87
Contact pressure at footing (q)	Mean	192.47	171.26
	Standard Deviation	118.16	124.91
Settlement (S)	Mean	23.32	25.92
	Standard Deviation	34.84	34.03

## 4. FORMULATION OF THE GENETIC MODEL

An initial population of 1000 chromosomes was used for the development of the Genetic model. Each chromosome contained an array of variables and an array of operators. The variable array contained the coefficients and power terms of each input variables to the model. The coefficients of the variables were assigned a random number between 0 and 500 and the power terms were assigned a random number between -3 and +3. The operator array contained eleven slots, six of them for placing the input variables and the remaining five positions in the 11-bit array were allocated to place the arithmetic operators connecting these variable terms. The type of operator as well as the position of operator was randomly generated. Post fixing was done to develop a tree structure of input variables, leading to the generation of 1000 random equations for measuring settlement.

The input variables from training dataset were substituted in all the random equations generated to obtain the settlement. The settlement calculated was compared with the actual measured settlement to determine the error in prediction of settlement. The sum of squares of errors of all the data in training dataset was calculated for the 1000 randomly generated equations of settlement. This error was divided by the number of observations in training dataset and the reciprocal of that value was selected as fitness as shown in equation 1. S represents the actual settlement and  $S_c$  represents the settlement calculated by the randomly generated equations of settlement and N represents the number of observations in the dataset used for training the genetic algorithm model.

$$\text{Fitness} = 1 / \left[ \sum \frac{(S - S_c)^2}{N} \right] \quad (1)$$

#### 4.1 Selection

The randomly generated equations are sorted in descending order of their fitness values. The equations having higher values of fitness are selected and are carried forward to the next generation while the ones having lower fitness scores will die out. Of the various selection methods available, the roulette wheel selection method was adopted. The selected individuals go into the mating pool for crossover.

#### 4.2 Crossover

Crossover refers to the process of exchange of genetic materials between parents. Half of the initial population was carried to the next generation without any crossover. Remaining half of the population for the next generation was obtained by crossover between any two randomly selected parents. The points of crossover were fixed at random. A crossover between two randomly selected parents generates two off springs of different character which may be fitter or less fit when compared to their parents. Crossover probability is generally fixed in the range of 0.7 to 0.8. A typical crossover is shown in Figure 2.

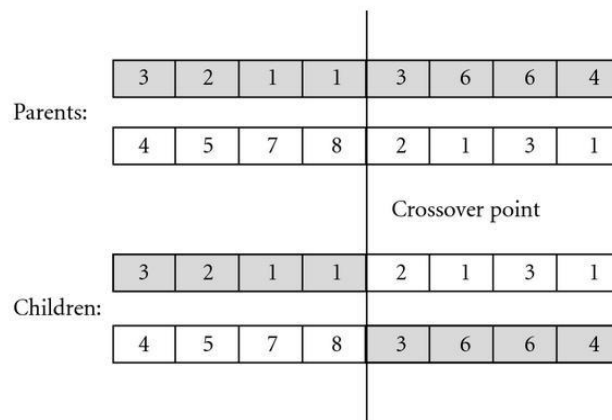


Figure 2: A typical crossover operation

#### 4.3 Mutation

The process of mutation involves replacing a random number in the variable array by another random number or replacing the type and position of operator in the operator array. Mutation can occur either in variable array, operator array, both arrays or not at all depending on the mutation rates. Since real valued genetic algorithm was used, higher mutation rates up to 0.5 were adopted. Mutation allows the program to search for a better solution in areas outside the local optimum. A typical mutation operation is shown in Figure 3.

