

# Nonlinear genetic-based simulation of soil shear strength parameters

SEYYED MOHAMMAD MOUSAVI<sup>1</sup>, AMIR HOSSEIN ALAVI<sup>2,\*</sup>,  
AMIR HOSSEIN GANDOMI<sup>3</sup> and ALI MOLLAHASANI<sup>4</sup>

<sup>1</sup>*Department of Geography and Urban Planning, Faculty of Humanities and Social Sciences,  
Science and Research Branch, Islamic Azad University, Tehran, Iran.*

<sup>2</sup>*Young Researchers Club, Mashhad Branch, Islamic Azad University, Mashhad, Iran.*

<sup>3</sup>*Department of Civil, Environmental and Material Engineering (DICAM),  
University of Bologna, Bologna, Italy.*

<sup>4</sup>*Young Researchers Club, Central Tehran Branch, Islamic Azad University, Tehran, Iran.*

*\*Corresponding author. e-mail: ah\_alavi@hotmail.com*

New nonlinear solutions were developed to estimate the soil shear strength parameters utilizing linear genetic programming (LGP). The soil cohesion intercept ( $c$ ) and angle of shearing resistance ( $\phi$ ) were formulated in terms of the basic soil physical properties. The best models were selected after developing and controlling several models with different combinations of influencing parameters. Comprehensive experimental database used for developing the models was established upon a series of unconsolidated, undrained, and unsaturated triaxial tests conducted in this study. Further, sensitivity and parametric analyses were carried out.  $c$  and  $\phi$  were found to be mostly influenced by the soil unit weight and liquid limit. In order to benchmark the proposed models, a multiple least squares regression (MLSR) analysis was performed. The validity of the models was proved on portions of laboratory results that were not included in the modelling process. The developed models are able to effectively learn the complex relationship between the soil strength parameters and their contributing factors. The LGP models provide a significantly better prediction performance than the regression models.

## 1. Introduction

One of the most important engineering properties of soil is its ability to resist sliding along internal surfaces within a mass. The stability of structures built on soil depends upon the shearing resistance offered by the soil along probable surfaces of slip-page. The shear strength of geotechnical materials is generally represented by the Mohr–Coulomb theory. According to this theory, the shear strength of soils varies linearly with the applied stress through two shear strength components known as the cohesion intercept and angle of shearing resistance.

The tangent to the Mohr–Coulomb failure envelopes is represented by its slope and intercept. The slope expressed in degrees is the angle of shearing resistance and the intercept is the cohesion (Arora 1988; Murthy 2008; Mollahasani *et al* 2011). The cohesion intercept and angle of shearing resistance are treated as constants over the range of normal stresses. The values of these empirical parameters for any soil depend upon several factors such as textural properties, stress history of soil, initial state, and permeability characteristics of soil (Murthy 2008). Figure 1 shows the Mohr circles and failure envelopes in terms of the total stresses.

**Keywords.** Soil shear strength parameters; soil physical properties; linear-based genetic programming; prediction.

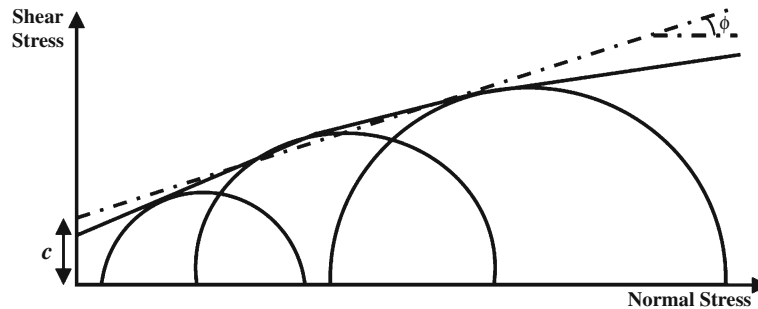


Figure 1. Mohr circles and failure envelopes in terms of total stresses (Murthy 2008).

If the cohesion intercept and angle of shearing resistance are determined using the total stresses, they are named as total or undrained cohesion intercept ( $c$ ) and angle of shearing resistance ( $\phi$ ). The effective stress is the difference between the total stress and the excess pore water pressure. If the pore water pressures are measured during the test, the effective circles can be plotted and the effective strength parameters ( $c'$  and  $\phi'$ ) can be obtained (Mollahasani *et al* 2011).

Accurate determination of the soil shear strength parameters is a major concern in the design of different geotechnical structures. These key parameters can be determined either in the field or in the laboratory. The triaxial compression and direct shear tests are the most common tests for determining the  $c$  and  $\phi$  values in the laboratory. The testing procedures of triaxial and direct shear tests have been standardized by ASTM WK3821 and ASTM-6528, respectively. The direct shear test is commonly used for sandy soils and its procedure is simpler, in comparison with the triaxial test. The tests employed in the field include vane shear test or any other indirect method (El-Maksoud 2006; Murthy 2008; Mollahasani *et al* 2011).

However, experimental determination of the strength parameters is extensive, cumbersome and costly. Further, it is not always possible to conduct the tests on every new situation. In order to cope with such problems, numerical solutions have been developed to estimate the  $c$  and  $\phi$  values. Despite the multivariable dependency of soils, such correlations are mostly developed on the basis of only one soil index property. Incorporating simplifying assumptions into the development of the statistical and numerical methods may also lead to very large errors (Panwar and Seimens 1972; Bowles 1992; Korayem *et al* 1996; Terzaghi *et al* 1996; Shahin *et al* 2001).

Several novel computer-aided data mining methods have been developed so far. An example is a pattern recognition system which learns adaptively from experiences and extracts various

discriminators. Artificial neural networks (ANNs) are one of the most widely-known pattern recognition methods. ANNs have successfully been applied to behavioural modelling of many civil engineering problems (Shahin *et al* 2001; Guven *et al* 2006; Alavi *et al* 2009; Heshmati *et al* 2009; Alavi and Gandomi 2011a). Recently, Mollahasani *et al* (2011) proposed a new model to estimate undrained cohesion intercept ( $c$ ) of soil using multilayer perceptron (MLP) of ANNs. Although ANNs are successful in prediction, they suffer from some shortcomings. The knowledge extracted by ANNs is stored in a set of weights that cannot be properly interpreted. Due to the large complexity to the network structure, ANNs do not give a transparent function relating the inputs to the corresponding outputs. Moreover, the structure and network parameters of ANNs (e.g., number of inputs, transfer functions, and number of hidden layers) should be identified in advance. This is usually done through a time consuming trial and error procedure (Javadi and Rezania 2009; Shahin *et al* 2009).

Genetic programming (GP) (Koza 1992; Banzhaf *et al* 1998) is a new approach for behavioural modelling of civil engineering problems. GP is an extension of genetic algorithms. It can be defined as a supervised machine learning technique that searches a program space instead of a data space. The programs generated by traditional GP are represented as tree structures and expressed in the functional programming language (Koza 1992; Gandomi *et al* 2010). Many researchers have employed GP and its variants to discover complex relationships by use of various types of experimental data (e.g., Guven *et al* 2007; Aytel and Alp 2008; Cevik and Cabalar 2009; Alavi *et al* 2010, 2011; Gandomi *et al* 2011a; Gandomi and Alavi 2011). Kayadelen *et al* (2009) have recently conducted a research with the specific objective of applying ANN and GP to the prediction of the  $\phi'$  of soils. Mousavi *et al* (2011) used a novel hybrid method coupling GP and orthogonal least squares

algorithm (OLS) to derive a prediction model for  $\phi'$ .

Linear genetic programming (LGP) (Brameier and Banzhaf 2007) is a new subset of GP with a linear structure similar to DNA molecule in biological genomes. LGP evolves programs of an imperative language or machine language. Based on the numerical experiments, the LGP approach is able to significantly outperform similar techniques such as gene expression programming (GEP) and multi-expression programming (MEP) (Oltean and Grosan 2003; Brameier and Banzhaf 2007; Baykasoglu *et al* 2008; Alavi and Gandomi 2011b). However, there have been limited scientific efforts aiming to apply LGP to civil engineering problems (e.g., Baykasoglu *et al* 2008; Guven 2009; Guven *et al* 2009; Gandomi *et al* 2010; Alavi and Gandomi 2011b).

In this paper, the LGP technique was utilized to obtain generalized models of  $c$  and  $\phi$  in terms of physical properties of soils including the coarse and fine-grained material contents, grain size characteristics, liquid limit, and soil unit weight. The proposed models were developed based on several unconsolidated-undrained triaxial tests conducted in this study. The paper is organized as follows: section 2 presents brief descriptions of the traditional GP and LGP algorithms. Section 3 outlines the model development using the LGP and regression methods and reviews the results. The detailed performance analyses of the proposed models are discussed in section 4. The verification phase is presented in section 5. The results of the sensitivity and parametric analyses are given in section 6. Finally, concluding remarks are outlined in section 7.

## 2. Genetic programming

GP uses the principle of Darwinian natural selection to create computer programs for solving a problem. It was introduced by Koza (1992) as an extension of genetic algorithms (GAs). Classical GP technique is also referred to as tree-based GP (Koza 1992). The main difference between the GA and GP approaches is that in GP, the evolving programs (individuals) are parse trees rather than fixed-length binary strings. The traditional optimization techniques, like GA, are generally used in parameter optimization to evolve the best values for a given set of model parameters. On the other hand, GP gives the basic structure of the approximation model along with the values of its parameters. GP is relatively a new field of pattern recognition methods in contrast with GA (Alavi *et al* 2011). A survey of the literature reveals the growing interest of research community in GP.

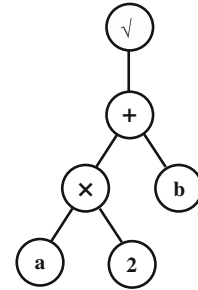


Figure 2. The tree representation of a GP model ( $\sqrt{(a \times 2 + b)}$ ).

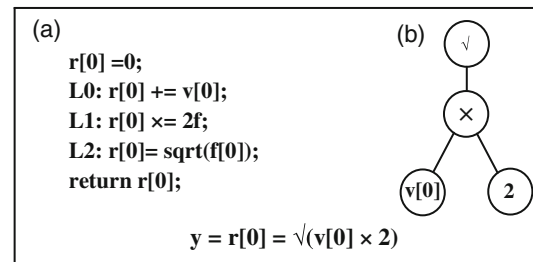


Figure 3. A comparison of the GP program structures: (a) LGP and (b) tree-based GP (Gandomi *et al* 2008).

In GP, a random population of computer programs is created for obtaining high diversity. A population member in GP is a hierarchically structured tree comprising functions and terminals. The functions and terminals are selected from a set of functions and a set of terminals. The functions and terminals are chosen at random and constructed together to form a computer model in a tree-like structure with a root point with branches extending from each function and ending in a terminal (Alavi *et al* 2011). Figure 2 illustrates a simple tree representation of a GP model.

In addition to the classical tree-based GP, there are other types of GP. These are linear and graph-based GP (Banzhaf *et al* 1998). The emphasis of the present study was placed on the linear-based GP techniques. There are some main reasons for using linear GP. Current computer architectures are fundamentally similar to those of 20 years old, which were used in first attempts of GP applications. Almost all architectures represent computer programs in a linear fashion. Also, computers do not naturally run tree-shaped programs. Hence, slow interpreters have to be used as part of tree-based GP. Conversely, by evolving the binary bit patterns actually obeyed by the computer, the use of an expensive interpreter (or compiler) is avoided and GP can run several orders of magnitude faster (Poli *et al* 2007; Alavi and Gandomi 2011b). LGP is a new linear-based GP method.

The linear variants of GP make a clear distinction between the genotype and the phenotype of an individual. Hence, the individuals are represented as linear strings that are decoded and expressed like nonlinear entities (trees) (Oltean and Grosan 2003; Alavi and Gandomi 2011b).

### 2.1 Linear genetic programming

The main characteristic of LGP in comparison with classical tree-based GP is that expressions of a functional programming language (like LISP) are substituted by programs of an imperative language (like C/C++) (Brameier and Banzhaf 2001, 2007). Figure 3 visualizes a comparison of the program structures in LGP and classical tree-based GP. A linear genetic program can be seen as a data flow graph generated by multiple usage of register content. That is, on the functional level, the evolved imperative structure denotes a special directed graph. In classic GP, the data flow is more rigidly determined by the tree structure of the program (Brameier and Banzhaf 2001; Gandomi et al 2011b).

Structurally non-effective codes coexist with effective codes in LGP. An instruction of a linear genetic program is called ‘effective’ at its position if it affects the program output. Non-effective codes in genetic programs, referred to as ‘intron’, represent instructions without affecting the software outcomes. An effective program is a linear genetic program in which introns are eliminated (Gandomi et al 2010). Details of the algorithm used for detecting and eliminating the structural introns can be found in Brameier and Banzhaf (2007). Because of the imperative program structure in LGP, the non-effective instructions can be

efficiently identified. This allows the corresponding effective instructions to be extracted from a program during runtime. Since only the effective programs are executed when testing the fitness cases, evaluation is significantly accelerated (see figure 4) (Brameier and Banzhaf 2007; Gandomi et al 2010).

The instructions from imperative languages are restricted to operations that accept a minimum number of constants or memory variables, called registers ( $r$ ), and assign the result to a destination register, e.g.,  $r_0 := r_1 + 1$ . A part of a linear genetic program in C code is represented as follows (Brameier and Banzhaf 2007; Gandomi et al 2010):

```
void LGP (double r[5])
r[0] = r[2] + 2;
r[2] = r[0] - 3;
if (r[0] > 0)
if (r[2] > 2)
r[3] = r[2] × r[0];
r[2] = r[5] + r[4];
r[0] = tan(r[2]);
```

where register  $r[0]$  holds the final program output. Automatic induction of machine code by genetic programming (AIMGP) is a particular form of LGP. In AIMGP, evolved programs are stored as linear strings of native binary machine code, which are directly executed by the processor. The absence of an interpreter and complex memory handling results in significant speed-up in the AIMGP execution compared to tree-based GP (Gandomi et al 2010, 2011b). The machine-code-based LGP uses the following steps to evolve a computer program (Brameier and Banzhaf 2007; Gandomi et al 2010, 2011b):

- 1) Initializing a population of randomly generated programs.
- 2) Running a tournament. In this step four programs are selected from the population randomly. Based on fitness, two programs are picked as the winners and two as the losers.
- 3) Transforming the winner programs. After that, two winner programs are copied and transformed probabilistically as follows:
  - a) Parts of the winner programs are exchanged with each other to create two new programs (crossover); and/or
  - b) each of the tournament winners are changed randomly to create two new programs (mutation).
- 4) Replacing the loser programs in the tournament with the transformed winner programs. The winners of the tournament remain without change.

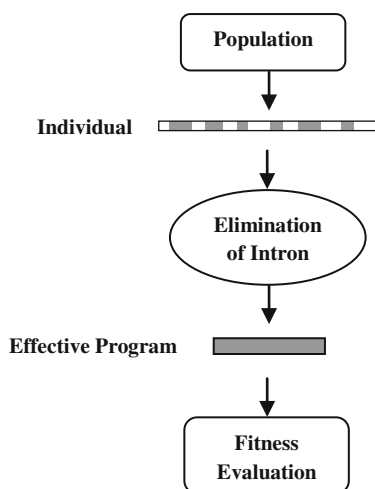


Figure 4. Elimination of non-effective code in LGP. Only effective program is executed (Brameier and Banzhaf 2007).

- 5) Repeating steps two through four until convergence. A program defines the output of the algorithm that simulates the behaviour of the problem to an arbitrary degree of accuracy.

