A Bayesian framework for probabilistic site investigation

Jinsong Huang
The University of Newcastle, Newcastle, Australia, jinsong.huang@newcastle.edu.au

Richard Kelly
SMEC, Australia & New Zealand Division, Newcastle, Australia, drrichard.kelly@smec.com

Jiawei Xie
The University of Newcastle, Newcastle, Australia, jiawei.xie@uon.edu.au

ABSTRACT: This paper proposes a probabilistic site characterization approach, based on Bayesian statistical methods, that allows a design profile involving uncertainty to be determined automatically. Based on random field theory and Bayes' theorem, the proposed method can integrate both geophysical and geotechnical datasets in a rigorous manner to derive a design profile automatically in two or more dimensions. The proposed approach is applied to a soft soil test site in Ballina, New South Wales, Australia, and compared with traditional approaches. The results show that by gradually incorporating more data into the Bayesian inversion, the uncertainty in the soil profile is greatly reduced.

Keywords: integrated site investigation; Bayesian statistics; spatial variability

1. Introduction

Accuracy and reliability of geotechnical field test are of great significance to the construction process. The electric cone penetration test (CPT) is widely used in the geotechnical site investigation because it is fast, repeatable and economical. Advances in geotechnical in situ testing like the seismic piezocone (SCPTU) can capture five independent readings in a single sounding. These soundings provide near-continuous data with depth on the stratigraphy, fundamental stiffness, stress state, strength, and flow characteristics of sub-soils, which can be used to determine soil stratigraphy, identify soil type and characterize soil properties. One approach is to use the chart developed by [1]. In this deterministic approach, the corrected cone tip resistance (q_t) and sleeve friction (f_s) are used to define a soil behavior type (SBT) index I_c , which is used to identify the soil. Some statistical methods have also been reported in the literature, e.g., Phoon et al. [2] proposed a modified-Bartlett method to statistically identify homogenous soil layers. Bayesian statistical methods were also used to identify soil stratigraphy by [3–5]. Apart from the Bayesian approach, Ching et al. [6] recently applied the Wavelet Transform Modulus Maxima method (WTMM) to identify soil boundaries while small-scale spurious layer boundaries were avoided. It is noted, however, that all these methods provide a single 1D stratigraphy at a particular location. In practice, these individual 1D stratigraphies are usually jointly interpreted with diverse geophysical surveys to derive a single model of the probed subsurface region. This approach is, however, largely qualitative in nature and hence the outcome depends heavily on the background, experience and preconceptions of the interpreter. To obtain 2D soil stratigraphy based on 1D piezocone tests (CPTUs) and avoid subjective decisions, Li et al. [7] used a Kriging interpolation technique to estimate the corrected cone tip resistance and sleeve friction at unsampled locations. The interpolated corrected cone tip resistance and sleeve friction were then used to identify 2D soil stratigraphy based on Roberson's chart. However, only one type of geotechnical test (e.g., CPTU) was employed. Geophysics provides a wide variety of tools which can help to identify subsoil stratigraphy for subsequent detailed assessment. Data is obtained on a two- (or three-) dimensional (2D/3D) section of the ground. For example, in comparison to a conventional drilling approach, the Multichannel Analysis of Surface Wave (MASW) test [8,9] is conducted from the ground surface (non-invasive), covers the subsurface continuously and provides more complete coverage. Integration of different methods and different datasets can provide a more reliable site characterization, reducing the uncertainties associated with a single measurement. Usually this integration is done manually and largely based on engineering judgement and experience, and results in additional uncertainties.

In conclusion, it is necessary to develop new tools that can integrate both geophysical and geotechnical datasets in a rigorous manner to derive a design profile automatically in two or more dimensions. Based on random field theory and Bayes' theorem, this paper proposes a rigorous statistical framework to combine geotechnical and geophysical test data. The 1D q_t measurement from CPTU tests and 2D shear wave velocity (V_s) data from Multichannel Analysis of Surface Wave (MASW) tests are used as an example to derive a 2D q_t profile. The real engineering data from the Ballina site, NSW, Australia [10] are used to show the potential of the proposed approach. Traditional approaches are also applied to show the differences in the final soil profiles. Only 2D problems are considered, although the proposed procedure can be extended easily to 3D.

2. Methodology

2.1. Probabilistic site characterization based on random field theory

Due to the weathering processes, pore fluids, and physical and chemical changes, the physical properties vary from place to place within a deposit. When the spatial variation in a soil property is assumed to be controlled by a random process [11], it can be modelled as the sum of a trend component and a residual component [2,12–16]. For example, the corrected cone tip resistance can be expressed as

$$q_{\mathrm{t}}(x,z) = \mu_{q_{\mathrm{t}}}(x,z) + \varepsilon(x,z) \tag{1}$$

where x and z are the horizontal and vertical coordinates respectively, $\mu_{q_{\mathrm{t}}}(x,z)$ is a known trend, and $\varepsilon(x,z)$ is the residual variability about that trend. Usually, the strength of soil increases with depth [16], and it is common that the trend function is assumed to be a linear or non-linear polynomial according to site-specific conditions.

