Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_01-004-cd

Bayesian Hierarchical Spatial Modeling of Soil Properties

Iason Papaioannou[1], Sebastian Geyer[1] and Daniel Straub[1]

[1]Engineering Risk Analysis Group, Technische Universität München, Arcisstr. 21, 80290 München, Germany.
E-mail: {iason.papaioannou,s.geyer,straub}@tum.de

Abstract: In geotechnical engineering, the mechanical properties of the soil at a specific site are usually inferred based on data obtained from in-situ or laboratory measurements. Results of geotechnical design and assessment depend on the proper quantification of the soil properties based on typically sparse site-specific data. We present a comprehensive hierarchical Bayesian model for incorporating measurement information in the probabilistic modeling of soil properties, which accounts for the inherent spatial variability of the properties and the spatial correlation of the measurements. We employ a conjugate prior approach to update the prior random field of the soil property such that closed-form expressions are obtained for the spatial posterior predictive distribution. The model requires specification of the parameters of the prior autocorrelation structure of the random field, which can be done through evaluating their maximum a-posteriori estimates. In this way, the proposed framework represents an efficient yet theoretically thorough approach to probabilistic site characterization. We illustrate the approach using real data from a geotechnical site.

Keywords: Site characterization; random fields; Bayesian analysis; hyperparameters; conjugate prior; predictive distribution.

1 Introduction

Geotechnical design and assessment require the specification of the mechanical properties of the soil material. Models of soil properties are governed by uncertainties related to the inherent variability of the soil formation process, measurement and transformation errors, as well as statistical uncertainty from limited availability of data (Phoon and Kulhawy 1999). Geotechnical engineers are often faced with the task of quantifying these uncertainties through combining experience with limited data collected at a specific site. Bayesian analysis provides an effective means for performing this task (Straub and Papaioannou 2015b, Wang et al. 2016). Thereby, information available prior to the project can be combined with in-situ data through application of Bayes' rule. Bayesian analysis is particularly suitable for modeling soil properties as it enables a consistent representation of the inherent spatial variability of the soil material (Papaioannou and Straub 2017). The prior distribution of the inherent variability is modeled by a random field and Bayesian analysis is applied to update the point-in-space distribution of the soil property given in-situ data.

Estimating the posterior distribution of random fields can be performed with sampling-based inference methods (e.g., Straub and Papaioannou 2015a, Jiang et al. 2018). Sometimes, it is possible to use analytical solutions through conjugate prior models (Papaioannou and Straub 2017). The approach of (Ching and Phoon 2017, Ching et al. 2020) combines sparse Bayesian learning with Markov chain Monte Carlo (MCMC) to draw samples from the posterior distribution of the random field. The authors of (Wang and Zhao 2017, Wang et al. 2018) apply Bayesian compressive sampling to represent spatially variable soil properties with sparse measurements. In (Geyer et al. 2021), the authors propose a hierarchical Bayesian model for representing spatially variable material properties that leads to a closed-form expression for the posterior predictive distribution of the random field given spatial data. The model enables the representation of non-Gaussian quantities through an isoprobabilistic transformation of the data. In cases where the choice of the prior correlation parameters of the random field is not straightforward, they can be selected as maximum a-posteriori estimates.

In this work, we study the performance of the hierarchical random field model of Geyer et al. (2021) in representing the spatial distribution of a geotechnical dataset. We employ the CPT data from a stiff clay site, reported in Jaksa et al. (1995, 1999), and study the influence of the data size as well as different prior choices on the analysis results.

2 Hierarchical random field model

In this section, we summarize the hierarchical random field modeling technique presented in Geyer et al. (2021), which enables learning the point-in-space distribution of material properties with spatially distributed data. The model assumes that the prior distribution of the soil property of interest is described by a homogeneous random field $Y(z)$, $z \in \Omega$, with Ω denoting a spatial domain of interest, and that $Y(z)$ can be expressed as a nonlinear transform of an underlying Gaussian random field $X(z)$. In this work, we employ a lognormal random field, which has a positive support and, hence, is commonly used for modeling soil material properties (e.g., Griffiths and Fenton 2001). In this case, the random field $Y(z)$ can be expressed through the transformation

$$Y(z) = \exp[X(z)]. \tag{1}$$

The Gaussian random field $X(z)$ is defined by the parameters of its marginal distribution $\boldsymbol{\theta} = [\mu_X, \lambda_X]$, where μ_X is the mean value and λ_X is the precision (inverse of the variance), and by its autocorrelation function $\rho(\Delta z|\boldsymbol{\tau})$, with Δz being the difference in location and $\boldsymbol{\tau}$ denoting the parameters of the chosen correlation model.

