

DOI: 10.11779/CJGE201907011

# 基于 CPTU 数据融合技术的黏性土应力历史与强度特性评价研究

刘晓燕, 蔡国军\*, 邹海峰, 李学鹏, 刘松玉

(东南大学岩土工程研究所, 江苏 南京 211189)

**摘要:** 超固结比 (OCR) 和不排水抗剪强度 ( $S_u$ ) 是土的基本力学参数, 对土体沉降变形分析和稳定性计算具有重要影响。采用数据融合技术, 结合孔压静力触探 (CPTU) 测试数据, 提出了江苏典型黏性土超固结比和不排水抗剪强度的预测模型。利用特征级数据融合技术 (回归树、模型树) 与决策级数据融合技术 (自举聚合、堆叠泛化) 对预测模型的可行性进行分析。将土的超固结比和不排水抗剪强度的预测值、室内试验所得到的参考值以及 CPTU 传统方法所得到的估计值进行对比分析。结果表明, 模型树预测结果比回归树要好, 决策级融合算法可以提高回归树的预测结果, 但对模型树的预测结果影响较小; 叠加回归树和模型树的预测结果会使其预测的不排水抗剪强度比回归树预测的结果要好, 但比模型树预测的结果要差; 对于几种数据融合模型, OCR 预测值大致相当, 回归树模型在预测 OCR 方面稍优于其他数据融合模型, 数据融合技术能更好地预测土的超固结比和不排水抗剪强度。

**关键词:** 孔压静力触探; 数据融合技术; 超固结比; 不排水抗剪强度; 黏性土

**中图分类号:** TU413      **文献标识码:** A      **文章编号:** 1000-4548(2019)07-1270-09

**作者简介:** 刘晓燕(1990—), 女, 博士研究生, 主要从事土工原位测试与能源岩土工程方面的研究工作。E-mail: happyliuxiaoyan@163.com。

## Prediction of stress history and strength of cohesive soils based on CPTU and data fusion techniques

LIU Xiao-yan, CAI Guo-jun, ZOU Hai-feng, LI Xue-peng, LIU Song-yu

(Institute of Geotechnical Engineering, Southeast University, Nanjing 211189, China)

**Abstract:** The overconsolidation ratio (OCR) and the undrained shear strength ( $S_u$ ) are the basic mechanical parameters of soils, which can influence the deformation analysis and strength calculation of soils. A prediction model for OCR and  $S_u$  of typical clay in Jiangsu Province is proposed by using the data fusion technique and the data of piezocene penetration test (CPTU). The feasibility of the prediction model is analyzed by using the feature-level data fusion techniques (regression tree, model tree) and decision-level data fusion techniques (bagging, stacking). The predicted OCR and  $S_u$ , the reference values obtained by the laboratory tests and the estimated values obtained by the existing calculation methods are compared and analyzed. The results show that the predicted results of the model tree are better than those of the regression tree. The decision and fusion algorithms can improve the predicted results of the regression tree, but they have little influences on the predicted results of the model tree. The superimposed regression tree and model tree can make the predicted  $S_u$  better than that of the regression tree, but worse than that of the model tree. For several data fusion models, the predicted OCR is approximately close. The regression tree model is slightly better than other data fusion models in predicting the OCR. Compared with other prediction methods, the data fusion model can better predict the OCR and  $S_u$ .