Residuals are statistically modeled as random fields, which are characterized by mean, standard deviation and fluctuation range.[12]. Because CPTU can provide almost continuous readings with depth [1] , it is usually used to estimate the scale of fluctuation in the vertical direction [14,16–18] . The spatial variation in the horizontal direction is usually less than that in the vertical direction [15]. In cases where there is not enough data, the spatial correlation structure in the horizontal direction can be inferred from similar sites.

Although random fields based on the three parameters (e.g., a mean, a standard deviation and a scale of fluctuation) can represent the heterogeneity statistically, a more realistic site characterization should be based on the site-specific conditions, i.e., the random fields should be conditioned on site investigation data [7,19,20]. In addition, the above site characterization only uses the information from CPTU. As mentioned previously, if multiple source of information can be used together, the uncertainties in the soil profile can be reduced.

2.2. Soil profile based on the Bayesian statistical methods

In this paper, Bayesian statistical methods are adopted to integrate different sources of information. To illustrate the proposed approach, the Multichannel Analysis of Surface Wave (MASW) and CPTU test results are used to obtain a 2D profile of corrected cone tip resistance. Suppose a 2D MASW test has been conducted and the corresponding shear wave velocity results are denoted as $\mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}}$. The 2D shear wave velocity profile $\mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}}$ is calculated over depth increments of say 0.5m to 1m and additional CPTUs are performed to increase the resolution of measurements along a line within the domain covered by the shear wave velocity data. The corresponding 1D q_{t} measurements are denoted as $\mathbf{q}^{\mathrm{obs}}_{\mathrm{t,1D}}$.

As shown in Fig. 1, in the proposed approach, the model parameters of q_{t} random fields, $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$, derived based on the method presented in section 2.1 are used as prior knowledge on the subsoil profile. The $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ uniquely define/control the 2D random filed $\mathbf{q}_{\mathrm{t,2D}}$. Both the $\mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}}$ and $\mathbf{q}^{\mathrm{obs}}_{\mathrm{t,1D}}$ (denoted by red block diagrams) are used to update the prior soil profiles. Once the $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ is updated, the corresponding random field and soil profile can be updated. Based on Bayes' theorem, the posterior probability distribution function (PDF) of parameters $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ for random field of q_{t} can be expressed as [21,22]:

$$P_{\mathrm{pos}}(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}} \,|\, \mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}}) = \alpha L(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}} \,|\, \mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}}) P_{\mathrm{pri}}(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}) \tag{2}$$

where $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ are the model parameters of random field $\mathbf{q}_{\mathrm{t,2D}}$; $P_{\mathrm{pri}}(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}})$ is the prior PDF of $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ according to prior information; $P_{\mathrm{pos}}(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}} \,|\, \mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}})$ is the posterior probability of $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ given the observations of $\mathbf{q}^{\mathrm{obs}}_{\mathrm{t,1D}}$ and $\mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}}$, α is a normalization constant, and $L(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}} \,|\, \mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}})$ is the likelihood of observing the measured data $\mathbf{q}^{\mathrm{obs}}_{\mathrm{t,1D}}$ and $\mathbf{V}^{\mathrm{obs}}_{\mathrm{s,2D}}$ given the $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$.

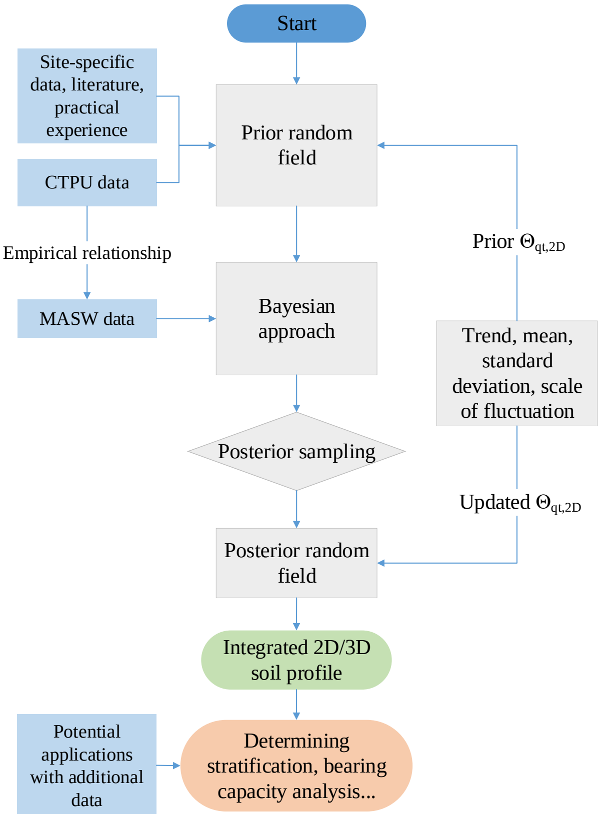

Figure 1. Flowchart of the proposed approach.