### 3. Modelling of soil shear strength parameters

Precise estimation of the soil shear strength parameters is an essential criterion in geotechnical design process. The angle of shearing resistance represents the interlocking between the soil particles. Cohesion is mainly due to the intermolecular bond between the adsorbed water surrounding each grain, especially in fine-grained soils (Murthy 2008; El-Maksoud 2006). Soils with high plasticity have lower angle of shearing resistance and higher cohesion. Conversely, as the soil grain size increases, the soil internal friction angle increases and its cohesion decreases. Therefore, in a rational manner the main parameters which affect the soil strength parameters are the soil type, soil plasticity, and soil density. This study aimed at obtaining meaningful LGP-based relationships between the undrained cohesion intercept ( $c$ ) and angle of shearing resistance ( $\phi$ ) and the soil physical properties. The most important factors representing the behaviour of  $c$  and  $\phi$  were detected on the basis of literature review (Barends *et al* 1999; El-Maksoud 2006; Murthy 2008; Kayadelen *et al* 2009).  $c$  (kg/cm<sup>2</sup>) and  $\phi$  (°) were considered to be functions of several parameters as follows:

$$c, \phi = f(FC, CC, D_{10}, D_{30}, D_{60}, C_u, C_c, LL, \omega, \tilde{a}, \tilde{a}_d) \quad (1)$$

where,

FC (%): Fine-grained content  
 CC (%): Coarse-grained content  
 $D_{10}$  (mm): Grain size for which 10% of the sample was finer  
 $D_{30}$  (mm): Grain size for which 30% of the sample was finer  
 $D_{60}$  (mm): Grain size for which 60% of the sample was finer  
 $C_u$ : Coefficient of uniformity ( $D_{60}/D_{10}$ )  
 $C_c$ : Coefficient of curvature ( $(D_{30})^2/(D_{60} \times D_{10})$ )  
 LL (%): Liquid limit  
 $\omega$  (%): Moisture content  
 $\gamma$  (gr/cm<sup>3</sup>): Soil unit weight  
 $\gamma_d$  (gr/cm<sup>3</sup>): Soil dry unit weight.

The significant influence of these parameters in determining  $c$  and  $\phi$  is well understood (Murthy 2008). These parameters have been successfully employed by Mollahasani *et al* (2011) and Mousavi

*et al* (2011) to develop empirical models for  $c$  and  $\phi$ , respectively. It is known that the soil shear strength parameters depend on the fabric characteristics, state and consolidation history of the soil. FC, CC,  $D_{10}$ ,  $D_{30}$ ,  $D_{60}$ ,  $C_u$ ,  $C_c$ , and LL represent the intrinsic soil properties.  $\gamma$ ,  $\omega$  and  $\gamma_d$  provide information on the state of the soil and its stress history. They are also indicators of void ratio. Over-consolidation ratio (OCR) could have been included in the analysis instead of  $\gamma$  or  $\gamma_d$ . OCR was not used as it should be obtained from time-consuming laboratory tests, while  $\gamma$  or  $\gamma_d$  can easily be calculated for a soil. The optimal models were selected after developing and controlling several models with different combinations of the input parameters. The following multi-objective strategy was considered to choose the best LGP models:

- Providing the best fitness value on the training set of data.
- Providing the best fitness value on the validation set of data.

#### 3.1 Parameters for performance evaluation

The correlation coefficient ( $R$ ), root mean squared error (RMSE) and mean absolute error (MAE) were used to evaluate the performance of the proposed models.  $R$ , RMSE and MAE are given in the form of equations as follows:

$$R = \frac{\sum_{i=1}^n (h_i - \bar{h}_i)(t_i - \bar{t}_i)}{\sqrt{\sum_{i=1}^n (h_i - \bar{h}_i)^2 \sum_{i=1}^n (t_i - \bar{t}_i)^2}}, \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (h_i - t_i)^2}, \quad (3)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |h_i - t_i|, \quad (4)$$

where  $h_i$  and  $t_i$  are respectively the actual and predicted output values for the  $i$ th output,  $\bar{h}_i$  and  $\bar{t}_i$  are, respectively, the average of the actual and predicted outputs.  $n$  is the number of samples.

#### 3.2 Experimental study and test results

Triaxial compression test is presently the most widely-used procedure for determining the shear strength parameters of soils. In the triaxial soil testing system, field conditions can be simulated in a way that surpasses other testing methods. This is due to the fact that the confining pressure is applied to the soil specimen according



to the uniform lateral *in situ* stresses in triaxial testing apparatus. With respect to the problem statement, three types of testing techniques can be applied, namely unconsolidated undrained (UU), consolidated undrained (CU), and consolidated drained (CD) (Kayadelen *et al* 2009). Within the scope of this study, a series of unconsolidated, undrained, and unsaturated triaxial (UU) tests were performed in accordance with ASTM D2850-87 to determine the shear strength parameters of 88 different undisturbed soil samples. A part of this database has been recently used by Mollahasani *et al* (2011) to develop ANN-based prediction model for  $c$ .

### 3.2.1 Sampling

A total of 50 drillings were performed at different locations in Khorasan and Khuzestan provinces of Iran. The soil samples were manually taken by divers from test pits using metal tubes, 4 and 6 inches in diameter. The samples were obtained at depths ranging from 5 to 30 m and contained no gravel or larger particles. After extracting, the cores were carefully taken to the geotechnical laboratory and maintained in a wet chamber to avoid losing of water content. Undisturbed sub-samples were then extracted from the cores for the geotechnical characterization tests. More, several disturbed soil samples were taken from the sites for the testing purposes.

### 3.2.2 Basic geotechnical characterization tests

An extensive geotechnical laboratory test program was carried out for the basic characterization of soils. These comprised classification tests; water (or moisture) content, defined as the ratio of the weight of water over the weight of dry soil; natural unit weight of the soil, which is simply the ratio of

over volume of an undisturbed soil sample; Atterberg limits (plastic and liquid limits), and grain size distribution. The grain size distribution of the disturbed samples was determined by sieving by use of numbers 4, 8, 16, 30, 50, 100, and 200 sieves. A sedimentation test was carried out throughout a hydrometer analysis. To avoid flocculation of the finer fraction, a dispersing agent (sodium hexametaphosphate) was added to the soil–water mixture before the sedimentation test (Mollahasani *et al* 2011). Figure 5 illustrates the lower and upper limits of the grain size distribution of the samples tested. Different soil types tested were gravelly silt with sand (ML), sandy silt (ML), silty clay with sand (CL-ML), silt with sand (ML), lean clay (CL), silty sand with gravel (SM), poorly-graded gravel with clay (GP-GC), sandy silty clay (CL-ML), lean clay with gravel (CL), lean clay with sand (CL), gravelly lean clay with sand (CL), sandy silt with gravel (ML), and silt (ML).

### 3.2.3 Triaxial compression tests

The soil samples with dimension of 38–50 mm in diameter and 76–100 mm in height were used in the unconsolidated, undrained, and unsaturated triaxial tests. An impermeable, rigid cap and base were used to prevent drainage of the specimens. Both the cap and base had a plane surface of contact and a circular cross section of diameter equal to that of the specimens. Rubber membranes were used to encase each specimen to provide reliable protection against leakage. The samples were closed tightly using plastic O-rings from the top platen and base pedestal. Then, they were placed inside a water-filled cell. The triaxial compression chamber included a bushing and piston aligned with the axis of the specimen. The chamber pressure application device was capable of applying and maintaining a constant chamber pressure at the desired value (within 10 kPa) throughout the test. This device

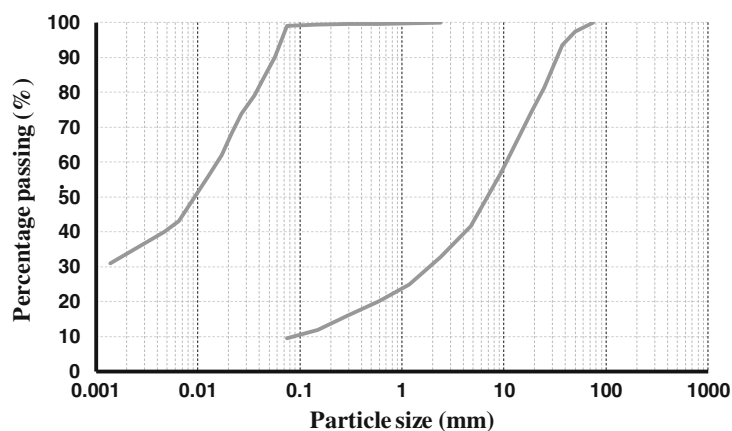


Figure 5. Lower and upper limits of grain size distribution of the soil samples.

Table 1. *Geotechnical properties of soils and test results used for the model development.*

| No. | FC<br>(%) | CC<br>(%) | D <sub>10</sub><br>(mm) | D <sub>30</sub><br>(mm) | D <sub>60</sub><br>(mm) | C <sub>u</sub> | C <sub>c</sub> | LL<br>(%) | $\gamma$<br>(gr/cm <sup>3</sup> ) | $\omega$<br>(%) | $\gamma_d$<br>(gr/cm <sup>3</sup> ) | $c$<br>(kg/cm <sup>2</sup> ) | $\phi$<br>(°) |
|-----|-----------|-----------|-------------------------|-------------------------|-------------------------|----------------|----------------|-----------|-----------------------------------|-----------------|-------------------------------------|------------------------------|---------------|
| 1   | 88        | 12        | 0.00300                 | 0.01700                 | 0.040                   | 13.333         | 2.408          | 25        | 1.5                               | 5.7             | 1.42                                | 0.41                         | 25            |
| 2   | 80.4      | 19.61     | 0.00010                 | 0.00480                 | 0.033                   | 330.000        | 6.982          | 30        | 1.82                              | 10.8            | 1.64                                | 1.07                         | 31            |
| 3   | 95        | 5         | 0.00120                 | 0.01100                 | 0.040                   | 33.333         | 2.521          | 25        | 2.01                              | 15.35           | 1.74                                | 0.77                         | 33            |
| 4   | 97.13     | 2.87      | 0.00100                 | 0.01300                 | 0.033                   | 33.000         | 5.121          | 27        | 2.02                              | 16.5            | 1.73                                | 0.57                         | 25            |
| 5   | 81.6      | 18.4      | 0.00110                 | 0.01100                 | 0.047                   | 42.727         | 2.340          | 26        | 2                                 | 19.8            | 1.67                                | 0.04                         | 23            |
| 6   | 43        | 57        | 0.00250                 | 0.02400                 | 1.000                   | 400.000        | 0.230          | 27        | 1.93                              | 12.8            | 1.71                                | 0.43                         | 29            |
| 7   | 45.25     | 54.75     | 0.00600                 | 0.03300                 | 0.200                   | 33.333         | 0.908          | 22        | 1.74                              | 13.8            | 1.53                                | 0.42                         | 27            |
| 8   | 45.9      | 54.1      | 0.00100                 | 0.03000                 | 0.500                   | 500.000        | 1.800          | 26        | 1.83                              | 13.13           | 1.62                                | 0.37                         | 13            |
| 9   | 65.7      | 34.3      | 0.01100                 | 0.03700                 | 0.060                   | 5.455          | 2.074          | 23        | 1.45                              | 3.4             | 1.40                                | 0.31                         | 23            |
| 10  | 84.5      | 15.5      | 0.00120                 | 0.01400                 | 0.033                   | 27.500         | 4.949          | 23        | 1.49                              | 4.5             | 1.43                                | 0.27                         | 23            |
| 11  | 55.2      | 44.8      | 0.00100                 | 0.01800                 | 0.105                   | 105.000        | 3.086          | 22        | 1.88                              | 14              | 1.65                                | 0.07                         | 18            |
| 12  | 90.6      | 9.4       | 0.00120                 | 0.01100                 | 0.026                   | 21.667         | 3.878          | 25        | 2.07                              | 16.3            | 1.78                                | 0.82                         | 30            |
| 13  | 68.8      | 31.2      | 0.00500                 | 0.02200                 | 0.058                   | 11.600         | 1.669          | 28        | 1.94                              | 17.9            | 1.65                                | 0.08                         | 32            |
| 14  | 72.7      | 27.3      | 0.00100                 | 0.00720                 | 0.043                   | 43.000         | 1.206          | 25        | 1.99                              | 18.5            | 1.68                                | 0.61                         | 26            |
| 15  | 88.5      | 11.5      | 0.00100                 | 0.01400                 | 0.031                   | 31.000         | 6.323          | 26        | 1.85                              | 15.6            | 1.60                                | 0.22                         | 26            |
| 16  | 94.4      | 5.6       | 0.00070                 | 0.01100                 | 0.035                   | 50.000         | 4.939          | 26        | 2.07                              | 16.4            | 1.78                                | 0.65                         | 33            |
| 17  | 52.2      | 47.8      | 0.00100                 | 0.03000                 | 0.100                   | 100.000        | 9.000          | 21        | 1.92                              | 16.3            | 1.65                                | 0.1                          | 17            |
| 18  | 92.6      | 7.4       | 0.00120                 | 0.00900                 | 0.029                   | 24.167         | 2.328          | 28        | 1.95                              | 24.7            | 1.56                                | 0.18                         | 11            |
| 19  | 93.6      | 6.4       | 0.00200                 | 0.01100                 | 0.024                   | 12.000         | 2.521          | 28        | 1.88                              | 23.4            | 1.52                                | 0.08                         | 15            |
| 20  | 92.6      | 7.4       | 0.00700                 | 0.02400                 | 0.040                   | 5.714          | 2.057          | 28        | 1.86                              | 20              | 1.55                                | 0.29                         | 26            |
| 21  | 80.3      | 19.7      | 0.00010                 | 0.01900                 | 0.040                   | 400.000        | 90.250         | 25        | 1.72                              | 9               | 1.58                                | 0.68                         | 31            |
| 22  | 74.6      | 25.4      | 0.00130                 | 0.02600                 | 0.055                   | 42.308         | 9.455          | 22        | 1.9                               | 15.9            | 1.64                                | 0.23                         | 29            |
| 23  | 93.8      | 6.2       | 0.00014                 | 0.00300                 | 0.022                   | 157.143        | 2.922          | 39        | 2.09                              | 15.4            | 1.81                                | 0.49                         | 33            |
| 24  | 59.3      | 40.7      | 0.00300                 | 0.02100                 | 0.080                   | 26.667         | 1.838          | 21        | 1.86                              | 13.25           | 1.64                                | 0.03                         | 21            |
| 25  | 94.8      | 5.2       | 0.00480                 | 0.02100                 | 0.041                   | 8.542          | 2.241          | 27        | 1.8                               | 12.8            | 1.60                                | 0.42                         | 28            |
| 26  | 83.1      | 16.9      | 0.00100                 | 0.01900                 | 0.060                   | 60.000         | 6.017          | 25        | 1.83                              | 15              | 1.59                                | 0.41                         | 31            |
| 27  | 84.1      | 15.9      | 0.00500                 | 0.02800                 | 0.050                   | 10.000         | 3.136          | 25        | 2.03                              | 22              | 1.66                                | 0.08                         | 29            |
| 28  | 47.8      | 52.2      | 0.01000                 | 0.03600                 | 0.300                   | 30.000         | 0.432          | 21        | 1.69                              | 11              | 1.52                                | 0.06                         | 24            |
| 29  | 93.3      | 6.7       | 0.00010                 | 0.00100                 | 0.013                   | 130.000        | 0.769          | 33        | 1.87                              | 29.4            | 1.45                                | 0.31                         | 22            |
| 30  | 87.3      | 12.7      | 0.00070                 | 0.00600                 | 0.030                   | 42.857         | 1.714          | 31        | 1.97                              | 20.4            | 1.64                                | 0.05                         | 32            |
| 31  | 58.4      | 41.6      | 0.00100                 | 0.02300                 | 0.080                   | 80.000         | 6.613          | 23        | 1.92                              | 11.1            | 1.73                                | 0.37                         | 8             |
| 32  | 84.5      | 15.5      | 0.01200                 | 0.04300                 | 0.059                   | 4.917          | 2.612          | 22        | 2.01                              | 16.6            | 1.72                                | 0.06                         | 14            |
| 33  | 89.9      | 10.1      | 0.00100                 | 0.01200                 | 0.033                   | 33.000         | 4.364          | 24        | 1.81                              | 10.8            | 1.63                                | 0.78                         | 24            |
| 34  | 95.1      | 4.9       | 0.00001                 | 0.00100                 | 0.010                   | 1000.000       | 10.000         | 29        | 1.96                              | 30              | 1.51                                | 0.07                         | 11            |
| 35  | 94        | 6         | 0.00700                 | 0.02100                 | 0.038                   | 5.429          | 1.658          | 27        | 1.87                              | 18.5            | 1.58                                | 0.41                         | 29            |
| 36  | 95.6      | 4.4       | 0.00790                 | 0.01400                 | 0.028                   | 3.544          | 0.886          | 26        | 1.94                              | 19.1            | 1.63                                | 0.37                         | 24            |
| 37  | 93.2      | 6.8       | 0.00100                 | 0.01300                 | 0.042                   | 42.000         | 4.024          | 24        | 1.81                              | 10.15           | 1.64                                | 0.32                         | 29            |
| 38  | 95.4      | 4.6       | 0.00100                 | 0.00940                 | 0.030                   | 30.000         | 2.945          | 27        | 2.08                              | 20.6            | 1.72                                | 0.13                         | 11            |
| 39  | 72.9      | 27.1      | 0.00100                 | 0.01800                 | 0.050                   | 50.000         | 6.480          | 24        | 1.62                              | 10.25           | 1.47                                | 0.09                         | 27            |
| 40  | 10.6      | 89.4      | 0.07000                 | 1.10000                 | 4.000                   | 57.143         | 4.321          | 39        | 1.55                              | 3.6             | 1.50                                | 0.32                         | 26            |
| 41  | 9.6       | 90.4      | 0.09000                 | 1.90000                 | 10.000                  | 111.111        | 4.011          | 31        | 1.86                              | 8.5             | 1.71                                | 1.12                         | 36            |
| 42  | 94        | 6         | 0.00120                 | 0.01500                 | 0.031                   | 25.833         | 6.048          | 36        | 1.78                              | 15.25           | 1.54                                | 0.26                         | 30            |
| 43  | 83.4      | 16.6      | 0.00100                 | 0.00640                 | 0.030                   | 30.000         | 1.365          | 40        | 1.89                              | 17.1            | 1.61                                | 0.64                         | 28            |
| 44  | 93.8      | 6.2       | 0.00050                 | 0.00600                 | 0.018                   | 36.000         | 4.000          | 28        | 2.01                              | 20.25           | 1.67                                | 0.13                         | 17            |
| 45  | 75.5      | 24.5      | 0.00140                 | 0.01000                 | 0.040                   | 28.571         | 1.786          | 27        | 1.78                              | 7.3             | 1.66                                | 1.15                         | 33            |
| 46  | 97.8      | 2.2       | 0.00100                 | 0.00820                 | 0.021                   | 21.000         | 3.202          | 29        | 1.97                              | 21.2            | 1.63                                | 0.11                         | 21            |
| 47  | 95.3      | 4.7       | 0.00600                 | 0.02100                 | 0.051                   | 8.500          | 1.441          | 26        | 1.84                              | 21.2            | 1.52                                | 0.02                         | 26            |
| 48  | 94.4      | 5.6       | 0.00900                 | 0.02600                 | 0.050                   | 5.556          | 1.502          | 28        | 1.98                              | 19.2            | 1.66                                | 0.18                         | 14            |
| 49  | 94.7      | 5.3       | 0.00070                 | 0.00700                 | 0.021                   | 30.000         | 3.333          | 27        | 2.06                              | 20              | 1.72                                | 0.04                         | 17            |
| 50  | 87.5      | 12.5      | 0.00100                 | 0.00610                 | 0.020                   | 20.000         | 1.861          | 39        | 2.07                              | 21.1            | 1.71                                | 0.58                         | 29            |
| 51  | 91        | 9         | 0.00110                 | 0.00800                 | 0.032                   | 29.091         | 1.818          | 41        | 1.97                              | 12.9            | 1.74                                | 1.16                         | 30            |
| 52  | 86.3      | 13.7      | 0.00300                 | 0.01400                 | 0.040                   | 13.333         | 1.633          | 38        | 2.07                              | 15.4            | 1.79                                | 1.75                         | 26            |
| 53  | 58.2      | 41.8      | 0.01000                 | 0.02800                 | 0.100                   | 10.000         | 0.784          | 37        | 1.87                              | 13.33           | 1.65                                | 0.84                         | 22            |
| 54  | 89.9      | 10.1      | 0.00600                 | 0.02000                 | 0.037                   | 6.167          | 1.802          | 27        | 1.54                              | 10.4            | 1.39                                | 0.59                         | 28            |

