Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_00-007-cd

Slope reliability analysis and risk assessment: a modern computational perspective

Dian-Qing Li[1], Shui-Hua Jiang[2], Te Xiao[3] and Xin Liu[4]

[1]State Key Laboratory of Water Resources and Hydropower Engineering Science, Institute of Engineering Risk and Disaster Prevention, Wuhan University, 299 Bayi Road, Wuhan 430072, China.
E-mail: dianqing@whu.edu.cn
[2]School of Infrastructure Engineering, Nanchang University, Nanchang, Jiangxi 330031, China.
E-mail: sjiangaa@ncu.edu.cn
[3]Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, China.
E-mail: xiaote@ust.hk
[4]Department of Architecture and Civil Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, China.
Email: xliu268-c@my.cityu.edu.hk

Abstract: Slope failures or landslides are a major geo-hazard worldwide. Quantitative slope risk assessment and control have been used as an effective method for mitigating landslide hazards. A key task of quantitative slope risk assessment is to evaluate probability and consequences of slope failure, which relies on the understanding of site conditions, slope stability modeling and analysis, uncertainty quantification and propagation, etc. It can be rationally accomplished under a probabilistic framework through slope reliability analysis and risk assessment. Slope reliability analysis and risk assessment have an advantage of incorporating various uncertainties inevitably encountered in slope engineering, such as soil stratification uncertainty, inherent spatial variability, transformation and model uncertainties, multiple failure modes, etc. With the advance of modern computational technologies and powers, slope reliability analysis and risk assessment has advanced rapidly in recent years. This study presents some recent developments of advanced computer-based methods for slope reliability and risk assessment. Major topics include efficient slope reliability analysis and risk assessment methods based on surrogate models and advanced simulation methods, and effects of the spatial variability of soil properties on the failure modes, probability of failure and risk of slopes. The presented methods can provide versatile and promising tools for slope reliability analysis and risk assessment in spatially variable soils.

Keywords: slope; reliability analysis; risk assessment; spatial variability; modern computational methods; uncertainty.

1 Introduction

Slope failures or landslides are a major geo-hazard worldwide. To assess and mitigate the risk of slope failure, it is crucial to evaluate the probabilities and consequences of slope failure. Conventional deterministic analysis methods are unable to properly account for various sources of uncertainties (e.g., soil stratification uncertainty, inherent spatial variability of geomaterials, transformation and model uncertainties, etc.) in geotechnical engineering, which may lead to inaccurate assessment of the probability and risk of slope failure (e.g., Duncan 2000; Juang et al. 2019; Phoon et al. 2022). In contrast, the probabilistic analysis of slope stability and risk assessment is able to not only consider these uncertainties in a rational and rigorous manner, but also offer a viable way to quantitatively evaluate the slope stability from a probabilistic perspective.

The probabilistic slope stability analysis has been widely conducted with the aid of random field theory, so as to investigate the effects of the spatial variability of soil parameters and geological uncertainty on the slope reliability (e.g., Griffiths and Fenton 2004; Deng et al. 2017). The direct Monte Carlo Simulation (MCS) is a conceptually simple and robust method for probabilistic slope stability analysis in spatially variable soils, but one obvious limitation of MCS is the high computational efforts, especially for the reliability problems at small probability levels (Jiang and Huang 2016). With the advance of computing technologies, two innovative approaches, surrogate model-based methods and advanced computer-based simulation methods, were developed to alleviate the computational burden of the MCS. The surrogate models, including Polynomial Chaos Expansion (PCE)- and machine learning or deep learning models-based response surface methods, implemented in a non-intrusive manner, can effectively approximate the implicit limit state function between the output responses of slopes associated with time-consuming finite element analyses and the input random variables (e.g., Wong 1985; Jiang et al. 2014, 2015; Li et al. 2016a; He et al. 2020; Wang and Goh 2021). Advanced computer-based simulation methods, including subset simulations (SS), Latin Hypercube Sampling (LHS) and line sampling, can reduce the sample size of MCS in the slope reliability and risk analyses (e.g., Au and Beck 2001). Besides, other methods such as response conditioning-based auxiliary analysis method (Li et al. 2016b) and dimension reduction method (Li et al. 2019a) can also overcome the computational burden of the MCS.

On the other hand, the risk of slope failure was usually evaluated by the product of the probability of failure and a constant consequence. However, there are many possible failure modes once the spatial variability of soil properties is considered. The consequences induced by different failure modes are distinctively different. Deep failure often leads to more severe consequence than shallow failure (e.g., Huang et al. 2013; Jiang et al. 2017a). To this end, the quantitative risk assessment of slope failure has attracted a lot of attentions in the past decade, in which the consequences are assessed individually for each failure mode (Huang et al. 2013). The volume (or area in the two-dimensional (2-D) case) of the sliding mass is taken as an equivalent index to quantify the consequence of slope failure (e.g., Zhang and Huang 2016; Jiang et al. 2017a; Cheng et al. 2018; Ng et al. 2021). The post-failure behavior was not captured, and the influence zone and runout distance were not quantitatively estimated for quantifying the consequence of slope failure in the above studies. This may be attributed to the limitation of Finite Element Method (FEM) in simulating the large deformation of soils during slope failure. To remove these shortcomings, Random Material Point Method (RMPM) was developed to simulate the evolution of slope failure with large deformation of soils and quantify the failure features (e.g., Wang et al. 2016; Liu et al. 2019; Liu and Wang 2021). The results may provide a reasonable basis for quantitative risk assessment. Although the computational burden of RMPM is relatively large, the computational efficiency can be significantly improved with the rapid development of modern computational technology and power.

This study aims to develop modeling approaches of geological and geotechnical uncertainties in geotechnical engineering, and to present recent developments of computer-based simulation methods for slope reliability analysis and risk assessment, which can take full advantage of modern computational technologies and powers. Major topics include efficient slope reliability analysis methods using surrogate models and advanced simulation methods, and quantitative risk assessment methods of slope failure. Finally, the effectiveness of the computer-based simulation methods developed by the authors is demonstrated using three slope examples. The effects of the spatial variability of soil properties on the failure modes, probability and risk of slopes are also investigated.

2 Basic Theory of Slope Reliability and Risk Analysis

The stability performance of a slope is commonly defined by a limit state function g. The slope failure occurs when $g < 0$. The probability of slope failure P_f can be estimated as follows by computing the volume of $f(\boldsymbol{x})$ within the failure domain defined by the limit state function (e.g., Ang and Tang 2007):

$$P_f = P\left[g(\boldsymbol{x}) < 0\right] = \int \cdots \iint_{g(\boldsymbol{x})<0} f(\boldsymbol{x}) d\boldsymbol{x} \tag{1}$$

where $f(\boldsymbol{x})$ is the joint probability density function of $\boldsymbol{x}$, in which $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ is the vector of random variables of soil parameters and n is the number of random variables; $g(\boldsymbol{x}) = FS(\boldsymbol{x}) - 1.0$ is the limit state function, in which $FS(\boldsymbol{x})$ is the estimated factor of safety (FS) at $\boldsymbol{x}$. The risk assessment of slope failure involves estimating the probability of slope failure as well as the corresponding failure consequence. Similarly, the overall risk of slope failure along multiple failure modes can be expressed as (Zhang and Huang 2016)

$$R = \int \cdots \iint_{g(\boldsymbol{x})<0} C_m(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x} \tag{2}$$

where $C_m(\boldsymbol{x})$ denotes the consequence induced by slope failure. It is practically impossible to evaluate the n-fold integrals in Eqs. (1) and (2) because the complete probabilistic information on the soil properties is often unavailable. Thus, approximation or simulation methods are needed to evaluate this integral. The risk in Eq. (2) can be simplified as the product of the probability of failure and a constant consequence:

$$R = P_f C \tag{3}$$

where C is the failure consequence. This method works well for the system that has a single failure consequence. There are many possible failure modes for a particular slope failure once the spatial variability is considered, obviously, each of which has an individual consequence (Huang et al. 2013). For the quantitative risk assessment of slope failure, the consequences are assessed individually for each potential failure mode as follows:

$$R = \sum_{i=1}^{n_f} P_{f_i} C_i \tag{4}$$

where P_{f_i} and C_i are the probability and consequence of the i-th failure mode, respectively; n_f is the number of slopes that fail during the MCS. In the MCS, P_{f_i} is approximately equal to the occurrence probability of that particular simulation, i.e.,

$$P_{f_i} = 1/N_{sim} \tag{5}$$

where N_{sim} is the total number of simulations. Eq. (4) can thus be rewritten as (Jiang et al. 2022)

$$R = \sum_{i=1}^{n_f} P_{f_i} C_i = \sum_{i=1}^{n_f} \frac{1}{N_{sim}} C_i = \frac{1}{N_{sim}} \sum_{i=1}^{n_f} C_i \tag{6}$$

Eq. (6) can be further rewritten as

$$R = \frac{1}{N_{sim}} \sum_{i=1}^{n_f} C_i = \frac{n_f}{N_{sim}} \frac{\sum_{i=1}^{n_f} C_i}{n_f} = P_f \overline{C} \tag{7}$$

where $\overline{C} = \sum_{i=1}^{n_f} C_i \Big/ n_f$ is the average consequence among the failures. In other words, Eq. (7) states that the risk is equal to the product of the total probability of slope failure and the average (the mean, in the limit) consequence. It is evident that the Eqs. (3) and (7) are consistent, the only difference between them is that the consequence C in Eq. (3) is a constant, while it is the mean consequence $\overline{C}$ among the failures in Eq. (7).

3 Modeling of Geological and Geotechnical Uncertainties

Soils and rocks are formed from a combination of geological, environmental, physical and chemical processes. In geotechnical engineering, there exist geological uncertainties (e.g., soil stratification uncertainty) and geotechnical uncertainties (e.g., inherent spatial variability of geomaterials, transformation and model uncertainties), which can affect the slope stability significantly. To enhance the slope reliability and risk analyses, these uncertainties and corresponding modeling approaches are briefly introduced as follows.

3.1 Soil stratification uncertainty

Geological heterogeneity can be classified into at least two categories (Elkateb et al. 2003), namely, the inherent spatial variability of the same material and the stratigraphic heterogeneity among different materials. To characterize the stratification of the whole site, the strata for every sampled location and a stratigraphic model linking the information between sampled and unsampled locations are both required. Two representative models are available in the literature, as shown in Figure 1. The first is a boundary-based model (e.g., Nobre and Sykes 1992; Zhang and Dasaka 2010; Li et al. 2016c). It assumes a continuous and single-valued boundary between two materials and predicts the boundary depth at unsampled locations. The boundary-based model is conceptually simple and can be easily integrated with engineering judgment; thus, it is widely used in geotechnical practice. Some material boundaries are determined with high confidence; others, with high uncertainty. This uncertainty, known as stratigraphic uncertainty, is highly related to the horizontal scale of fluctuation. The second stratigraphic model is a category-based model (e.g., Elfeki and Dekking 2001; Li et al. 2016d; Qi et al. 2016). In this model, the material categories at unsampled locations, which are finite and discrete, need to be predicted. Without the limit of material boundaries, this model is able to generate more complicated stratification, particularly in the form of one material embedded in another. The related stratification uncertainty can be simulated using several conditional simulation techniques and quantified in terms of information entropy. Xiao et al. (2017) proposed a heuristic combination model to integrate a boundary-based model and a category-based model in a consistent framework. It has the ability not only to generate almost arbitrary geotechnical strata but also to take into account the material spatial distribution trend and engineering judgment to a certain degree.

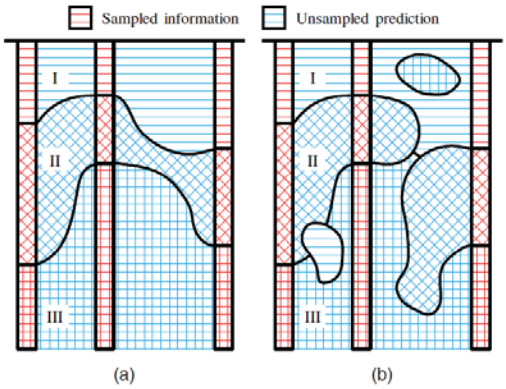


Figure 1. Two models for geological stratification modeling: (a) boundary-based model; (b) category-based model (adapted from Xiao et al. 2017).

