NEW COMPUTATIONAL MODELS FOR BETTER PREDICTIONS OF THE SOIL-COMPRESSION INDEX

Ahmet Demir

Osmaniye Korkut Ata University,
Civil Engineering Department
80000 Osmaniye, Turkey
E-mail: ahmetdemir@osmaniye.edu.tr

Keywords

compression index, statistical analysis, genetic expression programming, adaptive neuro-fuzzy, empirical equations

Abstract

The compression index is one of the important soil parameters that are essential for geotechnical designs. Because laboratory and in-situ tests for determining the compression index (C_c) value are laborious, time consuming and costly, empirical formulas based on soil parameters are commonly used. Over the years a number of empirical formulas have been proposed to relate the compressibility to other soil parameters, such as the natural water content, the liquid limit, the plasticity index, the specific gravity. These empirical formulas provide good results for a specific test set, but cannot accurately or reliably predict the compression index from various test sets. The other disadvantage is that they tend to use a single parameter to estimate the compression index (C_c), even though C_c exhibits spatial characteristics depending on several soil parameters. This study presents the potential for Genetic Expression Programming (GEP) and the Adaptive Neuro-Fuzzy (ANFIS) computing paradigm to predict the compression index from soil parameters such as the natural water content, the liquid limit, the plastic index, the specific gravity and the void ratio. A total of 299 data sets collected from the literature were used to develop the models. The performance of the models was comprehensively evaluated using several statistical verification tools. The predicted results showed that the GEP and ANFIS models provided fairly promising approaches to the prediction of the compression index of soils and could provide a better performance than the empirical formulas.

1 INTRODUCTION

The settlement of a structure is the vertical, downward movement due to a volume decrease of the soil on which the structure is built. Geotechnical engineers have a responsibility to calculate the extent of the possible settlements as completely as possible for the safety of particular projects. So, a study of the consolidation characteristics of soft, compressible geo-environmental materials is very useful for forecasting the magnitude and time of the settlement of the structure. The desired use of the structure may be damaged and the design life of the structure may be reduced if the settlement is not kept within a tolerable limit. To evaluate the spatial distribution of the consolidation settlement (s_c) in a large coastal reclamation area, geotechnical engineers need to correctly access the spatial characteristics of the soil's properties. However, it is difficult to evaluate the exact spatial characteristics of the soil's properties because the amount of geotechnical investigation data is insufficient in most cases.

In particular, it is well known that when compressible geo-materials, like silt or clay layers, are subjected to a stress, an increase in the pore-water pressure occurs immediately. Because the hydraulic conductivity of these soils is very small, the excess pore-water pressure generated by loading gradually dissipates over a long period. Consequently, the associated volume change (consolidation) of the soil continues a long time after the completion of the structure. In geotechnical engineering, the change in the void ratio versus the change in the effective pressure compressibility of the soils is defined as the coefficient of the compressibility index (C_c), which is generally determined directly using the e – logp curve.

Oedometer tests of which testing procedures have been standardized by ASTM-D-2435-96 [1] are commonly used for experimental determinations of the compression index in the laboratory. However, these are laborious, time-consuming and costly methods. In order to obtain the Cc value for soils with less effort and more economically, empirical equations based on the fundamental soil parameters, as specified using simpler laboratory tests, are generally preferred [2-9]. However, most of them are developed using limited experimental data and do not provide satisfactory and precise predictions.

The other disadvantage of these equations is that they generally use a single soil parameter or use multivariable equations based on linear approaches to predict the compression index [9-11]. The soils have reasonably complex structures, imprecise physical properties and a spatial variability (i.e., heterogeneities) associated with their formation. Therefore, their mechanical and dynamic features show an uncertain behavior in contrast to most other engineering materials [12]. Alternative methods such as GEP, ANNs and ANFIS allow the modelling of spatially complex systems and have recently emerged as commonly used and promising approaches [13-17]. Their importance is also related in all engineering areas as a result of the high-speed development of information and computer technologies. These methods have a capability for pattern recognition, classification, speech recognition, design of structures, automatic control, manufacturing process control, and the modeling of material behavior [18-20].

In this paper new approaches based on GEP and ANFIS were introduced for the prediction of the compression index of soils. The data sets for training and testing were obtained from the literature. Five basic soil properties that are accepted to be substantial parameters in geotechnical engineering, such as the natural water content (ω_n), the liquid limit (LL), the plasticity index (PI), the specific gravity (Gs) and the void ratio (e), were used for the GEP and ANFIS models as the input parameters.

2 OVERVIEW OF THE COMPRESSION INDEX OF SOILS

The compression index (C_c) shows the slope of the linear part of the curve of the void ratio versus the logarithm of the effective pressure (Fig. 1). In other words, it means a change in the void ratio due to the effective pressure change during the consolidation of soils.

For a layer of normally consolidated soil of thickness H, the initial void ratio e_o, the compression index C_c, the effective overburden pressure P'_0, and the total settlement S_t under an applied load Δp can be expressed as

$$S_t = \frac{C_c}{1+e_0} H \log \frac{P'_0 + \Delta p}{P'_0} \qquad (1)$$

where C_c is the slope of the virgin compression portion of the e – log p curve determined from a standard consolidation test on an undisturbed sample.
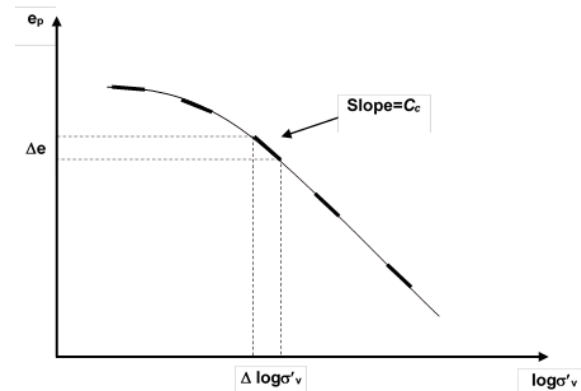

Figure 1. A typical e_p – log σ'_v curve for obtaining the value of C_c at any load increment.