Table 1. (*Continued.*)

| No. | FC<br>(%) | CC<br>(%) | D <sub>10</sub><br>(mm) | D <sub>30</sub><br>(mm) | D <sub>60</sub><br>(mm) | C <sub>u</sub> | C <sub>c</sub> | LL<br>(%) | $\gamma$<br>(gr/cm <sup>3</sup> ) | $\omega$<br>(%) | $\gamma_d$<br>(gr/cm <sup>3</sup> ) | $c$<br>(kg/cm <sup>2</sup> ) | $\phi$<br>(°) |
|-----|-----------|-----------|-------------------------|-------------------------|-------------------------|----------------|----------------|-----------|-----------------------------------|-----------------|-------------------------------------|------------------------------|---------------|
| 55  | 83.3      | 16.7      | 0.00600                 | 0.01600                 | 0.033                   | 5.500          | 1.293          | 28        | 1.81                              | 20.7            | 1.50                                | 0.08                         | 27            |
| 56  | 79        | 21        | 0.00300                 | 0.01200                 | 0.043                   | 14.333         | 1.116          | 23        | 1.54                              | 9.1             | 1.41                                | 0.18                         | 28            |
| 57  | 64.4      | 35.6      | 0.00610                 | 0.02200                 | 0.050                   | 8.197          | 1.587          | 22        | 1.7                               | 5.4             | 1.61                                | 0.5                          | 30            |
| 58  | 91.6      | 8.4       | 0.00900                 | 0.01300                 | 0.030                   | 3.333          | 0.626          | 23        | 1.75                              | 20.35           | 1.45                                | 0.03                         | 26            |
| 59  | 97.4      | 2.6       | 0.00900                 | 0.02600                 | 0.048                   | 5.333          | 1.565          | 25        | 1.43                              | 3.5             | 1.38                                | 0.44                         | 27            |
| 60  | 97.6      | 2.4       | 0.00200                 | 0.00810                 | 0.020                   | 10.000         | 1.640          | 28        | 1.67                              | 11.35           | 1.50                                | 0.23                         | 29            |
| 61  | 94.4      | 5.6       | 0.00180                 | 0.01000                 | 0.027                   | 15.000         | 2.058          | 29        | 1.62                              | 13.1            | 1.43                                | 0.04                         | 23            |
| 62  | 91.1      | 8.9       | 0.00200                 | 0.01700                 | 0.031                   | 15.500         | 4.661          | 25        | 1.92                              | 20.96           | 1.59                                | 0.04                         | 12            |
| 63  | 88        | 12        | 0.00300                 | 0.01500                 | 0.034                   | 11.333         | 2.206          | 28        | 1.82                              | 23.45           | 1.47                                | 0.06                         | 15            |
| 64  | 94.5      | 5.5       | 0.00450                 | 0.02000                 | 0.034                   | 7.556          | 2.614          | 25        | 1.69                              | 14.6            | 1.47                                | 0.07                         | 31            |
| 65  | 95.3      | 4.7       | 0.00060                 | 0.00600                 | 0.020                   | 33.333         | 3.000          | 27        | 1.9                               | 20.5            | 1.58                                | 0.07                         | 22            |
| 66  | 97.3      | 2.7       | 0.00050                 | 0.00200                 | 0.011                   | 22.000         | 0.727          | 29        | 1.96                              | 23.2            | 1.59                                | 0.41                         | 20            |
| 67  | 90.4      | 9.6       | 0.00100                 | 0.00400                 | 0.028                   | 28.000         | 0.571          | 30        | 1.96                              | 28.43           | 1.53                                | 0.1                          | 11            |
| 68  | 97.6      | 2.4       | 0.00150                 | 0.00700                 | 0.020                   | 13.333         | 1.633          | 27        | 1.86                              | 14.75           | 1.62                                | 0.78                         | 31            |
| 69  | 91.4      | 8.6       | 0.00700                 | 0.00230                 | 0.015                   | 2.143          | 0.050          | 29        | 1.95                              | 27.92           | 1.52                                | 0.23                         | 17            |
| 70  | 99.1      | 0.9       | 0.00060                 | 0.00190                 | 0.008                   | 13.333         | 0.752          | 31        | 1.93                              | 7.5             | 1.80                                | 0.74                         | 37            |
| 71  | 95.2      | 4.8       | 0.00070                 | 0.01400                 | 0.031                   | 44.286         | 9.032          | 27        | 1.77                              | 14.38           | 1.55                                | 0.59                         | 24            |
| 72  | 61        | 39        | 0.00170                 | 0.02300                 | 0.070                   | 41.176         | 4.445          | 33        | 1.98                              | 18.55           | 1.67                                | 0.22                         | 18            |
| 73  | 56.1      | 43.9      | 0.00100                 | 0.02000                 | 0.100                   | 100.000        | 4.000          | 34        | 1.9                               | 26.15           | 1.51                                | 0.3                          | 3             |
| 74  | 95.5      | 4.5       | 0.00070                 | 0.01600                 | 0.032                   | 45.714         | 11.429         | 27        | 2.01                              | 25.7            | 1.60                                | 0.13                         | 1             |
| 75  | 96        | 4         | 0.00140                 | 0.00690                 | 0.025                   | 17.857         | 1.360          | 28        | 2.05                              | 13.4            | 1.81                                | 0.39                         | 18            |
| 76  | 91.8      | 8.2       | 0.00300                 | 0.02000                 | 0.036                   | 12.000         | 3.704          | 30        | 2                                 | 19.8            | 1.67                                | 0.14                         | 24            |
| 77  | 93.8      | 6.2       | 0.00190                 | 0.01800                 | 0.040                   | 21.053         | 4.263          | 21        | 1.78                              | 14.9            | 1.55                                | 0.32                         | 18            |
| 78  | 99.1      | 0.9       | 0.00100                 | 0.01300                 | 0.030                   | 30.000         | 5.633          | 23        | 1.96                              | 14.9            | 1.71                                | 0.17                         | 28            |
| 79  | 97        | 3         | 0.00100                 | 0.00300                 | 0.009                   | 9.000          | 1.000          | 24        | 1.96                              | 24.5            | 1.57                                | 0.37                         | 3             |
| 80  | 52        | 48        | 0.00010                 | 0.01200                 | 0.115                   | 1150.000       | 12.522         | 24        | 2.06                              | 17.9            | 1.75                                | 0.56                         | 22            |
| 81  | 87        | 13        | 0.00280                 | 0.01600                 | 0.040                   | 14.286         | 2.286          | 22        | 1.96                              | 25.8            | 1.56                                | 0.26                         | 16            |
| 82  | 85.5      | 14.5      | 0.00430                 | 0.02000                 | 0.032                   | 7.442          | 2.907          | 23        | 1.87                              | 14.3            | 1.64                                | 0.5                          | 25            |
| 83  | 95.6      | 4.4       | 0.00200                 | 0.01800                 | 0.038                   | 19.000         | 4.263          | 26        | 1.88                              | 23.5            | 1.52                                | 0.13                         | 26            |
| 84  | 78        | 22        | 0.00060                 | 0.00680                 | 0.036                   | 60.000         | 2.141          | 30        | 1.74                              | 12.5            | 1.55                                | 0.62                         | 25            |
| 85  | 81.8      | 18.2      | 0.00190                 | 0.01200                 | 0.041                   | 21.579         | 1.849          | 29        | 1.76                              | 16.5            | 1.51                                | 0.16                         | 25            |
| 86  | 94.7      | 5.3       | 0.00100                 | 0.00700                 | 0.022                   | 22.000         | 2.227          | 29        | 2.05                              | 19.1            | 1.72                                | 0.46                         | 15            |
| 87  | 85.7      | 14.3      | 0.00100                 | 0.01000                 | 0.035                   | 35.000         | 2.857          | 28        | 1.91                              | 21.6            | 1.57                                | 0.33                         | 17            |
| 88  | 41.4      | 58.6      | 0.00400                 | 0.02800                 | 0.480                   | 120.000        | 0.408          | 25        | 2.04                              | 23              | 1.66                                | 0.32                         | 12            |

was connected to the triaxial chamber through pressure-control devices. Pressure was applied by compressed air. In order to obtain the soil failure, an axial stress was implemented on the top of the soil specimens by a frictionless ram through the top of the cell. For this aim, a screw jack driven by an electric motor through geared transmission device was used. The electric motor was capable of providing axial load with rates within 0.05–10 mm/min. A proving ring with sufficient accuracy was employed to measure the deviator load applied to the specimen. The axial deformation of the specimens was measured using a dial gauge reading to an accuracy of 0.001 mm (Mollahasani *et al* 2011). As the case of the tests was strain-controlled, the device was attached to the bottom plate of the machine which was moving up at a constant rate. Special care was taken in aligning

the axial load device, the axial load measuring device and the triaxial chamber to prevent application of lateral force to the piston during testing. The axial load was applied at the desired strain rate, approximately 10 min after the application of chamber pressure. Proving ring readings were recorded for intervals of axial deformation. Sufficient readings were taken to capture the stress-strain curve. If the sample did not fail showing a reduction in the deviator load, the loading was continued to 15% strain. If the residual strengths were required, the tests were continued further. At the end of the tests, the specimens were taken out, failure patterns were noted and moisture contents of the samples were determined. The triaxial tests were performed on at least two other identical samples at different chamber pressures to construct the failure



envelope and to determine the shear strength parameters (Mollahasani *et al* 2011). The test results and index properties of the tested samples are summarized in table 1. The information cited in this table includes FC, CC,  $D_{10}$ ,  $D_{30}$ ,  $D_{60}$ ,  $C_u$ ,  $C_c$ , LL,  $\omega$ ,  $\gamma$ , and  $\gamma_d$ .  $c$  and  $\phi$  are the measured soil shear strength parameters. A major part of the database contains the laboratory test results on fine-grained soil samples. The descriptive statistics of the data used in this study are given in table 2.

### 3.3 Data preprocessing

For the LGP analysis, the datasets were randomly divided into training, validating and testing subsets. Training data were used for learning (genetic evolution). The validation data were used to investigate the generalization capability of the evolved programs by the use of unseen data (model selection). In other words, the training and validation datasets were used to select the best evolved programs and included in the training process. The testing data were used to measure the performance of the program evolved by LGP on data that played no role in building the model (Gandomi *et al* 2010). Of the 88 data for the prediction of  $c$  and  $\phi$ , 52 and 18 datasets were used for training and validation purposes, respectively. The remaining 18 (20%) sets were taken for the testing of the models. In order to obtain a consistent data division, several combinations of the training, validation and testing sets were considered. The selection was such that the maximum, minimum, mean and standard deviation of parameters were consistent in the datasets (Gandomi *et al* 2010).

