Optimization of site investigation program for reliability assessment of undrained slope using Spearman rank correlation coefficient

Liang Zhang[1], Lei Wang[2*]

Abstract: Site investigation programs (e.g., boreholes) are crucial in characterizing soil properties and stratigraphic configurations. However, the traditional borehole patterns are generally of equally spaced distribution for the slope design, and the locations and total number of boreholes are considerably determined depending on engineers' experience, which may lead to cost-inefficient geotechnical design, especially considering the soil spatial variability. To address this dilemma, this paper presents a Spearman rank correlation coefficient-based scheme to optimize site investigation in slope design, where both locations and total number of boreholes are optimized. Conditional random field simulations are performed to consider the effect of the borehole data on the estimation of the soil property distribution. The superiority of the proposed method to the traditional method is illustrated by a comparison study in an undrained slope example. In this example, the accuracy of the characteristics of the slope (i.e., the factor of safety, location of slip surface, and sliding volume), robustness of the estimated characteristics of the slope, and risk reduction are examined. The comparison results show the effectiveness of the proposed method in accurately estimating the characteristics of the slope without prior knowledge about the slip surface, since the slip surface is unknown for most practical cases prior to the site investigation. The most robust estimate results and risk reduction are obtained using the proposed method. This study can also provide useful references to build an adaptive unequally spaced borehole pattern in practice.

Keywords: Optimization; Spatial variability; Slope stability; Spearman rank correlation coefficient.

---

[1] Research Assistant, Department of Civil and Architectural Engineering and Construction Management, University of Cincinnati, Cincinnati OH 45221, USA.

[2] Assistant Professor, Department of Civil and Architectural Engineering and Construction Management, University of Cincinnati, Cincinnati OH 45221, USA. (*Corresponding author: wang4li@ucmail.uc.edu)

1. Introduction

Natural soils are very complicated and highly variable geomaterials, and they are products of complex geological processes and depositional environments. To investigate the soil properties at geotechnical sites, site investigations (e.g., boreholes) are typically conducted in practice. However, with the restriction of time and budget for most geotechnical design projects, only a limited number of boreholes at scattered locations over a construction site are typically planned and executed, which results in significant uncertainties in the geotechnical characterization of the site (Jiang et al. 2018b&2020; Yang et al. 2019&2022). Furthermore, geotechnical properties for a given site can exhibit considerable spatial variability due to the natural fluctuation of material constituents, randomness in the depositional history, and variable historical loading conditions (Huang et al. 2020), which causes more challenges in the optimization of the borehole patterns for geotechnical design.

Some previous optimization studies aimed to accurately predict the soil properties at unsampled locations with measured borehole data in geotechnical profiles (e.g., Wang et al. 2017; Cai et al. 2019; Zhao et al. 2021), while other studies focused on uncertainty reduction in the characterization of the spatial variability (i.e., the mean, standard deviation, and scale of fluctuation of soil properties) at geotechnical sites (e.g., Lloret-Cabot et al. 2012; Li et al. 2016c; Xiao et al. 2018; Huang et al. 2020; Han et al. 2022). Although these borehole schemes provide useful means to characterize a given geological profile, they may not be effective in characterizing the performance of geotechnical systems (e.g., the slope, foundation, and tunnel). In a geological profile, all soil elements are of equal importance to provide information about the soil properties, and equally spaced borehole patterns can be acceptable. However, the optimal borehole patterns are generally related to the failure mechanism of geotechnical systems. The soil elements at the influence zones that control the failure mechanism of geotechnical systems are more influential in determining the optimal borehole patterns. For instance, Chwała (2021) investigated the effect of the space between two symmetrically distributed soil soundings on the bearing capacity of a rectangular footing foundation. The optimal space was found to depend on the normalized scale of fluctuation by the foundation length. It was concluded that the boreholes in the area with more dissipated energy in the foundation were more effective in reducing the uncertainty of the bearing capacity estimation. Li et al. (2016a) and Deng et al. (2017) showed that the boreholes at the place where the slip surface was extended resulted in a more accurate estimate of the mean and a smaller standard deviation of the factor of safety (*FS*) of the slope, since the soil elements in these areas determined how the slip surface could be formed. However, most current studies to optimize borehole patterns in geotechnical design have equally spaced borehole patterns that follow traditional site investigation programs (Gong et al. 2014 & 2017; Li et al. 2016a; Li et al. 2016b; Deng et al. 2017; Liu et al. 2020). Hence, the two optimized objectives in the borehole patterns, which are the locations and total number, can be considered a function of the borehole space, since sufficiently many boreholes will be fully distributed in the site for a given borehole space. As mentioned above, since the soil elements at the influence zone have a more considerable effect on the geotechnical system, more boreholes should be arranged at the most important influence zones in sequence, which implies that traditional equally spaced borehole patterns are more likely cost-inefficient.

This paper aims to propose an effective approach to optimize site investigation considering the spatial soil variability in slope engineering based on correlation analysis, where the influence zone can be automatically determined without prior knowledge about the slip surface. The effectiveness of the proposed method is validated in three aspects: the estimate accuracy, uncertainty reduction, and risk reduction, according to previous studies (Cai et al. 2019; Jiang et al. 2018b&2020; Yang et al. 2019&2022). The main advantages of this approach compared to the traditional method are: 1) The influence zone of the slope system can be automatically determined without prior knowledge about the slip surface; 2) The locations and total number of boreholes are separately optimized; 3) Most estimate accuracy and uncertainty reduction (in terms of the robustness) of the characteristics of the slope can be obtained; 4) Most risk reduction (i.e., expected loss cost) can be reached in the proposed method; 5) The proposed method is easy to implement due to its simple concept. The remainder of the paper is organized as follows. The optimization methodology is first briefly introduced. Then, an undrained slope example is taken to illustrate the effectiveness of the proposed method. A comparison study is conducted to evaluate the estimated accuracy of the characteristics of the slope (i.e., the factor of safety, location of slip surface, and sliding volume) between the proposed method and traditional methods. Afterwards, the robustness of the estimated results and risk reduction of the entire slope engineering system are comprehensively assessed. Finally, the concluding remarks are made based on the results.

2 Methodologies to optimize the site investigation program

Due to the restriction of time and budget for most slope engineering projects, only limited measured data (e.g., from boreholes) can be obtained. The soil properties at borehole locations are "known" without uncertainty, while other soil properties from unsampled locations are estimated by the borehole data with uncertainty. Since the spatial correlations of soil properties generally decrease with the relative distance, the constraint of the borehole data decreases with the relative distance to the boreholes, which leads to more uncertainty in the estimate of the soil properties far from existing boreholes. To characterize this feature and the spatial variability of soil properties, the conditional random field theory is employed to optimize the site investigation program (Li et al. 2016b; Liu et al. 2017). Conditional random field simulations can be realized by statistical methods such as the Bayesian method, Hoffman method, and kriging-based sampling method (Gong et al. 2018). As a linear unbiased estimation method, the kriging-sampling method uses a weighted linear average of nearby soil samples to predict soil properties at unsampled locations. The spatial autocorrelation function and unconditional random field simulations are also incorporated in the generation of conditional random fields. Thus, the soil property values at sampled locations always match the known data in the conditional random field simulations by the kriging method. The kriging method also ensures the uncertainty at unsampled locations in terms of the variance, which reduces with the distance to the borehole locations, and no uncertainty of soil samples at the sampled locations (i.e., the variance is zero), which is consistent with our basic knowledge. In addition, the kriging-sampling method is computationally efficient and easy to implement, since the high-dimensional matrix can be avoided, and the weight vector needs to be calculated only once for any number of MCS in the point-by-point prediction of unsampled soil samples. The kriging method has been validated to give sufficiently accurate and reliable predictions by both theoretical models and realistic models (Wang et al. 2017; Li et al. 2016b; Chen et al. 2018; Huang et al. 2019). Therefore, the kriging-sampling method is adopted in this study. Based on the constructed conditional random fields, correlation analysis using the Spearman rank correlation coefficient is performed to determine the optimal borehole patterns (i.e., the locations and total number of boreholes). Although the conditional random field simulations by the kriging method are not new, the conditional random field simulation procedures should be briefly introduced.

2.1 Conditional random field simulations of the soil properties

The conditional random fields by the kriging method are generated based on unconditional random field simulations, which are first reviewed as follows. The soil properties are generally assumed to be lognormally distributed because the soil properties have nonnegative values (Jiang et al. 2018a&2018b; Gong et al. 2018; Yang et al. 2019; Chen and Zhang 2021). For a lognormal random field soil property s with prior knowledge of the mean μ_s and coefficient of variation (COV) δ_s, the mean $\mu_{\ln s}$ and standard deviation $\sigma_{\ln s}$ of the equivalent normal random field ln$s$ are calculated as follows.

$$\sigma_{\ln s} = \sqrt{\ln(1+\delta_s^{\,2})} \tag{1a}$$

$$\mu_{\ln s} = \ln(\mu) - 0.5\sigma_{\ln s}^2 \tag{1b}$$

The anisotropic exponential autocorrelation structure is adopted to characterize the correlation coefficient ρ_{ij} between the normalized soil property ln$s$ at two different locations of (x_i, y_i) and (x_j, y_j), which is calculated as follows.

$$\rho_{ij} = \rho(|x_j - x_i|, |y_j - y_i|) = \exp\left(-\frac{2|x_j - x_i|}{\lambda_{\ln x}} - \frac{2|y_j - y_i|}{\lambda_{\ln y}}\right) \tag{2}$$

where $|x_j - x_i|$ and $|y_j - y_i|$ are the absolute distances between two positions (x_i, y_i) and (x_j, y_j) along the X and Y directions, respectively; $\lambda_{\ln x}$ and $\lambda_{\ln y}$ are the scales of fluctuation of the equivalent normal random field ln$s$ along the X and Y directions, respectively.

A fixed value is assigned to the soil element domain instead of at the mesh grids. The mean of the soil property $\mu_{\ln s\mathrm{E}}$ that should be averaged over the soil element domain is equal to that of the local soil property $\mu_{\ln s}$, while the standard deviation of the averaged soil property $\sigma_{\ln s\mathrm{E}}$ is reduced. For the autocorrelation structure established in Eq. (2), the variance reduction factor of the concerned element can be estimated by the equations in Knabe et al. (1998) and Huang and Griffiths (2015) with a range of 0-1. There are various sampling methods to generate unconditional random fields, such as the local average subdivision method, turning-band method, fast Fourier transformation method, and covariance matrix decomposition method (Fenton 1994; Yang and Ching 2021). In this study, the covariance matrix decomposition method is used for random field generation. For given mean, standard deviation, and autocorrelation structure, the $n_\mathrm{E} \times n_\mathrm{E}$ autocorrelation matrix $\boldsymbol{R}_{\ln s}$ of the soil property between every two soil elements can be constructed. A possible realization of the lognormal random field simulation can be generated as follows.

$$s_{ij} = \exp\left(\mu_{\ln s\mathrm{E}j} + \sigma_{\ln s\mathrm{E}j} \cdot \ln s_{ij}\right) \tag{3}$$

where s_{ij} is the j^{th} numerical element of the i^{th} realization of the random field (i = 1, 2, …, N_p; j = 1, 2, …, n_E), N_p is the number of realizations of the random field, and n_E is the number of discretized numerical elements of the slope; $\mu_{\ln s\mathrm{E}j}$ and $\sigma_{\ln s\mathrm{E}j}$ are the averaged mean and standard deviation of the soil property ln$s$ over the j^{th} numerical element, respectively; lns_{ij} is the j^{th} element of the i^{th} realization of the random field. The matrix ln$\boldsymbol{s}_i$ of the soil property for all numerical elements is derived as follows.

$$\ln \boldsymbol{s}_i = \boldsymbol{L}_{\ln s} \boldsymbol{\xi}_i \tag{4a}$$

$$\boldsymbol{R}_{\ln s} = \boldsymbol{L}_{\ln s} \times \boldsymbol{L}_{\ln s}^{T} \tag{4b}$$

where $\boldsymbol{\xi}_i$ is an $n_\mathrm{E} \times 1$ standard normal sample vector (i = 1, 2, …, N_p), which may be obtained with Latin hypercube sampling; $\boldsymbol{L}_{\ln s\cdot}$ is a lower triangular matrix of autocorrelation matrix $\boldsymbol{R}_{\ln s}$

derived by Cholesky decomposition technique.

