Engineering Applications of Artificial Intelligence 35 (2014) 143–157

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Function finding via genetic expression programming for strength and elastic properties of clay treated with bottom ash

Hamza Güllü *

Department of Civil Engineering, University of Gaziantep, 27310 Gaziantep, Turkey

ARTICLE INFO

Article history:
Received 7 November 2013
Received in revised form
30 May 2014
Accepted 18 June 2014
Available online 17 July 2014

Keywords:
Genetic expression programming
Clay
Bottom ash
Stabilization
Strength
Elasticity modulus

ABSTRACT

In order to understand the treatment of a marginal soil well, the underlying input–output relationship on the strength and elastic responses due to nonlinearity has always been a great importance in the stabilized mixtures for an optimal design. This paper employs a relatively new soft computing approach, genetic expression programming (GEP), to formulations for unconfined compressive strength (UCS) and elasticity modulus (Es) of clay stabilized with bottom ash, using a database obtained from the laboratory tests conducted in the study. The predictor variables included in the formulations are bottom ash dosage, dry unit weight, relative compaction, brittleness index and energy absorption capacity. The results demonstrate that the GEP-based formulas of UCS and Es are significantly able to predict the measured values to high degree of accuracy against the nonlinear behavior of soil ($p < 0.05$, $R > 0.847$). The GEP approach is found to have a better correlation performance as compared with the nonlinear regression as well. The sensitivity analysis for the parameter importance shows that the dominant influence on the predictions of UCS and Es is exerted by the variables of bottom ash dosage and energy absorption capacity. This study reveals that the GEP is a potential tool for establishing the functions and identifying the key variables for predicting the strength and elastic responses of the clay treated with bottom ash. Including a waste material in the proposed formulas can also serve to the environment for the development of recycling and sustainability.

1. Introduction

Clay is a marginal fine-grained soil that mostly requires to be amended due to its strength and elasticity problems to bearing capacity and settlement in the geotechnical applications such as foundation base of buildings, highway embankment, retaining wall backfill, road construction, dam fill etc. Stabilization is one of the improvement techniques popularly used to fix this issue by employing an appropriate stabilizer. Recently, some industrial waste materials have been intensively proposed for treatment of clay because of their good pozzolanic activity with the soil particles, economical benefits and environmental friendly considerations (Lin et al., 2007; Hossain and Mol, 2011; Güllü and Girişken, 2013). One of them is the bottom ash, an industrial by-product from the burning of coal produced in power generation plants, is relatively beneficial for improving the engineering properties of clay as a soil stabilizer due to its characteristics of particle size distribution as well as pozzolanic activity and chemical properties (Rifai et al., 2009; Kim and Do, 2012). It has a grain size distribution typically similar to a sandy material that enables to decrease plasticity and excessive settlement (or swelling) as well as increase strength and bearing capacity when added to clay (Huang and Lovell, 1990; Kayabali and Bulus, 2000; Kim, 2003; Kim and Do, 2012). In the stabilizations, the bottom ash is added to the natural soil at a proper proportion and mixed together with an optimum amount of water. The extent of the strength improvement in the treatment depends on compaction (packing density), mineralogy, amount of stabilizer (bottom ash) and curing period of mixture (Kim and Prezzi, 2008; Kim et al., 2011). As clearly can be understood from the nature of stabilization of clay, the strength and elasticity modulus are two essential parameters in the design considerations corresponding to bearing capacity and settlement, respectively. These are also two fundamental parameters that are commonly used in most of geotechnical designs (Bowles, 1996; Holtz et al., 2011). Due to their most priority role in the design, the strength and elastic responses relatively deserve to be investigated for the relationships with other engineering properties of geomaterial. More important, for optimal and effective utilization of bottom ash, a predictive model would be beneficial for understanding the gain in the strength and elastic considerations in terms of variables affecting the performance of treatment. Moreover, particularly at the conditions of limited laboratory facilities

* Tel.: +90 342 317 2433; fax: +90 342 360 1107.
E-mail address: hgullu@gantep.edu.tr

http://dx.doi.org/10.1016/j.engappai.2014.06.020
0952-1976/

and routine test requirements in the project applications, the formulas for strength and elasticity responses can provide with simpler, faster and more economical solution aspects to design. Due to the reasons arisen above, a predictive function that is able to formulate the strength and elasticity response of clay treated with bottom ash has been derived in this paper. However, it is reported that the performances of the formulas could be poor when derived by a conventional method of regression due to highly nonlinear behavior of clay to strength and elasticity (Horpibulsuk, 2001; Miura et al., 2001; Narendra et al., 2006).

