Reappraisal of reliability of a slope through hybridisation of regional and site-specific soil shear strength information

M.K. Lo[1], Y.F. Leung[1*], C.L. Chan[2], and E.H.Y. Sze[2]

[1]Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University
[2]Civil Engineering and Development Department, Hong Kong SAR Government, Hong Kong
[*]Corresponding author, email address: andy.yf.leung@polyu.edu.hk

Abstract

Uncertainty in soil parameters is usually characterised by probability density functions (PDFs), with the influence on system performance represented through the probability of failure. Difficulties in selecting representative PDFs for a project often arise from scarcity of site-specific information, even with ample previous knowledge and test data of similar soil types in the region. This paper proposes an approach to rationally assimilate regional and site-specific information. A newly-compiled regional database of shear strength information for saprolitic soils in Hong Kong is presented, based on results of multi-stage consolidated-undrained triaxial tests. A hierarchical Bayesian model is fitted to the regional database, followed by a Bayesian updating model that produces posterior predictive distributions of shear strength parameters. The posterior estimates incorporate site-specific features into regional information, leading to profound impacts on the evaluation of failure probability for a slope case. To further illustrate the significance of data hybridisation, four semi-hypothetical scenarios are created using the same slope geometry, by assuming that distributions of shear strength parameters are completely known at the site. With the proposed approach, the estimated failure probability approaches the true value with increasing amount of site-specific data, and is more robust than adopting regional data or site data alone.

Keywords: slope reliability; hierarchical Bayesian model; completely decomposed granite; triaxial test database

1 Introduction

The stability of a slope is often assessed through the deterministic factor of safety, evaluated using limit equilibrium methods, finite element or finite difference methods. However, the stability of a soil slope is in fact highly uncertain, as the shear strength parameters (e.g., effective cohesion, c', and effective friction angle, ϕ') may exhibit significant inherent variability. For example, Ching et al. (2017) compiled a global database with around 1000 data points of shear strength properties of sands (SAND/7/2794), and a coefficient of variation of 12.8% was reported for the friction angle. Other studies (Phoon and Kulhawy 1999; Parker et al. 2008; Rahardjo et al. 2012; Chan et al. 2022) have also revealed significant variability in material properties, even for similar soil types within a site.

To assess the landslide hazards arising from variability in shear strength parameters, the parameters can be characterised as probability density functions (PDFs), with the reliability of the slope represented through the probability of failure or reliability index, evaluated using probabilistic approaches such as Monte Carlo simulations (e.g., El-Ramly et al. 2005; Wang et al. 2011; Mori et al. 2020) or first-order reliability methods (e.g., Hasofer and Lind 1974; Low and Tang 1997; Duncan 2000). In addition, a number of recent studies adopted random field methods (e.g., Griffiths and Fenton 2004; Cho 2010; Lo and Leung 2017) to capture the spatial variation of $c' - \phi'$ parameters within a site. From the perspective of site-specific application, a common challenge for these reliability methods is to assign an appropriate PDF for shear strength parameters that reasonably reflects the site characteristics.

A general understanding on the geomaterial properties and their variability can be developed based on previous experience in the respective cities or regions. On some occasions, there may be sufficient test data to establish a 'regional' database

that captures the distribution of shear strength parameters in the territory. In this study, this is represented by the 'regional PDF', which can be considered as a pooled distribution of soil parameters from many sites of similar geologic history or conditions. On the other hand, a 'site-specific PDF' refers to the distribution of soil parameters at a particular site. This is typically unknown unless abundant site data are available, which is rarely the case in practice. This presents a dilemma regarding the choice of PDFs for probabilistic analyses. In principle, site-specific PDFs better reflect the site conditions, and non-parametric approaches had been proposed (Wang et al. 2015; Ching and Phoon 2019) to derive PDFs solely from site data, but these may suffer from significant statistical uncertainty. On the other hand, regional PDFs are less affected by the 'curse of small sample size' (Phoon 2017), but there is no guarantee that the regional data and ensuing analyses represent the distinctive site conditions.

This paper proposes an approach to rationally assimilate regional database and site-specific data to produce the posterior predictive distributions of shear strength parameters. Conceptually, with very limited site-specific data, these posterior estimates would resemble the regional PDF, and it would approach the site-specific PDF with increasing amount of site information. There are two main steps in the proposed approach. The first step involves modelling the variations in soil shear strength within individual sites and also among different sites, through the hierarchical Bayesian model (HBM) (Gelman and Hill 2006; Lunn et al. 2013). HBM has been previously applied to the characterisation and prediction of soil properties (Zhang et al. 2016; Lo et al. 2021; Xiao et al. 2021; Ching et al. 2021b; Ueda 2022), while Bozorgzadeh et al. (2019) offers the first comprehensive discussion of HBM as a promising and suitable approach to data analyses, using rock properties as an example. This was followed by HBM analyses of rock properties data by Feng et al. (2021) and Ching et al. (2021a), and of model factors for foundation design (Zhang et al. 2020). The second step is to utilize the fitted

HBM model to formulate informative priors for a new site. Data from the new site can be hybridised with the informative priors under the Bayesian framework, resulting in posterior predictive distribution of design parameters. Summarizing the a fitted HBM model as informative priors and communicating them for future analysis has recently been discussed by Bozorgzadeh et al. (2023).

This paper first describes the structure and fitting process of the HBM, utilising a newly compiled regional database of shear strength information containing hundreds of multi-stage consolidated-undrained triaxial compression test results on a saprolitic soil type encountered in Hong Kong. The probabilistic model for assimilation of site-data is then presented, followed by a case study in Hong Kong. The case involves a number of site-specific borehole information and triaxial test data, which facilitate comparisons between probabilistic analyses adopting regional PDF and the posterior predictive distribution. Through this process, the influence of data availability on the PDFs of parameters and the subsequent slope reliability assessments will be illustrated. Specifically, the novelty of this paper lies on the following aspects:

1. Formulation of a flexible HBM model which can be applied to data of various soil types. The model allows different types of PDFs for c' and ϕ' parameters, and can incorporate the correlation between c' and ϕ'.

2. Illustration of the practical significance of assimilating regional and site-specific data in the context of a slope case study, showing that the ensuing posterior predictive distributions of shear strength parameters lead to reduced statistical uncertainty and more robust reliability assessments.

3. Illustration of the effects of site-specific sample numbers on the transition from regional PDF to site-specific PDF of shear strength parameters, and the subsequent influence on slope reliability evaluations.

2 Hierarchical Bayesian modelling of soil shear strength database

2.1 Model Structure

A HBM can model information on multiple levels, such as data collected from multiple sites or regions, through a nested data structure with multiple layers. Fig. 1 shows the graphical structure of the HBM for the variations of shear strength information in a region. Each node in the graph represents the probability distribution of a variable, with arrows indicating the inter-dependence between variables. Through multiple nested layers, variability in shear strength parameters is decomposed into 'within-site variations' and 'between-sites variations'.

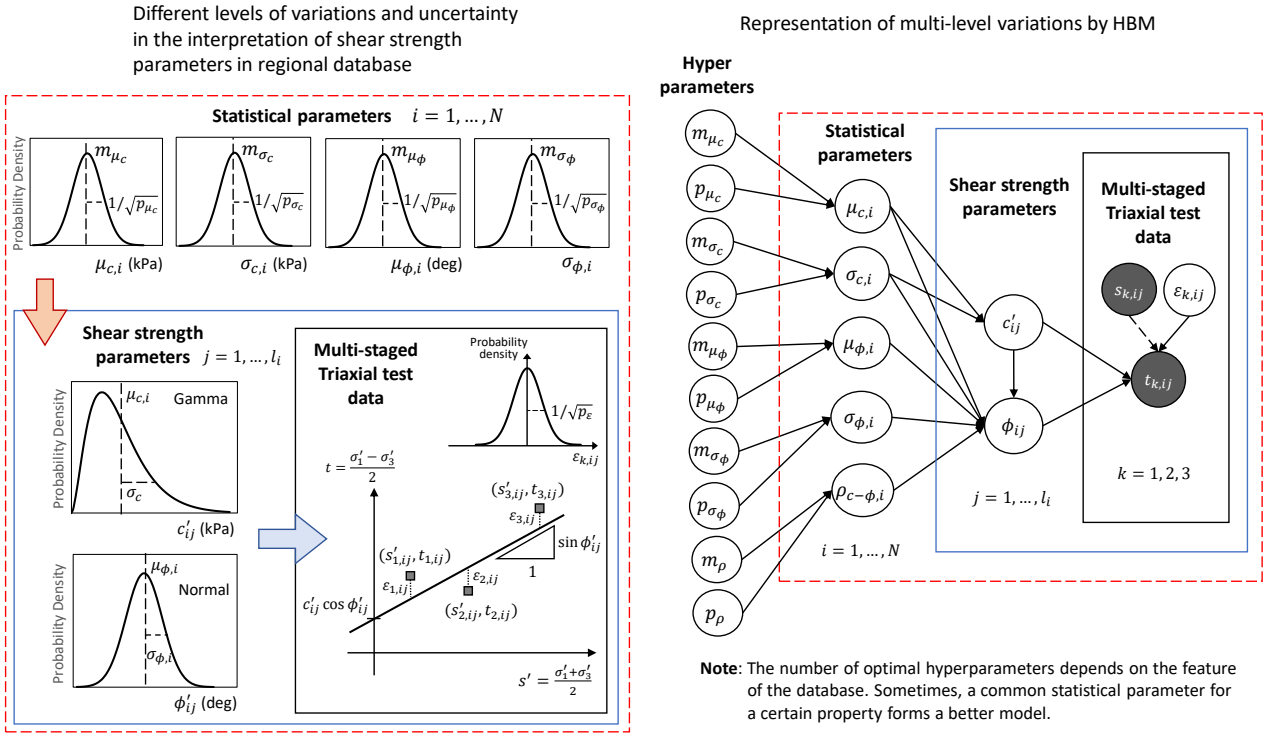

Figure 1: Hierarchical Bayesian model for the regional shear strength data. Solid arrows indicate statistical dependence; dashed arrows indicate functional dependence. Shaded nodes are the observed data, non-shaded nodes are parameters to be fitted.

Considering the scenario where a regional database is compiled using test data from

N sites across the region, the experimental data points at the i-th site can be denoted as $\boldsymbol{s'_{k,i}} = [s'_{k,i1}, \ldots, s'_{k,il_i}]$ and $\boldsymbol{t_{k,i}} = [t_{k,i1}, \ldots, t_{k,il_i}]$, where k (1, 2 or 3) denotes the individual stages in a multi-stage triaxial test, and l_i is the number of triaxial tests associated with that site. Meanwhile, the $c' - \phi'$ values at the i-th site are denoted as $\boldsymbol{c'_i} = [c'_{i1}, ..., c'_{il_i}]$ and $\boldsymbol{\phi'_i} = [\phi'_{i1}, ..., \phi'_{il_i}]$ respectively. The shear strength parameters (c', ϕ') are related to the stress points $(s' = (\sigma'_1 + \sigma'_3)/2,\ t = (\sigma'_1 - \sigma'_3)/2)$ through the Mohr-Coulomb criterion:

$$t_{k,ij} = c'_{ij} \cos \phi'_{ij} + s'_{k,ij} \sin \phi'_{ij} + \varepsilon_{k,ij} \tag{1}$$

where $\varepsilon_{k,ij}$ is a residual error term, assumed to be normally distributed with zero mean and precision (inverse variance) of p_ε, i.e. $p_\varepsilon = 1/\sigma_\varepsilon^2$. The current study adopts σ_ε as a 'global' parameter (i.e. common among all sites) as its value is found to be small (as shown later in Table 2) compared to the peak shear stresses. The practical value of varying σ_ε among the sites is not pronounced.

Within each individual site, the distribution of c' has a mean value of $\mu_{c,i}$ and standard deviation of $\sigma_{c,i}$, while ϕ' has a mean value of $\mu_{\phi,i}$ and standard deviation of $\sigma_{\phi,i}$. As will be shown later, based on the data in Hong Kong, c' is assumed to follow the gamma distribution, while ϕ' is assumed to follow the normal distribution. The correlation between c' and ϕ' is represented by the coefficient $\rho_{c-\phi}$. Accordingly,

$$\begin{aligned}
&c'_{ij}|\mu_{c,i}, \sigma_c \sim \text{Gamma}\left(\alpha_i = \frac{\mu_{c,i}^2}{\sigma_c^2},\ \beta_i = \frac{\mu_{c,i}}{\sigma_{c,i}^2}\right) \\
&\phi'_{ij}|c'_{ij}, \mu_{\phi,i}, \sigma_{\phi,i}, \mu_{c,i}, \sigma_c, \rho_{c-\phi,i} \\
&\sim N\left(\mu_{\phi,i} + \rho_{c-\phi,i}\sigma_{\phi,i}\Phi^{-1}[F_c(c'_{ij})], \sqrt{1-\rho_{c-\phi,i}^2}\sigma_{\phi,i}\right)
\end{aligned} \tag{2}$$

where α_i and β_i are the shape and rate parameters of the gamma distribution; and $X|Y$ indicates that the random variable X is conditioned by Y. Copula theory (Nelsen 1999) is used to model the conditional distribution of ϕ', with linear correlation coefficient $\rho_{c-\phi}$ introduced through the Gaussian copula. It should be noted that with necessary adjustments, the proposed HBM framework can also facilitate various types of PDF for c' and ϕ' other than those in Eq. (2), if warranted by the test data in the region.