Suppose that the borehole data are located at the points (x_1, y_1), (x_2, y_2)..., (x_p, y_p) and the unsampled locations are (x_{p+1}, y_{p+1}), (x_{p+2}, y_{p+2})..., ($x_{n\mathrm{E}}$, $y_{n\mathrm{E}}$). Based on the generated unconditional random fields and borehole data, the conditional random fields can be simulated by the kriging method as follows (Liu et al. 2017):

Step 1: Calculate the locally averaged mean $\mu_{\ln s\mathrm{E}}$ and standard deviation $\sigma_{\ln s\mathrm{E}}$ of the soil property s in normal space (see Eq. (1));

Step 2: Generate the unconditional random fields $\ln s^{\mathrm{UC}}$ of soil property s in normal space with obtained mean $\mu_{\ln s\mathrm{E}}$, standard deviation $\sigma_{\ln s\mathrm{E}}$, and scales of fluctuation $\lambda_{\ln x}$ and $\lambda_{\ln y}$ (see Eq. 2 and Eq. 4);

Step 3: Extract the values at sampled locations from the generated unconditional random fields as the "known data". The normalized soil properties $\ln s^{\mathrm{KU}}$ at unsampled locations can be estimated by the "known data" as

$$\begin{bmatrix} \boldsymbol{K} & \boldsymbol{I} \\ \boldsymbol{I}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\kappa} \\ 1 \end{bmatrix} \tag{5a}$$

$$K_{ij} = \rho_{ij} \cdot \sigma_{si} \cdot \sigma_{sj} \quad (i \text{ and } j = 1,2,\ \ldots,\ p) \tag{5b}$$

$$\boldsymbol{\beta} = \left[\beta_1, \beta_2, \ldots, \beta_p\right]^T \tag{5c}$$

$$\kappa_{ij} = \rho_{ij} \cdot \sigma_{si} \cdot \sigma_{sj} \; (i = 1,\ 2,\ \ldots,\ n_{\mathrm{E}}\text{-}p;\ j = 1,\ 2,\ \ldots,\ p) \tag{5d}$$

where $\boldsymbol{K}$ is the covariance matrix derived from borehole data, and each element K_{ij} of $\boldsymbol{K}$ can be calculated by Eq. (5b); $\boldsymbol{I}$ is a $p \times 1$ vector with all values equal to 1; $\boldsymbol{\beta}$ is a $p \times 1$ weight vector (see Eq. (5c)) with $\sum_{i=1}^{p} \beta_i = 1$; $\boldsymbol{\kappa}$ is the vector of covariance between the unsampled point and borehole data, and each element κ_{ij} of $\boldsymbol{\kappa}$ is derived from Eq. (5d); □ρ_{ij} in Eq. (5b) is the spatial correlation between the i^{th} and j^{th} borehole data (see Eq. (2)), while that in Eq. (5d) is the spatial correlation between the i^{th} unsampled point and j^{th} borehole data. With obtained weight vector $\boldsymbol{\beta}$, for example, the soil property value of the $(p+1)^{th}$ soil element at location (x_{p+1}, y_{p+1}) is estimated from borehole data as

$$\ln s^{K}(x_{p+1}, y_{p+1}) = \sum_{i=1}^{p} \beta_i \ln s(x_i, y_i) \tag{6}$$

Step 4: Estimate the normalized soil property values $\ln s^{KK}$ at unsampled locations with the real known data by repeating step 3.

Step 5: Obtain the normal conditional random fields as

$$\ln s^{c}(x,y) = \ln s^{KK}(x,y) + [\ln s^{UC}(x,y) - \ln s^{KU}(x,y)] \tag{7}$$

where $\ln s^{UC}(x,y)$ is the unconditional random field; $\ln s^{KK}(x,y)$ is the random field estimated by the kriging method for the given borehole data; $\ln s^{KU}(x,y)$ is the random field estimated by the kriging method, which takes the values at the borehole locations from the unconditional random field as the borehole data. According to Eq. (7), the soil properties at unsampled locations are estimated as kriging random fields $\ln s^{KK}(x,y)$ with a stochastic error of $|\ln s^{UC}(x,y) - \ln s^{KU}(x,y)|$, which increases with the distance between unknown and known data. Therefore, the discontinuity of the soil property distribution can be avoided, although the estimated soil properties at sampled locations always match the known data.

Step 6: Transfer the normal conditional random fields into lognormal conditional random fields using Eq. (3).

2.2 Correlation analysis using Spearman rank correlation coefficient to locate additional boreholes

With the constructed conditional random fields, the characteristics of the slope (i.e., the factor of safety, location of the slip surface, and sliding volume) can be captured by Monte Carlo simulations (MCS). Since stability assessment is the most crucial problem for the slope, the most effective boreholes are more likely to locate at the places where the soil elements are positively related to the factor of safety of the slope. Only the soil strength properties (e.g., the undrained shear strength c_u) are modeled by the random fields, and the large values of the strength properties tend to correlate with higher values of *FS*. The values of the correlation coefficient between each soil element and *FS,* which can be either positive or negative, can be implemented to characterize the contribution of the soil element to the *FS* instead of the absolute values of the correlation coefficient. As shown later in the following example application, there are very low negative relations between the soil elements far from the slip surface and *FS,* while the soil elements near the slip surface strongly positively correlate with *FS*. Various correlation coefficients may be applied for the correlation analysis, such as the Pearson correlation coefficient, Kendall correlation coefficient, and Spearman rank correlation coefficient. Although the Pearson correlation coefficient is much more popular, the Pearson correlation coefficient is generally effective in characterizing linear correlations and more likely to mischaracterize the relationships and cause bias due to the nonnormality of the data (Bishara and Hittner 2015). As an alternative (Bishara and Hittner 2015; De et al. 2016; Thirumalai et al. 2017), the Spearman rank correlation coefficient is 1) applicable for both normal and nonnormal distributed data; 2) effective in characterizing linear or nonlinear correlations; 3) more robust and insensitive to outliers. Compared to Kendall's tau correlation coefficient, the Spearman rank correlation coefficient is less computationally demanded, less complicated and sufficiently accurate to characterize the correlations in this study. Therefore, the Spearman rank correlation coefficient is adopted to characterize the correlation between soil elements and *FS*. Suppose that the soil property values of the j^{th} numerical element with N_p realizations of the random field are $\boldsymbol{s}_j = [s_{1j}, s_{2j}, s_{3j}, \ldots, s_{N_pj}]$ and the factors of safety with N_p random field simulations are $\boldsymbol{FS} = [FS_1, FS_2, \ldots, FS_{Np}]$. The Spearman rank correlation coefficient between the j^{th} soil element and $\boldsymbol{FS}$ is formulated as:

$$\rho_{Spearman}(\boldsymbol{s}_j, \boldsymbol{FS}) = \frac{\sum_{i=1}^{N_p}(n_{1i} - \bar{n}_1)(n_{2i} - \bar{n}_2)}{\sqrt{\sum_{i=1}^{N_p}(n_{1i} - \bar{n}_1)^2}\sqrt{\sum_{i=1}^{N_p}(n_{2i} - \bar{n}_2)^2}} \tag{8}$$

where n_{1i} and n_{2i} are the ascending or descending sorted positions determined by the values of each element in $\boldsymbol{s}_j$ and $\boldsymbol{FS}$ with N_p random field simulations, respectively; $\bar{n}_1$ is the mean of n_{1i} (i=1, 2, …, N_p); $\bar{n}_2$ is the mean of n_{2i} (i=1, 2, …, N_p). The Spearman rank correlation coefficient can be directly obtained using a command (corr($\boldsymbol{s}_j$, $\boldsymbol{FS}$, ‘type’, ‘spearman’)) in MATLAB, where all calculation procedures are involved. After the sensitive soil elements with high Spearman rank correlation coefficients are revealed, the stability of the slope can be accurately evaluated. Since each soil sample is equally mapped from the random field simulations and plays an equal role in building a geotechnical profile, the amount of information provided by each soil sample should be considered the same, while the importance of information (characterized by the Spearman rank correlation coefficient in this study) brought by each soil sample depends on the characteristics (e.g., failure mechanism) of the geotechnical systems. The importance of information provided by a borehole is believed to be well characterized by its statistics, i.e., the mean or sum of the data from the borehole. However, it may be problematic in some scenarios when the sum-based method is adopted.

For example, the important local area (e.g., slope toe) that controls the slope stability cannot be well captured by the sum-based method, since the optimal borehole location is more likely determined at the place where more data can be obtained by the sum-based method, which is not consistent with the fact that the boreholes at the slope domain are generally more effective, although fewer data are measured in this area (Li et al. 2016b). Therefore, the mean Spearman rank correlation coefficient is adopted to evaluate the effectiveness of the borehole data for the stability assessment of the slope.

The flowchart of the proposed method is illustrated in Figure 1. In the first step, N_p samples of unconditional random fields are generated with prior knowledge of the mean, standard deviation, and scales of fluctuation of the soil properties. Then, the degree of the influence of each soil element on the *FS* in terms of the Spearman rank correlation coefficient is calculated in step 2. In step 3, the effectiveness of the additional borehole along all potential horizontal locations is evaluated based on the mean Spearman rank correlation coefficient of the soil elements from the borehole. The optimal borehole is located where the borehole has the largest Spearman rank correlation coefficient. With the extracted borehole data, the conditional random fields are simulated to consider the constraint of the added borehole in step 4. In step 5, the *FS*, location of the slip surface, and sliding volume can be evaluated for the updated borehole pattern. This process will be repeated until the target number of boreholes (N_{BH}) is reached. Whether the target probability of failure or the accuracy of the target factor of safety is reached may be the optional ended conditions for the optimization. These values generally depend on the specific slope problems, and the target number of boreholes with a limited budget is assumed in this study.

