Computers and Geotechnics 74 (2016) 88–101

Contents lists available at ScienceDirect

Computers and Geotechnics

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

Research Paper

A generalized surrogate response aided-subset simulation approach for efficient geotechnical reliability-based design

Dian-Qing Li [a], Ke-Bo Shao [a], Zi-Jun Cao [a,*], Xiao-Song Tang [a], Kok-Kwang Phoon [b]

[a] State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, 8 Donghu South Road, Wuhan 430072, China
[b] Department of Civil and Environmental Engineering, National University of Singapore, Blk E1A, #07-03, 1 Engineering Drive 2, Singapore 117576, Singapore

ARTICLE INFO

Article history:
Received 31 July 2015
Received in revised form 9 December 2015
Accepted 19 December 2015
Available online 14 January 2016

Keywords:
Reliability-based design
Subset simulation
Driving variable
Computational efficiency
Geotechnical structures

ABSTRACT

This paper aims to develop an efficient geotechnical reliability-based design (RBD) approach using Monte Carlo simulation (MCS). The proposed approach combines a recently developed MCS-based RBD approach, namely expanded RBD approach, with an advanced MCS method called "Subset Simulation (SS)" to improve the computation efficiency at small probability levels that are often concerned in geotechnical design practice. To facilitate the integration of SS and expanded RBD, a generalized surrogate response f is proposed to define the driving variable, which is a key parameter in SS, for expanded RBD of geotechnical structures (e.g., soil retaining structures and foundations). With the aid of the proposed surrogate response, failure probabilities of all the possible designs in a prescribed design space are calculated from a single run of SS. Equations are derived for integration of the surrogate response-aided SS and expanded RBD, and are illustrated using an embedded sheet pile wall design example and two drilled shaft design examples. Results show that the proposed approach provides reasonable estimates of failure probabilities of different designs using a single run of the surrogate response-aided SS, and significantly improves the computational efficiency at small probabilities levels in comparison with direct MCS-based expanded RBD. The surrogate response-aided SS is able to, simultaneously, approach the failure domains of all the possible designs in the design space by a single run of simulation and to generate more complete design information, which subsequently yields feasible designs with a wide range of combinations of design parameters. This is mainly attributed to the strong correlation between the surrogate response and target response (e.g., factor of safety) of different designs concerned in geotechnical RBD.

1. Introduction

In the past few decades, several geotechnical reliability-based design (RBD) codes/methodologies have been developed around the world, such as the load and resistance factor design codes (e.g., [11,25,15] or multiple resistance factor design methodology [27,28] in North America, Eurocode 7 (e.g.,[24,12] in Europe, and the Geo-Code 21 (e.g., [18,17] in Japan. These codes/methodologies are usually calibrated for a pre-defined value of target reliability index β_T (or target failure probability p_T) through some calibration processes based on some assumptions and/or simplifications of uncertainty characterization and deterministic modeling, and provide tabulated load and/or resistance factors (or partial factors) for geotechnical designs. These calibration processes are, however, almost "invisible" to geotechnical practitioners in the sense that the assumptions and/or simplifications adopted in calibration processes are unknown to them. This may lead to potential misuse of the load and resistance factors because they are only valid for the assumptions and simplifications adopted in the calibration process.

To address the problem, a Monte Carlo Simulation (MCS)-based RBD approach, namely expanded RBD approach, for foundations is recently developed by Wang et al. [37] and Wang [35]. The expanded RBD approach makes use of direct MCS in design, and its entire design process (including uncertainty characterization and deterministic modeling) is transparent to design engineers so that they can make assumptions and/or simplifications deemed appropriate for a particular project during the design. Although direct MCS used in expanded RBD has the advantage of conceptual and mathematical simplicity, it suffers from a lack of efficiency and resolution at small probability levels [10], which are of great interest in design practice. This necessitates a large number (e.g., more than 1 million) of MCS samples for expanded RBD and may require

* Corresponding author. Tel.: +86 27 6877 2496; fax: +86 27 6877 4036.
E-mail address: zijuncao@whu.edu.cn (Z.-J. Cao).

http://dx.doi.org/10.1016/j.compgeo.2015.12.010

extensive computational efforts, which hampers the application of the approach in design practice.

