Probabilistic Stratification Modeling in Geotechnical Site Characterization

Te Xiao, S.M.ASCE[1]; Li-Min Zhang, Ph.D., F.ASCE[2]; Xue-You Li, Ph.D.[3]; and Dian-Qing Li, Ph.D., M.ASCE[4]

Abstract: Stratification in geologic profiles is one of the most significant uncertainties in geotechnical site characterization. In this paper, a three-level probabilistic framework is proposed for geotechnical stratification modeling considering stratigraphic uncertainty. The framework consists of model parameter identification, conditional simulation, and stratigraphic uncertainty quantification. Both boundary-based and category-based stratigraphic models are adopted in the framework, and a heuristic combination model is further recommended to combine the advantages of the boundary-based and category-based models. The geological stratification at a construction site in Hong Kong is characterized to illustrate the probabilistic framework. Results indicate that probabilistic stratification modeling quantifies stratigraphic uncertainty in a rigorous manner. Additionally, the heuristic combination model has the ability to generate almost arbitrary geotechnical strata and to account for material spatial distribution trends and engineering judgment to a certain degree. **DOI:** 10.1061/AJRUA6.0000924.

Author keywords: Geotechnical site characterization; Random field; Uncertainty analysis; Model identification; Reliability analysis.

Introduction

Geological materials are naturally heterogeneous, which is of great importance to the design of geotechnical structures (Vardanega and Bolton 2016). This 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. Because the number of boreholes or other site investigation points is often limited, many uncertainties are unavoidably involved in characterizing a heterogeneous site. In the geotechnical reliability community, the description of inherent spatial variability has gained increasing interest in recent years (Baecher and Christian 2003; Fenton and Griffiths 2008; Zhu and Zhang 2013; Li et al. 2016b; Xiao et al. 2016, 2017; Zhu et al. 2017), whereas studies on stratigraphic uncertainty are still limited. Although various probabilistic approaches have been developed for strata identification (Phoon et al. 2003; Liao and Mayne 2007; Cao and Wang 2013; Houlsby and Houlsby 2013; Wang et al. 2014; Ching et al. 2015), the majority focus on strata within a single sounding or borehole; the prediction of stratification at unsampled locations is still a difficult issue.

In practice, geotechnical engineers are more concerned about stratification of the whole site. To achieve this goal, the strata for every sampled location and a stratigraphic model linking the information between sampled and unsampled locations are both required. The former can be obtained directly if borehole logs are available; otherwise, the aforementioned probabilistic approaches can be used with other indirect investigation information. With respect to the stratigraphic model, two representative models are available in the literature, as shown in Fig. 1. The first is a boundary-based model (Nobre and Sykes 1992; Sitharam et al. 2008; Zhang and Dasaka 2010; Dasaka and Zhang 2012; Schöbi and Sudret 2015; Li et al. 2016c, d). 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. Fig. 2 is an example of stratification in a geologic profile (Liu et al. 2017). Some material boundaries are determined with high confidence; others, with high uncertainty. This uncertainty, known as the stratigraphic uncertainty, is determined by engineering judgment and usually ignored in conventional deterministic stratification modeling.

The second stratigraphic model is a category-based model (Elfeki and Dekking 2001, 2005; Li 2007; He et al. 2009; Park 2010; Liang et al. 2014; Li et al. 2016a, e; 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.

This study aims to develop a probabilistic framework for stratification modeling in geotechnical site characterization, consisting of model parameter identification, conditional simulation, and stratigraphic uncertainty quantification. Both boundary-based and category-based stratigraphic models are adopted in the framework, and their pros and cons are compared through an example in

[1]Ph.D. Candidate, State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan Univ., 8 Donghu South Rd., Wuhan 430072, P.R. China. E-mail: short_xiaote@whu.edu.cn

[2]Professor, Dept. of Civil and Environmental Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (corresponding author). E-mail: cezhangl@ust.hk

[3]Postdoctoral Fellow, Dept. of Civil and Environmental Engineering, Hong Kong Univ. of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. E-mail: lixueyou@ust.hk

[4]Professor, State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan Univ., 8 Donghu South Rd., Wuhan 430072, P.R. China. E-mail: dianqing@whu.edu.cn

Note. This manuscript was submitted on December 20, 2016; approved on April 17, 2017; published online on July 22, 2017. Discussion period open until December 22, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering***, © ASCE, ISSN 2376-7642.

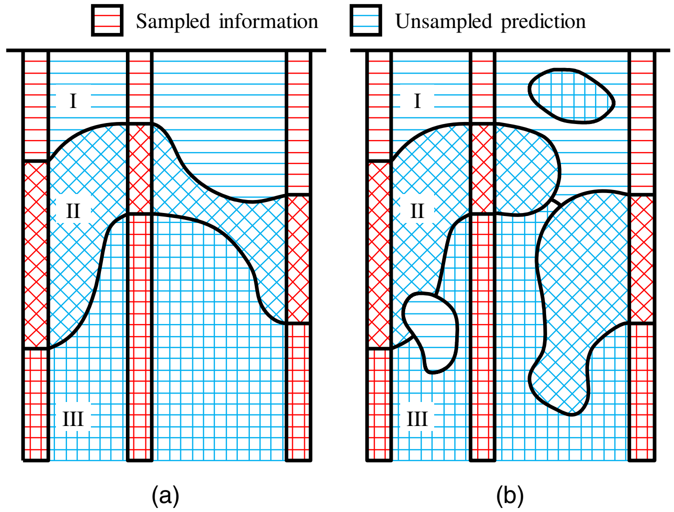

Fig. 1. Two models for geological stratification modeling: (a) boundary-based model; (b) category-based model

Hong Kong. Finally, a heuristic combination model is proposed to combine the advantages of both models for probabilistic stratification modeling.

Probabilistic Stratification Modeling Framework

In conventional deterministic stratification modeling, material zones are divided by several straight boundaries without considering any uncertainty. In contrast, probabilistic stratification modeling quantifies stratigraphic uncertainty in a rigorous manner and can be integrated with other sources of uncertainty and incorporated into geotechnical reliability analysis (Schöbi and Sudret 2015; Li et al. 2016a) and reliability-based design (Phoon et al. 2016).

The full probabilistic stratification modeling framework consists of three components: model parameter identification, conditional simulation, and stratigraphic uncertainty quantification. Herein, conditional simulation involves generating a series of possible geotechnical strata whose simulated values (e.g., boundary depths or material categories) at all sampled locations are consistent with observed values. Conditional simulation helps maximize the use of site investigation information and reduce stratigraphic uncertainty. Either a boundary-based stratigraphic model [Fig. 1(a)] or a category-based stratigraphic model [Fig. 1(b)] can be adopted in the framework, as described in the following sections.

Boundary-Based Model

In the context of the boundary-based model, the depth of the material boundary is usually considered to be spatially correlated and described by a random field with several parameters, $\boldsymbol{\theta}$, such as trend ($\boldsymbol{\beta}$), standard deviation (σ), and scale of fluctuation (δ). For a given set of $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma, \delta)$, the likelihood of observing the sequence of observations $\boldsymbol{X}$ (i.e., depths at sampled locations) can be written as (Baecher and Christian 2003)

$$L(\boldsymbol{X}|\boldsymbol{\theta}) = (2\pi\sigma^2)^{-n/2}|\boldsymbol{R}|^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{X}-\boldsymbol{F\beta})^T\boldsymbol{R}^{-1}(\boldsymbol{X}-\boldsymbol{F\beta})\right] \quad (1)$$

where n = number of observations; $\boldsymbol{F}$ and $\boldsymbol{\beta}$ = trend function matrix and coefficient vector, respectively (considering a linear trend function $X = a_1z + a_0$, for example, $\boldsymbol{F}$ is a $n \times 2$ matrix with n rows of $[z_i, 1]$, in which z_i is the location of the ith observation, and $\boldsymbol{\beta} = [a_1, a_0]^T$); and $\boldsymbol{R} = [\rho_{ij}]$ = correlation matrix [ρ_{ij} can be calculated according to δ and $\Delta_{ij} = |z_i - z_j|$ by a prescribed correlation function, such as the squared exponential correlation function $\rho_{ij} = \exp(-\pi\Delta_{ij}^2/\delta^2)$].

