Proc. of the 9th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Zhongqiang Liu, Jian Dai and Kate Robinson
Published by Research Publishing, Singapore.
doi: 10.3850/GRF-25280825_isgsr-173-P034-cd

ISGSR2025

HIERARCHICAL BAYESIAN MODEL FOR LOAD TEST DATABASE WITH EXTREMELY SPARSE DATA

Jianye Ching

Department of Civil Engineering, National Taiwan University, Taiwan. E-mail: jyching@gmail.com

For full-scale load tests (e.g., footing, driven pile, drilled shaft, etc.), it is common that a very limited number (e.g., 1 or 2) of load tests is conducted for a given site. The consequence is that a load test database usually contains sites with extremely sparse data. The hierarchical Bayesian model (HBM) originally proposed by Ching et al. (2021) may not be suitable to analyzea load-test database. This paper proposes a novel HBM, called the$\mathbf{C}_g$-HBM, to deal with a load-test database with extremely sparse data. The idea of this novel HBM and its model structure will be discussed and elaborated. The effectiveness of this novel HBM will be demonstrated using a spread foundation database.

Keywords: Load test database; hierarchical Bayesian model; data-driven geotechnics.

1. Introduction

"Site uniqueness" is a well-known fact in geotechnical engineering: data from one site cannot be directly transferred to another site without discretionfor design purpose. The site uniqueness poses a significant challenge to geotechnical design. Because site-specific data are usually sparse, there are significant uncertainties in the design process (e.g., spatial variability, statistical uncertainty, transformation uncertainty, model uncertainty, etc.). Phoon et al. (2022) framed the challenge pertaining to site uniqueness into the so-called "site recognition challenge". For soil/rock property prediction, the site recognition challenge can be cast into the following question: how to combine soil/rock propertydata from other sites (e.g., global database) with target-site data to assist site-specificsoil/rock propertyprediction.

Ching et al. (2021) proposed a hierarchical Bayesian model (HBM) to model the site uniqueness in soil/rock property cross correlation explicitly. In the HBM, each site has its own site-specific cross-correlation model. The site-specific parameters for different sites are distinct, but their distributions are governed by the same hyper-parameters. The hyper-parameters are calibrated by a database with "site labels". The calibrated hyper-parameters learn the site-uniqueness characteristics in the database and can be subsequently adopted to construct the prior distribution of the target site. This prior can be further updated by the target-site data into the posterior. This posterior model is quasi-site-specific because it combines the information in the database and target-site data. Moreover, because the site uniqueness is explicitly modeled by the HBM, the quasi-site-specific model may be suitable for the target site.

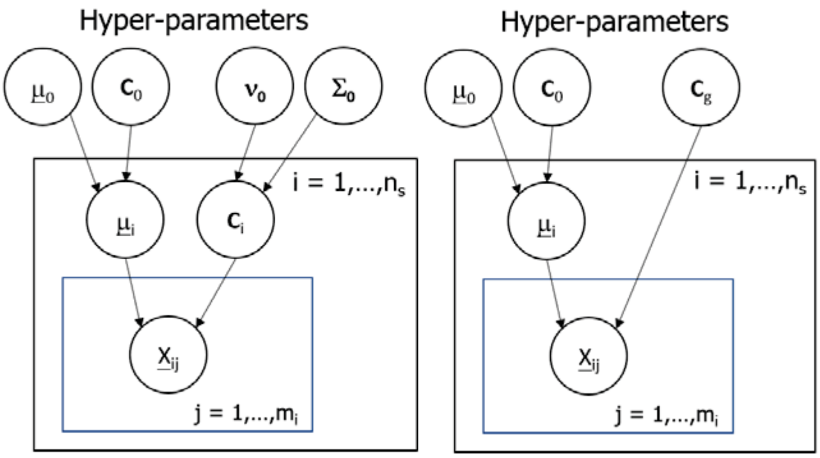
The site uniqueness is also evident in the cross correlations of load-test data.For foundation capacity prediction, the site recognition challenge can be cast into the following question: how to combine the database with the target-site data to assist site-specific foundation capacity prediction.Although it is possible to address this site recognition challenge by constructing a quasi-site-specific model based on HBM, our recent investigations indicate that the performance of the quasi-site-specific model degrades significantly if the load-test database is very sparse. A major difference between a load-test database and a soil/rock property database is that a load-test database contains much fewer sites, and each site with much fewer cases.

The main focus of the current paper is to develop a new HBM called the $\mathbf{C}_g$-HBM to deal with a load-test database with very sparse data. The idea of this novel HBM and its model structure will be discussed and elaborated. The effectiveness of this novel HBM will be demonstrated using a spread foundation database.