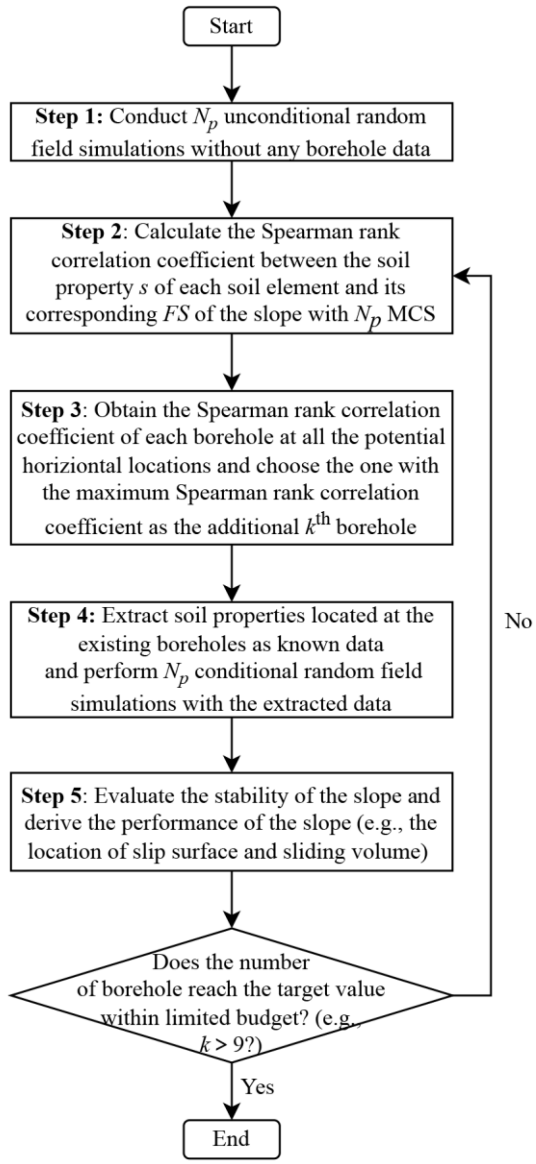


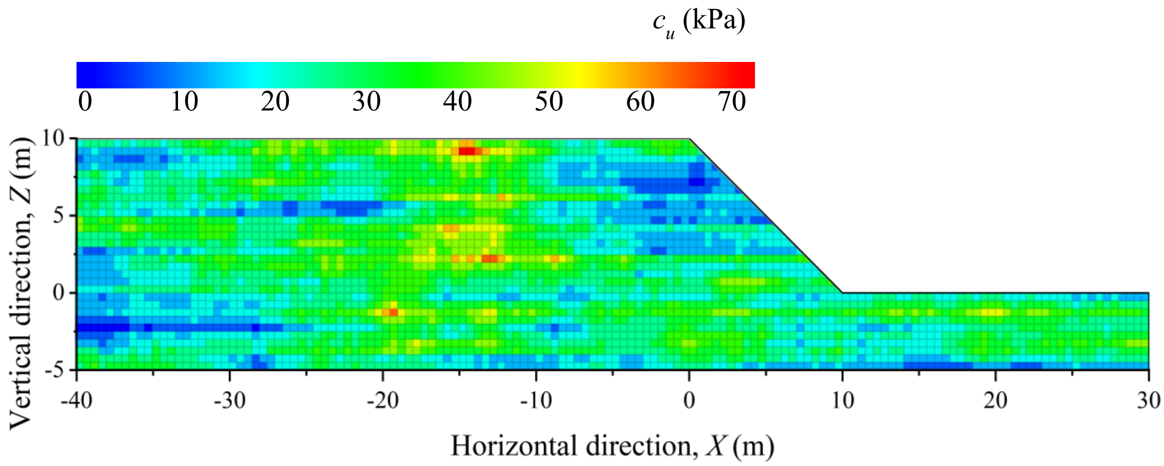
Figure 1. Flow chart of the proposed method for site investigation

3. Example application for the slope problem

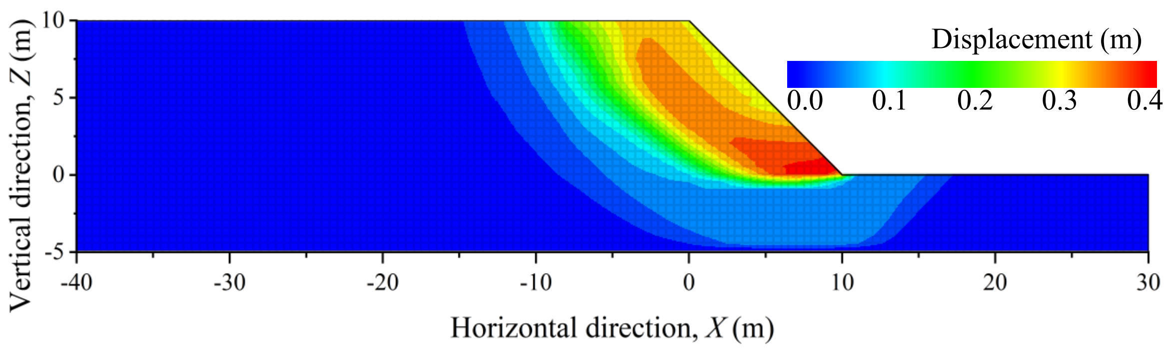
In this section, an undrained slope with a height of 10 m and a slope angle of 45° (1:1 slope) is adopted as an example to demonstrate the proposed method. The factor of safety is obtained using the strength reduction method built in the 3-D explicit finite difference program FLAC3D version 7.0 (Itasca 2022). The mesh size of the numerical model is 0.5 m by 0.5 m. The elastic-perfectly plastic Mohr–Coulomb model is adopted to model the soil behaviors. A fixed boundary is applied to the slope bottom, while the roller boundary is applied to the slope back and front faces. The undrained shear strength c_u is simulated as random fields, while other soil properties are set as constant values. The soil properties are tabulated in Table 1. The effectiveness of the proposed method is validated by a comparative study with traditional methods. To compare with the "true" site characterization, one of the generated unconditional random field simulations is taken as a "true" slope (Shen et al., 2018). Therefore, the characteristics of the "true" slope (i.e., the factor of safety, location of the slip surface, and sliding volume) can be derived to validate the proposed method. The spatially variable soil of the "true slope" with FS = 1.05 is illustrated in Figure 2(a), while the contours of the displacement and shear strain increment are shown in Figure 2(b) and Figure 2(c), respectively.

Table 1. Statistics of the soil properties for the example slope problem

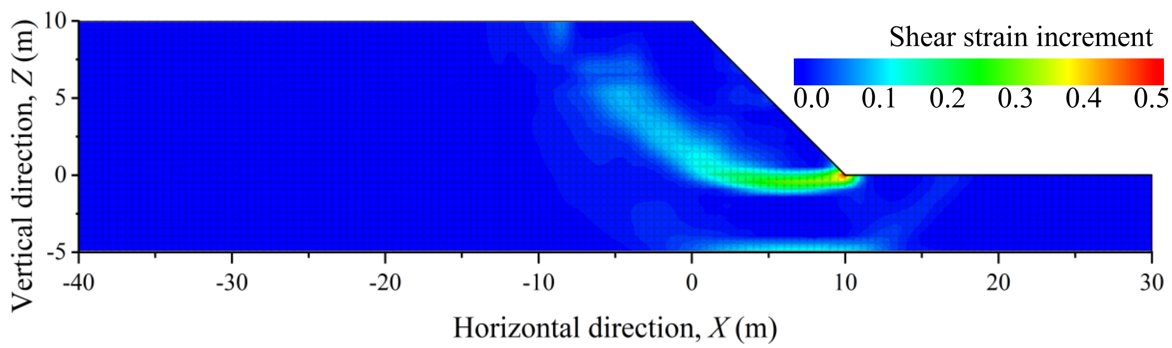
Parameter	**Value**
Density, ρ (kg/m^3)	2000
Young's modulus, E (MPa)	100
Poisson's ratio, v	0.30
Mean of undrained shear strength, c_u (kPa)	40
COV of undrained shear strength c_u	0.3
Horizontal scale of fluctuation, $\lambda_{\ln x}$ (m)	40
Vertical scale of fluctuation, $\lambda_{\ln v}$ (m)	4



(a) Spatially variable soil



(b) Contour of the displacement



(c) Contour of the shear strain increment

Figure 2. Characteristics of the “true” slope (FS = 1.05)

3.1 Illustration of the borehole patterns for the comparison study

In the traditional method, boreholes generally have an equally spaced distribution or are symmetrically distributed in an estimated influence zone (Gong et al. 2014; Yang et al. 2019&2022). However, the determination of the range of the influence zone considerably depends on the engineers' experience in practice, which leads to cost-inefficient design. For the slope in Figure 2, without loss of generality, 9 boreholes (N_{BH} = 9) with a limited budget are assumed for the slope problem. According to Li et al. (2016a) and Deng et al. (2017), boreholes that reveal the location of the slip surface are more effective. In the numerical analysis of slope stability by FLAC3D, two main methods are employed to automatically locate the slip surface (Wang et al., 2020). The first method is the shear strain increment-based (SSI-based) method, and the other method is the nodal displacement-based method. The latter is more extensively used for its simplicity, and it is now built in FLAC3D. This method can also be effective in identifying local failures in 3-D slopes with spatially variable soil (Zhang et al. 2022). Therefore, the nodal displacement-based method is adopted to locate the slip surface in this study. In the nodal displacement-based method, a threshold of the maximum nodal displacement of all mesh grids should be first determined. A soil element where the displacement of all nodes exceeds the threshold of the maximum nodal displacement will be considered a sliding soil element. The sliding surface is identified as the boundary between sliding soil elements and stable soil elements. To determine the threshold value of the maximum nodal displacement, the slip surface derived by the max shear strain increment can be taken as a benchmark (Hicks et al. 2014). The results in Figure 3 indicate that the contour of 35% maximum nodal displacement coincides with the contour of the maximum shear strain increment for the "true" slope. As shown in Figure 3, it is suggested that some advanced smoothing methods such as the polynomial spline technique should be applied to the slip surface to satisfy the kinematic demands for the slope (Wang et al. 2020). However, the unsmoothed error is negligible in the comparison results. Therefore, the fitting technique is not implemented in this study to simplify the calculation of the sliding volume.