Following Eq. (2), there are five statistical parameters characterising the distribution at each site, namely $\mu_{c,i}$, $\sigma_{c,i}$, $\mu_{\phi,i}$, $\sigma_{\phi,i}$ and $\rho_{c-\phi,i}$. In the general framework, these parameters are assumed to be similar but not necessarily identical across the sites, meaning that they may follow a common parametric distribution. Herein each statistical parameter is assumed to follow the truncated normally distribution across all sites, with the mean and precision (inverse variance) denoted by m and p, respectively. The parameters $\mu_{c,i}$, $\mu_{\phi,i}$, $\sigma_{c,i}$, $\sigma_{\phi,i}$ are truncated below zero, while $\rho_{c-\phi,i}$ are truncated below -1 and above +1. Expressing this in mathematical form,

$$
\begin{aligned}
\mu_{c,i} &\sim TN\left(m_{\mu_c}, 1/\sqrt{p_{\mu_c}}\right)(0,+\infty) \\
\mu_{\phi,i} &\sim TN\left(m_{\mu_\phi}, 1/\sqrt{p_{\mu_\phi}}\right)(0,+\infty) \\
\sigma_{c,i} &\sim TN\left(m_{\sigma_c}, 1/\sqrt{p_{\sigma_c}}\right)(0,+\infty) \\
\sigma_{\phi,i} &\sim TN\left(m_{\sigma_\phi}, 1/\sqrt{p_{\sigma_\phi}}\right)(0,+\infty) \\
\rho_{c-\phi,i} &\sim TN\left(m_{\rho}, 1/\sqrt{p_{\rho}}\right)(-1,+1)
\end{aligned} \tag{3}
$$

Depending on features of the specific dataset, certain statistical parameters may be better represented by common values across the sites, instead of fitting the distributions as in Eq. (3). The Bayesian model selection procedure for the reported database will be elaborated in later sections. The parameters governing the 'between-sites variation' (i.e. m_{μ_c}, p_{μ_c}, etc.) are known as hyperparameters in the HBM context. To represent an

Table 1: Priors for the hyperparameters

Uniform prior	Lower bound	Upper bound	Gamma prior	α	β
m_{μ_c} (kPa)	0.1	60	p_{μ_c}	10^{-4}	10^{-4}
m_{μ_ϕ} (°)	0.1	60	p_{μ_ϕ}	10^{-4}	10^{-4}
m_{σ_c} (kPa)	0.1	60	p_{σ_c}	10^{-4}	10^{-4}
m_{σ_ϕ} (°)	0.1	60	p_{σ_ϕ}	10^{-4}	10^{-4}
m_ρ	-0.99	0.99	p_ρ	10^{-4}	10^{-4}
			p_ε	10^{-4}	10^{-4}

initial lack of knowledge in their values, each hyperparameter is assigned a vague prior. The mean hyperparameters are assigned a wide uniform prior. Gamma distributions are commonly adopted in Bayesian literature as prior for the precision parameter (Lunn et al. 2013). The priors for p_{μ_c}, p_{μ_ϕ}, p_{σ_c}, p_{σ_ϕ}, p_ρ and p_ε use gamma distributions, with α and β set as 10^{-4}. Alternative choices of prior for standard deviation parameter include uniform distribution and half-T distribution (Gelman 2006; Huang and Wand 2013). Information of the priors are summarized in Table 1.

2.2 Model fitting

The objective of fitting the HBM is to infer the unobserved parameters based on the observed data points s' and t (grouped as vector $\boldsymbol{y}$). There are three types of unobserved parameters, namely $c' - \phi'$ parameters within each site (grouped as vector $\boldsymbol{\xi}$); statistical parameters of each site (grouped as vector $\boldsymbol{\theta}$), and the hyperparameters (grouped as vector $\boldsymbol{\psi}$). The joint probability density of the HBM is:

$$p(\boldsymbol{y}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\psi}) = p(\boldsymbol{y}|\boldsymbol{\xi})p(\boldsymbol{\xi}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\psi})p(\boldsymbol{\psi}) \tag{4}$$

According to Bayes' theorem, the posterior distribution of the unobserved parameters is:

$$p(\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\psi}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\psi})}{p(\boldsymbol{y})} \propto p(\boldsymbol{y}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\psi}) \tag{5}$$

where the $p(\boldsymbol{y})$ term is constant since it depends only on the observed data. In general, the posterior distribution is high-dimensional and analytically intractable. Therefore, posterior distribution is approximated by generating samples through the Gibbs sampling algorithm (Geman and Geman 1984). Gibbs sampling is implemented by the program 'Just Another Gibbs Sampler' (JAGS) (Plummer 2003, 2015). The first step of using JAGS is to declare the probabilistic model following the syntax in Plummer (2015), as shown in Appendix I for the proposed HBM model. Gibbs sampling is then performed through the 'rjags' package (Plummer et al. 2019) in the R software.

During Gibbs sampling, each unobserved parameter is sampled iteratively, with all other parameters being fixed. For example, the conditional distribution of a certain statistical parameter in $\boldsymbol{\theta}$ is $p(\theta_i|\boldsymbol{y}, \boldsymbol{\xi}, \boldsymbol{\theta}_{\sim \boldsymbol{i}}, \boldsymbol{\psi}) \propto p(\boldsymbol{\xi}|\theta_i, \boldsymbol{\theta}_{\sim \boldsymbol{i}}) p(\theta_i, \boldsymbol{\theta}_{\sim \boldsymbol{i}}|\boldsymbol{\psi})$. This illustrates a key feature of HBM, which is partial pooling, where the site-specific parameter is deduced both through information from its corresponding site ($\boldsymbol{\xi}$), and through indirect information from all other sites by the common hyperparameters ($\boldsymbol{\psi}$). In general, JAGS uses the slice sampling technique (Neal 2003) to sample from a conditional distribution. Furthermore, truncation can be imposed on the mean and standard deviation parameters, such that only positive values can be sampled. The iteration process continues until the distribution of Gibbs samples become stable. Based on the Gibbs samples, the sample mean approximates the mean of the posterior distribution, with the posterior uncertainty expressed as a credible interval (CI) or credible region (CR). After fitting the HBM model, it is important to examine the convergence of Gibbs samples, and to evaluate the fitting performance. More details can be found in Appendix II.

3 HBM of shear strength database for Hong Kong saprolitic soil

3.1 Database description

The regional database in this study involves shear strength information of completely decomposed granite (CDG), which are saprolitic soils that resulted from weathering of granitic rock. The database was compiled using the results of 747 multi-stage consolidated-undrained (CU) triaxial compression tests, conducted on samples retrieved from 208 slope projects across Hong Kong. Most of the tests were conducted as parts of the geotechnical investigation programme for slope upgrading works in Hong Kong. These slopes are all registered under the Hong Kong Slope Information System, which denotes 16 grids across the city, each sub-divided into northwest, northeast, southwest and southeast sub-grids. The sub-grids are further divided into four zones, and the slope projects considered in this study are located in 47 different zones, according to their locations registered in the system (Fig. 2). These 47 zones are treated as 'sites' in the context of HBM in this study. The size of each zone is 3.75 km $\times$ 2.4 km, therefore the threshold distance between two slope projects within a zone is 3.75 km. For the zone with the largest amount of slope projects (zone 11NW-A, 25 slopes), the mean distance between two slope projects is 1.3 km.

The procedures of multi-stage triaxial tests follow the guidelines stipulated in GEO (2017). For each test, the soil specimen underwent three stages of consolidation and compression, each under a different effective confining pressure. During the compression stage, the principal stresses that correspond to the maximum stress ratios (t/s') are adopted to represent the peak strength conditions. Since the tests in this database mostly arise from projects with slope heights of less than 10 m, for most of the triaxial tests the effective confining pressures are lower than 200 kPa. Also, potential outliers are identified and removed from the database. These outliers often involve soil specimens

containing pieces of corestones that led to exceptionally high shear strengths.

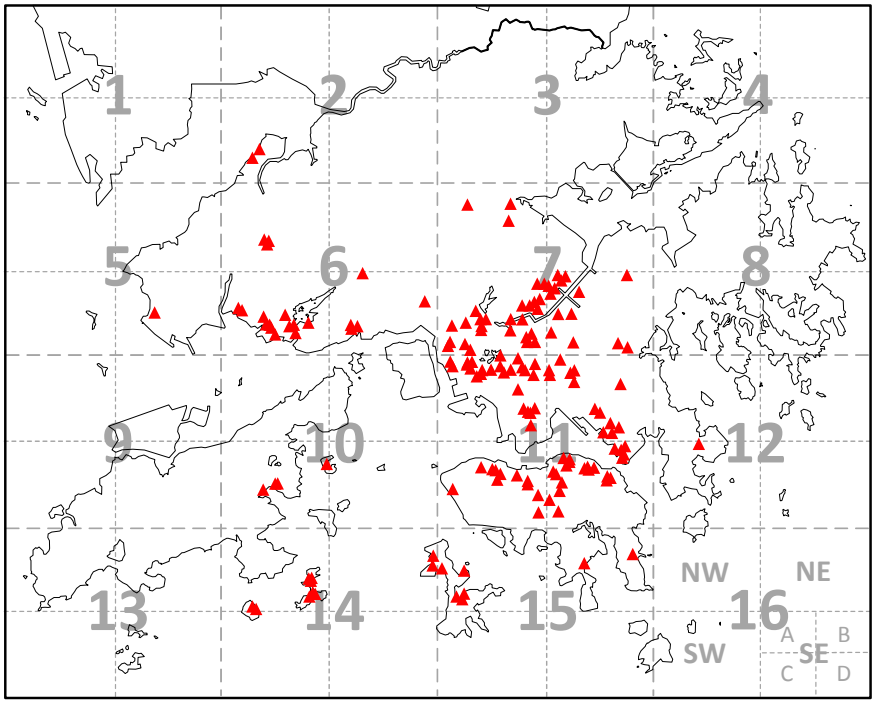

Figure 2: Locations of slope projects where CDG data were obtained.

3.2 Identical parameter model and individual parameter model

To gain a prior understanding of the differences in statistical properties between the sites, before fitting the hierarchical model, the "identical parameter model" (complete pooling) and the "individual parameter model" (no pooling) are fitted to the data, respectively. An identical parameter model means that the entire region follows an identical PDF with a single set of statistical parameters. An individual parameter model means the statistical parameters of each site are fitted separately from the data of each site. Fig. 3 presents the fitted statistical parameters and their 95% CIs. Under the identical parameter model, the CIs are narrow due to the large number of data. However, these intervals cannot capture the statistical properties of many individual

sites. For sites with sufficient data (>20), many of the individual estimates of μ_c, μ_ϕ and σ_ϕ deviate from the estimates of identical parameter model. Therefore, it is worthwhile to model "between-sites variations" of the statistical parameters.

Although the $c' - \phi'$ correlation of many sites are not significant (less than 0.5), for sites with more than 20 data, many CIs of $\rho_{c-\phi,i}$ tend towards the positive values. In addition, a slight positive correlation exists across the entire dataset ($\rho_{c-\phi} = 0.32$ under identical parameter model). Similar patterns of positive $c' - \phi'$ correlations in Hong Kong soils have been observed by Au (1993), who found that both parameters vary in a similar trend with the degree of weathering for the saprolitic soil, as both c' and ϕ' diminish with the process of decomposition. In particular, they attributed the reduction of cohesion to the breakdown of bonding in the parent rock through microfracturing during the weathering process. They also explained that the friction angle decreases with weathering process due to the decomposition of feldspar grains into halloysite and then kaolinite, and the particle alignment of kaolinite at the advanced stages of decomposition. A recent study by Chan et al. (2022) also indicated correlated $c' - \phi'$ values for other soil types in Hong Kong.

Fig. 4 shows the 95% CR and the scatterplots of $c' - \phi'$ values for each individual site. It provides another illustration that the $c' - \phi'$ distribution of each site may differ from the regional distribution, both in terms of the position of CR and the scattering of data points.

3.3 Fitting of the HBM model

The regional database is then analysed with the presented HBM framework, also referred to as partial pooling model here. Table 2 summarizes the posterior distributions of the hyperparameters, evaluated from the Gibbs samples. The m values represent the average statistical parameters among the sites; while the p values quantify the variability

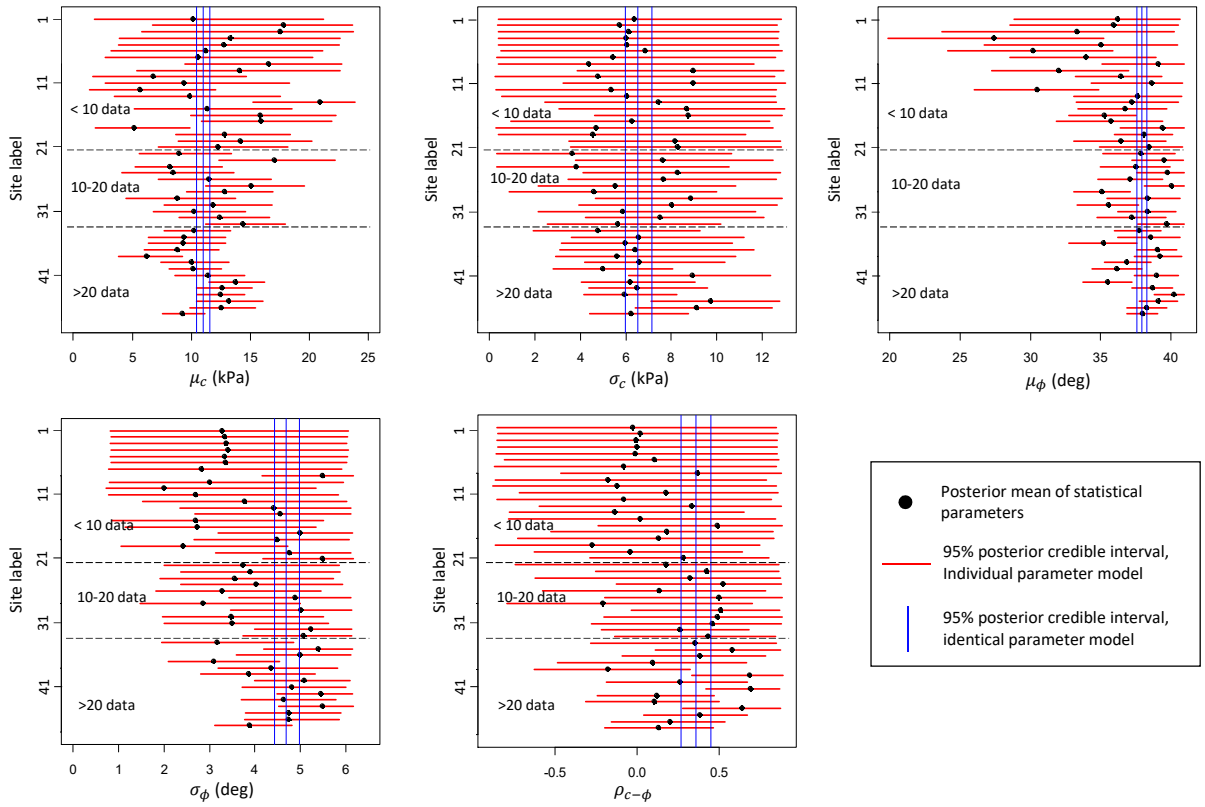

Figure 3: Statistical parameters under identical parameter model and individual parameter model