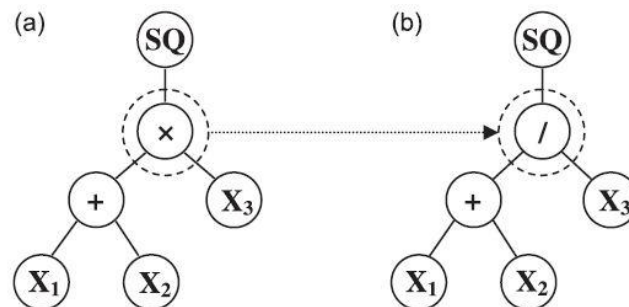


Figure 3: Typical mutation process

#### 4.4 Number of generations

The entire process of generating initial population, selection, crossover and mutation completes a generation. The selected population after crossover and mutation enter the next generation where they are sorted according to their fitness and the fitter individuals proceed to the next generations while the less fit individuals will die out. The solution obtained improves as the number of the generations increases. In this study an initial population of 1000 individuals was taken and 5000 generations were run in order to obtain a highly reliable solution. The full algorithm was implemented by coding in Scilab 5.5.1.

## 5. RESULTS AND DISCUSSIONS

The program was run several times by changing crossover and mutation rates keeping initial population and number of generations constant, to obtain a best solution to the problem. After analysis of the different solutions obtained the following equation was found to be the most robust and practical equation which can be used for settlement prediction.

$$S_c = \frac{231.68q^{0.58} * 430N^{-1.004}}{\left(368.23B^{-1.092} * 51.85L^{0.29}\right) - 105.57D_f^{-0.53} + 399.53H_w^{-0.031}}$$

where

$S_c$  = Predicted settlement (mm)

$N$  = SPT Blow count

$B$  = Footing width (m)

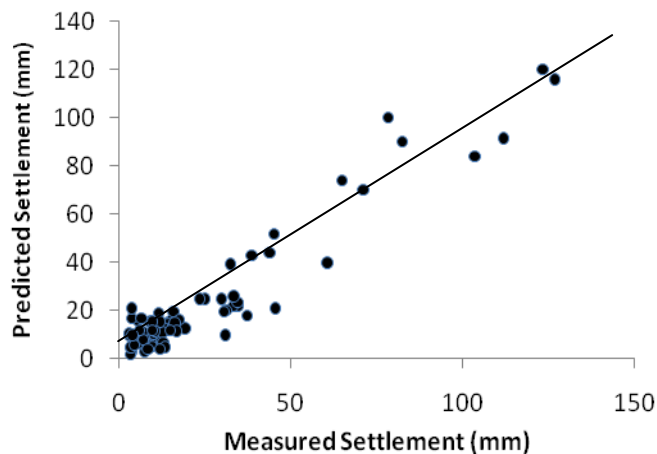
$L$  = Footing length (m)

$D_f$  = Footing embedment (m)

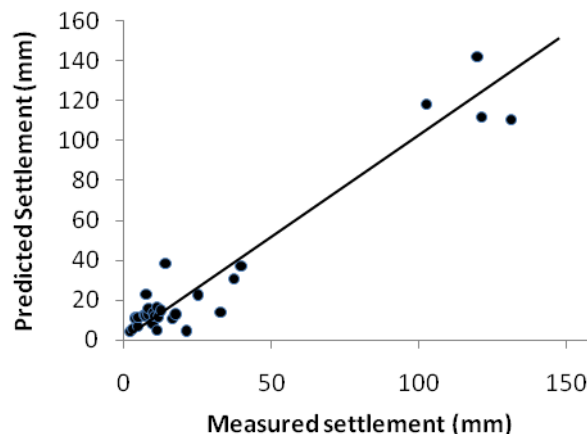
$H_w$  = Depth of water table below footing (m)

$q$  = contact pressure at footing ( $\text{kN/m}^2$ )

The performance of the model was tested by using the validation dataset, which was not used for developing the model. By testing the model with the validation dataset, it was seen that the genetic model was able to predict the settlement of unseen cases with a high level of accuracy. The variations of predicted data from the actual data for training and testing datasets are shown in Figure 4 and Figure 5 respectively. The performance of the model was also tested in terms of coefficient of correlation (R), root mean square error (RMSE) and mean absolute error (MAE).



