

DOI: 10.11779/CJGE20220299

场地有限数据条件下土体不排水抗剪强度的概率分布的贝叶斯估计研究

张德¹, 张泽超², 张璐璐^{*1,3,4}, 张洁⁵, 曹子君⁶

(1. 上海交通大学土木工程系海洋工程国家实验室, 上海 200240; 2. 中国长江三峡集团有限公司科学技术研究院, 北京 100038;
3. 高新船舶与深海开发装备协同创新中心, 上海 200240; 4. 上海市公共建筑和基础设施数字化运维重点实验室, 上海 200092;
5. 同济大学地下建筑与工程系, 上海 200092; 6. 西南交通大学智慧城市与交通学院高速铁路线路工程教育部重点实验室, 四川 成都 611756)

摘要: 针对勘察数据有限或不完备条件下设计参数取值不可靠的问题, 建立了包含 141 个场地共 1679 组数据的黏性土参数数据库, 分别采用基于单个场地数据的特定场地贝叶斯方法 (SBM) 和可综合利用多场地数据的层次贝叶斯方法 (HBM) 对黏性土不排水抗剪强度的概率分布进行参数估计。结果表明, 与 SBM 方法相比, 在场地仅有少量实测数据的情况下 HBM 方法能有效降低参数估计的不确定性, 且受数据量影响小。采用留一交叉验证法 (LOO-CV), 结合点预测密度的对数值 (lppd) 比较上述两种方法的准确度, 结果表明 HBM 方法的 $lppd_{loo-cv}$ 指标更大, 整体预测准确度更高。因此, HBM 方法比 SBM 方法更适用于场地有限数据情况下不排水抗剪强度概率分布的参数估计, 且 HBM 方法得到的后验统计值还可用于新场地的参数估计。

关键词: 层次贝叶斯方法; 多元正态分布; 参数估计; 交叉验证; 不排水抗剪强度

中图分类号: TU43 **文献标识码:** A **文章编号:** 1000-4548(2023)06-1259-10

作者简介: 张德(1998—), 男, 硕士研究生, 主要从事岩土工程可靠度相关的研究。E-mail: dezhang@sjtu.edu.cn。

Bayesian estimation of probability distributions of undrained shear strength of soils with limited site data

ZHANG De¹, ZHANG Zechao², ZHANG Lulu^{1,3,4}, ZHANG Jie⁵, CAO Zijun⁶

(1. State Key Laboratory of Ocean Engineering, Department of Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

2. Institute of Science and Technology, China Three Gorges Corporation, Beijing 100038, China; 3. Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration (CISSE), Shanghai 200240, China; 4. Shanghai Key Laboratory for Digital Maintenance of Buildings and Infrastructure, Shanghai 200240, China; 5. Department of Geotechnical Engineering, Tongji University, Shanghai 200092, China; 6. MOE Key Laboratory of High-Speed Railway Engineering, Institute of Smart City and Intelligent Transportation, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: To address the issue of poor reliability of the design parameters due to limited or incomplete geotechnical investigation data, a cohesive soil parameter database containing 1679 sets of data from 141 sites is established. The site-specific Bayesian method (SBM) and the hierarchical Bayesian method (HBM) are used to estimate the probability distribution of undrained shear strength of cohesive soils by utilizing the data from a specific site and multiple sites, respectively. The results show that compared with the SBM method, the HBM method can effectively reduce the uncertainty of parameter estimation when there is only limited measured data at the target site, and it is less affected by the number of measuring points at the target site. The leave-one-out cross-validation (LOO-CV) combined with the log pointwise predictive density (lppd) is used to compare the accuracy of the two methods. The results show that the $lppd_{loo-cv}$ index of the HBM method is larger, indicating that the overall prediction accuracy of the HBM method is higher. Therefore, the HBM method is more suitable for the estimation of undrained shear strength parameters in the case of limited site data, and the posterior means obtained by the HBM method can be used for parameter estimation of new sites.

Key words: hierarchical Bayesian method; multivariate normal

基金项目: 国家重点研发计划项目(2021YFB2600700); 中国长江三峡集团有限公司科研资助项目(WWKY-2020-0741); 国家自然科学基金项目(52025094, 51979158); 上海市教育委员会科研创新计划(2021-01-07-00-02-E00089)

收稿日期: 2022-03-21

***通信作者** (E-mail: lulu_zhang@sjtu.edu.cn)

distribution; parameter estimation; cross-validation; undrained shear strength

0 引言

参数确定是岩土工程设计的基础。土体性质不确定性显著, 工程实践中受经济和技术等因素制约, 单个场地的土工测试数量有限, 基于有限数据确定设计参数存在可靠性低的问题^[1-2]。以海上风电大直径单桩工程为例, 砂土内摩擦角 φ 和黏土不排水抗剪强度 s_u 是水平承载力设计 $p-y$ 曲线法的关键参数^[3]。由于海洋工程勘察取样和测试成本高, 数据量有限, 一般根据有限的实测数据结合工程经验确定设计值, 造成设计偏保守。利用有限数据对土体强度参数进行科学估计, 降低大直径单桩设计参数的不确定性, 实现设计优化, 是当前海上风电工程提质增效的迫切需要。

目前, 国内外在岩土体参数估计方面已开展了大量研究, 提出了一些有效的方法, 具体可分为 3 类。第一类方法将单个岩土体参数假定为随机变量, 采用参数估计或非参数估计确定其概率分布^[4-7]。然而, 这类方法在单个场地、有限小样本数据条件下可靠度往往较低, 且仅能利用待估计岩土体参数的测试数据, 忽略了不同类型岩土体参数之间的相关性, 因此不能利用不同类型的数据。第二类方法通过经验公式^[8-12]将其它岩土体参数的实测值转换为待估参数。但经验公式一般通过全球或地区范围内实测数据拟合得到, 具有显著的不确定性^[9], 不同经验公式的估计差异较大^[13]。第三类方法将多种类型的岩土体参数视为多元随机变量, 以多元随机变量概率分布表示参数之间的相互关系, 利用实测数据对概率分布的模型参数进行估计, 代表性的工作有, Ching 等^[14]针对 9 种岩体参数建立了岩体参数的多维联合分布; 李典庆等^[1]和 Tang 等^[15]基于多维 Gaussian Copula 建立了多维土体参数的联合分布; 汪海林等^[16]针对黏土的 9 个土性参数, 基于贝叶斯理论对土性参数的联合后验分布进行了研究。上述研究主要利用单个场地的数据, 当数据量有限时, 结果存在较大的不确定性, 为此一些学者通过融合不同场地的数据来降低不确定性。张广文等^[2]利用多场地数据建立黏聚力的先验分布, 再利用某场地的实测数据得到后验分布; Bozorgzadeh 等^[17]和 Xiao 等^[18]利用多场地数据分别对岩石力学特性参数和剪切波速参数进行了参数估计; Ching 等^[19-20]提出了多元土体参数的混合贝叶斯方法, 可综合利用场地数据; Ching 等^[21]采用层次贝叶斯方法和混合贝叶斯方法对不排水抗剪强度比 s_u / σ'_v 和归一化锥尖阻力 $(q_t - \sigma_v) / \sigma'_v$ 进行了二元正态分布的参数估计。整体而言, 当前针对黏土不排水抗剪强度的研究主要集中在前两种方

法, 采用第三类方法的还较为有限^[21], 也较少涉及多场地数据的融合, 而且现有多场地的研究主要考虑场地间的相似性, 忽略了场地间的差异性。

针对上述不足, 本文将采用基于单个场地数据进行多维参数估计的特定场地贝叶斯方法 (Site-specific Bayesian Method, SBM) 和可利用多场地数据进行多维参数估计的层次贝叶斯方法 (Hierarchical Bayesian Method, HBM) 对不排水抗剪强度的概率分布进行估计, 并对这两种方法的准确性和适用性进行定量评价, 为海上风电桩基工程设计优化和参数确定提供理论依据。