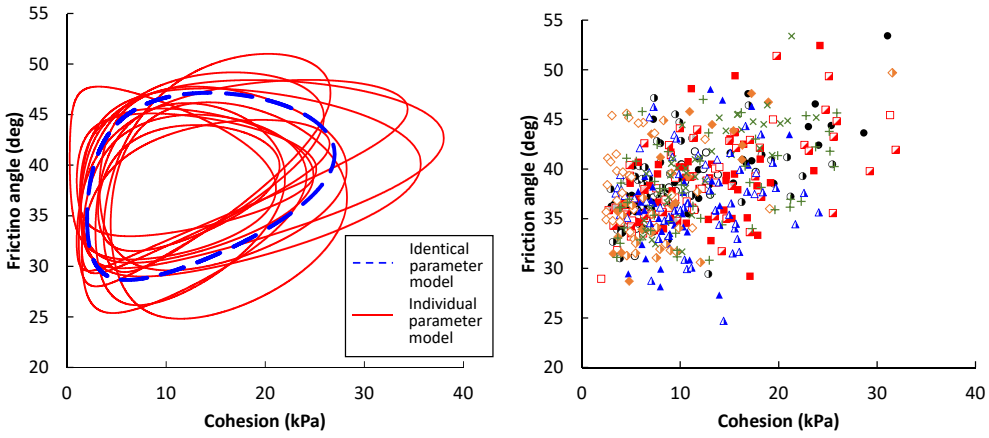

Figure 4: (a) 95% credible regions and (b) scatterplot of c'-ϕ' values under individual parameter model (for sites with more than 20 test data)

Table 2: Posterior distributions of the HBM parameters

Parameter	Posterior mean	Posterior SD	Posterior 95%CI
m_{μ_c} (kPa)	11.1	0.43	[10.2, 11.9]
p_{μ_c} (kPa^{-2})	0.387	0.279	[1.17, 3.07]
σ_c (kPa)	6.26	0.29	[5.7, 6.9]
m_{μ_ϕ} ($^\circ$)	37.7	0.31	[37.1, 38.3]
p_{μ_ϕ} ($^{\circ-2}$)	0.659	0.419	[0.238, 1.695]
m_{σ_ϕ} ($^\circ$)	4.40	0.21	[3.98, 4.80]
p_{σ_ϕ} ($^{\circ-2}$)	2.927	5.436	[0.544, 15.38]
m_ρ	0.328	0.060	[0.213, 0.445]
p_ρ	156.2	255.5	[9.743, 948.9]
p_ε (kPa^{-2})	0.261	0.013	[0.236, 0.287]

of statistical parameters among the sites.

Fig. 5 presents the fitted statistical parameters and the corresponding 95% CIs of each site. The mean values of c' and ϕ', and standard deviation of ϕ' of the individual sites show noticeable variations from their respective average values. This is more apparent for sites with more than 10 data. For example, the black symbols in Fig. 5 show that $\mu_{c,i}$ can vary between 8 and 13 kPa, from the average of 11.1 kPa; $\mu_{\phi,i}$ can vary between 36°and 40°, from the average of 37.7°. $\sigma_{\phi,i}$ can vary between 3.5°and 5.5°, from the average of 4.4°. $\rho_{c-\phi}$ of individual sites also display variation, with $\rho_{c-\phi,i}$ varying between 0.2 to 0.4, from the average of 0.328. Comparing HBM with the individual parameter model (i.e., Fig. 5 versus Fig. 3), application of HBM reduces the overall scatter in statistical estimates, where the CIs of each site converge towards the higher level parameters, i.e., hyperparameters $m_{\mu_c}, m_{\mu_\phi}, m_{\sigma_\phi}, m_\rho$. And as expected, the CIs are narrower for sites with more test data.

The CIs of statistical parameters for a site without any test data aren also evaluated, which are shown as vertical dashed lines in Fig. 5. From the CIs, the statistical parameters of $c' - \phi'$ for CDG at a Hong Kong site are expected to lie within the following ranges: $\mu_c \in [7.2, 15.0]$ kPa; $\sigma_c \in [5.7, 6.9]$ kPa; $\mu_\phi \in [34.8^\circ, 40.4^\circ]$; $\sigma_\phi \in [2.6^\circ, 6.1^\circ]$; and

$\rho_{c-\phi} \in [-0.027, 0.66]$. The coefficient of variation of c' ranges from 0.41 to 0.88, while the coefficient of variation of ϕ' ranges from 0.07 to 0.165. The ranges of μ_ϕ and coefficient of variation of ϕ' agree with that reported in Phoon and Kulhawy (1999), which are 35°-41°and 0.05-0.11 considering laboratory tests of multiple sand groups, although it should be noted that ϕ' in their study is likely to be secant friction angle.

Next, the posterior means of c' and ϕ' are evaluated, which are plotted in Fig. 6(a) as scatterplot, and in Figs. 6(b) and (c) as histograms. The 95% CR of c' and ϕ' for each site is also shown in Fig. 6(a). These CRs illustrate the variation of distributions of shear strength parameters for various zones across Hong Kong. The marginal PDFs of c' and ϕ' for each site are shown in Figs. 6(b) and (c). It can be seen that gamma distribution and normal distribution reasonably describe the variation of c' and ϕ', respectively. Furthermore, the regional distribution of $c' - \phi'$, which is the pooled distribution of all individual sites, is shown as blue in Fig. 6. Denoting the statistical parameters of regional distribution by subscript "r", $\{\mu_{c,r},\ \sigma_{c,r},\ \mu_{\phi,r},\ \sigma_{\phi,r},\ \rho_{c-\phi,r}\} = \{11.1$ kPa, 6.53 kPa, 37.6°, 4.76°, 0.286$\}$. The relationships between the regional statistics and site statistics can be derived through the law of total expectation and law of total covariance:

$$
\begin{aligned}
\mu_{c,r} &= E(\mu_c) \\
\sigma_{c,r} &= \sqrt{E(\sigma_c^2) + Var(\mu_c)} = \sqrt{E\left(\sigma_c\right)^2 + SD\left(\sigma_c\right)^2 + SD\left(\mu_c\right)^2} \\
\mu_{\phi,r} &= E(\mu_\phi) \\
\sigma_{\phi,r} &= \sqrt{E(\sigma_\phi^2) + Var(\mu_\phi)} = \sqrt{E\left(\sigma_\phi\right)^2 + SD\left(\sigma_\phi\right)^2 + SD\left(\mu_\phi\right)^2} \\
\rho_{c-\phi,r} &= \frac{E(\sigma_c \sigma_\phi \rho_{c-\phi}) + COV(\mu_c, \mu_\phi)}{\sigma_{c,r}\sigma_{\phi,r}} = \frac{E(\sigma_c)E(\sigma_\phi)E(\rho_{c-\phi})}{\sigma_{c,r}\sigma_{\phi,r}}
\end{aligned}
\tag{6}
$$

where $E(\ldots)$ and $SD(\ldots)$ denote the posterior mean and standard deviation of a site statistical parameter, and COV is the covariance. From Eq. (6), the regional means for c' and ϕ' are equal to the respective site average $E(\mu_c)$ and $E(\mu_\phi)$. The regional

standard deviations for c' and ϕ' are slightly larger than the respective site average $E(\sigma_c)$ and $E(\sigma_\phi)$. The variations of PDF of $c' - \phi'$ between sites, together with the regional PDF, will serve as references in the interpretation of presented approach in subsequent sections.

The residual precision associated with fitting $s' - t$ data to $c' - \phi'$ parameters, i.e. p_ε, is around of 0.261 kPa^{-2}. In other words, the residual standard deviation σ_ε is around 1.98 kPa, which is small compared to the peak stress values. To examine the fitted regression line for (s', t) data points to obtain $c' - \phi'$ values, peak stress points of four typical sets of trixaxial test are plotted in Fig. 7, with the fitted lines under the hierarchical model and no pooling model, respectively. Under the HBM model with partial pooling of data, the mean fitted line deviates slightly from the direct regression line of three stress points. In Fig. 7(a)(b), the c' intercept of the mean fitted line is higher, while ϕ' is lower than that of the no pooling model, while the opposite occurs in Fig. 7(c)(d). For all triaxial tests, the 95% prediction intervals under the HBM model are relatively narrow.

Finally, to select the optimal HBM model, the goodness-of-fit of various model structures is evaluated by the Widely applicable Bayesian information (WAIC) criterion (Watanabe 2013). WAIC is conceptually similar to the Akaike's information criterion (AIC), which consists of a model fitness term called expected deviance ($\bar{D}$), and a correction term (p_{WAIC}) for effective number of parameters to adjust for overfitting. The mathematical details and implementation of WAIC can be found in Appendix II. Models of increasing complexity are fitted, starting with the identical parameter model, followed by the HBM models with increasing numbers of site-specific parameters. For each fitted model, the WAIC score is evaluated, as summarized in Table 3. The WAIC score shows an obvious improvement compared with the identical parameter model, when site-specific μ_c and μ_ϕ are introduced. The improved WAIC score also

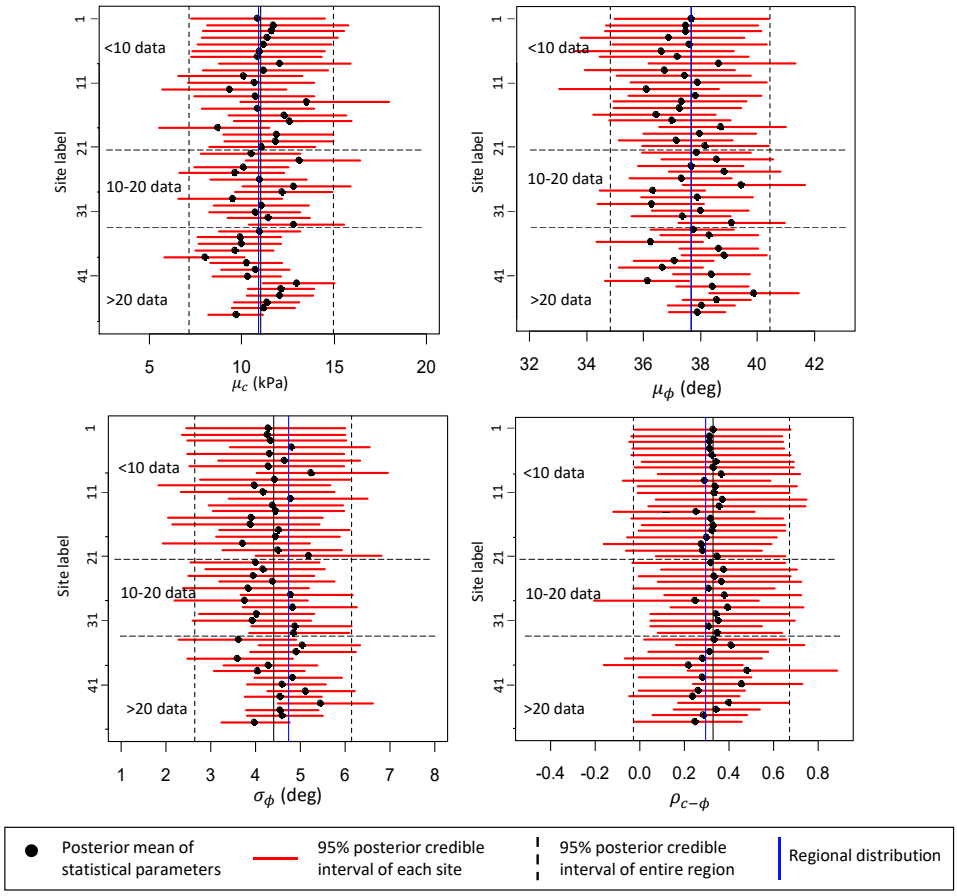

Figure 5: Statistical parameters for each site under the HBM model

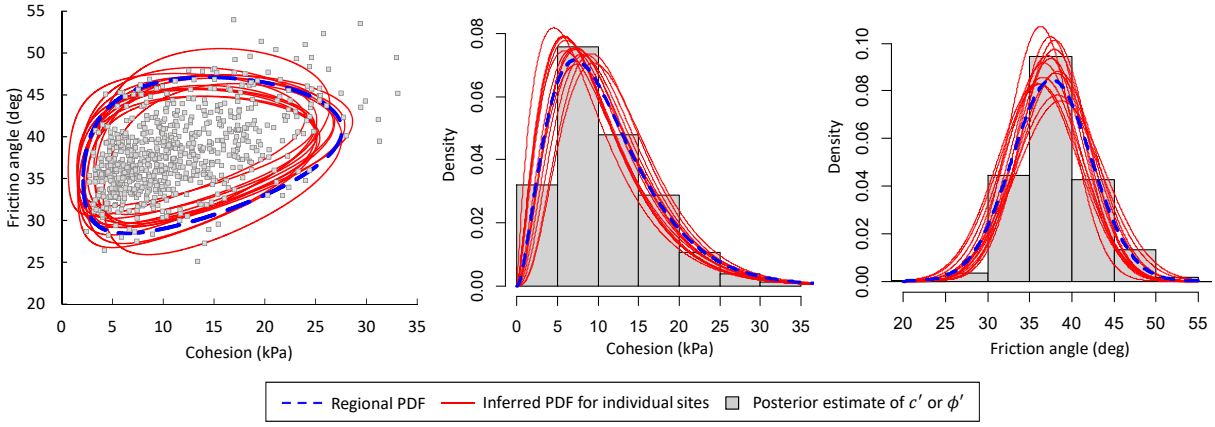

Figure 6: c' and ϕ' parameters for CDG in Hong Kong, under the HBM model

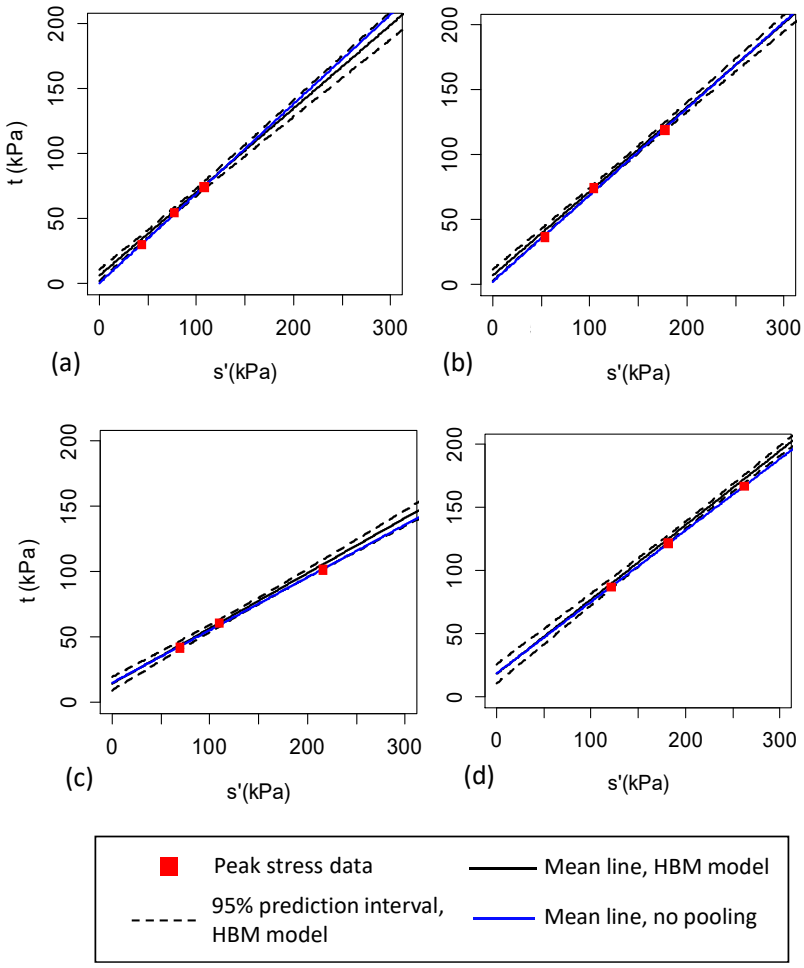

Figure 7: Fitted regression line for four sets of multi-staged triaxial test

serves as a quantitative evidence of the observation in Fig.3, where many estimates of μ_c and μ_ϕ deviates from the identical parameter model. Incorporation of site-specific μ_c and μ_ϕ also carries practical significance for slope reliability analysis, since changes in μ_c and μ_ϕ shift the entire distribution of $c' - \phi'$, which in turn affects the failure probability. Although not as significant, including site-specific σ_ϕ and $\rho_{c-\phi}$ in the model also improves the WAIC score. Past studies have shown that $c' - \phi'$ correlation can heavily influence the system reliability (Fenton and Griffiths 2003; Jiang et al. 2014). By including site-specific $\rho_{c-\phi}$ in the HBM model, $\rho_{c-\phi}$ of a site can be adjusted if there is strong site evidence, thus the reliability assessment better reflects the site condition.