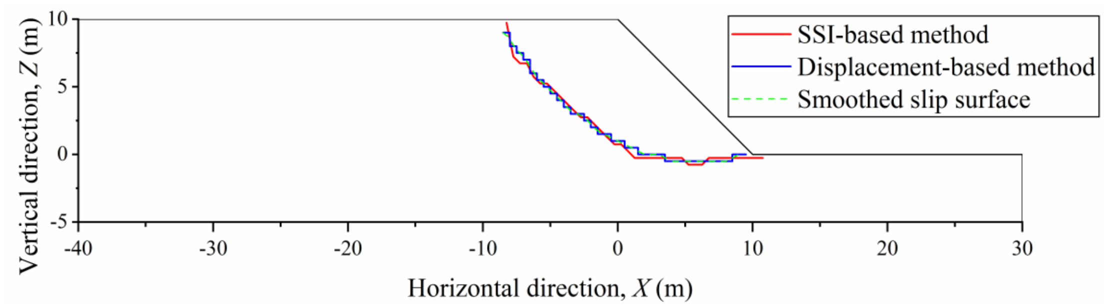
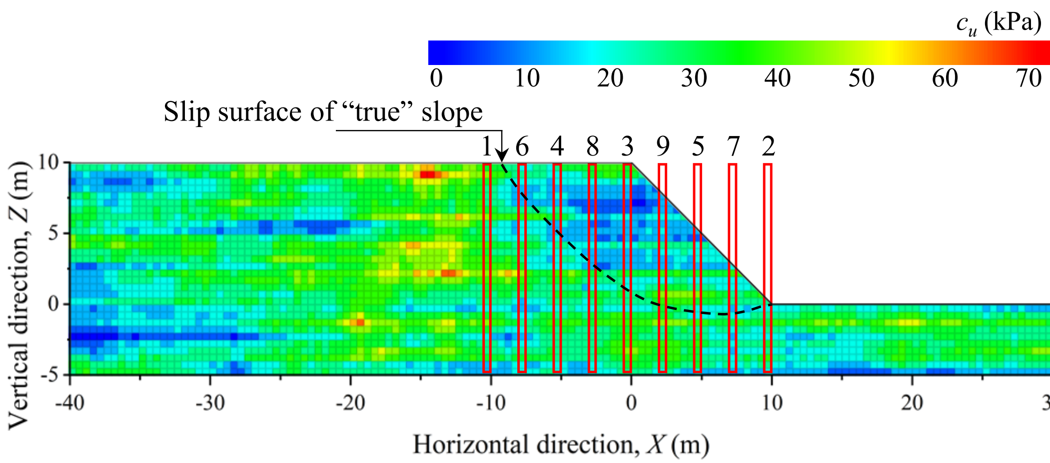
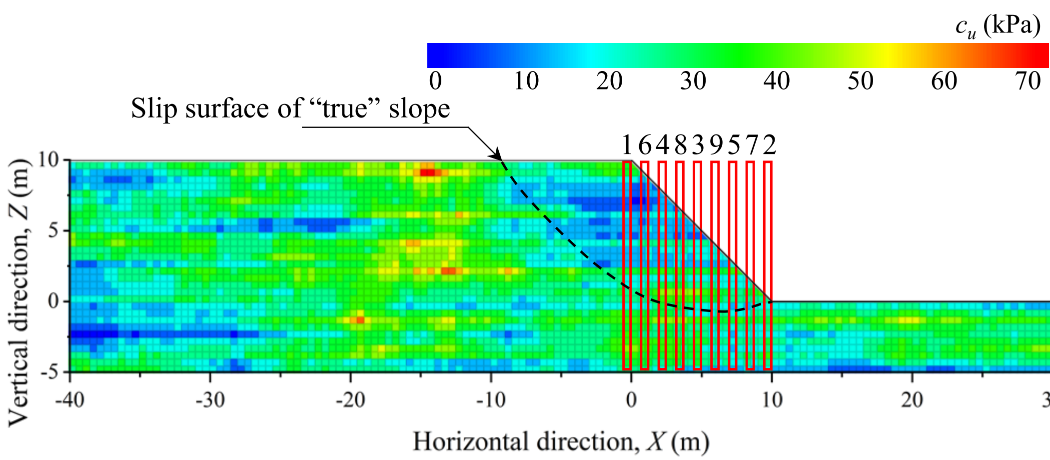


Figure 3. Location of the slip surface, which is determined by 35% of the maximum nodal displacement and maximum shear strain increment

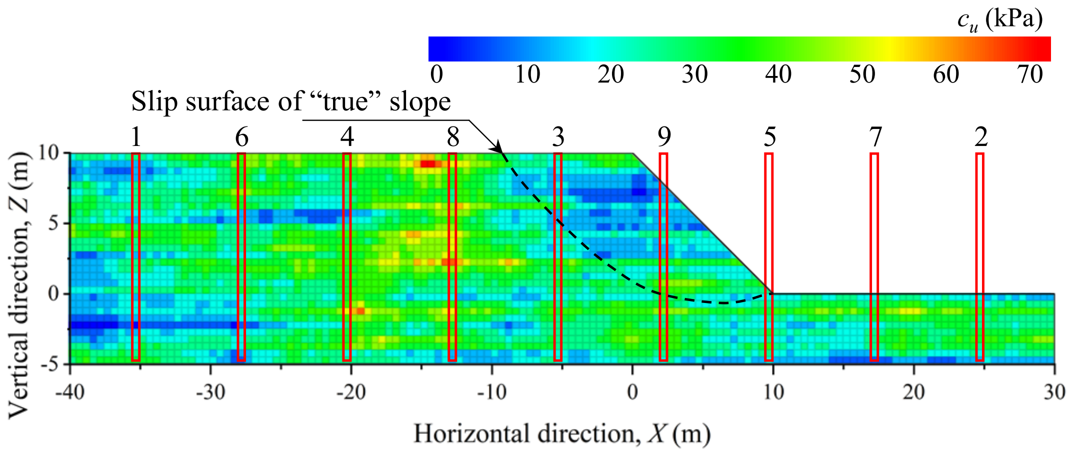
Based on the traditional method, several boreholes may be equally spaced and applied to the area where the slip surface is accurately estimated, which is denoted as the borehole pattern with good judgement (see Figure 4(a)). Since the slip surface is bound to go through the slope domain, all boreholes can be located at the slope domain, which is denoted as the borehole pattern with moderate judgement (see Figure 4(b)). To avoid the nonconservative design, a sufficiently larger area (of the model domain) may be estimated for the borehole distributions, which is denoted as the borehole pattern with poor judgement (see Figure 4(c)). With predefined traditional borehole patterns, Monte-Carlo simulation (MCS) is employed to capture the performance of the slope. As shown in Figure 5(a), the mean and standard deviation of *FS* converge when 800 MCS runs are performed. From Figure 5(b), the mean Spearman rank correlation coefficient of the soil elements from the entire numerical model and one borehole at the slope crest also becomes stable within 800 runs. Therefore, 1000 runs of MCS (N_p = 1000) are considered sufficient to obtain the slope system responses. With 1000 runs of MCS, both unconditional and conditional random field simulations are validated by the comparison between the preset statistics (e.g., the mean and standard deviation) of soil properties and those derived from the simulations (Gong et al. 2018; Huang et al. 2019; Johari and Fooladi 2020).



(a) Good judgement

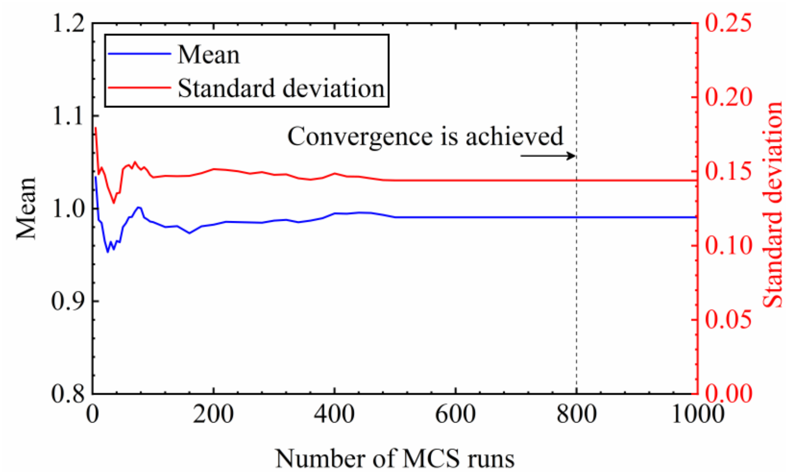


(b) Moderate judgement

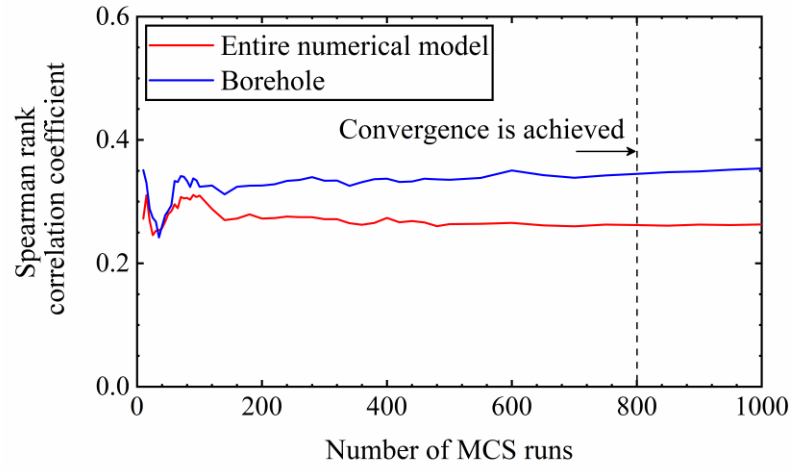


(c) Poor judgement

Figure 4. Borehole patterns from traditional methods



(a) Convergence of FS



(b) Convergence of the Spearman rank correlation coefficient

Figure 5. Convergence of the *FS* and Spearman rank correlation coefficient with the increase in number of MCS runs

Following the proposed site investigation method, the boreholes are added step by step. Figure 6 shows the contour of the Spearman rank correlation coefficient without boreholes. The soil elements with a large Spearman rank correlation coefficient are distributed at the bottom of the slope domain, since these soil elements may determine whether the deep mode or shallow mode of failure will occur if no borehole is applied. The first borehole is optimized at X = 9.75 m (i.e., at the slope toe), which is consistent with the conclusion from Jiang et al. (2018a). Thus, the final borehole pattern from the proposed method is illustrated in Figure 7. The characteristics of the slope (i.e., the factor of safety, location of the slip surface, and sliding volume) with the traditional methods and proposed method are studied in the following comparative study.