The hierarchical modeling approach treats the vector $\boldsymbol{\theta}$ as uncertain and employs direct observations of $Y(z)$ to learn its distribution through Bayesian analysis. In particular, a conjugate prior distribution is imposed on $\boldsymbol{\theta}$ such that a closed form expression for its posterior distribution can be obtained. Moreover, the posterior predictive distribution of the random field $Y(z)$ conditional on the correlation parameters $\boldsymbol{\tau}$ can also be derived in closed form. A point estimate of the parameters $\boldsymbol{\tau}$ can be obtained through determining the maximum of their posterior distribution. Next, we outline the steps of this updating process.

2.1 Prior distribution and likelihood function

The prior distribution of the parameters $\boldsymbol{\theta}$ of the underlying Gaussian random field $X(z)$ is chosen as the normal-gamma model, with probability density function (PDF)

$$f(\boldsymbol{\theta}) = \mathcal{NG}(\mu_X, \lambda_X|\mu_0, \kappa_0, \alpha_0, \beta_0) = \mathcal{N}(\mu_X|\mu_0, \kappa_0\lambda_X) \cdot \mathcal{G}(\lambda_X|\alpha_0, \beta_0), \tag{2}$$

where $\mathcal{N}(\mu_X|\mu_0, \kappa_0\lambda_X)$ is the PDF of the normal distribution with mean μ_0 and precision $\kappa_0\lambda_X$ and $\mathcal{G}(\lambda_X|\alpha_0, \beta_0)$ is the PDF of the gamma distribution with shape parameter α_0 and rate parameter β_0.

Consider a set of available spatially distributed measurements $\mathbf{M}$ of $Y(z)$, where $\mathbf{M} = [\mathbf{M}_1; \cdots; \mathbf{M}_n]$ and $\mathbf{M}_i = [y_{m,i}; z_{m,i}]$, with $y_{m,i}$ denoting the measurement outcome at location $z_{m,i}$. Neglecting the measurement uncertainty, the likelihood function of the parameters $\boldsymbol{\theta}$ given measurements $\mathbf{M}$ and correlation parameters $\boldsymbol{\tau}$ is the n-variate normal PDF of the transformed measurements $\mathbf{x}_m = \left[\log(y_{m,1}); \ldots; \log(y_{m,n})\right]$,

$$L(\boldsymbol{\theta}|\mathbf{M}, \boldsymbol{\tau}) = \frac{\lambda_X^{\frac{n}{2}}}{(2\pi)^{\frac{n}{2}}(\det \mathbf{R}_m)^{\frac{1}{2}}} \exp\left(-\frac{\lambda_X}{2}(\mathbf{x}_m - \mu_X\mathbf{1}_n)\mathbf{R}_m^{-1}(\mathbf{x}_m - \mu_X\mathbf{1}_n)^{\mathrm{T}}\right), \tag{3}$$

where $\mathbf{R}_m$ is the correlation matrix at the measurement locations, with element (i, j) equal to $\rho(z_{m,i} - z_{m,j}|\boldsymbol{\tau})$, and $\mathbf{1}_n$ denotes a $1 \times n$-vector of ones.

2.2 Posterior distribution

The posterior distribution of $\boldsymbol{\theta}$ is obtained by application of Bayes' rule as

$$f(\boldsymbol{\theta}|\mathbf{M}, \boldsymbol{\tau}) \propto L(\boldsymbol{\theta}|\mathbf{M}, \boldsymbol{\tau}) \cdot f(\boldsymbol{\theta}). \tag{4}$$