2. Review of Hierarchical Bayesian Model

Hierarchical Bayesian models (HBMs) (Gelman and Hill 2006) have been applied to geotechnical engineering in recent years (e.g., Zhang et al. 2016; Lu et al. 2018; Bozorgzadeh et al. 2019; Ching et al. 2021) to model site uniqueness. In this section, the HBM proposed by Ching et al. (2021) is briefly reviewed. This HBM can model the site uniqueness explicitly, so it can be trained by the database to learn its site-uniqueness characteristics. The trained HBM can construct a "prior model" for the target site, and this prior model is further updated by the target-site data to obtain the posterior model. The resulting posterior model is "quasi-site-specific" because it is based on the prior model learned from the database and then updated by the target-site data.

Let (Y_1, Y_2, … Y_n) be the parameters that are considered, where n is the parameter dimension. The parameters are transformed to normal variables (X_1, X_2, … X_n) (Ching and Phoon 2014) because the HBM assumes normality. The HBM proposed by Ching et al. (2021) addresses the site uniqueness with a three-level hierarchical model structure in Figure 1a. Let us denote the j-th case of the i-th site in the database by $\underline{X}_{ij} \in \mathbf{R}^{n\times1}$, and suppose that there are m_i cases at the i-th site. The hierarchy begins with the bottom level: $\underline{X}_{ij}$ is assumed to follow a multivariate normal probability density function (PDF), denoted by $N(\underline{x}_{ij};\underline{\mu}_i,\mathbf{C}_i)$, which is characterized by the site-specific mean vector = $\underline{\mu}_i \in \mathbf{R}^{n\times 1}$ and site-specific covariance matrix $\mathbf{C}_i \in \mathbf{R}^{n\times n}$ in the middle level. The mean vectors of different sites in the database $\{\underline{\mu}_1, \underline{\mu}_2, \ldots, \underline{\mu}_{ns}\}$ (n_s is the number of sites in the database) are distinct but assumed to follow a (hyper) multivariate normal PDF $N(\underline{\mu}_i;\underline{\mu}_0,\mathbf{C}_0)$, where $\underline{\mu}_0 \in \mathbf{R}^{n\times 1}$ and $\mathbf{C}_0 \in \mathbf{R}^{n\times n}$ are the hyper mean and hyper covariance for the site-specific mean vectors, respectively, in the top level. Similarly, the covariance matrices of different sites $\{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{ns}\}$ are distinct but assumed to follow a (hyper) inverse-Wishart PDF $IW(\mathbf{C}_i;\mathbf{\Sigma}_0,\nu_0)$, where $\mathbf{\Sigma}_0 \in \mathbf{R}^{n\times n}$ and $\nu_0 \in$ {the set of all integers}are the hyper scale matrix and hyper degree of freedom for the site-specific covariance matrices, respectively, in the top level. The parameters $\mathbf{\Theta} = \{\underline{\mu}_0, \mathbf{C}_0, \mathbf{\Sigma}_0, \nu_0\}$ are called the hyper-parameters. They govern the site-uniqueness characteristics in $\{\underline{\mu}_1, \underline{\mu}_2, \ldots, \underline{\mu}_{ns}\}$ and $\{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{ns}\}$. Let $\underline{X}_{sj} \in \mathbf{R}^{n\times 1}$ denote the j-th case of the target site, and it follows $N(\underline{x}_{sj};\underline{\mu}_s,\mathbf{C}_s)$, where $\underline{\mu}_s \in \mathbf{R}^{n\times 1}$ and $\mathbf{C}_s \in \mathbf{R}^{n\times n}$ are the target-site mean and covariance, respectively. It is also assumed that $\underline{\mu}_s$ and $\mathbf{C}_s$ are governed by the same hyper-parameters $\mathbf{\Theta}$.Let $\mathbf{D}_g = \{\underline{X}_{ij}: i = 1, \ldots, n_s; j = 1, \ldots, m_i\}$ denote the database and $\mathbf{D}_s = \{\underline{X}_{sj}: j = 1, \ldots, m_s\}$ denote the target-site data. The construction of the quasi-site-specific model for the target site consists of two stages: the learning stage and inference stage.



(a) Original HBM (b) $\mathbf{C}_g$-HBM

Figure 1 Model structures of (a) original HBM proposed by Ching et al. (2021) and (b) $\mathbf{C}_g$-HBM.