3.2 Inherent spatial variability

Because of the natural depositional and postdepositional processes, in situ soil parameters vary spatially even within homogeneous deposits (e.g., Lumb 1966; Li and Lumb 1987; Phoon and Kulhawy 1999a), as shown in Figure 2. The inherent spatial variability of soil properties has received considerable attention in the reliability and risk analyses of slopes. The random field theory is suitable for modeling the spatial variability of soil properties in statistically homogeneous soil layers (Vanmarcke 1977, 2010). The Coefficient of Variation (COV) of this spatial average can be much smaller than that of the soil property at a point. This uncertainty reduction can be calculated analytically using a variance reduction function. The variance reduction function is dependent on a key random field parameter called scale of fluctuation, which can be regarded as a characteristic length parameter that unifies various common autocorrelation models (Li and Lumb 1987). To determine a variance reduction factor from the point to the local average of soil properties, the autocorrelation function and scales of fluctuation should be accurately estimated based on hypothetical or statistical considerations (e.g., Vanmarcke 1977; Cami et al. 2020; Fei et al. 2022). However, it is not a trivial task because the available field data is generally Multivariate, Uncertain and Unique, Sparse, Incomplete, and potentially Corrupted with "3X" denoting 3D spatial/temporal variability (MUSIC-3X) as stated by Phoon et al. (2019) and Phoon and Ching (2021).

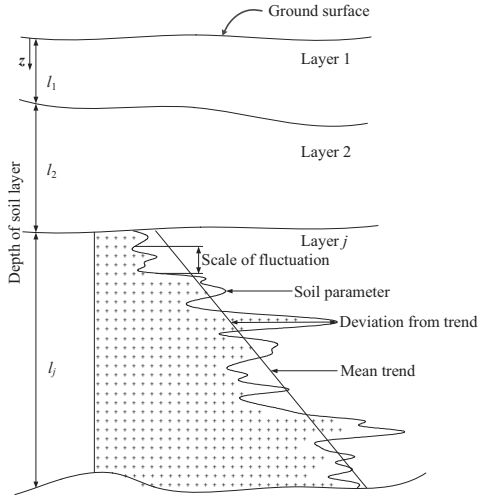


Figure 2. Inherent spatial variability of a typical soil parameter (adapted from Phoon and Kulhawy 1999a).

Additionally, the literature on modeling of the spatial variability of soil parameters is generally founded on unconditional random field theory wherein the site-specific data are not incorporated. Various unconditional random field simulation methods have been developed to directly model the spatial variability of soil properties, including the covariance matrix decomposition-based midpoint method (e.g., Der Kiureghian and Ke 1988; Li et al. 2015), local average subdivision (Fenton and Vanmarcke, 1990), series expansion method, including Karhunen-Loève Expansion (KLE) (e.g., Phoon et al. 2002; Jiang et al. 2014, 2015), Expansion Optimal Linear Estimation (EOLE), spectral representation method and Bayesian compressive sensing-based random field generator (Zhao and Wang, 2020). However, numerous field data collected from in situ geotechnical tests, including Vane Shear Test (VST), Cone Penetration Test (CPT) and Standard Penetration Test (SPT), have confirmed that the soil parameters exhibit an obvious depth-dependent nature (i.e., non-stationary characteristics). For example, the saturated hydraulic conductivity decreases with the depth, while the undrained shear strength (S_u) and cohesion (c) typically increase with the depth, and the friction angle (φ) can decrease or increase with the depth (Jiang et al. 2022). The non-stationary characteristics of soil properties should be depicted in the slope reliability and risk analyses.

3.3 Transformation uncertainty

Estimating the pertinent soil parameters from in situ geotechnical test is one of the most important tasks in slope stability and design, particularly the values governing the slope stability at a limit state. For the indirect data that is obtained from the CPT, SPT and drilling method, a transformation model is needed to transform the indirect measurement (e.g., tip resistance from a CPT test) to a concerned soil property (e.g., the undrained shear strength) (Phoon and Kulhawy 1999b). Although many correlation models have been developed to obtain an estimate of a soil parameter pertinent to design in practice using more commonly available data (e.g., Mayne et al. 2001), the correlations were generally developed by curve fitting based on laboratory or field data. Therefore, they tend to be case-specific and may not generalize to other or new soils/sites (Phoon et al. 2022). In using these correlations,

the caveat is to apply engineering judgment. To this end, bivariate correlation and multivariate correlation models are developed for the characterization of soil transformation uncertainty (e.g., Ching et al. 2016). Recently, Machine learning also has been adopted to depict the soil correlations to extract useful knowledge from big data (Zhang et al. 2020).

3.4 Model uncertainty

Due to inevitable assumptions and simplifications, a model cannot perfectly predict actual responses of a geotechnical structure (e.g., Tang and Phoon 2021). In other words, the model predictions often deviate from the measurements. The deviation between measured and predicted responses is called model uncertainty, which has a strong influence on the probability of failure and thus on the estimation of safety margin (e.g., Christian et al. 1994; Gilbert and Tang 1995; Lacasse and Nadim 1996). In geotechnical practice, the model uncertainty can be characterized in a relatively straightforward way by using a model factor (ISO 2015). The model factor itself is not constant but takes a range of values that may depend on the scenarios covered in the dataset used for evaluation. Unfortunately, it is not an easy task to establish the model factor, as the field data usually cover a limited range of influential parameters. An alternative way is to characterize the model factor as a random variable (e.g., Tang and Phoon 2021). Because of the model uncertainties of slope stability analysis methods (e.g., limit equilibrium method, FEM), the actual factor of safety (y) is commonly evaluated as (e.g., Zhang et al. 2009)

$$y = FS(\boldsymbol{x}) + \varepsilon \tag{8}$$

where ε denotes the model factor associated with the slope stability analysis. For example, the model factor of slope stability analysis using simplified Bishop method can be modeled as a normal distribution with mean μ_ε = 0.05 and standard deviation σ_ε = 0.07 (Zhang et al. 2009).

4 Deterministic Slope Stability and Mobility Analyses

Slope failures or landslides are often a complicated dynamic process with multiple stages, starting from triggering (e.g., rainfall), initiation, post-failure large deformation of soils, and final deposition. The first two-stages determine whether a landslide is initating, while the latter two-stages often control landslide consequences. In geotechnical practice, deterministic slope numerical analysis methods can be divided into slope stability analysis and slope mobility analysis, as shown by Figure 3. The former focuses on the first two stages by calculating a *FS* for representing a slope's stability performance, and it provides no information of slope movements, while the latter is able to simulate the post-failure large deformation of soils dynamically. For example, Material Point Method (MPM) has been used as a popular tool for simulating the large deformation of soils in recent years, and it bypasses the grid distortion problem encountered in the FEM through discretizing a domain into both grids and particles.

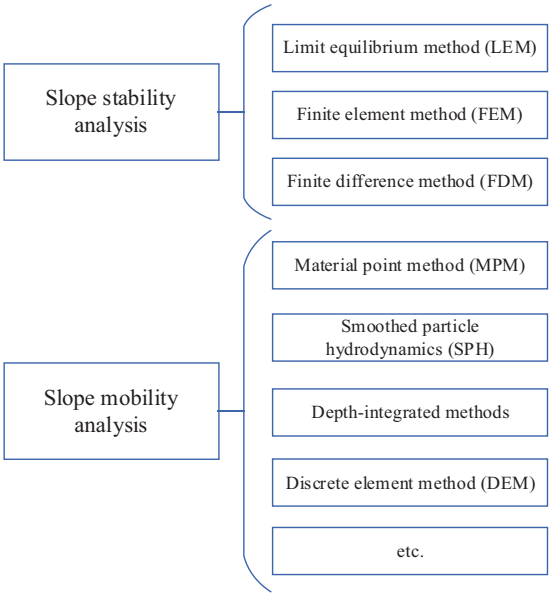


Figure 3. Commonly used methods for slope stability and mobility analyses.

4.1 Slope failure modes and post-failure behaviors

As shown in Figure 4, slope failures often exhibit complex failure modes, depending on specific geological and other site conditions. For a given slope, proper modeling of its slope failure modes is a prerequisite for slope stabiltiy analysis. Inclinometers can be buried in a slope to measure displacement along the depth and identify the location of failure surface. In the slope stabiltiy analysis using Limit Equilibrium Method (LEM), a slope failure mode is represented by a potential slip surface, which may be circular, planar, spiral, or arbitrary. When the location of failure slip surface is unknown, the most critical slip surface corresponding to the minimum *FS* is commonly identified by searching from a large number of possible slip surfaces. In contrast, the FEM has an advantage of automatically finding the most critical slip surface of arbitrary shape based on equilibrium conditions of stress-strain of soil and external forces (e.g., Cheng, 2007).

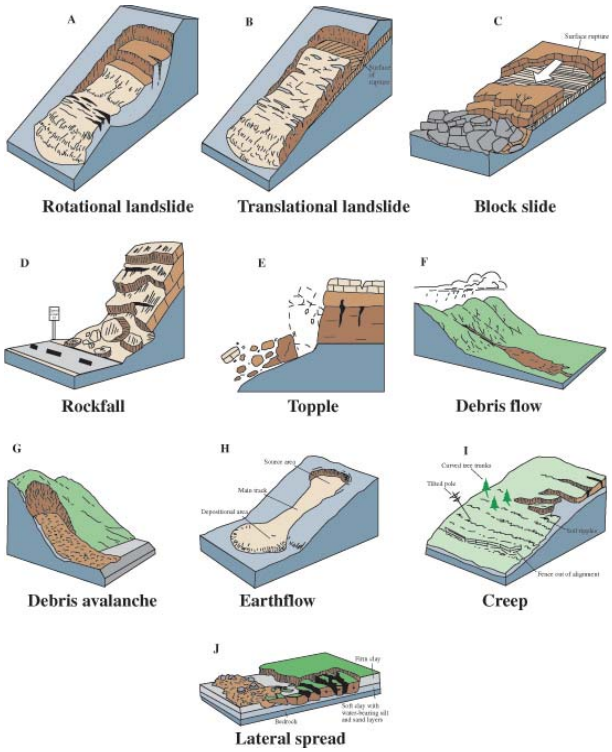


Figure 4. Illustration of typical slope failure modes (adapted from Varnes 1978).

A slope failure mode can also be characerized by its movement features, such as sliding volume, velocity, runout distance (or travel distance), slding depth, travel angle, etc (Liu et al. 2019). Figure 5 shows a typical slope failure mode obtained from slope mobility analysis. Observations of historical landslides show these quantitative features vary greatly at different locations and for different site conditions. For example, the sliding volume of rainfall-induced landslides occurred in Hong Kong often ranges from a few to tens of thousands of cubic meters (Kwong et al. 2004). The catastrophic Shenzhen landslide occurred on Decembr 20, 2015 that destroyed 33 houses and caused 77 deaths had a sliding volume of 2.73×10^6 m^3, a travel distance of 1100m, and a velocity estimated about 30 m/s (Yin et al. 2016). In some high mountains (e.g., Alpine, Himalayas), the sliding volume of rock avalanches can be several billions of cubic meters, and the velocity can be as high as 100 m/s (e.g., Weidinger 2006). Generally speaking, the larger the sliding volume or runout distance, the higher destructive power the debris flow contains, leading to a more hazadous consequences. These features are of great concern to risk assessment of landslides.

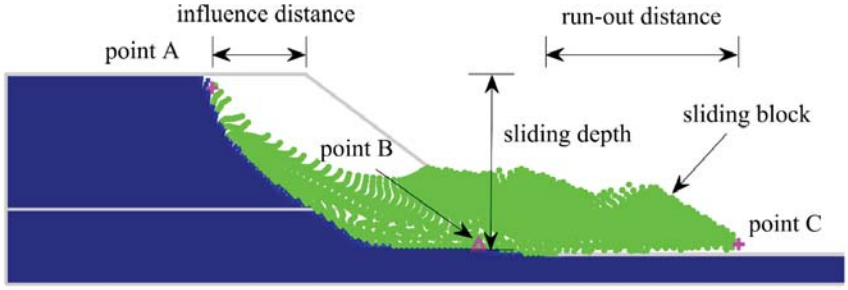


Figure 5. Illustration of a typical slope failure mode from slope mobility analysis (adopted from Liu et al. 2019).

4.2 Slope stability analysis methods

The LEM has been widely applied in geotechnical practice for deterministic slope stability analysis. The core of LEM is to develop the forces and/or momentum equilibrium conditions for a potential slip surface. The slip surface is often discretized into a number of soil slices for computational convenience. The LEM includes a series of methods of slices, such as ordinary method of slices, simplified Bishop method, Janbu's method, Spencer's method, etc. The difference of these methods lies in assumptions used for developing the equilibrium conditions. For example, Spencer's method satisfies horizontal force, vertical force and momentum equilibrium conditions, and it is deemed as a rigorous LEM. One difficulty in LEM is to search the most critical slip surface that corresponds to the minimum *FS*, especially when non-circular slip surfaces are considered. It often requires using an optimization algorithm and high computational efforts to do the search (e.g., Cheng et al. 2007).