**Key words:** CPTU; data fusion technique; OCR; undrained shear strength; clay

## 0 引言

数据融合技术是利用计算机技术, 在相应的计算准则下对获取的来自多个传感器的数据进行分析处理, 并遵循从假设中提取的规则或多个特征领域知识创建特征值, 能够剔除数据处理过程中的异常值, 保

证最终结果的准确性, 从而完成决策和评估的一种信

基金项目: 国家重点研发计划项目课题(2016YFC0800200); 国家自然科学基金项目(41672294, 41877231, 51578146)

收稿日期: 2018-05-04

\*通讯作者(E-mail: focuscai@163.com)

息处理技术。随着计算机技术、通信技术的快速发展,数据融合技术作为一种新兴技术得以迅速发展。早期的数据融合大部分应用在军事防御中,非军事数据融合的应用在20世纪90年代大量增加,目前涵盖的学科范围愈发广泛,主要包括农业生产<sup>[1]</sup>、通信传感网络应用<sup>[2]</sup>以及岩土工程<sup>[3-6]</sup>等多个学科。特别地,边坡工程和隧道工程是岩土工程中使用数据融合技术较多的项目,主要通过将来自多个传感器的信息融合计算,从而达到边坡、隧道稳定性监测和风险预报的目的<sup>[3-6]</sup>。

基于孔压静力触探(CPTU)测试的锥尖阻力 $q_c$ 、侧壁摩阻力 $f_s$ 、孔隙水压力 $u_2$ ,可以获得土体状态与强度等参数,例如超固结比和不排水抗剪强度<sup>[7]</sup>。目前,结合CPTU测试技术或室内试验,相关学者对超固结比和不排水抗剪强度进行了一定的研究。Senneiset等<sup>[8]</sup>提出采用修正的有效锥尖阻力 $q_e$ 对不排水抗剪强度进行计算,可以得到更准确可靠的结果。Lunne等<sup>[9]</sup>建议利用超静孔隙水压力 $\Delta u$ 对不排水抗剪强度进行计算。基于CPTU试验数据,童立元等<sup>[10]</sup>得到了桥梁基础中黏性土的不排水抗剪强度。Li等<sup>[11]</sup>采用CPTU测试技术和室内试验计算了福建地区黏性土的不排水抗剪强度,并得出CPTU测试技术能很好的估算黏性土不排水抗剪强度的结论。Mayne<sup>[12]</sup>已提出了OCR( $OCR = \sigma'_p / \sigma'_{v0}$ ,  $\sigma'_{v0}$ 为竖向有效应力)与CPTU测试结果的关系。Chen等<sup>[13]</sup>基于球形孔穴扩张和临界状态理论提出了预测OCR的方法,该法考虑了初始应力状态、强度各向异性、应变率、刚度指数和摩擦角的影响。刘松玉等<sup>[14]</sup>通过采用不同的预测方法对CPTU三个场地进行分析,提出采用有效锥尖阻力 $q_e$ 能更准确、可靠地确定先期固结压力。

用CPTU测试方法计算得到的黏性土的不排水抗剪强度可信度较高,而超固结比的可信度适中<sup>[15]</sup>。使用CPTU计算方法来推断土工程特性通常不考虑土结构、灵敏度、矿物学老化和地质成因等因素,使得推断结果不准确。数据融合技术能够迅速、准确处理大量数据,得到最可靠的结果。数据融合模型也可以修正样本数据之间复杂的非线性关系,从而克服现有CPTU方法的局限性。但目前并未有学者结合数据融合技术与CPTU测试技术来评价土体性质、估算土体的超固结比和不排水抗剪强度。

本文探讨了两种基于树的数据融合技术(特征级融合算法和决策级融合算法)利用CPTU试验数据来预测黏性土的超固结比和不排水抗剪强度值的可靠性。为了改善数据融合模型的性能,结合数学方程,将CPTU测试结果的原始特征集创建附加特征值。将所有采用数据融合模型获得的土体超固结比和不排水抗剪强度的预测值均与室内试验获得的参考值进行比

较,以评估每个数据融合算法的可靠度。并将超固结比和不排水抗剪强度的预测值与现有的计算方法获得的估计值进行比较,从而进一步确定是否可以通过使用数据融合技术来提高CPTU测试推断土体性质的准确性。

