Applied Soft Computing 45 (2016) 173–186

Contents lists available at ScienceDirect

Applied Soft Computing

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

Groutability estimation of grouting processes with cement grouts using Differential Flower Pollination Optimized Support Vector Machine

Nhat-Duc Hoang [a,*], Dieu Tien Bui [b], Liao Kuo-Wei [c]

[a] Faculty of Civil Engineering, Institute of Research and Development, Duy Tan University, P809-K7/25 Quang Trung, Danang 550000, Vietnam
[b] Geographic Information System Group, Dept. of Business Administration and Computer Science, University College of Southeast Norway, Hallvard Eikas Plass, N-3800 Bø i Telemark, 3925, Norway
[c] Dept. of Civil and Construction Engineering, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Da'an Dist., Taipei 106, Taiwan

ARTICLE INFO

Article history:
Received 15 September 2015
Received in revised form 25 March 2016
Accepted 25 April 2016
Available online 3 May 2016

Keywords:
Groutability estimation
Flower Pollenation Algorithm
Differential Evolution
Support Vector Machine
Model evaluation functions
Input factor selection

ABSTRACT

This research presents a soft computing methodology for groutability estimation of grouting processes that employ cement grouts. The method integrates a hybrid metaheuristic and the Support Vector Machine (SVM) with evolutionary input factor and hyper-parameter selection. The new prediction model is constructed and verified using two datasets of grouting experiments. The contribution of this study to the body of knowledge is multifold. First, the efficacies of the Flower Pollenation Algorithm (FPA) and the Differential Evolution (DE) are combined to establish an integrated metaheuristic approach, named as Differential Flower Pollenation (DFP). The integration of the FPA and the DE aims at harnessing the strength and complementing the disadvantage of each individual optimization algorithm. Second, the DFP is employed to optimize the input factor selection and hyper-parameter tuning processes of the SVM-based groutability prediction model. Third, this study conducts a comparative work to investigate the effects of different evaluation functions on the model performance. Finally, the research findings show that the new integrated framework can help identify a set of relevant groutability influencing factors and deliver superior prediction performance compared with other state-of-the-art approaches.

1. Introduction

Permeation grouting is an efficient approach for ground improvement in geotechnical engineering [1,2]. Different from chemical grouts, cement grouts are not only capable of meliorating the groutability for the target geomaterial, but also do not give rise to groundwater pollutions [3–5]. Since the cement grouts are increasingly applied in the construction industry, enhancing the predictive capability of the grouting process can be very helpful for geotechnical engineers in ground improvement projects.

Based on recent studies, traditional formula-based methods with their limited predictive capability and restricted flexibility have been shown to be inadequate for delivering satisfactory performances [5–8]. The reason is that various factors can affect the grouting process; moreover, the infiltration of cement into the grouted soil is governed by a complex transport process of cement grains through the soil matrix [6]. Thus, the task of learning the distinction boundary between groutable and non-groutable cases can be very challenging.

Recently, modern soft computing approaches have been proved to be viable and robust methods for learning and analyzing grouting processes with cement grouts. Tekin and Akbas [6] and Liao et al. [5] have constructed groutability estimation models based on the Artificial Neural Network (ANN). Although the ANN is viable for nonlinear classification including the problem of groutability estimation, its performance has been shown to be inferior to that of the Support Vector Machine (SVM) [7,9]. Instance-based learning approaches [10] and Bayesian frameworks [4] have also been attempted to construct groutability estimation models.

Among machine learning approaches, the SVM is a very efficient tool for solving learning and prediction problems in complex circumstances [11]. Verbiest et al. [12] state that the SVM algorithm is one of the most powerful, popular, and accurate classifiers. A particular capability of the SVM is that the method can generalize well even in the situation of limitted data samples. Mountrakis

* Corresponding author.
E-mail addresses: hoangnhatduc@dtu.edu.vn (N.-D. Hoang), Dieu.T.Bui@hit.no, BuiTienDieu@gmail.com (D. Tien Bui), kliao@mail.ntust.edu.tw (K.-W. Liao).

http://dx.doi.org/10.1016/j.asoc.2016.04.031
1568-4946/

et al. [13] summarize that the advantages of the this classification technique include self-adaptability in model structure, swift learning pace, and limited requirements on the size of training datasets. These advantages stem from the fact that the learning method of the SVM is established by the structural risk minimization which is less prone to overfitting than the empirical risk minimization employed by the ANN learning approach. Previous works have demonstrated the capability of the SVM in modeling grouting processes with microfine cement [7,9] and in other research domains [12,14–18]. Nevertheless, to the best of our knowledge, the application of the SVM in estimation of cement-based grouting has not yet been investigated.

Furthermore, the SVM implementation necessitates a proper determination of its hyper-parameters and a set of appropriate input factors. The reason is that the SVM hyper-parameters directly affect the model complexity which is crucial to achieve a balance between bias and variance. On the other hand, choosing a set of relevant input factors from a pool of candidates can help avoid the curse of dimensionality and potentially improves the model performance [19,20]. Nonetheless, the issue of feature selection for groutability prediction has rarely been addressed in the literature.