Although normalization is not strictly necessary in the GP-based analysis, better results are usually reached after normalizing the variables. Normalization also speeds up the process. Thus, both the input and the output variables were normalized. After controlling several normalization methods (Swingler 1996; Mesbahi 2000), the following method was used to normalize the variables to a range of  $[L, U]$ :

$$X_n = ax + b \quad (5)$$

where

$$a = \frac{(U - L)}{(X_{\max} - X_{\min})},$$

and

$$b = U - aX_{\max},$$

in which  $X_{\max}$  and  $X_{\min}$  are the maximum and minimum values of the variable and  $X_n$  is the normalized value. In the present study,  $L = 0.05$  and  $U = 0.95$ .

Table 2. Descriptive statistics of the variables used for the model development.

| Parameter          | FC (%)  | CC (%)  | $D_{10}$ (mm) | $D_{30}$ (mm) | $D_{60}$ (mm) | $C_u$    | $C_c$  | LL (%)  | $\gamma$ (gr/cm <sup>3</sup> ) | $\omega$ (%) | $\gamma_d$ (gr/cm <sup>3</sup> ) | $c$ (kg/cm <sup>2</sup> ) | $\phi$ (°) |
|--------------------|---------|---------|---------------|---------------|---------------|----------|--------|---------|--------------------------------|--------------|----------------------------------|---------------------------|------------|
| Mean               | 81.89   | 18.11   | 0.0045        | 0.049         | 0.224         | 72.62    | 4.16   | 27.40   | 1.87                           | 16.46        | 1.60                             | 0.36                      | 23.09      |
| Standard error     | 2.01    | 2.01    | 0.0013        | 0.025         | 0.122         | 18.78    | 1.03   | 0.49    | 0.02                           | 0.65         | 0.01                             | 0.03                      | 0.82       |
| Median             | 89.90   | 10.10   | 0.0013        | 0.014         | 0.036         | 27.08    | 2.33   | 27.00   | 1.90                           | 16.30        | 1.61                             | 0.31                      | 25.00      |
| Mode               | 94.40   | 5.60    | 0.0010        | 0.011         | 0.040         | 13.33    | 4.00   | 27.00   | 1.96                           | 10.80        | 1.67                             | 0.41                      | 26.00      |
| Standard deviation | 18.90   | 18.90   | 0.0120        | 0.231         | 1.142         | 176.14   | 9.62   | 4.59    | 0.16                           | 6.08         | 0.11                             | 0.32                      | 7.66       |
| Sample variance    | 357.16  | 357.16  | 0.0001        | 0.053         | 1.303         | 31023.65 | 92.55  | 21.05   | 0.02                           | 37.02        | 0.01                             | 0.10                      | 58.75      |
| Kurtosis           | 3.47    | 3.47    | 39.2224       | 53.111        | 64.619        | 24.38    | 75.89  | 1.36    | 0.56                           | -0.26        | -0.62                            | 3.67                      | 0.24       |
| Skewness           | -1.83   | 1.83    | 6.1209        | 7.171         | 7.813         | 4.74     | 8.44   | 1.20    | -0.97                          | -0.06        | -0.09                            | 1.61                      | -0.78      |
| Range              | 89.50   | 89.50   | 0.0900        | 1.899         | 9.992         | 1147.86  | 90.20  | 20.00   | 0.66                           | 26.60        | 0.43                             | 1.73                      | 36.00      |
| Minimum            | 9.60    | 0.90    | 0.0000        | 0.001         | 0.008         | 2.14     | 0.05   | 21.00   | 1.43                           | 3.40         | 1.38                             | 0.02                      | 1          |
| Maximum            | 99.10   | 90.40   | 0.0900        | 1.900         | 10.000        | 1150.00  | 90.25  | 41.00   | 2.09                           | 30.00        | 1.81                             | 1.75                      | 37         |
| Sum                | 7206.67 | 1593.33 | 0.3969        | 4.308         | 19.693        | 6390.16  | 366.03 | 2411.00 | 164.41                         | 1448.15      | 141.15                           | 31.98                     | 2032       |

Table 3. *Parameter settings for the LGP algorithm.*

| Parameter                     | Settings  |
|-------------------------------|---|
| Function set                  | +, −, ×, /, sqrt (√),<br>fabs (  ), power, sin,<br>cos, <, if ≤ |
| Population size               | 500–25000   |
| Maximum program size          | 256–1024  |
| Initial program size          | 80  |
| Fitness calculation           | Square error  |
| Crossover rate (%)            | 50, 95  |
| Homologous crossover (%)      | 95  |
| Mutation rate (%)             | 90  |
| Block mutation rate (%)       | 30  |
| Instruction mutation rate (%) | 30  |
| Data mutation rate (%)        | 40  |
| Number of demes               | 20  |

### 3.4 Model development using LGP

The available database was used for generating prediction models relating  $c$  and  $\phi$  to FC, CC,  $D_{10}$ ,  $D_{30}$ ,  $D_{60}$ ,  $C_u$ ,  $C_c$ , LL,  $\gamma$ ,  $\omega$ , and  $\gamma_d$ . Various parameters are involved in the LGP predictive algorithm. The parameter settings are shown in table 3. Several runs were conducted to obtain a parameterization of LGP that provided enough robustness and generalization to solve the problem. The LGP parameters were changed for different runs. The levels considered for the parameters were on the basis of both previously suggested values (Feldt and Nordin 2000; Francone 2001; Brameier and Banzhaf 2007; Gandomi *et al* 2010; Alavi and Gandomi 2011b) and making several preliminary runs and observing the performance behaviour. The population size sets the number of programs in the population that LGP will evolve. A run takes longer with a larger population size. The correct number for this parameter depends on the number of possible solutions and problem complexity. A fairly large number of initial populations were tested to find models with minimal error. The program was run until there was no longer significant improvement in the performance of the models. Mutation and crossover rates are the probabilities that an offspring is subjected to the mutation and crossover operations, respectively. Although most GP systems use a very low mutation rate, the machine-code-based LGP benefits from a much higher mutation rate (Francone 2001, 2004). Therefore, the mutation rate was set to 90%. At the low level the crossover rate was 50% and at the high level it was set to 95%. The lengths of the evolved programs in runs were controlled by initial and maximum program size parameters. The initial program size parameter sets the size of the programs in the first population at the start of each run. The maximum program size parameter sets

the maximum length of the other programs evolved during each run. The success of the LGP algorithm usually increases with increasing these parameters. In this case, the complexity of the evolved functions increase and the speed of the algorithm decreases (Gandomi *et al* 2011b). The initial program size was set to 80 bytes. The range of the maximum program sizes was between 256 and 1024 bytes. Demes are semi-isolated subpopulations in which evolution proceeds faster than a single population of equal size. The number of demes is related to the way that the population of programs is divided. In this study, the number of demes was set to 20. The average of the squared raw errors was used to evaluate the fitness of the evolved programs. To implement the LGP algorithm, a computer software, called Discipulus (Conrads *et al* 2004), was used which works on the basis of the AIMGP platform.

Two types of LGP models were used in this paper: a program model and a team model. A program model is a single solution. A team model combines single program models in order to produce better results than any of the single program models. A team solution employs a combination of an odd number of single solutions (minimum 1; maximum 9). After completing a project, the evolved single and team solutions are automatically written in Java, C++, or Intel assembler code. The resulting codes may be compiled into a DLL or COM object and called from the optimization routines (Deschaine 2000). Discipulus interactive evaluator mode or any C++ environment (TurboC) can be employed to run the codes evolved by LGP.

#### 3.4.1 LGP single and team solutions for predicting soil strength parameters

An extensive trial study was performed to select the most relevant input parameters for the LGP models. Several LGP models were developed using different combinations of the input parameters. Tables 4 and 5 summarize some of the optimal LGP models with their selected input parameters for predicting  $c$  and  $\phi$ , respectively. It can be seen from these tables that performances of the models are generally improved as the input parameters are increased. With the exception of the LGP single model for  $c$ , the best single and team LGP models for predicting the soil strength parameters were built using FC,  $D_{30}$ ,  $C_u$ , LL,  $\omega$ , and  $\gamma_d$  (model type IX). As can be seen in table 4, the best LGP single solution for predicting  $c$  was built using FC,  $C_u$ , LL,  $\omega$ , and  $\gamma_d$  (model type X). Since  $C_c$  did not improve the performance of the models,  $C_u$  and  $D_{30}$  seem to be sufficient representatives of the soil grain size distribution characteristics. Figures 6–9 present the  $c$  and  $\phi$  predictions made by the best

Table 4. Summary of the optimal LGP single and team solutions developed with various input parameters for  $c$  prediction.

| Model type | Input parameters  | R values for the LGP<br>single solutions |         | R values for the LGP<br>team solutions |         |
|------------|---|--|---------|--|---------|
|            |   | Training & validation                    | Testing | Training & validation                  | Testing |
| I          | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$ | 0.849                                    | 0.844   | 0.867                                  | 0.861   |
| II         | CC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$ | 0.826                                    | 0.807   | 0.838                                  | 0.837   |
| III        | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\gamma$              | 0.840                                    | 0.843   | 0.855                                  | 0.843   |
| IV         | CC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\gamma$              | 0.807                                    | 0.775   | 0.832                                  | 0.820   |
| V          | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , LL, $\omega$ , $\gamma_d$                  | 0.826                                    | 0.783   | 0.837                                  | 0.842   |
| VI         | D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$     | 0.784                                    | 0.774   | 0.775                                  | 0.752   |
| VII        | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , LL, $\omega$ , $\gamma_d$                                   | 0.809                                    | 0.788   | 0.797                                  | 0.783   |
| VIII       | FC, D <sub>60</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$  | 0.867                                    | 0.841   | 0.873                                  | 0.884   |
| IX         | FC, D <sub>30</sub> , C <sub>u</sub> , LL, $\omega$ , $\gamma_d$  | 0.876                                    | 0.848   | 0.896                                  | 0.892   |
| X          | FC, C <sub>u</sub> , LL, $\omega$ , $\gamma_d$  | 0.879                                    | 0.876   | 0.884                                  | 0.868   |
| XI         | D <sub>30</sub> , C <sub>u</sub> , LL, $\omega$ , $\gamma_d$  | 0.710                                    | 0.693   | 0.779                                  | 0.759   |
| XII        | FC, LL, $\omega$ , $\gamma_d$   | 0.780                                    | 0.758   | 0.800                                  | 0.807   |
| XIII       | FC, LL, $\gamma$  | 0.746                                    | 0.704   | 0.734                                  | 0.708   |

Table 5. Summary of the optimal LGP single and team solutions developed with various input parameters for  $\phi$  prediction.

| Model type | Input parameters  | R values for the LGP<br>single solutions |         | R values for the LGP<br>team solutions |         |
|------------|---|--|---------|--|---------|
|            |   | Training & validation                    | Testing | Training & validation                  | Testing |
| I          | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$ | 0.806                                    | 0.797   | 0.877                                  | 0.856   |
| II         | CC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$ | 0.775                                    | 0.769   | 0.838                                  | 0.843   |
| III        | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\gamma$              | 0.797                                    | 0.803   | 0.862                                  | 0.841   |
| IV         | CC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\gamma$              | 0.782                                    | 0.780   | 0.836                                  | 0.812   |
| V          | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , LL, $\omega$ , $\gamma_d$                  | 0.814                                    | 0.799   | 0.844                                  | 0.834   |
| VI         | D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , C <sub>u</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$     | 0.784                                    | 0.777   | 0.829                                  | 0.786   |
| VII        | FC, D <sub>10</sub> , D <sub>30</sub> , D <sub>60</sub> , LL, $\omega$ , $\gamma_d$                                   | 0.809                                    | 0.787   | 0.843                                  | 0.824   |
| VIII       | FC, D <sub>60</sub> , C <sub>c</sub> , LL, $\omega$ , $\gamma_d$  | 0.828                                    | 0.799   | 0.886                                  | 0.855   |
| IX         | FC, D <sub>30</sub> , C <sub>u</sub> , LL, $\omega$ , $\gamma_d$  | 0.838                                    | 0.823   | 0.892                                  | 0.863   |
| X          | FC, C <sub>u</sub> , LL, $\omega$ , $\gamma_d$  | 0.832                                    | 0.815   | 0.877                                  | 0.851   |
| XI         | D <sub>30</sub> , C <sub>u</sub> , LL, $\omega$ , $\gamma_d$  | 0.679                                    | 0.657   | 0.798                                  | 0.785   |
| XII        | FC, LL, $\omega$ , $\gamma_d$   | 0.732                                    | 0.727   | 0.788                                  | 0.756   |
| XIII       | FC, LL, $\gamma$  | 0.688                                    | 0.700   | 0.776                                  | 0.753   |

single and team solutions. The final LGP single programs obtained at the end of training in C++ are given in Appendix A. These programs can be run in the Discipulus interactive evaluator mode or in C++ environment. The team solutions for the  $c$  and  $\phi$  prediction were composed of nine single solutions. Due to the complexity, the best team solutions were not presented herein. The interested readers may contact the authors for the best LGP team solutions in C++ codes. Note that LGPs can be converted into a functional representation by successive replacements of variables starting with the last effective instruction (Oltean and Grosan 2003; Guven 2009; Gandomi *et al* 2010). In the present study, the optimal LGP single solutions were lengthy and contained various mathematical functions, comparison instructions, data transfer instructions, and conditional branches. Therefore,

they could not be converted into straightforward and practical design equations.

### 3.5 Model development using regression analysis

A multivariable least squares regression (MLSR) (Ryan 1997) analysis was performed to have an idea about the predictive power of the best LGP models, in comparison with a classical statistical approach. The method of LSR is extensively used in regression analysis primarily because of its interesting nature. The MLSR prediction equations relate  $c$  and  $\phi$  to the predictor variables as follows:

$$c, \phi = \acute{a}_1 FC + \acute{a}_2 D_{30} + \acute{a}_3 C_u + \acute{a}_4 LL + \acute{a}_5 \omega + \acute{a}_6 \gamma_d + \acute{a}_7, \quad (6)$$

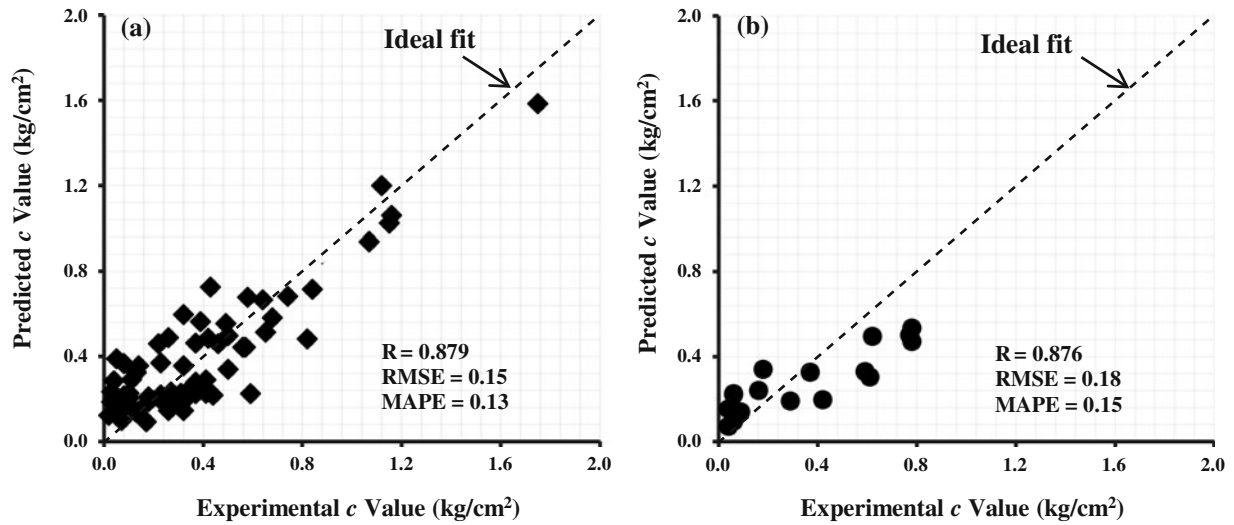


Figure 6. Experimental *versus* predicted  $c$  values using the best LGP single solution (model X): (a) training and validation data, and (b) testing data.