The chosen normal-gamma prior distribution model of Eq. (2) and the multivariate Gaussian likelihood function of Eq. (3) form a conjugate pair. Hence, the posterior distribution of Eq. (4) can be derived analytically and retains the same parametric form of the prior, i.e., it is a normal-gamma distribution, with parameters given as follows:

$$\mu_n = \frac{\kappa_0\mu_0 + \mathbf{1}_n\mathbf{R}_m^{-1}\mathbf{x}_m^{\mathrm{T}}}{\kappa_0 + \mathbf{1}_n\mathbf{R}_m^{-1}\mathbf{1}_n^{\mathrm{T}}}, \tag{5}$$

$$\kappa_n = \kappa_0 + \mathbf{1}_n\mathbf{R}_m^{-1}\mathbf{1}_n^{\mathrm{T}}, \tag{6}$$

$$\alpha_n = \alpha_0 + \frac{n}{2}, \tag{7}$$

$$\beta_n = \beta_0 + \frac{1}{2}\left(\mathbf{x}_m\mathbf{R}_m^{-1}\mathbf{x}_m^{\mathrm{T}} + \frac{\kappa_0\mu_0^2\mathbf{1}_n\mathbf{R}_m^{-1}\mathbf{1}_n^{\mathrm{T}} - 2\kappa_0\mu_0\mathbf{1}_n\mathbf{R}_m^{-1}\mathbf{x}_m^{\mathrm{T}} - (\mathbf{1}_n\mathbf{R}_m^{-1}\mathbf{x}_m^{\mathrm{T}})^2}{\kappa_0 + \mathbf{1}_n\mathbf{R}_m^{-1}\mathbf{1}_n^{\mathrm{T}}}\right). \tag{8}$$

A detailed derivation of this result can be found in (Geyer et al. 2021).

2.3 Posterior predictive random field

Having evaluated the posterior distribution of the random field parameters, we can now obtain the posterior predictive distribution of the random field $Y(z)$, which is required to perform updated predictions given the data $\mathbf{M}$. We first derive the predictive distribution of the underlying Gaussian random field $X(z)$ at locations $\mathbf{Z} = [z_1; \ldots; z_k]$. The joint PDF of X at these k locations given the measurements $\mathbf{M}$, also known as the kth order posterior predictive PDF of X, is given as

$$f(\mathbf{x}; \mathbf{Z}|\mathbf{M}, \boldsymbol{\tau}) = \int_\Theta f(\mathbf{x}; \mathbf{Z}|\boldsymbol{\theta}, \mathbf{M}, \boldsymbol{\tau}) f(\boldsymbol{\theta}|\mathbf{M}, \boldsymbol{\tau}) d\boldsymbol{\theta}, \tag{9}$$

where Θ is the domain of definition of $\boldsymbol{\theta}$. $f(\mathbf{x}; \mathbf{Z}|\boldsymbol{\theta}, \mathbf{M})$ is the kth order PDF of X given $\boldsymbol{\theta}$ and $\mathbf{M}$. Since $X(z)$ is a Gaussian random field, the conditional kth order PDF is also Gaussian with parameters (mean vector and precision matrix) available in closed form (e.g., see Papaioannou and Straub 2017). Substituting this result along with the

normal-gamma posterior distribution of the parameters $\boldsymbol{\theta}$ into Eq. (9) and solving the integral, one obtains the following result (Geyer et al. 2021)

$$f(\mathbf{x};\mathbf{Z}|\mathbf{M},\boldsymbol{\tau}) = f_t\left(\mathbf{x}|\boldsymbol{\mu}_{\mathbf{Z},t},\boldsymbol{\Lambda}_{\mathbf{Z},t},\nu_t\right), \tag{10}$$

where $f_t\left(\mathbf{x}|\boldsymbol{\mu}_{\mathbf{Z},t},\boldsymbol{\Lambda}_{\mathbf{Z},t},\nu_t\right)$ is the k-variate Student's t-distribution with location vector $\boldsymbol{\mu}_{\mathbf{Z},t}$, scale matrix $\boldsymbol{\Lambda}_{\mathbf{Z},t}$ and degrees of freedom ν_t, given by the following expressions

$$\boldsymbol{\mu}_{\mathbf{Z},t} = \mu_n \mathbf{1}_k^{\mathrm{T}} + \mathbf{R}_{\mathbf{Z},m}\mathbf{R}_m^{-1}(\mathbf{x}_m - \mu_n \mathbf{1}_n)^{\mathrm{T}}, \tag{11}$$