3 BRIEF REVIEW OF GENETIC EXPRESSION PROGRAMING (GEP)

Genetic Expression Programming (GEP), which is based on genetic algorithms (GAs) and genetic programming (GP), was developed for the first time by [21]. Its data-processing system is similar to the human genetic system and is a computer program encoded in linear chromosomes of fixed length. The fundamental concept of this approach is to find a mathematical function, defined as a chromosome with multi genes, by using the data presented to it. The mathematical expression is encoded as simple strings of fixed length, which are subsequently expressed as expression trees of different
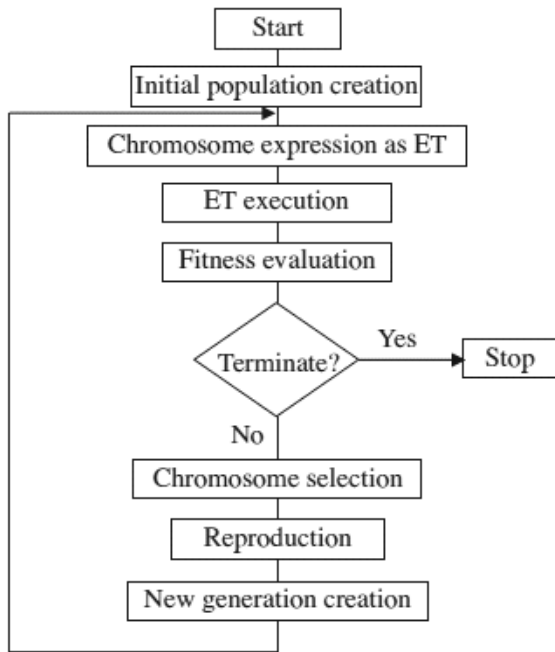
Figure 2. The algorithm of Genetic Expression Programming [25].

Chromosome with one gene

$$(a-b)+\sqrt{(a*b)}$$

Corresponding mathematical equation

Expression tree
(ET)

Figure 3. Schematic indication of a chromosome with one gene and its expression tree and the corresponding mathematical equation.

size and shape in GEP [22-24]. A typical GEP algorithm is sketched out in Fig. 2. Its processing initializes, selecting five elements, such as the function set, the terminal set, the fitness function, the control parameters and the stop condition. The GEP algorithm randomly generates an initial chromosome that symbolizes a mathematical function and then converts it into an expression tree (ET), as indicated in Fig. 3.

The later processing is to compare between the predicted values and the measured values. If the desired results are achieved using the initially selected error criteria , the GEP algorithm is terminated. When the expected results cannot be obtained, some chromosomes are selected by means of the method called roulette-wheel sampling. This method, with an elitism strategy, is employed by the GEP algorithm to select and copy the individuals. Single or several genetic operators, such as crossover, mutation and rotation, are used for introducing variations into the population. Note that the rotation operator rotates two subparts of the genome with respect to a randomly chosen point. Further descriptions of the GEP algorithm can be found in [21]. They are mutated to obtain new chromosomes. After the desired fitness score is obtained, this process terminates and then the knowledge coded in the genes in chromosomes is decoded for the best solution of the problem [25].
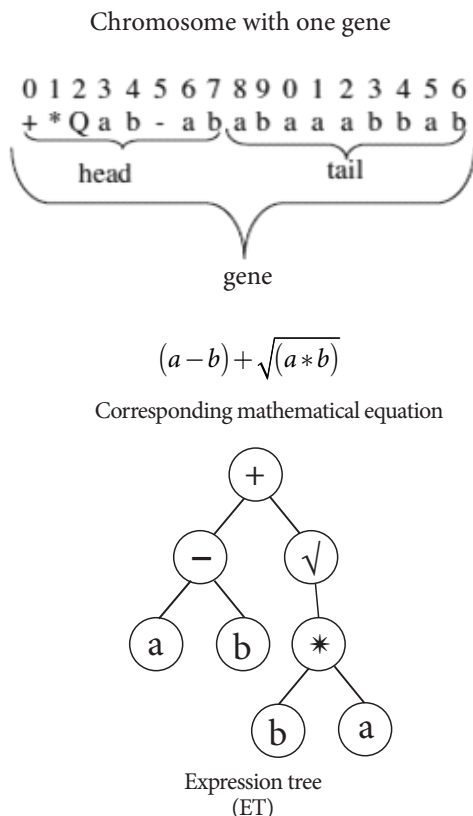
GEP has two main components that are defined as the chromosomes and the expression trees (ETs). The chromosomes that may have one or more genes are coded with some information using a special language about the problem. The mathematical information is translated to the ET using a bilingual and conclusive language called Karva Language (the language of the genes) and by means of the language of ETs. The genotype is accurately derived by using the Karva Language. The GEP genes are made up of two parts that are named as the head and the tail. The head of a gene includes the main variables needed to code any expression, such as some functions, variables and constants. The tail simply contains variables and constants, which may be required for additional terminal symbols. These symbols are used in the event that the variables in the head are insufficient to encipher a function. While the head of a gene contains arithmetic and trigonometric functions – like +, -, √, /, p, sin, cos – the tail includes the constants and the independent variables of the problem – like (1, a, b, c).