2.2.1. Dimensionality reduction and prior distribution

The resolution of $\mathbf{q}_{\mathrm{t,2D}}$ is usually 1–5 cm [23]. If each q_{t} in space is associated with a random variable, the dimension of the problem will be unavoidably high. This implies that the computational time will be significantly increased. For illustration, in this study, the Karhunen-Loève (KL) expansion method [24,25] is used to reduce the dimensionality according to

$$\mathbf{q}_{\mathrm{t,2D}} = \mu_{q_{\mathrm{t}}}(\mathbf{x},\mathbf{z}) + \sum_{i=1}^{M} \sigma_{\mathrm{de},q}\sqrt{\lambda_i}\, f_i(\mathbf{x},\mathbf{z})\xi_i \tag{3}$$

where $\{\xi_1, \xi_2, \ldots, \xi_M\}$ are the uncorrelated standard normal random variables; M is the number of random variables; $f_i(x,z)$ and λ_i are the eigenfunctions and eigenvalues of the auto-correlation function, which can be obtained by solving the homogeneous Fredholm integral equation of the second kind [25]; The horizontal and vertical scales of fluctuation are δ_x and δ_z, respectively, which control the spatial correlation of q_{t} at two random points; and $\sigma_{\mathrm{de},q}$ is the standard deviation of the de-trended q_{t}. Since Eq. (3) has to be truncated to a finite number of terms M, a significant concern is that the simulated variance will be reduced. In order to control this reduction, the eigenvalues are sorted in descending order and the number of terms is determined to ensure that the eigenvalues have decayed below a threshold [26].

The prior knowledge of $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ is critical in Bayesian updating and can be obtained from site data, practical experience, and literature [27]. Its PDF $P_{\mathrm{pri}}(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}})$ describes the prior knowledge of random field before Bayesian updating using data $\mathbf{q}_{\mathrm{t,1D}}^{\mathrm{obs}}$ and $\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}$. The $P_{\mathrm{pri}}(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}})$ is usually formulated in the form of joint distribution of model parameters [5,20,28–30] . The form is dependent on the technique (e.g., KL expansion) adopted for discretizing the random field and subsequent model parameterization approach [31]. For example, here the 2D $\mathbf{q}_{\mathrm{t,2D}}$ is modelled by KL expansion, thus the model parameters $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$ contain the $\mu_{q_{\mathrm{t}}}$, $\sigma_{\mathrm{de},q}$, δ_x, δ_z and $\xi_1,\xi_2,\ldots,\xi_M$ in Eq. (3). The $P_{\mathrm{pri}}(\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}})$ is the joint distribution of $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}=\{\mu_{q_{\mathrm{t}}}, \sigma_{\mathrm{de},q}, \delta_x, \delta_z, \xi_1,\xi_2,\ldots,\xi_M\}$. Typically, the $\mu_{q_{\mathrm{t}}}$, $\sigma_{\mathrm{de},q}$, δ_x, δ_z can either be deemed as normally distributed [28,29,32] or uniformly distributed [4,30,33] and $\xi_1,\xi_2,\ldots,\xi_M$ are standard normal variables [31]. It should be noted that the technique for discretizing the random field is case-dependent, and thus the corresponding model parameterization is also case-dependent. If other techniques, e.g., matrix decomposition [5], are adopted to discretize the random field, then model parameterization and prior PDF may be different.

2.2.2. Likelihoods and posterior random field sampling

The $L(\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}\mid\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}})$ is another likelihood of observing $\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}$ given $\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}$. It is noted that $\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}$ and $\mathbf{q}_{\mathrm{t,2D}}$ have different resolutions. The CPTU readings are usually taken every 1–5 cm, while the resolution of the seismic investigation values $\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}$ could be half to several meters [8]. Thus, the high resolution $\mathbf{q}_{\mathrm{t,2D}}$ is up-scaled to match the resolution of $\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}$ by averaging. The averaged q_{t} field is $\bar{\mathbf{q}}_{\mathrm{t,2D}}$.

Note that V_{s} and q_{t} are two different measurements of the engineering properties, and there is no analytical relationship between them. Some empirical relationships between the measured V_{s} and q_{t} are available in the literature [34,35]. These empirical regressions and transformation models, which can be evaluated by some pre-screening model selection procedure [36], are deemed appropriate at a specific site [34]. The model is derived from 481 data pairs worldwide and expressed as:

$$\ln V_{\mathrm{s}} = a\times\ln\bar{q}_{\mathrm{t}} + b + \varepsilon_{\mathrm{T}} \tag{4}$$

where a = 0.627, b = ln1.75; $\varepsilon_{\mathrm{T}} = \ln V_{\mathrm{s}} - a\times\ln\bar{q}_{\mathrm{t}} - b$ is the total transformation uncertainty [37] characterized in the regression model. The standard deviation (SD) of ε_{T} derived from the log regression is σ_{T} 0.146 [34].