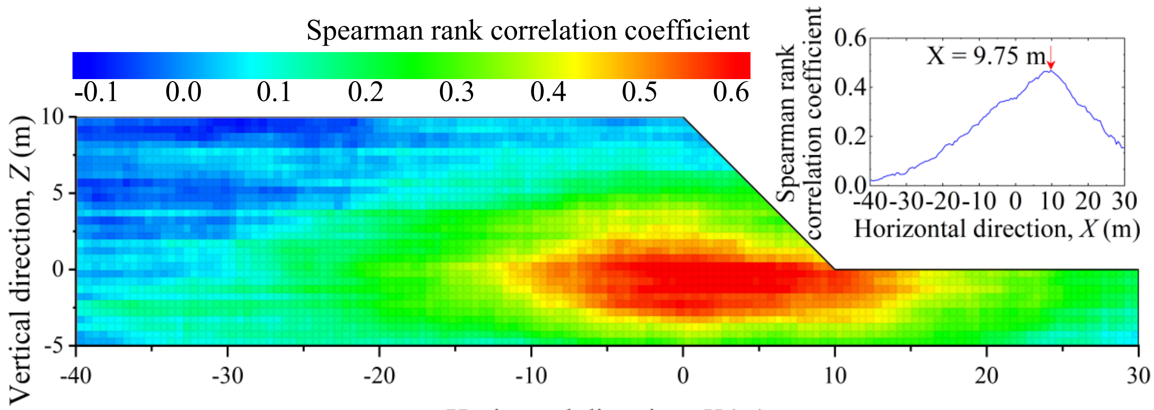


Figure 6. Determining the additional borehole from the Spearman rank correlation coefficient

($N_{BH} = 0$)

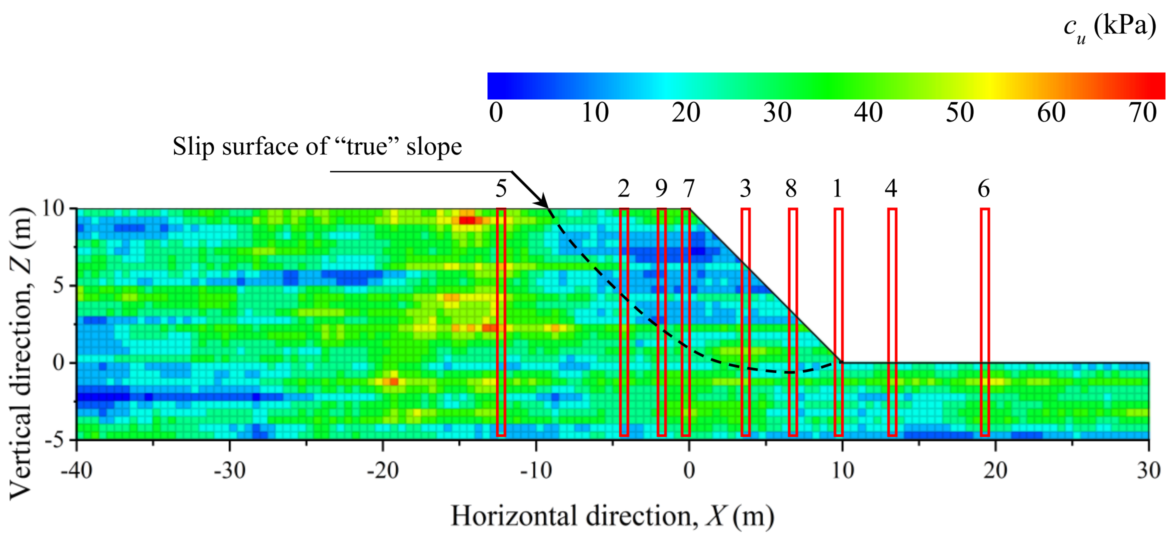


Figure 7. Borehole pattern obtained using the proposed method

3.2 Comparison study on the estimate of the slope characteristics

The mean *FS* with a (negative or positive) standard deviation from 1000 MCS is presented in Figure 8. As shown in Figure 8, the difference between *FS* of the "true" slope and mean estimated *FS* from all borehole patterns decreases with the number of boreholes. When the number of boreholes is larger than 3 (i.e., $N_{BH}>3$), the error is negligible, which implies that the mean *FS* can be effectively estimated for all borehole patterns. The standard deviation of the *FS* rapidly decreases with the increase in number of boreholes by applying the borehole patterns from the proposed method and traditional method with good judgement, which indicates that these two methods are superior to the other two in reducing the uncertainty of *FS*.

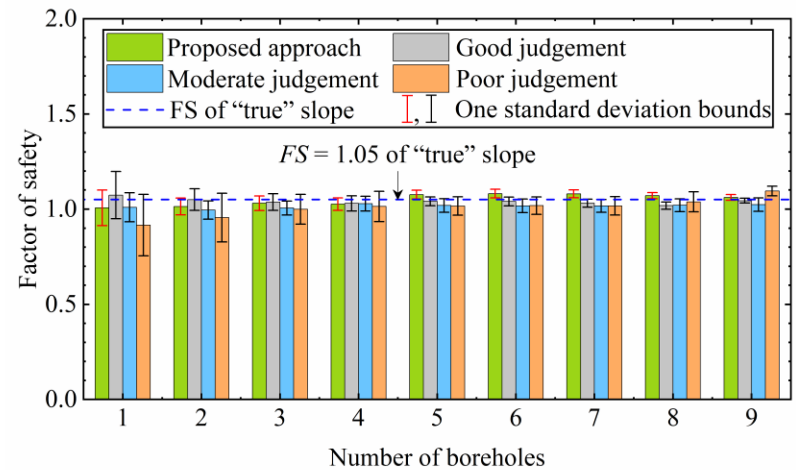


Figure 8. A comparison study of the estimated *FS* among different borehole patterns

The uncertainty of the location of the slip surface can be characterized by the area of the potential locations of the slip surface (Liu et al. 2017) or the uncertainty of the controlling points (Johari and Gholampour 2018; van den Eijnden and Hicks 2018). The latter is used in this study for its simplicity. As shown in Figure 9, there are three controlling points *A*, *B* and *C* at the slip surface of the slope. The leftmost and rightmost points *A* and *B* determine the range of the influence zone, which can be calculated by summing the horizontal distance from point *A* to slope crest d_{Begin} and that from point *B* to slope crest d_{End}. The deepest point *C* is related to the slope failure mechanism. If the vertical distance from point *C* to slope toe d_{Deep} is positive, a deep failure mode will occur for the slope. Otherwise, the shallow failure mode can be found. When the location of the slip surface is determined, the sliding volume can be easily calculated. With the adopted methods, the location of the slip surface of the "true" slope is derived as d_{Begin} = 8.5 m, d_{End} = 9.5 m, and d_{Deep} = 0.5 m, while the sliding volume of the "true" slope is V = 101.75 m^2/m. The estimated results of the location of the slip surface and sliding volume without boreholes are summarized in Figure 10. As shown in Figure 10, the estimated mean μ is far from the "true" value, and the standard deviation σ is large in the estimated results of the location of the slip surface and sliding volume, which indicates considerable estimate errors and uncertainties in the location of the slip surface when there is no site investigation effort (reflected by the total number of boreholes required).

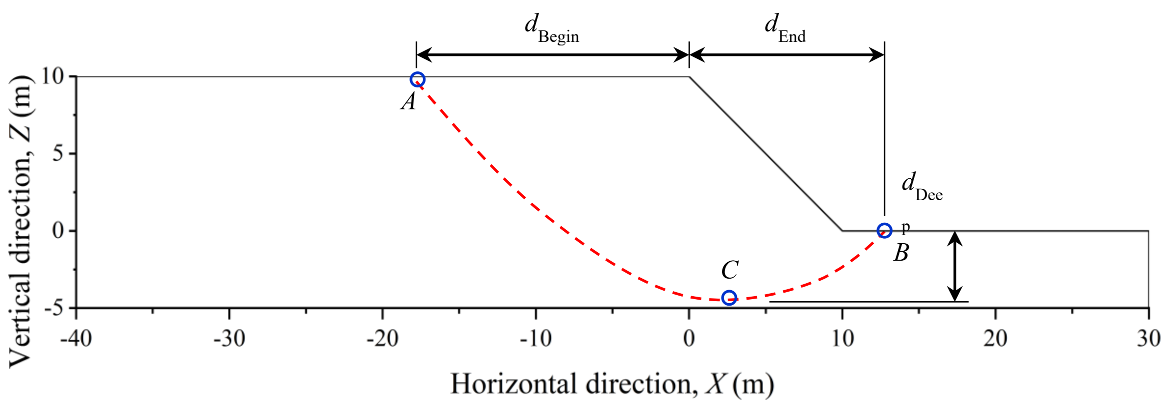
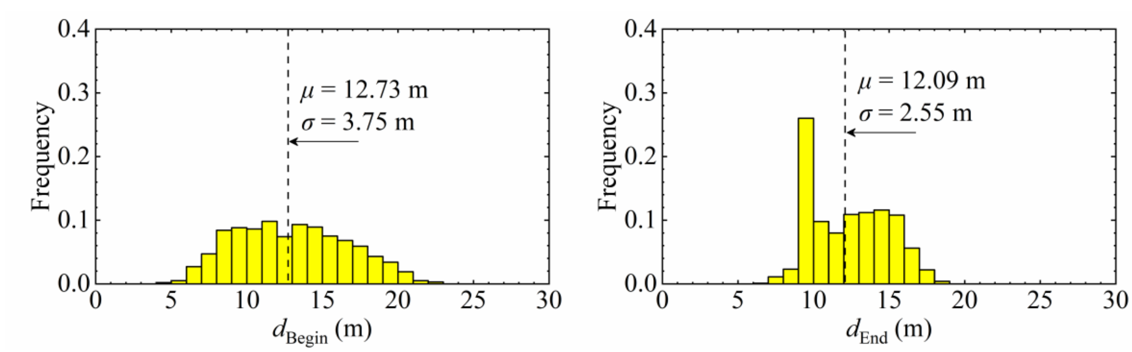
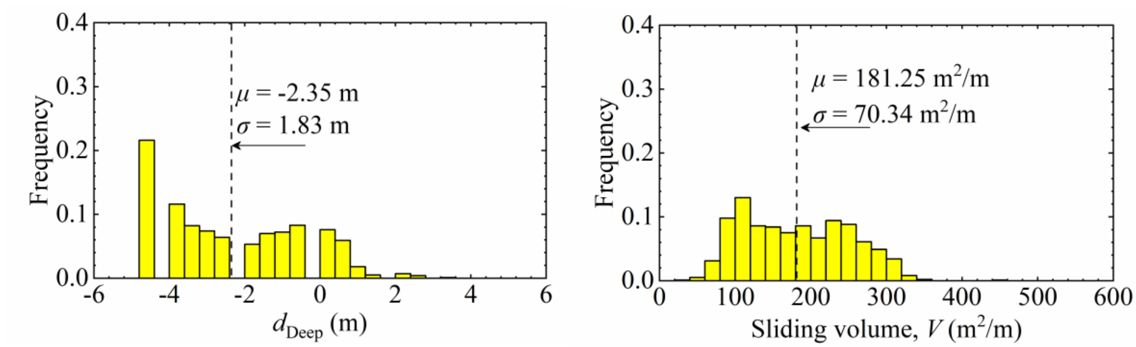


Figure 9. Location of the slip surface characterized using three controlling points



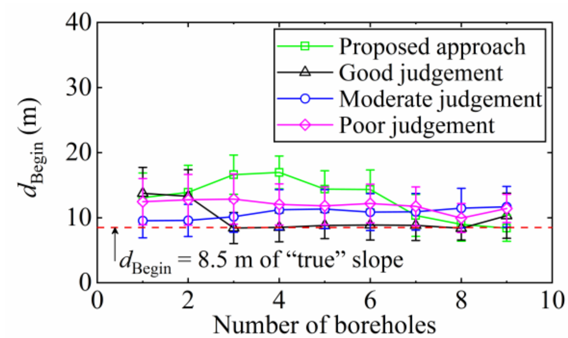
(a) Histogram of d_{Begin} (b) Histogram of d_{End}



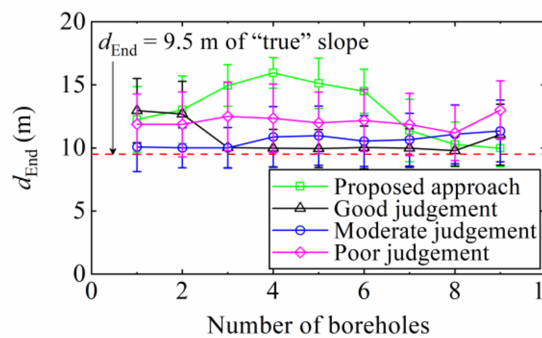
(c) Histogram of d_{Deep} (d) Histogram of sliding volume, V

