Computers and Geotechnics 79 (2016) 41–54

Contents lists available at ScienceDirect

Computers and Geotechnics

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

Research Paper

Efficient system reliability analysis of soil slopes using multivariate adaptive regression splines-based Monte Carlo simulation

Lei-Lei Liu, Yung-Ming Cheng *

Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong

ARTICLE INFO

Article history:
Received 21 January 2016
Received in revised form 23 March 2016
Accepted 1 May 2016

Keywords:
Slope stability
System reliability analysis
Monte Carlo simulation
Multivariate adaptive regression splines
Response surface method

ABSTRACT

System effects should be considered in the probabilistic analysis of a layered soil slope due to the potential existence of multiple failure modes. This paper presents a system reliability analysis approach for layered soil slopes based on multivariate adaptive regression splines (MARS) and Monte Carlo simulation (MCS). The proposed approach is achieved in a two-phase process. First, MARS is constructed based on a group of training samples that are generated by Latin hypercube sampling (LHS). MARS is validated by a specific number of testing samples which are randomly generated per the underlying distributions. Second, the established MARS is integrated with MCS to estimate the system failure probability of slopes. Two types of multi-layered soil slopes (cohesive slope and c–φ slope) are examined to assess the capability and validity of the proposed approach. Each type of slope includes two examples with different statistics and system failure probability levels. The proposed approach can provide an accurate estimation of the system failure probability of a soil slope. In addition, the proposed approach is more accurate than the quadratic response surface method (QRSM) and the second-order stochastic response surface method (SRSM) for slopes with highly nonlinear limit state functions (LSFs). The results show that the proposed MARS-based MCS is a favorable and useful tool for the system reliability analysis of soil slopes.

1. Introduction

Soil properties are uncertain due to the complex geologic origins of soil deposits and the various methods of exploration and testing. Thus, a conventional factor of safety (FS), which is usually calculated based on a set of deterministic soil strength parameters, does not accurately reflect the state of slope stability. For example, approximately 5% of the stabilized slopes in Hong Kong that are designed based on classical deterministic slope stability analyses, without consideration of the uncertainty in soil parameters, eventually fail [15]. For this problem, probabilistic methods provide a systematic and quantitative tool to address the effect of this uncertainty [31]. Fruitful contributions to slope reliability analyses have been achieved in recent decades, including studies by Li and Lumb [35], Christian et al. [10], Hassan and Wolff [19], Griffiths and Fenton [18], Low [38], Jimenez-Rodriguez et al. [27], Cho [7], Low et al. [39], Zhang et al. [49], Li et al. [32], Wang [44], Li et al. [36], Xu et al. [47], Gong et al. [17], Li et al. [33] and Jiang et al. [25]. These studies indicate that Monte Carlo simulation (MCS) is the most popular approach as it is conceptually simple and can provide unbiased reliability results. However, MCS is inefficient when directly evaluated based on the original deterministic stability model. To achieve an acceptable accuracy in MCS, a minimum of 10^4 direct deterministic stability analyses are required [42]. Hours and days are indispensable to performing a MCS, even for modern computer architecture, which is tremendously inconvenient for routine engineering design. This finding explains the low acceptance of slope reliability analysis by engineers.

The response surface method (RSM) integrated with MCS is considered to be a suitable alternative for slope reliability analysis. The merit of RSM is the construction of an explicit closed-form expression between the input variables and the corresponding output responses to accurately approximate the implicit limit state functions (LSF) of a slope. Thus, time-consuming deterministic stability models are not required. Instead, MCS is performed based on the constructed RSM, which is expected to be significantly more efficient without sacrificing the evaluation accuracy. Regarding the application of RSM in slope reliability, previous studies have addressed applications such as the polynomial-based RSM [37,46], Kriging methodology [40,48,51], the artificial neural network (ANN) [6] and the support vector machine (SVM) [30]. The polynomial-based RSM has been the most prevalent application, and Wong [45] was the first researcher to establish a linear RSM

* Corresponding author.
E-mail address: ceymchen@polyu.edu.hk (Y.-M. Cheng).

http://dx.doi.org/10.1016/j.compgeo.2016.05.001
0266-352X/

to approximate the finite element model of a homogenous slope. However, nonlinear stability problems are prevalent in slope engineering. Thus, the quadratic response surface method (QRSM) is becoming popular for more complicated system reliability analyses, of which the validity and capacity of the method have been extensively demonstrated by published studies [24,33,34,37,46,49]. In addition, an extension of the QRSM based on the Hermite polynomial chaos expansion, which is referred to as the stochastic response surface method (SRSM), has been proven to be a powerful approximation to the LSF of a slope [23,25,31]. A recent and more detailed review of the RSM is given by Li et al. [34]. A common feature of the QRSM and the SRSM is that they pertain to the category of parametric regression methods, which must be employed with a prior assumption on the order and type of polynomials. If a response surface (e.g., a QRSM) is erroneously assumed to approximate the true LSF, which may be a multimodal function with several peaks and troughs, and the corresponding results would not be credible as expected. For a specific slope system, the degree of nonlinear property is usually an unknown for the engineers or researchers, and the QRSM is often adopted based on the rule of thumb or even intuition, which may cause unsatisfactory estimation of the probability of failure (P_f). As reported by Zhang et al. [51], the traditional QRSM may not accurately approximate the LSFs for the system reliability analysis of soil slopes. High-order SRSMs seem to perform better than their low-order counterparts when small P_f (lower than or on the order of 10^{-3}) values are involved, as demonstrated by Li et al. [31]. Therefore, an acceptable RSM should be automatically determined by lending itself to available data samples without potentially wrong assumptions; this model is referred to as the data-driven RSM.