$$\boldsymbol{\Lambda}_{\mathbf{Z},t} = \frac{\alpha_n}{\beta_n}\left(\mathbf{R}_{\mathbf{Z}} - \mathbf{R}_{\mathbf{Z},m}\mathbf{R}_m^{-1}\mathbf{R}_{\mathbf{Z},m}^{\mathrm{T}} + \left(\mathbf{1}_k^{\mathrm{T}} - \mathbf{R}_{\mathbf{Z},m}\mathbf{R}_m^{-1}\mathbf{1}_n^{\mathrm{T}}\right)\kappa_n^{-1}\left(\mathbf{1}_k^{\mathrm{T}} - \mathbf{R}_{\mathbf{Z},m}\mathbf{R}_m^{-1}\mathbf{1}_n^{\mathrm{T}}\right)^{\mathrm{T}}\right)^{-1}, \tag{12}$$

$$\nu_t = 2\alpha_n. \tag{13}$$

Here, $\mathbf{R}_{\mathbf{Z},m}$ is a $k \times n$-matrix with element (i,j) equal to $\rho\left(z_i - z_{m,j}\middle|\boldsymbol{\tau}\right)$ and $\mathbf{R}_{\mathbf{Z}}$ is a $k \times k$-matrix with element (i,j) equal to $\rho\left(z_i - z_j\middle|\boldsymbol{\tau}\right)$. The k-variate posterior predictive distribution of the random field $Y(z)$ can be obtained from Eqs. (1) and (10) through probabilistic transformation, which results in a log-Student's t-distribution model. Geyer et al. (2021) discuss how to sample from the predictive distribution of $Y(z)$ through a nonlinear transform of a sample from an underlying Gaussian random field. Therein, it is also shown that the log-Student's t-distribution has diverging integer moments, i.e., it has infinite mean and variance.

2.4 Maximum-a-posteriori estimation of correlation parameters

The choice of the parameters $\boldsymbol{\tau}$ of the prior autocorrelation function of the random field $X(z)$ has significant influence on the predictions obtained by the proposed model. If available information on the modeled quantity is limited, the parameter vector $\boldsymbol{\tau}$ can be treated as uncertain and its full posterior distribution can be evaluated and accounted for in the computation of the predictive distribution of the random field. However, this posterior distribution cannot be evaluated in closed form and one needs to resort to numerical techniques, such as MCMC methods, as is done, e.g., in the sparse Bayesian learning approach of Ching and Phoon (2017). In (Geyer et al. 2021), the use of a point estimate of $\boldsymbol{\tau}$ is proposed, which is obtained as the maximum a-posteriori (MAP) estimate. That is, we solve:

$$\boldsymbol{\tau}^* = \arg\max_{\boldsymbol{\tau}\in\mathbf{T}} f(\boldsymbol{\tau}|\mathbf{M}), \tag{14}$$

where $f(\boldsymbol{\tau}|\mathbf{M})$ denotes the posterior PDF of $\boldsymbol{\tau}$ given the measurements and $\mathbf{T}$ is the domain of definition of $\boldsymbol{\tau}$. The program of Eq. (14) can be equivalently written as (Geyer et al. 2021)

$$\boldsymbol{\tau}^* = \arg\max_{\boldsymbol{\tau}\in\mathbf{T}} \ln\left(\kappa_n(\boldsymbol{\tau})\right) + 2\alpha_n \ln\left(\beta_n(\boldsymbol{\tau})\right) + \ln\left(\det\left(\mathbf{R}_m(\boldsymbol{\tau})\right)\right) - 2\ln f(\boldsymbol{\tau}), \tag{15}$$

where $f(\boldsymbol{\tau})$ is the chosen prior distribution of $\boldsymbol{\tau}$.

3 Numerical example

In this section, we demonstrate the performance of the proposed hierarchical random field method with a real geotechnical dataset. We employ the CPT data reported in Jaksa et al. (1995, 1999) from a relatively homogeneous stiff, overconsolidated clay, known as Keswick Clay, collected at a site in Adelaide, South Australia. In particular, we use the data from the vertical CPT at location $x = 40\mathrm{m}$, $y = 40\mathrm{m}$ (location ID 39) and at depths $z = 1.101\mathrm{m} - 5.054\mathrm{m}$. The soundings for depths $< 1.101\mathrm{m}$ were removed in accordance with Jaksa et al. (1999), as they are not consistent with Keswick Clay. We further assume that knowledge of the full data set is not available but only a subset of 16 measurements taken at equidistant locations, which are shown in panels (a) and (c) of Figure 1.