Figure 10. Location of the slip surface and sliding volume estimated from unconditional random field simulations

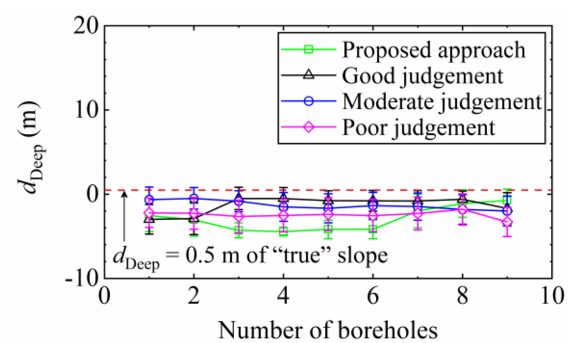
The comparison results of the estimate on the location of the slip surface and sliding volume are plotted in Figure 11. The borehole pattern with the traditional method from poor judgement causes significant deviation on the estimated slip surface (represented by three controlling points) and sliding volume. The deviation from the "true" value appears insensitive to the number of boreholes with poor judgement, since the location of the slip surface is not revealed by most boreholes from this method. The borehole pattern with moderate judgement provides a similar trend when there are 4 boreholes but with relatively less deviation. This result can be explained by the small space between every two boreholes, which makes the effect of the boreholes considerably overlapped. For borehole patterns with good judgement, very moderate effort (i.e., N_{BH} = 3) is sufficient to accurately characterize the location of the slip surface and sliding volume. Comparing the three borehole patterns from traditional methods, the borehole pattern from good judgement is the most cost-efficient. The difference among the three borehole patterns is the confidence in the estimated influence zone. Therefore, if the range of the influence zone is over- or underestimated, the borehole space may be problematically determined, and the increase in borehole number does not necessarily improve the accuracy in estimating the slip surface and sliding volume, which results in considerable challenges in the borehole configuration for the traditional method. For the proposed method, nine boreholes are required to accurately estimate the slip surface and sliding volume. The proposed method appears to require more effort than the results from good judgement. However, an implicit huge effort may be needed to guarantee the accurate estimation of the location of the slip surface for good judgement in practical slope problems, which is not evaluated in this comparison study. Therefore, the superiority of nonequal spacing planning of the site investigation based on the Spearman rank correlation coefficient is sufficiently demonstrated.



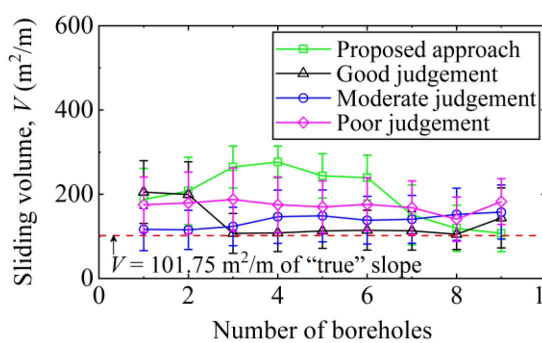
(a) d_{Begin}



(b) d_{End}



(c) d_{Deep}



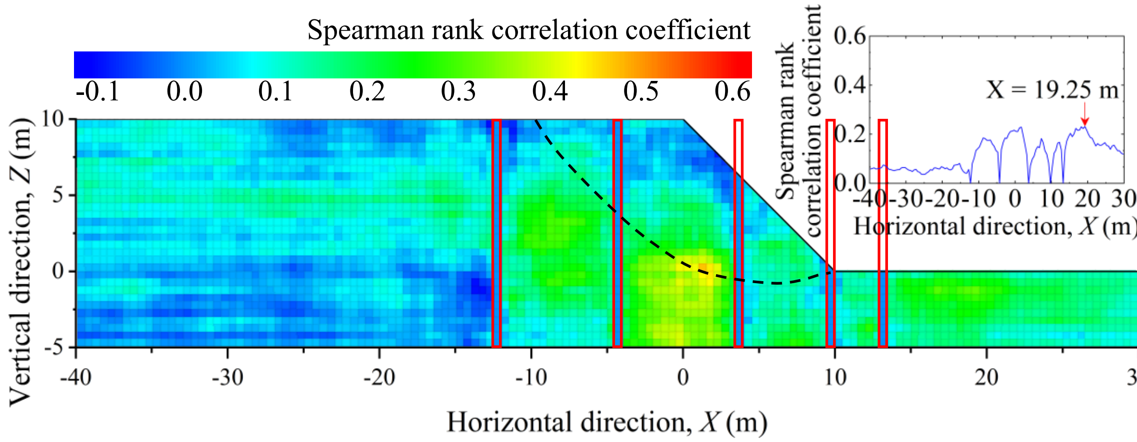
(d) Sliding volume, V

Figure 11. A comparison study of the characteristics of the slope with the proposed method

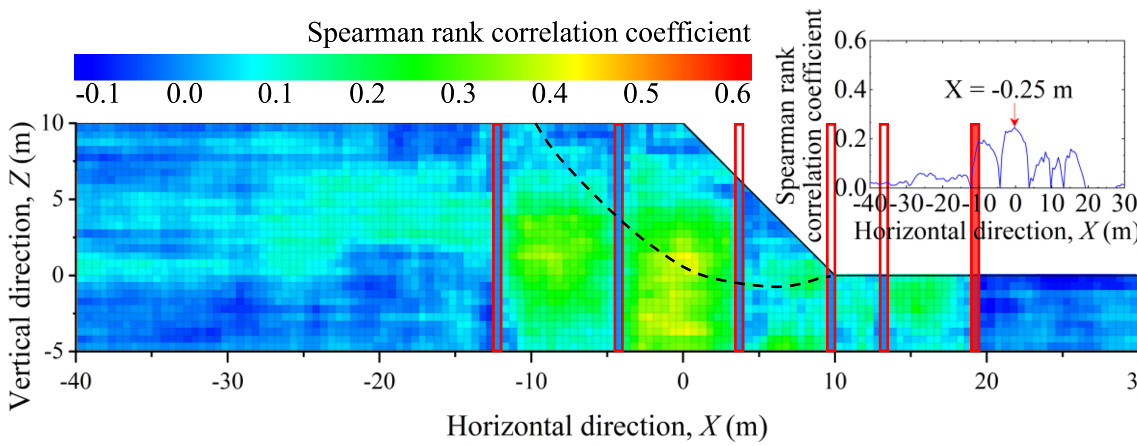
and traditional method

3.3 Discussion of the borehole pattern from the proposed method

As mentioned above, effective boreholes should be located to reveal the location of the slip surface, and the challenge for this objective is how to accurately estimate the range of the "true" sliding area. Figure 7 shows the final borehole pattern from the proposed method. It seems the fourth, the fifth and the sixth boreholes are ineffective in this borehole pattern, since the three boreholes are outside the "true" influence zone. However, the sliding area is well bracketed by the fourth and fifth boreholes. The two boreholes can be useful to estimate the influence zone, since a small error will be obtained. The effect of the sixth borehole can be illustrated by the change in contour of the Spearman rank correlation coefficient with different numbers of boreholes. The contour of the Spearman rank correlation coefficient with the fifth and sixth boreholes is plotted in Figure 12. As shown in Figure 12, with five boreholes applied to the slope, it is still difficult to examine if the area on the right of the fifth borehole is a low or high correlated soil zone because the large initial investigation area is adopted (from X = -40 m to X = 30 m). After applying the sixth borehole, this area is updated as a low-correlated soil zone. Therefore, the sixth borehole can be considered a part of the effort to automatically identify the influence zone. Figure 13 shows the contour of the Spearman rank correlation coefficient with nine boreholes from the proposed method. The final influence zone is derived as the area between the fifth and sixth boreholes, while the remaining area is a low-correlated soil zone. Therefore, it is effective to automatically estimate the influence zone for the slope with the proposed method, even if the initial investigation area is considerably conservatively selected.



(a) $N_{BH} = 5$



(b) $N_{BH} = 6$

Figure 12. Illustration of the effect of the sixth borehole from the proposed method

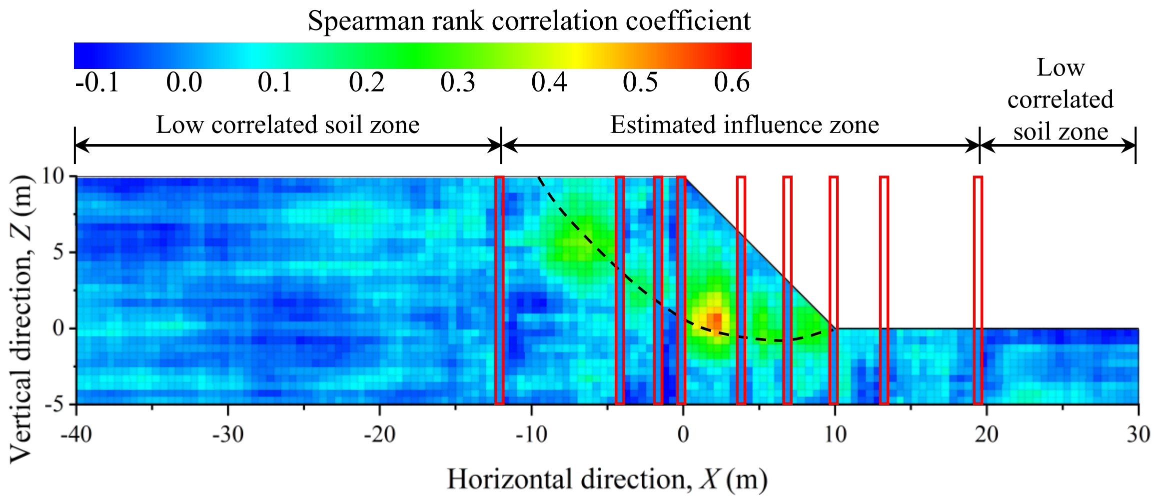


Figure 13. Estimated influence zone by the proposed method (N_{BH} = 9)

4. Robustness analysis and risk assessment

In addition to the characteristics of the slope, uncertainty reduction and risk reduction are aspects of slope design. In this section, robustness analysis and risk assessment are performed to validate the effectiveness of the proposed method in the comparison study.