## 1 数据融合技术

数据融合技术指的是利用计算机技术对按时序获取的若干传感器的观测信息在一定准则下加以自动分析、综合以完成需要的决策和估计任务而进行的信息处理的过程。

### 1.1 数据融合结构

身份数据融合是参数数据的融合,以确定观察对象的身份和属性。身份数据融合的体系结构通常分为两种方法:特征级融合和决策级融合<sup>[16]</sup>。

(1) 在特征级融合方法中,提取原始数据的特征值,以从每个传感器获得特征向量。然后将从所有传感器提取的特征向量连接在一起成为单个联合特征向量。使用特征级融合算法对单个联合特征向量进行身份声明,这些单独的身份声明随后通过使用决策级融合技术融合成一个身份声明。如图1(a)所示。

(2) 在决策级融合方法中,每个传感器仅基于自己的单个原数据进行独立的身份声明。随后,使用决策级融合技术将这些身份声明融合到一个单独的身份声明中。如图1(b)所示。

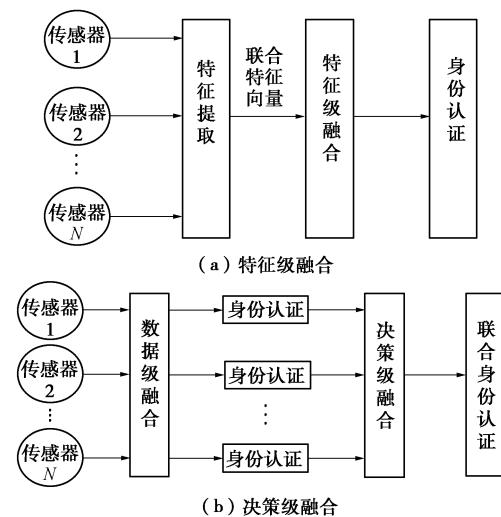


图1 数据融合架构图

Fig. 1 Architecture diagram of data fusion

### 1.2 特征级数据融合技术

回归树通过输入特征向量和输出定义之间的非线性变换进行特征级融合,其中输出的变量是数值。学习或泛化可以通过自适应过程来完成,回归树是由输入值和相应的输出或目标值组成的训练案例。如果数

据融合模型得到适当的训练，那么它能够根据输入变量与输出变量相关性建立相应的函数模型，并且可以通过测试对一组给定的未知的输入模式进行预测，此时的输出值也是未知的。

基于树的算法是一种特征级融合技术，能够用易于理解的流程图描述变量之间复杂的相互作用或依赖关系。基于树的方法可以处理大量的输入变量，并为最终模型选择最优变量。回归树能够在训练和测试期间处理丢失的数据。但是个别异常值容易对处理结果产生干扰，因此回归树分析应该基于大样本进行，这样才能保证各节点内有足够的样本量从而进行准确分类。模型树对异常值的影响非常稳健。由于不会使用异常值分裂节点，因此在输入变量中出现的异常值不影响树的生长过程<sup>[4]</sup>。

### 1.3 决策级数据融合技术

决策级融合技术，能够提高特征级融合算法的预测结果。自举聚合融合了相同算法生成的多个数据融合模型的结果。堆叠泛化融合了由不同算法构建的多个数据融合模型的结果。

由于训练集的微小变化会使生成的树存在很大的差异，因此选用标准的自顶向下的树归纳过程本质上是不稳定的。自举聚合通过改变原始训练集的形式反复诱导树来减轻这种不稳定性，并通过对训练数据的替换进行重采样，将其中一部分原始训练案例省略，复制其他的案例创建新的、相同大小的训练集。当所有形式的训练集的树建立之后，取不同树预测值的平均值，将结果简单地融合。堆叠泛化本质上试图了解哪些基础数据融合模型最可靠，以便在最终融合预测中对这些预测值进行更大程度的加权。