The FEM, finite difference method (FDM) or Limit Analysis Method (LAM) provide a strict and universal tool for the deterministic modeling of slope stability because they can seek out the weakest failure path for the slopes in spatially variable soils or including weak seams without a need to assume the location and shape of the sliding surface a priori (Griffiths and Lane, 1999). Shear Strength Reduction (SSR) technique-based FEM or FDM are widely-used methods for slope stability analysis (e.g., Griffiths and Lane 1999). The use of the SSR technique within the FEM or FDM frameworks can effectively capture the development of realistic compound failure surface and allow failure to occur by progressively reducing the strengths of the materials until the equilibrium in the system is disturbed. As a result, a more accurate estimation of the probability and risk of slope failure can be obtained in the subsequent analyses. This technique has been incorporated in some commercial software, such as PLAXIS, FLAC3D and ABAQUS. The bisection method, due to its improved efficiency, is usually used to search the SSR factor as follows:

$$\begin{cases} \tan\varphi_f = \dfrac{\tan\varphi}{SR} \\ c_f = \dfrac{c}{SR} \end{cases} \tag{9}$$

where *SR* is the SSR factor; φ and c are the soil friction angle and cohesion, respectively; φ_f and c_f are the reduced friction angle and cohesion, respectively. When the equilibrium is disturbed, the reduction factor *SR* is deemed as the *FS* of slope.

4.3 Slope mobility analysis methods

Traditional numerical methods such as the FEM, FDM and LAM are not suitable for simulating the slope mobility since they will suffer from severe mesh distortion once the large soil deformation occurs during the slope failure. The slope mobility analysis (i.e., post-failure behavior analysis) is quite crucial to the assessments of influence zone and runout distance in the risk assessments of slope failure (e.g., Wang et al. 2016; Zhou and Sun 2020). In particular, the MPM originated from the fluid mechanics called particle-in-cell, and was first applied to tackle solid mechanical problems by Sulsky et al. (1994). The basic idea of MPM is to discretize a continuum into both a set of material points and a set of background meshes, in which the particles carry state variables and material properties, and the mesh is employed to solve governing equations. Other methods such as the Smoothed Particle Hydrodynamics (SPH) method can also be used to simulate the slope mobility problems. The SPH method is a popular Lagrangian meshfree method capable of modeling the large deformations of geomaterials associated with slope movements. The SPH method that was first developed in the astronomy (Gingold and Monaghan 1977) has also been successfully applied to simulate the landslides, river levees, embankments and other geotechnical problems (e.g., Li et al. 2019b; Mori et al. 2020).

5 Slope Reliability and Risk Analysis

The direct MCS, although being robust and versatile, suffers from a lack of efficiency due to the large number of realizations needed to estimate sufficiently accurate probability and risk of slope failure in spatially variable soils. To alleviate the computational burden, many strategies, approximation methods such as first-order second-moment, first-order reliability method and probability density evolution method have been applied for efficient slope reliability analysis considering the spatial variation of soil properties (e.g., El-Ramly et al. 2002; Ji et al. 2012; Low 2015). However, the approximation methods may suffer from the curse of dimensionality when thousands of random variables are involved to model the soil spatial variability. Additionally, the approximation methods are inaccurate for slope reliability problems involving a compound failure surface (e.g., Papaioannou and Kiureghian 2010; Ji et al. 2012). To this end, the surrogate model-based response surface methods (including PCE- and CNN-based response surface methods) and advanced computer-based simulation methods are

proposed to provide efficient strategies for dealing with the reliability and risk problems of high dimensionality, small probabilities, nonlinear limit state functions and multiple failure modes.

5.1 PCE-based response surface method

It is well recognized that the output responses of slopes including the factor of safety, pore water pressure and displacement cannot be explicitly expressed as the functions of uncertain input parameters because analytical solutions of these output responses typically do not exist. One has to resort to numerical methods to evaluate these responses. However, the slope reliability analyses usually require computing these responses for many sets of inputs, which are quite time-consuming. To reduce the computational cost of the direct MCS, Hermite Polynomial Chaos Expansion (PCE) is often employed to construct a response surface of factor of safety, *FS*, for each key failure mode. Using the Hermite PCE, the factor of safety for a given key failure mode can be calculated as (e.g., Jiang et al. 2014, 2015; Li et al. 2016a)

$$FS_{j_r}\left(\boldsymbol{\xi}\right)=a_0\Gamma_0+\sum_{i_1=1}^{N}a_{i_1}\Gamma_1\left(\xi_{i_1}\right)+\sum_{i_1=1}^{N}\sum_{i_2=1}^{i_1}a_{i_1,i_2}\Gamma_2\left(\xi_{i_1},\xi_{i_2}\right)+\sum_{i_1=1}^{N}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2}a_{i_1,i_2,i_3}\Gamma_3\left(\xi_{i_1},\xi_{i_2},\xi_{i_3}\right)+\cdots \tag{10}$$

in which j_r = 1, 2, …, N_r, in which N_r is the number of key failure modes; $n = M\times N_F$ is the number of random variables in standard normal space, where N_F is number of random fields involved in slope reliability analysis and M is the number of the truncated KLE terms; $a_0, a_{i_1}, a_{i_1,i_2}, a_{i_1,i_2,i_3}\cdots$ are the unknown coefficients; $\Gamma_{j_p}(\cdot)$, j_p = 1, 2, 3, $\cdots$ are Hermit polynomials with j_p degrees of freedom; $\boldsymbol{\xi}=\left(\xi_1,\xi_2,\ldots,\xi_n\right)$ are a set of independent standard normal random variables corresponding to those used to discretize the random fields using KLE. From a physical point of view, the response surface given by Eq. (10) is a surrogate of the slope stability analysis model that involves uncertain input parameters.

For the n_{PCE}-th order Hermite PCE, there are a total of $(n+n_{PCE})!/(n!\times n_{PCE}!)$ unknown coefficients (i.e., $a_0, a_{i_1}, a_{i_1,i_2}, a_{i_1,i_2,i_3}\cdots$) in Eq. (10), which are required to be determined for construction of the response surfaces. N_p realizations of the random fields of soil parameters and their corresponding N_p factors of safety are needed to determine the unknown coefficients for each failure mode. The N_p realizations of random fields can be generated by using the LHS-based KLE. Based on the N_p random samples and their corresponding factors of safety for a given key failure mode, N_p linear equations are obtained using Eq. (10). Then, the unknown coefficients are determined by solving these N_p linear equations. After that, the response surface for the j_r-th key failure mode concerned is obtained. The procedure described above is repeated for the N_r failure modes to obtain their respective response surfaces. Finally, the N_r response surfaces are, collectively, used as a surrogate of the deterministic slope stability analysis to, explicitly and efficiently, evaluate the factor of safety, *FS*, for each random sample. The MCS, LHS, SS or other probabilistic methods are then employed to estimate the probability of slope failure:

$$P_f=\frac{1}{N_{sim}}\sum_{i=1}^{N_{sim}}I[FS(\boldsymbol{x}_i)<1.0] \tag{11}$$

where N_{sim} is the total number of samples; $I(.)$ is the indicator function. For a given random sample, the indicator function is taken as the value of 1 when $FS < 1.0$. Otherwise, it is equal to zero; $\boldsymbol{x}_i$ is the i-th realization of random fields. In this way, the computational cost used for each random sample is substantially reduced.

5.2 CNN-based response surface method

Convolutional Neural Network (CNN) is a particular deep-learning architecture that adopts the operations such as convolution and pooling to process and interpret high-dimensional input data. A typical CNN architecture consists of the input layer, convolutional layer, activation layer, pooling layer, fully-connected layer, dropout layer and output layer, as shown in Figure 6. The CNN-based response surface method has been gradually applied in the reliability analysis of spatially variable slopes due to its good generalization ability (e.g., Wang and Goh 2021; Ji et al. 2022).

The configuration of the CNN input layer for handling random fields is somewhat different from the conventional configuration used to process digital images. For the conventional configuration, a digital image is made of pixels. As explained by Wang and Goh (2021) and Ji et al. (2022), in finite-element analyses, the values of material parameters are assigned to the Gauss integration points within each finite element. When using the CNNs to process the random fields, the Gauss integration points in the discretized mesh are the counterparts of the pixels in conventional image processing. The random field of a spatially variable property of interest corresponds to the “channel”. The magnitude of this spatially variable property at a Gauss integration point is analogous to the “channel intensity” of a “pixel”. Once the CNN model is well trained and validated, the MCS,

LHS, SS or other probabilistic methods are also employed to efficiently estimate the probability of slope failure using Eq. (11). Note that the above surrogate model-based response surface methods are both implemented in a non-intrusive manner because the deterministic numerical analysis and probabilistic analysis of slope stability are deliberately decoupled, which thus can provide a practical and effective tool for tackling the reliability problems involving complex numerical analyses.

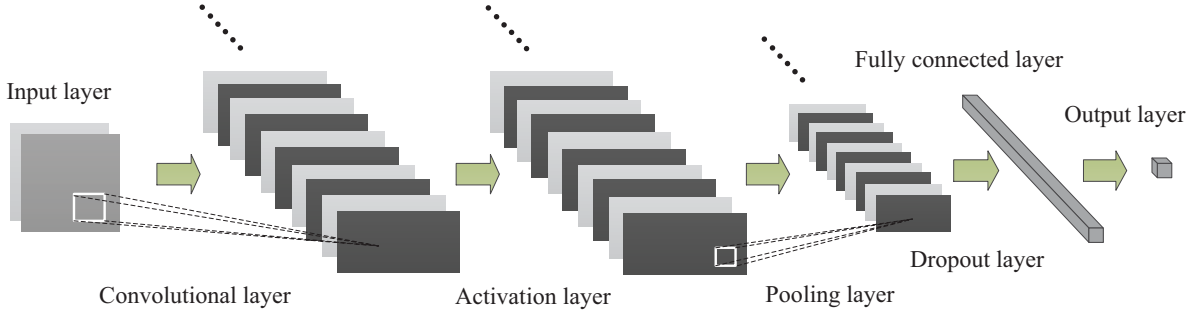


Figure 6. A typical architecture of the CNN model.

5.3 Subset simulation and response conditioning method

To improve the computational efficiency of MCS for small probability problems and inherit its powerful ability, several advanced computer-based simulation techniques have been developed, among which the SS (e.g., Au and Beck 2001; Li et al. 2016b,e) is the most widely-used technique in geotechnical applications. The SS stems from the idea that a small probability can be expressed as a product of larger conditional probabilities of some intermediate failure events, thereby converting a rare event into a sequence of more frequent ones as follows:

$$P_f = P(F_m) = P(F_1)\prod_{k=2}^{m} P(F_k|F_{k-1}) \tag{12}$$

where $F_k = \{g(\boldsymbol{x}) < g_k, k = 1, 2, \ldots, m\}$ are a set of intermediate failure events defined by a decreasing sequence of intermediate threshold values $g_1 > g_2 > \ldots > g_m = 0$, respectively; $P(F_1) = P[g(\boldsymbol{x}) < g_1]$ and $P(F_k|F_{k-1}) = P[g(\boldsymbol{x}) < g_k|g(\boldsymbol{x}) < g_{k-1}]$, $k = 2, 3, \ldots, m$. During the subset simulation, the sample space is divided into m+1 mutually exclusive and collectively exhaustive subsets Ω_k, $k = 0, 1, \ldots, m$, by the m intermediate threshold values, where $\Omega_0 = \{g(\boldsymbol{x}) \geq g_1\}$, $\Omega_k = \{g_{k+1} \leq g(\boldsymbol{x}) < g_k\}$ for $k = 1, 2, \ldots, m-1$, and $\Omega_m = \{g(\boldsymbol{x}) < g_m\}$. To implement the subset simulation, $g_1, g_2, \ldots, g_{m-1}$ can be determined adaptively so that the sample estimates of $P(F_1)$ and $P(F_k|F_{k-1})$, $k = 2, 3, \ldots, m-1$, always correspond to a common specified value of a conditional probability p_0 and $P(F_m|F_{m-1})$ can be estimated by the rate of failure samples ($g_m = 0$) in the last subset.

Random samples in different subsets have different probability weights, w_k, which is quantified by the occurrence probability of each subset $P(\Omega_k)$ [i.e., $P(\Omega_k) = p_0^k(1-p_0)$ for $k = 0, 1, \ldots, m$-1, and p_0^m for $k = m$] to the sample size N_k in each subset [i.e., $N_k = (1-p_0)N$ for $k = 0, 1, \ldots, m-1$, and N for $k = m$], namely $w_k = P(\Omega_k)/N_k$. This is different to MCS, in which samples have the same weight of $1/N_t$. According to the total probability theorem, Eq. (12) can be rewritten as:

$$P_f = \sum_{k=0}^{m} P(F|\Omega_k)P(\Omega_k) = \sum_{k=0}^{m}\left[\frac{\sum_{j=1}^{N_k} I_k^{(j)}}{N_k}\right]P(\Omega_k) = \sum_{k=0}^{m}\sum_{j=1}^{N_k} I_k^{(j)} w_k^{(j)} \tag{13}$$

where $I_k^{(j)} = I(\boldsymbol{x}_k^{(j)})$ is the indicator of failure of sample $\boldsymbol{x}_k^{(j)}$ [i.e., $I_k^{(j)} = 1$ if $g(\boldsymbol{x}_k^{(j)}) < 0$; $I_k^{(j)} = 0$ otherwise]; $\boldsymbol{x}_k^{(j)}$ is the jth sample falling in Ω_k; and $w_k^{(j)} = w_k$ is the weight of $\boldsymbol{x}_k^{(j)}$.