At the beginning of the model's construction the user specifies the length of the head (i.e., the number of symbols), which is the most significant parameter in the GEP process. The encoding process takes place by reading the ET from left to right in the top line of

the tree and from the top to the bottom, and the ET is converted to Karva Language. The GEP genes include a non-coding part similar to the coding and non-coding sequences of biological genes. There are four primary operators (such as selection, mutation, transposition, and cross-over) during the GEP processing. When the mathematical equation obtained from the GEP model is not suitable for the problem, the chromosomes should be modified by means of GEP operators to obtain the next generation. The operators given above are applied with the operator rate that shows a certain probability for a chromosome. The operator rates are specified by the user prior to the analysis. The mutation rate is generally used between 0.001 and 0.1. On the other hand, it is suggested that the transposition rate and cross-over rate are 0.1 and 0.4, respectively [25].

4 FUZZY INFERENCE SYSTEMS

Zadeh [26] proposed the first fuzzy approach. This method exhibits some differences from the traditional cluster theory (TCT), which is the crisp definition for an element belonging to a cluster. According to TCT, an element either belongs to a cluster or not. However, the fuzzy approach does not decide completely with respect to the belonging. This is because the degree of membership of an element is important for the fuzzy approach. It is defined partially by the continuous membership functions that take a value between 0 and 1 [27, 28]. Takagi and Sugeno [29] proposed two models known as the Mamdani and Tagagi-Sugeno (TS) models. There are, however, some differences between them. The Mamdani model employs human expertise and the linguistic knowledge's to build the membership functions and the if-then rules. However, in the TS model the optimization and adaptive techniques are used and it also uses a smaller number of if-then rules. So, the advantage of some aspects of the TS model is that it is more suitable for the mathematical and computational modeling, and therefore it is mostly preferred by researchers [14, 24]. The other advantage is that it also makes it possible to design the output function as either linear or constant [13, 30 and 31]. ANFIS is a kind of fuzzy-logic approach proposed by Jang [32]. It has more advanced properties than other fuzzy models, such as learning and parallelism, like that of ANNs, which allow the fuzzy rules and membership functions to be generated adaptively using a neural training process with the data set presented. In the Sugeno-type fuzzy approach, the if-then rules given below are used:

$$\text{If } x = A_1 \text{ and } y = B_1 \text{ then } f_{1(x,y)} = p_1 x + q_1 y + k_1 \quad (2)$$

$$\text{If } x = A_2 \text{ and } y = B_2 \text{ then } f_{2(x,y)} = p_2 x + q_q y + k_2 \quad (3)$$

where x (or y) = the input node; i, p, q and k = the consequence parameters obtained from the training; A and B = are labels of the fuzzy set defining a suitable membership function. A backpropagation learning algorithm and a hybrid learning algorithm are utilized to update the membership functions during the fuzzy process [33]. In Fig. 4, the main concept of the ANFIS approach is illustrated. As shown, the computational process of the ANFIS is performed in five steps. The initial parameters that define the membership functions are specified in the first step. For example, the initial parameters of the generalized bell-shaped are widely used as a membership function, as given below:

$$\eta_A(X) = \frac{1}{1 + \left| \dfrac{x - c_i}{a_i} \right|^{2b}} \quad (4)$$

where η = the membership function; a, b and c = the parameter set known as the initial (or premise) parameters. In second step, every node in this layer is a fixed node labelled Π, representing the firing strength of each rule. The firing strength means the degree to which the antecedent part of the rule is satisfied. It represents the product of the incoming signals that is calculated using the following equation:

$$U_{2,i} = w_i = \eta A_i(x) \times \eta B_i(y) \quad i = 1,2 \quad (5)$$

The firing strengths computed in the second step are normalized using the following equation in the third step,

$$U_{3,i} = \bar{w}_l = \frac{w_i}{w_1 + w_2} \quad i = 1,2 \quad (6)$$

In the fourth step, the effect of each rule on the output is computed with an adaptive node function using the following equation:

$$U_{4,i} = \bar{w}_l f_i = \bar{w}_i (p_i x + q_i y + k_i) \quad (7)$$

where w is the normalized firing strength calculated in step 3; and p_i, q_i, and k_i are the resulting parameters. The final output of the ANFIS model is found in the fifth step using the equation shown below:

$$U_{5,i} = \sum \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (8)$$
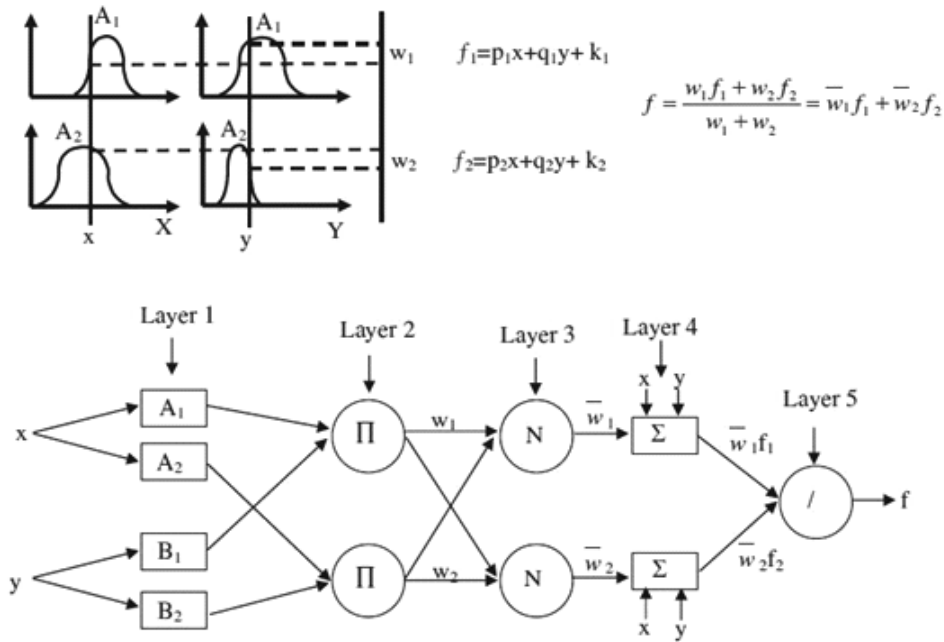
Figure 4. The simple ANFIS architecture [34].