## 2 CPTU 数据融合技术

CPTU是一个多传感器装置，由圆锥形尖端和圆柱形摩擦套筒组成，当套管推入地面时，贯入阻力在摩擦套筒上产生摩擦阻力（图2）。通过压力传感器，确定孔压静力触探仪的贯入阻力。除此之外，CPTU还包含一个或多个孔压过滤环，压力传感器以及连接元件和传感器的流体通道，以估算孔隙水压力。当进行CPTU试验时，用计算机提取锥尖阻力、侧壁摩阻力和孔隙水压力的特征值（图2）。其中，孔隙水压力是根据锥体底部 $u_2$ 位置处的传感器获得的数据确定的。

将提取的特征值连同其他相关信息，合并成一个联合特征向量，并使用特征级融合算法做相应的定义。如图2所示，修正的锥尖阻力 $q_t$ 、在 $u_2$ 位置处引起的孔隙水压力以及垂直总应力和静水压力作为联合特征向量输入到回归树算法中。随后，数据融合算法则会给出相应的超固结比和不排水抗剪强度值。

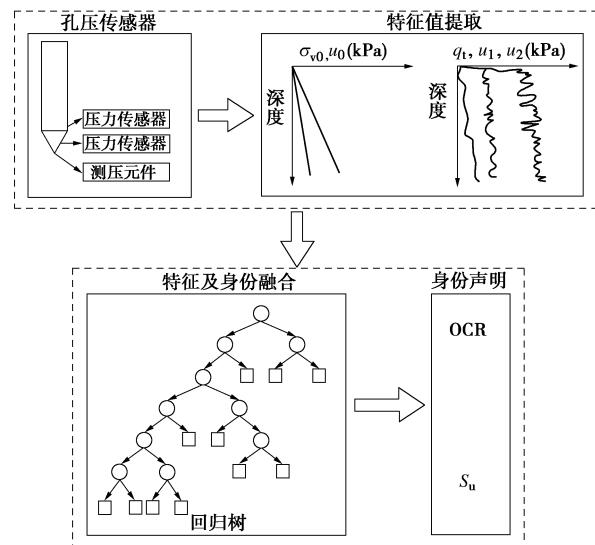


图2 基于CPTU的数据融合技术

Fig. 2 Data fusion technique based on CPTU

回归树是一个逻辑模型，可以根据输入变量的相应值预测目标或输出变量的值。回归树用于预测估计问题中的数值变量。树中的每个节点（图2中的圆）代表来自原始训练集的一个子集，最上面的节点代表所有的训练案例，称为根节点。有子节点的节点称为内部节点；没有子节点的节点称为外部或叶节点（图2中的矩形）。每个节点都有与它相关的数值，该数值等于到达节点的训练案例的目标值的平均值，在回归树中叶节点的数值表示模型的最终预测值。

回归树和模型树使用基本的分而治之的方法来实现节点的分裂，即将一个复杂的问题递归地分解成两个或更多更容易的问题，直到问题可以直接解决掉。分裂基于分裂变量的输入变量的值来执行，为分裂变量的不同值创建不同分支，且根节点中的每个训练案例都被传送到下面其中一个分支，根据分支变量的值，将其传送到相应的子节点中，然后对每个子节点重复处理，只考虑该节点中的情况。当一个节点不能被分割时，它被称为叶节点；当只保留叶节点时，用于构建树的递归分区终止。大多数树归纳算法构建完全树，然后修剪它，以防止训练数据过度拟合。事实上，许多在训练数据中反映噪声或异常值的分支都是从最终模型中删除的<sup>[15]</sup>，这样能得到一个更加紧凑、复杂度更低的模型树。

## 3 数据融合技术特征创建与测试

### 3.1 数据库

根据蔡国军等<sup>[17-18]</sup>和邹海峰等<sup>[19]</sup>获得的数据库（共有240例），包含软、正常固结黏性土到坚硬的超固结土。包含CPTU的实测值和相应的超固结比和不排水抗剪强度计算值，这些值取自于江苏省的7个黏土