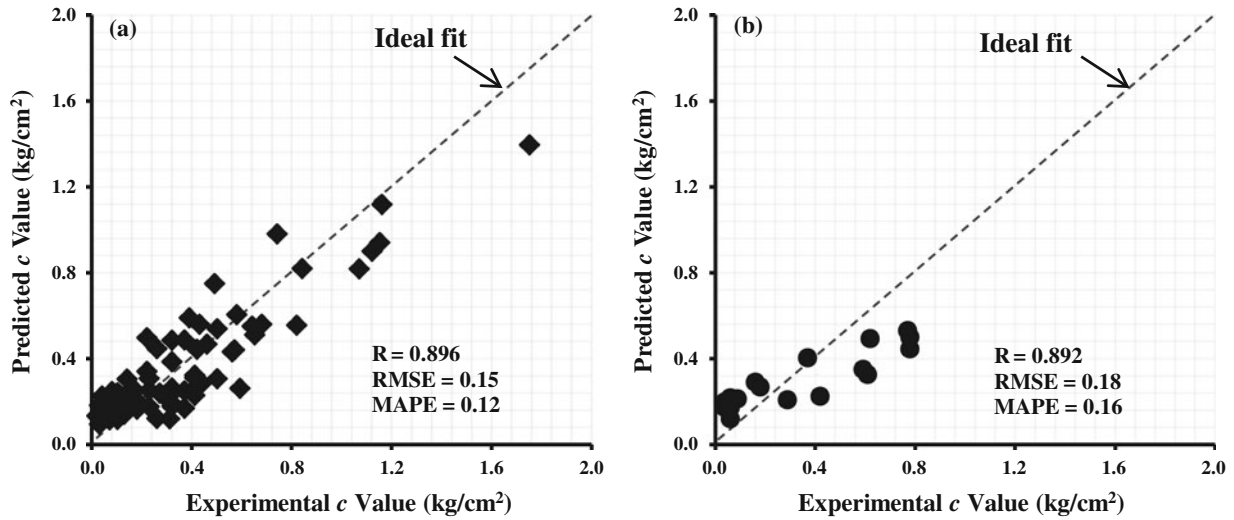


Figure 7. Experimental *versus* predicted  $c$  values using the best LGP team solution (model IX): (a) training and validation data, and (b) testing data.

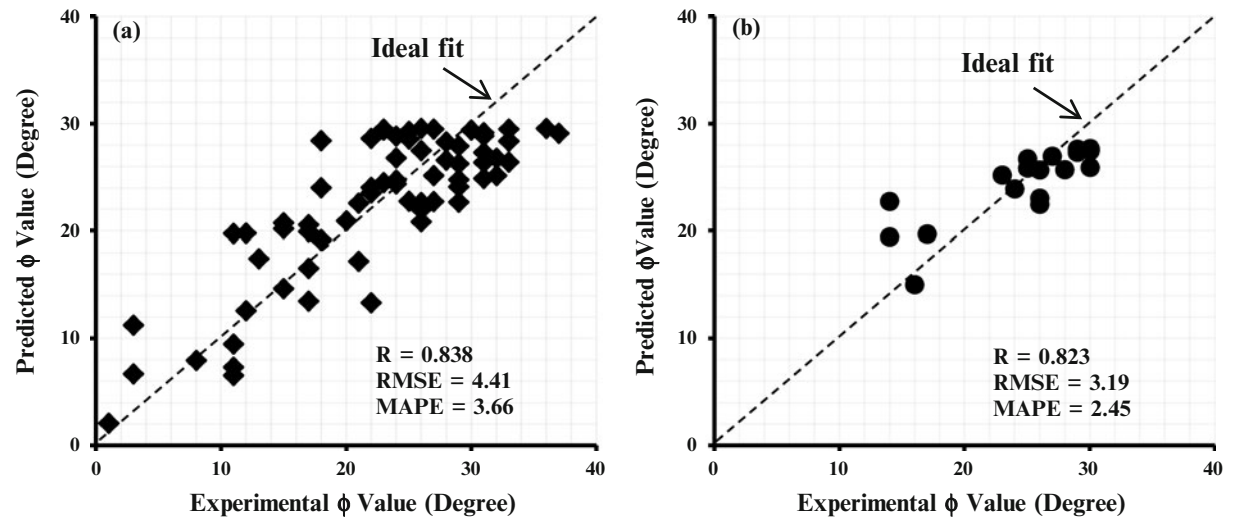


Figure 8. Experimental *versus* predicted  $\phi$  values using the best LGP single solution (model IX): (a) training and validation data, and (b) testing data.

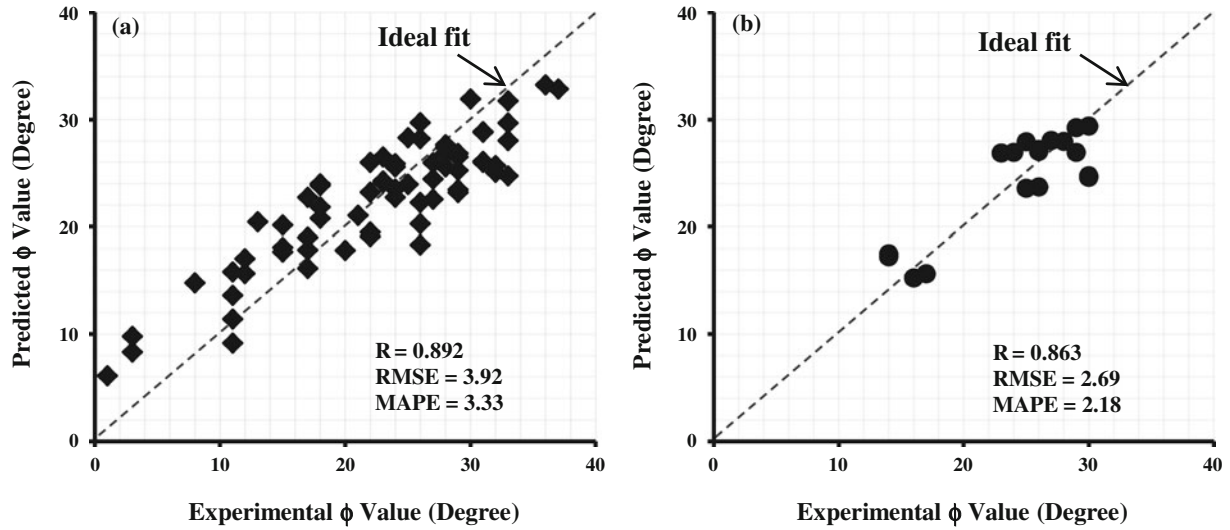


Figure 9. Experimental *versus* predicted  $\phi$  values using the best LGP team solution (model IX): (a) training and validation data, and (b) testing data.

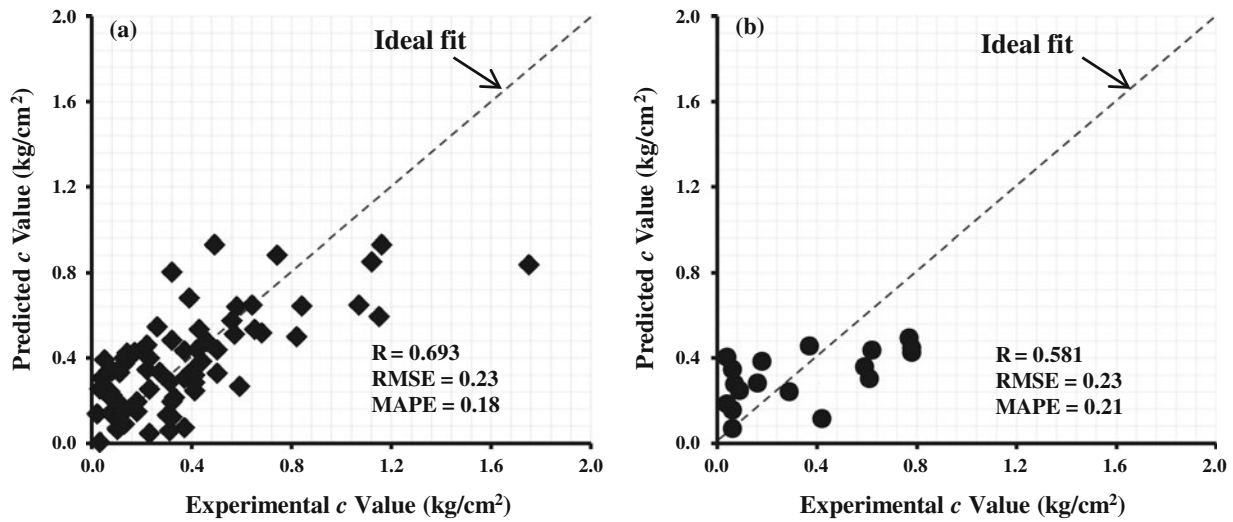


Figure 10. Experimental *versus* predicted  $c$  values using the MLSR model: (a) training data, and (b) testing data.

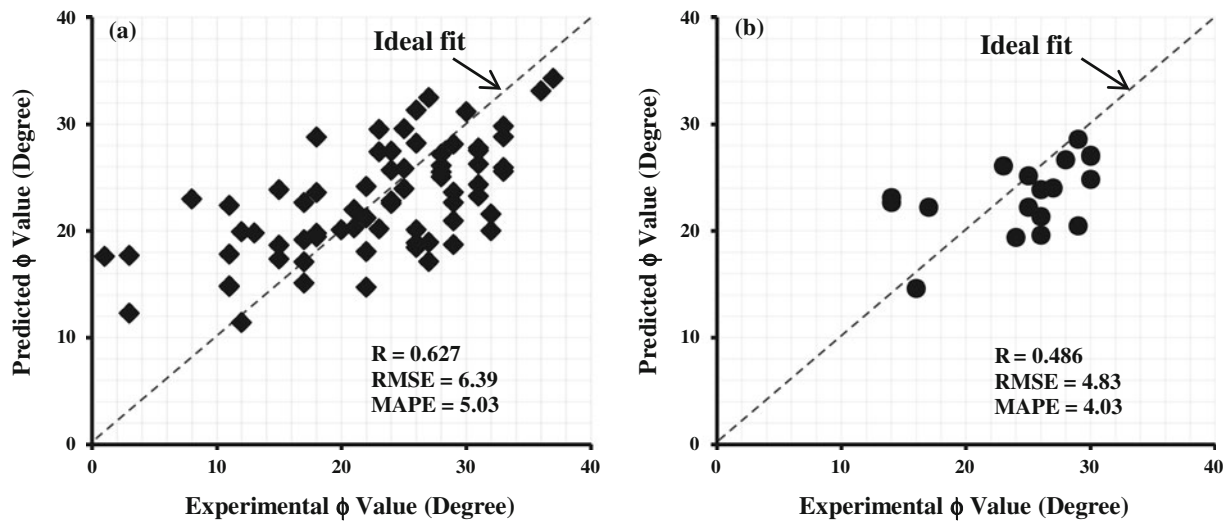


Figure 11. Experimental *versus* predicted  $\phi$  values using the MLSR model: (a) training data, and (b) testing data.



where  $a$  denotes coefficient vector. LSR minimizes the sum-of-squared residuals for each equation, accounting for any cross-equation restrictions on the parameters of the system. If there are no such restrictions, this technique is identical to estimating each equation using single-equation ordinary least squares. The MLSR models were trained using the data in training and validation datasets, which were previously considered for developing the LGP models. The same testing data as LGP was also used for the testing of the regression models. Eviews software package (Maravall and Gomez 2004) was used to perform the regression analysis.

### 3.5.1 MLSR models of soil strength parameters

The MLSR-based formulations of the soil strength parameters ( $c$  and  $\phi$ ) are as given below:

$$\begin{aligned} c_{\text{MLSR}}(\text{kg/cm}^2) = & 0.0023\text{FC} + 0.1394\text{D}_{30} \\ & + 0.0002\text{C}_u + 0.0241\text{LL} \\ & - 0.0233\omega + 1.0761\tilde{a}_d - 1.8553, \end{aligned} \quad (7)$$

$$\begin{aligned} \phi_{\text{MLSR}}(^{\circ}) = & 0.1541\text{FC} + 7.3270\text{D}_{30} \\ & + 0.0008\text{C}_u + 0.2508\text{LL} \\ & - 0.7534\omega + 7.8936\gamma_d + 2.7273. \end{aligned} \quad (8)$$

Figures 10 and 11 show a comparison of the experimental  $c$  and  $\phi$  values and the values predicted by MLSR, respectively. The resulting Fisher values of the performed regression analyses for  $c$  and  $\phi$  are equal to 10.07 and 6.73, respectively.

## 4. Performance analysis

Comparisons of the ratio of the predictions made by the LGP and MLSR models over the experimental  $c$  and  $\phi$  values are shown in figure 12. No rational models for the prediction of  $c$  and  $\phi$  was found that encompass the influencing variables considered in this study. Thus, it was not possible to conduct a comparative study between the results of this research and those in hand. Based on a rational hypothesis, Smith (1986) and Kasabov (1998) suggested the following criterion for judging the performance of a model:

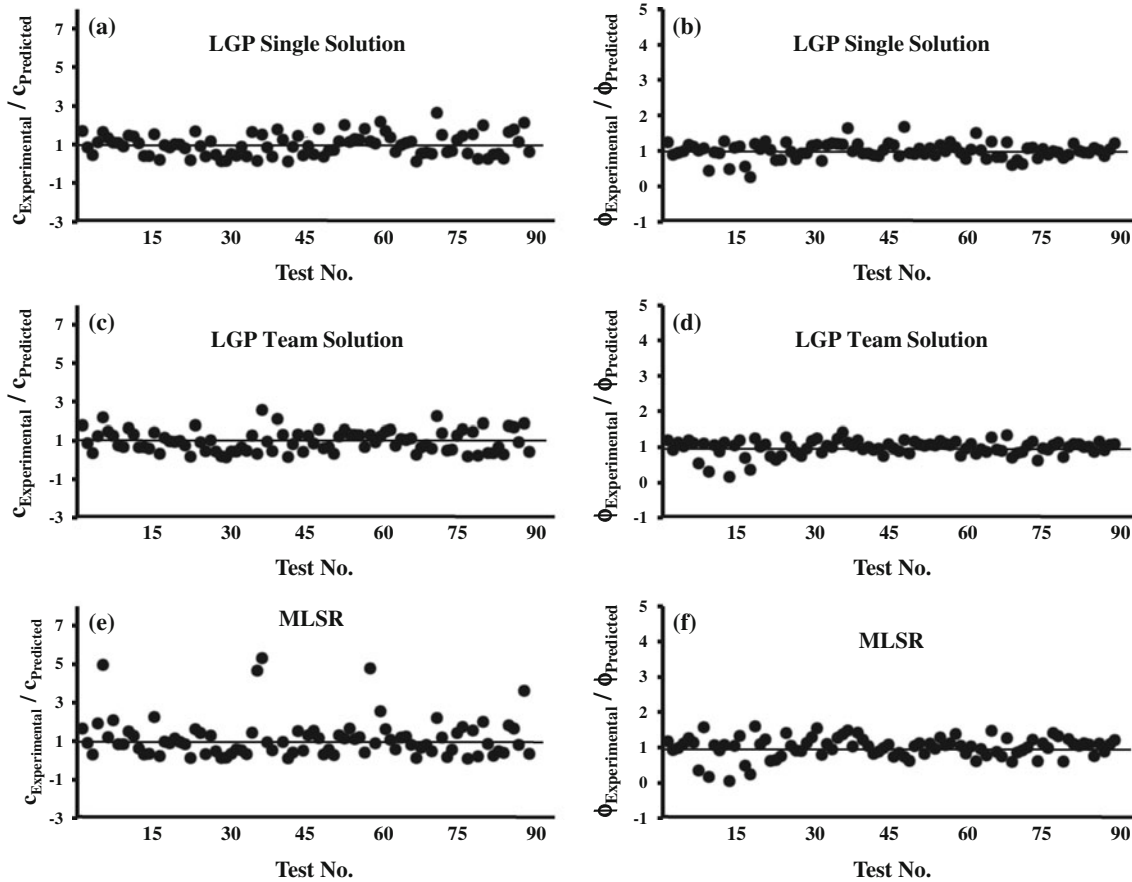


Figure 12. Comparison of the  $c$  and  $\phi$  predictions made by different models.