This paper aims to develop an efficient geotechnical RBD approach that combines the expanded RBD approach with an advanced MCS method called "Subset Simulation (SS)". SS was originally developed by Au and Beck [5] in structural engineering, and has been successfully applied to perform reliability analyses in geotechnical engineering (e.g., [39,38,30,2,19]) and other disciplines (e.g., nuclear engineering [41,42,20,13]). However, it should be noted that the reliability analysis aims to estimate the failure probability P_f of a system with pre-defined design parameters (e.g., an existing geotechnical structure). This is the inverse of RBD, which aims to determine an optimal design of a system (e.g., geotechnical structures) that satisfies a series of pre-defined performance requirements (e.g., p_T). Research that directly uses SS in geotechnical RBD is relatively limited except for a few recent studies [40,34,33]. Wang and Cao [40] showed that the implementation of SS in geotechnical RBD is not a trivial task and relies on the choice of the system response used to define the driving variable that is a key parameter in SS. Although Wang and Cao [40] suggested a heuristic method to choose the system response in SS for foundation designs, how to choose a proper system response in SS for general purposes of geotechnical RBD remains an open question and has not been systematically explored.

This paper proposes a generalized surrogate response to aid in the application of SS in expanded RBD and systemically explores its performance in geotechnical RBD. The proposed surrogate response is general and applicable to various geotechnical designs (e.g., soil retaining structures and foundations). With the aid of the proposed surrogate response, the P_f values of all the possible designs in a prescribed design space are calculated from a single run of SS for expanded RBD, avoiding repeatedly performing SS for each design. The paper starts with a brief introduction of expanded RBD approach and SS, followed by discussion on the difficulty in defining the driving variable of SS for expanded RBD and formulation of the generalized surrogate response for SS-based expanded RBD. Then, equations are derived for integration of the surrogate response-aided SS with expanded RBD and evaluation of the correlation coefficient between the surrogate response and the system response of intrinsic interest (i.e., target response) in design. Finally, the proposed approach is illustrated by three design examples.

2. Expanded reliability-based design (expanded RBD) of geotechnical structures

In the context of expanded RBD, design parameters $\underline{\theta_D}$ of geotechnical structures are artificially considered as independent discrete random variables with uniformly distributed probability mass function $P(\underline{\theta_D})$ (e.g., [35,37]. Herein, the design parameters $\underline{\theta_D} = \{\theta_{D,i},\ i = 1, 2, \ldots, n_D\}$ are referred to n_D geometric dimensions of geotechnical structures needed to be determined in designs, such as the embedded depth D_{spw} of the sheet pile wall, and the length D and diameter B of drilled shafts. They shall be distinguished from the uncertain system parameters $\underline{\theta_S} = \{\theta_{S,j},\ j = 1, 2, \ldots, n_S\}$ (e.g., uncertain soil parameters) defined in this study, which are n_S uncertain parameters involved in geotechnical reliability analyses for a given design (i.e., a given combination of $\underline{\theta_D}$). In expanded RBD, the design process of geotechnical structures is revised as a process of evaluating conditional failure probabilities $P(F|\underline{\theta_D})$ corresponding to all the possible designs with various combinations of $\underline{\theta_D}$ by a single run of direct MCS and comparing them with p_T. Feasible designs are those with $P(F|\underline{\theta_D}) \leqslant p_T$. Using Bayes' Theorem, $P(F|\underline{\theta_D})$ is calculated as [37,35]

$$P(F|\underline{\theta_D}) = P(\underline{\theta_D}|F)P_f/P(\underline{\theta_D}) \quad (1)$$

where $P(\underline{\theta_D}|F)$ = conditional joint probability of $\underline{\theta_D}$ given failure; $P(\underline{\theta_D}) = 1/\prod_{i=1}^{n_D} k_{D,i}$, in which $k_{D,i}$ is the number of possible values of $\theta_{D,i}$. In Eq. (1), $P(\underline{\theta_D})$ is known before the calculation and reflects design decisions of geotechnical practitioners. Meanwhile, $P(\underline{\theta_D}|F)$ and P_f are also needed for evaluating $P(F|\underline{\theta_D})$, and they are calculated by a single run of SS with the aid of a generalized surrogate response in this study, as discussed in the following two sections.