Despite this, thousands of deterministic analyses are still needed in subset simulation for a problem with the probability of failure being less than 10^{-4}. A more advanced computer-based simulation technique, Response Conditioning Method (RCM) (Au 2007; Xiao et al. 2016), was proposed to introduce an additional but efficient simple-model-based preliminary reliability analysis to improve the computational efficiency of complex-model-based target reliability analysis. The simple model can be a simplified analytical model (Li et al. 2016b), a response surface model or a coarse finite element model (Xiao et al. 2016). The preliminary reliability analysis can be performed using the SS. After that, a part of samples is randomly selected as the representative samples in small sample space, which is referred as the sub-binning strategy (Au 2007). By this way, Ω_k can be further divided into N_s sub-bins Ω_{kj}, $j = 1, 2, \ldots, N_s$, which have the same number of random samples. In each Ω_{kj}, one of N_k/N_s samples is randomly selected as the representative sample to judge whether Ω_{kj} belongs target failure

domain or not, as shown in Figure 7 schematically. Since Ω_{kj}, j = 1, 2, …, N_s, are mutually exclusive and collectively exhaustive sub-bins of Ω_k, the target probability of slope failure, $P_{f,t}$, can be updated as

$$P_{f,t} = \sum_{k=0}^{m} P\left(F_t \middle| \Omega_k\right) P\left(\Omega_k\right) = \sum_{k=0}^{m} \sum_{j=1}^{N_s} P\left(F_t \middle| \Omega_{kj}\right) P\left(\Omega_{kj}\right) = \sum_{k=0}^{m} \sum_{j=1}^{N_s} I_{k,t}^{(j)} w_k^{(j)} \tag{14}$$

where $w_k^{(j)} = P(\Omega_{kj}) = P(\Omega_k)/N_s$ due to the equal division; $P(F_t|\Omega_k)$ and $P(F_t|\Omega_{kj})$ are conditional target probabilities of failure given sampling in Ω_k and Ω_{kj}, respectively; $P(F_t|\Omega_{kj})$ can be estimated by $I_{k,t}^{(j)} = I(\mathbf{x}_k^{(j)}|\Omega_{kj})$, which is the indicator function of slope failure for the representative sample in Ω_{kj} using the complex model. Compared with Eq. (13) that needs $mN(1-p_0)+Np_0$ analyses of complex model, Eq. (14) significantly reduces the computational amount to only $(m+1)N_s$ analyses. For a problem with probability of failure being less than 10^{-4}, only hundreds of deterministic analyses are required; if a high correlation exists between the simple model and complex model, tens of analyses have already been sufficient, similar to the response surface methods. More details can be referred to Xiao et al. (2016).

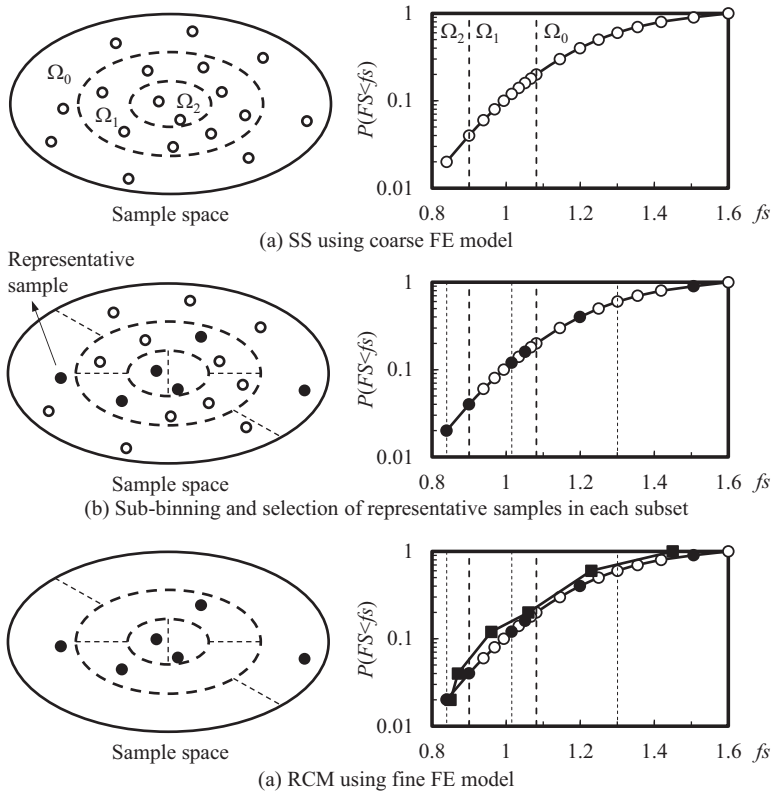


Figure 7. Schematic diagram of SS and RCM (adapted from Xiao et al. 2016).

5.4 Quantitative risk assessment method of slope failure

Risk analysis of slope failure has been developed into the quantitative risk assessment wherein the individual consequence is assessed for each failure mode (Huang et al. 2013). To avoid evaluating the n-fold integral in Eq. (2) for the overall risk assessment, the R in Eq. (2) can be approximately estimated using the direct MCS:

$$R \approx \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} C_m\left(\boldsymbol{x}_i\right) I\left[FS\left(\boldsymbol{x}_i\right) < 1.0\right] = \frac{1}{N_{sim}} \sum_{j_r=1}^{N_r} C_m^{j_r} n_f^{j_r} \tag{15}$$

where N_{sim} is the number of MCS samples for estimating the R; $C_m^{j_r}$ is the consequence corresponding to the j_r-th key failure mode, j_r = 1, 2, …, N_r; $n_f^{j_r}$ is the number of the failure samples associated with the j_r-th key failure mode (Jiang et al. 2017a). Additionally, the contribution of each key failure mode to the R can be quantified as

$$w_R^{j_r} = \frac{C_m^{j_r} n_f^{j_r}}{N_{sim} R}, \quad j_r = 1, 2, \mathrm{K}\ , N_r \tag{16}$$

The coefficient of variation of R, COV_R, is proposed to measure the accuracy in the R, which is defined as (e.g., Zhang and Huang 2016; Jiang et al. 2017a)

$$\mathrm{COV}_R = \sqrt{\frac{N_{sim} - n_f + \sum_{j_r=1}^{N_r} n_f^{j_r} \left(\frac{C_m^{j_r}}{R} - 1 \right)^2}{N_{sim}\left(N_{sim} - 1\right)}} \tag{17}$$

where n_f is the total number of failure samples, $n_f = \sum_{j_r=1}^{N_r} n_f^{j_r}$. To yield a good estimate of R, N_{sim} should be selected large enough such that the COV_R is below a common specified value (e.g., 10%).

Based on the probabilities of failure of N_r key failure modes and the correlation coefficients among them, the bimodal bounds of P_f considering the correlations among the slope failure modes can be obtained as follows (Ditlevsen, 1979):

$$P_{f_1} + \sum_{i=2}^{N_r} \max\left(P_{f_i} - \sum_{j=1}^{i-1} P_{f_{ij}}, 0 \right) \le P_f \le \sum_{i=1}^{N_r} P_{f_i} - \sum_{i=2}^{N_r} \max_{j<i} P_{f_{ij}} \tag{18}$$

where P_{f_1} is the probability of slope failure corresponding to the i-th failure mode; $P_{f_{ij}}$ is the joint probability of the i-th and j-th failure modes. The reader is referred to Ditlevsen (1979) for estimation of $P_{f_{ij}}$. By incorporating the consequences associated with different failure modes, the bimodal bounds of R considering the correlations among the slope failure modes can be derived as

$$P_{f_1} C_m^1 + \sum_{i=2}^{N_r} \max\left(P_{f_i} C_m^i - \sum_{j=1}^{i-1} P_{f_{ij}} C_m^{ij}, 0 \right) \le R \le \sum_{i=1}^{N_r} P_{f_i} C_m^i - \sum_{i=2}^{N_r} \max_{j<i} \left(P_{f_{ij}} C_m^{ij} \right) \tag{19}$$

where C_m^{ij} is the consequence corresponding to slope failing along the i-th and j-th slip surfaces simultaneously, which is approximately taken as $\frac{C_m^i + C_m^j}{2}$. The volume (or area in the 2-D case) of the sliding mass, was often taken as an equivalent index to quantify the consequence of slope failure C_m^i. To enable more complete and realistic assessment of the consequence of slope failure, recently, the progressive failure of the slope is modeled using the MPM or SPH so that the true consequence of slope failure, including the runout distance, retrogression distance and sliding volume of landslide can be assessed, which has greatly facilitated the quantitative risk assessment of slope failure.

6 Numerical Examples

6.1 Example 1: stability of a *c*-*φ* slope

A *c*-*φ* slope example is investigated to demonstrate the effectiveness of the proposed surrogate model-based response surface methods (including HPCE- and CNN-based response surface methods). This slope example has been studied by Cho (2010), Jiang et al. (2015, 2017b), Li et al. (2015), Deng et al. (2021) and Zhu et al. (2021). As shown in Figure 8, the slope is 10 m in height and the slope inclination is 45°. To minimize the boundary effects, the right-end model boundary is extended 10 m beyond the slope toe while the left-end model boundary is extended 10 m beyond the slope crest. The bottom of the model is 5 m below the slope toe.

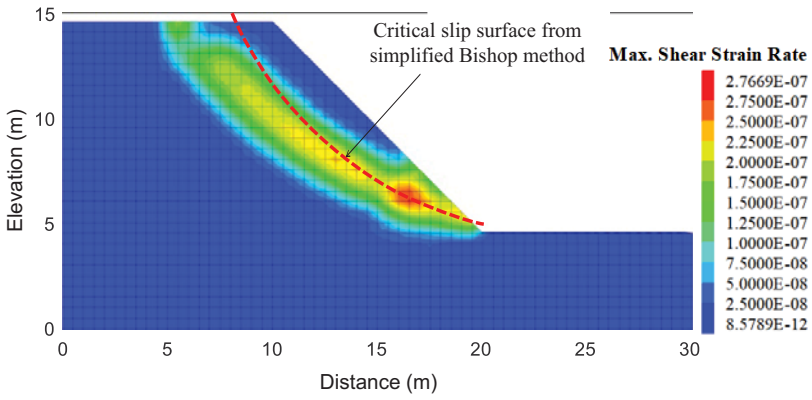


Figure 8. *c*-*φ* slope model and deterministic analysis results.

The model consists of a pool of triangular and square finite-difference elements with an edge length of 0.5 m. There are a total of 1,210 elements and 1,281 modes in the discretized model. The discretized random field elements are consistent with the finite-difference elements. Based on a unit weight of 20 kN/m^3 and a 2-D Gaussian autocorrelation function, the cross-correlated c-φ stationary random fields are generated using the KLE technique. The cohesion c is associated with a mean of 10 kPa and a COV of 0.3 while the friction angle φ is associated with a mean of 30° and a COV of 0.2. The horizontal and vertical scales of fluctuation are, respectively, 40 m and 4 m for both the cohesion and friction angle. In addition, the cross-correlation coefficient $\rho_{c,\varphi}$ between the cohesion and friction angle is taken as -0.7. Based on the KLE and the required maximum error in the random field discretization (i.e., expected energy ratio $\geq 95\%$) (Jiang et al. 2014), $M = 20$ KLE terms are needed for the discretization of each random field. With an elastic Young's modulus of 100 MPa and a Poisson's ration of 0.3, the finite-difference SSR technique in FLAC3D and simplified Bishop method are used, respectively, to calculate the factor of safety underlying a base case with the mean values of cohesion and friction angle. The factor of safety calculated using the SSR technique is 1.219, which is slightly larger than that (1.206) calculated using the simplified Bishop method. As shown in Figure 8, the maximum shear strain contour obtained from the SSR technique match with the most critical slip surface obtained from the simplified Bishop method.

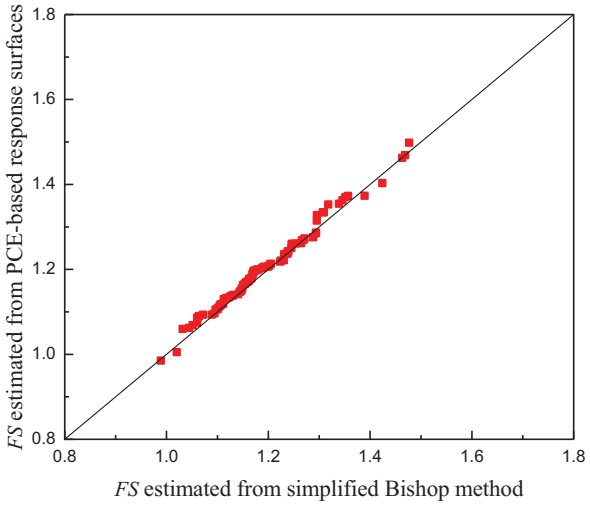


Figure 9. Validation of the PCE-based response surfaces using 100 random samples.