Although $\boldsymbol{\theta}$ can be directly determined by maximizing the likelihood function $L(\boldsymbol{X}|\boldsymbol{\theta})$, experience indicates that $L(\boldsymbol{X}|\boldsymbol{\theta})$ is inclined to be unidentifiable when the number of observations is limited and δ is small. To address this issue, prior information about $\boldsymbol{\theta}$ is introduced to facilitate parameter identification through a Bayesian approach (Wang et al. 2016a, b):

$$f(\boldsymbol{\theta}|\boldsymbol{X}) = C \times L(\boldsymbol{X}|\boldsymbol{\theta}) \times f(\boldsymbol{\theta}) \quad (2)$$

where C = normalizing constant to make the probability distribution valid; $f(\boldsymbol{\theta})$ = prior distribution of $\boldsymbol{\theta}$ according to engineering experience; and $f(\boldsymbol{\theta}|\boldsymbol{X})$ = posterior distribution of $\boldsymbol{\theta}$ given the observations.

After identifying the model parameters, the kriging approach (Forrester et al. 2008; Schöbi et al. 2017), which is a widely used spatial interpolation technique, can be adopted to predict boundary depths at unsampled locations. Although kriging cannot be used for conditional simulation directly (i.e., no random variations),

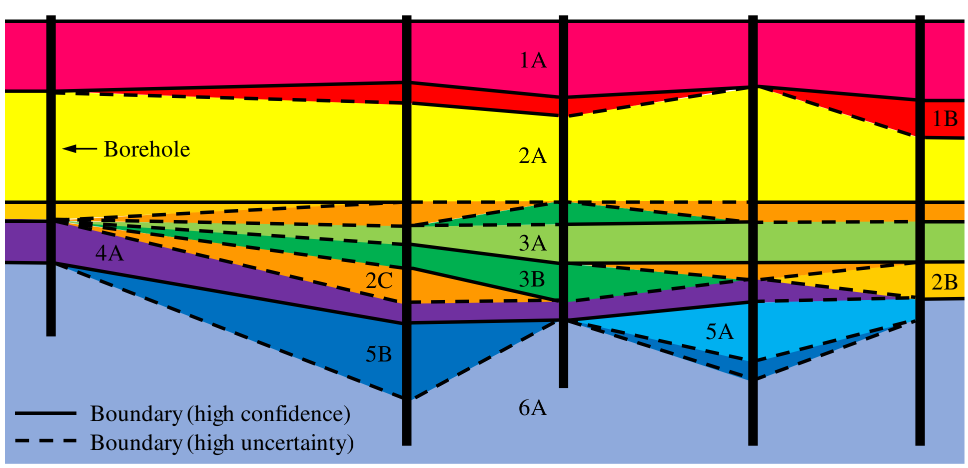

Fig. 2. Example of stratification in a geologic profile (adapted from Liu et al. 2017)

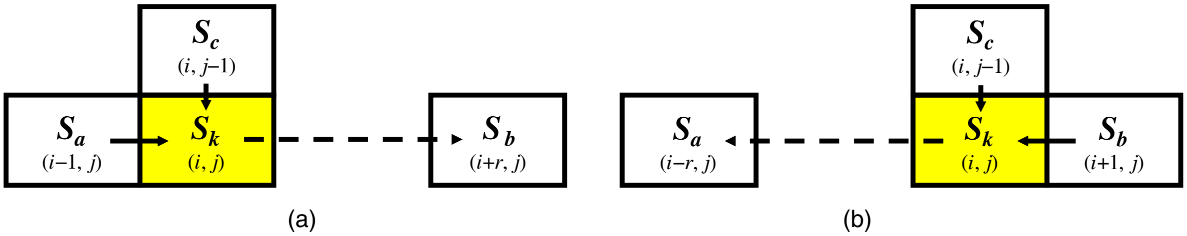

Fig. 3. Transition direction in the coupled Markov chain model: (a) forward; (b) backward

conditional simulation can be realized through a superposition algorithm (Fenton and Griffiths 2008):

$$\boldsymbol{X}_c(\boldsymbol{z}) = \boldsymbol{X}_u(\boldsymbol{z}) + \boldsymbol{X}_m(\boldsymbol{z}) - \boldsymbol{X}_s(\boldsymbol{z}) \tag{3}$$

where $\boldsymbol{z} = (z_1, z_2)$ = location, among which z_1 and z_2 are sampled and unsampled, respectively; $\boldsymbol{X}_c$ = conditional simulation; $\boldsymbol{X}_u$ = unconditional simulation generated by various random field generation techniques (Fenton and Griffiths 2008); $\boldsymbol{X}_m$ = kriging prediction based on measured values at location z_1; and $\boldsymbol{X}_s$ = kriging prediction based on unconditional simulation values at location z_1.

The uncertainty associated with the boundary can be quantified analytically, with its mean and variance estimated as (Forrester et al. 2008)

$$E[\boldsymbol{X}_c(z_2)] = \boldsymbol{F}_2\boldsymbol{\beta} + \boldsymbol{R}_{21}\boldsymbol{R}_{11}^{-1}[\boldsymbol{X}_m(z_1) - \boldsymbol{F}_1\boldsymbol{\beta}] \tag{4}$$

$$\mathrm{Var}[\boldsymbol{X}_c(z_2)] = \sigma^2 \times diag(\boldsymbol{R}_{22} - \boldsymbol{R}_{21}\boldsymbol{R}_{11}^{-1}\boldsymbol{R}_{12}) \tag{5}$$

where the subscripts 1 and 2 correspond to z_1 and z_2, respectively; and $diag(\cdot)$ = diagonal elements of a matrix.

The uncertainty depends not only on the identified model parameters (i.e., σ and δ) but also on the sampled locations. It reflects the degree of belief in boundary identification.

Category-Based Model

The category-based model assumes that geotechnical stratification is formed by material transition in space. Compared with the boundary-based model, it is more in accordance with the natural weathering or deposition process. The transition from one material category to another is typically characterized by a transition matrix or transiograms. Several category-based models have been developed, such as coupled Markov chain (Elfeki and Dekking 2001), Markov chain random field (Li 2007), and indicator kriging (He et al. 2009). In this study, the coupled Markov chain model is applied to category-based probabilistic stratification modeling for illustration because of its ability to capture the naturally anisotropic material transition in vertical and horizontal directions. Other category-based models can be applied to the framework also if more complicated anisotropy is observed (Li et al. 2016e).