- if a model gives  $|R| > 0.8$ , a strong correlation exists between the predicted and measured values.

In all cases, the error values (e.g., RMSE and MAE) should be at the minimum. It can be observed from figures 6–9 that the LGP models with high  $R$  and low RMSE and MAE values are able to predict the target values to an acceptable degree of accuracy. The results indicate that the best team solutions found by LGP have produced the best results for the training, validation and testing data. The number of datasets considered for the training process of ANNs, GP, or other soft computing tools is an important issue, as it heavily bears on the reliability of the final models. To cope with this issue, Frank and Todeschini (1994) suggested that the minimum ratio of the number of objects over the number of selected variables for model acceptability is 3. They also suggest that considering a ratio equal to 5 is more reasonable. In the present study, this ratio is higher and is at least equal to  $88/9 = 9.8$ . The above facts ensure that the derived prediction models are valid, have the prediction power and are not established by chance.

In all cases, the best single and team solutions evolved by LGP have better performance than the MLSR models. Empirical modelling based on statistical regression techniques has significant limitations. Most commonly used regression analysis can have significant uncertainties. In most cases, the best models developed using the commonly used statistical approaches are obtained after controlling only few equations established in advance. Thus, such models cannot efficiently consider the interactions between the dependent and independent variables. On the other hand, LGP introduces completely new characteristics and traits. A major distinction of LGP for determining the  $c$  and  $\phi$  values lies in its powerful ability to model the mechanical behaviour without requesting prior form of the

existing relationships. The best solutions evolved by this technique are determined after controlling numerous preliminary models, even millions of linear and nonlinear models (Gandomi *et al* 2011b). For instance, the proposed single solutions for the estimation of  $c$  and  $\phi$  were, respectively, selected among a total of 1101259811 and 923596710 programs evolved and evaluated by LGP during the conducted runs. Furthermore, transparent human-interpretable solutions in Java, C++ or assembler codes can be developed through the LGP analysis. The evolved solutions are a knowledge-based representation of the underlying data information, which can be used for further analysis of the shear strength parameters. A further feature of LGP is the high level of interactivity between the user and the methodology. The user physical insight can be used to make propositions on the elements in the evolved functions and on their structures. The LGP method is especially practical for cases where the behaviour is too complex and conventional models are unable to effectively describe various aspects of the behaviour.

However, one of the main aims of introducing the GP-based approaches into the design processes is better handling of the information in the pre-design phase (Kraslawski *et al* 1999; Alavi *et al* 2011). It is idealistic to have some initial estimates of the outcome before performing any extensive laboratory or field work. The LGP approach employed in this research is based on the data alone to determine the structure and parameters of the model. Thus, the derived models are considered to be mostly valid for use in preliminary design stages and should cautiously be used for final decision-making. The proposed LGP models are suggested to be used to check the general validity of the laboratory test results. Further, these solutions are good alternatives to determine the soil shear strength parameters when testing is not possible.

Table 6. *Geotechnical properties of soils and test results used for model verification.*

| FC<br>(%) | CC<br>(%) | D <sub>30</sub><br>(mm) | C <sub>u</sub> | LL<br>(%) | $\omega$<br>(%) | $\gamma_d$<br>(gr/cm <sup>3</sup> ) | $c$<br>(kg/cm <sup>2</sup> ) | $\phi$ (°) | $c_{LGP,Single}$ | $c_{LGP,Team}$ | $\phi_{LGP,Single}$ | $\phi_{GP,Team}$ |
|-----------|-----------|-------------------------|----------------|-----------|-----------------|-------------------------------------|------------------------------|------------|------------------|----------------|---------------------|------------------|
| 63.1      | 36.9      | 0.012                   | 20.22          | 20        | 20              | 1.5                                 | 0.06                         | 22         | 0.21             | 0.14           | 23.8                | 19.9             |
| 66.8      | 33.2      | 0.012                   | 62.00          | 20        | 6.15            | 1.71                                | 0.55                         | 28         | 0.40             | 0.43           | 26.6                | 26.5             |
| 53.7      | 46.3      | 0.022                   | 150.03         | 20        | 15.3            | 1.78                                | 0.1                          | 11         | 0.18             | 0.20           | 19.4                | 13.5             |
| 82        | 18        | 0.006                   | 21.01          | 46        | 16              | 1.55                                | 0.44                         | 27         | 0.63             | 0.50           | 27.6                | 28.4             |
| 81.3      | 18.7      | 0.0078                  | 22.11          | 43        | 18.05           | 1.37                                | 0.02                         | 22         | 0.22             | 0.16           | 25.4                | 24.8             |
| 80.1      | 19.9      | 0.004                   | 36.67          | 42        | 16.85           | 1.73                                | 0.65                         | 30         | 0.65             | 0.72           | 26.5                | 27.1             |
| 64.5      | 35.5      | 0.015                   | 60.17          | 42        | 21.6            | 1.44                                | 0.42                         | 25         | 0.23             | 0.38           | 23.2                | 24.3             |
|           |           |                         |                |           |                 |                                     |                              | R          | 0.776            | 0.946          | 0.900               | 0.939            |
|           |           |                         |                |           |                 |                                     |                              | RMSE       | 0.155            | 0.092          | 3.828               | 2.115            |
|           |           |                         |                |           |                 |                                     |                              | MAE        | 0.139            | 0.085          | 2.968               | 1.970            |

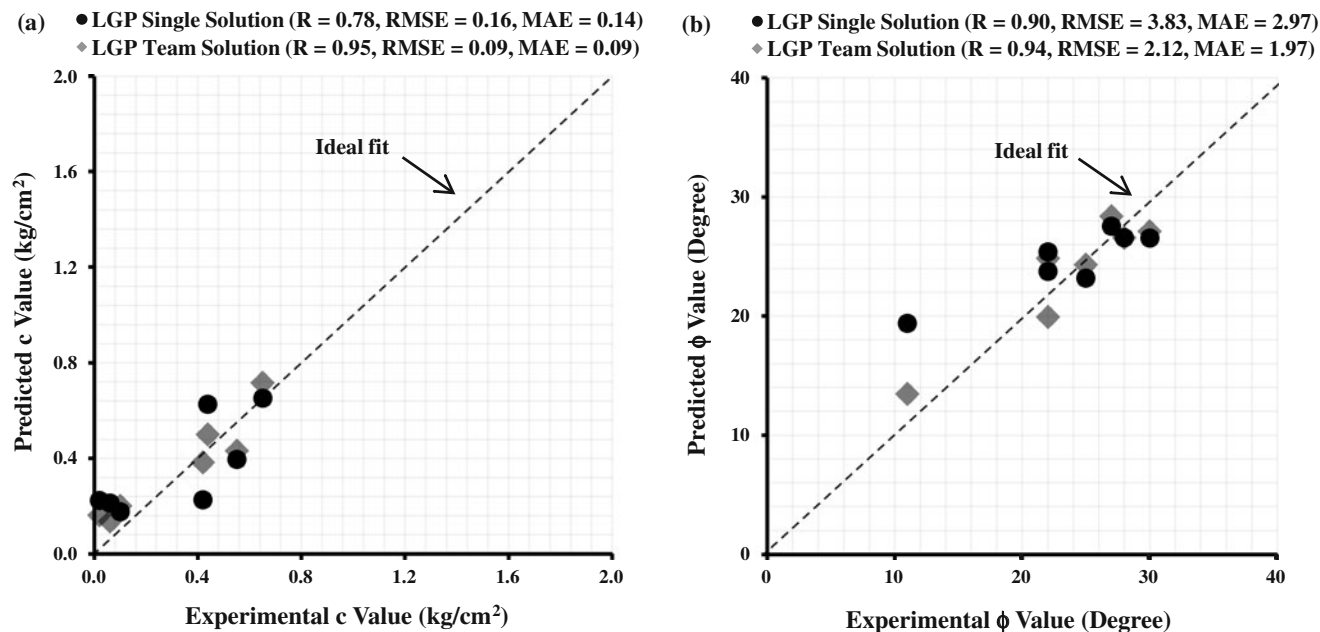


Figure 13. Experimental *versus* predicted shear strength parameters using the LGP models for the verification data: (a)  $c$ , and (b)  $\phi$ .

## 5. Validity verification

The generalization of the best LGP models was further verified by validating them by prediction of the shear strength parameters determined using new series of laboratory data. Table 6 shows the data used for the model verification. Of the seven soil samples shown in this table, the first three samples were taken from Sabzevar, Khorasan province and the next four datasets were obtained from Mahshahr, Khouzestan province in Iran. The percentages of the liquid limit for these data are lower or higher than the liquid limit range in the database used for the training of the models (see table 2). The prediction results obtained by the best LGP single and team solutions are shown in figure 13. It can be seen from this figure that the predicted  $c$  and  $\phi$  values are in good agreement with the corresponding experimental values. This confirms that the models can provide fairly acceptable results for the data beyond the ranges of the calibration database. The LGP team model outperforms the single model, in particular for the prediction of  $c$ .

## 6. Sensitivity and parametric analyses

Sensitivity analysis is of utmost concern for selecting the important input variables. The contribution of each input parameter in the LGP models was evaluated through a sensitivity analysis. For this purpose, frequency values (Francone 2001) of the input parameters were obtained. A frequency value

equal to 1.00 for an input indicates that this input variable has been appeared in 100% of the best thirty programs evolved by LGP. This is a common approach in the GP-based analyses (Francone 2001; Gandomi *et al* 2010).

The frequency values of the input parameters of the LGP models are presented in figure 14. According to these results, it can be found that  $c$  and  $\phi$  are more dependent on  $LL$  and  $\omega$  compared with the other soil properties. Another important observation from the results of the sensitivity study is that  $D_{30}$  and  $C_u$  are much less important in explaining the variations of the  $c$  and  $\phi$  values than the other variables.

For further verification of the LGP models, a parametric analysis was performed in this study. The main goal was to find the effect of each parameter on the soil strength parameters. The methodology was based on the variation of only one input variable at a time while the other variables were kept constant at the average values of their entire datasets. A set of synthetic data for the single varied parameter was generated by increasing the value of this in increments. These inputs were presented to the models and the  $c$  and  $\phi$  values were calculated. This procedure was repeated using another variable until the responses of the models were tested for all input variables. Discipulus interactive evaluator mode was utilized to run the resulting C++ codes of the best LGP single and team models. Figures 15 and 16 present the tendency of the  $c$  and  $\phi$  predictions to the variations of the soil physical properties ( $FC$ ,  $D_{30}$ ,  $C_u$ ,  $LL$ ,  $\omega$ , and  $\gamma_d$ ).

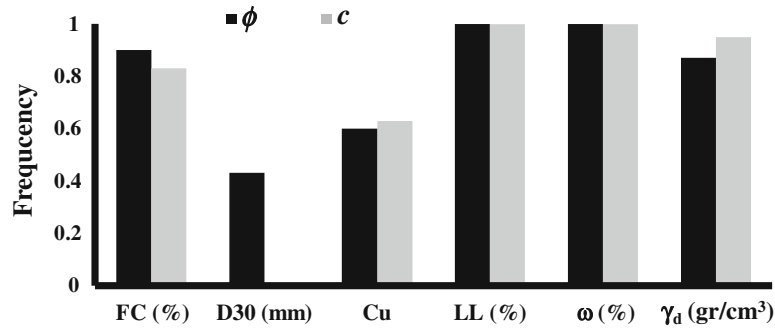


Figure 14. Contributions of the input parameters in the LGP models.

As can be seen in figure 15(a–d and f),  $c$  increases due to increasing FC,  $D_{30}$ ,  $C_u$  and  $\gamma_d$ . The results of the parametric analysis indicate that  $c$  continuously increases when  $\omega$  increases up to 8%

and thereafter it starts decreasing. The results for  $\phi$  are more complex than those for  $c$ . As shown in figure 16(c and e),  $\phi$  continuously decreases due to increasing  $C_u$  and  $\omega$ . The results for the LGP single

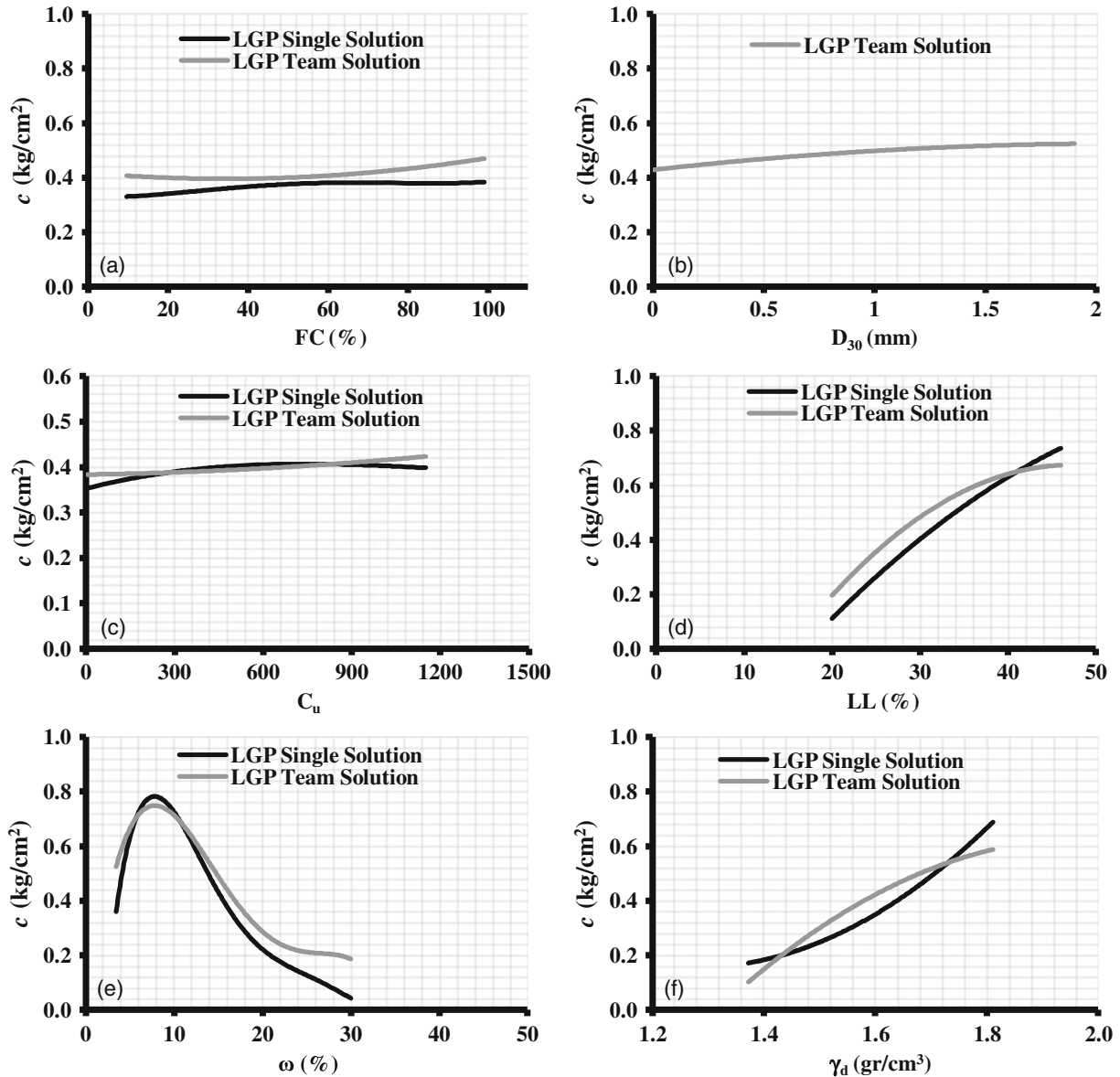


Figure 15. Parametric analysis of the cohesion intercept in the LGP models.