Table 3: WAIC scores of various probabilistic models

Model	D ($\times 10^5$)	p_{WAIC} ($\times 10^5$)	WAIC ($\times 10^5$)
Identical parameter model	4.143	4.032	8.175
Site-specific μ_c, μ_ϕ; Common $\sigma_c, \sigma_\phi, \rho_{c-\phi}$	3.726	3.625	7.351
Site-specific $\mu_c, \mu_\phi, \sigma_\phi$; Common $\sigma_c, \rho_{c-\phi}$	3.696	3.596	7.293
Site-specific $\mu_c, \mu_\phi, \sigma_c, \sigma_\phi$; Common $\rho_{c-\phi}$	3.700	3.600	7.300
Site-specific $\mu_c, \mu_\phi, \sigma_\phi, \rho_{c-\phi}$; Common σ_c	3.691	3.594	7.285
Site-specific $\mu_c, \mu_\phi, \sigma_c, \sigma_\phi, \rho_{c-\phi}$	3.698	3.599	7.297

However, further inclusion of site-specific σ_c does not improve the WAIC score. Based on the WAIC criteria, the HBM model with site-specific $\mu_c, \mu_\phi, \sigma_\phi, \rho_{c-\phi}$ and a common σ_c is selected. The common σ_c value of each site is around 6.26 kPa.

4 Reliability assessment with fitted HBM

4.1 Hybridising regional and site-specific information

Hierarchical Bayesian modelling has the potential to augment the limited site-specific data to improve reliability assessment. The fitted HBM model can be utilised to establish informative prior for the statistical parameters of a new site. The newly available site data can be asssimilated with the informative prior through Bayesian approach, leading to a posterior predictive distribution of shear strength parameters. For meaningful assimilation of information, the site under investigation should belong in the same region with similar soil types as the regional database. This concept of posterior predictive distributions (Bozorgzadeh et al. 2023) were sometimes referred to as 'Bayesian equivalent samples' (Wang et al. 2016) or 'quasi-site-specific PDF' (Ching et al. 2021b,a) in the literature.

Specifically, the informative prior is evaluated from the posterior distribution of the hyperparameters, where $\mu_c, \mu_\phi, \sigma_\phi, \rho_{c-\phi}$ are simulated from the Gibbs samples of the hyperparameters. Meanwhile, the informative prior of σ_c is evaluated directly from the Gibbs samples of σ_c. It is found that normal distribution reasonably describes the

Table 4: Non-informative priors and informative priors for the statistical parameters of $c' - \phi'$. Non-informative priors for σ_ϕ and $\rho_{c-\phi}$ are extended to accommodate the features of CDG database in this study

Statistical parameter	Non-informative prior based on Wang and Akeju (2016)	Informative prior from HBM fitting
μ_c, in kPa	U(0.1, 24)	N(11.1, 1.94)
σ_c, in kPa	U(0.1, 13.2)	N(6.26, 0.29)
μ_ϕ	U(17°, 41°)	N(37.7°, 1.41°)
σ_ϕ	U(0.7°, 6.2°)	N(4.41°, 0.86°)
$\rho_{c-\phi}$	U(-0.9, 0.5)	N(0.327, 0.167)

informative prior. These informative priors govern the variation of statistical parameters for the new site, and Table 4 summarizes their properties.

The probabilistic model for assimilation of site information is shown in Fig. 8. Considering l sets of triaxial test data available at the new site, the stress points of the triaxial test results are denoted by $\boldsymbol{s'_k} = [s'_{k,1}, \ldots, s'_{k,l}]$ and $\boldsymbol{t_k} = [t_{k,1}, \ldots, t_{k,l}]$, with the underlying c' and ϕ' values denoted by $\boldsymbol{c'_s} = [c_{s,1}, ..., c_{s,l}]$ and $\boldsymbol{\phi'_s} = [\phi_{s,1}, ..., \phi_{s,l}]$, respectively. c'_p and ϕ'_p represent the shear strength parameters at the unsampled locations of the new site, which follow the same statistical parameters as $\boldsymbol{c'_s}$ and $\boldsymbol{\phi'_s}$. In

mathematical form,

$$
\begin{aligned}
&\mu_c \sim TN\left(E(\mu_c), SD(\mu_c)\right)(0, +\infty) \\
&\sigma_c \sim TN\left(E(\sigma_c), SD(\sigma_c)\right)(0, +\infty) \\
&\mu_\phi \sim TN\left(E(\mu_\phi), SD(\mu_\phi)\right)(0, +\infty) \\
&\sigma_\phi \sim TN\left(E(\sigma_\phi), SD(\sigma_\phi)\right)(0, +\infty) \\
&\rho_{c-\phi} \sim TN\left(E(\rho_{c-\phi}), SD(\rho_{c-\phi})\right)(-1, +1) \\
&c'_{s,j}|\mu_c, \sigma_c \sim \text{Gamma}\left(\alpha = \frac{\mu_c^2}{\sigma_c^2},\ \beta = \frac{\mu_c}{\sigma_c^2}\right) \\
&\phi'_{s,j}|c'_{s,j}, \mu_\phi, \sigma_\phi, \mu_c, \sigma_c, \rho_{c-\phi} \\
&\sim N\left(\mu_\phi + \rho_{c-\phi}\sigma_\phi \Phi^{-1}[F_c(c'_{s,j})], \sqrt{1-\rho_{c-\phi}^2}\sigma_\phi\right) \\
&t_{k,j} \sim N\left(c'_{s,j}\cos\phi'_{s,j} + s'_{k,j}\sin\phi'_{s,j}, \sigma_\varepsilon\right) \\
&c'_p|\mu_c, \sigma_c \sim \text{Gamma}\left(\alpha = \frac{\mu_c^2}{\sigma_c^2},\ \beta = \frac{\mu_c}{\sigma_c^2}\right) \\
&\phi'_p|c'_p, \mu_\phi, \sigma_\phi, \mu_c, \sigma_c, \rho_{c-\phi} \\
&\sim N\left(\mu_\phi + \rho_{c-\phi}\sigma_\phi \Phi^{-1}[F_c(c'_p)], \sqrt{1-\rho_{c-\phi}^2}\sigma_\phi\right)
\end{aligned} \tag{7}
$$

To perform assimilation of site data, s' and t are fixed to their observed values in the probabilistic model. Statistical parameters of the new site, $\boldsymbol{c'_s}$, $\boldsymbol{\phi'_s}$, c'_p and ϕ'_p are then updated to their posterior distributions. The posterior distributions are evaluated by Gibbs sampling, and the corresponding JAGS code can be found in Appendix III. The Gibbs samples of c'_p and ϕ'_p are used to evaluate the posterior predictive distribution of c' and ϕ'.

In the case where a regional database is unavailable, an alternative way to formulate priors for $c' - \phi'$ is to adopt non-informative uniform priors over reasonable ranges (Cao et al. 2016), which implies that there is no preference within the lower and upper bounds. For example, Wang and Akeju (2016) formulated a non-informative prior of

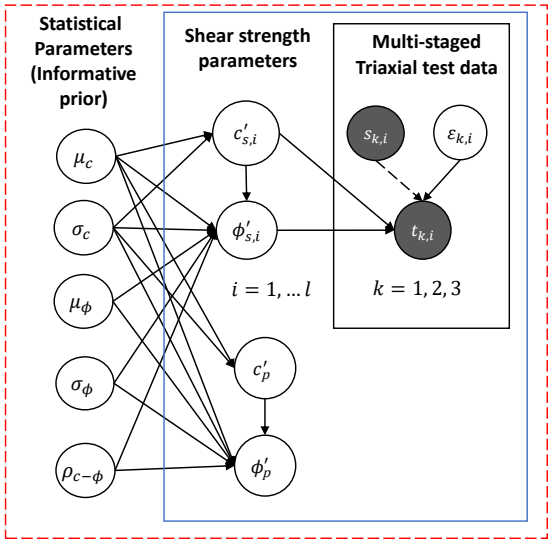

Figure 8: Probabilistic model for assimilation of information at new site

$c' - \phi'$ for sandy soils based on reported data in the literature. The non-informative priors and informative priors are compared in Table 4, with the latter generally resulting in narrower distributions. Under limited site-specific data, it is expected that using informative priors reduces the statistical uncertainty of $c' - \phi'$ more effectively, leading to more robust slope reliability estimates.

4.2 Slope reliability analysis with assimilation of site data

This study adopts the SLOPE/W software (GEO-SLOPE 2017) for slope reliability analysis, where the factor of safety (FOS) is evaluated by limit equilibrium method (LEM). Unlike finite element method or finite difference method which automatically searches for the critical slip surface, the SLOPE/W software allows multiple potential failure surfaces to be pre-defined, and includes the functionality to conduct Monte Carlo simulations for each individual failure surface. In this study, potential failure surfaces are specified through the 'entry-exit' method. For each combination of entry and exit points, multiple circular slip surfaces are generated, with each slip surface having a

different radius.

To evaluate the system reliability of the slope, the FOS values are extracted for each slip surface in the Monte Carlo process, and combined in a Microsoft Excel spreadsheet, through which the minimum FOS and the corresponding critical slip surface are determined for each $c' - \phi'$ combination. The system failure probability is then determined through the distribution of minimum FOS among all slip surfaces in the Monte Carlo simulations.

Although these procedures associated with LEM and post-processing of results are less common than finite element simulations in the literature of slope reliability research, there are certain advantages with the use of SLOPE/W. The approach does not require additional scripting or subroutines, and the post-processing procedures can be performed using spreadsheets which are easily accessible to practitioners. It also allows the failure probability of each individual slip surface (with different failure volumes) to be clearly elucidated. This proposed approach will be demonstrated through the reappraisal of the reliability of a slope case.

5 Ching Cheung Road slope case

5.1 Site description

The case study involves a cut slope located above Ching Cheung Road, Hong Kong. A major slope failure occurred on 3 August 1997, shortly after a major rainfall event associated with Typhoon Victor (Halcrow 1998). Although the slope failed eventually, this study does not intend to perform a back analysis of the failed slope to calibrate the $c' - \phi'$ parameters. Instead, the aim is to reassess its reliability using available information prior to the event. The geological cross-section before slope failure is shown in Fig. 9. The majority of soil materials below the slope surface were completely

decomposed granite (CDG), followed by highly decomposed granite (HDG). At the slope surface is a layer of colluvium up to 2.5 m in thickness, and a thin layer of residual soil. Basalt dykes were identified in several boreholes in the area, with one basalt dyke possibly spanning across tens of metres near the base of the slope. The permeability of the basalt dyke was much lower than the surrounding granite and it possibly acted as an aquitard that inhibited the downward movement of groundwater. Natural erosion pipes of quartz-rich sands were identified throughout the CDG, which might have caused piezometric pressure to build up above the pipes (Hencher 2010). Based on field observations, the most likely piezometric head before the occurrence of landslide was about 18 m above the basalt dyke at the base. Ground investigations was carried out in nearby areas in the 1980s, with 25 specimens of CDG retrieved from 14 boreholes. Multi-stage consolidated-undrained triaxial compression tests were conducted on those soil specimens, under similar procedures described earlier in the establishment of regional database. Table 5 summarises the shear strength information on the specimens. It should be noted that the shear strength information of this site had not been included in the regional database.

5.2 Slope reliability assessments based on regional PDF and posterior predictive distribution

To assess the slope reliability, CDG parameters are treated as random variables, while the parameters of colluvium and residual soil are treated as deterministic, as the slip surfaces only pass through small portions of these layers. The HDG parameters are also assumed deterministic, since a database for HDG is not yet available. Based on geotechnical investigation reports of the site, c' and ϕ' of colluvium are taken as 7.9 kPa and 34.2°; c' and ϕ' of residual soil are taken as 7.6 kPa and 32.4°; c' and ϕ' of HDG are taken as 15.0 kPa and 39.0°. The unit weight of all soil layers are taken as 19 kN/m^3.