The total transformation uncertainty may be spatially correlated, which has been proved in the literature [6,37]. These transformation uncertainties at different points in space shall be treated as correlated random variables if the characteristic dimension of a geotechnical structure is larger than the estimated vertical SOF of the transformation uncertainty [37]. Recall that ε_{T} in Eq. (4) is the total transformation uncertainty. According to [37], it can be decomposed as

$$\varepsilon_{\mathrm{T}} = w_s + w_r + e' \tag{5}$$

in which $w_r + e'$ is serving as zero-mean uncorrelated white noise with SD of σ_r, w_s is slowly fluctuating component which can be modelled as zero-mean stationary correlated random field. Thus the likelihood of observing $\boldsymbol{\varepsilon}_{\mathrm{T}} = [\varepsilon_{\mathrm{T},1},\varepsilon_{\mathrm{T},2},\ldots,\varepsilon_{\mathrm{T},N_v}]$ at a 2D domain can be given as:

$$\begin{aligned} P_2(\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}\mid\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}) &= P(\boldsymbol{\varepsilon}_{\mathrm{T}}\mid\boldsymbol{\Theta}_{\mathbf{q}_{\mathrm{t,2D}}}) \\ &= \frac{1}{(2\pi)^{N_V/2}\det(\boldsymbol{\Sigma}_{\mathrm{T}})^{1/2}}\times\exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}_{\mathrm{T}}{}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathrm{T}}{}^{-1}\boldsymbol{\varepsilon}_{\mathrm{T}}\right\} \end{aligned} \tag{6}$$

where $\boldsymbol{\Sigma}_{\mathrm{T}}$ is the covariance matrix for $\boldsymbol{\varepsilon}_{\mathrm{T}}$:

$$\boldsymbol{\Sigma}_{\mathrm{T}} = \begin{bmatrix} \sigma_s^2 & \sigma_s^2\rho(\tau_{x_{12}},\tau_{z_{12}}) & \cdots & \sigma_s^2\rho(\tau_{x_{1N_v}},\tau_{z_{1N_v}}) \\ & \sigma_s^2 & \cdots & \sigma_s^2\rho(\tau_{x_{2N_v}},\tau_{z_{2N_v}}) \\ & & \ddots & \vdots \\ \mathrm{SYM.} & & & \sigma_s^2 \end{bmatrix} + \begin{bmatrix} \sigma_r^2 & 0 & \cdots & 0 \\ & \sigma_r^2 & \cdots & 0 \\ & & \ddots & \vdots \\ \mathrm{SYM.} & & & \sigma_r^2 \end{bmatrix} \tag{7}$$

N_V is the number of V_{s} in $\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}$; $\rho(\tau_{x_{ij}},\tau_{z_{ij}})$ is the autocorrelation coefficient between spatial quantities (i.e., transformation uncertainties) at any two points, in which the lags $\tau_{x_{ij}} = |x_i - x_j|$ and $\tau_{z_{ij}} = |z_i - z_j|$ are the absolute distances between coordinates of the ith point and the jth point in domain (i, j = 1,2, …, N_v).

In summary, the proposed approach combines the $\mathbf{q}_{\mathrm{t,1D}}^{\mathrm{obs}}$ and $\mathbf{V}_{\mathrm{s,2D}}^{\mathrm{obs}}$ based on rigorous Bayesian statistical principles to derive a 2D q_{t} profile. Different sources of uncertainty: heterogeneity of the real data, transformation uncertainties for $q_{\mathrm{t}} - V_{\mathrm{s}}$ empirical regression model, measurement errors and scale changes are consistently taken into account.

Because random field theory is adopted in Eq. (3) and a large number of components in $\boldsymbol{\Theta}_{\mathbf{q}_{t,2D}}$ are involved into Bayesian updating, it is impossible to derive an analytical expression for the posterior random fields. In this study, the Differential Evolution Adaptive Metropolis Method (DREAM$_{\text{(ZS)}}$) [38] is used to obtain the posterior samples of $\boldsymbol{\Theta}_{\mathbf{q}_{t,2D}}$ according to [39,40]. Once the posterior samples of $\boldsymbol{\Theta}_{\mathbf{q}_{t,2D}}$ are obtained, the updated soil profiles in the form of updated random fields will be determined.

3. Case study

The Australian Research Council Centre of Excellence for Geotechnical Science and Engineering (CGSE). is operating the National Field Test Facility (NFTF) for soft soils in Ballina, New South Wales (NSW), Australia. In June 2013, prior to embankment construction (embankments A and B in Fig. 2), initial geophysics surveys were carried out at the site. The MASW tests were undertaken along two lines, an East-West alignment and a North-South alignment, which are shown in Fig. 2. The East-West geophysical line (MASW line-1) started at a chainage of approximately 80 m along the North-South line (denoted by the start and end points). CPT-6A is not along this alignment but is closely located south of this MASW line-1, as shown by the small red triangle in Fig. 2.