The hierarchical random field model presented in Section 2 requires the choice of the hyperparameters of the prior normal-gamma distribution of the mean μ_X and precision λ_X. Here, we employ an uninformative prior by choosing the hyperparameters as $[\mu_0, \kappa_0, \alpha_0, \beta_0] = [/, 0, -1/2, 0]$ (DeGroot 1969). Moreover, we model the autocorrelation function of the underlying Gaussian field by the exponential model, i.e.,

$$\rho(\Delta z|\delta_v) = \exp\left(-\frac{2\Delta z}{\delta_v}\right), \tag{16}$$

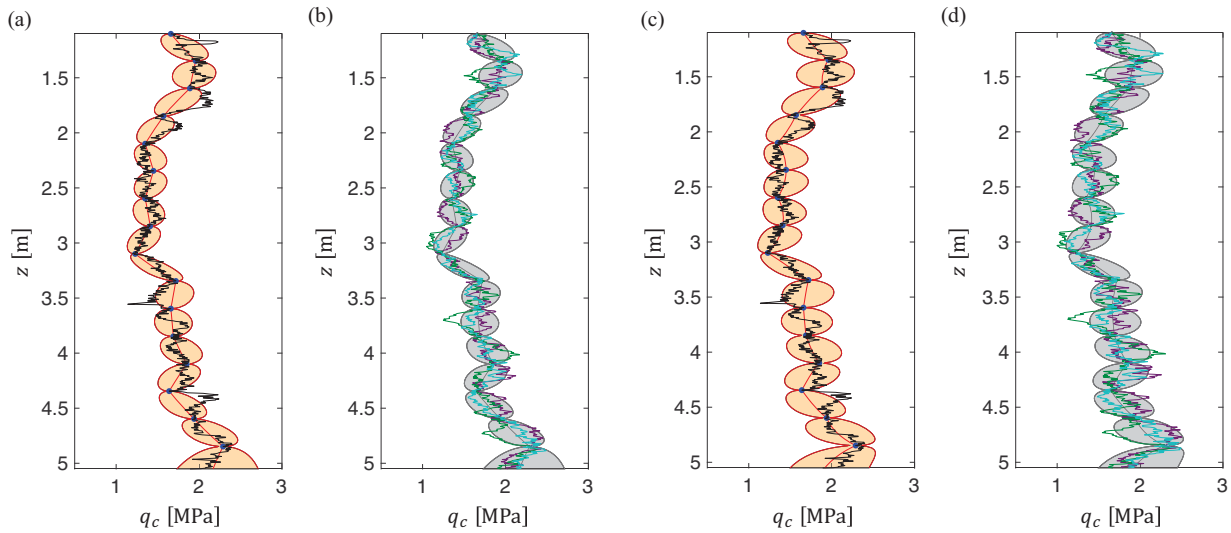

Figure 1. Posterior predictive random field of the cone tip resistance q_c; (a), (c) median (red line) and two-sided 90% credible interval of the log-Student's t-distribution for (a) $\delta_v = \delta_v^*$ and (c) $\delta_v = 0.3$ m. The 16 blue dots represent the used measurements and the black line is the full dataset; (b), (d) three realizations of the posterior predictive random field for (b) $\delta_v = \delta_v^*$ and (d) $\delta_v = 0.3$ m.

where δ_v represents the vertical scale of fluctuation. We consider two different choices of the prior scale of fluctuation, $\delta_v = 0.3$ m, which is the averaged observed value at a number of clay sites in the studies reported by Phoon and Kulhawy (1999), and $\delta_v = \delta_v^*$, where $\delta_v^* = 1.89$ m is the solution of the MAP optimization problem of Eq. (15) imposing a uniform prior on δ_v. Figure 1 shows the median and two-sided 90% credible intervals along with three independent realizations of the posterior predictive random field for the two considered scales of fluctuation. The credible intervals for both cases are narrow close to the measurement locations and increase away from the measurements. The intervals are wider for the choice of $\delta_v = 0.3$ m, which is much smaller than the MAP value, since for a low prior scale of fluctuation the influence of the measurements concentrates around the measurement locations and the variability away from these locations remains high. The random realizations obtained for $\delta_v = \delta_v^*$ compare well with the full dataset in terms of the magnitude and number of local variations from the posterior predictive median. Hence, the proposed approach provides a sufficient approximation of both the trend and associated uncertainty of the soil profile.