5 DATA COLLECTION

This study aims to estimate the compression index by taking account of the physical properties of soils. The data sets used to develop the GEP and ANFIS models were obtained from Kalantary and Kordnaeij [35]. They reported on data from Iran and suggested an ANN model for the prediction of the compression index. The input parameters used herein were selected in such a manner that the phenomenon of the compression index is defined by these parameters in accordance with the

methods used extensively in practical engineering. Therefore, the void ratio (e_o), the natural water content (ω_n), the liquid limit (LL), the plastic index (PI), and the specific gravity (Gs) were chosen as the input parameters. Also, based on the previous trend of studies, the compression index of the soils is assumed to be affected by them. Fig. 5 shows histograms of the inputs and the target parameters and, also, the statistical parameters of the input and output variables for each data set are given in Table 1. As is clear from Table 1, there is a high correlation between the input parameters (ω_n and e_o) and the target parameter (C_c).

Table 1. Statistical parameters of the input and output variables for each data set.

Data Set	Variable	x_{ort}	σ	C_v	C_{sx}	C_k	x_{maks}	x_{min}	Range	Correlation coefficient with the C_c
Training	ω_n	28.39	6.94	48.16	1.01	2.44	57.40	11.10	46.30	0.76
	LL	40.02	9.83	96.31	1.26	0.92	96.31	24.00	72.31	0.48
	PI	18.83	8.50	71.88	1.37	0.36	71.88	4.00	67.88	0.46
	e_o	0.76	0.15	0.02	0.72	4.95	4.95	0.41	4.54	0.85
	Gs	2.64	0.06	0.00	-9.77	0.50	2.80	2.43	0.37	-0.16
Testing	ω_n	28.93	6.80	46.28	0.46	-0.04	46.40	14.50	31.90	0.80
	LL	37.41	7.46	55.60	0.62	-0.40	57.00	25.00	32.00	0.30
	PI	16.53	6.79	46.07	0.44	-0.68	34.00	6.00	28.00	0.27
	e_o	0.77	0.15	0.02	0.68	0.63	1.23	0.48	0.76	0.87
	Gs	2.64	0.05	0.00	-0.70	1.29	2.74	2.49	0.25	0.00

C_v: variation coefficient, C_{sx}: skewness coefficient, C_k: kurtosis coefficient.
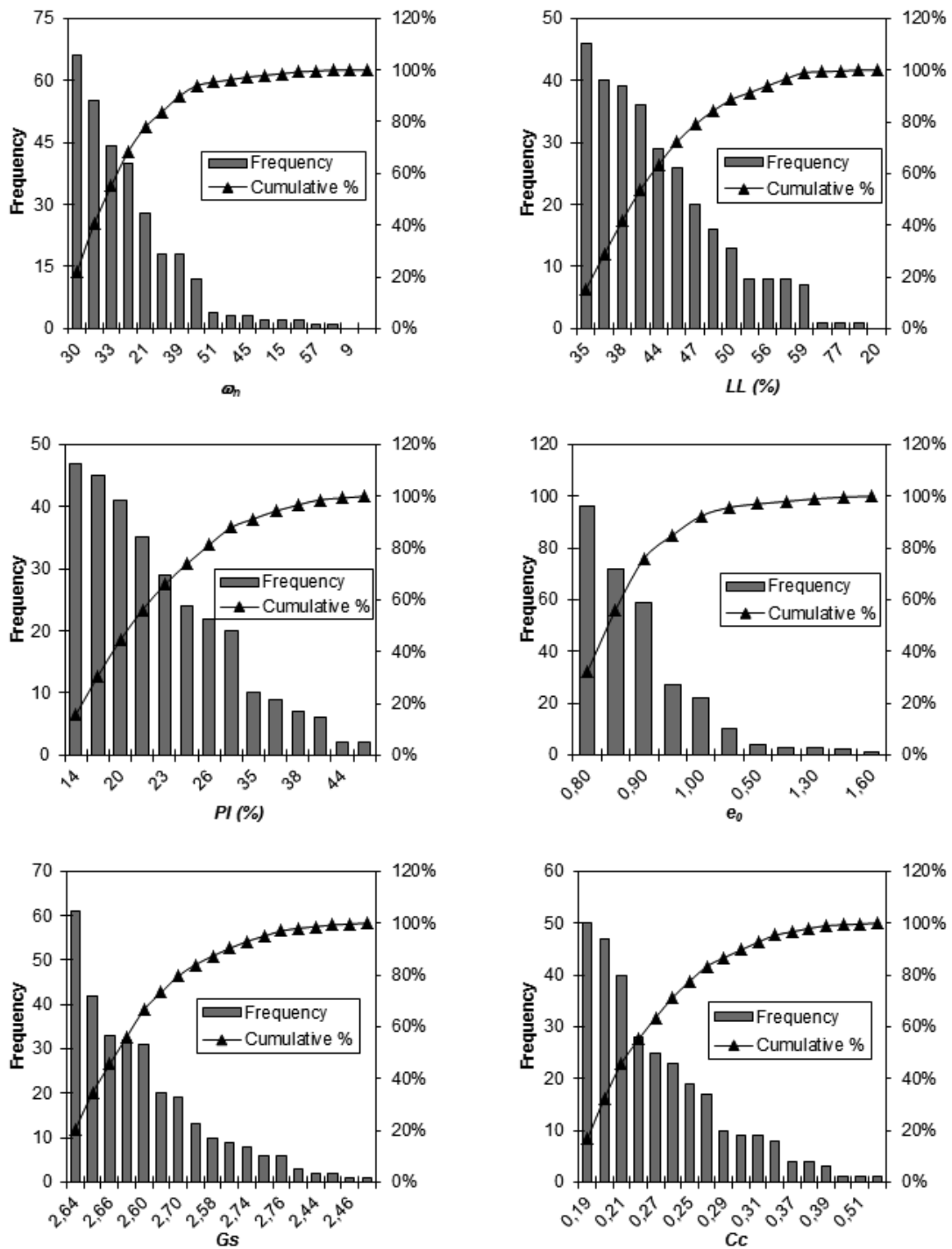
Figure 5. The histograms of the input variables and the output variable.