The coupled Markov chain model is an inherent conditional simulation approach in that material transition initially occurs from the observed states. Herein, a state in the coupled Markov chain model represents a material category. It assumes that (1) the material transition follows the first-order Markovian property—that is, the current state depends only on the previous one state; and (2) the total transition in space can be considered a coupling of independent horizontal transition and vertical transition. With additional conditioning on a future state, as shown in Fig. 3, the total transition equations can be written as

$$\begin{aligned} P[Y(i,j) = S_k | Y(i-1,j) = S_a, Y(i+r,j) \\ = S_b, Y(i,j-1) = S_c] = C \times p_{ak}^h \times p_{kb}^{h(r)} \times p_{ck}^v \end{aligned} \tag{6a}$$

$$\begin{aligned} P[Y(i,j) = S_k | Y(i-r,j) = S_a, Y(i+1,j) \\ = S_b, Y(i,j-1) = S_c] = C \times p_{ka}^{h(r)} \times p_{bk}^h \times p_{ck}^v \end{aligned} \tag{6b}$$

where $Y(z)$ = material category at location z [e.g., $z = (i,j)$ for a two-dimensional (2D) numbered lattice]; S_k, S_a, S_b, S_c = material categories; C = normalizing constant; and p = transition probability obtained from transition probability matrix [e.g., $p_{kb}^{h(r)}$ is the probability of the r-step horizontal transition from material category S_k to material category S_b, and p_{ck}^v is the probability of the one-step vertical transition from material category S_c to material category S_k].

The vertical transition always occurs from top to bottom, similar to the geotechnical weathering process, and the direction of the horizontal transition can be either forward [Eq. (6a) and Fig. 3(a)] or backward [Eq. (6b) and Fig. 3(b)]. To avoid the directional effect caused by uniform horizontal transition in the coupled Markov chain model, the horizontal transition direction in each step is randomly determined in this study. An example for a random three-step horizontal transition (e.g., forward, backward, and backward) is shown in Fig. 4.

Characterizing the vertical and horizontal transition probability matrices is a primary task before using the coupled Markov chain model. The vertical transition matrix can be estimated directly by statistical analysis based on borehole data. In contrast, the horizontal transition matrix is more difficult to obtain because of discontinuous sampling in the horizontal direction. According to the assumption of Walther's law (Elfeki and Dekking 2005), the horizontal transition may undergo a similar process but at a larger scale compared with the vertical transition. Mathematically, this implies that the horizontal and vertical transition probability matrices may satisfy $p_{kk}^h = \eta p_{kk}^v / K$ and $p_{ki}^h = p_{ki}^v / K$, where $K = \eta p_{kk}^v + 1 - p_{kk}^v$ and η is a scaling factor larger than 1. By this means, the task of model parameter identification is significantly simplified to find the most probable value of η in the coupled Markov chain model. Given the transition probability matrices and two boundary boreholes, the likelihood of observing the sequence of $\boldsymbol{Y}$ in an intermediate borehole can be written as (Qi et al. 2016)

$$L(\boldsymbol{Y}|\eta) = \prod_k P[Y(z_k) = S_{mk}|_\eta] \tag{7}$$

where z_k = kth location; and S_{mk} = measured material category at z_k.

To estimate the probability $P[Y(z_k) = S_{mk}|_\eta]$, the conditional simulation based on Eq. (6) is repeatedly performed in a Monte Carlo simulation manner. The scaling factor η is then determined by maximizing the likelihood $L(\boldsymbol{Y}|\eta)$. Unlike in the boundary-based model, an analytical solution for category-based stratigraphic

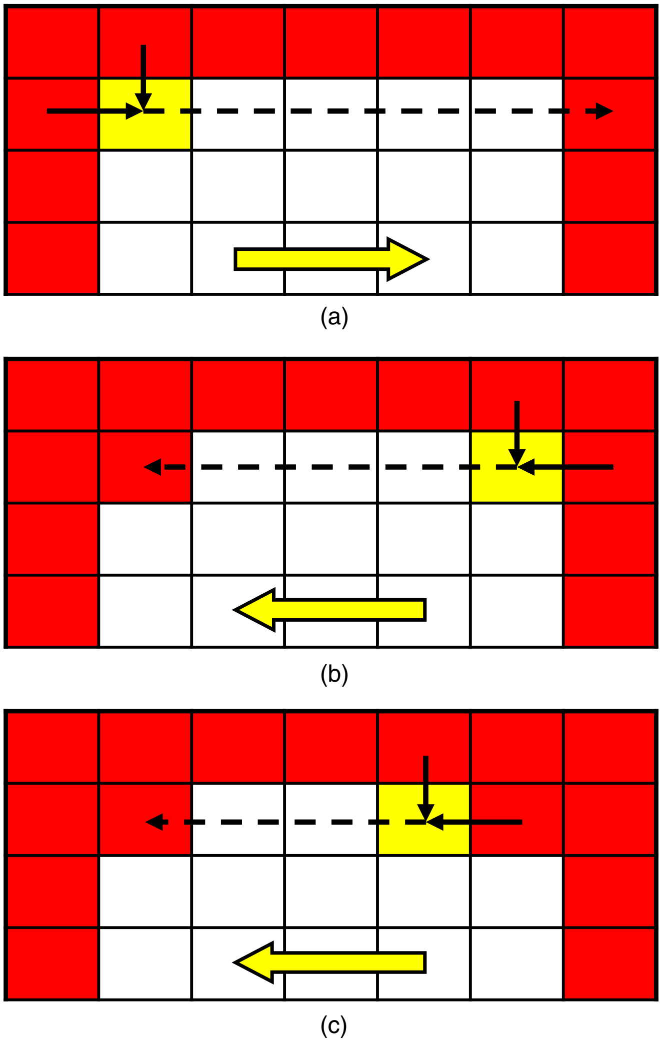

Fig. 4. Examples of random horizontal transition: (a) forward; (b) backward; (c) backward again

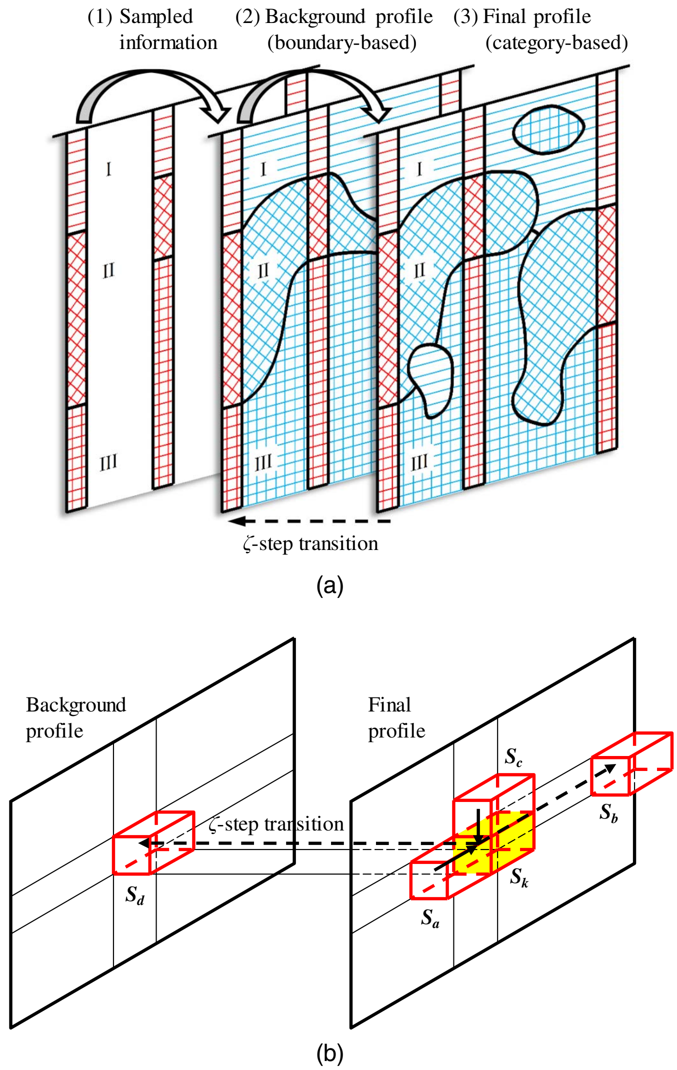

Fig. 5. Heuristic combination model: (a) introduction of hypothetical background profile; (b) material transition (forward)

uncertainty estimation does not exist. Instead, the Monte Carlo simulation and the entropy theory (Elfeki and Dekking 2005) can be adopted to estimate the occurrence probability of different material categories in space.