1,000 realizations (i.e., N_p = 1,000) of random fields of c and φ are generated using the LHS-based KLE, and their corresponding critical slip surfaces are determined using the simplified Bishop method and taken as key failure modes, resulting in a total of 79 key failure modes. After the 79 key failure modes are obtained, one response surface is constructed for each failure mode using the second order Hermite PCE accordingly, resulting in 79 response surfaces. The 79 PCE-based response surfaces are then verified by comparing the *FS* of slope stability obtained from the PCE-based response surfaces and the original deterministic analysis of slope stability. Figure 9 shows the *FS* values obtained from the 79 response surfaces using additional 100 sets of random samples versus those obtained from the simplified Bishop method. The *FS* values obtained from the two approaches agree well with each other. This indicates that the PCE-based response surfaces can replace the original deterministic analysis to evaluate the *FS* in this example. Based on these response surfaces, a MCS run with 500,000 random samples is performed to calculate P_f. It should be pointed out that the computational cost used for the MCS run is minimal and negligible once the PCE-based response surfaces are constructed.

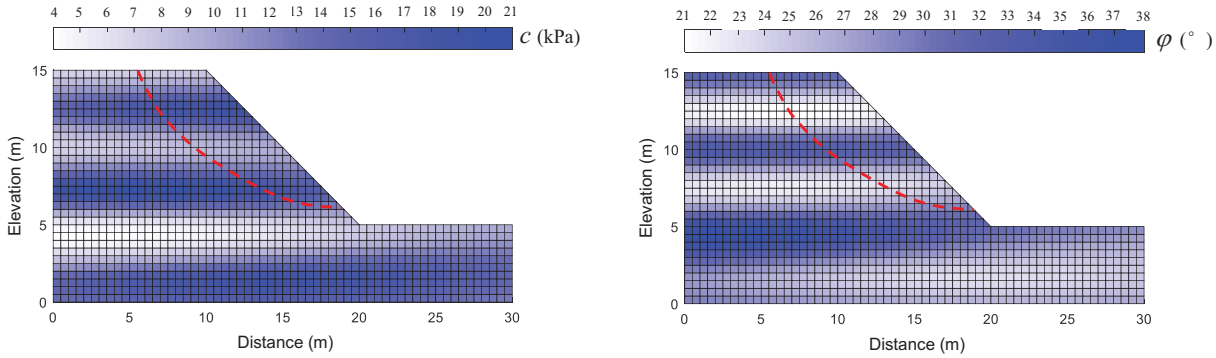
Table 1 Comparison of slope reliability results with different methods ($\rho_{c,\varphi}$ = -0.7).

Reliability method	Slope stability analysis method	No. of simulations	Probability of failure	Reference
Direct MCS	Simplified Bishop method	50,000	3.9×10^{-3}	Cho (2010)
Direct LHS	Simplified Bishop method	10,000	4.4×10^{-3}	Jiang et al. (2015)
PCE+MCS	Simplified Bishop method	500,000	4.9×10^{-3}	This study
CNN+LHS	Finite-difference SSR technique	100,000	1.01×10^{-3}	This study
Direct LHS	Finite-difference SSR technique	10,000	1.8×10^{-3}	This study

To further validate the proposed method, the LHS with 10,000 random samples is conducted to calculate the probability of failure, in which the *FS* is evaluated using the simplified Bishop method with a Gaussian

autocorrelation function used to simulate random fields. As shown in Table 1, the probabilities of failure vary across different studies. The value of P_f obtained from the proposed PCE-based response surface method is 4.9×10^{-3}, which agrees well with the results (i.e., 4.4×10^{-3} and 3.9×10^{-3}) obtained from the direct LHS with 10,000 samples and the direct MCS with 50,000 samples, respectively.

Additionally, a combined dataset that consists of 400 initial samples (realizations of random fields of c and φ) and 6,000 additionally generated samples generated using a data augmentation technique are used to train the CNN model of the factor of safety. Figure 10 presents one typical realization of random fields of c and φ and corresponding slope stability analysis results evaluated using the SSR technique. The light and dark shaded regions indicate areas of small and high shear strength, respectively. Then, 70% of the samples are treated as the training dataset while the remaining 30% of the samples are treated as the validation dataset. Both the training dataset and validation dataset are involved in the training of the CNN model. The input layer consists of images of size of 60×30×2. The first convolutional layer consists of 20 kernels of size of 5×10 and a padding of 0 and a stride of 1. A ReLu layer and a Batch Normalization layer are used to process the information before an average pooling layer is used. Then, the extracted information is fed to a full-connected layer. Based a dropout rate of 50%, which aims to reduce overfitting, a regression between the information in the fully-connected layer and the output layer that contains the factor of safety information is constructed. Hyperparameters used in the training of the CNN model are taken as: learning rate of 0.01, batch size of 256, validation frequency of 10, and the maximum number of iterations of 2,000. An early termination criterion is applied. When the performance of the CNN does not improve for 1,000 consecutive iterations, the training will terminate. In this regard, 1,000 additional samples are generated and calculated using the SSR technique in FLAC3D, which are used as the testing dataset to evaluate the performance of the trained CNN model, as shown in Figure 11. It can be observed from Figure 11 that the data also clusters tightly along the 1:1 line with a determination coefficient R^2 of 0.961. The results effectively confirm that the trained CNN-based response surface is also sufficiently accurate to replace the original deterministic analysis with unseen data.



(a) Cohesion (b) Friction angle

Figure 10. One typical realization of random fields of c and φ and slope stability analysis results (FS = 1.234).

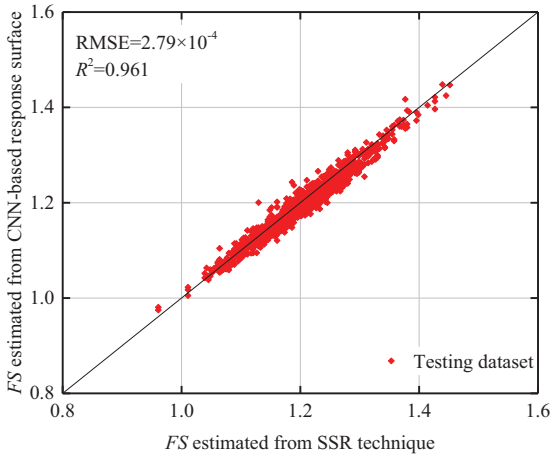


Figure 11. Validation of the CNN-based response surfaces using 1000 random samples.

Based on the CNN model, the probability of failure calculated using the LHS with 100,000 samples is 1.01×10^{-3}, which agrees reasonably well that (1.8×10^{-3}) obtained from the direct LHS with 10,000 samples and SSR technique, as shown in Table 1, which further demonstrates the validity of the trained CNN-based response surface in carrying out practical slope reliability calculations in the presence of high-dimensional random fields. With reference to Table 1, the finite-difference SSR technique yields smaller probabilities of failure compared to the simplified Bishop method. This may be attributed to the fact that the SSR carried out in the finite-difference scheme is a more stable technique than the limit equilibrium analysis using the simplified Bishop method, resulting in larger calculated factors of safety.

In addition, analyses are also carried out for the case with a cross-correlation coefficient of $\rho_{c,\varphi}$ = -0.5 between c and φ, and the reliability results are summarized in Table 2. It is observed that the probabilities of failure obtained from the CNN-based response surface method with the SSR technique compare favourably with those obtained from the PCE-based response surface method with the simplified Bishop method. In addition, the comparison with the results reported by Deng et al. (2021) and Jiang et al. (2017b) indicates that the CNN-response surface method is likely to be more accurate than the SS and the Sliced Inverse Regression-based Multivariate Adaptive Regression Spline (SIR-MARS) technique. Taking the benchmark result (1.60×10^{-2}) obtained using 1,000 direct LHS, the relative errors in the proposed CNN-based response surface method (i.e., 33.8%) is also much smaller than the errors reported by Deng et al. (2021) and Jiang et al. (2017b) (i.e., 63.1% and 50.0%). The results further validate the effectiveness of the proposed CNN-based response surface method.

Table 2 Comparison of slope reliability results with different methods ($\rho_{c,\varphi}$ = -0.5).

Reliability method	Slope stability analysis method	No. of simulations	Probability of failure	Reference
Direct MCS	Simplified Bishop method	50,000	1.71×10^{-2}	Cho (2010)
Direct LHS	Simplified Bishop method	10,000	2.08×10^{-2}	Jiang et al. (2015)
PCE+MCS	Simplified Bishop method	500,000	2.28×10^{-2}	This study
Gaussian Process Regression+MCS	Simplified Bishop method	390	1.81×10^{-2}	Zhu et al. (2021)
SIR-MARS+MCS	Finite-difference SSR technique	120	2.61×10^{-2}	Deng et al. (2021)
SS	Finite-difference SSR technique	950	2.4×10^{-2}	Jiang et al. (2017b)
CNN+LHS	Finite-difference SSR technique	100,000	1.06×10^{-2}	This study
Direct LHS	Finite-difference SSR technique	1,000	1.60×10^{-2}	This study

It is worth highlighting that a single finite-difference-based SSR calculation takes approximately 138 seconds on on a desktop computer with an Intel (R) Core (TM) i5-6500 processor with 3.2 GHz main frequency and a RAM of 8 GB. Therefore, 100,000 simulations would take 159.7 days to complete. In contrast, the generation of the initial 400 samples requires only 55,080 seconds while the additional samples generated through the data augmentation technique only takes approximately 20 seconds. After considering the computational time spent to (i) generate 100,000 random field samples and pre-process the input data (approximately 23 hours), (ii) train the CNN model (approximately 2,880 seconds), and (iii) make prediction for the 100,000 samples (approximately 37 seconds), the total computational cost required by the proposed CNN-based response surface method to carry out the slope reliability analysis is approximately 2 days, which is significantly shorter than the 159.7 days required by the direct LHS.

6.2 Example 2: stability of a three-dimensional undrained slope

This section applies the RCM to evaluate the probability of failure of a three-dimensional (3-D) undrained slope. As shown in Figure 12, the slope has a height (H) of 6 m, a slope angle (α) of about 26.6°, and a length (B) of 100 m. Two finite-element (FE) models are developed in ABAQUS. The FE mesh size measures 2m×2m×5m for the coarse FE model (simple model) (see Figure 13(a)) and 1m×1m×1m for the fine one (complex model) (see Figure 13(b)). For soil property, the elastic-perfectly plastic constitutive model with a Mohr-Coulomb failure criterion is used in both FE analyses.

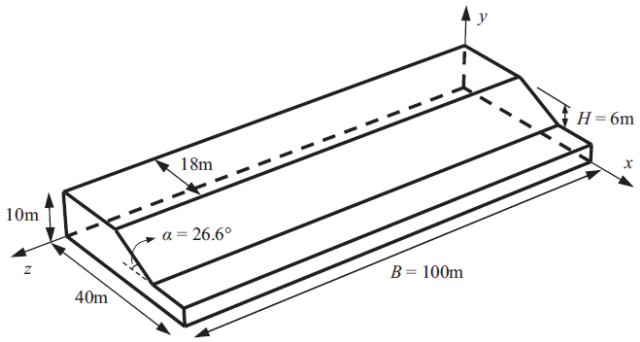
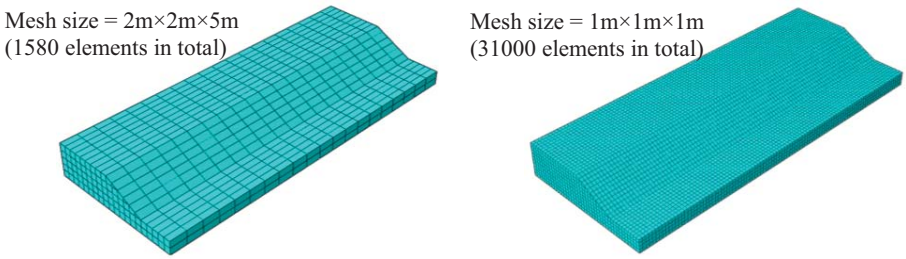


Figure 12. 3-D undrained slope model.



(a) FE mesh for coarse FE model (b) FE mesh for fine FE model

Figure 13. Two FE models for the 3-D undrained slope example: (a) coarse model; (b) fine model.