1 多维土体参数概率分布的贝叶斯估计方法

1.1 贝叶斯参数估计的基本原理

岩土工程勘察往往涉及多个土体参数^[1], 如重度、含水率、稠度指标、不排水抗剪强度等。假定土体参数 $\mathbf{X} = (X_1, X_2, \dots, X_d)^T$ 为 d 维随机变量, 其概率密度函数的模型参数设为 $\boldsymbol{\theta}$ 。例如, 当 \mathbf{X} 服从多元正态分布时, 记 $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$, 其中 $\boldsymbol{\mu}$ 为均值列向量, \mathbf{C} 为协方差矩阵, 则模型参数 $\boldsymbol{\theta} = (\boldsymbol{\mu}, \mathbf{C})$ 。

由贝叶斯理论, 可由 \mathbf{X} 的观测数据得到 $\boldsymbol{\theta}$ 的后验分布:

$$f(\boldsymbol{\theta} | \mathbf{Data}) = \frac{f(\mathbf{Data} | \boldsymbol{\theta}) f(\boldsymbol{\theta})}{f(\mathbf{Data})} \propto f(\mathbf{Data} | \boldsymbol{\theta}) f(\boldsymbol{\theta}) \quad (1)$$

式中: \mathbf{Data} 为 \mathbf{X} 的观测数据; $f(\mathbf{Data} | \boldsymbol{\theta})$ 为似然函数, 即给定 $\boldsymbol{\theta}$ 条件下得到 \mathbf{Data} 发生的概率; $f(\boldsymbol{\theta})$ 为 $\boldsymbol{\theta}$ 的先验分布; $f(\mathbf{Data})$ 是使后验概率总和为 1 的归一化常数。

需要指出的是, 式 (1) 为通式。在不同数据条件下, 需采用不同方法得到具体形式。特定场地贝叶斯方法利用本场地的数据对模型参数进行概率估计。层次贝叶斯方法通过构建多层次的贝叶斯模型, 考虑不同场地的数据之间的相似性和差异性, 利用多个场地的数据进行模型参数的概率估计。下面介绍特定场地贝叶斯方法和层次贝叶斯方法的基本原理。

1.2 特定场地贝叶斯方法

如图 1 所示, 假定土体参数 \mathbf{X} 为 d 维随机变量, 服从多元正态分布 $\mathcal{N}(\boldsymbol{\mu}_s, \mathbf{C}_s)$, 模型参数 $\boldsymbol{\theta} = \boldsymbol{\theta}_s = (\boldsymbol{\mu}_s, \mathbf{C}_s)$, 下标 s 表示 SBM。当在某一测点得到数据 $\mathbf{x} = (x_1, x_2, \dots, x_d)^T$ 时, 式 (1) 中 $f(\mathbf{Data} | \boldsymbol{\theta})$ 可表示为

$$\begin{aligned} f(\mathbf{x}|\boldsymbol{\theta}_s) &= N(\mathbf{x}|\boldsymbol{\mu}_s, \mathbf{C}_s) \\ &= |\mathbf{C}_s|^{-1/2} (2\pi)^{-d/2} \cdot \\ &\exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_s)^T \mathbf{C}_s^{-1} (\mathbf{x} - \boldsymbol{\mu}_s)\right] \quad . \end{aligned} \quad (2)$$

式中: $N(\mathbf{x}|\boldsymbol{\mu}_s, \mathbf{C}_s)$ 为多元正态分布; $(\cdot)^T$ 表示转置。

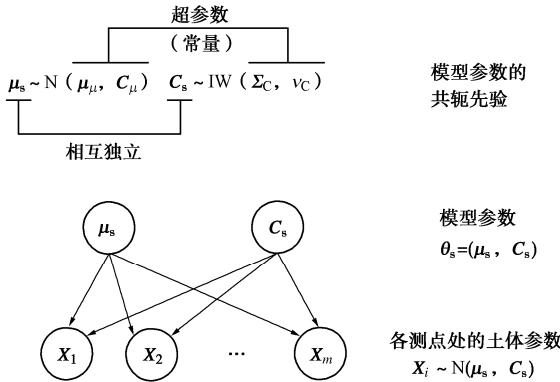


图 1 特定场地贝叶斯方法示意图

Fig. 1 Diagram of site-specific Bayesian method

如在 m 个测点获得观测数据 $\mathbf{x}_s = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$ 时, \mathbf{x}_s 为 $d \times m$ 矩阵。假设各测点土体参数独立同分布, 由式 (2) 可得 $f(\mathbf{Data}|\boldsymbol{\theta})$ 为

$$f(\mathbf{x}_s|\boldsymbol{\theta}_s) = f(\mathbf{x}_s|\boldsymbol{\mu}_s, \mathbf{C}_s) = \prod_{j=1}^m N(\mathbf{x}_j|\boldsymbol{\mu}_s, \mathbf{C}_s) \quad . \quad (3)$$

假定 $\boldsymbol{\mu}_s$ 和 \mathbf{C}_s 相互独立, 则由式 (1) 可得 $\boldsymbol{\theta}_s$ 的后验分布为

$$\begin{aligned} f(\boldsymbol{\theta}_s|\mathbf{x}_s) &= f(\boldsymbol{\mu}_s, \mathbf{C}_s|\mathbf{x}_s) \\ &\propto f(\mathbf{x}_s|\boldsymbol{\mu}_s, \mathbf{C}_s) \cdot [f(\boldsymbol{\mu}_s)f(\mathbf{C}_s)] \quad . \end{aligned} \quad (4)$$

式中: $f(\boldsymbol{\mu}_s, \mathbf{C}_s|\mathbf{x}_s)$ 为 $\boldsymbol{\mu}_s$ 和 \mathbf{C}_s 的联合后验分布; $f(\boldsymbol{\mu}_s)$ 和 $f(\mathbf{C}_s)$ 分别为 $\boldsymbol{\mu}_s$ 和 \mathbf{C}_s 的先验分布。这里分别取二者的共轭先验, 使后验分布与先验分布属于同一分布族。其中 $\boldsymbol{\mu}_s$ 的共轭先验为多元正态分布 $N(\boldsymbol{\mu}_\mu, \mathbf{C}_\mu)$, \mathbf{C}_s 的共轭先验为逆威希特分布^[22], 记为 $IW(\Sigma_C, v_C)$, 则 $\boldsymbol{\mu}_s$ 和 \mathbf{C}_s 的先验概率密度函数分别为

$$\begin{aligned} f(\boldsymbol{\mu}_s) &= N(\boldsymbol{\mu}_s|\boldsymbol{\mu}_\mu, \mathbf{C}_\mu) \\ &= |\mathbf{C}_\mu|^{-1/2} (2\pi)^{-d/2} \cdot \end{aligned}$$

$$\exp\left[-\frac{1}{2}(\boldsymbol{\mu}_s - \boldsymbol{\mu}_\mu)^T \mathbf{C}_\mu^{-1} (\boldsymbol{\mu}_s - \boldsymbol{\mu}_\mu)\right] \quad , \quad (5)$$

$$\begin{aligned} f(\mathbf{C}_s) &= IW(\mathbf{C}_s|\Sigma_C, v_C) \\ &= \frac{|\Sigma_C|^{v_C/2}}{2^{dv_C/2} \cdot \Gamma_d(v_C/2)} |\mathbf{C}_s|^{-\frac{v_C+d+1}{2}} \cdot \\ &\exp\left[-\frac{1}{2}\text{tr}(\Sigma_C \times \mathbf{C}_s^{-1})\right] \quad . \end{aligned} \quad (6)$$

式中: $\boldsymbol{\mu}_\mu$, \mathbf{C}_μ 分别为 $\boldsymbol{\mu}_s$ 的均值列向量和协方差矩阵;

Σ_C , v_C 分别为逆威希特分布的尺度矩阵和自由度; $\Gamma_d(\cdot)$ 表示自由度为 d 的多元伽马分布; $\text{tr}(\cdot)$ 为矩阵的迹。一般称 $(\boldsymbol{\mu}_\mu, \mathbf{C}_\mu)$ 和 (Σ_C, v_C) 为超参数。

对式 (4) 中 $\boldsymbol{\theta}_s$ 的联合后验分布积分可得 $\boldsymbol{\mu}_s$ 和 \mathbf{C}_s 的边缘后验分布分别为

$$f(\boldsymbol{\mu}_s|\mathbf{x}_s) = \int f(\boldsymbol{\mu}_s, \mathbf{C}_s|\mathbf{x}_s) d\mathbf{C}_s \quad , \quad (7)$$

$$f(\mathbf{C}_s|\mathbf{x}_s) = \int f(\boldsymbol{\mu}_s, \mathbf{C}_s|\mathbf{x}_s) d\boldsymbol{\mu}_s \quad . \quad (8)$$