4.1 Robustness analysis

The robustness is defined as the sensitivity of the system response to the variation in input parameters. The higher robustness of the geotechnical system implies that the system can better resist the uncertainty of input parameters. Various robustness measurements are formulated in geotechnical engineering, and the signal-to-noise ratio (SNR) is a commonly used robustness measurement for slope problems (Gong et al. 2015&2017&2020). In this paper, SNR is adopted to assess the robustness of the slope system, and a higher SNR value indicates a higher system robustness. The *FS*, location of the slip surface, and sliding volume are treated as the response of concern for the slope problem. The robustness of the *FS*, location of the slip surface, and sliding volume can be calculated using Eqs. (9-11), respectively. The robustness of the location of the slip surface is evaluated using the average SNR of the three controlling points of the slip surface. With the obtained SNR values for the *FS*, location of the slip surface and sliding volume, the robustness of the entire slope system can be calculated using Eq. (12) based on a weighted average of three SNR values from Eqs. (9-11) (Gong et al. 2017).

$$SNR_{FS} = 10\log_{10}\left(\frac{\mu_{FS}^2}{\sigma_{FS}^2}\right) \tag{9}$$

$$SNR_L = \frac{1}{3}(SNR_A + SNR_B + SNR_C) \tag{10a}$$

$$SNR_A = 10\log_{10}\left(\frac{\mu_{d_{\text{Begin}}}^2}{\sigma_{d_{\text{Begin}}}^2}\right) \tag{10b}$$

$$SNR_B = 10\log_{10}\left(\frac{\mu_{d_{\text{End}}}^2}{\sigma_{d_{\text{End}}}^2}\right) \tag{10c}$$

$$SNR_C = 10\log_{10}\left(\frac{\mu_{d_{\text{Deep}}}^2}{\sigma_{d_{\text{Deep}}}^2}\right) \tag{10d}$$

$$SNR_V = 10\log_{10}\left(\frac{\mu_V^2}{\sigma_V^2}\right) \tag{11}$$

$$S = w_{FS}SNR_{FS} + w_L SNR_L + w_V SNR_V \tag{12}$$

where SNR_{FS}, SNR_L, and SNR_V are the robustness for the estimation of the *FS*, location of slip surface (*L*) and sliding volume (*V*), respectively; SNR_A, SNR_B, and SNR_C are the robustness for the estimation of the horizontal distance of point *A* to the slope crest (d_{Begin}), horizontal distance of point *B* to the slope crest (d_{End}) and vertical distance of point *C* to the slope toe (d_{Deep}), respectively. The robustness estimation of the location of the slip surface can be characterized by the averaged SNR_A, SNR_B and SNR_C. The overall robustness *S* for the slope is a weighted summation of the three terms in Eqs. (9-11). The weighted factors w_{FS}, w_L and w_V may be determined by their corresponding contributions to the given geotechnical problems. Here, $w_{FS} = w_L = w_V = 1$ is assumed for simplicity (Gong et al. 2017).

The robustness analysis results for different borehole patterns are plotted in Figure 14. The highest overall SNR value S can be obtained from the proposed method, which indicates that this method finds the most robust estimated characteristics of the slope. The overall SNR values S from traditional methods with good judgement and poor judgement increase with the number of boreholes, while those from the traditional method with moderate judgement converge at $N_{\mathrm{BH}} = 3$. It seems that the borehole pattern with good judgement is the most effective in the robustness analysis, followed by the borehole pattern with moderate judgement, since all boreholes with good judgement are located at the influence zone and considerably reduce the uncertainty of the location of the slip surface and sliding volume. However, although the effect of the boreholes from moderate judgement is overlapped due to small borehole space and more additional boreholes cannot improve the robustness of the estimated results of the characteristics of the slope after $N_{\mathrm{BH}} = 3$, more investigation effort is required to reach the higher level of robustness for the slope system, which is consistent with the fact that most boreholes based on poor judgement are outside the "true" influence zone. Therefore, the superiority of the robustness performance with the proposed method is sufficiently demonstrated.

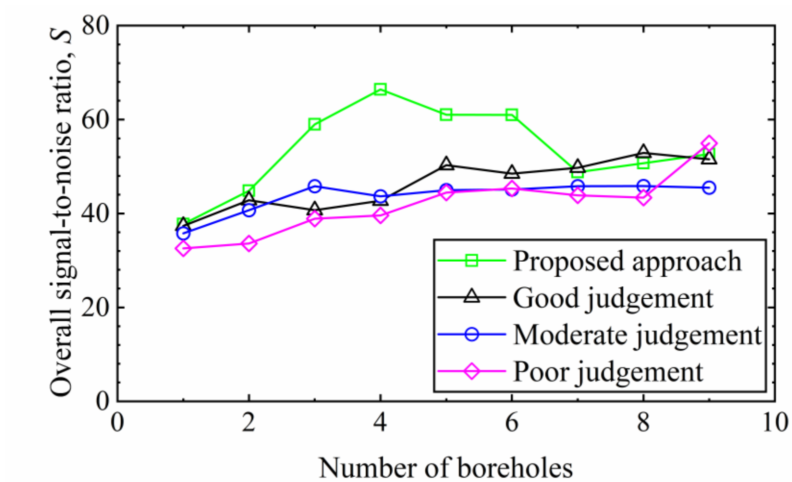


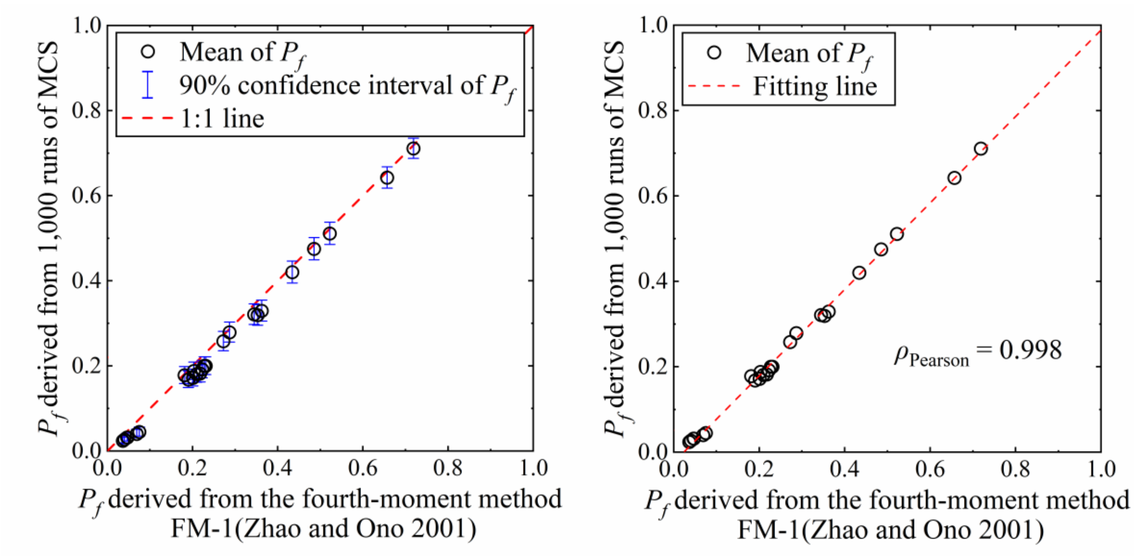
Figure 14. Overall robustness with the increase in number of boreholes

4.2 Risk assessment

Risk assessment can provide information for risk-informed decision-making. Herein, the risk assessment for site investigation of slope problems is conducted following Yang et al. (2019). According to Yang et al. (2019), the total loss cost C_{total} with different numbers of boreholes can be described as

$$C_{total} = N_{BH} \bullet C_{BH} + P_f \bullet C_{false} \tag{13}$$

where N_{BH} is the number of boreholes; C_{BH} is the average cost of one borehole; P_f is the probability of failure of the slope; the C_{false} is the of making a false decision. To improve the computational efficiency of the calculation of P_f, the MCS-based moment method FM-1 is adopted to estimate the probability of failure (Zhao and Ono 2001; Zhang et al. 2022). The MCS is first performed to derive the dimensionless moments of the limit state function that define the slope failure based on *FS*. Then, the formulas related to dimensionless moments are utilized to estimate the probability of failure. Detailed descriptions and formulas can be found in Zhao and Ono (2001) and Zhang et al. (2022). Figure 15 shows the validations of the moment method FM-1 to estimate the probability of failure in this slope example, where the analyses are performed for 25 slope scenarios with different numbers of boreholes. As shown in Figure 15(a), the probabilities of failure from the FM-1 moment method are well bracketed by the 90% confidence intervals of the probabilities of failure from the MCS. The Pearson correlation coefficient between P_f obtained from the MCS and that from the FM-1 moment method is 0.998 in Figure 15(b), which implies a strong linear correlation between them and validates the accuracy of the FM-1 moment method. The cost of each borehole is assumed to be C_{BH} = \$AUD 5,000, and the loss of making the false unsafe assessment of the slope stability is assumed to be C_{false} = \$AUD 150,000 following Yang et al. (2019), although the measured data may be suggested to be more acceptable for the risk assessment if there are available data.



(a) P_f from MCS versus that from FM-1 (b) Pearson correlation coefficient analysis

Figure 15. Validation of the moment method FM-1 (Zhao and Ono 2001) in estimating the probability of slope failure in this study

Figure 16 shows that the expected total loss cost first decreases and subsequently increases with the number of boreholes for the proposed method. A minimum expected loss cost of approximately \$AUD 25,000 with five boreholes ($N_{BH} = 5$) is obtained in the proposed method. From Eq. (13), the expected total loss cost consists of two parts: the cost of the boreholes and the expected loss cost of making a false decision. With the increase in number of boreholes applied to the slope, the mean *FS* will approach the "true" *FS,* and the standard deviation of *FS* will decrease. Since the "true" *FS* is 1.05, which is larger than 1, P_f will be close to 0 when sufficient boreholes are located at the slope, so the loss cost of making a false decision approaches 0. The total cost of the loss cost will be dominated by the cost of the boreholes. A minimum loss cost of approximately \$AUD 31,000 with the same site investigation effort ($N_{BH} = 5$) is reached in the borehole pattern with good judgement. Compared to the risk assessment results from good judgement, an approximate 19% loss cost can be avoided from the proposed method. A similar trend of the loss cost to the proposed method can be obtained in the traditional method with moderate judgement, and the minimum loss cost is reached with four boreholes ($N_{BH} = 4$). However, the loss cost from moderate judgement is higher than these two methods, since a small borehole space results in the overlapped effect of the applied boreholes. For poor judgement, the total cost decreases with the increase in number of boreholes, but it generally yields the highest amount of total cost, which can be explained by the fact that the uncertainty of *FS* cannot be effectively reduced based on the method from poor judgement (see Figure 8). Therefore, the advantage of the proposed method in risk reduction is sufficiently validated.

The estimated accuracy of the characteristics of the slope, robustness of the estimated results and risk reduction of the proposed method are evaluated in an undrained slope example. It is suggested to adopt five boreholes (N_{BH} = 5) for the slope problem when applying the proposed method, since the overall performance of the slope system can be well evaluated from many perspectives. For instance, when five boreholes are configured in this slope, the "true" influence zone can be estimated with tolerable error, and the risk reduction is maximal. For the robustness analysis, the borehole pattern with N_{BH} = 5 is the second highest cost-effective borehole pattern in Figure 14. Thus, N_{BH} = 5 is considered the optimal number for the proposed method based on the gain-sacrifice relationship between cost and investigation effort.