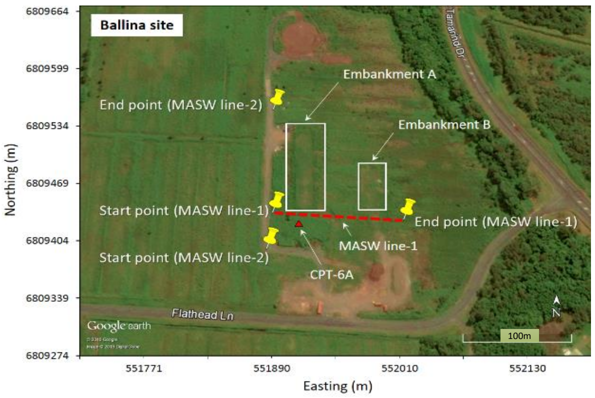

Figure 2. Location of the Ballina site, and the location of CPT-6A and MASW-1

3.1. Prior soil profile

The MASW-1 field is specified on a 75.2 m long by 19.2 m deep domain which is discretized using 0.376 m × 0.196 m cells to yield a grid size of 200 × 100. The 1D q_t data of CPT-6A and 2D V_s data of MASW-1 are integrated together using the proposed approach to derive the 2D q_t profile.

Although site investigation is carried at the Ballina site [10,41], the sampling and drilling location is remote to the MASW line-1 (orthogonal distance is larger than 20 m). These site data may not provide enough indicative information for determining the 2D stratigraphy along this section due to spatial variability [42]. Thus, the stratification is not considered in this example.

The non-stationary random fields can be turned into stationary random fields by removing the deterministic mean trend. The order of the trend is usually not higher than quadratic [43–45]. Because the CPT-6A q_t profile shows obvious nonlinearity with depth, a second-order mean trend is assumed. Based on Eq. (8), the mean trend is calculated as

$$\mu_{q_t}(x,z) = 0.525 - 0.198z + 0.021z^2 \tag{8}$$

After the deterministic trend has been removed, the standard deviation of q_t is found to be $\sigma_{\text{de},q}$ = 0.54 MPa and the vertical scale of fluctuation δ_z is estimated to be 0.92 m. According to [46], the horizontal scale of fluctuation δ_x is set to be 9.5 m and the squared exponential auto-correlation function is assumed. The uncertainties of the trend μ_{q_t}, auto-correlation function model, and parameters $\sigma_{\text{de},q}$, δ_x, δ_z are not considered for simplification. The vector $\xi_1, \xi_2, \ldots, \xi_M$ contains M = 250 components, and is formulated as an uncorrelated standard multi-normal distribution.

3.2. Posterior soil profile by incorporating CPT-6A and MASW-1

Compared to field tests such as the CPT, geophysical tests such as MASW can provide more economical, quick, but indirect information on the conditions of subsoils in two dimensions. The data from CPT-6A and MASW-1 are integrated herein. Because the length of the domain is larger than the SOF = 28.81 m of CPTU-Su model [37] , the spatial correlation of transformation uncertainty $\boldsymbol{\varepsilon}_{\text{T}}$ should be considered. Thus, the Eq. (7) is used as the covariance matrix $\boldsymbol{\Sigma}_{\text{T}}$ of $\boldsymbol{\varepsilon}_{\text{T}}$ in the likelihood (Eq. (6)). Following [37], in Eq. (7) single exponential model is adopted and given as:

$$\rho(\tau_x, \tau_z) = \exp(-2\tau_x / \delta_s - 2\tau_z / \delta_s) \tag{9}$$

in which δ_s = 27.81 m. Suppose that the SDs are σ_r = 0.05 and $\sigma_s = 0.137$, the $\sqrt{\sigma_r^2 + \sigma_s^2}$ equals to the SD of the total transformation $\sigma_{\text{T}} = 0.146$ given by [34].

Fig. 3 shows the posterior mean and standard deviation of the posterior field q_t by incorporating both the CPT-6A and MASW-1 measurements. For the purpose of comparison, the in situ 2D shear wave velocity profile from MASW-1 and the 1D corrected cone tip resistance profile from CPT-6A are also included in Fig. 3. As shown in Fig. 3b, both the information from MASW-1 and CPT-6A measurements greatly contributes to updating the configuration of the underground soil profile, i.e., the whole 2D domain, including the areas far away from CPT-6A, have clearly identified strata and details of corrected cone tip resistance.