**Figure 4:** Performance of genetic model with training dataset



**Figure 5:** Performance of genetic model with testing dataset

### 5.1 Comparison with traditional methods

The performance of the model was also assessed by comparing with different empirical equations commonly used in practice for measuring settlement. The genetic model was compared with empirical equations developed by Terzaghi and

Peck (1948), Meyerhoff (1956), Bowles (1968), Schultze and Sherif (1973). These empirical methods were selected for comparison as the database used for model development supports the use of these equations. The performances of model as well as the different empirical equations in comparison with the actual data are presented in table 3.

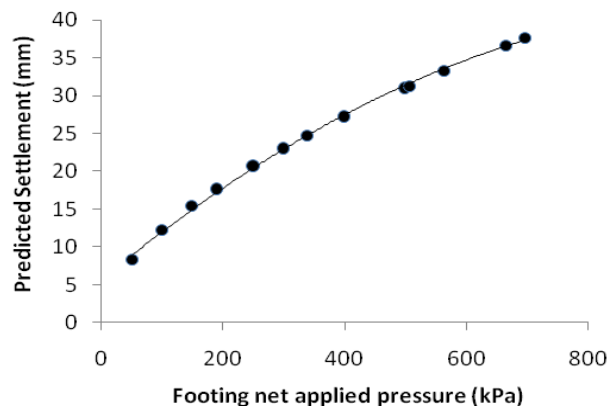
**Table 3:** Performance comparison of Genetic model

Method	Performance Criteria		
	Correlation <i>R</i>	RMSE (mm)	MAE (mm)
GA model testing	0.95	10.36	7.98
Terzaghi and Peck (1948)	0.36	34.92	28.33
Meyerhoff (1956)	0.42	26.43	13.02
Bowles (1968)	0.35	26.20	18.76
Schultze and Sherif (1973)	0.29	28.59	14.59

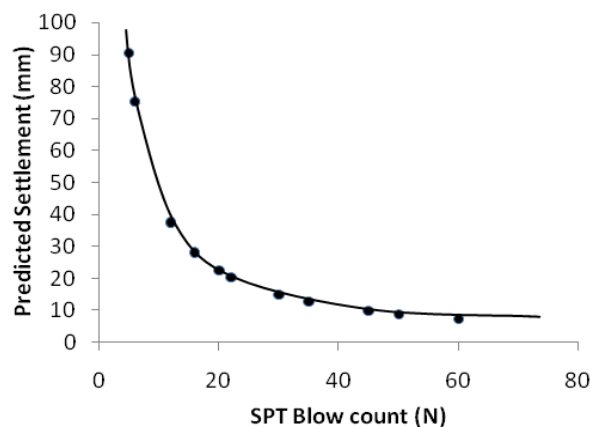
It was found out that the settlement predicted by the genetic algorithm model was much more reliable and accurate when compared to the empirical equations commonly used in practice. It was observed that the method of Terzaghi and Peck and other methods tend to overestimate the settlement.

### 5.2 Sensitivity Analysis

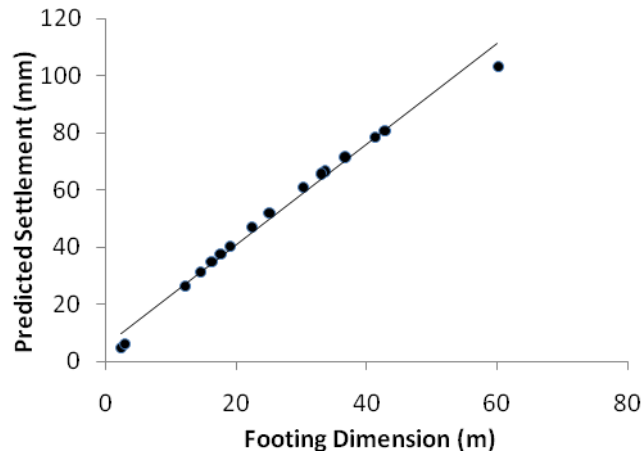
Parametric study was done to evaluate the predictive capability of the obtained formula as well as the effect of varying the input parameters on the output response. This was done by fixing all except one of the input variables to their mean value and varying the other parameter between its range of minimum and maximum values. The results of sensitivity analysis and the influence of different input variables on settlement are shown in figures 6 to 10.