Table 5: Site-specific shear strength data of CDG

Borehole	Distance from slope (m)	Sample depth (m)	Triaxial test data ($s'-t$ at max. stress ratio) s'_1 (kPa)	s'_2 (kPa)	s'_3 (kPa)	t_1 (kPa)	t_2 (kPa)	t_3 (kPa)	Fitted values c'_s (kPa)	ϕ'_s (deg)
DH1	224	4.5	74	118	164	46	70	89	9.5	30.5
DH1	224	9.0	121	192	291	81	121	183	9.8	36.8
DH1	224	9.0	144	267	402	95	168	248	11.8	36.4
DH2	194	3.0	46	71	113	29	42	68	4.8	34.0
DH2	194	9.0	155	214	310	105	137	191	19.0	34.6
DH3	161	4.5	102	149	202	67	93	124	10.0	35.0
DH4	148	4.5	63	103	151	40	65	93	6.2	35.5
DH4	148	9.0	201	312	429	139	213	275	24.4	37.1
DH5	143	4.5	95	156	216	67	103	137	13.8	35.8
DH5	143	9.0	135	223	323	92	141	198	16.9	34.8
DH7	74	6.8	171	279	399	124	195	269	19.9	39.6
DH8	101	9.0	181	286	426	127	193	274	22.6	37.2
DH8	101	17.0	189	261	351	120	156	199	25.4	30.7
DH9	54	6.0	127	238	356	101	173	251	22.8	41.0
DH10	48	1.5	114	139	181	76	90	112	13.1	34.4
DH10	48	6.0	164	262	379	118	184	260	14.5	41.1
DH10	48	6.0	143	254	339	97	106	215	4.7	33.7
DH11	26	3.0	73	122	180	50	82	119	7.9	38.5
DH11	26	9.0	211	291	406	149	201	272	19.9	39.3
DH12	37	5.0	137	224	330	99	152	219	16.5	38.6
DH13	37	4.0	86	132	203	59	87	133	8.3	38.1
DH13	37	7.0	148	262	383	104	181	257	12.5	40.4
DH14	75	5.5	111	179	253	77	118	163	12.3	37.2
DH15	107	5.5	101	160	216	66	96	123	14.2	31.4
DH15	107	8.5	186	313	462	133	211	301	23.4	37.8

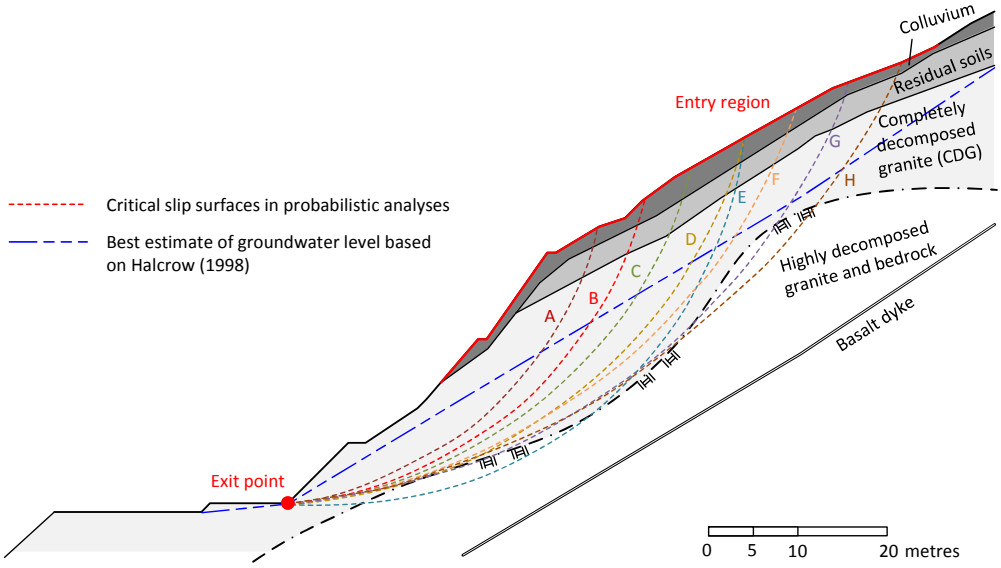

Figure 9: SLOPE/W model of the Ching Cheung Road slope case

The entry region and exit point of the slip surfaces for SLOPE/W analyses are shown in Fig. 9. The Morgenstern-Price method with half-sine interslice force function is adopted in the analyses.

The first set of probabilistic analysis utilises the regional database assuming there are no site-specific data, with the $c' - \phi'$ parameters of CDG follow the regional PDF. Monte Carlo simulation is performed with a simulation size of 4000. The resulting probability distribution of FOS, and the contour plot of FOS overlying the $c' - \phi'$ distribution are in Fig. 10 and Fig. 11(a) respectively. The mean and standard deviation of FOS are 0.896 and 0.134, respectively. The system failure probability is 0.787, calculated as the ratio of the number of cases with FOS < 1 to the total number of cases. Out of 95 potential slip surfaces, 8 of them are identified as dominant and are labelled as Slips A to H in Fig. 9. In general, shallower slip surfaces are associated with small c' values, while the relatively deep slip surfaces are associated with higher c' and lower ϕ' values. Failure cases are mainly contributed by Slips B, C and D, which account for 15.5%, 27.1% and

52.1% of the cases, respectively.

To demonstrate the significance of assimilation of site-specific data, the site information shown in Table 5 are incorporated to produce the posterior predictive distribution of $c' - \phi'$ parameters. This distribution, together with the fitted $c' - \phi'$ values from the triaxial test data, are plotted in Fig. 11(b). The statistics of the posterior distribution are $\{\mu_{c,p}, \sigma_{c,p}, \mu_{\phi,p}, \sigma_{\phi,p}, \rho_{c-\phi,p}\} = \{13.3 \text{ kPa}, 6.5 \text{ kPa}, 36.6^\circ, 3.7^\circ, 0.352\}$. Compared to the regional PDF, the posterior predictive distribution involves higher mean c' and lower mean ϕ', and smaller standard deviations for both c' and ϕ'.

Reliability assessment of the slope is then performed based on the posterior predictive distribution of $c' - \phi'$, with the resulting FOS distribution and FOS contour shown in Fig. 10 and Fig. 11(b). The mean FOS estimates are similar considering regional or posterior predictive distribution, but a smaller standard deviation of FOS is observed for the latter, mainly due to the smaller standard deviations in the posterior distribution of c' and ϕ'. As a result, the failure probability becomes slightly higher when site-data is assimilated. Comparison between the FOS contours also reveals the difference in failure mechanisms. Considering both regional and site data, the deeper slip surfaces are expected to be more influential, as Slips B, C, D now contribute to 10.8%, 21.3%, 60.5% of the failure cases, respectively.

5.3 Effects of site data size on slope reliability assessment

This section investigates the influence in slope reliability when incorporating different amounts of site data, through a resampling method called bootstrapping, which has been recently applied to probabilistic analyses of various civil engineering problems (Ching and Phoon 2013; Li et al. 2015; Wang et al. 2016; Liu et al. 2020). Bootstrapping involves random sampling with replacement within the pool of available data, which can evaluate the effect of various sampling patterns. For example, to evaluate the

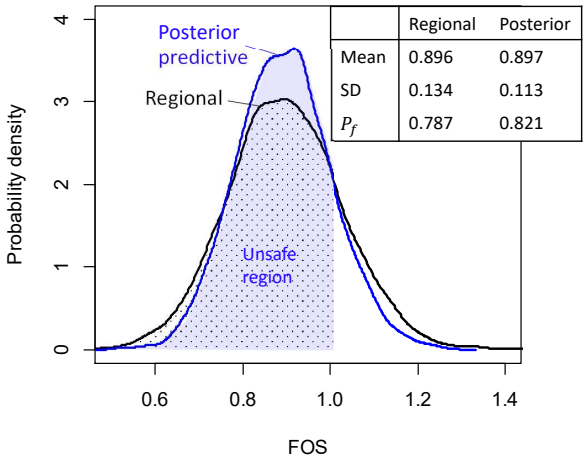

Figure 10: FOS distribution evaluated from regional or posterior predictive $c' - \phi'$.

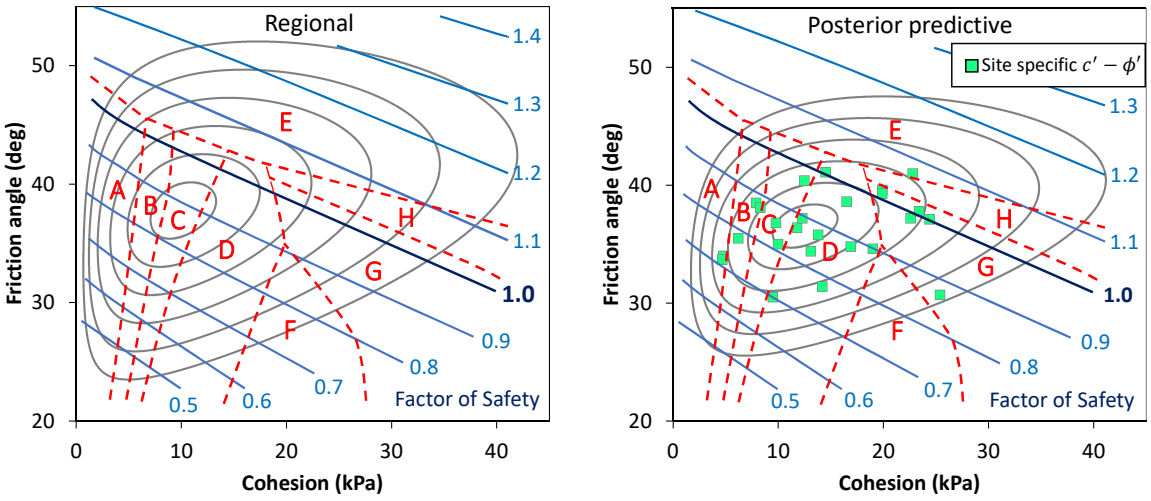

Figure 11: (a)(b) FOS contour plot under posterior predictive $c' - \phi'$. The interval between the PDF contours is 0.5SD. The slip surfaces are indicated in red.

variability that arises from the sampling of 5 site data, 5 data samples can be drawn with replacement from the $c' - \phi'$ combinations presented in Table 5. The 5 samples then form a set of 'bootstrap samples', which are utilised to evaluate a posterior predictive distribution of $c' - \phi'$. In this study, the bootstrapping process is repeated for 50 times, which produces 50 posterior distributions to illustrate the variability that arises from how (or where) the 5 site data are sampled.

Fig. 12 shows the posterior predictive distributions of c' and ϕ' by incorporating 1, 5 or 25 site-specific data. With a single site-specific data, the posterior distributions only deviate slightly from the regional PDF. When 5 site data points are considered, a systematic deviation of posterior distributions from the regional PDF can be observed, which becomes even more pronounced when 25 site data are incorporated. The resulting FOS distributions from slope analyses also shift away from the regional distribution as more site data are incorporated.

Fig. 13 displays the statistics of the posterior predictive distributions of $c' - \phi'$, when 1 to 25 site-specific data points are incorporated. As more site data are incorporated, $\mu_{c,p}$ deviates towards higher values compared with the regional statistics, while $\mu_{\phi,p}$ deviates towards lower values, leading to more influences from the deep slip surfaces. Fig. 14 shows the mean values and standard deviations of FOS, and the corresponding probability of failure, evaluated from the posterior predictive distributions of $c' - \phi'$. The ranges of slope reliability estimates arise from the different possible sampling patterns under the different sets of bootstrap samples. On average, a steady trend of reliability estimates can be observed. As the standard deviation of FOS gradually reduces, the probability of failure increases with the amount of site data, deviating from the regional failure probability in this particular case.

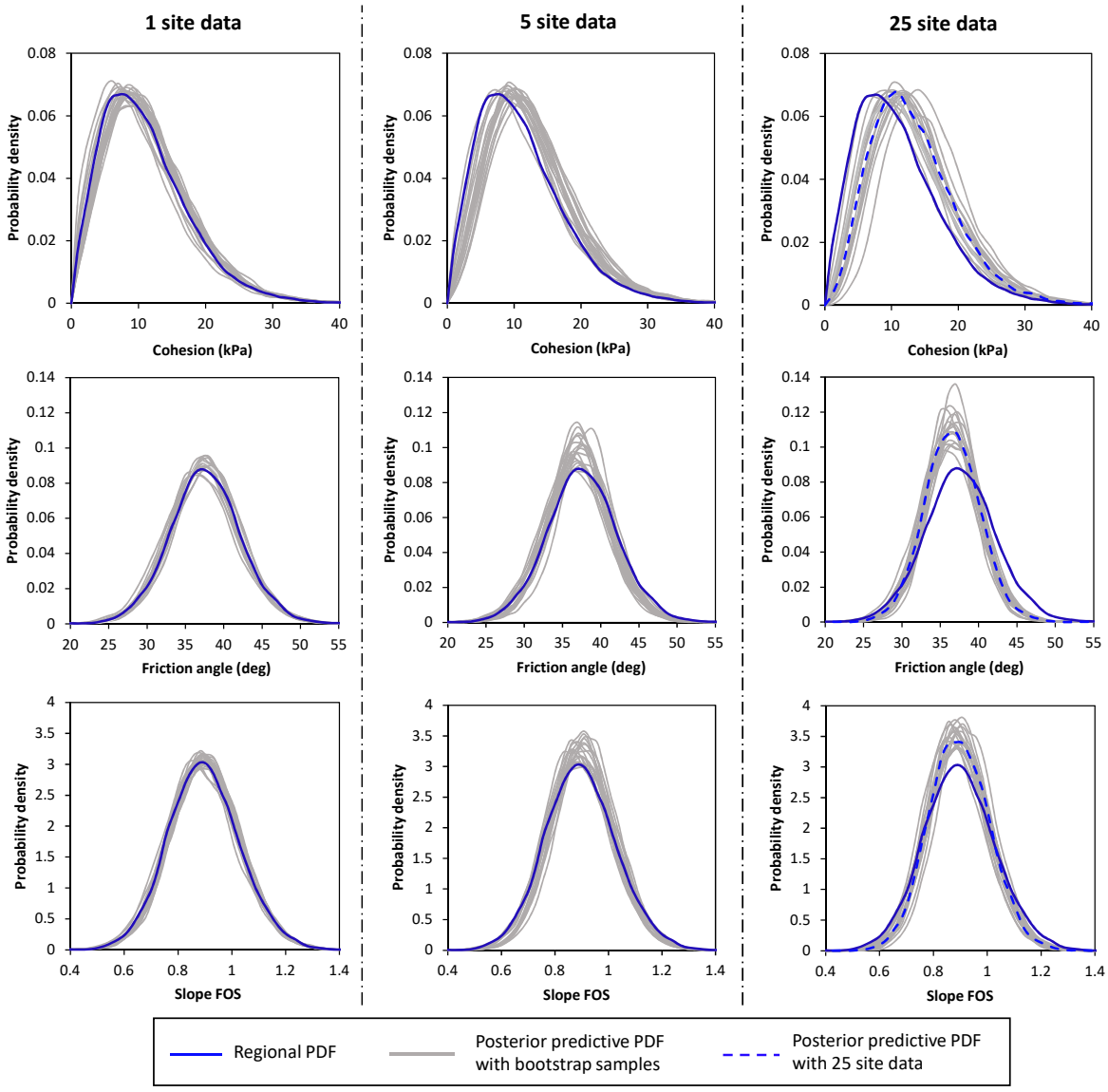

Figure 12: Posterior predictive distributions of $c' - \phi'$ and slope FOS considering various amounts of site data

5.4 Slope reliability based on non-informative priors

In the absence of a regional database, PDFs of c' and ϕ' may be estimated solely from the site-specific data, and considering the non-informative priors in Table 4. Bayesian analysis is performed following a similar probabilistic model as Fig. 8, except that the statistical parameters $(\mu_c, \sigma_c, \mu_\phi, \sigma_\phi, \rho_{c-\phi})$ are now governed by the non-informative

uniform priors. The same set of bootstrap samples as in the earlier section is used to obtain the posterior distributions of $c' - \phi'$. The posterior statistics of $c' - \phi'$ are then compared with those obtained from informative priors in Fig. 13. Without the contributions from regional database (i.e., prior knowledge of site statistics), adopting non-informative priors leads to larger statistical uncertainty, which is manifested into wider deviations of the $c' - \phi'$ statistics. Also, $\rho_{c-\phi}$ is slightly negative under limited site data, and then increases as more site data are included. This is because the non-informative prior of $\rho_{c-\phi}$ is more on the negative side, as negative $\rho_{c-\phi}$ was widely reported in literature. Conversely, the $\rho_{c-\phi}$ estimates under the informative prior remain nearly unchanged.

