

工程勘察测试数据分析处理过程

李 超

(上海岩土工程勘察设计研究院有限公司, 上海 200002)

摘 要 针对岩土的工程特性的不确定性分析原因, 提出解决土的指标不确定性的处理步骤, 论述计算模式不确定性的处理办法, 最后通过具体工程实例, 演示整个分析处理过程, 证明方法的可靠性和科学性。

关键词 工程勘察 样本容量 数据优选 可靠度

1 前言

工程地质问题具有高度的复杂性, 研究深度往往处于定性或半定量范畴。用概率统计理论因为土的变异性而被应用于对工程测试数据的整理、分析与统计中。从勘察提取工程测试数据子样, 到子样预处理后统计参数估计值, 再将量化的统计参数值运用于量化的计算模式中, 这是目前勘察领域中常用的“概率分析方法”。在这种途径中, 从勘察到试验、从参数选择到设计计算、从现场监测到反演分析等各个环节都有“随机”的概念。但目前而言, 现行的一些规程规范或技术标准, 缺乏对试验样本容量和精度要求确定方法的规定, 缺乏关于设计参数可靠性分析的检测规定; 也没能将工程经验和概率统计手法有机的结合起来。本文将就具体工程实例, 尝试将数据优选、置信评估、经验参照纳入概率设计途径之中。

2 勘察工程的不确定性及其解决途径

2.1 土的性质不确定性

土的性质随空间或时间改变, 反映其特性参数指标大多是不完全定量甚至是随机的, 模糊的。这种性质不仅难以人为地控制, 要正确深入地认识它亦非易事。正如耗散结构理论创始人普利高津所言: “那许许多多塑造着自然之形的基本过程本来是不可逆的和随机的, 而那些描述基本相互作用的决定性和可塑性的定律不可能告诉人们自然界的全部真情”。这种不确定性主要来自于如下三方面:

(1) 土的性质天然可变性是固有的。土体不同于其他结构材料有较好的“均质”性, 以往常规设计方法将土体作为“均质”材料看待也是在一定的

假设和范围内的近似。而且土的变异性易受外界环境的影响, 具有时空效应。

(2) 勘测取样方法与试验方法的误差是人为因素影响的体现。现行很多土工试验和原位测试技术由于装置不完善, 技术水平、操作经验等因素、试验或测试方法上的疏忽很容易造成数据失真, 缺乏代表性、真实性、可信性。

(3) 试验数量和勘察数量不足直接影响概率设计所需子样的容量, 勘察工作本身就是以少量资料推断全部情况的手段, 是随机性地层特性的反映与体现。

2.2 土的不确定性的解决途径

基于上述土的不确定性, 实践中不可能找到完全真实反映土层特性的方法。概率统计方法将代替常规的定值设计方法, 包括对数据的整理、计算、分析、判断。从观测或试验数据中寻求事物的规律性。笔者认为工程设计所需的地质参数的确定, 应该综合考虑。包括样本容量优化确定、数据优选预处理、计算可靠度验算三个部分:

(1) 样本容量优化

样本容量和数据本身的离散性决定工程地质参数的精度。毫无疑问, 指标精度越高, 置信概率越大, 但工作量和耗费也越大。从有限的测试样本中统计出来的指标的可靠性如何, 其精度标准将决定

收稿日期: 2004-05-26

第一作者简介: 李超, 男, 1979 年生, 硕士。2003 年毕业于哈尔滨工业大学土木工程学院。主要从事岩土工程勘察、设计、咨询及科研工作。

样本容量的大小。变异法作为参数精度估计方法介绍如下:

根据样本求置信上限或下限,关键在于求给定 α 下的单侧置信区间 ε (即相对偏差率或均值确定),计算式如下:

$$\varepsilon = \frac{\delta \cdot t_{\alpha}}{\sqrt{N-1}} \quad (1)$$

式中, δ —变异系数;

N —样本容量;

t_{α} —置信系数。

显然 ε 越小,精度愈高。则指标可靠度可定义为

$$P_r = 1 - \varepsilon = 1 - \frac{\delta \cdot t_{\alpha}}{\sqrt{N-1}} \quad (2)$$