Fig. 3c shows the posterior standard deviation of the q_t field. Compared to the case where only CPTU data is involved in the probabilistic inversion and only local uncertainty is reduced near CPT-6A, the standard deviation of the whole q_t field has been greatly reduced. This is because the MASW tests provide information at un-sampled places. It is also noted that the standard deviations of q_t at the locations far away from CPTU are still lightly higher than the ones near the CPTU. This may be because that the CPT-6A cannot provide reliable and enough information for areas remote from the CPTU.

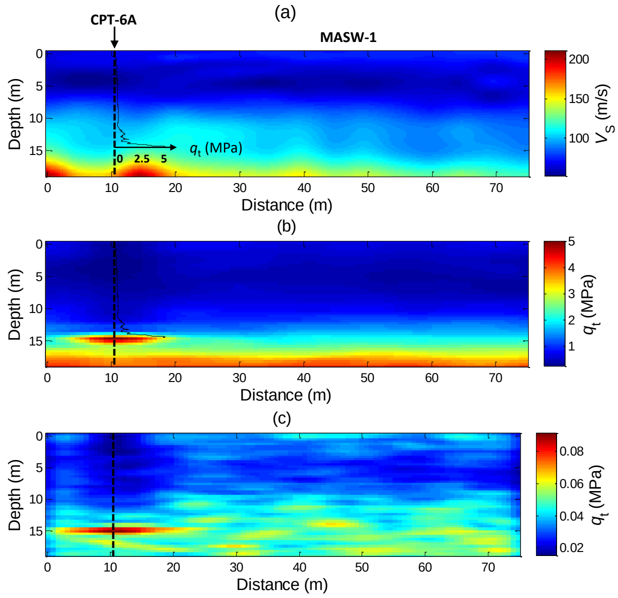

Figure 3. (a) In situ 2D shear wave velocity profile from MASW-1 and the 1D corrected cone tip resistance profile from CPT-6A, (b) Mean and (c) standard deviation of the posterior realizations of the 2D corrected cone tip resistance field integrating field data $\mathbf{q}_{t,1D}^{obs}$ and $\mathbf{V}_{s,2D}^{obs}$.

4. Conclusions

The main conclusion of this study is that the combination of geotechnical tests and geophysical tests can reduce uncertainties of 2D soil profiles in geotechnical site characterization.

Based on random field theory and Bayes' theorem, a rigorous probabilistic site characterization framework is proposed. The results of the case study of Ballina station show that this novel methodology is fully applicable to field data and can be used to estimate the spatial distribution of the corrected cone tip resistance field on a regional scale.

The method also provides a quantitative measure of uncertainty throughout the domain. This data-combining method allows a geotechnical analyst to develop a probabilistic model of the soil domain conditioned on site investigation data which can be utilized in subsequent engineering calculations. Therefore, this approach is a fundamental precursor to the probabilistic method being used routinely in engineering practice. Although only cone tip resistance field is considered, the framework is also suitable for other engineering properties of soils.

Acknowledgement

The project presented in this article is supported by the National Natural Science Foundation of China (Projects 51679117, 41972280) and the Discovery project DP190101592 from Australian Research Council.

References

[1] Robertson, P. K., "Interpretation of cone penetration tests — a unified approach" Canadian Geotechnical Journal, 46, pp. 1337–1355, 2009. https://doi.org/10.1139/T09-065.

[2] Phoon Kok-Kwang, Quek Ser-Tong, An Ping, "Identification of Statistically Homogeneous Soil Layers Using Modified Bartlett Statistics" Journal of Geotechnical and Geoenvironmental Engineering, 129, pp. 649–659, 2003. https://doi.org/10.1061/(ASCE)1090-0241(2003)129:7(649).

[3] Houlsby, N. M. T., Houlsby, G. T., "Statistical fitting of undrained strength data" Geotechnique; London, 63, pp. 1253–1263, 2013. https://doi.org/10.1680/geot.13.P.007

[4] Wang, Y., Huang, K., Cao, Z., "Bayesian identification of soil strata in London clay" Geotechnique; London, 64, pp. 239–246, 2014. https://doi.org/10.1680/geot.13.T.018

[5] Wang, Y., Huang, K., Cao, Z., "Probabilistic identification of underground soil stratification using cone penetration tests" Canadian Geotechnical Journal, 50, pp. 766–776, 2013. https://doi.org/10.1139/cgj-2013-0004.

[6] Ching, J., Wang, J.-S., Juang, C. H., Ku, C.-S., "Cone penetration test (CPT)-based stratigraphic profiling using the wavelet transform modulus maxima method" Canadian Geotechnical Journal, 52, pp. 1993–2007, 2015. https://doi.org/10.1139/cgj-2015-0027.

[7] Li, D.-Q., Qi, X.-H., Cao, Z.-J., Tang, X.-S., Phoon, K.-K., Zhou, C.-B., "Evaluating slope stability uncertainty using coupled Markov chain" Computers and Geotechnics, 73, pp. 72–82, 2016. https://doi.org/10.1016/j.compgeo.2015.11.021.