6 GEP MODEL DEVELOPMENT

The GEP models enhanced herein are mainly designed to generate mathematical functions for the prediction of the compression index of the soil. Three GEP models (GEP Model I, GEP Model II and GEP Model III) were developed. Whilst five input parameters were selected, such as ω_n, LL, PI, e_o and Gs for the GEP Model I, three and two input parameters ((LL, PI and e_o), (ω_n and e_o)) were used in the GEP Model II and GEP Model III, respectively. In other words, the ω_n and Gs parameters were not taken into account for the inputs of the GEP

Model II. For that reason, three mathematical functions in the form $y = f(\omega_n, LL, PI, e_o \text{ and } Gs)$, $y = f(LL, PI \text{ and } e_o)$ and $y = f(\omega_n \text{ and } e_o)$ were generated for the prediction of the compression index of the soil. The model parameters used for both models are given in Table 2. DTREG software is used for the GEP algorithm [36]. The functions obtained from the GEP Models are given below:

Model I

$$C_c = \tan\left[\exp\left[\left(\tan\left(LL \times G_s\right)^{1/9}\right)\right]^{1/3}\right]^2$$

$$+ \text{atan}\left[\frac{\sin\left(\cos\left(G_s\right) \times \left(2e_0 + G_s\right)\right)}{\omega_n}\right]$$

$$- \text{atan}\left[\text{atan}\left(\cot\left(LL\right) - 2LL\right) + \left(G_s - e_0\right)^{1/2}\right]$$

$$- \cos\left[\frac{\sec\left(PI\right)^3}{G_s^2 \times \left(G_s + G_s^3\right)}\right]$$

Model II

$$C_c = \cos\left[\ln\left(\sqrt{\left(e_0 \times LL\right)}\right) + \text{atan}\left(LL\right)\right]$$

$$+ \sin\left[\sqrt{\ln\left[2.80\left[2\left(LL\right) - \left(LL \times e_0\right)\right]\right]}\right]$$

$$+ \text{atan}\left[-1.88\left(\ln\left[\sqrt{\left(\frac{LL}{e_0}\right)}\right]\right)\right]$$

$$+ \text{atan}\left[\sqrt{e_0\left[\left(\left(PI - 3.44\right) \times e_0\right) \times \left(LL + 6.15\right)\right]}\right]$$

Model III

$$C_c = \cos\left[\left(\omega_n \times e_0 + \omega_n\right)\right] / \left(\sqrt{\omega_n} \times 2\omega_n\right)$$

$$+ \cos\left(\sqrt{\omega_n} \times e_0\right) \times \left[\left(e_0\right)^2 / 2\omega_n\right] + \frac{\left(e_0\right)^2}{\left(\omega_n \times e_0\right)^{1/3}}$$

7 ANFIS MODEL DEVELOPMENT

Three ANFIS Models (ANFIS Model I, ANFIS Model II and ANFIS Model III) were built using the same inputs as in the GEP. The membership functions of each of the input variables were generated using the grid-partition method. The triangular membership function was chosen for both models and the hybrid learning algorithm was executed for optimizing the parameters that can perform a rapid identification of the parameters, substantially reducing the time needed to reach convergence.

Table 2. GEP parameters of the developed models .

Population Size	50
Generation	194174
Number of The Genes	4
Length of The Gene Head	8
Max. Generation	227405
Linking Function	+
Function Set	+, -, *, /, √, exp, ln, sin, cos, atan
Mutation Rate	0.044
Inversion Rate	0.1
One-Point Recombination Rate	0.3
Two-Point Recombination Rate	0.3
Gene-Transposition Rate	0.1

In order to avoid over fitting, the stopping criterion was adopted as the minimum validation error. The ANFIS Model I has 243 linear parameters, 45 nonlinear parameters, 524 nodes and 243 fuzzy rules. On the other hand, ANFIS Model II has 64 linear parameters, 48 nonlinear parameters, 158 nodes and 64 fuzzy rules. Also, ANFIS Model III has 9 linear parameters, 18 nonlinear parameters, 35 nodes and 9 fuzzy rules. The fuzzy toolbox of the MATLAB computer-aided software was used for the model development [33].

8 RESULTS AND DISCUSSION

This paper, to a large extent, intends to investigate the potential use of GEP and ANFIS for the prediction of the compression index of soils, which has great significance for soil mechanics and foundation engineering. The results obtained from these approaches were comprehensively evaluated in terms of statistics for a quantitative assessment of the model's predictive abilities. Of the 299 data sets, 233 were used for training the models and 66, which are not used in training stage, were presented for the testing of the models. In order to learn the performance of the developed models, several statistical verification criteria were used, such as the coefficient of correlation (R), the root-mean-square error (RMSE) and the standard deviation (σ) of the errors. The definition of these evaluation criteria are given as follows:

$$R = \frac{\sum_{i=1}^{n}\left(u_i^m - \overline{u}^m\right)\left(u_i^c - \overline{u}^c\right)}{\sqrt{\sum_{i=1}^{n}\left(u_i^m - \overline{u}^m\right)^2}\sqrt{\sum_{i=1}^{n}\left(u_{i=1}^n - \overline{u}^c\right)^2}} \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i}^{N}\left(u_i^m - u_i^c\right)^2}{N}} \qquad (13)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}\left(e - \overline{e}\right)^2}{(N-1)}} \qquad (14)$$