然后,根据给定 α 值,绘制 $P_r \sim N$ 图。根据图形曲线走势,在曲线平缓处对应的 N 值通常认为是子样基本容量。

(2) 数据优选预处理

剔除样本中异常数据既要避免取伪,又要防止弃真。由于工程勘察一般属于小子样。剔除效果将直接影响均值的可靠性,常用方法^[3]有拉依达(Pauta)法、肖维耐(Chauvenet)法、格拉布斯法、t 检验准则法、狄克逊(Dixon)法,其中 t 检验准则法理论上最为严格^[1],其步骤如下:

① 在 N 个测试值中找出与平均值相比误差最大的值 X_k ,作为可疑值;

② 对不包括可疑值 X_k 在内的 $N-1$ 个值计算平均值 \bar{x} ,标准差 \bar{s} ;

③ 选定风险率 α 值,从 $t(\alpha, n)$ 表中查出 t 检验准则参数 $K(\alpha, n)$;

④ 检查 X_k ,如果有 $|X_k - \bar{x}| > t(\alpha, n) \bar{s}$ 成立,则 X_k 为异常值,并剔除。否则不剔除而终止检查;

⑤ 若 X_k 为异常值,剔除后,对余下 $N-1$ 个值重复①~④步骤,直至终止检查。

(3) 计算可靠度检验

由式(1)可知, α 越大, t_{α} 则越小,则 ε 越小, P_r 越大,那么置信概率 $C(C=1-\alpha)$ 与 P_r 的关系成反比。因此,工程上不能单纯追求过高的可靠度,而确定一个适当的置信水平是重要的。如果按给定的置信概率 C 计算相应的 P_r 值不能达到标准规定,则需要增加样本容量进行补充试验:根据工程经验,如表 1 给出不同设计阶段 C 、 P_r 的目标值。

表 1 不同设计阶段 C 、 P_r 的目标值

Tab. 1 Target values of C and P_r for different design stages

设计阶段 特征指标	可行性研究	初步设计	技施设计
置信概率 C	0.85	0.90	0.95
可靠度 P_r	≥ 0.80	≥ 0.85	≥ 0.90

2.3 计算模式的不确定性及其处理方法

由于土体结构的复杂性,许多工程地质问题的根本机理并不能完全明确,且从实践到创建理论的过程中,大量的随机现象将出现在土工问题中。因而,对一个工程作推断,建立行之有效的数学模型,通常有如下步骤^[2]:1)明确现场条件和环境;2)加以抽象和简化;3)假设和建立基本反映机理的数学模型;4)制定分析方法和明确参数;5)推断描述。所以,计算模式本身具有经验性,加之研究的岩土工程问题具有区域性的特点,因而计算模式又要受到不同区域的限制和影响。

对于计算模式不确定性,其解决无疑是建立更加科学和合理以及较为简易易于推广应用的计算方法,而现在一般的处理方法表现为分地域适用性或者反映在被测试数据子样的变异系数中。

3 工程实例

3.1 工程概况

工程位于上海张江高科技园区青桐路南侧,创新河东侧,本场地 B 型别墅中(2 层框架),有 6 幢处于自南向北贯穿的暗浜区域。施工单位对暗浜进行了换填处理,监理方要求对暗浜区域换填后的地基土采用轻型动力触探试验进行测试,评价其均匀性,并根据 N_{10} 值结合经验估算地基承载力。轻型动力触探试验成果资料如表 2 所示。

表 2 回填土轻型触探测试成果表

Tab. 2 Light dynamic penetration test of backfill

深度 (m)	N_{10} (击)									
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1.0-1.3	18	22	19	34	20	59	22	18	18	#
1.3-1.6	50	65	18	51	22	43	32	10	11	#
1.6-1.9	26	29	19	40	15	18	8	11	11	#
1.9-2.2	38	19	25	15	19	20	11	18	28	30
2.2-2.5	27	31	32	11	18	10	21	22	22	36
2.5-2.8	15	35	29	22	19	12	>100	25	28	31
2.8-3.1	22	14	31	25	22	18		30	21	11