Reliability of the slope is assessed again based on the $c' - \phi'$ PDFs obtained from the non-informative priors. The results are then compared with those from the informative priors in Fig. 14. In theory, with an abundant amount of site data, reliability estimates should be similar under both informative and non-informative priors, as the importance of priors diminishes. However, with limited site-specific data, the patterns of reliability estimates are substantially different. Under the non-informative priors, the reliability estimates appear to be less robust under the various sampling patterns. The mean FOS display large fluctuations under limited amount of site data, hence leading to a wide range of failure probability estimates. For example, the first and third quartiles of failure probability are 0.67 to 0.92 under 1 site-specific data, and 0.71 to 0.86 under 5 site-specific data.

On the other hand, under the informative priors, a steady trend of slope reliability can be observed, with the reliability estimates gradually deviating from the regional conditions. The approach also stabilises the reliability estimates by reducing the statistical uncertainty that arises from the limited amount of site-specific data. The robustness brought about by this approach is noteworthy, considering that many slope

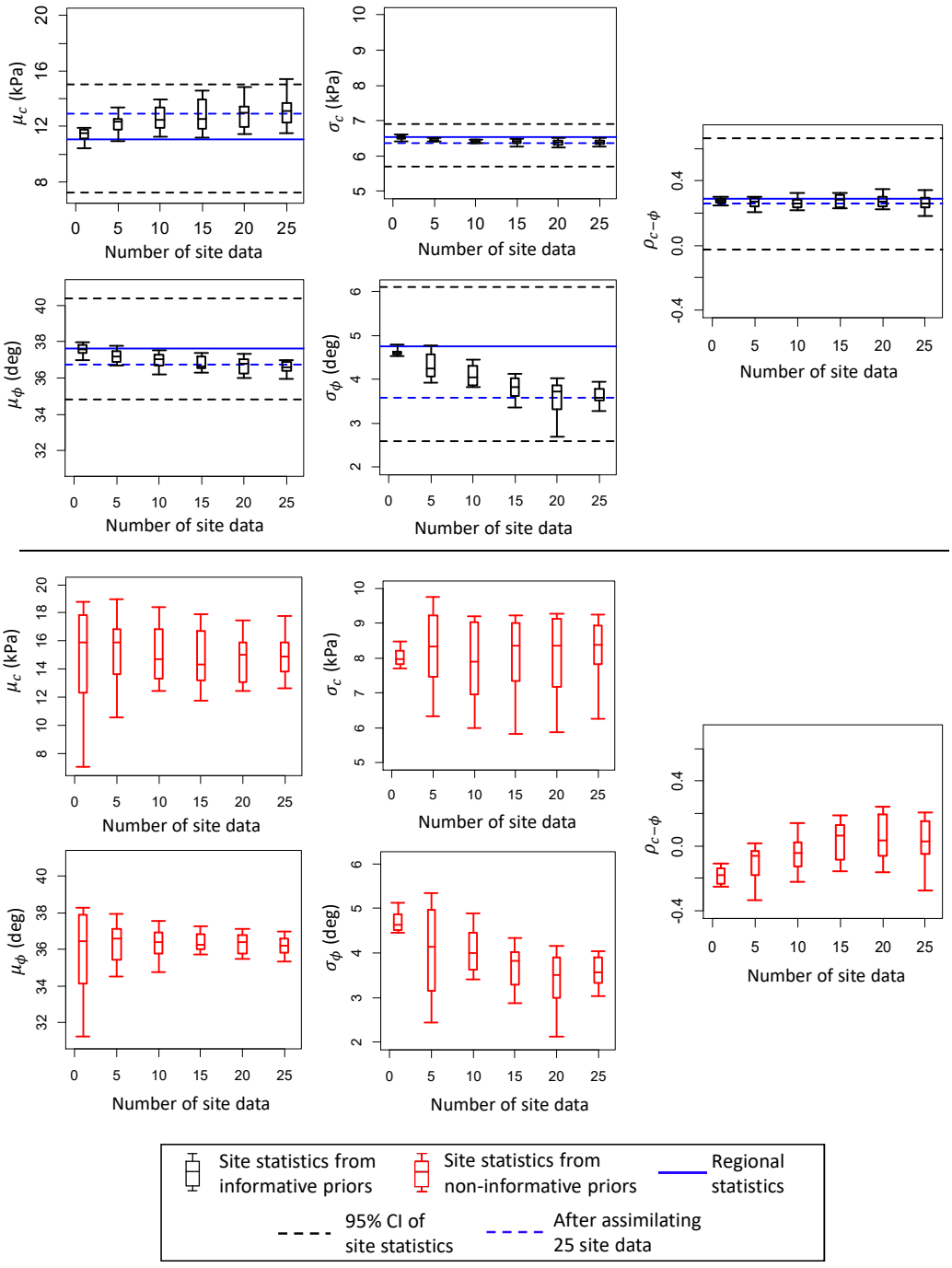

Figure 13: Statistical parameters of $c' - \phi'$ considering various amounts of site data

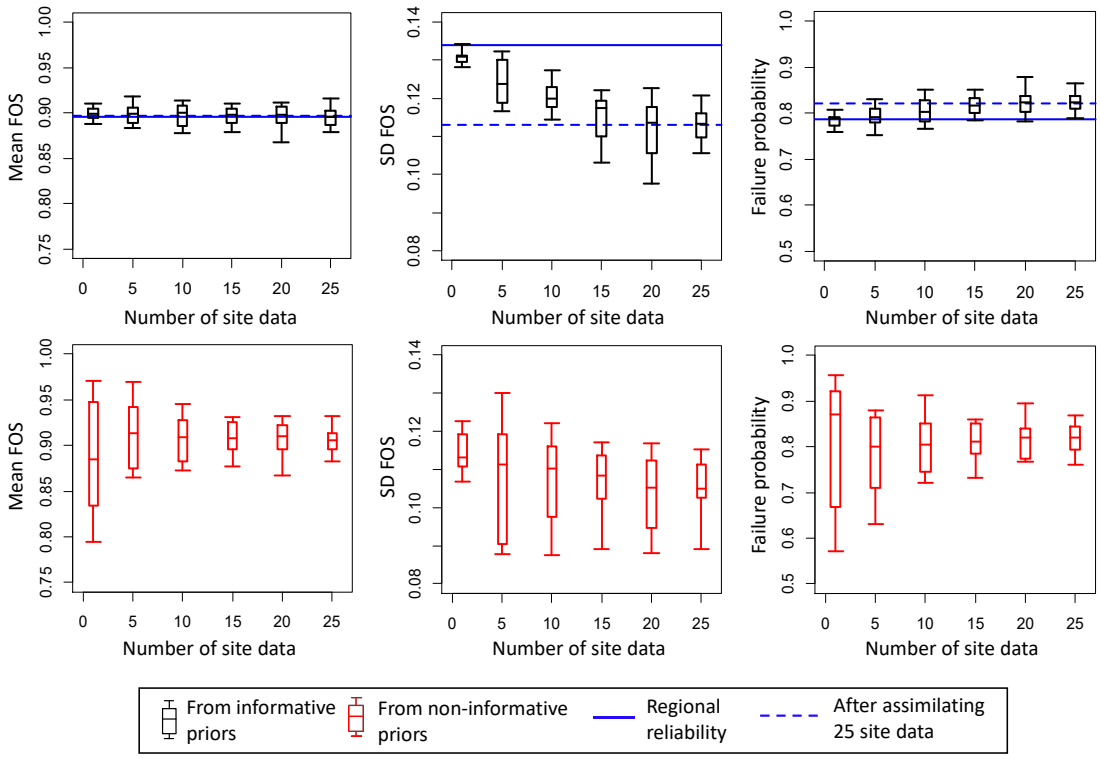

Figure 14: Slope reliability estimates considering various amounts of site data

projects involve less than 10 to 25 sets of site-specific triaxial test data. Solely using the limited site data may reduce the accuracy and value of reliability analyses.

6 Alternative validation using semi-hypothetical scenarios

As in most project sites, the 'ground truth' of probabilistic distributions of soil parameters, and hence the precise 'true' failure probability of the slope, can never be known. However, using the relevant data and the same slope geometry, it is possible to create 'semi-hypothetical' scenarios to validate the approach and demonstrate its benefits. To this end, this section illustrates how the proposed model leads to the transition of reliability estimates towards site-specific conditions with increasing amount of site data, and how the resulting estimates become more robust and accurate than adopting regional data or site data alone. Four semi-hypothetical sites are created herein, assuming the true

Table 6: Statistical parameters of $c' - \phi'$ assigned to the four semi-hypothetical sites and the associated failure probabilities

Site	μ_c (kPa)	σ_c (kPa)	μ_ϕ	σ_ϕ	$\rho_{c-\phi}$	$P_{f,site}$
1	9.8	6.2	36.2°	4.8°	0.311	0.854
2	10.0	6.2	37.0°	4.0°	0.480	0.849
3	12.2	6.2	38.5°	4.5°	0.237	0.717
4	11.9	6.2	39.9°	5.4°	0.397	0.625
Regional	11.1	6.53	37.6°	4.8°	0.286	0.787

site-specific PDFs of shear strength parameters are 'completely known'. The associated statistical parameters and the PDFs are shown in Table 6 and Fig. 15(a)(b). These PDFs are extracted from individual sites (zones) that contribute to the regional database of shear strength information, with their statistical parameters markedly different from the average values. Since these PDFs are assumed to represent the 'ground truths' of soil variability, a 'true' value of failure probability can also be obtained for each of the scenario, denoted as $P_{f,site}$ in Table 6.

'Hypothetical' test data are then generated by fixing the statistical parameters in the probabilistic model in Fig. 8, and with $(s'_1, s'_2, s'_3) = (90, 160, 250)$ kPa. For each case, 10,000 sets of triaxial data are generated, from which the site data are sampled. Moreover, to investigate the influence of different sampling patterns, the sampling process is repeated for 50 times under each data size, which means 50 different possibilities are considered regarding the sampling patterns in the hypothetical 'retrieved samples'. Therefore, the proposed approach results in 50 posterior predictive distributions of the shear strength parameters and 50 estimates of failure probability, denoted as $P_{f,p}$.

Fig. 15(c) displays the trends of the average $P_{f,p}$ for the four cases. Without any site data, estimates of failure probability can only be made using regional PDF of soil parameters; with increasing amounts of site data, $P_{f,p}$ generally transitions towards the corresponding $P_{f,site}$. The amount of site data required for $P_{f,p}$ to approach the true

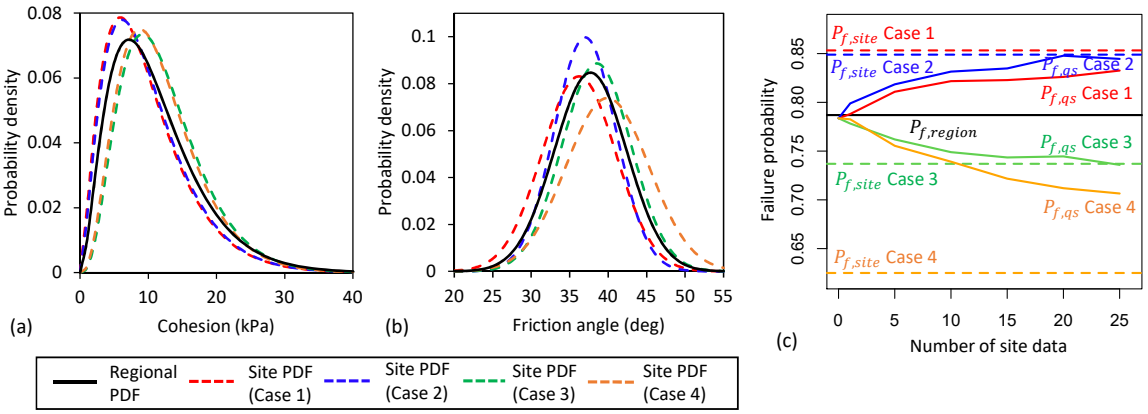

Figure 15: (a)(b) Site-specific PDFs assigned to the semi-hypothetical cases; (c) Average trends of failure probabilities estimated using posterior predictive distributions.

value of $P_{f,site}$ seems to be site-dependent, with Sites 2 and 3 requiring less amount of site data.

Fig. 16 compares the ranges of reliability estimates for the four semi-hypothetical sites, obtained through the PDFs from the informative priors and non-informative priors. The ranges of failure probabilities arise from the 50 different sampling patterns, where the error bars indicate the 10th and 90th percentiles of estimates. Similar to the previous case study, using informative priors produce more robust results, with a steady trend of failure probability observed. When 10 to 15 site-specific data points are assimilated, $P_{f,p}$ systematically deviate from the regional probability, and are mostly within ± 0.10 from the true value of $P_{f,site}$. In contrast, analyses under non-informative priors result in larger statistical uncertainty under limited amount of site data, leading to less robust estimates of failure probability.