Undrained shear strength, S_u, is considered to be log normally distributed with a mean of 30 kPa and a COV of 0.3. The spatial variability of S_u is also modeled using the Gaussian autocorrelation function with horizontal and vertical autocorrelation distances of l_h = 20 m and l_v = 2 m, respectively, and generated using the EOLE approach. The deterministic *FS* values calculated by the coarse and fine FE models are 1.651 and 1.593, respectively, and the corresponding computational time is 48 seconds and 35 minutes, respectively. One typical random field realization of the slope is shown in Figure 14. The corresponding *FS* of 3-D slope stability analysis calculated by the fine FE model is 0.741, which implies the slope fails. Its slip surface is nearly spherical with a small sliding mass length (i.e., 24 m) located from 19.5 m to 43.5 m in the axial direction. The 3-D heterogeneous slope considering spatial variability of soil properties models the real slope failure event more realistically than the 3-D homogeneous slope in terms of the shape, location and length of slip surface. A series of cross sections are extracted from the 3-D realization to perform 2-D FE analyses for comparison. Although 2-D analysis could be more conservative than 3-D analysis based on the cross section with minimal 2-D *FS*, the location of the 3-D critical slip surface remains unknown if the 3-D analysis is not performed.

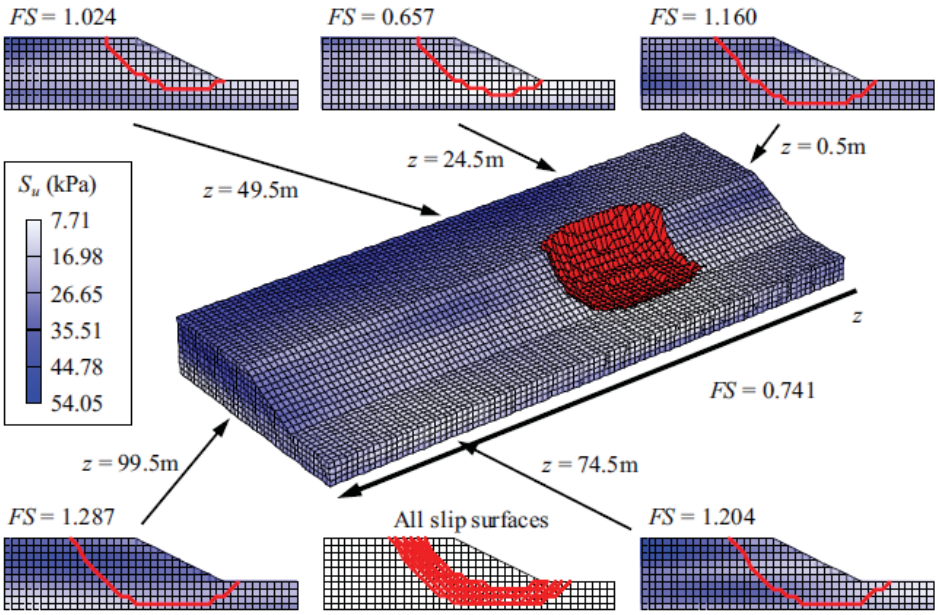
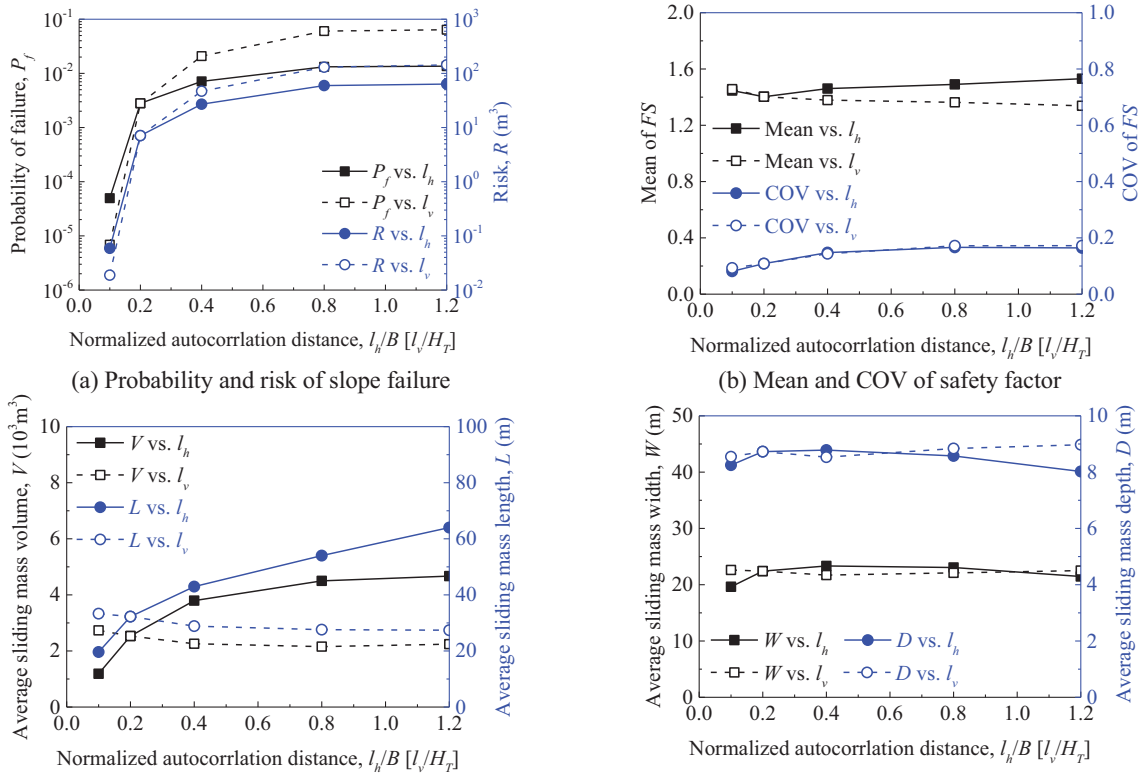


Figure 14. Typical slope stability analysis results in spatially varying soils.

To estimate the P_f for this slope example, $m = 4$, $N = 500$, and $p_0 = 0.1$ are taken in the preliminary analysis using the coarse FE model and $N_s = 25$ is used in the target analysis using the fine FE model. The preliminary analysis gives $P_{f,p} = 8.84\times10^{-4}$ with 1,850 coarse FE analyses and requires about 7 hours by parallel computing, while the target analysis updates $P_{f,t} = 2.80\times10^{-3}$ with the 125 fine FE analyses in about 27 hours using parallel computing. In total, approximate 34 hours (or 1.4 days) is required using the RCM. To validate the results, the direct MCS with 10,000 samples is carried out to calculate the P_f of the considered slope, where the fine FE model is directly used to perform deterministic slope stability analysis. The estimate of P_f is 3.20×10^{-3} and 89.9 days are required to finish the reliability analysis. The RCM provides an unbiased estimation of the probability of failure with significantly reduced computational efforts.

The effects of spatial variability on 3-D slope reliability are explored using the RCM. In addition to the nominal case with l_h = 20 m and l_v = 2m, eight cases with different autocorrelation distances are also considered, including four cases with l_h = [10, 40, 80, 120] m and l_v = 2m and four cases with l_h = 20m and l_v = [1, 4, 8, 12] m. To make a fair comparison, l_h and l_v are normalized by slope length B and nominal height H_T (see Figure 12), respectively. Figure 15 shows the variation of probability of slope failure as a function of normalized autocorrelation distance. When normalized autocorrelation distance increases from 0.1 to 1.2, P_f increases by several orders of magnitude, and the influence weakens when l_h exceeds half of the slope length or l_v exceeds the slope height. Besides, the vertical spatial variability has a greater impact on P_f than the horizontal spatial variability. With respect to slope failure mechanisms, the variation of average volume, length, width and depth of sliding mass are shown in Figures 15(c) and (d). Apparently, the horizontal spatial variability and vertical spatial variability have opposite influences on average sliding mass volume and length in this example. Both increase as the normalized horizontal autocorrelation distance increases, and slightly decrease as the normalized vertical autocorrelation distance increases. In addition, since average sliding mass width and depth are almost unchanged as normalized autocorrelation distance varies in this example (see Figure 15(d)), the variation of sliding mass volume is dominated by the variation of sliding mass length. On the one hand, this indicates that the horizontal spatial variability in the axial direction, instead of that in the lateral direction, affects 3-D slope failure mechanism and average failure consequence. On the other hand, it also indicates that the horizontal spatial variability has a greater impact on slope failure mechanism than the vertical spatial variability. At least in this example, the vertical spatial variability has limited influence on all characteristics (i.e., volume, length, width and depth) of sliding mass. The location of sliding mass is dominated by the horizontal spatial variability as well.



(a) Probability and risk of slope failure

(b) Mean and COV of safety factor

(c) Average sliding mass volume and length

(d) Average sliding mass width and depth

Figure 15. Variations of probaility of failure, *FS*, and sliding mass characteristics with autocorrelation distances.

6.3 Example 3: large deformation of a two-layer undrained slope

A two-layer undrained slope example from previous studies (e.g., Li et al. 2016b,e; Liu, et al. 2019) is adopted to illustrate the proposed landslide risk assessment method based on large deformation of soil simulated from the MPM, as shown by Figure 16. The spatial variability of undrained shear strengths of two layers is considered and modelled by two lognormal random fields. The soil properties and statistics can be found in Liu et al. (2019). The Random Limit Equilibrium-Material Point Method (RLE-MPM) proposed by Liu et al. (2019) is adopted to simulate the landslides with large deformation of soils. For the MPM analysis, a Drucker-Prager soil constitutive model with straining softening is used to model the soil strength. The residual soil strength is taken as 50% of the undrained shear strength. A total of 40,000 random field samples are generated using the covariance matrix decomposition-based midpoint method. Based on the RLE-MPM method, the 40,000 random field samples are repeatedly used to perform RLE analyses, and only 1,512 field samples are used for random material point analyses. Slope failure is defined as when the maximum displacement exceeding a threshold (i.e., 1m) based on the histogram of maximum displacement (Liu et al. 2019). A total of 520 failure samples are obtained, while the remaining samples show a stable slope. These failure samples are of great concern for landslide risk assessment.

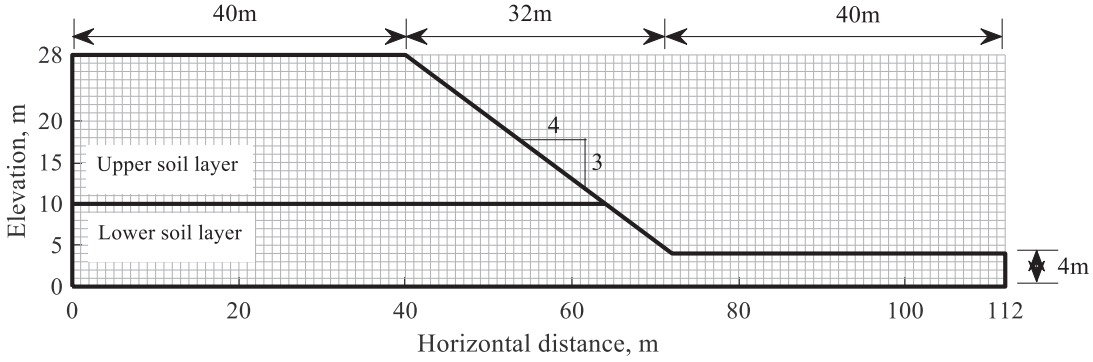
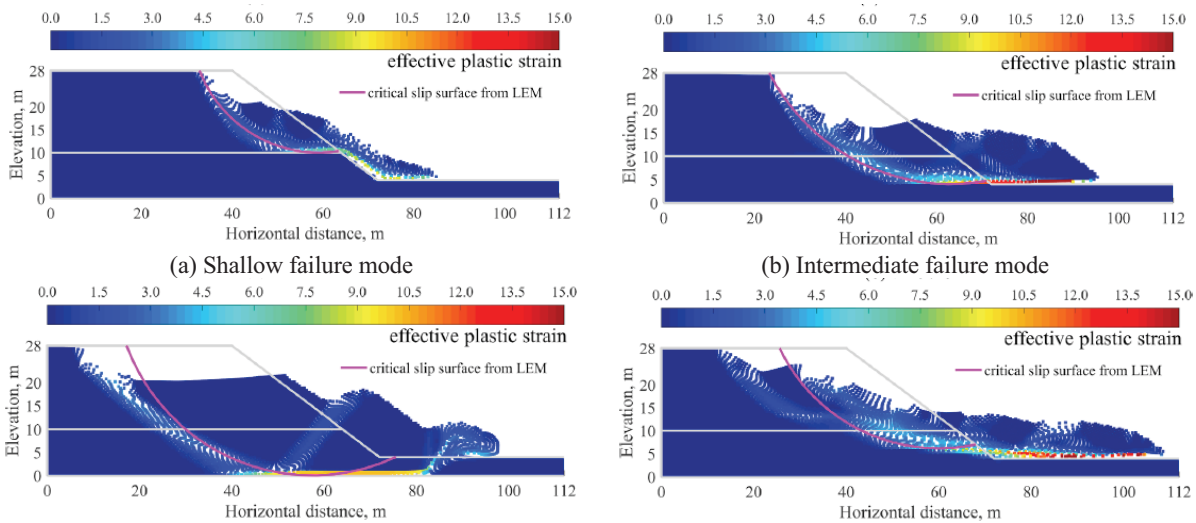


Figure 16. A two-layer undrained slope model (after Li et al. 2016b,e).