注:表中“#”表示 Q_{10} 起始点自室外排水管道坑底记起(相当于室内标高处以下 1.9m)。

3.2 数据分析处理

(1) 样本容量优化

规定失效概率 $\alpha = 0.10$, 查 t 分布上侧分位数表^[1]则根据式 1、2, 绘制 $Pr-N$ 图, 如图 1 所示。

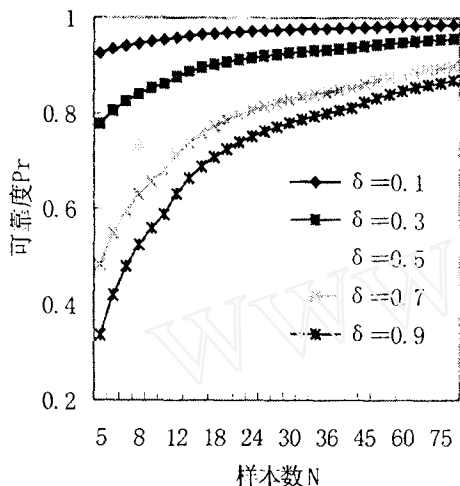


图 1 $Pr-N$ 图

Fig. 1 $Pr-N$ nomogram

由图 1 知, 在规定 α 下, 各个变异系数约在 $N = 45$ 处趋于平缓, 即超过 45 后对提高精度影响不大。本工程如表 2 所示子样数共有 65 个, 足见能够满足测试精度所需要的未进行预处理的最初子样数。若样本优化实际只需要取 45 个数据, 多余的子样数会增加精度, 但效果不甚明显。

(2) 数据优选预处理(t 检验准则法判别)

由于数据离散性较大, 首先经验删除一些失真数据, 收集各地轻型触探资料, 粘性素填土一般 $N_{10} < 30$, 所以首先删除 $N_{10} > 40$ 伪数据。删除后剩下子样数为 61 大于 45, 满足要求。

首先选出最大值 $N_{10\text{怀疑}} = 38$ 作为怀疑值, 由余下的 60 个测试值计算得到:

$$\bar{N}_{10} = 21.02; \quad \bar{s} = 7.17$$

取失效概率 $\alpha = 0.10$, 查 $t(\alpha, n)$ 表得:

$$t(0.1, 61) = 1.282, \text{ 故有 } t(0.1, 61) * \bar{s} = 9.19$$

$$\text{而 } |N_{10\text{怀疑}} - \bar{N}_{10}| = |38 - 21.02| = 16.98.$$

显然有 $|x_k - \bar{x}| > t(\alpha, n) \bar{S}$ 成立, 则删除 $N_{10\text{怀疑}} = 38$. 如此按照①~④步骤重复检查, 期间要注意照顾删除子样最大怀疑值及最小怀疑值。最终得, 当最大 $N_{10\text{怀疑}} = 22$, 最小 $N_{10\text{怀疑}} = 18$ 的时(余下数据中有 8 个 22, 8 个 18), 子样数共为 26。

$$\bar{N}_{10} = 19.81; \quad \bar{s} = 1.72$$

$$t(0.1, 26) \bar{S} = 1.315 \times 1.72 = 2.262 \quad (3)$$

$$|N_{10\text{怀疑}} - \bar{N}_{10}| = |22 - 19.81| = 2.19 < t(0.1, 34) \bar{S} \quad (4)$$

$$|N_{10\text{怀疑}} - \bar{N}_{10}| = |18 - 19.81| = 1.81 < t(0.1, 34) \bar{S} \quad (5)$$

$$\delta = \bar{s} / \bar{N}_{10} = 0.09 \quad (6)$$

实际上根据上海地区经验, $N_{10} > 30$ 都可以首先删除, 则剩下 47 个子样数, 仍然大于 45, 满足要求。

(3) 计算可靠度检验

由步骤得计算结果可知, 当 $n = 34$, $\delta = 0.09$ 得较真值。查图 1 或按式(2)计算可得 $pr = 0.9763$, 对照表 1 可知, 符合技施设计阶段要求, 数据置信度高。