1.3 层次贝叶斯方法

如图 2 所示, 假设共有 r 个场地。其中第 i 个场地的土体参数为多维随机变量 \mathbf{X}_i , \mathbf{X}_i 服从多元正态分布, 记为 $\mathbf{X}_i \sim N(\boldsymbol{\mu}_i, \mathbf{C}_i)$, $i=1, 2, \dots, r$ 。 $\boldsymbol{\mu}_i$ 和 \mathbf{C}_i 分别为 \mathbf{X}_i 的均值列向量和协方差矩阵, 记 $\boldsymbol{\theta}_h = (\boldsymbol{\mu}_h, \mathbf{C}_h)$, 其中 $\boldsymbol{\mu}_h = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_r)$, $\mathbf{C}_h = (\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_r)$, 下标 h 表示 HBM。假设各场地的 $\boldsymbol{\mu}_i$ 和 \mathbf{C}_i 相互独立, 但具有相同的统计特性, 其中 $\boldsymbol{\mu}_i$ 服从多元正态分布 $N(\boldsymbol{\mu}_\mu, \mathbf{C}_\mu)$, \mathbf{C}_i 服从逆威希特分布 $IW(\Sigma_C, v_C)$ 。假定超参数 $\boldsymbol{\mu}_\mu$, \mathbf{C}_μ 和 Σ_C 为随机变量, 超参数 v_C 为常数, 记超参数随机变量 $\boldsymbol{\phi}_h = (\boldsymbol{\mu}_\mu, \mathbf{C}_\mu, \Sigma_C)$ 。

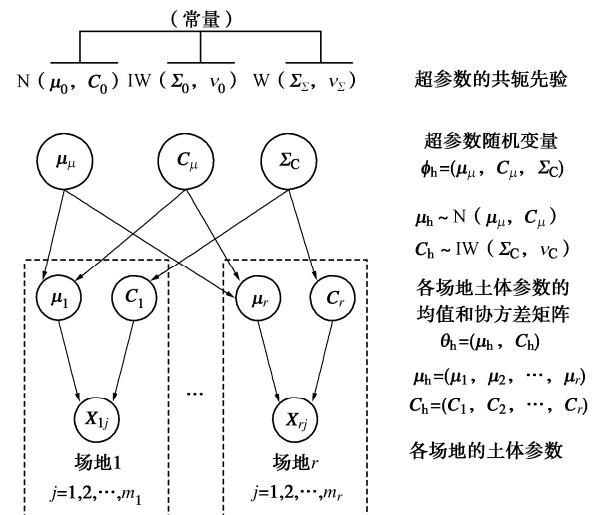


图 2 层次贝叶斯方法示意图

Fig. 2 Diagram of hierarchical Bayesian method

HBM 方法中的模型参数 $\boldsymbol{\theta}$ 既包含 $\boldsymbol{\theta}_h$, 也包含 $\boldsymbol{\phi}_h$ 。由式 (1), 模型参数的联合后验分布可改写为

$$\begin{aligned} f(\boldsymbol{\theta}_h, \boldsymbol{\phi}_h | \mathbf{Data}) &\propto f(\mathbf{Data} | \boldsymbol{\theta}_h, \boldsymbol{\phi}_h) f(\boldsymbol{\theta}_h, \boldsymbol{\phi}_h) \\ &= f(\mathbf{Data} | \boldsymbol{\theta}_h) f(\boldsymbol{\theta}_h, \boldsymbol{\phi}_h) \quad . \end{aligned} \quad (9)$$

式中: $f(\mathbf{Data} | \boldsymbol{\theta}_h)$ 为似然函数; $f(\boldsymbol{\theta}_h, \boldsymbol{\phi}_h)$ 为联合先验分布,

$$f(\boldsymbol{\theta}_h, \boldsymbol{\phi}_h) = f(\boldsymbol{\theta}_h | \boldsymbol{\phi}_h) f(\boldsymbol{\phi}_h) \quad . \quad (10)$$

式中: $f(\boldsymbol{\theta}_h | \boldsymbol{\phi}_h)$ 为 $\boldsymbol{\theta}_h$ 的先验分布; $f(\boldsymbol{\phi}_h)$ 为超参数的先验分布。式 (9) 中第二行等式成立是因为与 \mathbf{Data} 直接相关的参数只有 $\boldsymbol{\theta}_h = (\boldsymbol{\mu}_h, \mathbf{C}_h)$, 而 $\boldsymbol{\phi}_h$ 只能通过 $\boldsymbol{\theta}_h$ 间

接地影响 **Data**^[22]。由此体现了 HBM 方法的层级特性。

当第 i 个场地有 m_i 个测点, 且各测点都有完整的 d 维数据时, 则 **Data** 可表示为 $\mathbf{x}_h = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r)$, 为 $d \times (\sum_{i=1}^r m_i)$ 矩阵, 其中 $\mathbf{x}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{im_i})$, 为 $d \times m_i$ 维矩阵, $i=1, 2, \dots, r$ 。此时, 式(9)中 $f(\mathbf{Data}|\boldsymbol{\theta}_h)$ 对应为

$$\begin{aligned} f(\mathbf{x}_h|\boldsymbol{\mu}_h, \mathbf{C}_h) &= \prod_{i=1}^r f(\mathbf{x}_i|\boldsymbol{\mu}_i, \mathbf{C}_i) \\ &= \prod_{i=1}^r \left[\prod_{j=1}^{m_i} N(x_{ij}|\boldsymbol{\mu}_i, \mathbf{C}_i) \right] \quad (11) \end{aligned}$$

与式(4)的推导类似, 假定 $\boldsymbol{\mu}_h$, \mathbf{C}_h 相互独立, 且超参数 $\boldsymbol{\mu}_\mu$, \mathbf{C}_μ , $\boldsymbol{\Sigma}_c$ 相互独立, 则可得 $\boldsymbol{\theta}_h$, $\boldsymbol{\phi}_h$ 的联合后验分布为

$$\begin{aligned} &f(\boldsymbol{\theta}_h, \boldsymbol{\phi}_h | \mathbf{x}_h) \\ &\propto f(\mathbf{x}_h | \boldsymbol{\theta}_h) \cdot f(\boldsymbol{\theta}_h | \boldsymbol{\phi}_h) \cdot f(\boldsymbol{\phi}_h) \\ &\propto f(\mathbf{x}_h | \boldsymbol{\mu}_h, \mathbf{C}_h) \cdot \\ &\quad [f(\boldsymbol{\mu}_h | \boldsymbol{\mu}_\mu, \mathbf{C}_\mu) f(\mathbf{C}_h | \boldsymbol{\Sigma}_c, v_c)] \cdot \\ &\quad [f(\boldsymbol{\mu}_\mu) f(\mathbf{C}_\mu) f(\boldsymbol{\Sigma}_c)] \quad (12) \end{aligned}$$

式中: $f(\boldsymbol{\mu}_h | \boldsymbol{\mu}_\mu, \mathbf{C}_\mu)$ 为 $\boldsymbol{\mu}_h$ 的先验分布, $\boldsymbol{\mu}_i \sim N(\boldsymbol{\mu}_\mu, \mathbf{C}_\mu)$; $f(\mathbf{C}_h | \boldsymbol{\Sigma}_c, v_c)$ 为 \mathbf{C}_h 的先验分布, $\mathbf{C}_i \sim IW(\boldsymbol{\Sigma}_c, v_c)$; $f(\boldsymbol{\mu}_\mu)$, $f(\mathbf{C}_\mu)$, $f(\boldsymbol{\Sigma}_c)$ 分别为超参数 $\boldsymbol{\mu}_\mu$, \mathbf{C}_μ , $\boldsymbol{\Sigma}_c$ 的先验分布。这里取 $\boldsymbol{\mu}_\mu$ 和 \mathbf{C}_μ 的共轭先验分别为 $N(\boldsymbol{\mu}_0, \mathbf{C}_0)$, $IW(\boldsymbol{\Sigma}_0, v_0)$, $\boldsymbol{\Sigma}_c$ 的共轭先验为威希特分布^[22], 记为 $W(\boldsymbol{\Sigma}_\Sigma, v_\Sigma)$, 则 $\boldsymbol{\Sigma}_c$ 的先验分布为

$$\begin{aligned} f(\boldsymbol{\Sigma}_c) &= W(\boldsymbol{\Sigma}_c | \boldsymbol{\Sigma}_\Sigma, v_\Sigma) \\ &= \frac{|\boldsymbol{\Sigma}_c|^{\frac{v_\Sigma-d-1}{2}}}{\frac{d v_\Sigma}{2} |\boldsymbol{\Sigma}_\Sigma|^{\frac{v_\Sigma}{2}} \Gamma_d\left(\frac{v_\Sigma}{2}\right)} \\ &\exp\left[-\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}_c \times \boldsymbol{\Sigma}_\Sigma^{-1})\right] \quad (13) \end{aligned}$$