(a) Shallow failure mode

(b) Intermediate failure mode

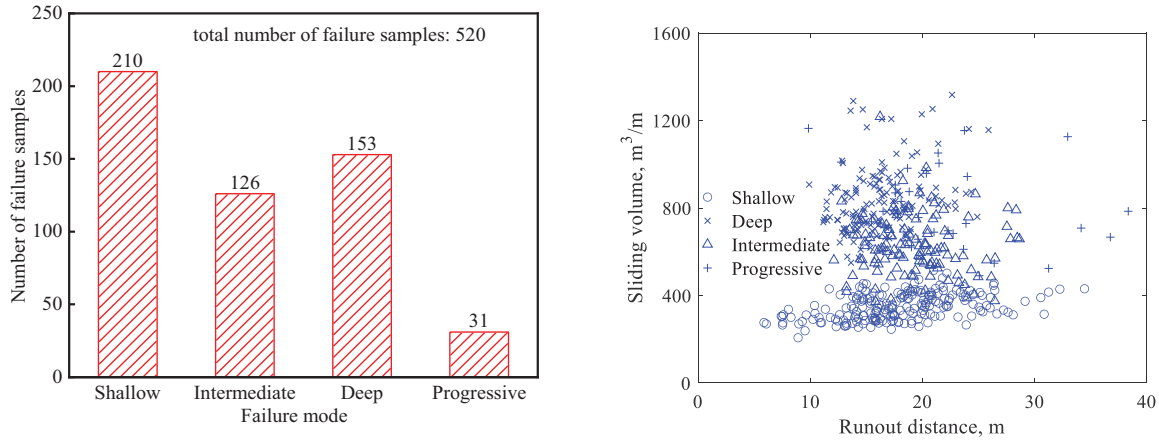
(c) Deep failure mode

(d) Progressive failure mode

Figure 17. Four typical slope failure modes obtained from RLE-MPM (adopted from Liu et al. 2019).

Figures 17(a)~(d) show four typical slope failure modes in spatially variable soils obtained from the RLE-MPM, respectively. In Figure 17(a), only the upper soil layer slides, and it is categorized as shallow failure mode. Figures 17(b) and (c) show an intermediate and a deep failure mode based on the sliding depth, respectivley. In addition, Figure 17(d) shows a progressive failure mode, which contains a shallow and an intermediate failure modes that occurs successively. The results indicate that multiple types of slope failure modes may occur in spatially variable soils, and they are significantly affected by the spatial variability of soil properties.

Figure 18(a) shows the number of failure samples corresponding to these four failure modes. The number of samples for shallow failure mode is 210, the highest among four failure modes. Then, there are 153 and 126 failure samples respectivley for deep and intermediate slope failure modes. Progressive failure mode accounts for only 31 samples among 520 failure samples. Dividing by the total number of samples (i.e., 40,000), the probabilties of failure underlying the shallow, intermediate, deep, and progressive failure modes can be calcualted. In addition, quantitative features are characterized from the simulated slope failure modes. Figure 18(b) shows the runout distance and sliding volume of failure samples for these four failure modes. Shallow failure modes generally have a small sliding volume and a wide distribution of runout distance. In contrast, deep failure modes often correspond to a large sliding volume. When compared with shallow and deep failure modes, the sliding volume and runout distance are intermediate for the intermediate failure mode. Progressive failure mode may have the largest runout distance and sliding volume simultaneously.



(a) Number of samples of slope failure modes (b) Runout distance vs sliding volume

Figure 18. Quantitative features of slope failure modes.

The abovementioned slope failure modes and their quantitative features can provide a reasonable basis for quantitative risk assessment of landslides, including landslide probability assessment, identification of the elements at risk, and vulnerability assessment. Suppose that there is a multi-story building located 20 m away from the slope toe, and there are 60 residents in the building. Their lives are at risk of the landslide when the landsldie occurs with the large deformation of soils and destructive power. The quantified risk of lives may be calculated as the product of probability, consequences of slope failure and life vulnerability. The life vulnerability for a resident living in the buidling represents the degree of loss of a elemnt at risk affected by a landslide (Fell et al. 2008). The vulnerability can be affected by many factors, such as sliding volume, runout distance, time (e.g., day or night), sliding velocity, emergency skills (Corominas et al. 2014). In this study, the vulnerability of a resident living in the building is empirically assessed based on the sliding volume and runout distance obtained from numerical simualtions and recommended values from previous studies (e.g., Finlay, 1996; Dai et al. 2002), as shown in Table 3. The vulnerability increases as sliding volume or runout distance increases. According to Hong Kong recommended values for death from landslide debris (Dai et al. 2002), vulnerability is taken as 1.0 if the building is inundated with debris and the person is buried, or the vulnerability is taken as 0.2 if if the building is inundated with debris and the person is not buried. For this example, the maximum value of vulnerability is taken as 0.6 based on the average of 1.0 and 0.2. In addition, the minimum vulnerability is taken as 0.01 considering that there may be an accidental rockfall during the landslide.

Table 3 Vulnerability matrix of a resident living in the building near to the slope (adapted from Finlay 1996; Dai et al. 2002).

Vulnerability of a person		Runout distance, m		
		<20	20~30	30~40
Sliding volume, m^3/m	<400	0.01	0.15	0.30
	400~800	0.10	0.25	0.40
	800~1200	0.15	0.35	0.50
	1200~1600	0.20	0.45	0.60

Based on the abovementioned results and vulnerability matrix in Table 3, Table 4 shows quantitative risk assessment results considering different failure modes. The risk for each failure sample is calculated as the

product of probability of failure (i.e., 1/40,000), vulnerability (see Table 1), and the consequences denoted by total lives (i.e., 60). These risks are then aggregated to produce the overall risk and respevtivley risks for these four failure modes. As indicated in Table 5, although shallow failure mode is most likely to occur, its risk is smaller than intermediate and deep failure modes, as the shallow failure has less hazardous consequences than the latter two. Progressive failure mode has the least risk, mainly due to its smallest probability of failure. In this case, deep failure mode has the highest risk and it contributes to 33.8% of the overall landslide risk. The quantitative risk assessment results may provide useful reference for making risk-informed decisions for landslide risk mitgation measures and site planning.

Table 4 Quantitative landslide risk assessment results.

Slope failure modes	Failure probability	Mean runout distance (m)	Mean sliding volume (m3/m)	Risk	Risk contribution (%)
Shallow failure mode	5.25×10^{-3}	17.9	348.7	0.026	25.1
Intermediate failure mode	3.15×10^{-3}	19.3	626.0	0.032	30.5
Deep failure mode	3.83×10^{-3}	16.4	839.6	0.035	33.8
Progressive failure mode	7.75×10^{-4}	22.2	808.7	0.011	10.6
Overall	1.30×10^{-2}	18.0	588.3	0.104	100.0

7 Conclusions

In this paper, the geological and geotechnical uncertainties (including soil stratification uncertainty, inherent spatial variability of geomaterials, transformation and model uncertainties) and their modeling approaches are briefly introduced. To alleviate the limitation of low computational efficiency of the MCS for small probability problems, recent developments of two innovative approaches, surrogate model-based methods and advanced computer-based simulation methods, for slope reliability analysis and quantitative risk assessment of landslides are presented. Major topics including efficient slope reliability analysis methods using the PCE- and CNN-based response surface methods and subset simulation, and quantitative risk assessment methods of slope failure are systematically addressed. Finally, the effectiveness of the computer-based simulation methods developed by the authors is demonstrated using three slope examples. The effects of the spatial variability of soil properties on the failure modes, probability and risk of slopes are systematically investigated. It is found that these modern computational techniques can provide a versatile and promising tool for slope reliability analysis and risk assessment in spatially variable soils.

Acknowledgments

This work was supported by the Key Program of the National Natural Science Foundation of China (Grant No. U2240211), National Natural Science Foundation of China (Grant Nos. 52222905, 52179103 and 41972280) and Visiting Researcher Fund Program of State Key Laboratory of Water Resources and Hydropower Engineering Science (Grant No. 2019SGG03). The financial supports are gratefully acknowledged.

References

Ang, H. S., and Tang, W. H. (2007). *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*. 2nd edition. New York, John Wiley and Sons.

Au, S. K., and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4), 263-277.

Au, S. K. (2007). Augmenting approximate solutions for consistent reliability analysis. *Probabilistic Engineering Mechanics,* 22(1), 77-87.

Cami, B., Javankhoshdel, S., Phoon, K. K., and Ching, J. (2020). Scale of fluctuation for spatially varying soils, Estimation methods and values. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A, Civil Engineering*, 6(4), 03120002.

Cheng, H., Chen, J., Chen, R., Chen, G., and Zhong, Y. (2018). Risk assessment of slope failure considering the variability in soil properties. *Computers and Geotechnics*, 103, 61-72.

Cheng, Y. M., Lansivaara, T., and Wei, W. B. (2007). Two-dimensional slope stability analysis by limit equilibrium and strength reduction methods. *Computers and Geotechnics*, 34(3), 137-150.

Ching, J., Li, D. Q., and Phoon, K. K. (2016). Statistical characterization of multivariate geotechnical data. Chapter 4, *Reliability of Geotechnical Structures in ISO2394*, CRC Press/Balkema, 89-126.

Cho, S. E. (2010). Probabilistic assessment of slope stability that considers the spatial variability of soil properties. *Journal of Geotechnical and Geoenvironmental Engineering*, 136(7), 975-984.

Christian, J. T., Ladd, C. C., and Baecher, G. B. (1994). Reliability applied to slope stability analysis. *Journal of Geotechnical Engineering*, 120(12), 2180-2207.

Corominas, J., van, Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M. G., Pastor, M., Ferlisi, S., Tofani, V., Hervás, J., and Smith, J. T.

(2014) Recommendations for the quantitative analysis of landslide risk. *Bulletin of Engineering Geology and the Environment*, 73(2), 209-263.

Dai, F. C., Lee, C. F., and Ngai, Y. Y. (2002). Landslide risk assessment and management, an overview. *Engineering Geology*, 64(1), 65-87.

Deng, Z. P, Li, D. Q, Qi, X. H, Cao, Z. J, and Phoon, K. K. (2017). Reliability evaluation of slope considering geological uncertainty and inherent variability of soil parameters. *Computers and Geotechnics*, 92, 121-131.

Deng, Z. P., Pan, M., Niu, J. T., Jiang, S. H., and Qian, W. W. (2021). Slope reliability analysis in spatially variable soils using sliced inverse regression-based multivariate adaptive regression spline. *Bulletin of Engineering Geology and the Environment*, 80(9), 7213-7226.

Der, Kiureghian A, and Ke, J. B. (1988). The stochastic finite element method in structural reliability. *Probabilistic Engineering Mechanics*, 3(2), 83-91.

Ditlevsen, O. (1979). Narrow reliability bounds for structural systems, *J. Struct. Mech.* 7(4), 453-472.

Duncan, J. M. (2000). Factors of safety and reliability in geotechnical engineering. *Journal of Geotechnical and Geoenvironmental Engineering*, 126(4), 307-316.

Elfeki, A., and Dekking, M. (2001). A Markov chain model for subsurface characterization, theory and applications. *Mathematical Geology*, 33(5), 569-589.

Elkateb, T., Chalaturnyk, R., and Robertson, P. K. (2003). An overview of soil heterogeneity, quantification and implications on geotechnical field problems. *Canadian Geotechnical Journal* 40(1), 1-15.

El-Ramly, H, Morgenstern, N. R, and Cruden, D, M. (2002). Probabilistic slope stability analysis for practice. *Canadian Geotechnical Journal*, 39(3), 665-683.

Fei, S., Tan, X., Lin, X., Xiao, Y., Zha, F., and Xu, L. (2022). Evaluation of the scale of fluctuation based on variance reduction method. *Engineering Geology*, 308, 106804.

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., and Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Engineering Geology*, 102(3-4), 85-98.

Fenton, G. A., and Vanmarcke E. H. (1990). Simulation of random fields via local average subdivision. *Journal of Engineering Mechanics*, 116 (8), 1733-1749.

Finlay, P. J. (1996). *The Risk Assessment of Slopes*. University of New South Wales. PhD thesis.

Gilbert, R. B., and Tang, W. H. (1995). Model uncertainty in offshore geotechnical reliability. *Offshore Technology Conference*, Houston, Paper No. OTC-7757-MS.

Gingold, R. A., and Monaghan, J. J. (1977). Smoothed particle hydrodynamics, theory and application to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, 181(3), 375-389.