Heuristic Combination Model

Both boundary-based and category-based models can be applied in the probabilistic stratification modeling framework. Each model has advantages and disadvantages. The boundary-based model is applicable to the uniform stratification case. It is conceptually simple and can be easily integrated with engineering experience, and it is considered an extension of the deterministic stratification modeling that engineers are more familiar with. The boundary-based model has the ability to incorporate the material spatial distribution trend and to analytically estimate stratigraphic uncertainty. It can also be easily extended to three-dimensional (3D) stratification modeling (Sitharam et al. 2008; Li et al. 2016d). However, it may fail to characterize complicated stratification forms, such as one material embedding in another material or several staggered materials. In addition, the uncertainty involved in the potential boundary identification for a given borehole cannot yet be considered.

In contrast, the category-based model is applicable more to the embedding stratification case than to the uniform case. The latter may be unfavorable for estimation of the transition probability matrix. The category-based model matches the natural weathering or deposition process by accounting for the material transition, and it is capable of generating almost arbitrary geotechnical strata. However, applying it to 3D geotechnical strata (He et al. 2009; Park 2010; Liang et al. 2014) is relatively impractical because of its high demand for sampled data (e.g., amount and locations). Besides, the material spatial distribution trend cannot be incorporated, which may lead to some physically abnormal realizations.

To better model geotechnical stratification in the probabilistic framework, a heuristic combination model is proposed in this study to combine the boundary-based and category-based models, as shown in Fig. 5. The procedures are summarized in the following paragraphs.

The model parameters involved in both models should be identified first, as described in previous sections. Thereafter, a conditional random realization of material boundaries is generated using the boundary-based model [Eq. (3)] and is further converted into a realization of material categories, referred to as the background profile. Finally, the category-based model is applied to simulate the strata of the final profile. In addition to the borehole information,

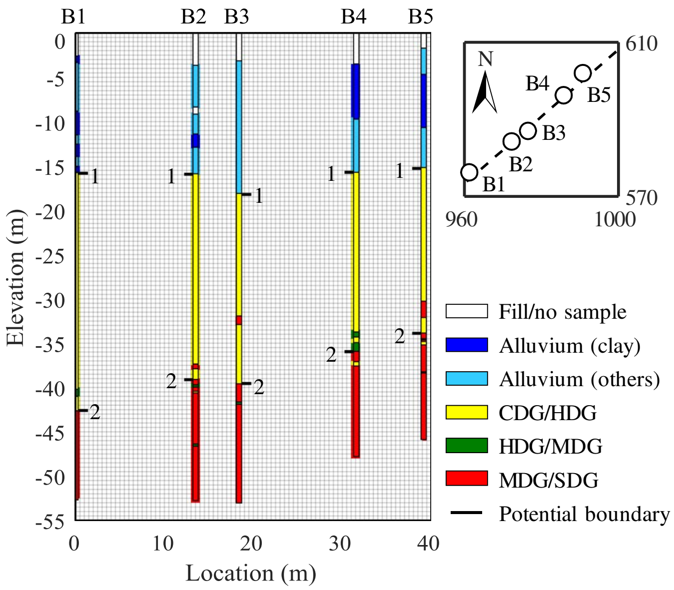

Fig. 6. Borehole details for a selected geologic profile

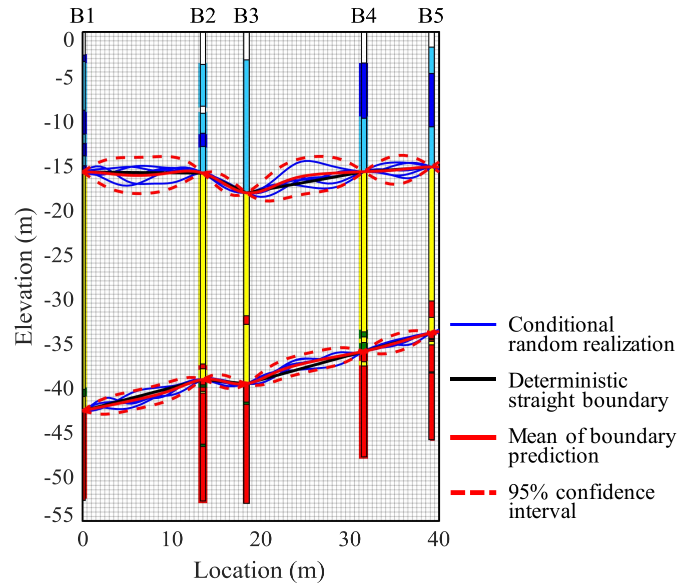

Fig. 8. Boundary-based probabilistic stratification modeling

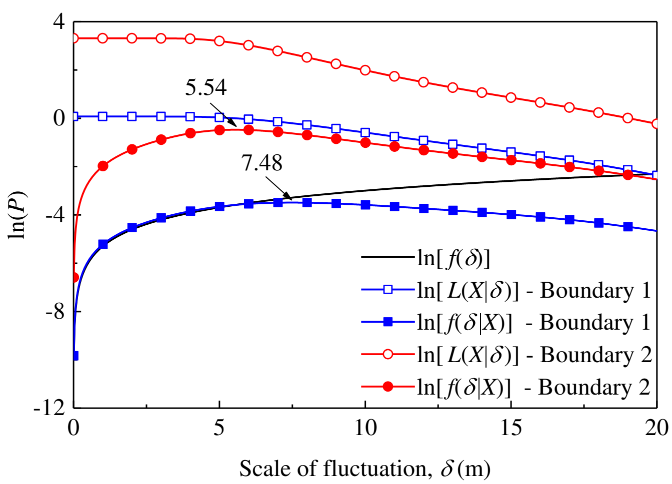

Fig. 7. Parameter identification for the boundary-based model

Table 1. Identified Model Parameters for the Boundary-Based Model

Boundary	Trend	Standard deviation, σ (m)	Scale of fluctuation, δ (m)
Boundary 1	$X = 0.014z - 16.34$	1.05	7.48
Boundary 2	$X = 0.22z - 42.68$	0.54	5.54

the hypothetical background profile is taken as a constraint during the material transition. Hence the total transition equation can be written as Eq. (6) multiplied by $p_{kd}^{h(\zeta)}$, where ζ is the number of transition steps from the final profile (S_k) to the background profile (S_d). Different from the other model parameters, ζ is related to the horizontal transition probability matrix and determined by engineers to reflect the degree of belief concerning the material spatial distribution trend. The larger the value of ζ is, the less information is read from the boundary-based model. When ζ increases to infinity, the heuristic combination model provides the same results that the category-based model does. For reference, if ζ takes the

Table 2. Estimated Vertical Transition Probability Matrix

	Material 1	Material 2	Material 3	Material 4	Material 5
Material 1	0.833	0.148	0.019	0	0
Material 2	0.048	0.903	0.048	0	0
Material 3	0	0	0.945	0.017	0.038
Material 4	0	0	0.231	0.385	0.385
Material 5	0	0	0.035	0.018	0.947