式中: $W(\boldsymbol{\Sigma}_c | \boldsymbol{\Sigma}_\Sigma, v_\Sigma)$ 为威希特分布, 其中 $\boldsymbol{\Sigma}_\Sigma$, v_Σ 分别为尺度矩阵和自由度。

需要指出的是, 式(12)所得为模型参数 $\boldsymbol{\theta}_h$, $\boldsymbol{\phi}_h$ 的联合后验分布, 各边缘后验分布可通过类似式(7), (8)的积分得到。

1.4 土体参数的估计

上述两种方法除了可以得到各模型参数的后验分布, 还可对土体参数本身进行估计。

(1) 测点缺失数据的估计

由于条件限制, 在某一测点处不一定能获得土体参数 $\mathbf{X} = (X_1, X_2, \dots, X_d)^T$ 的所有 d 维实测数据^[9, 21]。例如, 对某测点的土样只开展了含水率等基本物性试

验, 但并未开展不排水抗剪强度试验, 则该测点缺失 s_u 的实测数据。此时可利用已知的数据, 对缺失数据的土体参数进行估计。

假设在某测点 d 个土体参数中前 k 个有实测值, 记为 \mathbf{x}_k^o , 对应式(1)中的 **Data**; 后 $d-k$ 个参数无实测数据, 记为 \mathbf{x}_{d-k}^u , 是待估计的参数。下文简化为实测值 \mathbf{x}^o 和缺失项 \mathbf{x}^u 。

基于观测数据得到模型参数 $\boldsymbol{\theta}$ 的后验分布 $f(\boldsymbol{\theta} | \mathbf{Data})$ 后, 通过积分可得 \mathbf{x}^u 的后验分布为

$$\begin{aligned} &f(\mathbf{x}^u | \mathbf{Data}) \\ &= \int f(\mathbf{x}^u | \mathbf{Data}, \boldsymbol{\theta}) f(\boldsymbol{\theta} | \mathbf{Data}) d\boldsymbol{\theta} \\ &= \int f(\mathbf{x}^u | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \mathbf{Data}) d\boldsymbol{\theta} \quad (14) \end{aligned}$$

式中: $f(\boldsymbol{\theta} | \mathbf{Data})$ 由式(4)或(12)得到 $f(\mathbf{x}^u | \boldsymbol{\theta})$ 为 \mathbf{x}^u 的似然函数:

$$\begin{aligned} &f(\mathbf{x}^u | \boldsymbol{\mu}, \mathbf{C}) \\ &= N\left[\mathbf{x}^u | \boldsymbol{\mu}^u + \mathbf{C}^{uo} (\mathbf{C}^o)^{-1} (\mathbf{x}^o - \boldsymbol{\mu}^o), \right. \\ &\quad \left. \mathbf{C}^u - \mathbf{C}^{uo} (\mathbf{C}^o)^{-1} \mathbf{C}^{ou} \right] \quad (15) \end{aligned}$$

式中: $\boldsymbol{\mu}$ 分块记为 $\boldsymbol{\mu} = (\boldsymbol{\mu}^o, \boldsymbol{\mu}^u)^T$, \mathbf{C} 分块记为

$$\mathbf{C}_s = \begin{bmatrix} \mathbf{C}^o & \mathbf{C}^{ou} \\ \mathbf{C}^{uo} & \mathbf{C}^u \end{bmatrix} \quad (16)$$

式中: $\boldsymbol{\mu}^o$ 为 k 维列向量, $\boldsymbol{\mu}^u$ 为 $d-k$ 维列向量; \mathbf{C}^o 为 $k \times k$ 矩阵, \mathbf{C}^{uo} 为 $k \times (d-k)$ 矩阵, \mathbf{C}^{ou} 为 $(d-k) \times k$ 矩阵, $\mathbf{C}^{ou} = (\mathbf{C}^{uo})^T$, \mathbf{C}^u 为 $(d-k) \times (d-k)$ 矩阵。

(2) 未知点处土体参数的估计

工程中有时还需要对未知点位(无任何实测数据的位置)的土体参数 \mathbf{x}^{new} 进行估计。类似式(14), 对模型参数进行积分可得 \mathbf{x}^{new} 的后验分布为

$$f(\mathbf{x}^{\text{new}} | \mathbf{Data}) = \int f(\mathbf{x}^{\text{new}} | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \mathbf{Data}) d\boldsymbol{\theta} \quad (17)$$

式中: $f(\mathbf{x}^{\text{new}} | \boldsymbol{\theta})$ 为似然函数, 即多元正态分布。

式(7), (8), (14), (17)的积分可通过马尔科夫链-蒙特卡洛模拟方法(Markov Chain Monte Carlo, MCMC)进行求解。MCMC 的核心思想是构造一条平稳分布为待估参数后验分布的马尔科夫链, 再利用马尔科夫链达到平稳分布时的样本进行蒙特卡洛积分^[22]。为提高计算效率, 本文采用 MCMC 的吉布斯抽样算法, 具体可见文献[22]。

2 工程应用: 不排水抗剪强度的参数估计

不排水抗剪强度 s_u 是风电大直径单桩设计的重要参数。由于含水率等物理性质试验较强度试验更易开展, 实践中常利用液性指数 L_I 来估计 s_u ^[13]。本文以 $X_1 = LI$ 和 $X_2 = \ln(s_u / \sigma'_v)$ 为正态随机变量, 假定

$\mathbf{X} = (X_1, X_2)^T$ 服从二元正态分布, 利用 LI 对 $\ln(s_u / \sigma'_v)$ 进行估计。

2.1 数据集的建立

从 CLAY/10/7490 数据库^[20, 23]中提取 141 个场地共 1679 组 LI 和 $\ln(s_u / \sigma'_v)$ 数据建立数据集, 各场地的数据量至少 5 组。液性指数 I_L 和竖向有效应力 σ'_v 分别由室内试验和现场孔压静力触探测定, s_u 为不排水抗剪强度修正值, 由各类不排水剪切试验的实测值转换得到, 转换公式详见文献^[23]。原数据库^[20, 23]中部分场地仅有编号, 无具体位置信息, 故本文也仅用编号代表各场地。本文数据集中有 13 个场地含 20 组及以上数据, 数据量较大, 故将这 13 个场地作为目标场地进行参数估计, 分别标记为场地 1~场地 13。如图 3 所示为各场地土体的塑性图, 其中高液限黏土 CH 和低液限黏土 CL 分别占 49.1%, 34.7%, 高液限粉土 MH 和低液限粉土 ML 分别占 13.1%, 2.6%。