Since the task of parameter determination and input factor selection can be formulated as an integrated optimization problem [21,22], employing metaheuristics coupled with soft computing-based learning methods can be very promising to establish a new groutability estimation model. It is because metaheuristics have been proved to be effective tools for dealing with complicated optimization tasks [23–27]. Furthermore, this research aims at contributing to the body of knowledge by putting forward a hybrid metaheuristic algorithm, named as Differential Flower Pollination (DFP), which inherits the advantages of the Differential Evolution (DE) [28] and the recently developed Flower Pollination Algorithm (FPA) [29].

The DE is undoubtedly one of the most powerful metaheuristic in current use [27,30]. Nevertheless, as pointed out by Das and Suganthan [27] and Neri and Tirronen [31], the DE is sensitive to the position of the initial population. Experimental simulations done by Sa et al. [32] revealed undesirable behaviors of the DE when all individuals of the population are in a basin of attraction of a local optimum; in this situation, the population cannot escape from it. Furthermore, as the distance between two randomly chosen solutions within the population gradually shrinks, the step size becomes smaller and the searching process is basically performed in the neighborhood of the current solutions [33]. That is the reason why the DE is only explorative at the beginning phase of the optimization process and focuses on exploitation during the latter phase. Thus, it is beneficial to enhance the explorative capability of the DE in latter phase of the evolutionary process. The reason is that this enhancement can help the algorithm to achieve a better trade-off between exploration and exploitation and alleviate the risk of stagnation and premature convergence [34].

In addition, the FPA, proposed by Yang [29], is inspired by the pollination process of flowering plants. This algorithm features two mechanisms of searching: local and global pollinations. These two mechanisms are complementary and facilitate the FPA to effectively explore the information of the search space [35]. The FPA possesses an excellent way of controlling the balance between explorations and exploitations by employing a switch probability p. Notably, the Lévy distribution used in the global pollination process has the ability to generate new candidate solutions with bigger mutation step size [36]. Experiments carried out by Yang [29] demonstrate that the FPA is more efficient than the Genetic Algorithm and Particle Swarm Optimization. A recent review work carried out by Chiroma et al. [37] points out that there is much room for the extension of the FPA. From our view, the current local pollination operator of the FPA is still relatively simple and this operator can be enhanced to meliorate the performance of the algorithm. Therefore, our research proposes to integrate the DE and the FPA to harness the strength and complement the weakness of each individual algorithm. The merits of the DFP are verified by a set of 15 benchmark functions and the new algorithm is applied to optimize the SVM-based groutability prediction model.

Furthermore, since the aim of learning is to find a model that features the trade-off between model generalization and complexity, the objective function used to quantify the fitness of a learning model is indeed crucial and it should receive more attention in hybrid soft computing models. Investigating the effect of different forms of the objective function on the model performance may reveal interesting facts regarding the studied datasets and the employed soft computing technique. The remaining part of this paper is organized as follows. Section 2 presents the research method. The DFP algorithm is presented in Section 3. Section 4 describes the proposed groutability estimation approach followed by the experimental results. Section 6 summarizes conclusion of this research.

2. Research methodology

2.1. Datasets of grouting experiments

The first data set (Dataset 1) of grouting experiments used in this paper includes 87 laboratory tests [6]. The following nine factors are characterized as influencing factors: water/cement ratio of the grout, relative density of the soil (%), grouting pressure (kPa), content of soil passing through a 0.6 mm sieve (%), the diameter of a sieve through which 10% of the soil passes (mm), the diameter of a sieve through which 15% of the soil passes (mm), the diameter of a sieve through which 85% of the cement grout passes (mm), the diameter of a sieve through which 90% of the cement grout passes (mm), and the diameter of a sieve through which 95% of the cement grout passes (mm). The detailed statistical description of the dataset is shown in Table 1. It is noted that the numbers of 'groutable' and 'non-groutable' cases in the dataset are 51 and 36, respectively.

The second dataset (Dataset 2) used in this study consists of 240 in-situ permeation grouting cases for sandy silt soil obtained during the construction processes of highway projects in Kaohsiung city and Mass Rapid Transit projects in Taipei city of Taiwan [5,38]. A mixture of microfine cement and micro-slag in equal proportions was employed as the grouting material. The diameters through which 95%, 90%, and 85% of the total grout passes are 7.4 μm, 6.4 μm, and 4.5 μm, respectively. In addition, the diameter through which 70% of the total grout passes is less than 1 μm. The grout is considered to be semi-nanometer material; thus, the grouting material used in this dataset is microfine cement. Seven factors are recorded to characterize the grouting process (see Table 2). In Dataset 2, the numbers of 'groutable' and 'non-groutable' cases are 155 and 85, respectively.

2.2. Metaheuristics approaches

2.2.1. Flower Pollenation Algorithm (FPA)

The FPA [29] is inspired from the characteristics of the flower pollination process of flowering plants. In the natural world, pollination is the process by which pollen is transported from the male section (anther) to the female section (stigma) of the plant. Typically, there are two means of pollination: biotic and abiotic [39]. The first form of pollination is often achieved by the assistance of pollinators such as insects, birds, and other animals. The abiotic pollination does not involve pollinators; instead, wind and water flow are means of pollen delivery.