where u_i^m and u_i^c are the measured and predicted values, respectively. $\overline{u}^m$ and $\overline{u}^c$ are the mean of the measured and predicted values, e is the absolute error ($\left|u_i^m - u_i^c\right|$), $\overline{e}$ is the mean of the absolute error, and N is the size of the sample. The C_c values estimated from all the models through the training and testing process were graphically compared with the case records in Fig.6. It is clear from the figure that the results from the GEP and ANFIS are in good agreement with the case records. This also shows that all the models were found to be able to learn the complex relationship between the input parameters relating to soils and the value of C_c. Moreover, the statistical performances of the models are presented in Table 3. With respect to this table, all the models for compression index C_c give a satisfactory agreement in terms of the statistical evaluation criteria. The best results with regard to the R values were 0.910 and 0.900 for the GEP Model I and the ANFIS Model I, respectively. However, the GEP Model II and the ANFIS Model II give relatively high R values, i.e., 0.870 and 0.870, respectively. In statistics, the overall error performances of the relationship between the two groups can be interpreted from the R values. According to Smith (1986), if a proposed model gives R>0.8, there is a strong correlation between the measured and the predicted values for all the data available in the database.

As the models are compared with regard to the RMSE, which is a measurement of the deviation around the regression line, it is clear that the lowest RMSE is obtained from Eq. (9) generated from GEP Model I, i.e., 0.029. On the other hand, Eq. (10) generated from GEP Model II gives an RMSE value of 0.034. The other models yield low RMSE values, ranging from 0.032 to 0.090. The RMSE value has great significance for the statistics in addition to the R value, because although the relationship provides a high R value, it may also give a high RMSE value.

As seen from Table 4, the predictability of the GEP and ANFIS models is also statistically compared with the empirical formulas, the mean and the standard deviation of the ratio $C_{c,pred}/C_{c,mea}$, which is often used for a statistical analysis. The mean (μ) and the standard deviation (σ) of $C_{c,pred}/C_{c,mea}$ are important indicators of

the accuracy and the precision of the prediction method. Under ideal conditions, an accurate and precise method gives a mean value of 1.0 and a standard deviation of 0. A μ value greater than 1.0 indicates an overestimation and an underestimation, otherwise. The best model is represented by a μ value close to 1.0 and a σ value close to 0. Based on the μ value, the GEP Model I shows a good prediction when using the soil parameters such as ω_n, LL, PI, e_o and Gs. Other empirical formulas yield a μ value in the range 0.820–14.574. This means that, on average, they considerably underestimate or overestimate the compression index. The value of σ is also found to be a minimum for the GEP Model I.

Table 3. Performance statistics of the models.

	GEP Models			ANFIS Models		
	I	II	III	I	II	III
R	0.910	0.870	0.866	0.900	0.870	0.852
RMSE	0.029	0.034	0.035	0.032	0.090	0.037
σ	0.018	0.023	0.024	0.021	0.081	0.026

Table 4. Statistical results for conventional empirical formulas.

| Equation	Average, μ	Standard Deviation, σ	References
$C_c=0.01\omega_n-0.05$	1.154	0.224	Azzouz [4]
$C_c=0.01\omega_n$	1.411	0.262	Koppula [37]
$C_c=0.013\omega_n-0.115$	0.820	0.206	Park and Lee [9]
$C_c=0.54e_0-0.19$	1.058	0.204	Nishida [38]
$C_c=0.75e_0-0.38$	0.874	0.352	Sowers [39]
$C_c=0.006\omega_L-0.054$	0.858	0.292	Azzouz [4]
$C_c=0.009\omega_L-0.090$	1.241	0.429	Terzaghi and Peck [3]
$C_c=0.014\omega_L-0.168$	1.786	0.642	Park and Lee [9]
$C_c=0.2343(\omega_L/100)G_s$	1.167	0.352	Nagaraj and Murthy [5]
$C_c=2.926(\omega_L/100)G_s$	14.574	4.399	Park and Lee [9]
$C_c=0.009\omega_n+0.005\omega_L$	2.216	0.443	Koppula [37]
GEP Model I	0.993	0.131	This study
GEP Model II	0.980	0.151	This study
GEP Model III	1.015	0.150	This study
ANFIS Model I	1.016	0.140	This study
ANFIS Model II	0.992	0.138	This study
ANFIS Model III	1.010	0.149	This study