表1 试验场地的孔压静力触探试验数据总结表

Table 1 Summary of CPTU data for test sites

试验场地	深度/m	$q_t/\text{MPa}$	$f_s/\text{kPa}$	$u_2/\text{kPa}$	测试案例编号
南京	3.0	220.9~246.7	18.2~21.2	45.9~55.4	1~10
	6.0	210.7~230.6	6.8~8.2	105.6~124.3	
	17.0	1008.9~1230.6	22.3~26.8	389.4~410.2	
连云港	1.2	653.2~687.4	21.9~25.6	108.9~127.3	11~15
	3.0	146.6~178.1	17.6~20.4	198.2~210.7	
	9.0	249.3~267.8	16.4~19.9	243.8~289.7	
盐城	3.0	98.2~118.7	21.3~24.5	89.7~100.2	16~20
	8.0	286.3~314.4	23.7~25.1	231.7~254.3	
	12.0	1187.6~1321.7	10.1~13.1	387.6~401.3	
苏州	4.0	218.6~239.1	7.5~9.6	112.9~127.6	21~25
	9.0	379.1~401.2	9.5~10.7	243.1~264.8	
	12.0	486.3~505.7	8.1~10.9	389.7~408.4	
常州	3.0	498.6~728.1	52.3~84.3	79.2~92.7	26~30
	7.0	1127.3~1512.7	16.3~27.1	122.3~145.1	
	13.0	1814.3~2011.3	45.1~70.2	172.1~207.5	
泰州	6.0	672.1~987.1	9.1~12.7	87.6~139.4	31~35
	12.0	1349.7~1785.1	15.6~20.7	172.9~200.1	
	18.0	1932.1~2276.5	21.3~27.6	210.8~264.2	
扬州	8.0	335.7~357.9	2.1~9.4	563.4~672.1	35~40
	14.5	1763.1~1983.7	40.9~51.3	721.3~795.6	
	17.5	1498.2~1787.9	51.2~64.8	501.3~589.6	

场地: 南京<sup>[17~19]</sup>、连云港<sup>[17~19]</sup>、盐城<sup>[17~19]</sup>、苏州<sup>[17~19]</sup>、常州<sup>[17~19]</sup>、泰州<sup>[17~19]</sup>和扬州<sup>[17~19]</sup>。 $q_t$ 、 $f_s$ 根据现场CPTU测得(见表1); OCR由一维固结仪测试确定; 不排水抗剪强度值由各向同性固结不排水三轴压缩(CIUC)测试确定(见表2)。

数据库中的噪声源(即由实验室和现场测试过程中出现的错误或异常值)可能导致试验值和现场CPTU预测的土体参数存在显著差异。虽然CPTU试验地点和取样位置相邻,但在某一测试深度记录的CPTU数据与从附近钻孔位置的相同深度采集的样品的岩土参数之间可能存在差异。换句话说,从钻孔中获取的土体样本可能与附近探孔的CPTU数据不一致,固有的地质变化或地点的差异也可能导致CPTU预测数据不准确或不一致。正如蔡国军等<sup>[17~18]</sup>和邹海峰等<sup>[19]</sup>获得的不排水抗剪强度,测量的是邻近探孔中的锥尖阻力和孔隙压力,而不是同一探孔。

### 3.2 特征创建

为提高数据融合模型的预测精度,可以从原始的特征集中创建附加特征值。尽管这些附加特征值不会改变原有模型,但是所创建的特征值比现有特征值更接近于目标变量,因此可能对模型更有利<sup>[20]</sup>。需要注意的是,使用额外创建的特征既可以帮助也能阻碍基于树的数据融合算法。有许多特性使回归树和模型树具有更强的辨别力,因为基于树的方法在理论上只对最合适的特性进行划分,特别是在根节点上,而不是不相关的特性上。然而,在实践中,将无用的特性添加到数据集会使基于树的算法产生混淆,而且,由于