In real site settings, although the true statistics of shear strength parameters at a site are never completely known, the presented analyses demonstrated that hybridisation of regional and site data leads to more rational and robust characterisations of slope reliability.

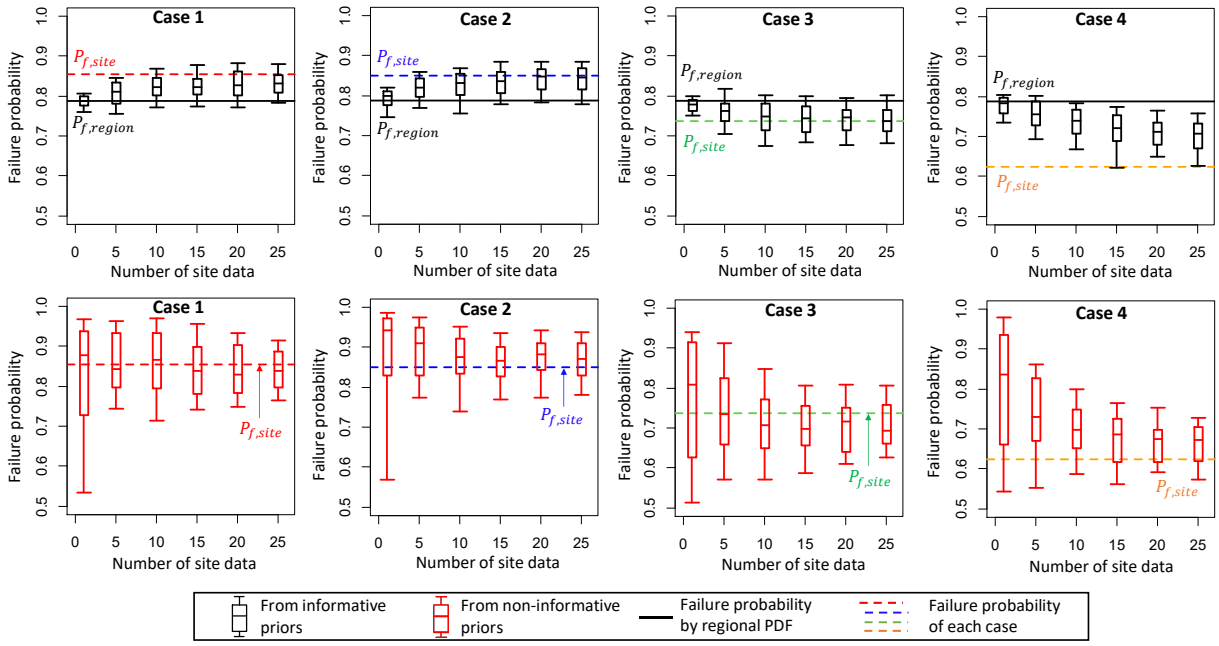

Figure 16: Probabilities of failure estimated with various amounts of site data for four cases

7 Influence of site data on reliability of a hypothetical slope

The Ching Cheung Road slope case involves a relatively high probability of failure, whether or not site-information is assimilated. In this section, a hypothetical scenario is presented to demonstrate how the proposed analysis may affect decision-making through changes to P_f in slope reliability assessments. This hypothetical case involves a CDG soil slope, with a slope height of 20 m and slope angle of about 36°, and the groundwater level is located around one-third of the slope height. Considering the regional distribution of $c' - \phi'$, E(FOS) = 1.644, SD(FOS) = 0.349, and $P_{f,region}$ = 0.0168.

Two sets of hypothetical site characteristics outlined in Table 6, namely Site 2 and Site 3, are considered for this slope geometry. These two sites are chosen because they entail μ_c or μ_ϕ values close to the regional statistics, with Site 2 having slightly lower

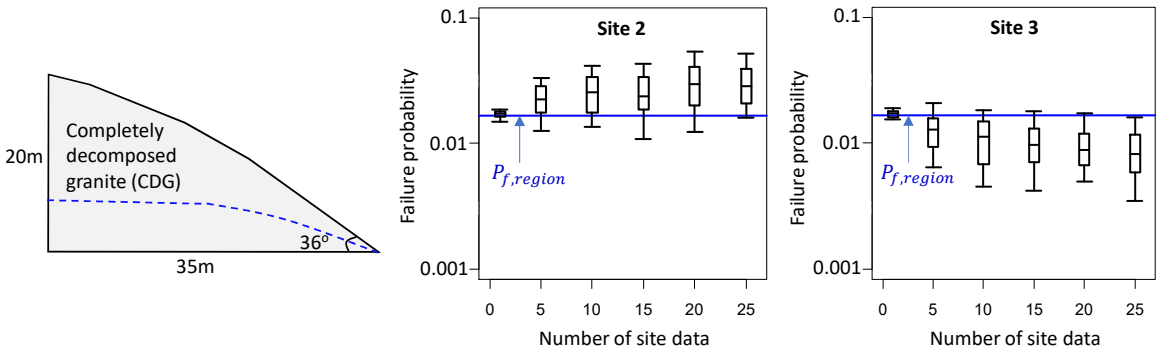

Figure 17: Probabilities of failure for hypothetical slope estimated with various amounts of site data

mean shear strength parameters, and Site 3 slightly higher values. Failure probabilities are estimated again using the corresponding posterior predictive distribution of $c' - \phi'$ parameters, and the resultant trends are shown in Fig. 17. Similar to the previous section, the variability of P_f arises from the 50 different bootstrap sampling patterns of site data. At Sites 3, higher μ_c and μ_ϕ values (compared with regional PDF) lead to an increase in E(FOS), while the reduced statistical uncertainty leads to reduction in SD(FOS). These effects combine and reduce the estimated P_f compared with $P_{f,region}$. In contrast, Site 2 involves lower μ_c and μ_ϕ than regional PDF. Even with reduced statistical uncertainty, the site features outweigh the reduced SD(FOS) and result in higher estimates of P_f.

8 Discussions

The above-mentioned case studies demonstrated the capabilities of the proposed approach and benefits brought by the rational assimilation of regional data with site-specific information, and the resulting posterior predictive distributions of $c' - \phi'$ and reliability estimates lie between their regional and site-specific counterparts. It is worth noting that hybridisation is meaningful only between shear strength data of the same soil type

and testing method, and where the site conditions (e.g., geology) are not categorically different.

This study adopts the zoning scheme in Fig. 2 as the grouping criterion. If the data were grouped according to individual slope projects, the average number of test data in each group would be very low (only 3.2), with more than 82% of the slope projects having fewer than 5 test data. It is difficult to fit meaningful statistical parameters for many of them under this grouping criterion. The current zoning scheme attempts to strike a balance between number of data in a group and potential dissimilarity among individual project locations. Apart from using locations as grouping criteria, other possible criteria include geological origins or particle gradation (Chan et al. 2022), which may form the scope for future research. Another potential modification is related to spatial correlations of $c' - \phi'$ parameters, where soil samples closer to the investigated slope may be given greater importance. For example, Ching and Phoon (2020) proposed a framework to evaluate the multivariate site-specific PDF of shear strength parameters exhibiting stationary spatial autocorrelations. Similar concept may be adopted to extend the current approach.

This study focuses on the assimilation of shear strength information and the corresponding influences to slope reliability. Uncertainty in groundwater table could also be an important source of uncertainty in slope reliability. During rainfall events, the infiltration of rainwater into the slope is uncertain due to uncertainty in soil permeability, or existence of hydrogeological features such as soil pipes, dykes and faults. While this is not considered under the scope of the current work, this factor can be included into reliability assessments provided that the relevant distributions can be reasonably characterised.

9 Conclusions

This paper presents an approach to rationally assimilate regional and site-specific soil shear strength information. As the first step, a hierarchical Bayesian model is fitted to a regional database for completely decomposed granite in Hong Kong, to capture the between-sites variations and within-site variations of shear strength parameters. The regional database is compiled with the results of 747 triaxial tests from 208 slopes. The fitted HBM is then utilized to formulate informative priors, which is applied to the Ching Cheung Road slope case study to produce posterior predictive distributions of $c' - \phi'$ parameters, for reassessments of landslide hazards at the site. As more site data are incorporated, the posterior distributions shift away from the regional PDF, with higher c' and lower ϕ' values than the regional average. The shift in posterior distributions is also reflected by the increase in estimated failure probabilities.

Compared with slope reliability assessed without the regional database, reliability estimates based on assimilation of regional and site-specific data is more robust, as the statistical uncertainty due to small sample size is reduced. This is further demonstrated through the semi-hypothetical scenarios, where the 'true' values of failure probability are approached as increasing numbers of site data are assimilated with information from the regional database. The approach can therefore serve as a useful tool for reliability assessments of many geotechnical projects where site-specific samples are scarce.

10 Acknowledgements

This paper is published with permission of the Head of the Geotechnical Engineering Office and the Director of Civil Engineering and Development, the Government of the Hong Kong Special Administrative Region. The work presented in this paper is partially supported by the Research Grants Council of the Hong Kong Special

609 Administrative Region (Project No. 15222021). Mr. Hansong Pang provided assistance
610 to the compilation of JAGS code.

611 The authors declare there are no competing interest. The triaxial test database
612 reported in this study is available in the supplementary materials. Other data gener-
613 ated or analyzed during this study are available from the corresponding author upon
614 reasonable request.

11 Notation

c'	Effective cohesion
c'_p	Effective cohesion under posterior predictive distribution
c'_{ij}	j-th cohesion value at the i-th site
$c'_{s,j}$	j-th cohesion value at the investigated site
l	Number of data points at the investigated site
l_i	Number of data points at the i-th site
m_{μ_c}, p_{μ_c}	Mean and precision of μ_c
$m_{\mu_\phi}, p_{\mu_\phi}$	Mean and precision of μ_ϕ
$m_{\sigma_\phi}, p_{\sigma_\phi}$	Mean and precision of σ_ϕ
m_ρ, p_ρ	Mean and precision of $\rho_{c-\phi}$
$P_{f,p}$	Failure probability evaluated by posterior distributions of shear strength parameters
$P_{f,region}$	Failure probability evaluated by regional PDF
$P_{f,site}$	Failure probability evaluated by site-specific PDF
s'	Average effective stress
t	Deviator stress
$\boldsymbol{y}$	Vector of observed data
α	Shape parameter of gamma distribution
β	Rate parameter of gamma distribution
ε	Residual error in determination of shear strength parameters
$\boldsymbol{\theta}$	Vector of statistical parameters
μ_c	Mean cohesion
μ_ϕ	Mean friction angle
$\boldsymbol{\xi}$	Vector of $c' - \phi'$ parameters
$\rho_{c-\phi}$	Correlation coefficient between c' and ϕ'
σ_c	Standard deviation of cohesion
σ_ϕ	Standard deviation of friction angle
σ_ε	Standard deviation of ε
ϕ'	Effective friction angle
ϕ'_p	Effective friction angle under posterior predictive distribution
ϕ'_{ij}	j-th friction angle value at the i-th site
$\phi'_{s,j}$	j-th friction angle value at the investigated site
$\boldsymbol{\psi}$	Vector of hyperparameters

12 Appendix I: JAGS code for Hierarchical Bayesian model

```
##########################################
## In JAGS, dnorm(m,p) refers to normal distribution with mean m and precision p
##########################################
model{
###################################
## Hyperparameters
###################################
m_mu_c ~ dunif(0.1,60);
p_mu_c ~ dgamma(10e-4,10e-4);
m_mu_phi ~ dunif(0.1,60);
p_mu_phi ~ dgamma(10e-4,10e-4);
m_sig_phi ~ dunif(0.1,60);
p_sig_phi ~ dgamma(10e-4,10e-4);
m_rho ~ dunif(-0.99,0.99);
p_rho ~ dgamma(10e-4,10e-4);
p_eps ~ dgamma(10e-4,10e-4);
p_c ~ dgamma(10e-4,10e-4);
sig_c <- 1/sqrt(p_c)^2;
###################################
## Statistical parameters of each site
###################################
for (i in 1:N){
   mu_c[i] ~ dnorm(m_mu_c,p_mu_c)T(0,);
   mu_phi[i] ~ dnorm(m_mu_phi,p_mu_phi)T(0,);
   sig_phi[i] ~ dnorm(m_sig_phi,p_sig_phi)T(0,);
   rho_cphi[i] ~ dnorm(m_rho,p_rho)T(-0.99,0.99);
```

```
    alpha_c[i] <- (mu_c[i]^2)/(sig_c^2);
    beta_c[i] <- mu_c[i]/(sig_c^2);
  }
  ###################################
  ## Shear strength data within each site
  ###################################
  pi <- 3.141593;
  for (i in 1:num_sites){
    for (j in 1:l[i]){
      c[i,j] ~ dgamma(alpha_c[i],beta_c[i]);
      phi[i,j] ~ dnorm(mu_phi[i]+(rho_cphi[i]*sig_phi[i]*
        qnorm(pgamma(c[i,j],alpha_c[i],beta_c[i]),0,1);
      t1[i,j] ~ dnorm(c[i,j]*cos(phi[i,j]*pi/180)+s1[i,j]*sin(phi[i,j]*pi/180),p_eps);
      t2[i,j] ~ dnorm(c[i,j]*cos(phi[i,j]*pi/180)+s2[i,j]*sin(phi[i,j]*pi/180),p_eps);
      t3[i,j] ~ dnorm(c[i,j]*cos(phi[i,j]*pi/180)+s3[i,j]*sin(phi[i]*pi/180),p_eps);
    }
  }
  } ##End of Model
```

13 Appendix II: Convergence and model selection for HBM model

After fitting the HBM model, the Gibbs samples should be examined to ensure they have converged to the target distribution. This involves producing multiple Gibbs sampling chains, with each chain having a different starting position. A visual inspection can be performed to check if all sampling chains reach the same part of the parameter space after sufficiently long iterations. A statistical test can also be performed, such

as calculating the scale reduction factor ($\hat{R}$) for each of the variables (Gelman and Rubin 1992). $\hat{R}$ indicates the reduction of between-chains variance compared to the within-chain variance, and convergence is indicated when both variances are similar and $\hat{R}$ is close to 1.0. In this paper, the proposed HBM is fitted with 5 Gibbs sampling chains, with each chain having a burn-in period of 50000, a further chain length of 200000 and thinning interval of 20. Upper confidence limit of $\hat{R}$ is evaluated for each of the variables. The average and maximum $\hat{R}$ are 1.000 and 1.007 respectively, indicating convergence is reached.

