Advanced Engineering Informatics 38 (2018) 593–604

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Full length article

Prediction of soil compression coefficient for urban housing project using novel integration machine learning approach of swarm intelligence and Multi-layer Perceptron Neural Network

Dieu Tien Bui[a,b,*], Viet-Ha Nhu[c], Nhat-Duc Hoang[d]

[a] Geographic Information Science Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[b] Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[c] Department of Geological-Geotechnical Engineering, Hanoi University of Mining and Geology, No. 18 Pho Vien, Duc Thang, Bac Tu Liem, Hanoi, Viet Nam
[d] Faculty of Civil Engineering, Institute of Research and Development, Duy Tan University, P809 – 03 Quang Trung, Da Nang 550000, Viet Nam

ARTICLE INFO

Keywords:
Artificial neural network
Particle Swarm Optimization
Compression coefficient
Metaheuristic
Geotechnical engineering
Vietnam

ABSTRACT

In many engineering projects, the soil compression coefficient is an important parameter used for estimating the settlement of soil layers. The common practice of determining the soil compression coefficient via the oedometer test is time-consuming and expensive. This study proposes a machine learning solution to replace the conventional tests used for obtaining the coefficient of soil compression. The new approach is an integration of the Multi-Layer Perceptron Neural Network (MLP Neural Nets) and Particle Swarm Optimization (PSO). These two computational intelligence methods work synergistically to establish a prediction model of soil compression coefficient. The PSO metaheuristic is employed to optimize the MLP Neural Nets model structure. To train and validate the proposed method, named as PSO-MLP Neural Nets, a dataset of 154 soil samples featuring 12 influencing factors has been collected from the geotechnical investigation process of a high-rise building project. Experimental results show that the proposed PSO-MLP Neural Nets has attained the most accurate prediction of the soil compression coefficient performance with RMSE = 0.0267, MAE = 0.0145, and R^2 = 0.884. The result of the proposed model is significantly better than those obtained from other benchmark methods including the backpropagation neural network, the radial basis function neural network, the support vector regression, the random forest, and the Gaussian process. Based on the experimental results, the newly constructed PSO-MLP Neural Nets is very potential to be a new alternative to assist geotechnical engineers in design phase of civil engineering projects.

1. Introduction

Soil compression is a phenomenon that disturbs the soil structure, changes the soil properties such as porosity, water drainage, and the arrangement of soil particles [1]. In principle, the definition of the soil compression can be stated as the decrease in volume of the soil under pressure caused by the drainage of pore water [2]. Therefore, the estimation of soil compressibility is mostly useful for fine-grained soils with low permeability such as clays. In recent studies on soil compression, the Compression Coefficient (C_c) is the main indicator of the compressibility of soil and this index is usually employed for computing the settlement of soil layers [3–5]. It is also noted that the determination of soil compressibility is very challenging for fine-grained soils due to their low permeability and also due to the fact that the rate of pore water drainage in such types of soil is a complex time-dependent process [2].

The accurate estimation of the degree of soil compression is a crucial task in the design of foundation structures. It is because the soil compressibility is a key factor in the calculation of the settlement of the soil layers beneath the structure [5,6]. As pointed out by Li [7], the estimation of soil layers' post-construction settlement is an important issue for the serviceability of various construction projects. In practice, the settlement caused by the increments of loads is derived by the employment of the logarithm of the normal compressive stress against soil void ratio curve [3]. To obtain the final calculation result of settlement for a soil layer of normally consolidated clay, besides the input variables of the initial void ratio, the effective overburden pressure, the applied load, the thickness of the layer, the index of C_c needs to be accurately determined [2].

In fact, the compression coefficient C_c is widespread utilized in the

* Corresponding author.
E-mail address: buitiendieu@tdt.edu.vn (D. Tien Bui).

https://doi.org/10.1016/j.aei.2018.09.005
Received 28 April 2018; Received in revised form 15 September 2018; Accepted 25 September 2018

design phases of various civil engineering structures including dams, retaining walls, pavements, landfill liners, and building foundations. This coefficient is usually obtained in the laboratory via the oedometer test [8]. As stated by Polidori [9], the result of C_c value determined by the oedometer test is time-consuming and expensive. One of the reasons is that specific laboratory equipment is required. Another reason is that experienced geotechnical engineers as well as highly skilled technicians are necessary to perform the test and attain reliable results. Hence, it is of practical need to come up with intelligent data-driven methods to determine the C_c values based on existing oedometer tests [10–12].

To alleviate the difficulties of the laboratory tests, various approaches have been proposed to infer the soil compressibility expressed by C_c. In essence, the C_c value is estimated through a mapping function in which a set of soil properties is used as input variables. The simplest forms of mapping functions are empirical equations that attempt to correlate C_c with basic properties of soils including the liquid limit, natural water content, plasticity index, specific gravity, and void ratio [13–19]. The aforementioned empirical equations are frequently used in the geotechnical engineering field to avoid extensive laboratory or field testing [10]. Nevertheless, the soil materials are highly complex due to their stress–strain time-conditioning response, the nonlinear stress–strain relationships, as well as elasto-plastic performance under loading and unloading conditions [11]. Hence, traditional empirical equations are insufficient to describe the mapping relationship between soil properties and C_c accurately.