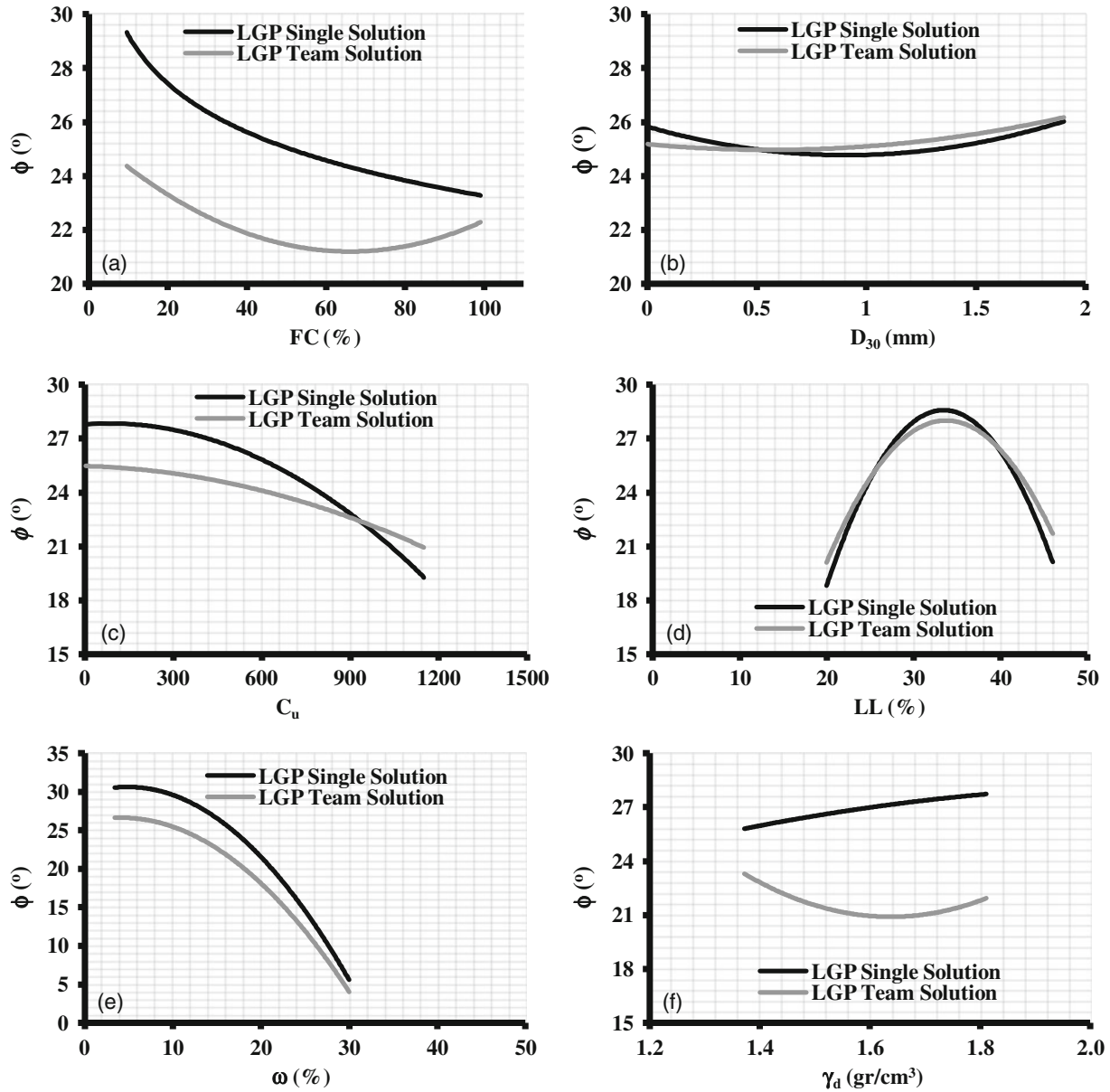


Figure 16. Parametric analysis of the angle of shearing resistance in the LGP models.

solution indicate that  $\phi$  decreases due to increasing FC. The relevant results for the team model depict that  $\phi$  initially decreases when FC increases up to about 70% and thereafter it starts increasing.  $\phi$  decreases with increasing  $D_{30}$  up to 1 mm in the LGP single model and for larger values of  $D_{30}$ , it starts increasing.  $\phi$  and  $D_{30}$  are positively correlated with each other in the LGP team model. As presented in figure 16(d),  $\phi$  is positively correlated with LL up to a peak about 35% and thereafter the correlation of these parameters becomes negative. The parametric analysis results for the LGP single solution indicate that  $\phi$  continuously increases with increasing  $\gamma_d$ . In the team model,  $\phi$  decreases when  $\gamma_d$  increases up to about 1.65 gr/cm<sup>3</sup> and thereafter it starts increasing. The parametric analysis results for the soil strength parameters, in par-

ticular for  $c$ , are generally expected cases from a geotechnical engineering viewpoint (Murthy 2008).

## 7. Conclusions

In this research, the LGP paradigm was used to derive high-precision models for assessing the soil shear strength parameters ( $c$  and  $\phi$ ). The proposed models were developed based on well established and widely dispersed triaxial test results obtained through an experimental study. The best models were selected after developing and controlling several models with different combinations of the input parameters. The following principal conclusions may be drawn based on the results presented:



- The developed LGP single and team models figured out reliable estimations of the  $c$  and  $\phi$  values. The best team solutions evolved by LGP produced better results than the best LGP single solutions. The proposed models are obviously superior in comparison with the developed regression models.
- Contrary to the conventional models, the LGP models simultaneously took into account the effect of several important factors ( $FC$ ,  $D_{30}$ ,  $C_u$ ,  $LL$ ,  $\omega$ , and  $\gamma_d$ ) representing the behaviour of the shear strength parameters.  $\omega$  and  $\gamma_d$  were found to be efficient representatives of the initial state and stress history of the soil.
- In addition to the acceptable accuracy, the results of the LGP analysis provided transparent programs of an imperative language for further use in the  $c$  and  $\phi$  prediction, as well as optimization purposes. The evolved solutions, in this case C++ codes, represent a mapping of the input and output data. These programs can easily be inspected and evaluated. The interested readers may consult the corresponding author for free C++ codes of the best LGP team solutions.
- The validity of the LGP models was further tested for a portion of laboratory results beyond the training data domain. The validation phase confirmed the efficiency of the models for their application to estimation of the shear strength parameters.
- The proposed models are mostly suitable for fine-grained soils with physical properties similar to the soil samples used in this study.
- An important finding from the sensitivity analysis results is that the soil unit weight and liquid limit are the most important parameters governing the behaviour of the shear strength parameters.
- The shear strength parameters can easily be estimated from the soil physical properties by the use of derived models. Thus, there is no need to go through sophisticated and time-consuming laboratory tests.

### Appendix A. The optimum LGP programs

The optimum LGP programs can be run in the Discipulus interactive evaluator mode or be compiled in C++ environment. (Note: The input variables of the following programs are in the normalized forms.)

Table A1. *The optimum LGP program for the prediction of  $c$ .*

|   |   |   |
|---|---|---|
| float DiscipulusCFunction(float FC,<br>C <sub>u</sub> , LL, $\omega$ , $\gamma_d$ )<br>{<br>long double f[8];<br>long double tmp = 0;<br>int cflag = 0;<br>f[0]=f[1]=f[2]=f[3]=f[4]=f[5]=f[6]=<br>f[7]=0;<br><br>L0:f[0] -= $\omega$ ;<br>L1:tmp=f[3]; f[3]=f[0]; f[0]=tmp;<br>L2:f[0] += 0.1387641429901123f;<br>L3:f[2] += f[0];<br>L4:f[0]=sin(f[0]);<br>L5:f[1] += f[0];<br>L6:f[0]/=0.9955191612243652f;<br>L7:tmp=f[3]; f[3]=f[0]; f[0]=tmp;<br>L8:f[0]/=f[0];<br>L9:f[0]=-f[0];<br>L10:f[0]/=FC;<br>L11:f[0]=-f[0];<br>L12:f[0]/=f[1];<br>L13:f[0] += 0.6342074871063232f;<br>L14:f[0]=sin(f[0]);<br>L15:f[0]=fabs(f[0]);<br>L16:f[0] -= f[3];<br>L17:f[0]*= $\gamma_d$ ;<br>L18:f[2]*=f[0]; | L19:f[0]/=f[1];<br>L20:f[0]=cos(f[0]);<br>L21:f[0]*<br>= - 0.1401152610778809f;<br>L22:cflag=(f[0] < f[1]);<br>L23:f[0] += f[2];<br>L24:f[0]*=f[0];<br>L25:f[3] += f[0];<br>L26:f[0] -= LL;<br>L27:f[0] -= $\gamma_d$ ;<br>L28:f[0] += f[2];<br>L29:f[0] -= f[0];<br>L30:f[0]*= $\gamma_d$ ;<br>L31:f[2]*=f[0];<br>L32:f[0] -= f[0];<br>L33:if (!cflag) f[0] = f[3];<br>L34:f[0]*=f[1];<br>L35:f[0] -= f[0];<br>L36:f[0]=sin(f[0]);<br>L37:f[0]=sin(f[0]);<br>L38:f[0] += f[2];<br>L39:f[0]=sin(f[0]);<br>L40:f[0]/= $\omega$ ;<br>L41:f[0]=sin(f[0]);<br>L42:f[0]=sin(f[0]);<br>L43:cflag=(f[0] < f[1]);<br>L44:f[0]*=f[1];<br>L45:if (cflag) f[0] = f[1]; | L46:f[0]=sin(f[0]);<br>L47:f[0] += C <sub>u</sub> ;<br>L48:f[0] -= LL;<br>L49:f[0]/= $\omega$ ;<br>L50:f[0]=fabs(f[0]);<br>L51:cflag=(f[0] < f[1]);<br>L52:f[0]*= $\gamma_d$ ;<br>L53:f[2]*=f[0];<br>L54:if (!cflag) f[0] = f[2];<br>L55:f[3] -= f[0];<br>L56:f[0] -= LL;<br>L57:f[0]/= $\omega$ ;<br>L58:f[0] += LL;<br>L59:f[0] -= f[3];<br>L60:f[0]=sin(f[0]);<br>L61:f[0]=fabs(f[0]);<br>L62:f[0] -= f[2];<br>L63:f[0]*= $\gamma_d$ ;<br>L64:cflag=(f[0] < f[3]);<br>L65:f[0] -= f[2];<br>L66:f[0] -= f[2];<br>L67:if (cflag) f[0] = f[3];<br>L68:<br>if (!finite(f[0])) f[0]=0;<br>return f[0];<br>} |
|---|---|---|

Table A2. The optimum LGP program for the prediction of  $\phi$ .