[8] Park, C., "MASW (Horizontal Resolution in 2D Shear-Velocity (Vs) Mapping" , 12, .

[9] Park, C. B., Miller, R. D., Xia, J., "Multichannel analysis of surface waves" GEOPHYSICS, 64, pp. 800–808, 1999. https://doi.org/10.1190/1.1444590.

[10] Pineda, J. A., Suwal, L. P., Kelly, R. B., Bates, L., Sloan, S. W., "Characterisation of Ballina clay" Géotechnique, 66, pp. 556–577, 2016. https://doi.org/10.1680/jgeot.15.P.181.

[11] Vanmarcke, E., "Random fields: Analysis and synthesis" Cambridge, MA, MIT Press, 1983, 394 p., 1983.

[12] Vanmarcke, E. H., "Probabilistic modeling of soil profiles" Journal of the geotechnical engineering division, 103, pp. 1227–1246, 1977.

[13] Baecher, G., "Simplified Geotechnical Data Analysis" In; "Simplified Geotechnical Data Analysis;" 1983; pp. 257–277. https://doi.org/10.1007/978-94-009-6896-7_9.

[14] DeGroot Don J., Baecher Gregory B., "Estimating Autocovariance of In-Situ Soil Properties" Journal of Geotechnical Engineering, 119, pp. 147–166, 1993. https://doi.org/10.1061/(ASCE)0733-9410(1993)119:1(147).

[15] Phoon, K.-K., Kulhawy, F. H., "Characterization of geotechnical variability" Canadian Geotechnical Journal, 36, pp. 612–624, 1999. https://doi.org/10.1139/t99-038.

[16] Uzielli, M., Vannucchi, G., Phoon, K., "Random field characteris ation of stress-nomalised cone penetration testing parameters" Ge otechnique, 55, pp. 3–20, 2005. https://doi.org/10.1680/geot.2005 .55.1.3

[17] Fenton Gordon A., "Random Field Modeling of CPT Data" Journal of Geotechnical and Geoenvironmental Engineering, 125, pp. 486–498, 1999. https://doi.org/10.1061/(ASCE)1090-0241(1999)125:6(486).

[18] Vanmarcke, E., Fenton, G. A., Eds.; "Probabilistic Site Characterization at the National Geotechnical Experimentation Sites;" American Society of Civil Engineers: Reston, VA; https://doi.org/10.1061/9780784406694.

[19] Lloret-Cabot, M., Fenton, G. A., Hicks, M. A., "On the estimatio n of scale of fluctuation in geostatistics" Georisk: Assessment an d Management of Risk for Engineered Systems and Geohazards, 8,pp.129–140, 2014. https://doi.org/10.1080/17499518.2013.871 189.

[20] Cao, Z.-J., Wang, Y., Li, D.-Q., "Site-specific characterization of soil properties using multiple measurements from different test procedures at different locations – A Bayesian sequential updating approach" Engineering Geology, 211, pp. 150–161, 2016. https://doi.org/10.1016/j.enggeo.2016.06.021.

[21] Tarantola, A., "Inverse Problem Theory and Methods for Model Parameter Estimation;" SIAM;

[22] Ang, A., Tang, W., Translators, Chen, J.-B., Peng, Y., Liu, W., Ai, X., Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering 2nd",.2007, John