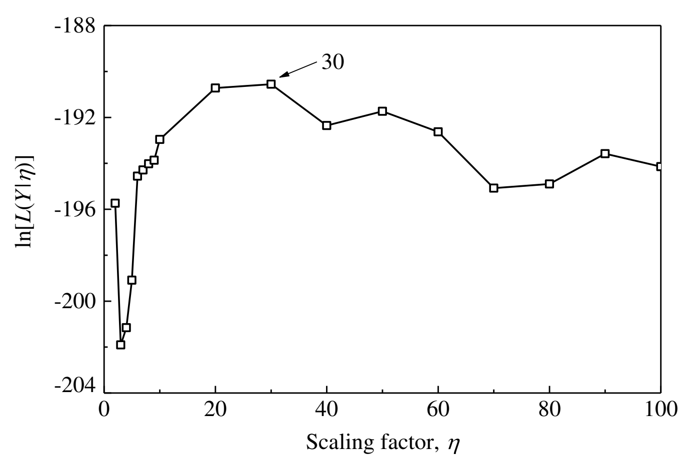

Fig. 9. Parameter identification for the category-based model

Table 3. Estimated Horizontal Transition Probability Matrix

	Material 1	Material 2	Material 3	Material 4	Material 5
Material 1	0.993	0.006	0.000	0	0
Material 2	0.002	0.996	0.002	0	0
Material 3	0	0	0.998	0.000	0.001
Material 4	0	0	0.019	0.949	0.032
Material 5	0	0	0.001	0.000	0.998

same order of magnitude as the number of transition steps among neighbor boreholes, the material spatial distribution trend is given the same importance as the borehole information. The proposed heuristic combination model takes advantage of the boundary-based and category-based models. 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.

Illustrative Examples

For illustration, this section applies the proposed probabilistic stratification modeling framework to a construction site in Hong Kong. The example was investigated by Zhang and Dasaka (2010), Dasaka and Zhang (2012), and Li et al. (2016d) through a boundary-based stratigraphic model. For simplicity, this study only considers 2D strata for a selected geologic profile, which is approximately 40 m in width and 55 m in depth and contains five boreholes (i.e., B1–B5) almost along a straight line. Fig. 6 shows the details of the five boreholes. Five geotechnical material categories, referred to as Materials 1–5, are identified—namely, clay alluvium, other alluvium, completely to highly decomposed granite (CDG/HDG), highly to moderately decomposed granite (HDG/MDG), and moderately to slightly decomposed granite (MDG/SDG). Both material boundaries and material categories are predicted from a probabilistic perspective.

Material Boundary Prediction

As shown in Fig. 6, the clay alluvium and other alluvium are mutually embedded, resulting in an unidentifiable boundary between them. However, with the aid of engineering judgment, the boundary between alluvium and granite and the boundary between CDG/HDG and MDG/SDG are basically identifiable. They are referred to as Boundary 1 and Boundary 2, respectively, as noted in Fig. 6.

To apply the boundary-based model, the spatial correlation of boundary depth is described by a linear trend function and a squared exponential correlation function in this study. According to Eq. (1), the marginal distribution of the likelihood function as a function of scale of fluctuation δ is shown in Fig. 7. Because the number of observations is limited (i.e., only five), the likelihood functions for Boundaries 1 and 2 are unidentifiable when δ is small. Notably, δ is the most important parameter for the boundary-based model and has a significant impact on stratification modeling. Therefore, the Bayesian approach is adopted to facilitate model parameter identification, with an uninformative prior on β and σ (i.e., a uniform distribution) and a relatively strong prior on δ (i.e., a triangular distribution ranging from 0 to 20 m, with the largest probability at 20 m). Herein, the strong prior on δ reflects, preliminarily, the engineering judgment that δ is inclined to be large. Different prior distributions of δ can also be applied if other prior knowledge is available (e.g., the range of δ is from 5 to 40 m). Interested readers are referred to Cao et al. (2016) for more details on the determination of prior distribution. By this means, the most probable values of δ for Boundaries 1 and 2 are identified as 7.48 and 5.54 m, respectively, as shown in Fig. 7, with statistical uncertainties (in the form of standard deviations) of 4.90 and 3.07 m, respectively. The relatively large statistical uncertainties imply that the number of observations may be insufficient. The most probable values for the other boundary parameters are summarized in Table 1. Boundary 2 has a more evident trend and a smaller standard deviation than Boundary 1, the spatial variation in which is mainly represented by the trend.

Taking the identified boundary parameters into Eq. (3), several conditional simulation realizations for the two boundaries can be generated. Fig. 8 shows five random realizations as examples. At every borehole location, the conditional simulation gives material boundaries consistent with those observed in the borehole logs. For reference, the deterministic straight boundaries, without consideration of stratigraphic uncertainty, and the mean and 95% confidence interval of the boundary predictions evaluated by Eqs. (4) and (5) are also shown in Fig. 8. The deterministic boundary

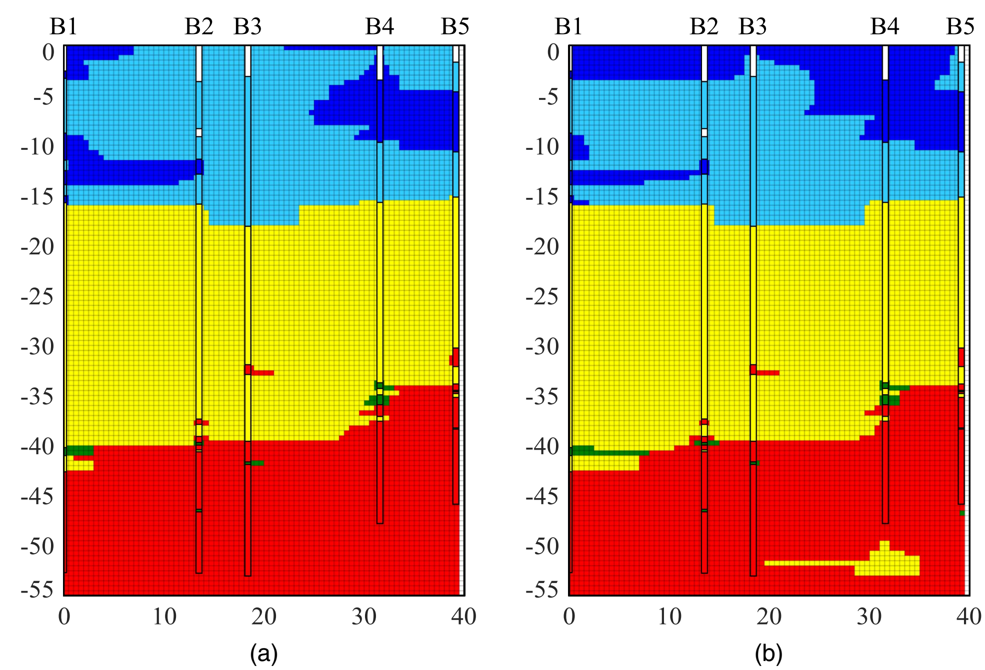

Fig. 10. Conditional random realizations using the category-based model: (a) realization 1; (b) realization 2

appears reasonable to a certain degree because it agrees well with the mean prediction from a probabilistic viewpoint. Furthermore, the 95% confidence interval for Boundary 2 is narrower than that for Boundary 1, indicating that Boundary 2 has lower stratigraphic uncertainty. For both boundaries, the uncertainty at unsampled locations decreases as the distance to boreholes shortens, and approaches zero at the borehole locations. This is because the relative uncertainty reduction is related only to relative location and scale of fluctuation, as interpreted by Eq. (5). For the same reason, the shortest distance between B2 and B3 leads to the minimal uncertainty in space. The stratigraphic uncertainty can be properly quantified by the boundary-based model, rather than determined by engineering judgment, as in conventional deterministic stratification modeling. However, the uncertainty involved in the potential boundary identification for a given borehole (i.e., identification of the actual locations of Boundaries 1 and 2 as shown in Fig. 6) cannot be considered at the current stage. Such uncertainty can be avoided in the category-based model because the boundary itself is not required any more.