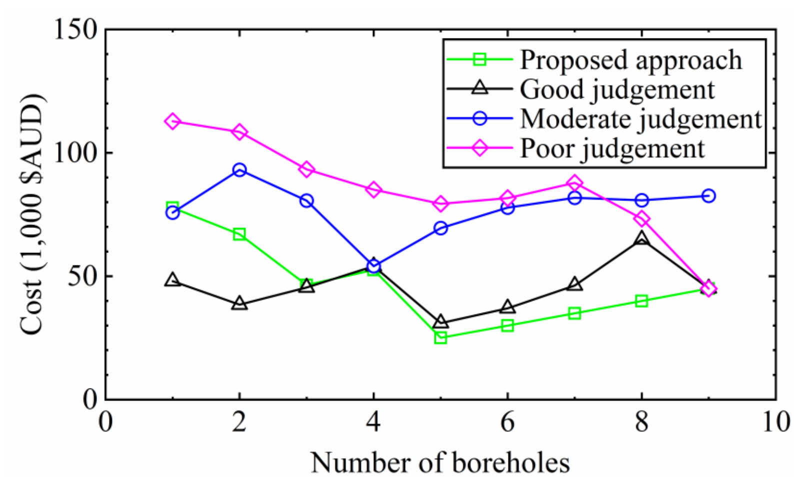


Figure 16. A comparison study of risk assessment of different borehole patterns

5. Summary and conclusions

This paper proposed an optimization method for geotechnical site investigation to minimize the risk and associated site investigation effort and maximize the robustness of the slope system. The proposed method can optimize the location and number of boreholes without prior knowledge about the slip surface, which results in adaptive patterns of borehole planning based on the Spearman rank correlation coefficient with unequal space for a given slope site. Compared to the traditional method, the advantages of the proposed method are:

1) The location and number of boreholes can be optimized considering the synthesized system responses (e.g., *FS*, location of slip surface and sliding volume, robustness, and risk) in the proposed method. The proposed method accurately estimates *FS*, the location of the slip surface, and the sliding volume if sufficient boreholes are applied.

2) The proposed method can similarly reduce the uncertainty of *FS* compared with the traditional method with good judgement and tends to obtain a more robust site investigation program than traditional borehole patterns.

3) The proposed method minimizes the risk with the optimized number of boreholes. The effectiveness of this optimized borehole pattern on the estimate of the range of influence zone and robustness of the slope system can also be reached in this scenario.

4) The proposed method is straightforward and easy to implement to automatically identify the range of the sliding area with unequally spaced borehole patterns, which provides a reference to build an adaptive unequally spaced borehole pattern without prior knowledge about the slip surface in practice.

CRediT authorship contribution statement

Liang Zhang: Conceptualization, Methodology, Software, Visualization, Writing – original draft. **Lei Wang:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The study on which this paper is based was partially supported by the National Science Foundation through Grants 1818649 and 1900445. The results and opinions expressed in this paper do not necessarily reflect the views and policies of the National Science Foundation.

References

Bishara, A.J., & Hittner, J.B. (2015). Reducing bias and error in the correlation coefficient due to nonnormality. *Educational and Psychological Measurement*, 75(5), 785-804.

Cai, Y., Li, J., Li, X., Li, D., & Zhang, L. (2019). Estimating soil resistance at unsampled locations based on limited CPT data. *Bulletin of Engineering Geology and the Environment*, 78(5), 3637-3648.

Chen, F., & Zhang, W. (2021). Influence of spatial variability on the uniaxial compressive responses of rock pillar based on 3D random field. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 7(3), 04021035.

Chen, G., Zhu, J., Qiang, M., & Gong, W. (2018). Three-dimensional site characterization with borehole data–a case study of Suzhou area. *Engineering Geology*, 234, 65-82.

Chwała, M. (2021). Optimal placement of two soil soundings for rectangular footings. *Journal of Rock Mechanics and Geotechnical Engineering*, 13(3), 603-611.

De Winter, J.C., Gosling, S.D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods*, 21(3), 273.

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

Fenton, G.A. (1994). Error evaluation of three random-field generators. *Journal of Engineering Mechanics*, 120(12), 2478-2497.

Gong, W., Luo, Z., Juang, C.H., Huang, H., Zhang, J., & Wang, L. (2014). Optimization of site exploration program for improved prediction of tunneling-induced ground settlement in clays. *Computers and Geotechnics*, 56, 69-79.

Gong, W., Wang, L., Khoshnevisan, S., Juang, C.H., Huang, H., & Zhang, J. (2015). Robust geotechnical design of earth slopes using fuzzy sets. *Journal of Geotechnical and Geoenvironmental Engineering*, 141(1), 04014084.

Gong, W., Tien, Y.M., Juang, C.H., Martin, J.R., & Luo, Z. (2017). Optimization of site investigation program for improved statistical characterization of geotechnical property based on random field theory. *Bulletin of Engineering Geology and the Environment*, 76(3), 1021-1035.

Gong, W., Juang, C.H., Martin II, J.R., Tang, H., Wang, Q., & Huang, H. (2018). Probabilistic analysis of tunnel longitudinal performance based upon conditional random field simulation of soil properties. *Tunnelling and Underground Space Technology*, 73, 1-14.

Gong, W., Tang, H., Juang, C.H., & Wang, L. (2020). Optimization design of stabilizing piles in slopes considering spatial variability. *Acta Geotechnica*, 15(11), 3243-3259.

Han, L., Wang, L., Zhang, W., Geng, B., & Li, S. (2022). Rockhead profile simulation using an improved generation method of conditional random field. *Journal of Rock Mechanics and Geotechnical Engineering*, 14(3), 896-908.

Hicks, M.A., Nuttall, J.D., & Chen, J. (2014). Influence of heterogeneity on 3D slope reliability and failure consequence. *Computers and Geotechnics*, 61, 198-208.

Huang, J., and Griffiths, D.V. (2015). Determining an appropriate finite element size for modeling the strength of undrained random soils. *Computers and Geotechnics*, 69, 506-513.

Huang, L., Cheng, Y.M., Leung, Y.F., & Li, L. (2019). Influence of rotated anisotropy on slope reliability evaluation using conditional random field. *Computers and Geotechnics*, 115, 103133.

Huang, L., Huang, S., & Lai, Z. (2020). On the optimization of site investigation programs using centroidal Voronoi tessellation and random field theory. *Computers and Geotechnics*, 118, 103331.

Itasca Consulting Group, Inc. (2022). FLAC3D - Fast Lagrangian Analysis of Continua in 3 Dimensions, Version 7.0. Minneapolis: Itasca.

Jiang, S.H., Huang, J., Huang, F., Yang, J., Yao, C., & Zhou, C.B. (2018a). Modelling of spatial variability of soil undrained shear strength by conditional random fields for slope reliability analysis. *Applied Mathematical Modelling*, 63, 374-389.

Jiang, S.H., Papaioannou, I., & Straub, D. (2018b). Bayesian updating of slope reliability in spatially variable soils with in-situ measurements. *Engineering Geology*, 239, 310-320.

Jiang, S.H., Papaioannou, I., & Straub, D. (2020). Optimization of site-exploration programs for slope-reliability assessment. *ASCE-ASME Journal of Risk and Uncertainty in*

Engineering Systems, Part A: Civil Engineering, 6(1), 04020004.

Johari, A., & Gholampour, A. (2018). A practical approach for reliability analysis of unsaturated slope by conditional random finite element method. *Computers and Geotechnics*, 102, 79-91.

Johari, A., & Fooladi, H. (2020). Comparative study of stochastic slope stability analysis based on conditional and unconditional random field. *Computers and Geotechnics*, 125, 103707.

Knabe, W., Przewłócki, J., and Różyński, G. (1998). Spatial averages for linear elements for two-parameter random field. *Probabilistic Engineering Mechanics*, 13(3), 147-167.

Li, D.Q., Qi, X.H., Cao, Z.J., Tang, X.S., Phoon, K.K., & Zhou, C.B. (2016a). Evaluating slope stability uncertainty using coupled Markov chain. *Computers and Geotechnics*, 73, 72-82.

Li, Y.J., Hicks, M.A., & Vardon, P.J. (2016b). Uncertainty reduction and sampling efficiency in slope designs using 3D conditional random fields. *Computers and Geotechnics*, 79, 159-172.

Li, X.Y., Zhang, L.M., & Li, J.H. (2016c). Using conditioned random field to characterize the variability of geologic profiles. *Journal of Geotechnical and Geoenvironmental Engineering*, 142(4), 04015096.

Liu, L.L., Cheng, Y.M., & Zhang, S.H. (2017). Conditional random field reliability analysis of a cohesion-frictional slope. *Computers and Geotechnics*, 82, 173-186.

Liu, L.L., Cheng, Y.M., Pan, Q.J., & Dias, D. (2020). Incorporating stratigraphic boundary uncertainty into reliability analysis of slopes in spatially variable soils using one-dimensional conditional Markov chain model. *Computers and Geotechnics*, 118, 103321.

Lloret-Cabot, M.H.M.A., Hicks, M.A., & van den Eijnden, A.P. (2012). Investigation of the reduction in uncertainty due to soil variability when conditioning a random field using Kriging. *Géotechnique letters*, 2(3), 123-127.

Shen, M., Chen, Q., Juang, C.H., Gong, W., & Tan, X. (2018). Bi-objective optimization of site investigation program for liquefaction hazard mapping. *In GeoShanghai International Conference* (pp. 86-93). Springer, Singapore.

Thirumalai, C., Chandhini, S.A., & Vaishnavi, M. (2017). Analysing the concrete compressive strength using Pearson and Spearman. *In 2017 international conference of Electronics, Communication and Aerospace Technology* (*iCECA*) (Vol. 2, pp. 215-218). IEEE.

van den Eijnden, A.P., & Hicks, M.A. (2018). Probability-dependent failure modes of slopes and cuts in heterogeneous cohesive soils. *Géotechnique Letters*, 8(3), 214-218.

Xiao, T., Li, D.Q., Cao, Z.J., & Zhang, L.M. (2018). CPT-based probabilistic characterization of three-dimensional spatial variability using MLE. *Journal of Geotechnical and Geoenvironmental Engineering*, 144(5), 04018023.

Yang, Z., & Ching, J. (2021). Simulation of three-dimensional random field conditioning on incomplete site data. *Engineering Geology*, 281, 105987.

Yang, R., Huang, J., Griffiths, D.V., Meng, J., & Fenton, G.A. (2019). Optimal geotechnical site investigations for slope design. *Computers and Geotechnics*, 114, 103111.

Yang, R., Huang, J., & Griffiths, D.V. (2022). Optimal geotechnical site investigations for slope reliability assessment considering measurement errors. *Engineering Geology*, 297, 106497.

Wang, J., Yang, R., & Feng, Y. (2017). Spatial variability of reconstructed soil properties and the optimization of sampling number for reclaimed land monitoring in an opencast coal mine. *Arabian Journal of Geosciences*, 10(2), 1-13.

Wang, Y., Huang, J., and Tang, H. (2020). Automatic identification of the critical slip surface of slopes. *Engineering Geology*, 273, 105672.

Zhang, L., Gong, W.P., Li, X.X., Tan, X.H., Zhao, C., & Wang, L. (2022). A comparison study between 2D and 3D slope stability analyses considering spatial soil variability. *Journal of Zhejiang University-SCIENCE A*, 23(3), 208-224.

Zhao, T., Wang, Y., & Xu, L. (2021). Efficient CPT locations for characterizing spatial variability of soil properties within a multilayer vertical cross-section using information entropy and Bayesian compressive sensing. *Computers and Geotechnics*, 137, 104260.

Zhao, Y.G., and Ono, T. (2001). Moment methods for structural reliability. *Structural Safety*, 23(1), 47-75.

Lists of Tables

Lists of Figures