Genetic approaches such genetic algorithm (Goldberg, 1989) and genetic programming (Koza, 1992) present some opportunities to the solution of nonlinear problems. In addition to these previous ones, a new soft computing technique called genetic expression programming (GEP) that has recently been developed by Ferreira (2001) relatively offers a good process of evolution for finding function and identifying key variables in an efficient way. Applying the GEP method to a given finite data set (or measured data) provides with a fitting to an appropriate mathematical formula by an advantage that can be easily manipulated in practical circumstances. Determination or identification of key variables and the variable combinations in the comprehensive developed models are greatly benefited from the GEP approach with its good experience in choosing the main factors for a model. This would be very helpful to explain the complex relationships between the variables (Ferreira, 2001; Baykasoglu et al., 2008; Canakci et al., 2009; Güllü, 2012, 2013). Genetic approaches (GEP and previous genetic methods) are well suited to the geotechnical problems in many applications (see Güllü (2012) for referred applications). For instance, the nonlinear characteristic of soil–water relations was successfully defined by Johari et al. (2006) using genetic principles. The genetic-based-prediction of strength and elastic property of geomaterial has been researched to some extent. However, most of the researches incorporated the functions of variables involving with rock material only. Baykasoglu et al. (2008) applied the genetic approaches (multi-expression, genetic expression and linear genetic programming) to estimate compressive and tensile strength of limestone for the first time, and found the mathematical functions of strength with good predictions. A similar study has been successfully performed for deriving the functions to the strength prediction of basalt (Canakci et al., 2009). While simple functions from genetic approach may not be adequate in explaining strength relations with index properties, the genetic approach is found to be a potential tool for identifying the key optimal variables for function derivation to the prediction of the elasticity modulus and the strength of granitic rocks (Karakus, 2011). To the author's knowledge from the literature that there is a lack of research about the function identification for the strength and elastic properties of clay, in particular to the stabilization with bottom ash.

This paper presents a preliminary attempt about finding some mathematical functions using GEP approach for the estimation of strength and elastic properties of clay treated with bottom ash for soil stabilizations. Based on some engineering aspects of bottom ash in the past studies (Kayabali and Bulus, 2000; Kim et al., 2011; Kim and Do, 2012), the independent variables (or input variables) to the function derivation have been included by the engineering parameters of treatments that are bottom ash dosage, dry unit weight, relative compaction, brittleness index and energy absorption capacity. As for the dependent variables (or output variable), they are taken unconfined compressive strength and elasticity modulus into consideration. In order to construct a database of computations, an extensive experimental study was conducted primarily based on unconfined compression test at the soil+bottom ash mixtures including the bottom ash dosages from 0 to 30%. The formulas evolved by the GEP proposed in this paper are introduced first time. It is believed that they could be relatively employed in the assistance of engineering characterization of soil stabilization.

2. Genetic expression programming

In this section, some fundamentals (definition, component, structure, etc.) of genetic expression programming (GEP) to model production will be introduced for motivation and the algorithm employed in this study will be presented. For a comprehensive background of GEP, the reader is referred to Ferreira (2001). GEP can be considered as a natural extension of previous genetic approaches, genetic algorithm (GA) genetic programming (GP). It was first invented and developed by Ferreira (2001). Similar to GA and GP; GEP run a process of using populations of individuals, selecting them according to fitness and presenting genetic variation by one or more genetic operators. The main difference between the three methods is coming from the nature of individuals. While the individuals in GAs are linear strings of fixed length (chromosomes), they are nonlinear entities of different sizes and shapes (parse tress) in GP. As for the individuals in GEP models, they are encoded as linear strings of fixed length (chromosomes or genome) which are expressed as nonlinear entities of different size and shapes (i.e., expression trees or simple diagram representations) (Ferreira, 2001). Based on numerical applications, the GEP approach is able to significantly outperform the conventional evolutionary approaches (Ferreira, 2001; Baykasoglu et al., 2008; Canakci et al., 2009; Güllü, 2012, 2013). The advantage of GEP method is attributed to simple entities of chromosomes and expression tree (ET)s. The structures of chromosomes are linear, compact, relatively small and easy to genetically operate (i.e., replication, mutation, recombination, transposition, etc.) for the solution of problem considered. The ETs, object of selection, are the presentation of their chromosomes exclusively, and they are selected to reproduce with modification according to fitness. The evolution operation through these separate entities (i.e., chromosomes and ETs) with distinct functions allows the GEP algorithm to find mathematical expressions with high efficiency as compared with existing adaptive techniques (Ferreira, 2001). The evolutions in GEP are conducted via the main components including: (i) function set (i.e., arithmetic operators such as +, −, *, /, Sqrt, Exp, Ln, etc. or Boolean functions such as OR, AND, IF, etc.), (ii) terminal set (i.e., variables like *a*, *b*, *c*, etc., and constants like 1, 2, 3, etc.), (iii) fitness function (i.e., definition of expression by a fitness case within a certain error or correlation coefficient, (iv) control parameter and (v) stop condition. The goal by the fitness function is to obtain an optimal solution for the expression that performs well for the fitness case. Individuals in the expression of problem according to fitness function are mostly selected by roulette-wheel sampling coupled with the cloning of the best individual (Canakci et al., 2009). The fitness function is mathematically expressed by the error chosen, absolute error (Eq. (1a)) or relative error (Eq. (1b)) as follow (Ferreira, 2001)

$$\text{Fitness}_i = \sum_{j=1}^{C_t} (M - |C_{(i,j)} - T_{(j)}|) \tag{1a}$$

$$\text{Fitness}_i = \sum_{j=1}^{C_t} \left(M - \left| \frac{C_{(i,j)} - T_{(j)}}{T_{(j)}} 100 \right| \right) \tag{1b}$$

where "M" is the range of selection, "$C_{(i,j)}$" is the value returned by the individual an individual chromosome "i" for fitness case "j", "C_t" is the total number of fitness cases and "T_j" is the target value for fitness case "j". It is noted that the system with this kind of fitness function can find the optimal solution for itself. When the fitness varies within small limits, GEP runs can be stopped.