Material Category Prediction

With regard to material category prediction, the geologic profile is first discretized into thousands of cells as shown in Fig. 6, each cell being treated as the minimum material transition unit with 0.5×0.5 m in dimensions. By counting the material transition numbers from top to bottom in all boreholes, the vertical transition probability matrix is estimated as shown in Table 2. The matrix, which is approximately diagonally dominant, indicates that the geological material category (except HDG/MDG) has a high probability of

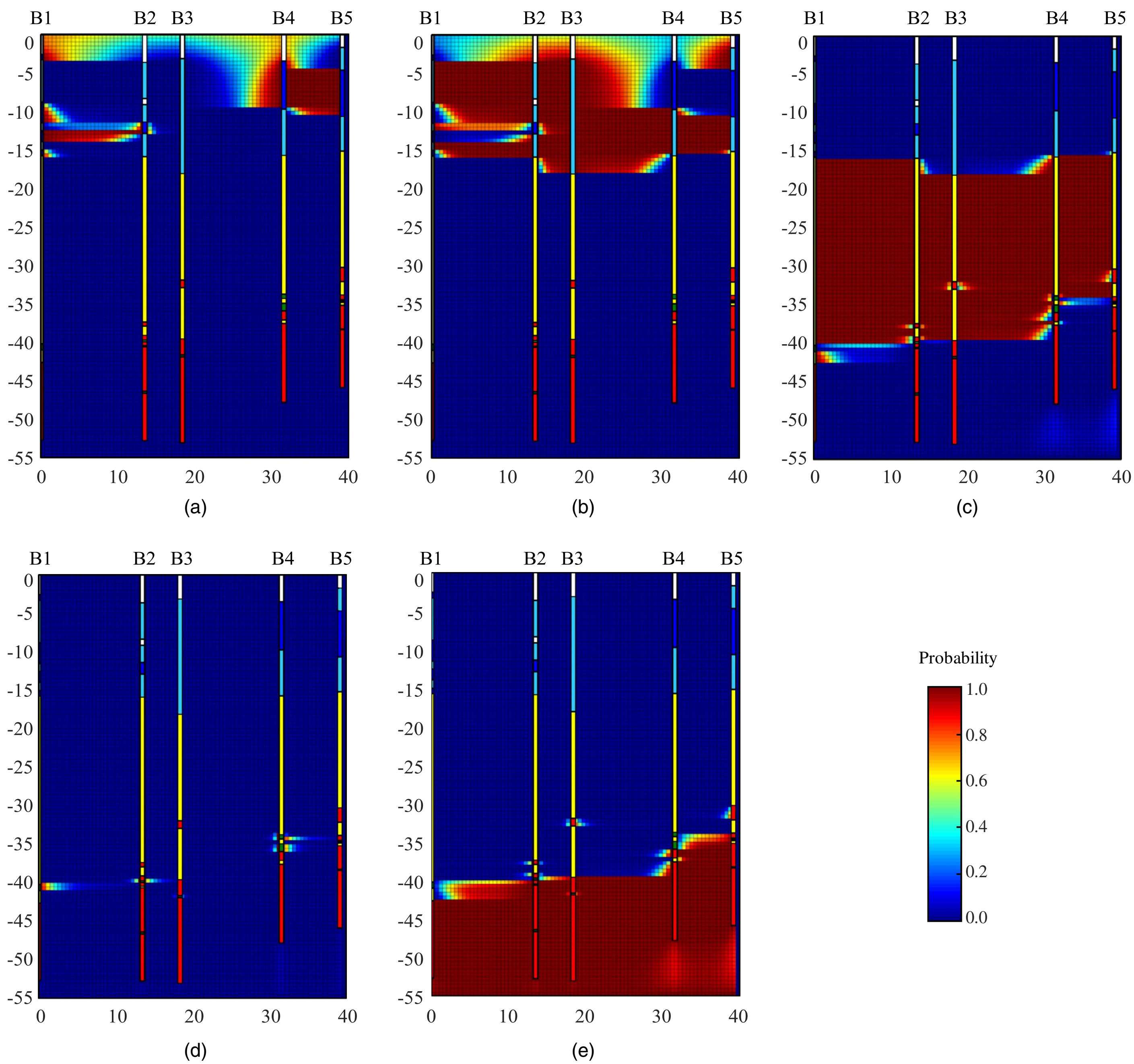

Fig. 11. Occurrence probability for material categories: (a) alluvium (clay); (b) alluvium (other) (c) CDG/HDG; (d) HDG/MDG; (e) MDG/SDG

remaining unchanged during the transition. No transition occurs from Materials [1, 2] to [4, 5] or from Materials [3, 4, 5] to [1, 2]. This implies that the vertical transition sequence should be $[1,2] \rightarrow [3] \rightarrow [4,5]$. In other words, as the depth increases the material category transits from alluvium to completely decomposed granite and then to slightly decomposed granite. This is consistent with the natural decomposition process of weathered rocks.

To estimate the horizontal transition probability matrix using Eq. (7), the profile is divided into two zones. In the left zone, B1 and B3 are treated as the boundary boreholes and B2 is viewed as an intermediate borehole. Similarly, in the right zone B4 is an intermediate borehole bounded by B3 and B5. Eighteen possible values of scaling factor η, ranging from 2 to 100, are considered in this study. For each η value, a Monte Carlo simulation with 10,000 realizations is performed to estimate the likelihood of observing B2 and B4 simultaneously. As illustrated in Fig. 9, $\eta = 30$ is the most probable value for the obtained borehole information. The corresponding horizontal transition probability matrix is provided in Table 3. Each material has a much higher probability of transiting to itself in the horizontal direction than in the vertical direction, which contributes to relatively horizontal stratification in space.

After the two transition probability matrices have been obtained, all five boreholes are viewed as the boundary boreholes and the profile is divided into four zones. Two examples of random realization conditioned on all borehole data are shown in Fig. 10. As expected, the material strata are relatively horizontal. Moreover, because no borehole data are available in the bottom-right zone, there is a small chance to mathematically form some pockets of CDG/HDG in a more uniform MDG/SDG [e.g., Fig. 10(b)]. Such a case is often regarded as abnormal from a physical viewpoint. In addition to the boundary between alluvium and granite and the boundary between CDG/HDG and MDG/SDG, the category-based model can also capture the complicated boundary between clay alluvium and other alluvium and various complicated embedding forms (e.g., HDG/MDG).

With respect to stratigraphic uncertainty, Fig. 11 shows the occurrence probabilities of the different material categories at every location which are evaluated, statistically and cell by cell, based on 10,000 realizations through Monte Carlo simulation. It is clear that the stratigraphic uncertainty between clay alluvium and other alluvium is the most significant one and that it influences the largest region. For the stratification between alluvium and granite (i.e., Boundary 1 in the boundary-based model), the largest uncertainty exists in the region between B3 and B4, followed by the region between B2 and B3; almost no uncertainty exists at other locations. Furthermore, the stratification between CDG/HDG and MDG/SDG (i.e., Boundary 2 in the boundary-based model) is more uncertain than that between alluvium and granite. Compared with the boundary-based model, the different findings may be attributed to the embedding of HDG/MDG and the fact that the material spatial distribution trend cannot be incorporated into the category-based model.