2.1. *Learning Stage*

In the learning stage, the purpose is to estimate the hyper-parameters $\mathbf{\Theta}$ based on $\mathbf{D}_g$ and subsequently to construct the prior model for the target site. The estimation of $\mathbf{\Theta}$ can be achieved by drawing $\mathbf{\Theta}$ samples ~ $f(\mathbf{\Theta}|\mathbf{D}_g)$, denoted by $\{\mathbf{\Theta}_k: k = 1, \ldots, N\}$ (N is the total number of the $\mathbf{\Theta}$ samples). Note that $\{\mathbf{\Theta}_k: k = 1, \ldots, N\}$ have learned the site-uniqueness characteristics in $\mathbf{D}_g$. Because the $(\underline{\mu}_s,\mathbf{C}_s)$ parameters for the target site are governed by the same hyper-parameters, $\{\mathbf{\Theta}_k: k = 1, \ldots, N\}$can be adopted to construct the prior PDF for the target site.

2.2. *Inference Stage*

In the inference stage, the prior PDF is further updated to the posterior (quasi-site-specific) PDF using the target-site data $\mathbf{D}_s$:

$$\begin{aligned} f\left(\underline{x}\middle|\mathbf{D}_g,\mathbf{D}_s\right) &= \int f\left(\underline{x}\middle|\underline{\mu}_s,\mathbf{C}_s\right) f\left(\underline{\mu}_s,\mathbf{C}_s\middle|\mathbf{\Theta},\mathbf{D}_s\right) f\left(\mathbf{\Theta}\middle|\mathbf{D}_g,\mathbf{D}_s\right) d\underline{\mu}_s d\mathbf{C}_s d\mathbf{\Theta} \\ &\approx \int f\left(\underline{x}\middle|\underline{\mu}_s,\mathbf{C}_s\right) f\left(\underline{\mu}_s,\mathbf{C}_s\middle|\mathbf{\Theta},\mathbf{D}_s\right) f\left(\mathbf{\Theta}\middle|\mathbf{D}_g\right) d\underline{\mu}_s d\mathbf{C}_s d\mathbf{\Theta} \approx \frac{1}{N}\sum_{k=1}^{N} f\left(\underline{x} \mid \underline{\mu}_{s|\mathbf{\Theta}_k,\mathbf{D}_s}, \mathbf{C}_{s|\mathbf{\Theta}_k,\mathbf{D}_s}\right) \end{aligned} \tag{1}$$

where $\underline{\mu}_{s|\mathbf{\Theta}k,\mathbf{D}s}$ is a sample drawn from $f(\underline{\mu}_s|\mathbf{\Theta}_k,\mathbf{D}_s)$ and $\mathbf{C}_{s|\mathbf{\Theta}k,\mathbf{D}s}$ is a sample drawn from $f(\mathbf{C}_s|\mathbf{\Theta}_k,\mathbf{D}_s)$, where$f(\underline{\mu}_s|\mathbf{\Theta}_k,\mathbf{D}_s)$ and $f(\mathbf{C}_s|\mathbf{\Theta}_k,\mathbf{D}_s)$ can be sampled using the Gibbs sampler (Ching et al. 2021). It is fair to say that the samples $\{(\underline{\mu}_{s|\mathbf{\Theta}k,\mathbf{D}s}, \mathbf{C}_{s|\mathbf{\Theta}k,\mathbf{D}s}): k = 1, \ldots, N\}$ characterize the posterior (quasi-site-specific) model for the target site. Note that in Eq. (1), $f(\mathbf{\Theta}|\mathbf{D}_g,\mathbf{D}_s)$ can be approximated by $f(\mathbf{\Theta}|\mathbf{D}_g)$if n_s is much greater than 1. The unseen parameter for the target site can be predicted by $f(x_u|\underline{x}_o,\mathbf{D}_g,\mathbf{D}_s)$:

$$f\left(x_5\middle|\underline{x}^o,\mathbf{D}_g,\mathbf{D}_s\right) \approx \frac{1}{N}\sum_{k=1}^{N} f\left(x_5 \mid \underline{x}^o, \underline{\mu}_{s|\mathbf{\Theta}_k,\mathbf{D}_s}, \mathbf{C}_{s|\mathbf{\Theta}_k,\mathbf{D}_s}\right) \tag{2}$$

where X_u is the unseen parameter to be predicted, and $\underline{x}_o$ contains the observed parameters. The effectiveness of the original HBM has been verified for soil/rock property prediction (e.g., Ching et al. 2021). However, the HBM has not be implemented to a load-test database. In the next section, we will show that the original HBM faces some challenges when implemented to a load-test database.