a) GEP Models
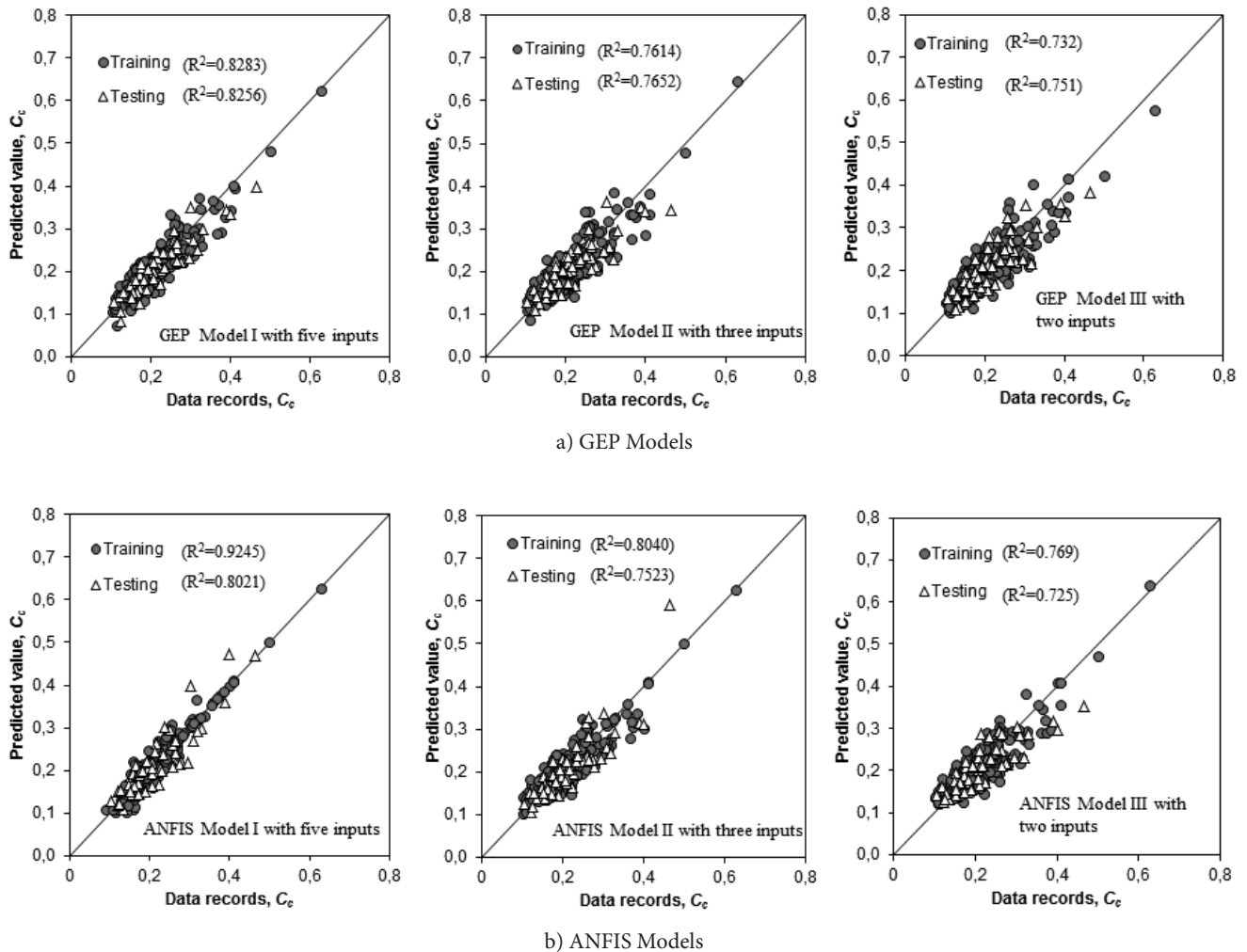
b) ANFIS Models

Figure 6. Comparison of the case records with the predicted C_c values from the GEP and ANFIS Models.

The evaluations given above clearly reveal that Eqs. (9), (10) and (11) generated by the GEP Models and the ANFIS Models have a good prediction ability. The prediction accuracy of the models appears to be statistically sufficient in terms of the prediction of C_c. Since the laboratory and the *in-situ* tests for the determination of the C_c value are laborious, time consuming and costly, it is better to use Eqs. (9), (10) and (11) to estimate the compression index (C_c) of soils.

9 CONCLUSIONS

This study looks at the capability of Genetic Expression Programming (GEP) and Adaptive Neuro-Fuzzy (ANFIS) for the prediction of the compression index (C_c) in soils. Data for the development and the testing of the models were obtained from the literature. At the end of the analyses, three mathematical equations from the GEP and three ANFIS Models were developed for a better prediction of the soil-compression index. The results of these compression indexes of the soils predicted by the GEP Models are compared with those obtained from the ANFIS and the conventional empirical formulas.

The comparison between soft computing systems mentioned above indicated that the GEP Model I markedly outperformed the other models. The satisfactory agreement was obtained as a result of the testing procedures of Eqs. (1) and the GEP Models. This was evidenced by some statistical performance criteria used for evaluating the models. Eqs. (1) produced the GEP Model I and GEP Model II, respectively, and gave a high correlation coefficient (0.910 and 0.870, respectively) and low RMSE values (0.029 and 0.035). On the other hand, the ANFIS Models produced satisfactory results

with R values that are ranging from 0.900 to 0.850 and higher RMSE values ranging from 0.032 to 0.900 than the GEP Models. The overall evaluation of the results obtained throughout the paper revealed that the soft-computing techniques used herein are very encouraging for the cases tested.

It was also observed that the best results were obtained from Model I with five input parameters. The other models that were developed to see the effects of different soil properties also produced satisfactory results. It was concluded from all the findings herein that the use of basic soil properties, such as the void ratio (e_o), the natural water content (ω_n), the liquid limit (LL), the plastic index (PI), and the specific gravity (G_s), appears to be reasonable for the prediction of the compression index (C_c) of soils.

REFERENCES

[1] ASTM D 2435-96 1998. Standard test method for one-dimensional consolidation properties of soils. Annual Book of ASTM standards. Vol. 04.08, Soil and Rock (I). Standard. Pennsylvania pp.207-216.

[2] Helenelund K.V. 1951. On consolidation and settlement of loaded soil layers. Dissertation, Finland Technical Institute, Helsinki, Finland.

[3] Terzaghi, K., Peck, R.B. 1967. Soil mechanics in engineering practice. 2nd ed. John Wiley & Sons Inc., New York.

[4] Azzouz, A.S., Krizek R.J., Corotis, R.B. 1976. Regression analysis of soil compressibility. Soils Foundations 16, 2, 19-29.

[5] Nagaraj, T.S., Murthy, B.R.S. 1986. A critical reappraisal of compression index equations. Geotechnique 36, 1, 27–32.

[6] Nakase, A., Kamei, T., Kusakabe, O. 1988. Constitute parameters estimated by plasticity index. Journal of Geotechnical Engineering, ASCE, 114, 844–58.