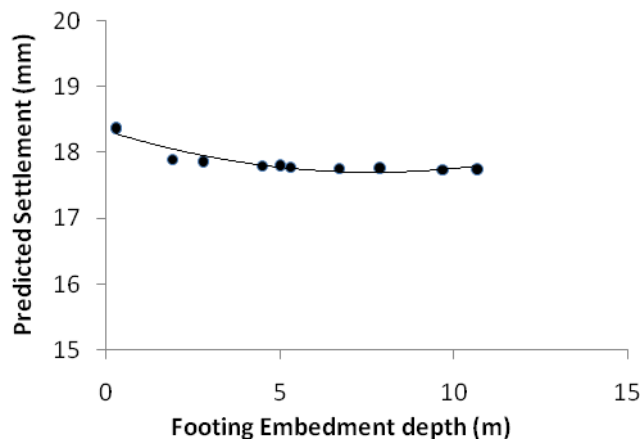
**Figure 6:** Effect of net applied pressure on settlement



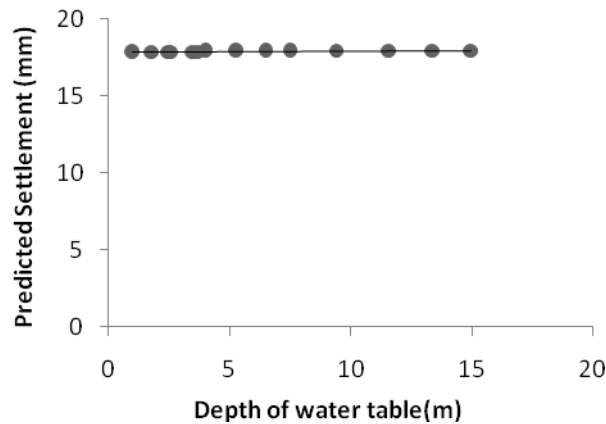
**Figure 7:** Effect of SPT Blow count on settlement



**Figure 8:** Effect of Footing width on settlement



**Figure 9:** Effect of Footing embedment on settlement



**Figure 10:** Effect of water table depth on settlement

From the parametric analysis, it could be observed that input variables like SPT Blow count, footing net applied pressure and footing dimensions are most important parameters in determining the settlement of shallow foundations, while other parameters like depth to water table and foundation embedment depth have very little effect on settlement. It could also be observed that the settlement increases with an increase in net applied pressure and footing width and that the settlement decreases with an increase in blow count. The other factors like depth to water table and depth of embedment do not have significant effect on settlement.

## 6. CONCLUSIONS

Settlement determination is a complex process due to the heterogeneous nature of soils. The theoretical relationships and empirical equations available for measuring settlement do not yield reliable results. Hence there is a need for alternative methods of prediction of settlement from existing field records which can be used widely for a large number of cases with high reliability and accuracy.

The genetic algorithm model presented in this paper proves to be a handy tool in predicting settlement of shallow foundations. The values predicted by the model are in good agreement with the measured values. It was also seen that the model performs much better than the different empirical equations which are currently used for determining settlement.

## 7. ACKNOWLEDGEMENT

The authors would like to thank Dr. Mark B. Jaksa and Dr. Harvey Wahls for providing the necessary data for preparation of database for the development of genetic model.