The objective of this paper is to suggest a data-driven RSM for the system reliability analysis of soil slopes, and MARS is selected for this purpose. The application of MARS to geotechnical engineering is a relatively recent development that primarily involves soil liquefaction assessment [28,53], tunnel convergence prediction [1], and other geotechnical problems [52]. To the best of our knowledge, MARS as a tool for slope reliability analysis has not been well appreciated. It is adopted to automatically establish the response surface to approximate the implicit LSF of a slope based on a limited number of training samples generated by Latin hypercube sampling (LHS). A hundred randomly generated testing samples are obtained to validate the MARS model. Then, MCS is performed based on the constructed MARS for the system reliability analysis of soil slopes. Four examples are examined to demonstrate the capacity and validity of the proposed approach. For convenient comparison purposes, the results obtained by the proposed approach are also compared with those evaluated by QRSM, SRSM and several other available reliability approaches.

The remainder of this paper is organized as follows: In Section 2, we will provide a brief review of the most popular RSMs: QRSM and SRSM. In Section 3, a detailed theory of MARS is presented. MARS-based MCS for system reliability analysis of soil slopes is introduced in Section 4. In Section 5, the applicability and validity of the proposed MARS-based MCS are illustrated using four practical examples from literatures. The conclusions are presented in Section 6.

2. Review of QRSM and SRSM

2.1. QRSM

For conventional engineering design, the FS of a slope is usually calculated using a limit equilibrium method (LEM) or a finite element method (FEM), which only requires the definitions of the strength parameters. The basic idea of the QRSM is to establish an explicit expression between the FS and soil shear strength parameters to approximate the implicit LSF of a slope using quadratic polynomials. The quadratic polynomials without cross terms, which are commonly utilized by researchers [3,33,37,49], are defined as

$$\mathrm{FS}(\boldsymbol{X}) = a_0 + \sum_{i=1}^{p} a_i x_i + \sum_{i=1}^{p} a_{i+p} x_i^2 \tag{1}$$

where $\mathrm{FS}(\boldsymbol{X})$ is the FS for a given vector of input variables $\boldsymbol{X} = (x_1, x_2, \ldots, x_p)$; p is the number of variables; and $\boldsymbol{a} = (a_0, a_1, \ldots, a_{2p})^T$ is the vector of unknown coefficients. To calibrate Eq. (1), discrete training data sets must be obtained from direct evaluation of the true LSF, and the number of training samples should not be less than $2p + 1$. Suitable training samples should be as small as possible and contain effective information (e.g., inflexion) about the true LSF. Design of experiments (DOE), such as central composite design and LHS can be employed as references. The calibrated QRSM can serve as a substitution of the true LSF, with probabilistic methods such as MCS, to perform the probabilistic analysis of slope stability.

2.2. SRSM

Unlike the QRSM, the premise of the SRSM is to approximate the potential relationship between the FS and the soil strength parameters in terms of random variables by a polynomial chaos expansion [21,31]. The expansion usually proceeds in the following steps: First, input variables are represented by selected random variables, such as standard normal variables. Second, the FS is expressed in the form of polynomial chaos expansion with the foregoing random variables. Third, unknown coefficients in the polynomial chaos expansion are determined as described in the QRSM. Last, a reliability analysis of slope stability is performed based on the SRSM using a probabilistic approach, such as MCS. A Hermite polynomial chaos expansion, which is the most common expansion in terms of independent standard normal variables, is usually adopted. The SRSM is expressed as

$$\begin{aligned}\mathrm{FS}(\xi) = {} & a_0\Gamma_0 + \sum_{i_1=1}^{p} a_{i_1}\Gamma_1(\xi_{i_1}) + \sum_{i_1=1}^{p}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Gamma_2(\xi_{i_1}, \xi_{i_2}) \\ & + \sum_{i_1=1}^{p}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}\Gamma_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) \\ & + \cdots + \sum_{i_1=1}^{p}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2}\cdots\sum_{i_p}^{i_{p-1}} a_{i_1 i_2,\ldots,i_p}\Gamma_p(\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_p})\end{aligned} \tag{2}$$

where $\mathrm{FS}(\xi)$ is the FS for a given vector of independent standard normal variables $\boldsymbol{\xi} = (\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_p})$ that represents the uncertainty in the input variables; p is the total number of random variables; $\boldsymbol{a} = (a_0, a_{i_1}, \ldots, a_{i_1 i_2,\ldots,i_p})^T$ is the vector of unknown coefficients to be determined; and $\Gamma_p(\cdot)$ is the multi-dimensional Hermite polynomial of order p (assumed to be two by default in this study), which is expressed as

$$\Gamma_p(\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_p}) = (-1)^p e^{\frac{1}{2}\xi^T\xi} \frac{\partial^p}{\partial\xi_{i_1}\partial\xi_{i_2}\ldots\partial\xi_{i_p}} e^{-\frac{1}{2}\xi^T\xi} \tag{3}$$

3. MARS-based RSM

3.1. Basic theory of MARS

MARS was introduced by Friedman [14] as a flexible statistical strategy to represent the relationship between a group of input variables and their dependent outputs. Generally, this underlying functional relationship is identified by lending itself to and