Combination of Boundary and Category Predictions

In addition to the boundary-based and category-based models, the proposed heuristic combination model is applied to the probabilistic stratification modeling of the considered example. The adopted model parameters are in accordance with those identified in the previous sections. Consider the number of transition steps $\zeta = 10$ for example. A conditional random realization using the heuristic combination model is given in Fig. 12. The background profile shown in Fig. 12(a) is first generated by the boundary-based model and then used as an additional future state to obtain the final profile shown in Fig. 12(b) based on the category-based model. In comparison with Fig. 10, Fig. 12(b) not only simulates the complicated embedding forms for alluvium and HDG/MDG but also learns partial information from the material spatial distribution trend.

As in the category-based model, the stratigraphic uncertainty associated with the heuristic combination model can be estimated by Monte Carlo simulation. Fig. 13 shows the occurrence probability for CDG/HDG with different ζ values based on 10,000 realizations. On the one hand, the material spatial distribution trend can be heavily utilized when $\zeta = 0$. As a result, the stratigraphic

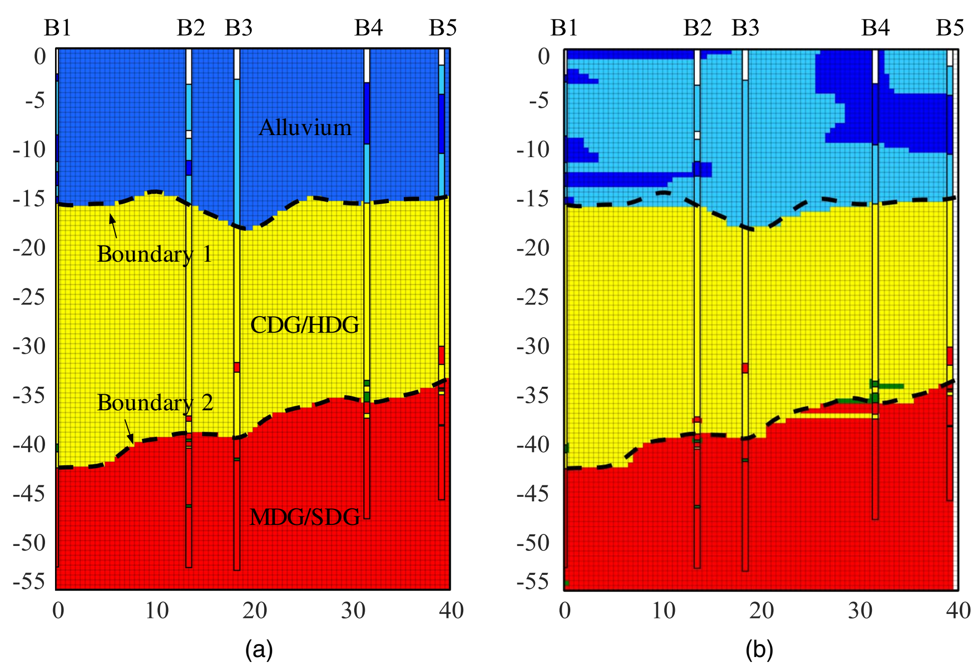

Fig. 12. Conditional random realization using the heuristic combination model ($\zeta = 10$): (a) background profile; (b) final profile

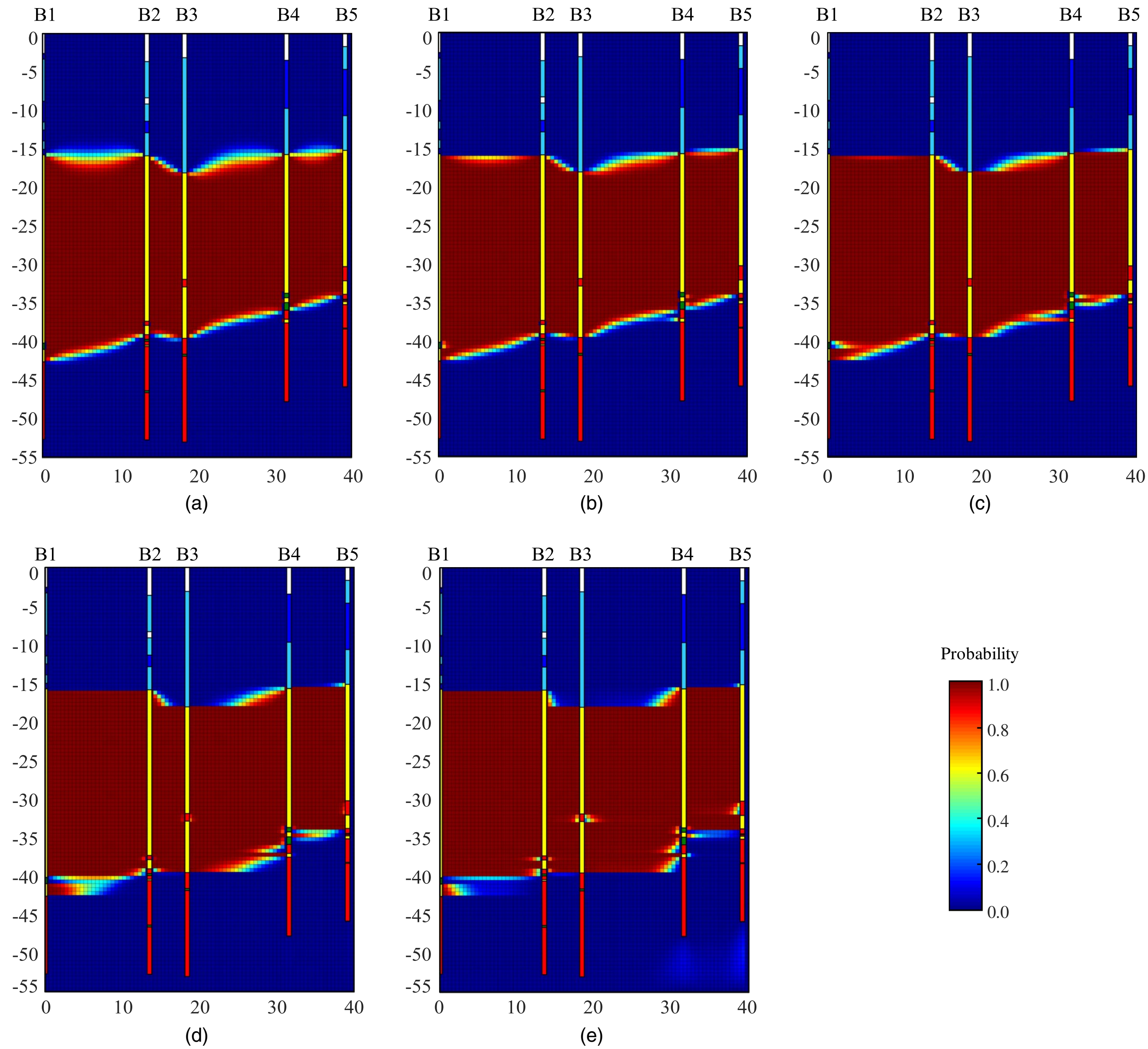

Fig. 13. Occurrence probability for CDG/HDG using the heuristic combination model: (a) $\zeta = 0$; (b) $\zeta = 1$; (c) $\zeta = 10$; (d) $\zeta = 100$; (e) $\zeta = 10{,}000$

uncertainties for Boundaries 1 and 2 are close to those obtained from the boundary-based model (i.e., Fig. 8). On the other hand, $\zeta = 10{,}000$ brings about stationary transition so that nothing can be learned from the material distribution trend. The stratigraphic uncertainties for Boundaries 1 and 2 are thus equal to those obtained from the category-based model [Fig. 11(c)]. The heuristic combination model becomes a link between the boundary-based and category-based models via the ζ-step transition from the final profile to the background profile. The choice of ζ depends on engineering judgment and reflects the degree of belief regarding the material spatial distribution trend.