## 8. REFERENCES

- [1] A. R. S. S. Bazaraa, "Use of the standard penetration test for estimating settlements of shallow foundations on sand", PhD thesis, Univ. of Illinois, Urbana, 1967.
- [2] Ali R. Zolfaghari, Andrew C. Heath, Paul F. McCombie, "Simple genetic algorithm search for critical non-circular failure surface in slope stability analysis", *Journal of Computers and Geotechnics* 32 (2005) 139–152, 2004.
- [3] Bowles, Joseph, "Foundation Analysis and Design", 2nd edn, McGraw-Hill, New York, 1977.
- [4] C. Chan and L. Zhan, "Optimization of Pile Groups Using Hybrid Genetic Algorithms." *J. Geotech. Geoenviron. Eng.*, 135(4), 497–505, 2009.
- [5] C. M. Chan, L. M. Zhang, and T. Jenny, "Optimization of Pile Groups Using Hybrid Genetic Algorithms." *Journal of Geotechnical and Geoenvironmental Engineering.*, 135(4), pp 497–505, 2009.
- [6] D. J. Douglas, State-of-the-art, *Ground Engineering*, Vol 19, pp 2-6, 1986.
- [7] E. Momeni, E. Nazir, D. Jahed Armaghani, H. Maizir, "Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN", *J. Geotech. Geoenviron. Eng.*, 137(5), 462-468, 2014.
- [8] G. G. Meyerhof. "Shallow foundations", *Journal of the Soil Mechanics and Foundations Division, ASCE*, Vol. 91(SM2), pp. 21–31, 1965.
- [9] H. E. Wahls. "Settlement analysis for shallow foundations on sand", *Proc., 3rd Int. Geotechnical Engineering Conf., Cairo, Egypt*, 7–28, 1997.
- [10] H. Shahnazari, M. A. Shahin, M. A. Tutunchian, "Evolutionary-based approaches for settlement prediction of shallow foundations on cohesionless soils", *International journal of civil engineering*, No. 4, Vol. 7, pp. 236-247, 2013.
- [11] J. B. Burland, and M. C. Burbidge, "Settlement of foundations on sand and gravel", *Proc., Inst. Civ. Eng., Part I*, 78(6), 1325–1381, 1985.
- [12] J. Koza, "Genetic programming on the programming of computers by means of natural selection", Cambridge (MA): MIT Press, 813 p, 1992.
- [13] J. L. Briaud, and R. Gibbens "Behavior of five large spread footings in sand", *Journal of Geotechnical & Geoenvironmental Engineering*, 125(9), 787–796, 1999.
- [14] Lijie Cui, Daichao Sheng, "Genetic algorithms in probabilistic finite element analysis of geotechnical problems", *Journal of Computers and Geotechnics* 32 (2005) 555–563, 2005.
- [15] M. A. Shahin, H. R. Maier, M. B. Jaksa. "Predicting settlement of shallow foundations using neural networks", *Journal of Geotechnical and Geoenvironmental Engineering, ASCE*, No. 9, Vol. 128, pp. 785–793, 2002.
- [16] M. C. Burbidge, "A case study review of settlements on granular soil", MSc. thesis, Imperial College of Science and Technology, Univ. of London, London, 1982.
- [17] M. Picornell, and E. Del Monte, "Prediction of settlement of cohesive granular soils", *Proc., Measured Performance of Shallow Found., Nashville, Tenn.*, 55–72, 1988.
- [18] M. Rezaia, A. Javadi. "A new genetic programming model for predicting settlement of shallow foundations", *Canadian Geotechnical Journal*, No. 12, Vol. 44, pp. 1462-1473, 2007.
- [19] Paul McCombie, Philip Wilkinson, "The use of simple genetic algorithm in finding the critical factor of safety in slope stability analysis", *Computers and Geotechnics*, Volume 29, Issue 8, Dec, pp 699–714, 2002.
- [20] S. Levasseur, Y. Malécot, Y. Boulon and E. Flavigny, "Soil parameter identification using a genetic algorithm", *International Journal for Numerical and Analytical Methods in Geomechanics*, Volume 32, Issue 2, pp 189–213, 2008.