树归纳算法对可接受的解决方案执行无止境搜索,因此,它不会回顾以前的拆分选择,也就是说最终的树模型可能不会代表最优解。

采用正确的数据融合方法,从CPTU和现场应力数据中创建6个额外的特征值或输入变量,以训练数据融合模型。自20世纪70年代早期第一次推出CPTU以来,已经提出了许多涉及CPTU测试值的关系或特征。本文采用的附加特征包括:垂向有效应力 $\sigma'_{v0} = \sigma_{v0} - u_0$ ;超孔隙压力 $\Delta u_2 = u_2 - u_0$ ;净锥尖阻力 $q_n = q_t - \sigma_{v0}$ ;净压力 $Q_t = (q_t - \sigma_{v0}) / \sigma'_{v0}$ ;有效锥尖阻力 $q_e = q_t - u_2$ ;孔压比 $B_q = \Delta u_2 / (q_t - \sigma_{v0})$ 。

### 3.3 数据融合模型的训练与测试

利用两个特征级融合技术、回归树和模型树开发了多个特征级数据融合模型,将特征级融合技术和决策级融合技术联合到一起,开发了混合数据融合模型。训练的数据融合模型是使用Weka机器学习软件包生成的。测试集由120个案例组成,训练集则由剩下的120个案例组成。按照标准的测试程序,测试用例是从每个CPTU现场随机挑选的且在训练过程中从未使用过。

从CPTU传感器数据和原位应力(输入值)中提取/创建的特征值以及超固结比和不排水抗剪强度的相应目标值,用于开发两种特征级融合模型:回归树和模型树。使用Weka的REPTree算法来诱导回归树,而使用Weka的M5'算法诱导模型树。由于基于树的算法只能接受一个目标变量,所以必须为每个目标变量引入单独的树模型。即两个独立的回归树和两个单独