3. Implementation of HBM to Load-test Database

Tang et al. (2020) compiled a load-test database named NUS/SpreadFound/919 that contains 919 load tests for spread foundations. In particular, it contains 170 in-situ axial compressive load tests in cohesionless and cohesive soils with interpretable foundation capacities. For each case, the following parameters are considered:

$$\begin{aligned} &Y_1 = B\ (\text{m}) \\ &Y_2 = L\ (\text{m}) \\ &Y_3 = D_f\ (\text{m}) \\ &Y_4 = q_{uc}\ (\text{kPa}) \\ &Y_5 = q_{um}\ (\text{kPa}) \end{aligned} \qquad (3)$$

where B and L are the width and length of the spread foundation; D_f is the embankment depth; q_{um} is the measured axial compressive capacity interpreted based on the L_1-L_2 criterion proposed by Akbas and Kulhawy (2009); q_{uc} is the axial compressive capacity calculated by AASHTO (2017). These 170 load-test cases from 47 sites are adopted to form the database to illustrate the implementation of the HBM in this section.

A cross validation exercise is conducted in this section. The 15 load-test cases from the site of Brie Plateau in France (a site mainly consisting of silts), as shown in Table 1, are pulled out of the database. Note that the dimensions for most cases are the same (B = L = 1 m), except Cases #10 (B = L = 0.71 m), #11 (B = 0.71 m, L = 1.4 m), and #12 (B = 0.71 m, L = 2.1 m). The remaining 155 cases from 46 sites are adopted to train the HBM in the learning stage. To demonstrate the q_{um} prediction, let us treat the q_{um} value for the 9-th case as known, and the q_{um} values for the remaining 14 cases are hidden (the hidden values are shown in the parentheses in Table 1). Their q_{um} values are predicted in the inference stage and compared to the hidden actual values for validation. The prediction results are shown in Figure **2**a, where the red dot is for the 9-th case (treated as known), and the yellow dots are for the 14 cases (hidden).The vertical bars represent the 95% Bayesian confidence (or credible) intervals (CIs) of the prediction. Note that some cases (e.g., Cases 2-8) share the same 95% CIs because they have identical input features.It is clear that the 95% CIs for most cases are relatively narrow, but the 95% CI for Case #10 is relatively wide, and those for #11 and #12 explode (ranging from 10 to 10^6 kPa). Recall that Cases #10, #11, and #12 have distinct dimensions from Case #9.

Table 1 Load-test cases from the Brie-Plateau site.

Case	B (m)	L (m)	D_f (m)	q_{uc}(kPa)	q_{um}(kPa)
1	1	1	0	428	(461)
2	1	1	0	362	(386)
3	1	1	0	362	(377)
4	1	1	0	362	(368)
5	1	1	0	362	(335)
6	1	1	0	362	(305)
7	1	1	0	362	(400)
8	1	1	0	362	(296)
9	1	1	0	362	390
10	0.71	0.71	0	421	(438)
11	0.71	1.4	0	353	(352)
12	0.71	2.1	0	371	(346)
13	1	1	0	374	(383)
14	1	1	0	339	(433)
15	1	1	0	362	(435)

It is noteworthy that this 95% CI explosion for distinct unseen cases (e.g., Cases #11 and #12 in Figure 2a) did not occur when the HBM was previously implemented to soil/rock databases (e.g., Ching et al. 2021). As mentioned earlier, a major difference between load test and soil/rock property is that load tests are expensive, so most sites in the database have only 1 to 3 load tests. With only 1 to 3 load tests, it is not possible to reliably train the site-specific covariance $\mathbf{C}_i$ in the HBM. Because $\{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_{ns}\}$ are poorly trained, the hyper-parameters {£$_0$, ⅛} are also poorly trained, so the posterior PDF of {£$_0$, ⅛} is still highly uncertain. This high uncertainty/variability is the cause of the 95% CI explosion in Figure 2a. Note that this issue of poorly trained {**£**$_0$, ⅛} still persists for the scenario of many sites (n_s is large) with very sparse data (e.g., 1 to 3 load tests).

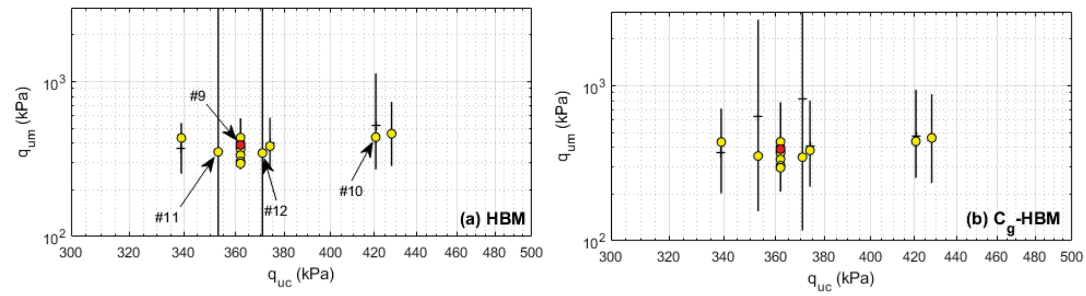


Figure 2 Prediction results for the q_{um} values: (a) original HBM; (b) $\mathbf{C}_g$-HBM.