Goodness-of-fit of the HBM model is evaluated by the Widely applicable Bayesian information (Watanabe 2013). WAIC is a Bayesian metric for estimating the posterior predictive density of the out-of-sample expectation, while adding a correction for effective number of parameters to adjust for overfitting. WAIC should be carefully defined according to the "model focus" (Merkle et al. 2019), which is related to the aim of predictions. If the aim is to generalize the predictions beyond the specific groups that were observed in the database, which agrees with this study, then the "model focus" should be on the hyperparameters ψ. The formula for WAIC is:

$$\begin{aligned} WAIC &= \bar{D} + p_{WAIC} \\ \bar{D} &= -2E_{\psi|y}\left[\ln p\left(y|\psi\right)\right] \\ p_{WAIC} &= 2\ln E_{\psi|y}\left[p\left(y|\psi\right)\right] + \bar{D} \end{aligned} \tag{8}$$

where $\bar{D}$ is the expected deviance, which is -2 times the expected log predictive density of the data. p_{WAIC} is a measure of model complexity, which is 2 times the log of expected predictive density plus the expected deviance. Compared to another widely

used Bayesian metric, deviance information criterion (DIC) (Spiegelhalter et al. 2002), DIC also comprises of an expected deviance term $\bar{D}$ and a model complexity term p_{DIC}. The difference is that DIC evaluates the complexity term through a point estimate $\ln[p(y|\widetilde{\psi})]$, where $\widetilde{\psi}$ is the posterior mean of hyperparameters; but WAIC evaluates the complexity term by averaging over the entire posterior distribution of hyperparameters. According to Gelman et al. (2014), WAIC works better for probabilistic models with hierarchical structures, which involves a large number of parameters, and point estimates often do not make sense. Watanabe (2013) also shown that WAIC is better related to Bayesian leave-one out cross-validation, and they are asymptotically equivalent. Based on the WAIC definition, the following implementation is proposed by Gelman et al. (2014):

$$\begin{aligned} E_{\psi|y}\left[\ln p\left(y|\psi\right)\right] &= \sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T}\ln p(y_i|\psi^t)\right) \\ \ln E_{\psi|y}\left[p\left(y|\psi\right)\right] &= \sum_{i=1}^{n}\ln\left(\frac{1}{T}\sum_{t=1}^{T}p(y_i|\psi^t)\right) \end{aligned} \tag{9}$$

where y_i are the individual data points, and ψ^t are the Gibbs samples of the hyperparameters. To evaluate $p(y_i|\psi^t)$, the latent variables ξ and θ (i.e. $c' - \phi'$ values and statistical parameters) are simulated from $p(\xi, \theta|\psi^t)$, which are denoted as ξ^s and θ^s, and then $p(y_i|\psi^t) \approx p(y_i|\xi^s)$. In other words, the Gibbs samples of ξ and θ should not be directly applied to evaluate $p(y_i|\psi^t)$. An R code is written by the authors to implement WAIC for the HBM models in this study.

14 Appendix III: JAGS code for assimilating site data through informative prior

```
model{
####################################
## Informative priors
####################################
E_mu_c <- 11.1;
SD_mu_c <- 1.94;
E_sig_c <- 6.26;
SD_sig_c <- 0.29;
E_mu_phi <- 37.7;
SD_mu_phi <- 1.41;
E_sig_phi <- 4.41;
SD_sig_phi <- 0.86;
E_rho_cphi <- 0.327;
SD_rho_cphi <- 0.163;
E_sig_eps <- 1.96;
SD_sig_eps <- 0.048;
####################################
## Statistical parameters of the new site
####################################
mu_c ~ dnorm(E_mu_c,1/SD_mu_c^2)T(0,);
sig_c ~ dnorm(E_sig_c,1/SD_sig_c^2)T(0,);
mu_phi ~ dnorm(E_mu_phi,1/SD_mu_phi^2)T(0,);
sig_phi ~ dnorm(E_sig_phi,1/SD_sig_phi^2)T(0,);
```

```
rho_cphi ~ dnorm(E_rho_cphi,1/SD_rho_cphi^2)T(-0.99,0.99);
alpha_c <- (mu_c^2)/(sig_c^2);
beta_c <- mu_c/(sig_c^2);
sig_eps ~ dnorm(E_sig_eps,1/SD_sig_eps^2);
#####################################
## Shear strength data of the new site
#####################################
pi <- 3.141593;
for (i in 1:data_size){
   c[i] ~ dgamma(alpha_c,beta_c);
   phi[i] ~ dnorm(mu_phi+rho_cphi*sig_phi*qnorm(pgamma(c[i],alpha_c,beta_c),0,1), 1/((1-
rho_cphi^2)*sd_phi^2));
   t1[i] ~ dnorm(c[i]*cos(phi[i]*pi/180)+s1[i]*sin(phi[i]*pi/180),1/sig_eps^2);
   t2[i] ~ dnorm(c[i]*cos(phi[i]*pi/180)+s2[i]*sin(phi[i]*pi/180),1/sig_eps^2);
   t3[i] ~ dnorm(c[i]*cos(phi[i]*pi/180)+s3[i]*sin(phi[i]*pi/180),1/sig_eps^2);
}
#####################################
## c-phi distribution of the new site
#####################################
c_p ~ dgamma(alpha_c,beta_c);
phi_p ~ dnorm(mu_phi+rho_cphi*sig_phi*qnorm(pgamma(c_p[i],alpha_c,beta_c),0,1), 1/((1-
rho_cphi^2)*sig_phi^2));
} ##End of Model
```

References

Au, S. W. C. 1993. "Reversal shear box test for hong kong saprolitic soils." *Quarterly Journal of Engineering Geology and Hydrogeology*, 26(3), 233–237.

Bozorgzadeh, N., Harrison, J. P., and Escobar, M. D. 2019. "Hierarchical Bayesian modelling of geotechnical data: application to rock strength." *Géotechnique*, 69(12), 1056–1070.

Bozorgzadeh, N., Liu, Z., Nadim, F., and Lacasse, S. 2023. "Model calibration: A hierarchical Bayesian approach." *Probabilistic Engineering Mechanics*, 71, 103379.

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

Chan, C., Wong, L., Leung, W., Chung, P., and Lo, M.K. Leung, Y. 2022. "Development of regional soil shear strength database and its application in probabilistic analysis of slope stability." *Proceedings of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR), Newcastle, Australia*.

Ching, J., Lin, G. H., Chen, J. R., and Phoon, K. K. 2017. "Transformation models for effective friction angle and relative density calibrated based on generic database of coarse-grained soils." *Canadian Geotechnical Journal*, 54(4), 481–501.

Ching, J. and Phoon, K. K. 2013. "Multivariate distribution for undrained shear strengths under various test procedures." *Canadian Geotechnical Journal*, 50(9), 907–923.

Ching, J. and Phoon, K. K. 2019. "Constructing site-specific multivariate probability distribution model using Bayesian machine learning." *Journal of Engineering Mechanics*, 145(1), 04018126.

Ching, J. and Phoon, K. K. 2020. "Constructing a site-specific multivariate probability

distribution using sparse, incomplete, and spatially variable (MUSIC-X) data." *Journal of Engineering Mechanics*, 146(7), 04020061.

Ching, J., Phoon, K. K., Ho, Y. H., and Weng, M. C. 2021a. "Quasi-site-specific prediction for deformation modulus of rock mass." *Canadian Geotechnical Journal*, 58(7), 936–951.

Ching, J., Wu, S., and Phoon, K. K. 2021b. "Constructing quasi-site-specific multivariate probability distribution using hierarchical Bayesian model." *Journal of Engineering Mechanics*, 147(10), 04021069.

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

Duncan, J. M. 2000. "Factors of safety and reliability in geotechnical engineering." *J. Geotech. Geoenviron. Eng.*, 126(4), 307–316.

El-Ramly, H., Morgenstern, N. R., and Cruden, D. M. 2005. "Probabilistic assessment of stability of a cut slope in residual soil." *Géotechnique*, 55(1), 77–84.

Feng, Y., Gao, K., Mignan, A., and Li, J. 2021. "Improving local mean stress estimation using bayesian hierarchical modelling." *International Journal of Rock Mechanics and Mining Sciences*, 148, 104924.

Fenton, G. A. and Griffiths, D. V. 2003. "Bearing-capacity prediction of spatially random c-ϕ soils." *Canadian geotechnical journal*, 40(1), 54–65.

Gelman, A. 2006. "Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)." *Bayesian Anal.*, 1(3), 515–534.

Gelman, A. and Hill, J. 2006. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Gelman, A., Hwang, J., and Vehtari, A. 2014. "Understanding predictive information criteria for Bayesian models." *Statistics and computing*, 24, 997–1016.

Gelman, A. and Rubin, D. B. 1992. "Inference from iterative simulation using multiple sequences." *Statistical science*, 7(4), 457–472.

Geman, S. and Geman, D. 1984. "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." *IEEE Transactions on pattern analysis and machine intelligence*, 6, 721–741.

GEO 2017. *Model Specification for Soil Testing (Geospec 3)*. Geotechnical Engineering Office, Civil Engineering and Development Department, HKSAR Government.

GEO-SLOPE 2017. *Stability Modeling with Geostudio*. Canada.

Griffiths, D. V. and Fenton, G. A. 2004. "Probabilistic slope stability analysis by finite elements." *Journal of geotechnical and geoenvironmental engineering*, 130(5), 507–518.

Halcrow 1998. *Report on the Ching Cheung Road Landslide of 3 August 1997, GEO Report No. 78*. Hong Kong.

Hasofer, A. M. and Lind, N. C. 1974. "Exact and invariant second-moment code format." *Journal of the Engineering Mechanics division*, 100(1), 111–121.

Hencher, S. R. 2010. "Preferential flow paths through soil and rock and their association with landslides." *Hydrological processes*, 24(12), 1610–1630.

Huang, A. and Wand, M. 2013. "Simple marginally noninformative prior distributions for covariance matrices." *Bayesian Anal.*, 8(2), 439–452.

Jiang, S., Li, D., Cao, Z., Zhou, C., and Phoon, K. 2014. "Efficient system reliability analysis of slope stability in spatially variable soils using monte carlo simulation." *Journal of Geotechnical and Geoenvironmental Engineering*, 141(2), 04014096.

Li, D. Q., Tang, X. S., and Phoon, K. K. 2015. "Bootstrap method for characterizing the effect of uncertainty in shear strength parameters on slope reliability." *Reliability Engineering & System Safety*, 140, 99–106.

Liu, X. F., Tang, X. S., Li, D. Q., and Qi, X. H. 2020. "Jackknifing for modeling

sampling properties of soil statistics for geotechnical reliability analysis." *Computers and Geotechnics*, 125, 103685.

Lo, M. K. and Leung, Y. F. 2017. "Probabilistic analyses of slopes and footings with spatially variable soils considering cross-correlation and conditioned random fields." *J. Geotech. Geoenviron. Eng.*, 143(9), 04017044.

Lo, M. K., Wei, X., Chian, S. C., and Ku, T. 2021. "Bayesian network prediction of stiffness and shear strength of sand." *Journal of Geotechnical and Geoenvironmental Engineering*, 147(5), 04021020.

Low, B. K. and Tang, W. H. 1997. "Probabilistic slope analysis using janbu's generalized procedure of slices." *Computers and Geotechnics*, 21(2), 121–142.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. 2013. *The BUGS book: A practical introduction to Bayesian analysis*. Chapman Hall, London.

Merkle, E. C., Furr, D., and Rabe-Hesketh, S. 2019. "Bayesian comparison of latent variable models: Conditional versus marginal likelihoods." *Psychometrika*, 84, 802–829.

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

Neal, R. M. 2003. "Slice sampling." *The annals of statistics*, 31(3), 705–767.

Nelsen, R. 1999. *An Introduction to Copulas*. Springer, New York.

Parker, C., Simon, A., and Thorne, C. R. 2008. "The effects of variability in bank material properties on riverbank stability: Goodwin Creek, Mississippi." *Geomorphology*, 101(4), 533–543.

Phoon, K. and Kulhawy, F. H. 1999. "Characterization of geotechnical variability." *Canadian geotechnical journal*, 36(4), 612–624.

Phoon, K.-K. 2017. "Role of reliability calculations in geotechnical design." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 11(1), 4–21.

Plummer, M. 2003. "Jags: A program for analysis of bayesian graphical models using gibbs sampling." *Proceedings of the 3rd international workshop on distributed statistical computing*, 124(125.10), 1–10.

Plummer, M. 2015. "Jags version 4.0.0 user manual.

Plummer, M., Stukalov, A., Denwood, M., and Plummer, M. M. 2019. "Package 'rjags'.

Rahardjo, H., Satyanaga, A., Leong, E. C., Ng, Y. S., and Pang, H. T. C. 2012. "Variability of residual soil properties." *Engineering Geology*, 141, 124–140.

Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. 2002. "Bayesian measures of model complexity and fit (with discussion)." *Journal of the Royal Statistical Society Series B*, 64, 583–639.

Ueda, K. 2022. "Hierarchical bayesian approach to estimating variability of liquefaction resistance of sandy soils considering individual differences in laboratory tests." *Journal of Geotechnical and Geoenvironmental Engineering*, 148(2), 04021188.

Wang, Y. and Akeju, O. V. 2016. "Quantifying the cross-correlation between effective cohesion and friction angle of soil from limited site-specific data." *Soils and Foundations*, 56(6), 1055–1070.

Wang, Y., Akeju, O. V., and Cao, Z. 2016. "Bayesian equivalent sample toolkit (best): an Excel VBA program for probabilistic characterisation of geotechnical properties from limited observation data." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 10(4), 251–268.

Wang, Y., Cao, Z., and Au, S. K. 2011. "Practical reliability analysis of slope stability by advanced Monte Carlo simulations in a spreadsheet." *Canadian Geotechnical Journal*, 48(1), 162–172.

Wang, Y., Zhao, T., and Cao, Z. 2015. "Site-specific probability distribution of geotechnical properties." *Computers and Geotechnics*, 70, 159–168.

Watanabe, S. 2013. "A widely applicable Bayesian information criterion." *Journal of Machine Learning Research*, 14, 867–897.

Xiao, S., Zhang, J., Ye, J., and Zheng, J. 2021. "Establishing region-specific N–Vs relationships through hierarchical Bayesian modeling." *Engineering Geology*, 287, 106105.

Zhang, J., Hu, J., Li, X., and Li, J. 2020. "Bayesian network based machine learning for design of pile foundations." *Automation in Construction*, 118, 103295.

Zhang, J., Juang, C. H., Martin, J. R., and Huang, H. W. 2016. "Inter-region variability of Robertson and Wride method for liquefaction hazard analysis." *Engineering Geology*, 203, 191–203.