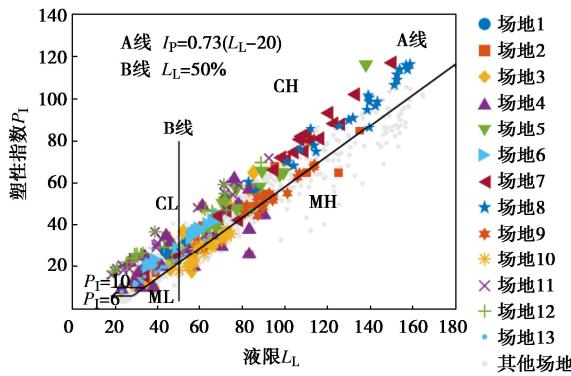


图 3 各场地土体的类型

Fig. 3 Plasticity chart

注: 数据集中少量 LL 或 PI 数据缺失。

经 Shapiro-Wilk 检验^[24], I_L 和 $\ln(s_u / \sigma'_v)$ 的 p 值分别为 1.33 和 4.85, 均大于 0.05, 即在置信度为 95% 的条件下, I_L 和 $\ln(s_u / \sigma'_v)$ 的边缘分布均满足正态性假设。采用表 1 所示 4 种 Copula 函数对数据进行拟合, 得到 Copula 参数 θ_c , 再利用 AIC 指标和 BIC 指标比较各 Copula 函数的拟合优度^[15, 25], 最终确定数据的最优拟合 Copula 函数为 Gaussian Copula 函数。因此, 本文数据满足二元正态分布假设。

表 1 Copula 函数拟合
Table 1 Copula function fitting

Copula	$C(u_1, u_2)$	θ_c	AIC	BIC
Gaussian	$\Phi_2 \left[\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta_c \right]$	-0.05	-3.50	1.93
Clayton	$(u_1^{-\theta_c} + u_2^{-\theta_c} - 1)^{-1/\theta_c}$	1.45×10^{-6}	2.00	7.43
Frank	$-\frac{1}{\theta_c} \ln \left[1 + \frac{(e^{-\theta_c u_1} - 1)(e^{-\theta_c u_2} - 1)}{e^{-\theta_c} - 1} \right]$	-0.07	1.80	7.22
Gumbel	$\exp \left\{ - \left[(-\ln u_1)^{\theta_c} + (-\ln u_2)^{\theta_c} \right]^{1/\theta_c} \right\}$	1.00	2.00	7.43

注: u_1 和 u_2 分别为变量 X_1 和 X_2 的边缘累积分布; Φ_2 为标准二元正态分布的累积分布函数; $\Phi^{-1}(\cdot)$ 为标准正态分布累积分布的逆函数; 加粗值为 AIC 和 BIC 值的最小值, 对应最优拟合 Copula 函数。

各目标场地的实测数据如图 4 所示, $X_1 = I_L$ 的分布区间为 $[-0.44, 3.25]$, $X_2 = \ln(s_u / \sigma'_v)$ 的分布区间为 $[-3.68, 0.91]$ 。场地的均值 \bar{x}_1 、 \bar{x}_2 和协方差 S_{11} 、 $S_{12} (= S_{21})$ 、 S_{22} 汇总于表 2。

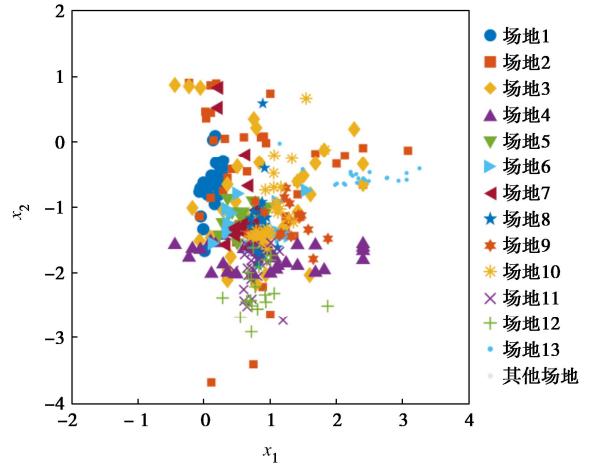


图 4 目标场地实测数据的散点图

Fig. 4 Scatter plots of measured data from target sites

2.2 先验分布的参数确定

先验分布是贝叶斯估计的重要组成部分, 前文已设定各先验均为共轭先验, 本节确定先验分布中的参数。Cao 等^[26]基于文献总结了 LI 均值和标准差的先验范围分别为 0.5~2.5, 0.025~2.2, s_u / σ'_v 均值和标准差的先验范围分别为 0.23~1.4, 0.01~1.26。为降低先验分布对后验分布影响, 本文参考 Cao 等^[26]并采用更为平坦的先验分布: 假定 $X_1 = I_L$ 和 $X_2 = \ln(s_u / \sigma'_v)$ 的均值都为 0, 标准差都为 5。

SBM 方法中需要确定 μ_s 和 C_s 的先验分布 $N(\mu_s, C_s)$, $IW(\Sigma_c, v_c)$ 中的参数取值。根据上述假定和相关研究^[22, 27], 确定如表 3 所示的参数值。HBM 方法中需要确定超参数 μ_μ , C_μ , Σ_c 的先验分布 $N(\mu_0, C_0)$, $IW(\Sigma_0, v_0)$, $W(\Sigma_\mu, v_\mu)$ 中的参数, 以及超参数 v_c 。这里假定 μ_μ 的标准差为 μ_i 标准差的一半, 结合威希特分布期望的性质, 确定先验分布参数的取值, 汇总在表 3。

SBM 方法中需要确定 μ_s 和 C_s 的先验分布

表 2 目标场地的样本统计量
Table 2 Sample statistics of measured data

编号	m_i	\bar{x}_1	\bar{x}_2	S_{11}	S_{12}	S_{22}
1	22	0.131	-0.632	0.015	0.027	0.163
2	39	0.839	-0.666	0.603	-0.021	1.222
3	30	0.861	-0.761	0.637	-0.005	0.806
4	32	0.932	-1.752	0.648	-0.008	0.031
5	20	0.526	-1.176	0.036	0.011	0.028
6	21	0.728	-1.306	0.139	0.015	0.065
7	20	0.598	-1.067	0.032	-0.072	0.464
8	25	0.847	-1.256	0.007	0.006	0.265
9	21	1.236	-1.213	0.093	-0.017	0.066
10	34	1.019	-1.096	0.055	0.091	0.268
11	22	0.742	-1.969	0.037	0.002	0.105
12	20	0.827	-2.253	0.092	0.009	0.120
13	22	2.368	-0.526	0.246	-0.026	0.018

注: m_i 为第 i 个场地的数据量 (测点数量)。

表 3 先验分布的参数取值
Table 3 Parameter values of prior distributions

SBM				HBM						
μ_μ	C_μ	Σ_c	v_c	μ_0	C_0	Σ_0	v_0	Σ_Σ	v_Σ	v_c
$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}$	$\begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}$	4	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 6.25 & 0 \\ 0 & 6.25 \end{pmatrix}$	$\begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}$	4	$\begin{pmatrix} 6.25 & 0 \\ 0 & 6.25 \end{pmatrix}$	4	4

2.3 预测准确度的评价

采用留一交叉验证法 (Leave-One-Out Cross-Validation, LOO-CV) 对两种方法的预测准确度进行比较。该方法通过每次留出目标场地一个测点的 X_2 实测值作为验证集, 利用其他测点的实测数据和该测点的 X_1 值对 X_2 进行后验估计, 循环直到场地内各测点都完成验证为止。

LOO-CV 可结合点预测密度的对数值 (log pointwise predictive density, lppd) [22, 28] 作为评价指标。假设目标场地是第 i 个场地, m_i 为该场地的数据量, 则目标场地 lppd_{loo-cv} 指标的计算如下:

$$\begin{aligned} \text{lppd}_{\text{loo-cv}} &= \sum_{j=1}^{m_i} \lg f_{\text{post}(\setminus j)}(x_{ij,2} | x_{ij,1}) \\ &= \sum_{j=1}^{m_i} \lg \left(\frac{1}{T-t_b} \sum_{t=t_b+1}^T f(x_{ij,2} | x_{ij,1}, \theta^t) \right) \quad (18) \end{aligned}$$