3. Surrogate response-aided subset simulation for expanded RBD

3.1. Algorithm of SS

SS is an advanced MCS method that uses conditional probability and Markov Chain Monte Carlo (MCMC) method to efficiently compute small tail probability [5,6,9]. It expresses a rare event E with a small probability as a sequence of intermediate events $\{E_1, E_2, \ldots, E_m\}$ with larger conditional probability and employs specially designed Markov chains to generate conditional samples of these intermediate events until the target sample domain is achieved. Let Y be the output parameter that is of interest and increases monotonically, and define the rare event E as $E = \{Y > y\}$, where y is a given threshold value for determining whether E occurs. Let $y = y_m > y_{m-1} > \cdots > y_2 > y_1 > 0$ be a decreasing sequence of intermediate threshold values. Then, the intermediate events $\{E_l, l = 1, 2, \ldots, m\}$ are defined as $E_l = \{Y > y_l, l = 1, 2, \ldots, m\}$. By sequentially conditioning on the event $\{E_l, l = 1, 2, \ldots, m\}$, the probability of event E, i.e., $P(E = \{Y > y\})$, can be written as:

$$P(E) = P(E_m) = P(E_1)\prod_{l=2}^{m} P(E_l|E_{l-1}) \quad (2)$$

where $P(E_1)$ is equal to $P(Y > y_1)$ and $P(E_l|E_{l-1})$ is equal to $\{P(Y > y_l| Y > y_{l-1}), l = 2, \ldots, m\}$. In implementations, $y_1, y_2, \ldots, y_m$ are generated adaptively using information from simulated samples so that the sample estimates of $P(E_1)$ and $\{P(E_l|E_{l-1}), l = 2, \ldots, m\}$ always correspond to a common specified value of conditional probability p_0 (e.g., 0.1) [5,6].

The efficient generation of conditional samples is pivotal to the success of SS, and it is made possible through MCMC. In this study, a modified Metropolis–Hasting (MM–H) algorithm [6,9] is used, in which the candidate sample of a random vector is generated component by component in a Markov chain and is accepted or rejected according to not only the acceptance ratio but also the occurrence of intermediate events (i.e., $E_1, E_2, \ldots, E_m$). Previous studies [7,31] have demonstrated that using the MM–H algorithm in SS allows generating conditional samples in high dimensional space and solving high dimensional reliability problems efficiently. In contrast, the classical Metropolis–Hasting (CM–H) algorithm [21,16] suffers from the curse of dimension because the acceptance ratio in the algorithm often decreases exponentially as the dimension of uncertain parameters space increases, leading to many repeated samples in the Markov chain and significant reduction of computational efficiency in high dimensional problems [9,26].

As shown in Fig. 1(a), SS starts with direct MCS, in which N MCS samples are generated. The Y values of the N samples are calculated and ranked in an ascending order. The $(1 - p_0)N$-th value in the ascending list of Y values is chosen as y_1 (see Fig. 1(b)), and hence, the sample estimate for $P(E_1) = P(Y > y_1)$ is always p_0. In other words, there are p_0N samples with $E_1 = \{Y > y_1\}$ among the samples generated by direct MCS. Then, starting from the p_0N samples with $E_1 = \{Y > y_1\}$, MCMC is used to simulate $(1 - p_0)N$ additional conditional samples given $E_1 = \{Y > y_1\}$ so that there are a total of N samples with $E_1 = \{Y > y_1\}$, as shown in Fig. 1(c). The Y