3. New HBM for Load-test Database

In this study, we consider a new HBM called the $\mathbf{C}_g$-HBM, and its model structure is shown in Figure 1b. The $\mathbf{C}_g$-HBM allows site-specific means $\{\underline{\mu}_1, \underline{\mu}_2, \ldots, \underline{\mu}_{ns}\}$, which follow a common distribution $N(\underline{\mu}_i;\underline{\mu}_0,\mathbf{C}_0)$. However, it only allows the generic $\mathbf{C}_g$ to prevent poorly trained variance/covariance. In the learning stage, the hyper-parameters $\mathbf{\Theta} = \{\underline{\mu}_0, \mathbf{C}_0, \mathbf{C}_g\}$ samples ~ $f(\underline{\mu}_0,\mathbf{C}_0,\mathbf{C}_g|\mathbf{D}_g)$ are drawn, denoted by $\{(\underline{\mu}_{0,k}, \mathbf{C}_{0,k},\mathbf{C}_{g,k})$: k = 1, …, N}. Note that the same database $\mathbf{D}_g$ that trains the original HBM is used here to train the new HBM. In the inference stage, $\underline{\mu}_{s|\Theta k,\mathbf{D}s}$ is first drawn from $f(\underline{\mu}_s|\mathbf{\Theta}_k,\mathbf{D}_s)$, then the unseen foundation capacity (q_{um}) for a target-site case can be predicted based on the $\{(\underline{\mu}_{s|\Theta k,\mathbf{D}s},\mathbf{C}_{g,k})$: k = 1, …, N} samples:

$$f\left(x_5 \middle| \underline{x}^o, \mathbf{D}_g, \mathbf{D}_s\right) \approx \frac{1}{N}\sum_{k=1}^{N} f\left(x_5 \mid \underline{x}^o, \underline{\mu}_{s|\mathbf{\Theta}_k,\mathbf{D}_s}, \mathbf{C}_{g,k}\right) \quad (4)$$

The prediction results based on the $\mathbf{C}_g$-HBM are shown in Figure 2b. It is clear that now the 95% CIs for Cases #11 and #12 do not explode, but the 95% CIs for other cases become wider than those in Figure 2a. Because all sites share the same $\mathbf{C}_g$, data from all sites are combined to train $\mathbf{C}_g$. The issue of poorly trained variance/covariance disappears, so the 95% CI explosion disappears as well.

4. Discussion

In machine learning, the bias-variance tradeoff describes the relationship between the model complexity and prediction accuracy. In general, a complex model with many parameters (such as the HBM) may better fit the training data, but it tends to produce higher variability in the prediction for an unseen case (such as Cases #10, #11, and #12). In contrast, a simple model with limited parameters (such as the $\mathbf{C}_g$-HBM) may not fit the training data well, but it tends to produce lower variability in the prediction for an unseen case. More research is needed to clarify the bias-variance tradeoff in the context of HBM vs. $\mathbf{C}_g$-HBM.

References

AASHTO (2017). LRFD Bridge Design Specifications. 8th ed. Washington, DC, AASHTO.

Akbas, S.O. and Kulhawy, F.H. (2009). Axial compression of footings in cohesionless soils. I: Load-settlement behavior. ASCE Journal of Geotechnical and Geoenvironmental Engineering, 135(11), 1562-1574.

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.

Ching, J. and Phoon, K.K. (2014). Correlations among some clay parameters–the multivariate distribution. Canadian Geotechnical Journal, 51(6), 686-704.

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

Gelman, A. and Hill, J. (2006). Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Lu, S., Zhang, J., Zhou, S., and Xu, A. (2018). Reliability prediction of the axial ultimate bearing capacity of piles: A hierarchical Bayesian method. Advances in Mechanical Engineering, 10(11), 1-11.

Tang, C., Phoon, K.K., Li, D.Q., and Akbas, S.O. (2020). Expanded database assessment of design methods for spread foundations under axial compression and uplift loading. ASCE Journal of Geotechnical and Geoenvironmental Engineering, 146(11), 04020119.

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.