式中: $f_{\text{post}(\setminus j)}(\cdot)$ 为基于训练集 $\mathbf{x}_{\setminus j}$ 得到的后验概率密度函数, 相应的验证集为 2×1 维列向量 $x_{ij} = (x_{ij,1}, x_{ij,2})^T$ 。SBM 和 HBM 两种方法的训练集不同, SBM 方法中 $\mathbf{x}_{\setminus j} = \mathbf{x}_{\setminus j}^s = (x_{i1}, \dots, x_{i,j-1}, x_{i,j+1}, \dots, x_{im_i})$ 为 $2 \times (m_i - 1)$ 维矩阵; HBM 方法中 $\mathbf{x}_{\setminus j} = \mathbf{x}_{\setminus j}^h = (\mathbf{x}_1, \dots, \mathbf{x}_{i,j}, \dots, \mathbf{x}_r)$ 为 $2 \times (\sum_{i=1}^r m_i - 1)$ 维矩阵。其中 r 为场地总数; $\mathbf{x}_k (k = 1, \dots, i-1, i+1, \dots, r)$ 为第 k 个场地所有测点数据的集合, 为 $2 \times m_k$ 维矩阵; $\mathbf{x}_{i,j} = (x_{i1}, \dots, x_{i,j-1}, x_{i,j+1}, \dots, x_{im_i})$ 为第 i 个场地除验证集

x_{ij} 以外测点数据的集合, 为 $2 \times (m_i - 1)$ 维矩阵。

3 结果与讨论

3.1 数据缺失情况下的参数估计

为模拟实际工程中单个场地 LI 数据量大而 $\ln(s_u / \sigma_v')$ 数据量少的情形, 只取本文数据集中的部分数据作为贝叶斯参数估计的 Data, 利用 SBM 方法和 HBM 方法分别对 $\ln(s_u / \sigma_v')$ 缺失数据 x_2^u 、未知点处的 x_2^{new} 和场地 $\ln(s_u / \sigma_v')$ 均值 μ_2 进行后验估计。

以场地 1 为例, 场地 1 共 22 个测点, 即数据集中有 22 组 $\mathbf{X} = (X_1, X_2)^T$ 实测数据。采用 SBM 方法时, 取该场地 22 个 X_1 数据和 n 个 X_2 数据构成 Data, 其中 X_2 数据对应 n 个 X_1 分别 (近似) 为该场地 X_1 数据的 $n+1$ 分位数。利用式 (14), (17), (7) 可分别得到 x_2^u , x_2^{new} 和 μ_2 的后验分布。采用 HBM 方法时, Data 包括 SBM 方法中构造的本场地数据和数据集中其余 140 个场地的全部数据。

首先分析讨论给定数据量, 即给定 n 值时, x_2^u 的后验估计结果。这里以 $n=5$ 为例。如图 5 (a) 所示, 场地 9 共有 21 个测点, 图中黑色圆点表示 Data 中 5 个信息完备的测点, 白色圆点表示 16 个假设 X_2 值缺失的测点 (待估点), 线段表示 x_2^u 的 95% 后验置信区间 (posterior confidence interval, 简称 PCI)。由图 5 (a) 可见, 测点真实值均位于 PCI 中。SBM 方法所得 PCI 的范围 (红线) 大, 甚至超过了图 4 中 X_2 的实际分布区间 $[-3.68, 0.91]$, 表明 SBM 方法无法提供有

效估计。HBM 方法的 PCI 范围 (蓝线) 相比 SBM 的显著缩小, 恰好覆盖真实值的范围, 表明 HBM 方法可有效降低土体参数估计的不确定性。

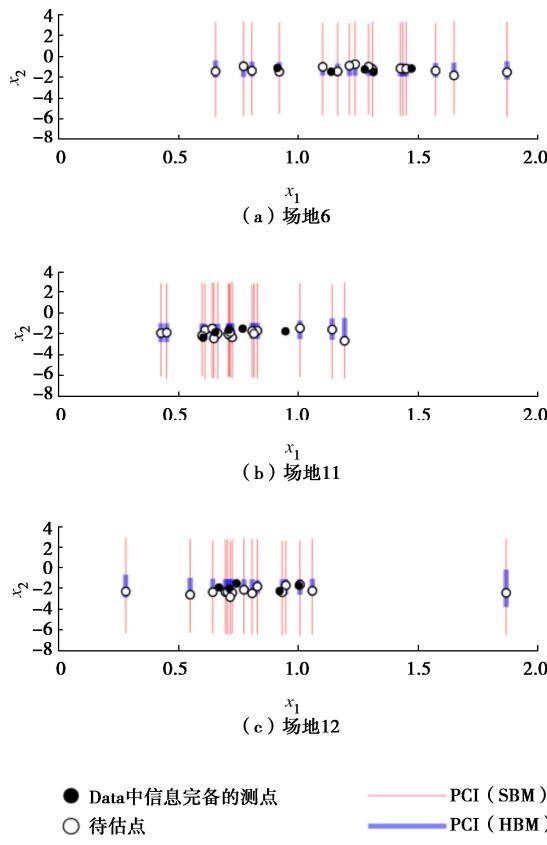


图 5 x_2^u 的 95% 后验置信区间

Fig. 5 95% posterior confidence interval of x_2^u

场地 11 和场地 12 的实测数据和 PCI 分别如图 5 (b), 5 (c) 所示, 与上述结论基本一致, 其余场地也得到了与图 5 类似的结果, 限于篇幅不再展示。各目标场地的估计中, MCMC 抽样次数均为 10000, 由于抽样 2000 次后马尔科夫链即达到平稳, 因此取后 8000 组样本作为后验分布的稳定样本, 下文抽样过程与此一致。

下面讨论数据量 n 对场地未知点处 x_2^{new} 和 μ_2 后验估计的影响, 并对两种方法进行比较。如图 6 所示, x 轴表示 **Data** 中有 n 组 X_2 数据。以场地 9 (图 6 (a)) 为例, 当 $n=0$ 时, SBM 方法的 PCI 范围很大, x_2^{new} 的 PCI 范围为 $[-11.2, 13.4]$, μ_2 的 PCI 范围为 $[-9.1, 11.2]$ 。这是因为 SBM 方法仅利用本场地的信息, 当没有实测数据时, 后验分布完全由先验分布决定。随着 n 的增大, PCI 的范围大幅减小, 表明 SBM 方法受数据量的影响显著。但即使 **Data** 使用了该场地的全部数据 ($n=21$), PCI 的范围仍较大, x_2^{new} 和 μ_2 的不确定性也较大。

对于 HBM 方法, 即使目标场地没有任何数据

($n=0$), x_2^{new} 和 μ_2 的 PCI ($[-2.8, 1.1]$ 和 $[-2.5, 0.9]$) 已经明显小于 SBM 方法的, 说明 HBM 参数估计的不确定性更低, 这是因为 HBM 方法可利用其余 140 个场地的数据降低估计的不确定性。当数据量 n 增大到 5 之后, HBM 所得 PCI 的范围已基本稳定, 表明 HBM 方法的结果受目标场地数据量的影响较小。场地 11 和场地 12 的结果分别如图 6 (b), 6 (c) 所示, 与上述结论一致。

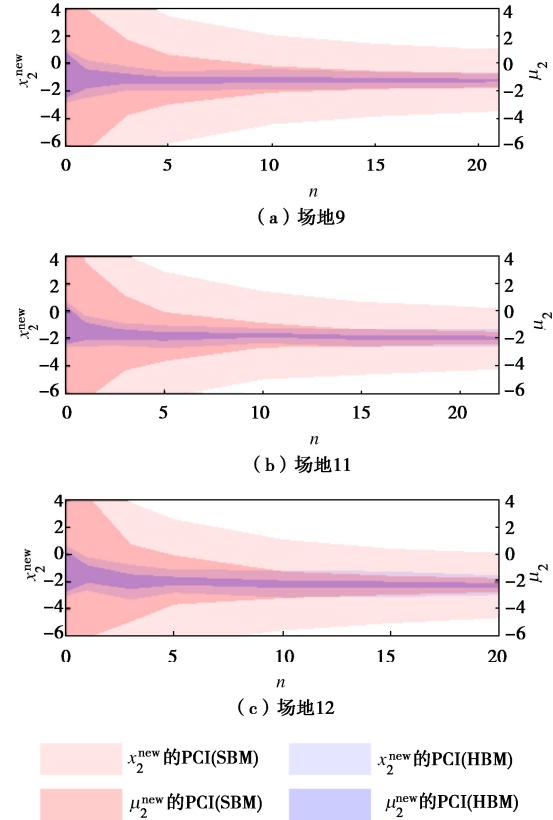


图 6 数据量对参数估计的影响

Fig. 6 Influences of data volume on prediction of parameters

3.2 不同场地后验估计和预测准确度的比较

下面分别利用 13 个目标场地的全部数据进行后验估计, 得到各场地 x_2^{new} 和 μ_2 的 PCI 如图 7 所示。结果表明, SBM 方法仅利用了单个场地的数据, 因此 PCI 的范围较大; 而 HBM 方法可结合多个场地的数据进行估计, 因此 PCI 的范围显著缩小, x_2^{new} 和 μ_2 的不确定性大幅降低。