Figure 2 plots the posterior density of the scale of fluctuation together with the MAP estimate δ_v^*. It is shown that the posterior density is relatively flat over a wide range of δ_v values, but attains a distinct maximum at $\delta_v^* = 1.89$ m. We note that this value is close to the value of 1.83 m, which is obtained through a regression analysis with the full data set by Jaksa et al. (1999). For values of δ_v larger than δ_v^*, the posterior density is also high, suggesting that a further increase of the prior scale of fluctuation has low influence on the posterior predictive random field. The reason for this is that the fine spacing of the measurement locations. The posterior density decreases fast for $\delta_v < \delta_v^*$, implying that the choice of a small scale of fluctuation is not appropriate. This can be verified by observing the random realizations obtained for $\delta_v = 0.3$ m, which show higher local variability than the full dataset.

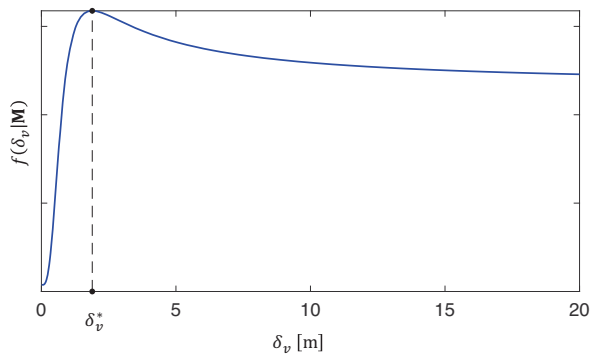

Figure 2. Posterior density of the vertical scale of fluctuation δ_v and the correspond MAP estimate δ_v^*.

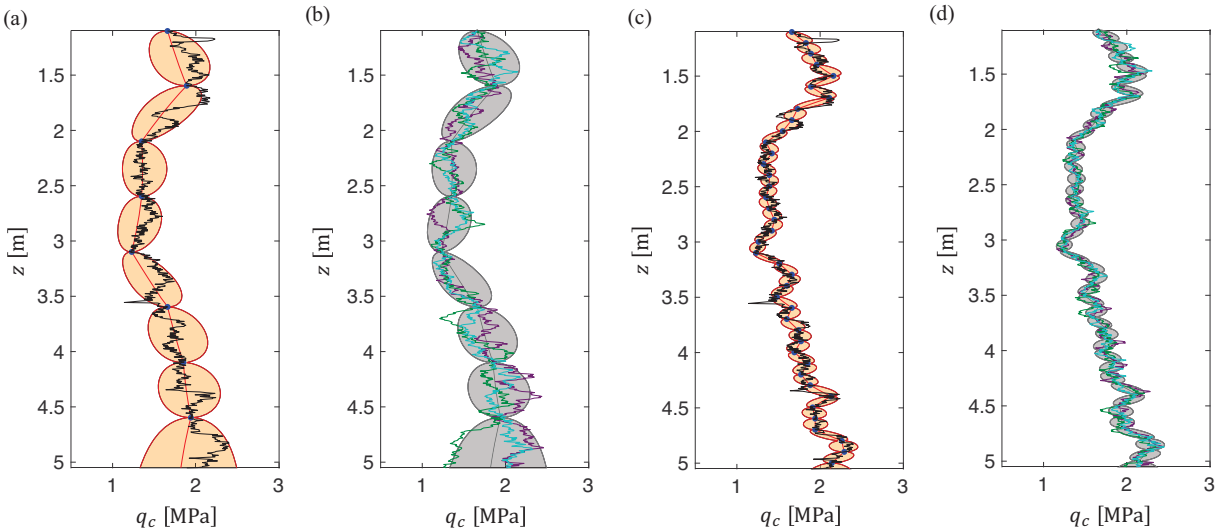

Figure 3. Posterior predictive random field of the cone tip resistance q_c for $n = 8$ and $n = 40$; (a), (c) median (red line) and two-sided 90% credible interval of the log-Student's t-distribution for (a) $n = 8$ and (c) $n = 40$. The 16 blue dots represent the used measurements and the black line is the full dataset; (b), (d) three realizations of the posterior predictive random field for (b) $n = 8$ and (d) $n = 40$.