(4) 测试参数在计算模式中的应用

依照国标 GB50021-2001 中 14.2.4 条修正标准值 N_{10} 可得计算统计修正系数 γ_s 为:

$$\gamma_s = 1 - \left\{ \frac{1.704}{\sqrt{n}} + \frac{4.678}{n^2} \right\} \delta = 0.969 \quad (7)$$

$$\text{统计修正标准值 } N_{10}: N_{10} = \gamma_s \cdot \bar{N}_{10} = 19.2 \quad (8)$$

轻型动力触探 $N_{10} - f_0$ 的关系式^[4], 一般是用静载试验或者土工试验资料进行数理统计建立的回归方程式表示的, 具有地区性和经验性的特点, 由于工程属上海地区, 则依照上海规范《岩土工程勘察规范》DGJ08-37-2002 中表 13.3.4 规定计算地基承载力:

$$f_0 = 2.2 N_{10} + 45 = 87.24 (\text{kPa}) \quad (9)$$

当自然地坪标高为 4.2m, 基础埋深为 1.3m, 地下水位为 0.5m 时, 考虑基础埋深对地基承载力的影响, 根据上海规范 DGJ08-37-2002 第 13.3.4-1 条估算回填土的承载力设计值 f_d 为:

$$f_d = 1.1 f_0 + \eta_d \gamma_0 (d - 0.5) = 105 (\text{kPa}) \quad (10)$$

观测子样可知, 子样离散性较大, 说明回填土不均匀, 应辅以后期沉降观测。

4 结语

综上, 针对岩土勘察测试工程, 首选应通过样本容量优化分析确定最低子样数目, 达到工程经济性的目的; 然后运用概率统计方法对子样进行优选预处理, 去除伪值; 接着检查勘察成果的精度水平, 评价设计参数的可靠性; (下转 40 页)

Spatial Data Mining Technology

Song Bei¹ Dai Xiaoyan² Yu Jianhua³

(¹ Shanghai JiaoTong University, Shanghai, 200030

² Lab of Geo - information Science, East China Normal University, Shanghai, 200062

³ ShenYong Construction Contract Co. Ltd. 200433)

Abstract: Recently, with the development of spatial data acquisition, such as remote sensing, network technology, the number and the size of different spatial databases are rapidly growing, knowledge discovery from spatial databases become more and more important, therefore, spatial data mining is developed, it can extract hidden but useful information from large amount of data. In this paper, we review spatial data mining rules, including association rules, characteristic rules, classification rules, clustering rules et al. and different spatial data mining methods, such as neural network, decision trees are discussed, the results indicate that some problems, for example, precision, efficiency, need be improved in the course of spatial data mining. Based on these, we discuss the existed problems and future research objectives in spatial data mining.

Key words: spatial data mining; mining rules; knowledge discovery

(上接 35 页)

最后兼顾计算不确定性影响和地域性的特点,将设计参数应用到适宜的计算模式中解决岩土工程问题。

参考文献

[1] 马良驹,袁灿勤:岩土工程勘察数据统计分析[M],南

京:南京大学出版社,1991

[2] 钱家欢,殷宗泽:土工原理与计算(第2版)[M],北京:中国水利水电出版社,1996

[3] 胡上序,陈德钊:观测数据的分析与处理[M],浙江大学出版社,1996

[4] 周筱滨:动力触探[M],北京:中国铁道出版社,1986

Testing Data Analyzing & Processing Procedure of Engineering Exploration

Li Chao

(Shanghai Geo - technical Engineering Exploration Institute, Shanghai 20002)

Abstract: According to the analysis for the uncertainty of engineering characteristics of rock and soil, the paper has put forward the processing procedure for the uncertainty of soil index. It has stated the processing method for the uncertainty of calculating model and also demonstrated the whole analyzing and processing procedure by a few specific illustrations in order to prove the reliability and science of this method.

Key words: engineering exploration, sample capacity, data optimization, reliability