Summary and Conclusions

A three-level probabilistic framework is proposed for geotechnical stratification modeling that considers stratigraphic uncertainty, and a heuristic combination model is recommended to combine the boundary-based and category-based stratigraphic models. The proposed approach is applied to a construction site in Hong Kong. The following conclusions can be drawn:

- Probabilistic stratification modeling quantifies stratigraphic uncertainty in a rigorous manner and can be integrated with other sources of uncertainty and incorporated into geotechnical reliability analysis and reliability-based design;
- The Bayesian approach takes full advantage of both observations and prior information, and facilitates parameter identification for the boundary-based model when the number of observations is limited; the boundary between CDG/HDG and MDG/SDG exhibits a more evident trend, a smaller standard deviation, and a smaller scale of fluctuation than the boundary between alluvium and granite;
- The category-based model captures various complicated embedding forms between clay alluvium and other alluvium, and

between HDG/MDG and other decomposed granite; the stratigraphic uncertainty between clay alluvium and other alluvium is the most significant and influences the largest region; and

- The heuristic combination model takes advantage of both boundary-based and category-based models; 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.

Acknowledgments

This work was supported by the Research Grants Council of Hong Kong (Nos. 16202716 and C6012-15G) and the National Natural Science Foundation of China (Nos. 51329901, 51509188, and 51579190). The first author wishes to thank the Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology, for hosting his visit as an exchange Ph.D. student.

References

Baecher, G. B., and Christian, J. T. (2003). *Reliability and statistics in geotechnical engineering*, Wiley, Hoboken, NJ.

Cao, Z., and Wang, Y. (2013). "Bayesian approach for probabilistic site characterization using cone penetration tests." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0000765, 267–276.

Cao, Z., Wang, Y., and Li, D. (2016). "Quantification of prior knowledge in geotechnical site characterization." *Eng. Geol.*, 203, 107–116.

Ching, J., Wang, J. S., Juang, C. H., and Ku, C. S. (2015). "Cone penetration test (CPT)-based stratigraphic profiling using the wavelet transform modulus maxima method." *Can. Geotech. J.*, 52(12), 1993–2007.

Dasaka, S. M., and Zhang, L. M. (2012). "Spatial variability of in situ weathered soil." *Géotechnique*, 62(5), 375–384.

Elfeki, A., and Dekking, M. (2001). "A Markov chain model for subsurface characterization: Theory and applications." *Math. Geol.*, 33(5), 569–589.

Elfeki, A. M. M., and Dekking, F. M. (2005). "Modelling subsurface heterogeneity by coupled Markov chains: Directional dependency, Walther's law and entropy." *Geotech. Geol. Eng.*, 23(6), 721–756.

Elkateb, T., Chalaturnyk, R., and Robertson, P. K. (2003). "An overview of soil heterogeneity: Quantification and implications on geotechnical field problems." *Can. Geotech. J.*, 40(1), 1–15.

Fenton, G. A., and Griffiths, D. V. (2008). *Risk assessment in geotechnical engineering*, Wiley, Hoboken, NJ.

Forrester, A., Sobester, A., and Keane, A. (2008). *Engineering design via surrogate modelling: A practical guide*, Wiley, Chichester, U.K.

He, Y., et al. (2009). "Sequential indicator simulation and indicator kriging estimation of 3-dimensional soil textures." *Soil Res.*, 47(6), 622–631.

Houlsby, N. M. T., and Houlsby, G. T. (2013). "Statistical fitting of undrained strength data." *Géotechnique*, 63(14), 1253–1263.

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

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

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

Li, W. (2007). "Markov chain random fields for estimation of categorical variables." *Math. Geol.*, 39(3), 321–335.

Li, X. Y., Zhang, L. M., and Li, J. H. (2016d). "Using conditioned random field to characterize the variability of geologic profiles." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0001428, 04015096.

Li, Z., Wang, X., Wang, H., and Liang, R. Y. (2016e). "Quantifying stratigraphic uncertainties by stochastic simulation techniques based on Markov random field." *Eng. Geol.*, 201, 106–122.

Liang, Y. R., Wang, Z. Z., and Guo, J. H. (2014). "Reservoir lithology stochastic simulation based on Markov random fields." *J. Cent. South Univ.*, 21(9), 3610–3616.

Liao, T., and Mayne, P. W. (2007). "Stratigraphic delineation by three-dimensional clustering of piezocone data." *Georisk*, 1(2), 102–119.

Liu, Y., Jiang, Y. J., Xiao, H., and Lee, F. H. (2017). "Determination of representative strength of deep cement-mixed clay from core strength data." *Géotechnique*, 67(4), 350–364.

Nobre, M. M., and Sykes, J. F. (1992). "Application of Bayesian kriging to subsurface characterization." *Can. Geotech. J.*, 29(4), 589–598.

Park, E. (2010). "A multidimensional, generalized coupled Markov chain model for surface and subsurface characterization." *Water Resour. Res.*, 46(11), W11509.

Phoon, K. K., et al. (2016). "Some observations on ISO2394: 2015 annex D (reliability of geotechnical structures)." *Struct. Saf.*, 62, 24–33.

Phoon, K. K., Quek, S. T., and An, P. (2003). "Identification of statistically homogeneous soil layers using modified Bartlett statistics." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)1090-0241(2003)129:7(649), 649–659.

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." *Eng. Geol.*, 207, 129–140.

Schöbi, R., and Sudret, B. (2015). "Application of conditional random fields and sparse polynomial chaos expansions to geotechnical problems." *Proc., 5th Int. Symp. on Geotechnical Safety and Risk*, T. Schweckendiek, et al., eds., IOS Press, Rotterdam, Netherlands, 441–446.

Schöbi, R., Sudret, B., and Marelli, S. (2017). "Rare event estimation using polynomial-chaos kriging." *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A Civil Eng.*, 3(2), D4016002.

Sitharam, T. G., Samui, P., and Anbazhagan, P. (2008). "Spatial variability of rock depth in Bangalore using geostatistical, neural network and support vector machine models." *Geotech. Geol. Eng.*, 26(5), 503–517.

Vardanega, P. J., and Bolton, M. D. (2016). "Design of geostructural systems." *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A Civil Eng.*, 2(1), 04015017.

Wang, Y., Akeju, O. V., and Cao, Z. J. (2016a). "Bayesian equivalent sample toolkit (BEST): An excel VBA program for probabilistic characterization of geotechnical properties from limited observation data." *Georisk*, 10(4), 251–268.

Wang, Y., Cao, Z., and Li, D. (2016b). "Bayesian perspective on geotechnical variability and site characterization." *Eng. Geol.*, 203, 117–125.

Wang, Y., Huang, K., and Cao, Z. (2014). "Bayesian identification of soil strata in London clay." *Géotechnique*, 64(3), 239–246.

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." *Comput. Geotech.*, 79, 146–158.

Xiao, T., Li, D. Q., Cao, Z. J., and Tang, X. S. (2017). "Full probabilistic design of slopes in spatially variable soils using simplified reliability analysis method." *Georisk*, 11(1), 146–159.

Zhang, L. M., and Dasaka, S. M. (2010). "Uncertainties in geologic profiles versus variability in pile founding depth." *J. Geotech. Geoenviron. Eng.*, 10.1061/(ASCE)GT.1943-5606.0000364, 1475–1488.

Zhu, H., and Zhang, L. M. (2013). "Characterizing geotechnical anisotropic spatial variations using random field theory." *Can. Geotech. J.*, 50(7), 723–734.

Zhu, H., Zhang, L. M., Xiao, T., and Li, X. Y. (2017). "Generation of multivariate cross-correlated geotechnical random fields." *Comput. Geotech.*, 86, 95–107.