表 2 试验场地黏性土物理力学指标

Table 2 Physical and mechanical indices of cohesive soils for test sites

试验场地	深度/m	$I_p$	w/%	$\gamma /(\text{kN}\cdot\text{m}^{-3})$	$S_t$	OCR	$S_u/\text{kPa}$
南京1	2.6~17.1	14.2~16.1	38.1~44.6	16.1~17.3	2.1~20.5	0.6~1.1	30.1~42.8
南京2	2.3~17.5	15.3~16.5	35.9.1~48.8	16.3~17.5	3.8~17.5	1.3~2.5	33.1~45.3
南京3	3.0~17.2	14.9~17.2	37.7~44.5	15.8~18.1	1.2~8.9	0.7~1.4	34.6~41.8
南京4	2.8~17.9	12.3~18.9	38.6~45.7	14.9~17.6	1.0~7.6	1.2~2.6	37.5~45.3
南京5	2.6~17.0	13.7~17.0	36.1~49.3	16.3~18.4	2.3~6.9	1.4~3.2	33.5~46.1
南京6	2.4~17.6	12.7~19.1	39.0~44.6	16.9~18.1	1.7~10.4	0.9~1.6	36.2~46.3
南京7	2.7~17.3	14.3~17.6	36.8~47.1	17.1~17.9	1.3~8.6	0.8~2.3	33.9~46.8
南京8	2.8~17.4	14.5~17.3	30.5~50.3	16.8~17.9	1.7~9.1	1.1~3.1	35.7~45.8
南京9	2.7~17.0	14.3~16.9	38.6~44.1	17.1~18.2	1.8~9.4	0.7~1.6	35.1~45.7
南京10	3.0~17.2	11.9~18.3	33.4~46.7	16.3~17.8	1.4~6.4	1.6~2.6	33.9~46.2
连云港1	0.5~10.1	20.1~27.3	70.1~74.3	15.2~16.8	9.6~15.7	0.8~1.7	16.0~36.5
连云港2	0.5~9.5	19.8~26.1	69.8~75.1	11.6~16.8	10.4~15.3	1.2~3.6	23.7~45.8
连云港3	1.0~9.5	19.8~27.6	72.6~75.6	10.9~15.3	98.6~17.2	0.9~2.3	18.9~50.2
连云港4	1.0~10.5	20.7~26.5	74.2~76.9	12.4~17.1	11.3~16.4	1.3~5.6	23.8~45.7
连云港5	1.0~9.5	18.6~25.3	70.6~76.5	11.6~15.9	10.7~15.4	0.9~4.3	33.1~52.9
盐城1	1.5~12.5	18.9~23.4	46.1~53.2	14.5~16.2	1.8~5.9	0.9~1.9	21.8~32.7
盐城2	1.0~12.5	17.6~24.1	48.3~54.1	15.1~18.2	2.3~7.6	1.3~3.2	18.9~40.3
盐城3	2.0~13.0	19.3~25.7	45.9~57.6	12.3~18.4	1.0~13.8	0.8~2.9	23.8~56.8
盐城4	2.0~12.5	18.3~26.1	44.1~55.8	14.9~16.7	2.8~17.9	1.6~4.3	19.2~44.5
盐城5	3.0~12.5	19.7~26.4	43.2~56.8	13.9~16.8	2.3~6.9	2.4~5.6	28.7~51.3
苏州1	2.0~12.0	12.5~18.6	52.3~64.1	12.7~16.2	1.5~24.1	1.2~3.1	27.3~33.1
苏州2	2.5~12.5	10.3~14.7	50.6~63.4	13.4~17.3	1.3~3.6	1.1~2.3	21.6~35.4
苏州3	3.0~13.0	13.5~20.1	49.8~55.7	14.2~17.8	2.3~11.7	0.8~1.7	19.8~40.1
苏州4	1.0~13.0	17.2~23.4	46.7~67.4	12.6~16.5	1.8~22.1	1.6~4.1	22.5~36.8
苏州5	1.5~13.5	16.8~21.3	44.8~68.1	15.1~17.2	3.2~15.6	1.3~2.9	30.1~36.1
常州1	2.0~13.5	12.2~16.3	19.4~25.8	16.1~19.3	2.3~4.7	1.0~3.2	22.1~43.1
常州2	1.5~13.0	10.6~17.8	18.6~27.8	15.8~18.9	1.3~3.6	0.9~1.7	31.7~54.2
常州3	2.5~14.0	11.3~18.1	18.1~27.4	16.1~18.7	2.1~5.1	1.1~3.5	30.8~56.1
常州4	1.0~13.5	13.5~17.5	20.1~26.5	15.6~19.1	2.6~4.4	2.1~4.5	28.9~36.9
常州5	3.0~13.5	12.7~18.9	18.2~28.1	16.7~19.3	1.7~3.9	1.2~3.3	40.1~56.8
泰州1	1.0~18.0	6.1~10.3	17.5~23.6	14.3~18.1	1.9~6.6	0.7~1.6	49.8~56.7
泰州2	2.0~19.0	5.6~8.6	20.4~36.7	15.7~18.3	1.1~19.3	0.9~2.1	52.3~62.3
泰州3	2.5~19.0	4.3~10.4	19.5~33.6	15.4~19.1	2.5~9.4	1.3~5.6	48.6~66.1
泰州4	3.0~18.0	6.1~8.9	30.1~40.5	14.9~19.7	3.1~5.6	3.4~10.6	50.3~75.4
泰州5	3.5~18.5	5.3~9.8	20.6~35.7	16.3~18.7	1.2~23.2	2.3~6.8	62.1~76.8
扬州1	1.0~18.0	14.2~16.8	59.8~76.1	17.5~18.2	1.2~4.3	1.1~2.3	27.4~51.2
扬州2	1.5~19.0	14.6~17.3	56.2~69.8	15.1~19.3	1.5~3.3	1.4~3.2	25.6~48.1
扬州3	1.0~18.5	15.1~18.4	60.1~79.1	15.6~18.8	2.3~4.5	2.6~4.9	30.2~55.3
扬州4	0.5~18.5	16.4~18.2	55.3~77.4	17.2~18.5	3.2~10.9	3.7~7.1	40.1~56.9
扬州5	2.0~19.5	13.3~17.9	50.7~74.2	17.1~19.1	1.7~17.3	4.2~10.9	35.1~52.7