Thus, recent studies have resorted to advanced machine learning algorithms as potential alternatives to prediction C_c of fine-grained soils. Machine learning algorithms, with their competence in nonlinear modeling as well as the capability of generalizing inference rules from data, provide feasible tools for understanding and simulating various complex geotechnical process [1,10,20–24] including C_c predictions [11,12].

Park and Lee [13] and Kurnaz, Dagdeviren, Yildiz and Ozkan [12] proposed artificial neural network approach for predicting compressibility parameters of soils; these parameters have been proved to be useful for computing the settlement for fine-grained soil layers. Although neural network is a capable tool for studying nonlinear processes, the conventional training method of neural network relies on backpropagation algorithm which is often suffered from local optimal stagnation. This fact considerably reduces the performance of the neural network methods.

Various curve-fitting methods associated with data collection have been proposed by Wang, Qiu, Hao and Zhang [25] to estimate compression characteristics of artificially mixed soil. Chu, Wu, Deng, Chen and Wang [26] examined the intrinsic compression behavior of remolded sand–clay mixture using regression analysis. The basic drawback of curve-fitting or traditional regression analysis approaches is their limitation in handling multivariate dataset. These approaches also require the expert judgement in the degree of nonlinearity of the collected data. Therefore, regression analysis methods often feature low generalization capability. A regression equation can perform well with a data set but may perform poorly with another one.

Mohammadzadeh, Bolouri and Alavi [2] introduced the application of multi-expression programming in automatically constructing a C_c predictor from experimental data; input variables required by the predictors are the liquid limit, the plastic limit and the void ratio of fine-grained soils. Multi-gene genetic programming, an extension of the traditional genetic programming, has been used by Mohammadzadeh, Bolouri, Vafaee and Alavi [4] for the same task of interest. Recently, the group method of data handling and the genetic algorithms were utilized to predict compressibility indices of saturated clays [11]. Although these approaches of genetic programming can explicitly show the constructed models used for C_c, their physical meaning is often hard to interpret, the reason is that those models are established via the concept of 'survival of the fittest' without the aim of model interpretability. The aim of the model construction process is to provide an equation which best fits the collected data. Moreover, the performance of genetic programming based approach can be only comparable to other black-box machine learning methods [27,28].

In the field of machine learning as well as geotechnical engineering, Multi-Layer Perceptron Neural Network (MLP Neural Nets) remains a powerful method for simulating complex processes [29]. The structure of a MLP Neural Nets comprises a network of interconnected neurons. Each neuron is an individual information processing unit. Due to its universal learning capability [30], MLP Neural Nets is able to generalize the complex functional relationships between the soil properties and the output of C_c. Before the prediction phase of C_c, it is mandatory to learn the MLP Neural Nets model structure from the training data set via the process of adapting connection weights among the neurons. This process of connection weight adjustment can be considered as a complex optimization problem. The conventional method used for the learning an MLP Neural Nets model relies on gradient descent algorithms with backpropagation [31]. Although this conventional algorithm can help MLP Neural Nets to be trained quickly with acceptable performance in many applications [32–35], MLP Neural Nets models trained by backpropagation are susceptible to premature convergence and being entrapped in local optimal solutions [36].

To alleviate the aforementioned drawback in the MLP Neural Nets training process, various studies have resorted to metaheuristic algorithms to replace the conventional approach of backpropagation. Metaheuristic provides a flexible and effective method for solving complex and large-scale optimization problems [37]. These features make MLP Neural Nets a very suitable tool for training MLP Neural Nets. Metaheuristic approaches such as Biogeography-Based Optimization [38], Differential Evolution [39], Particle Swarm Optimization [40], Cuckoo Search [41], Imperialist Competitive Algorithm [42], Kidney-Inspired Algorithm [43], Symbiotic Organisms Search [44], Ant colony optimization [45], Dynamic Group Optimisation [46], Flower Pollination Algorithm [47] have been successfully employed to construct various MLP Neural Nets models to solve complex problems in many engineering disciplines. A recent review work done by Ojha, Abraham and Snášel [30] clearly showed an increasing trend of MLP Neural Nets-metaheuristic integration in complex process modeling. Nevertheless, the applicability of metaheuristic-trained MLP Neural Nets has not yet been investigated in the problem of C_c estimation.

Therefore, this study attempts to fill this gap in the literature by proposing a new C_c prediction model by hybridizing MLP Neural Nets and the Particle Swarm Optimization (PSO). The newly constructed model is therefore named as Particle Swarm Optimized Multi-Layer Perceptron neural network (PSO-MLP Neural Nets). Moreover, a dataset of 156 testing samples including 12 influencing factors has been collected to train and verify the PSO-MLP Neural Nets approach. The subsequent part of this paper is organized as follows: the second section reviews the the employed algorithms of MLP Neural Nets and PSO, followed by the third section that describes the study site and the collected data; the fourth section depicts the proposed model structure in detail; the next section reports the experimental results, followed by several remarks on the study at the final section.

2. Background of the employed algorithms

2.1. Multilayer Perceptron Neural Network (MLP Neural Nets)

MLP Neural Nets is a supervised machine learning method that resembles biological neural networks in the natural world. This machine learning method is designed for knowledge acquisition and inference. These two processes are formally termed as the training and prediction phases. With the capability of universal approximation, MLP Neural Nets has been extensively used for modeling nonlinear and complex processes occurred in the real world [32,34,48–50]. The structure of MLP Neural Nets can be divided into three connected layers: input, hidden, and output layers. The input layer has the number of nodes that