Griffiths, D. V., and Lane, P. A. (1999). Slope stability analysis by finite elements. *Geotechnique*, 49(3), 387-403.

Griffiths, D. V., and Fenton, G. A, (2004). Probabilistic slope stability analysis by finite elements. *Journal of Geotechnical and Geoenvironmental Engineering*, 130(5), 507-518.

He, X., Xu, H., Sabetamal, H., and Sheng, D. (2020). Machine learning aided stochastic reliability analysis of spatially variable slopes. *Computers and Geotechnics*, 126, 103711.

Huang, J., Lyamin, A. V., Griffiths, D. V., Krabbenhoft, K., and Sloan, S. W. (2013). Quantitative risk assessment of landslide by limit analysis and random fields. *Computers and Geotechnics*, 53, 60-67.

ISO, 2015. General Principles on Reliability of Structures. ISO 2394.

Ji, J., Liao, H. J., and Low, B. K. (2012). Modeling 2-D spatial variation in slope reliability analysis using interpolated autocorrelations. *Computers and Geotechnics*, 40, 135-146.

Ji, J., Jiang, Z., Yin, X., Wang, T., Cui, H., and Zhang, W. (2022). Slope reliability analysis based on deep learning of digital images of random fields using CNN. *Chinese Journal of Geotechnical Engineering*, 44(8), 1463-1473.

Jiang, S. H., Li, D. Q., Zhang, L. M., and Zhou, C. B. (2014). Slope reliability analysis considering spatially variable shear strength parameters using a non-intrusive stochastic finite element method. *Engineering Geology*, 168, 120-128.

Jiang, S. H., Li, D. Q., Cao, Z. J., Zhou, C. B., and Phoon, K. K. (2015). Efficient system reliability analysis of slope stability in spatially variable soils using Monte Carlo simulation. *Journal of Geotechnical and Geoenvironmental Engineering*, 141(2), 04014096.

Jiang, S. H, and Huang, J. S. (2016). Efficient slope reliability analysis at low-probability levels in spatially variable soils. *Computers and Geotechnics*, 75(5), 18-27.

Jiang, S. H., Huang, J., Yao, C., and Yang, J. H. (2017a). Quantitative risk assessment of slope failure in 2-D spatially variable soils by limit equilibrium method. *Applied Mathematical Modelling*, 47, 710-725.

Jiang, S. H., Papaioannou, I., Li, C. G., and Straub, D. (2017b). Integrating LEM with FEM through model correction factor method in reliability analysis of spatially variable slopes. *The 15th International Conference of the International Association for Computer Methods and Advances in Geomechanics (15th IACMAG)*, October 19-23, Wuhan, China. pp. 1-7.

Jiang, S. H., Huang, J., Griffiths, D. V., and Deng, Z. P. (2022). Advances in reliability and risk analyses of slopes in spatially variable soils, A state-of-the-art review. *Computers and Geotechnics*, 141, 104498.

Juang, C. H., Zhang, J., Shen, M. F., and Hu, J. Z. (2019). Probabilistic methods for unified treatment of geotechnical and geological uncertainties in a geotechnical analysis. *Engineering Geology*, 249, 148-161.

Kwong, A. K. L., Wang, M., Lee, C. F., and Law, K. T. (2004). A review of landslide problems and mitigation measures in Chongqing and Hong Kong, similarities and differences. *Engineering Geology*, 76, 27-39.

Lacasse, S., and Nadim, F. (1996). Model uncertainty in pile axial capacity calculations. *Offshore Technology Conference*, Houston, Paper No. OTC-7996-MS.

Li, D. Q., Jiang, S. H., Cao, Z. J., Zhou, W., Zhou, C. B., and Zhang, L. M. (2015). A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties. *Engineering Geology*, 187, 60-72.

Li, D. Q., Zheng, D., Cao, Z. J., Tang, X. S., and Phoon, K. K. (2016a). Response surface methods for slope reliability analysis, Review and comparison. *Engineering Geology*, 203, 3-14.

Li, D. Q., Xiao, T., Cao, Z. J., Phoon, K. K., and Zhou, C. B, (2016b). Efficient and consistent reliability analysis of soil slope stability using both limit equilibrium analysis and finite element analysis. *Applied Mathematical Modelling*, 40(9-10), 5216-5229.

Li, D. Q., Xiao, T., Cao, Z. J., Zhou, C. B., and Zhang, L. M. (2016e). Enhancement of random finite element method in reliability analysis and risk assessment of soil slopes using Subset Simulation. *Landslides*, 13(2), 293-303.

Li, D. Q., Zheng, D., Cao, Z. J., Tang, X. S., and Qi, X. H., (2019a). Two-stage dimension reduction method for meta-model based slope reliability analysis in spatially variable soils. *Structural Safety*, 81, 101872.

Li, J., Cassidy, M. J., Huang, J., Zhang, L., and Kelly, R. (2016c). Probabilistic identification of soil stratification. *Géotechnique*, 66(1), 16-26.

Li, K. S, and Lumb, P, (1987). Probabilistic design of slopes. *Canadian Geotechnical Journal*, 24(4), 520-535.

Li, L., Wang, Y., Zhang, L., Choi, C., and Ng, C. W. W, (2019b). Evaluation of critical slip surface in limit equilibrium analysis of slope stability by smoothed particle hydrodynamics. *International Journal of Geomechanics*, 19(5), 04019032.

Li, Z., Wang, X., Wang, H., and Liang, R. Y. (2016d). Quantifying stratigraphic uncertainties by stochastic simulation techniques based on Markov random field. *Engineering Geology*, 201, 106-122.

Liu, X., Wang, Y., and Li, D. Q. (2019). Investigation of slope failure mode evolution during large deformation in spatially variable soils by random limit equilibrium and material point methods. *Computers and Geotechnics*, 111, 301-312.

Liu, X, and Wang, Y. (2021). Probabilistic simulation of entire process of rainfall-induced landslides using random finite element and material point methods with hydro-mechanical coupling. *Computers and Geotechnics*, 132, 103989.

Low, B. K. (2015). Reliability-based design, practical procedures, geotechnical examples, and insights. In Phoon, K.K., Ching, J. (Eds.), *Risk and Reliability in Geotechnical Engineering*. CRC Press, Boca Raton, Florida, USA, pp. 385-424.

Lumb, P. (1966). The variability of natural soils. *Canadian Geotechnical Journal*, 3(2), 74-97.

Mayne, P. W., Christopher, B. R., and DeJong, J. (2001). *Manual on Subsurface Investigations*, National Highway Institute Publication No. FHWA NHI-01-031, Federal Highway Administration.

Mori, H., Chen, X., Leung, Y. F., Shimokawa, D., and Lo, M. K. (2020). Landslide hazard assessment by smoothed particle hydrodynamics with spatially variable soil properties and statistical rainfall distribution. *Canadian Geotechnical Journal*, 57(12), 1953-1969.

Ng, C. W. W., Qu, C., Cheung, R. W. M., Guo, H., Ni, J., Chen, Y., and Zhang, S. (2021). Risk assessment of soil slope failure considering copula-based rotated anisotropy random fields. *Computers and Geotechnics*, 136, 104252.

Nobre, M. M., and Sykes, J. F. (1992). Application of Bayesian kriging to subsurface characterization. *Canadian Geotechnical Journal*, 29(4), 589-598.

Papaioannou, I, and Kiureghian, A. D, (2010). Reliability-based design of slope angle considering spatial variability of soil material. *Computational Stochastic Mechanics – Proc. of the 6th International Conference (CSM-6)* G. Deodatis and P.D. Spanos (eds.) Rhodos, Greece, June 13-16, 2010.

Phoon, K. K., Huang, S. P., and Quek, S T. (2002). Implementation of Karhunen-Loeve expansion forsimulation using a wavelet-Galerkin scheme. *Probabilistic Engineering Mechanics*, 17(3), 293-303.

Phoon, K. K., and Kulhawy, F. H. (1999a). Characterization of geotechnical variability. *Canadian Geotechnical Journal*, 36(4), 612-624.

Phoon, K. K., and Kulhawy, F. H. (1999b). Evaluation of geotechnical property variability. *Canadian Geotechnical Journal*, 36(4), 625-639.

Phoon, K K., Ching, J., and Wang, Y. (2019). Managing risk in geotechnical engineering – from data to digitalization. *Proceedings, 7th International Symposium on Geotechnical Safety and Risk (ISGSR 2019)*, Taipei, Taiwan, 13-34.

Phoon, K. K., and Ching, J. Y. (2021). Project deepgeo - data-driven 3D subsurface mapping. *Journal of GeoEngineering*, 16(2), 61-73.

Phoon, K. K., Cao, Z. J., Ji, J., Leung, Y. F., Najjar, S., Shuku, T., Tang, C., Yin, Z. Y., Yoshida, I., and Ching, J. (2022). Geotechnical uncertainty, modeling, and decision making. *Soils and Foundations*, 62, 101189.

Qi, X. H., Li, D. Q., Phoon, K. K., Cao, Z. J., and Tang, X. S. (2016). Simulation of geologic uncertainty using coupled Markov chain. *Engineering Geology*, 207, 129-140.

Sulsky, D., Chen, Z., and Schreyer, H, L. (1994). A particle method for history-dependent materials. *Computer Methods in Applied Mechanics and Engineering*, 118 (1-2), 179-196.

Tang, C., and Phoon, K. K., (2021). *Model Uncertainties in Foundation Design.* CRC Press.U.S. Department of the Interior (USDI) & United States Geological Survey (USGS). (2004). *Landslide Types and Processes*. United States Geological Survey.

Vanmarcke, E. H. (1977). Probabilistic modeling of soil profiles. *Journal of the Geotechnical Engineering Division*, 103(11), 1227-1246.

Vanmarcke, E. H. (2010). *Random fields, analysis and synthesis*. Revised and Expanded New Edition. World Scientific Publishing, Beijing.

Varnes, D. J. (1978). Slope movement types and processes, in Schuster, R.L., and Krizek, R.J., eds., *Landslides - Analysis and control*, National Research Council, Washington, D.C., Transportation Research Board, Special Report 176, 11-33.

Wang, B., Hicks, M. A., and Vardon, P. J. (2016). Slope failure analysis using the random material point method. *Géotechnique Letters*, 6(2), 113-118.

Wang, Z. Z., and Goh, S. H. (2021). Novel approach to efficient slope reliability analysis in spatially variable soils. *Engineering Geology*, 281, 105989.

Weidinger, J. T. (2006). Predesign, failure and displacement mechanisms of large rockslides in the Annapurna Himalayas, Nepal. *Engineering Geology*, 83(1-3), 201-216.

Wong, F. S. (1985). Slope reliability and response surface method. *Journal of Geotechnical Engineering*, 111(1), 32-53.

Xiao, T., Li, D. Q., Cao, Z. J., Au, S. K., and Phoon, K. K., (2016). Three-dimensional slope reliability and risk assessment using auxiliary random finite element method. *Computers and Geotechnics*, 79, 146-158.

Xiao, T., Zhang, L., M, Li., X, Y., Li, D. Q. (2017). Probabilistic stratification modeling in geotechnical site characterization. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems*, Part A, Civil Engineering, 3(4), 04017019.

Yin, Y., Li, B., Wang, W., Zhan, L., Xue, Q., Gao, Y., Zhang, N., Chen, H., Liu, T., and Li, A. (2016). Mechanism of the December 2015 Catastrophic Landslide at the Shenzhen Landfill and Controlling Geotechnical Risks of Urbanization. *Engineering*, 2(2), 230-249.

Zhang, J., Zhang, L., and Tang, W. H. (2009). Bayesian framework for characterizing geotechnical model uncertainty. *Journal of Geotechnical and Geoenvironmental Engineering*, 135(7), 932-940.

Zhang, J., and Huang, H. W. (2016). Risk assessment of slope failure considering multiple slip surfaces. *Computers and Geotechnics*, 74, 188-195.

Zhang, L. M., and Dasaka, S. M. (2010). Uncertainties in geologic profiles versus variability in pile founding depth. *Journal of Geotechnical and Geoenvironmental Engineering*, 136(11), 1475-1488.

Zhang, P., Yin, Z. Y., Jin, Y. F., and Chan, T. H. T. (2020). A novel hybrid surrogate intelligent model for creep index prediction based on particle swarm optimization and random forest. *Engineering Geology*, 265, 105328.

Zhao, T. and Wang, Y. (2020). Interpolation and stratification of multilayer soil property profile from sparse measurements using machine learning methods. *Engineering Geology*, 265, 105430.

Zhou, X., and Sun, Z. (2020). Quantitative assessment of landslide risk using Monte Carlo material point method. *Engineering Computations*, 37(5), 1577-1596.

Zhu, B., Hiraishi, T., Pei, H., and Yang, Q. (2021). Efficient reliability analysis of slopes integrating the random field method and a Gaussian process regression-based surrogate model. *International Journal for Numerical and Analytical Methods in Geomechanics*, 45(4), 478-501.