[7] Sridharan, A. Nagaraj, H.B. 2000. Compressibility behaviour of remoulded, finegrained soils and correlation with index properties. Canadian Geotechnical Journal 37, 3, 712–722.

[8] Giasi, C.I., Cherubini, C., Paccapelo, F. 2003. Evaluation of compression index of remoulded clays by means of Atterberg limits. Bull. Eng. Geol. Environ. 62, 4, 333–340.

[9] Park, H.I., Lee, S.R. 2011. Evaluation of the compression index of soils using an artificial neural network. Computers and Geotechnics 38, 472–481.

[10] Bowles, J. E. 1992. Engineering properties of soils and their measurement (fourth ed.). New York, NY: McGraw-Hill.

[11] Yoon, G.L., Kim, B.T., Jeon, S.S. 2004. Empirical correlations of compression index for marine clay from regression analysis. Canadian Geotechnical Journal 41, 6, 1213–1221.

[12] Jaksa, M. B. 1995. The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay. Ph.D. thesis, Faculty of Engineering, The University of Adelaide, 469 pp.

[13] Shahin, M. A., Maier, H. R., Jaksa, M. B. 2003. Settlement prediction of shallow foundations on granular soils using B-spline neurofuzzy models. Computers and Geotechnics 30, 637–647.

[14] Tutmez, B., Tercan, A. E. 2007. Spatial estimation of some mechanical properties of rocks by fuzzy modeling. Computers and Geotechnics 34, 10–18.

[15] Cobaner, M., Haktanir, T., Kisi, O. 2008. Prediction of hydropower energy using ANN for the feasibility of hydropower plant installation to an existing irrigation dam. Water Resources Management 22, 6, 757-774.

[16] Baykasoglu, A., Gullu, H., Canakcı, H., Ozbakır, L.L. 2008. Prediction of compressive and tensile strength of limestone via genetic programming. Expert Systems with Applications 35, 111–123.

[17] Kayadelen, C., Gunaydın, O., Fener, M., Demir, A., Ozvan, A. 2009. Modeling of the angle of shearing resistance of soils using soft computing systems. Expert Systems with Applications 36, 11814–11826.

[18] Najjar, Y. M., Basheer, I. A. 1996. Discussion of stress-strain modelling of sands using artificial neural networks. Journal of Geotechnical Engineering 122, 11, 949–950.

[19] Adam, E. G. 2003. Data-driven linguistic modeling using relational fuzzy rules. IEEE Transactions on Fuzzy Systems 11, 1, 121–134.

[20] Kim, Y.S., Kim, B.T. 2008. Prediction of relative crest settlement of concrete-faced rockfill dams analyzed using an artificial neural network model. Computers and Geotechnics 35, 3, 313–322.

[21] Ferreira, C. 2001. Gene expression programming: A new adaptive algorithm for solving problems. Complex Systems 13, 2, 87–129.

[22] Munoz, D. G. 2005. Discovering unknown equations that describe large data sets using genetic programming techniques. Master thesis in electronic systems at Linkoping Institute of Technology, LITH-ISY-EX–05/3697.

[23] Cevik, A. 2007. A new formulation for longitudinally stiffened webs subjected to patch loading. Journal of Constructional Steel Research 63, 1328–1340.

[24[Kayadelen, C. 2011. Soil liquefaction modeling by Genetic Expression Programming and Neuro-Fuzzy. Expert Systems with Applications 38, 4080–4087.

[25] Teodorescu, L., Sherwood, D. 2008. High energy physics event selection with gene expression programming. Computer Physics Communications 178, 409–419.

[26] Zadeh, L. A. 1965. Fuzzy sets. Information and Control 8, 338–353.

[27] Dubois, D., Prade, H. 1980. Fuzzy sets and systems. New York: Academic Press.

[28] Topcu, I. B., Sarıdemir, M. 2007. Prediction of rubberized concrete properties using artificial neural network and fuzzy logic. Construction and Building Materials, in press

[29] Takagi, T., Sugeno, M. 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions Systems, Man, and Cybernetics 15, 116–132.

[30] Abdulsamet, H., Mehmet, Y., Omer, C., Ismail, E. 2004. Adaptive neuro-fuzzy modeling of transient heat transfer in circular duct air flow. International Journal of Thermal Sciences 43, 1075–1090.

[31] Yılmaz, Y., Mahmut, B. 2006. Adaptive neuro-fuzzy based modelling for prediction of air pollution daily levels in city of Zonguldak. Chemosphere 63, 1575–1582.

[32] Jang, J. S. R. 1993. ANFIS: Adaptive-network-based fuzzy inference systems. IEEE Transactions Systems, Man, and Cybernetics 23, 3, 665–685.

[33] Demuth, H., Beale, M. 2001. Neural network toolbox for use with Matlab. Natick, Mass: The MathWorks Inc. 840 pp.

[34] Padmini, D., Ilamparuthi, K., Sudheer, K. P. 2008. Ultimate bearing capacity prediction of shallow foundations on cohesionless soils using neurofuzzy models. Computers and Geotechnics 35, 33–46.

[35] Kalantary, F., Kordnaeij, A. 2012. Prediction of compression index using artificial neural network. Scientific Research and Essays 7, 31, 2835-2848.

[36] Sherrod, P.H. 2008. DTREG predictive modeling software. <http://www.dtreg.com/>.

[37] Koppula S.D. 1981. Statistical estimation of compression index. Geotechnical Testing Journal 4, 2, 68–73.

[38] Nishida Y. 1956. A brief note on compression index of soils. J SMFE Div. ASCE 1, 14, 1027.

[39] Sowers G.B. 1970. Introductory soil mechanics and foundation. 3rd ed. London: The Macmillan company of Collier-Macmillan Ltd., pp. 102.