To demonstrate the influence of the number of measurements, we repeat the analysis for $n = 8$ and $n = 40$ equidistant measurements. The measurements together with the posterior predictive 90% credible intervals are depicted in panels (a) and (c) of Figure 3; panels (b) and (d) show again three independent realizations from the posterior predictive random field. In both cases, we choose δ_v as the corresponding MAP estimate, $\delta_v^* = 1.86$ m for $n = 8$ and $\delta_v^* = 2.18$ m for $n = 40$. The figure shows that the variability of the posterior predictive random field around its median decreases with increase of the number of measurements. For small number of measurements, the credible intervals around the posterior median are larger, which is reflected by a larger variability between the independent posterior random field realizations. However, even for a small number of measurements the posterior predictive median captures the global trend of the full dataset.

4 Conclusions

This paper presents a hierarchical Bayesian approach for probabilistic site investigation. The approach models the soil property by a random field and employs a conjugate prior model to learn the parameters of its marginal distribution given spatially distributed data. The posterior predictive random field conditional on the prior correlation parameters takes the form of a log-Student's t-distribution with parameters computable in closed form. We propose to estimate the prior correlation parameters through solving a maximum a-posteriori (MAP) estimation problem. The method is applied to learn the spatial distribution of the cone tip resistance with data collected at a stiff clay site. The results demonstrate the ability of the proposed model to represent the local variability due to the measurements and reflect the uncertainty away from the measurement locations. They also show that the MAP estimation approach for choosing the prior scale of fluctuation may lead to more consistent predictive uncertainty bounds as compared to a choice based on data from other sites.

Acknowledgments

The authors would like to thank the members of the TC304 Committee on Engineering Practice of Risk Assessment & Management of the International Society of Soil Mechanics and Geotechnical Engineering for developing the database 304dB used in this study and making it available for scientific inquiry. We also wish to thank Mark Jaksa for contributing this database to the TC304 compendium of databases.

References

Ching, J., Huang, W.-H., and Phoon, K. K. (2020). 3D probabilistic cite characterization by sparse Bayesian learning. *Journal of Engineering Mechanics*, 146(12), 04020134.

Ching, J., and Phoon, K. K. (2017). Characterizing uncertain site-specific trend function by sparse Bayesian learning. *Journal of Engineering Mechanics*, 143(7), 04017028.

DeGroot, M. H. (1969). Optimal statistical decisions. *In McGraw-Hill Series in Probability and Statistics. McGraw-Hill, New York City, NY*.

Geyer, S., Papaioannou, I., and Straub, D. (2021). Bayesian analysis of hierarchical random fields for material modeling. *Probabilistic Engineering Mechanics*, 66, 103167.

Griffiths, D. V., and Fenton, G. A. (2001). Bearing capacity of spatially random soil: the undrained clay Prandtl problem revisited. *Géotechnique*, 51(4), 351-359.

Jaksa, M.B. (1995). The Influence of Spatial Variability on the Geotechnical Design Properties of a Stiff, Overconsolidated Clay. *Ph.D. Dissertation, University of Adelaide, Australia.*

Jaksa, M.B., Kaggwa, W.S., and Brooker, P.I. (1999). Experimental evaluation of the scale of fluctuation of a stiff clay. *Proc. 8th Int. Conf. on the Application of Statistics and Probability*. Sydney, AA Balkema, Rotterdam, Vol. 1, 415-422.

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

Papaioannou, I., and Straub, D. (2017). Learning soil parameters and updating geotechnical reliability estimates under spatial variability–theory and application to shallow foundations. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 11(1), 116-128.

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

Straub, D., and Papaioannou, I. (2015a). Bayesian updating with structural reliability methods. *Journal of Engineering Mechanics*, 141(3), 04014134.

Straub, D., and Papaioannou, I. (2015b). Bayesian analysis for learning and updating geotechnical parameters and models with measurements. *In Risk and Reliability in Geotechnical Engineering, (eds. K.-K. Phoon and J.Y. Ching), CRC Press, Boca Raton, FL* .

Wang, Y., Cao, Z., and Li, D. (2016). Bayesian perspective on geotechnical variability and site characterization. *Engineering Geology*, 203, 117-125.

Wang, Y., and Zhao, T. (2017). Statistical interpretation of soil property profiles from sparse data using Bayesian compressive sampling. *Géotechnique*, 67(6), 523-536.

Wang, Y., Zhao, T., and Phoon, K. K. (2018). Direct simulation of random field samples from sparsely measured geotechnical data with consideration of uncertainty in interpretation. *Canadian Geotechnical Journal*, 55(6), 862-880.