13 个目标场地的 $\text{lppd}_{\text{loo-cv}}$ 指标如表 4 所示。 $\text{lppd}_{\text{loo-cv}}$ 越大, 表示预测准确度越高。由表 4 可知, 各场地中 HBM 方法对应的 $\text{lppd}_{\text{loo-cv}}$ 基本都大于 SBM 方法的结果。其中场地 2, 3 的数据离散、变异性较大 (表 2), 造成了两种方法的结果差异较小。将全部 13 个场地对应的评价指标求和, 得到 HBM 方法的整体 $\text{lppd}_{\text{loo-cv}}$ 指标为 -76.7, SBM 方法的整体 $\text{lppd}_{\text{loo-cv}}$

指标为-164.6, 表明 HBM 方法的预测准确度更高, 更适于有限数据下土体参数的概率估计。

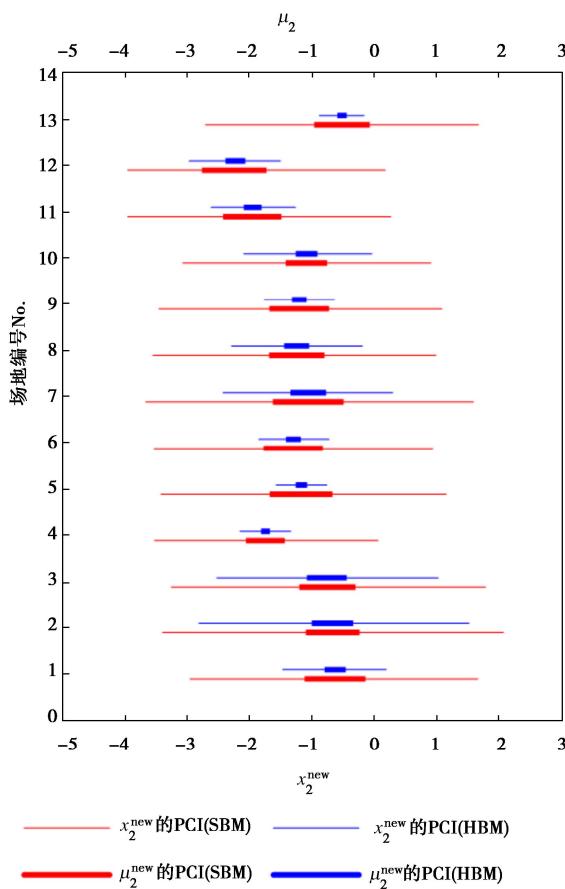


图 7 x_2^{new} 和 μ_2 的 95% 后验置信区间

Fig. 7 95% posterior confidence intervals of x_2^{new} and μ_2

表 4 两种方法的 $\text{lppd}_{\text{loo-cv}}$ 评价指标比较

Table 4 Comparison of $\text{lppd}_{\text{loo-cv}}$ between SBM and HBM

场地编号	$\text{lppd}_{\text{loo-cv}}$ (SBM)	$\text{lppd}_{\text{loo-cv}}$ (HBM)
1	-10.6	-4.3
2	-26.9	-26.9
3	-18.7	-18.2
4	-11.6	3.4
5	-9.1	2.7
6	-9.6	-1.5
7	-11.5	-8.8
8	-12.4	-10.1
9	-9.6	-1.3
10	-15.4	-6.8
11	-10.2	-4.4
12	-9.7	-4.5
13	-9.5	4.2
整体求和	-164.6	-76.7

3.3 HBM 方法统计量后验均值的工程应用

HBM 方法所得 13 个目标场地统计量的后验均值汇总在表 5, 与表 2 对照可知, HBM 方法所得统计量

的后验均值与实测数据的统计量基本吻合。实际应用中, 若需对这些目标场地中任一未知点的强度参数进行取值, 则可直接根据 μ_2 和测点处的 σ'_v 求得不排水抗剪强度。本文重要研究成果之一是利用 HBM 方法得到了超参数 μ_μ , \mathbf{C}_μ , Σ_c 的后验均值, 分别为 $\mu_\mu = (1.040, -0.931)^T$, $\mathbf{C}_\mu = \begin{pmatrix} 0.539 & -0.018 \\ -0.018 & 0.585 \end{pmatrix}$, $\Sigma_c = \begin{pmatrix} 0.088 & -0.008 \\ -0.008 & 0.313 \end{pmatrix}$ 。当需对一个新场地进行设计时, 可直接利用上述结果确定新场地的不排水抗剪强度的设计值或者概率分布。

表 5 HBM 所得场地统计量的后验均值

Table 5 Posterior means of site statistics obtained by HBM

编号	μ_1	μ_2	C_{11}	C_{12}	C_{22}
1	0.132	-0.635	0.018	0.026	0.170
2	0.843	-0.678	0.589	-0.021	1.198
3	0.871	-0.771	0.619	-0.007	0.790
4	0.936	-1.751	0.631	-0.008	0.040
5	0.529	-1.174	0.038	0.010	0.042
6	0.733	-1.304	0.137	0.014	0.077
7	0.598	-1.064	0.034	-0.069	0.457
8	0.847	-1.251	0.010	0.006	0.266
9	1.234	-1.211	0.093	-0.017	0.077
10	1.020	-1.093	0.056	0.087	0.269
11	0.743	-1.959	0.039	0.001	0.115
12	0.829	-2.238	0.092	0.008	0.130
13	2.342	-0.524	0.240	-0.025	0.032

4 结 论

针对场地有限数据条件下黏土不排水抗剪强度不确定性的大的问题, 建立了一个 LI 和 $\ln(s_u / \sigma'_v)$ 双变量参数数据集, 应用特定场地贝叶斯 SBM 方法和层次贝叶斯 HBM 方法分别对 13 个目标场地进行参数估计, 并对两种方法的估计结果进行了对比, 得到 2 点结论。

(1) 在测点缺失数据的情况下, SBM 方法的 PCI 范围大, 而 HBM 方法的 PCI 范围较小, 刚好覆盖真实值的分布范围; 随着数据量的增加, 两种方法得到的 PCI 范围均有减小。SBM 方法受测点数据量的影响更显著, 且参数估计的不确定性较大。HBM 方法受本场地数据量的影响较小, 当数据量大于 5 时, PCI 的范围就基本稳定, 且参数估计的不确定性小。

(2) 分别利用数据集中 13 个目标场地的全部实测数据对 $\ln(s_u / \sigma'_v)$ 进行估计, 结果表明 HBM 方法得到 x_2^{new} 和 μ_2 的 PCI 范围均小于 SBM 方法对应的结果。采用留一交叉验证法对两种方法的预测准确度进行比

较, 得到 HBM 方法的整体 $lppd_{loo-cv}$ 指标为 -76.7, SBM 方法的整体 $lppd_{loo-cv}$ 指标为 -164.6, 表明 HBM 方法的预测准确度更高。

综上, HBM 方法更适用于场地有限数据下土体参数的概率估计。此外, HBM 方法得到的超参数 μ_μ , C_μ 和 Σ_c 的后验均值还可用于新场地的设计和可靠度计算。需要指出的是, 上述结论依赖于本文数据集, 在其他类型土体参数上的适用性有待后续研究。

参考文献:

- [1] 李典庆, 吕天健, 唐小松. 基于多维 Gaussian Copula 的岩土体设计参数概率转换模型构建方法[J]. 岩土工程学报, 2021, 43(9): 1592-1601. (LI Dianqing, LÜ Tianjian, TANG Xiaosong. Establishing probabilistic transformation models for geotechnical design parameters using multivariate Gaussian Copula[J]. Chinese Journal of Geotechnical Engineering, 2021, 43(9): 1592-1601. (in Chinese))
- [2] 张广文, 刘令遥. 确定随机变量概率分布参数的推广 Bayes 法[J]. 岩土工程学报, 1995, 17(3): 91-94. (ZHANG Guangwen, LIU Lingyao. Extended Bayes method for determining probability distribution parameters of random variables[J]. Chinese Journal of Geotechnical Engineering, 1995, 17(3): 91-94. (in Chinese))
- [3] American Petroleum Institute. ANSI/API RECOMMENDED PRACTICE 2GEO Geotechnical and Foundation Design Considerations[M]. Washington: API Publishing Services, 2014.
- [4] LUMB P. The variability of natural soils[J]. Canadian Geotechnical Journal, 1966, 3(2): 74-97.
- [5] LACASSE S, NADIM F. Uncertainties in characterising soil properties[C]//Uncertainty in the Geologic Environment: from Theory to Practice. New York, 1996.
- [6] 宫凤强, 李夕兵, 邓建. 小样本岩土参数概率分布的正态信息扩散法推断[J]. 岩石力学与工程学报, 2006, 25(12): 2559-2564. (GONG Fengqiang, LI Xibing, DENG Jian. Probability distribution of small samples of geotechnical parameters using normal information spread method[J]. Chinese Journal of Rock Mechanics and Engineering, 2006, 25(12): 2559-2564. (in Chinese))
- [7] 骆飞, 罗强, 蒋良淮, 等. 小样本岩土参数的 Bootstrap 估计及边坡稳定分析[J]. 岩石力学与工程学报, 2017, 36(2): 370-379. (LUO Fei, LUO Qiang, JIANG Liangwei, et al. Bootstrap estimation for geotechnical parameters of small samples and slope stability analysis[J]. Chinese Journal of Rock Mechanics and Engineering, 2017, 36(2): 370-379. (in Chinese))
- [8] KULHAWY F H, MAYNE P W. Manual on Estimating Soil Properties for Foundation Design: EL-6800[R]. Palo Alto: Electric Power Research Institute, 1990.
- [9] PHOON K K, KULHAWY F H. Evaluation of geotechnical property variability[J]. Canadian Geotechnical Journal, 1999, 36(4): 625-639.
- [10] MESRI G. Discussion of "New design procedure for stability of soft clays"[J]. Journal of the Geotechnical Engineering Division, 1975, 101(4): 409-412.
- [11] MESRI G. A reevaluation of $S_u(mob) = 0.22\sigma'_p$ using laboratory shear tests[J]. Canadian Geotechnical Journal, 1989, 26(1): 162-164.
- [12] CHANDLER R J. The in-situ measurement of undrained shear strength of clays using the field vane[C]//Vane Shear Strength Testing in Soils: Field and Laboratory Studies (ASTM STP 1014). Baltimore, 1988.
- [13] CAO Z J, WANG Y. Bayesian model comparison and characterization of undrained shear strength[J]. Journal of Geotechnical and Geoenvironmental Engineering, 2014, 140(6): 04014018.
- [14] CHING J, PHOON K K, LI K H, et al. Multivariate probability distribution for some intact rock properties[J]. Canadian Geotechnical Journal, 2019, 56(8): 1080-1097.
- [15] TANG X S, LI D Q, RONG G, et al. Impact of copula selection on geotechnical reliability under incomplete probability information[J]. Computers and Geotechnics, 2013, 49: 264-278.
- [16] 汪海林, 刘航宇, 顾晓强, 等. 基于多元概率分布模型的珠海黏土多参数预测[J]. 岩土工程学报, 2021, 43(增刊2): 193-196. (WANG Hailin, LIU Hangyu, GU Xiaoqiang, et al. Multi-parameter prediction of Zhuhai clay based on multivariate probability distribution model[J]. Chinese Journal of Geotechnical Engineering, 2021, 43(S2): 193-196. (in Chinese))
- [17] BOZORGZADEH N, HARRISON J P, ESCOBAR M D. Hierarchical Bayesian modelling of geotechnical data: application to rock strength[J]. Géotechnique, 2019, 69(12): 1056-1070.
- [18] XIAO S H, ZHANG J, YE J M, et al. Establishing region-specific $N-V_s$ relationships through hierarchical Bayesian modeling[J]. Engineering Geology, 2021, 287:

- 106105.
- [19] CHING J, PHOON K K. Constructing site-specific multivariate probability distribution model using Bayesian machine learning[J]. *Journal of Engineering Mechanics*, 2019, **145**(1): 04018126.
- [20] CHING J, PHOON K K. Correlations among some clay parameters—the multivariate distribution[J]. *Canadian Geotechnical Journal*, 2014, **51**(6): 686-704.
- [21] CHING J, WU S, PHOON K K. Constructing quasi-site-specific multivariate probability distribution using hierarchical Bayesian model[J]. *Journal of Engineering Mechanics*, 2021, **147**(10): 04021069.
- [22] GELMAN A, CARLIN J B, STERN H S, et al. *Bayesian Data Analysis*[M]. 3rd ed. New York: Chapman and Hall/CRC, 2013.
- [23] CHING J, PHOON K K. Transformations and correlations among some clay parameters—the global database[J].
- Canadian Geotechnical Journal, 2014, **51**(6): 663-685.
- [24] WU X Z. Quantifying the non-normality of shear strength of geomaterials[J]. *European Journal of Environmental and Civil Engineering*, 2020, **24**(6): 740-766.
- [25] TANG X S, WANG J P, YANG W, et al. Joint probability modeling for two debris-flow variables: copula approach[J]. *Natural Hazards Review*, 2018, **19**(2) 05018004.
- [26] CAO Z J, WANG Y, LI D Q. Quantification of prior knowledge in geotechnical site characterization[J]. *Engineering Geology*, 2016, **203**: 107-116.
- [27] LUNN D, JACKSON C, BEST N, et al. *The BUGS Book: A Practical Introduction to Bayesian Analysis*[M]. 1st ed. Chapman and Hall/CRC, 2012.
- [28] BOZORGZADEH N, BATHURST R J. Hierarchical Bayesian approaches to statistical modelling of geotechnical data[J]. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 2022, **16**(3): 452-469.

《岩土工程学报》征订启事

《岩土工程学报》创办于 1979 年, 是我国水利、土木、力学、建筑、水电、振动等六个全国性学会联合主办的学术性科技期刊。由南京水利科学研究院承办, 国内外公开发行。主要刊登土力学、岩石力学领域中能代表当今先进理论和实践水平的科学的研究和工程实践成果等。报道新理论、新技术、新仪器、新材料的研究和应用。欢迎国家自然科学基金项目及其他重要项目的研究成果向本刊投稿, 倡导和鼓励有实践经验的作者撰稿, 并优先利用这些稿件。主要栏目有黄文熙讲座、综述、论文、短文、工程实录、焦点论坛、讨论和简讯等。

本刊被《中文核心期刊要目总览》连续多年确认为核心期刊, 并在建筑类核心期刊中位于前列; 本刊被科技部“中国科技论文统计源期刊”(中国科技核心期刊)及“中国科技论文与引文数据库”收录; 本刊被中国科学院“中国科学引文数据库”收录; 本刊被“中国知网”全文收录及“中国学术期刊综合评价数据库”收录; 本刊被“工程索引 Ei Compendex 数据

库”和“日本科学技术振兴机构数据库 JST”等国际数据库收录。本刊网址 (www.cgejournal.com) 全文发布本刊所有刊载文章。

本刊读者对象为土木建筑、水利电力、交通运输、矿山冶金和油气资源等行业中从事岩土力学与岩土工程专业的科研、设计、施工、监理和学习人员。

本刊为月刊, A4 开本, 双栏排版, 220 页, 每月中旬出版, 每期定价 40 元, 全年定价 480 元。

本刊国际标准刊号 ISSN 1000-4548, 国内统一刊号 CN 32-1124/TU, 国内发行代号 28-62, 国外发行代号 MO 0520。欢迎广大读者在全国各地邮局订购, 也可在编辑部订购(不另收邮费)。编辑部订购地址: 南京虎踞关 34 号《岩土工程学报》编辑部; 邮编: 210024; 联系电话: 025-85829534, 85829543, 85829553, 85829556; 传真: 025-85829555; E-mail: ge@nhri.cn。

(本刊编辑部)