Willey & Sons. Inc, Hoboken, New Jersey. https://doi.org/10.1080/15732470802027894
[23] Robertson P. K., “CPT-DMT Correlations” Journal of Geotechnical and Geoenvironmental Engineering, 135, pp. 1762–1771, 2009. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000119.
[24] Ghanem, R. G., Spanos, P. D., “Representation of Stochastic Processes” In Stochastic Finite Elements: A Spectral Approach; “Representation of Stochastic Processes;” Springer New York: New York, NY, 1991; pp. 15–65. https://doi.org/10.1007/978-1-4612-3094-6_2.
[25] Phoon, K. K., Huang, S. P., Quek, S. T., “Implementation of Karhunen–Loeve expansion for simulation using a wavelet-Galerkin scheme” Probabilistic Engineering Mechanics, 17, pp. 293–303, 2002. https://doi.org/10.1016/S0266-8920(02)00013-9.
[26] Huang, J., Lyamin, A. V., Griffiths, D. V., Krabbenhoft, K., Sloan, S. W., “Quantitative risk assessment of landslide by limit analysis and random fields” Computers and Geotechnics, 53, pp. 60–67, 2013. https://doi.org/10.1016/j.compgeo.2013.04.009.
[27] Cao, Z., Wang, Y., Li, D., “Quantification of prior knowledge in geotechnical site characterization” Engineering Geology, 203, pp. 107–116, 2016. https://doi.org/10.1016/j.enggeo.2015.08.018.
[28] Wang, Y., Au, S.-K., Cao, Z., “Bayesian approach for probabilist ic characterization of sand friction angles” Engineering Geology, 114, pp. 354–363, 2010. https://doi.org/10.1016/j.enggeo.2010.0 5.013.
[29] Cao Zijun, Wang Yu, “Bayesian Approach for Probabilistic Site Characterization Using Cone Penetration Tests” Journal of Geotechnical and Geoenvironmental Engineering, 139, pp. 267–276, 2013. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000765.
[30] Cao, Z., Wang, Y., “Bayesian model comparison and selection of spatial correlation functions for soil parameters” Structural Safet y, 49, pp. 10–17, 2014. https://doi.org/10.1016/j.strusafe.2013.06 .003.
[31] Laloy, E., Rogiers, B., Vrugt, J. A., Mallants, D., Jacques, D., “Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion” Water Resources Research, 49, pp. 2664–2682, 2013. https://doi.org/10.1002/wrcr.20226.
[32] Wang, B., Vardon, P. J., Hicks, M. A., “The Random Material Po int Method” In; “The Random Material Point Method;” America n Society of Civil Engineers, 2017; pp. 460–466. https://doi.org/1 0.1061/9780784480717.044.
[33] Tian, M., Li, D.-Q., Cao, Z.-J., Phoon, K.-K., Wang, Y., “Bayesi an identification of random field model using indirect test data” E ngineering Geology, 210, pp. 197–211, 2016. https://doi.org/10.1 016/j.enggeo.2016.05.013.
[34] Mayne, P. W., Rix, G. J., “Correlations Between Shear Wave Ve -locity and Cone Tip Resistance in Natural Clays” Soils and Fou ndations, 35, pp. 107–110, 1995. https://doi.org/10.3208/sandf19 72.35.2_107.
[35] L’Heureux Jean-Sebastien, Long Michael, “Relationship between Shear-Wave Velocity and Geotechnical Parameters for Norwegi an Clays” Journal of Geotechnical and Geoenvironmental Engine ering, 143, 04017013, 2017. https://doi.org/10.1061/(ASCE)GT.1 943-5606.0001645.
[36] Wang, Y., Aladejare, A. E., “Selection of site-specific regression model for characterization of uniaxial compressive strength of ro ck” International Journal of Rock Mechanics and Mining Science s, 75, pp. 73–81, 2015. https://doi.org/10.1016/j.ijrmms.2015.01. 008.
[37] Ching, J., Phoon, K.-K., Wu, T.-J., “Spatial correlation for transf ormation uncertainty and its applications” Georisk: Assessment a nd Management of Risk for Engineered Systems and Geohazards , 10, pp. 294–311, 2016. https://doi.org/10.1080/17499518.2016. 1205749.
[38] Laloy, E., Vrugt, J. A., “High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing” Water Resources Research, 482012. https://doi.org/10.1029/2011WR010608.
[39] Robertson, P. K., “Soil classification using the cone penetration test” Canadian Geotechnical Journal, 27, pp. 151–158, 1990. https://doi.org/10.1139/t90-014.
[40] Robertson, P. K., Wride, C. E., “Evaluating cyclic liquefaction potential using the cone penetration test” Canadian Geotechnical Journal; Ottawa, 35, pp. 442–459, 1998. https://doi.org/10.1139/cgj-35-3-442
[41] Kelly, R., Pineda, J., Bates, L., Suwal, L., Fitzallen, A., “Site characterisation for the Ballina field testing facility” Faculty of Engineering and Information Sciences - Papers: Part B, pp. 279–300, 2017. https://doi.org/10.1680/jgeot.15.P.211.
[42] Ching, J., Wu, S.-S., Phoon, K.-K., “Statistical characterization of random field parameters using frequentist and Bayesian approaches” Canadian Geotechnical Journal, 53, pp. 285–298, 2015. https://doi.org/10.1139/cgj-2015-0094.
[43] Jaksa Mark B., Brooker Peter I., Kaggwa William S., “Inaccuracies Associated with Estimating Random Measurement Errors” Journal of Geotechnical and Geoenvironmental Engineering, 123, pp. 393–401, 1997. https://doi.org/10.1061/(ASCE)1090-0241(1997)123:5(393).
[44] Stuedlein Armin W., Kramer Steven L., Arduino Pedro, Holtz Robert D., “Geotechnical Characterization and Random Field Modeling of Desiccated Clay” Journal of Geotechnical and Geoenvironmental Engineering, 138, pp. 1301–1313, 2012. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000723.
[45] Kenarsari, A. E., Chenari, R. J., Eslami, A., “Characterization of the correlation structure of residual CPT profiles in sand deposits” International Journal of Civil Engineering, 11, 9, 2013.
[46] Li, D.-Q., Qi, X.-H., Phoon, K.-K., Zhang, L.-M., Zhou, C.-B., “ Effect of spatially variable shear strength parameters with linearl y increasing mean trend on reliability of infinite slopes” Structura l Safety, 49, pp. 45–55, 2014. https://doi.org/10.1016/j.strusafe.2 013.08.005.