|  |                                    |                                     |
|--|------------------------------------|-------------------------------------|
| float DiscipulusCFFunction(float FC,<br>D30, Cu, LL, $\omega$ , $\gamma_d$ ) | L50:f[0]*=f[0];                    | L106:f[0]=sqrt(f[0]);               |
| {long double f[8];   | L51:tmp=f[1]; f[1]=f[0]; f[0]=tmp; | L107:cflag=(f[0] < f[0]);           |
| long double tmp = 0;   | L52:f[0]=D30;                      | L108:if (!cflag) f[0] = f[2];       |
| int cflag = 0;   | L53:f[1]*=f[0];                    | L109:f[1]/=f[0];                    |
| f[0]=f[1]=f[2]=f[3]=f[4]=f[5]=f[6]=f[7]=0;                                   | L54:f[0]/= $\gamma_d$ ;            | L110:f[0]=sqrt(f[0]);               |
| L0:f[0]+= $\gamma_d$ ;   | L55:tmp=f[1]; f[1]=f[0]; f[0]=tmp; | L111:f[0]/=-4.178634643554688f;     |
| L1:tmp=f[1]; f[1]=f[0]; f[0]=tmp;  | L56:f[0]=fabs(f[0]);               | L112:f[0]=-f[1];                    |
| L2:f[0]=cos(f[0]);   | L57:f[1]*=f[0];                    | L113:tmp=f[1]; f[1]=f[0]; f[0]=tmp; |
| L3:tmp=f[1]; f[1]=f[0]; f[0]=tmp;  | L58:f[1]-=f[0];                    | L114:f[0]=-f[0];                    |
| L4:f[0]+=f[1];   | L59:if (!cflag) f[0] = f[1];       | L115:f[0]+=-3.035147190093994f;     |
| L5:f[0]*=8.718416213989258f;   | L60:f[0]*=f[1];                    | L116:f[0]*=f[0];                    |
| L6:f[0]=sin(f[0]);   | L61:f[0]-=f[1];                    | L117:f[0]=sin(f[0]);                |
| L7:f[0]=cos(f[0]);   | L62:f[0]=sqrt(f[0]);               | L118:f[0]=-f[0];                    |
| L8:f[0]=cos(f[0]);   | L63:f[0]=cos(f[0]);                | L119:if (cflag) f[0] = f[2];        |
| L9:f[0]/=0.9754431247711182f;  | L64:cflag=(f[0] < f[2]);           | L120:f[0]/=-3.076897859573364f;     |
| L10:f[0]=sin(f[0]);  | L65:f[0]*=f[2];                    | L121:f[0]+= $\gamma_d$ ;            |
| L11:f[0]+=f[1];  | L66:if (cflag) f[0] = f[1];        | L122:f[0]=FC;                       |
| L12:f[0]/=-0.9327507019042969f;  | L67:tmp=f[1]; f[1]=f[0]; f[0]=tmp; | L123:f[0]=-4.970226287841797f;      |
| L13:f[0]-= $\omega$ ;  | L68:f[0]=cos(f[0]);                | L124:cflag=(f[0] < f[1]);           |
| L14:f[0]/= $\omega$ ;  | L69:f[2]*=f[0];                    | L125:tmp=f[1]; f[1]=f[0]; f[0]=tmp; |
| L15:f[0]=sin(f[0]);  | L70:f[0]=FC;                       | L126:f[0]/=-2.971264123916626f;     |
| L16:f[0]=sin(f[0]);  | L71:f[0]/=f[2];                    | L127:f[0]*=8.718416213989258f;      |
| L17:f[0]-=D30;   | L72:cflag=(f[0] < f[2]);           | L128:f[0]=f[2];                     |
| L18:f[0]=sin(f[0]);  | L73:f[0]=cos(f[0]);                | L129:if (cflag) f[0] = f[2];        |
| L19:f[1]/=f[0];  | L74:f[1]=f[0];                     | L130:f[0]*=f[1];                    |
| L20:f[2]=f[0];   | L75:f[1]/=f[0];                    | L131:f[1]+=f[0];                    |
| L21:f[0]=fabs(f[0]);   | L76:f[0]=cos(f[0]);                | L132:f[0]-=f[0];                    |
| L22:f[1]*=f[0];  | L77:if (!cflag) f[0] = f[2];       | L133:f[0]-= $\omega$ ;              |
| L23:f[1]-=f[0];  | L78:f[0]*=8.718416213989258f;      | L134:f[0]/= $\omega$ ;              |
| L24:f[0]/=-0.9327507019042969f;  | L79:f[0]*= $\omega$ ;              | L135:f[0]+=6.721087455749512f;      |
| L25:f[0]-=LL;  | L80:f[0]-= $\omega$ ;              | L136:f[0]+= $\gamma_d$ ;            |
| L26:f[0]=cos(f[0]);  | L81:f[0]*= $\omega$ ;              | L137:f[0]=sin(f[0]);                |
| L27:f[0]=sin(f[0]);  | L82:f[0]-= $\omega$ ;              | L138:f[0]=sqrt(f[0]);               |
| L28:f[0]*=f[2];  | L83:f[0]=fabs(f[0]);               | L139:f[0]/= $\omega$ ;              |
| L29:f[0]+= $\gamma_d$ ;  | L84:f[0]=sin(f[0]);                | L140:f[0]*=f[0];                    |
| L30:f[0]*=f[0];  | L85:f[0]=-f[0];                    | L141:f[0]+= $\gamma_d$ ;            |
| L31:f[1]*=f[0];  | L86:f[0]=D30;                      | L142:f[0]=-2.447271347045898f;      |
| L32:f[0]=sqrt(f[0]);   | L87:f[0]=cos(f[0]);                | L143:f[0]=-4.970226287841797f;      |
| L33:tmp=f[1]; f[1]=f[0]; f[0]=tmp;   | L88:f[2]*=f[0];                    | L144:cflag=(f[0] < f[1]);           |
| L34:f[0]=cos(f[0]);  | L89:cflag=(f[0] < f[2]);           | L145:tmp=f[1]; f[1]=f[0]; f[0]=tmp; |
| L35:f[2]*=f[0];  | L90:f[1]-=f[0];                    | L146:f[0]=fabs(f[0]);               |
| L36:f[0]=f[1];   | L91:f[0]/=f[2];                    | L147:f[1]-=f[0];                    |
| L37:f[0]+= $\omega$ ;  | L92:f[0]=sin(f[0]);                | L148:f[0]=-f[0];                    |
| L38:f[0]*=8.718416213989258f;  | L93:f[0]=cos(f[0]);                | L149:f[0]=sin(f[0]);                |
| L39:f[0]=sin(f[0]);  | L94:f[0]*=f[0];                    | L150:f[0]+= $\gamma_d$ ;            |
| L40:f[0]*=f[1];  | L95:f[0]=Cu;                       | L151:f[0]=sin(f[0]);                |
| L41:f[0]=cos(f[0]);  | L96:f[0]*=-4.464609622955322f;     | L152:f[0]/=FC;                      |
| L42:f[0]*=-5.272376537322998f;   | L97:f[0]=f[0];                     | L153:f[0]=f[2];                     |
| L43:f[0]-=-4.178634643554688f;   | L98:f[0]=f[2];                     | L154:tmp=f[1]; f[1]=f[0]; f[0]=tmp; |
| L44:f[0]*=pow(2,TRUNC(f[1]));  | L99:f[0]=sin(f[0]);                | L155:f[0]+=6.721087455749512f;      |
| L45:f[0]=sin(f[0]);  | L100:f[1]-=f[0];                   | L156:f[0]-=-0.9327507019042969f;    |
| L46:f[0]*=Cu;  | L101:f[0]*=f[1];                   | L157:f[0]+=-2.367838621139526f;     |
| L47:f[0]=f[0];   | L102:f[0]/=-3.035147190093994f;    | L158:if (cflag) f[0] = f[2];        |
| L48:f[0]=f[0];   | L103:f[0]-=9.826432228088379f;     | L159:f[0]=fabs(f[0]);               |
| L49:tmp=f[1]; f[1]=f[0]; f[0]=tmp;   | L104:f[0]/=LL;                     | L160:f[0]=-f[0];                    |
|  | L105:f[0]=cos(f[0]);               | L161:tmp=f[2]; f[2]=f[0]; f[0]=tmp; |

Table A2. (Continued.)

|                                     |                              |                                |
|-------------------------------------|------------------------------|--------------------------------|
| L162:f[0]/= $\omega$ ;              | L177:f[0]/= $\omega$ ;       | L192:f[0]=-f[0];               |
| L163:f[0]+=f[0];                    | L178:if (cflag) f[0] = f[2]; | L193:f[0]=cos(f[0]);           |
| L164:f[0]+=f[0];                    | L179:f[0]=cos(f[0]);         | L194:f[2]*=f[0];               |
| L165:f[0]=sin(f[0]);                | L180:f[0]/= $\omega$ ;       | L195:f[0]=-7.077741622924805f; |
| L166:f[0]*=f[0];                    | L181:cflag=(f[0] < f[1]);    | L196:f[0]*=8.718416213989258f; |
| L167:f[0]=sqrt(f[0]);               | L182:f[0]*=f[0];             | L197:f[0]*=f[0];               |
| L168:f[0]*=D30;                     | L183:f[1]+=f[0];             | L198:cflag=(f[0] < f[1]);      |
| L169:f[2]-=f[0];                    | L184:f[0]-=f[0];             | L199:if (!cflag) f[0] = f[2];  |
| L170:f[0]+=f[0];                    | L185:if (cflag) f[0] = f[2]; | L200:f[0]=fabs(f[0]);          |
| L171:f[0]/=f[1];                    | L186:f[0]-=D30;              | L201:f[0]=sqrt(f[0]);          |
| L172:tmp=f[1]; f[1]=f[0]; f[0]=tmp; | L187:f[0]=fabs(f[0]);        | L202:f[0]=sin(f[0]);           |
| L173:f[0]=cos(f[0]);                | L188:f[1]-=f[0];             | L203:                          |
| L174:cflag=(f[0] < f[1]);           | L189:f[0]/=f[1];             | if (!finite(f[0])) f[0]=0;     |
| L175:f[0]*=f[1];                    | L190:f[0]/=f[2];             | return f[0]; }                 |
| L176:f[0]-=2.447271347045898f;      | L191:f[0]*=f[0];             |                                |

## References

- Alavi A H and Gandomi A H 2011a Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing; *Comput. Struct.*, doi: 10.1016/j.compstruc. 2011.08.019.
- Alavi A H and Gandomi A H 2011b A robust data mining approach for formulation of geotechnical engineering systems; *Int. J. Comput. Aided Meth. Eng.-Eng. Computations* **28(3)** 242–274.
- Alavi A H, Gandomi A H, Gandomi M and Sadat Hosseini S S 2009 Prediction of maximum dry density and optimum moisture content of stabilized soil using RBF neural networks; *The IES J Part A: Civil. Struct. Eng.* **2** 98–106.
- Alavi A H, Gandomi A H, Sahab M G and Gandomi M 2010 Multi expression programming: A new approach to formulation of soil classification; *Eng. Comput.* **26(2)** 111–118.
- Alavi A H, Ameri M, Gandomi A H and Mirzahosseini M R 2011 Formulation of flow number of asphalt mixes using a hybrid computational method; *Constr. Build. Mater.* **25(3)** 1338–1355.
- Arora K R 1988 Introductory soil engineering; text book. Nem Chand Jane (Prop), Standard Publishers Distributors, Nai Sarak, Delhi.
- ASTM D 6528 Consolidated undrained direct simple shear testing of cohesive soils.
- ASTM D2850-87 Standard Test Method for Unconsolidated, Undrained Compressive Strength of Cohesive Soils in Triaxial Compression.
- ASTM WK3821 New test method for consolidated drained triaxial compression test for soils.
- Aytek A and Alp M 2008 An application of artificial intelligence for rainfall-runoff modeling; *J. Earth Syst. Sci.* **117(2)** 145–155.
- Banzhaf W, Nordin P, Keller R and Francone F 1998 Genetic programming – An introduction on the automatic evolution of computer programs and its application; dpunkt/Morgan Kaufmann, Heidelberg/San Francisco.
- Barends F B J, Lindenberg J L, De Quelerij L, Verruijt A and Luger H J 1999 *Geotechnical Engineering for Transportation Infrastructure: Theory and Practice, Planning and Design, Construction and Maintenance*; Balkema Publishers, Netherlands.
- Baykasoglu A, Gullub H, Canakci H and Ozbakir L 2008 Prediction of compressive and tensile strength of limestone via genetic programming; *Expert. Syst. Appl.* **35(1–2)** 111–123.
- Bowles J E 1992 Engineering properties of soils and their measurement; McGraw-Hill, New York.
- Brameier M and Banzhaf W 2001 A comparison of linear genetic programming and neural networks in medical data mining; *IEEE Trans. Evol. Com.* **5(1)** 17–26.
- Brameier M and Banzhaf W 2007 *Linear Genetic Programming*; Springer Science + Business Media, New York.
- Cevik A and Cabalar A F 2009 Modelling damping ratio and shear modulus of sand–mica mixtures using genetic programming; *Expert. Syst. Appl.* **36(4)** 7749–7757.
- Conrads M, Dolezal O, Francone F D and Nordin P 2004 Discipulus – fast genetic programming based on AIM learning technology; Register Machine Learning Technologies Inc, Littleton, CO.
- Deschaine L M 2000 Using Genetic Programming to Develop a C/C++ Simulation Model of a Waste Incinerator Science Applications; International Corp, Draft Technical Report.
- El-Maksoud M A F 2006 Laboratory determining of soil strength parameters in calcareous soils and their effect on chiseling draft prediction; In: *Proceedings of Energy Efficiency and Agricultural Engineering International Conference*, Rousse, Bulgaria.
- Feldt R and Nordin P 2000 Using Factorial Experiments to Evaluate the Effect of Genetic Programming Parameters; In: *Proceedings of EuroGP 2000 LNCS 1802* pp. 271–282.
- Francone F D 2001 Discipulus Pro<sup>TM</sup> Software Owner's Manual; Register Machine Learning Technologies Inc: Littleton, CO.
- Francone F D 2004 Discipulus Lite<sup>TM</sup> Software Owner's Manual, Littleton, CO, USA, Machine Learning Technologies Inc.
- Frank I E and Todeschini R 1994 The data analysis handbook; Elsevier, Amsterdam, The Netherlands.
- Gandomi A H and Alavi A H 2011 Multi-stage genetic programming: A new strategy to nonlinear system modeling; *Inf. Sci.* **181(23)** 5227–5239.
- Gandomi A H, Alavi A H and Sadathosseini S S 2008 A discussion on genetic programming for retrieving missing information in wave records along the west coast of India; *Appl. Ocean Res.* **30** 338–339.

- Gandomi A H, Alavi A H and Sahab M G 2010 New formulation for compressive strength of CFRP confined concrete cylinders using linear genetic programming; *Mater. Struct.* **43**(7) 963–983.
- Gandomi A H, Alavi A H, Mirzahosseini R and Moqaddas Nezhad F 2011a Nonlinear genetic-based models for prediction of flow number of asphalt mixtures; *J. Mater. Civil Eng. ASCE* **23**(3) 1–18.
- Gandomi A H, Alavi A H and Yun G J 2011b Nonlinear modeling of shear strength of sfrcb beams using linear genetic programming; *Struct. Eng. Mech.* **38**(1) 1–25.
- Güven A 2009 Linear Genetic Programming for time-series modelling of daily flow rate; *J. Earth Syst. Sci.* **118**(2) 137–146.
- Güven A, Günel M and Cevik A 2006 Prediction of pressure fluctuations on stilling basins; *Can. J. Civ. Eng.* **33**(11) 1379–1388.
- Güven A, Aytekin A, Yüce M I and Aksoy H 2007 Genetic programming-based empirical model for daily reference evapotranspiration estimation; *CLEAN–Soil Air Water J.* **36**(10–11) 905–912.
- Güven A, Azamathullah H M and Zakaria N A 2009 Linear Genetic Programming for prediction of circular pile scour; *Ocean Eng.* **36**(12–13) 985–991.
- Heshmati A A R, Alavi A H, Keramati M and Gandomi A H 2009 A radial basis function Neural Network Approach for compressive strength prediction of stabilized soil; *Geotech. Spec. Publ. ASCE* **191** 147–153.
- Javadi A A and Rezaei M 2009 Applications of artificial intelligence and data mining techniques in soil modelling; *Geomech. Eng.* **1**(1) 53–74.
- Kasabov N K 1998 *Foundations of neural networks fuzzy systems and knowledge engineering*; Cambridge, MIT Press.
- Kayadelen C, Günaydin O, Fener M, Demir A and Özvan A 2009 A modeling of the angle of shearing resistance of soils using soft computing systems; *Expert. Syst. Appl.* **36** 11,814–11,826.
- Korayem A Y, Ismail K M and Sehari S Q 1996 Prediction of soil shear strength and penetration resistance using some soil properties; *Mis. J. Agr. Res.* **13**(4) 119–140.
- Koza J 1992 *Genetic programming, on the programming of computers by means of natural selection*; Cambridge (MA), MIT Press.
- Kraslawski A, Pedrycz W and Nyström L 1999 Fuzzy Neural Network as instance generator for case-based reasoning system: An example of selection of heat exchange equipment in mixing; *Neu. Comput. Appl.* **8**(2) 106–113.
- Maravall A and Gomez V 2004 Eviews Software; Version 5. Quantitative Micro Software, LLC, Irvine CA.
- Mesbahi E 2000 Application of artificial neural networks in modelling and control of diesel engines; PhD Thesis, University of Newcastle, UK.
- Mollahasani A, Alavi A H, Gandomi A H and Rashed A 2011 Nonlinear neural-based modeling of soil cohesion intercept; *KSCE J. Civil. Eng.* **15**(5) 831–840.
- Mousavi S M, Alavi A H, Gandomi A H and Mollahasani A 2011 Formulation of soil angle of shearing resistance using a hybrid GP and OLS Method; *Eng. Comput.* in press. doi: 10.1007/s00366-011-0242-x
- Murthy S 2008 *Geotechnical Engineering: Principles and Practices of Soil Mechanics*; 2nd edn, Taylor & Francis, CRC Press, UK.
- Oltean M and Grosan C 2003 A comparison of several linear genetic programming techniques; *Complex Syst.* **14**(4) 1–29.
- Panwar J S and Seimens J C 1972 Shear strength and energy of soil failure related to density and moisture; *T. ASAE* **15** 423–427.
- Poli R, Langdon W B, McPhee N F and Koza J R 2007 Genetic programming: An introductory tutorial and a survey of techniques and applications; Technical report [CES-475], University of Essex, UK.
- Ryan T P 1997 *Modern Regression Methods*; Wiley, New York.
- Shahin M A, Jaksa M B and Maier H R 2009 Recent advances and future challenges for artificial neural systems in geotechnical engineering applications; *Adv. Artif. Neur. Syst.*, **Article ID 308239**.
- Shahin M A, Maier H R and Jaksa M B 2001 Artificial neural network applications in geotechnical engineering; *Aus. Geomech.* **36**(1) 49–62.
- Smith G N 1986 *Probability and statistics in civil engineering*; Collins, London.
- Swingler K 1996 *Applying neural networks: A practical guide*; Academic Press, New York.
- Terzaghi K, Peck R B and Mesri G 1996 *Soil mechanics in engineering practice*, 3rd edn, Wiley & Sons, Inc, New York.