的模型树, 用来预测超固结比和不排水抗剪强度。表3是用于训练的数据融合模型特征向量的汇总表, 其中包括CPTU试验参数和额外创建的特征值。测定不排水抗

剪强度的试验方法(三轴压缩试验), 用  $s_{u(TC)}$  表示。

决策级融合算法(自举聚合和堆叠泛化)可以提高任意不稳定树的稳定性或提高数据融合模型的性能

(即预测精度)。在自举聚合中, 针对所有版本的重采样训练集合生成10个个体回归树或模型树, 并通过对每个单独树的数值预测进行平均来融合测试结果。在堆叠泛化过程中, 利用线性回归技术将一回归树和一模型树的结果进行融合。换句话说, 堆叠泛化后的输出是一个线性回归方程, 它结合了回归树和模型树的预测, 其中每个预测通常根据每个基本数据融合模型的精度或可靠性而不同地加权。

表3 数据融合模型的特征值

Table 3 Eigenvalues of data fusion model

数据融合模型	特征向量
OCR	$\sigma_{v0}$ , $\sigma'_{v0}$ , $q_t$ , $u_2$ , $\Delta u_2$ , $q_n$ , $q_e$ , $Q_t$ , $B_q$
$S_u$	$\sigma_{v0}$ , $\sigma'_{v0}$ , $q_t$ , $u_2$ , $\Delta u_2$ , $q_n$ , $q_e$ , $Q_t$ , $B_q$ , $S_u$ (三轴压缩试验)

图3~6是预测超固结比和不排水抗剪强度的回归树模型和模型树的数据融合模型。如前所述, 根和内部节点用圆表示, 叶节点用正方形表示, 分支上的值是相应父节点中分裂变量的值。将数值分配给回归树中的叶节点, 并将线性回归方程(或线性模型)分配给模型树中的叶节点。回归树预测目标变量(超固结比和不排水抗剪强度)的离散值, 而模型树预测目标变量(超固结比和不排水抗剪强度)的连续值。在自举聚合中分别输出10种不同的回归树的平均值和10种不同的模型树的平均值。堆叠泛化一个回归树(REPTree)和一个模型树(M5型)的结果, 作为目标变量的一个线性回归模型。本文中, 通过堆叠泛化的方法将回归树和模型树融合后的数据融合模型结果如下:

$$\text{OCR} = (0.3015 \cdot \text{回归树}) + (0.637 \cdot \text{模型树}(M5\text{型})) + 0.1713,$$

$$S_u = (-0.1901 \cdot \text{回归树}) + (1.3027 \cdot \text{模型树}(M5\text{型})) - 2.01864.$$

其中, LM1:  $\text{OCR} = -0.0045 \cdot \sigma_{v0} + 0.0004 \cdot q_t + 0.0013 \cdot u_2 + 0.0041 \cdot \sigma'_{v0} + 0.0226 \cdot Q_t - 0.0865 \cdot B_q - 0.0689 \cdot \text{NEPP} - 0.0009 \cdot q_n + 0.0006 \cdot q_e - 0.0251$ ; LM2:  $\text{OCR} = -0.0061 \cdot \sigma_{v0} + 0.0004 \cdot q_t + 0.0013 \cdot u_2 + 0.0038 \cdot \sigma'_{v0} + 0.0226 \cdot Q_t + 0.0141 \cdot B_q - 0.0891 \cdot \text{NEPP} - 0.0007 \cdot q_n + 0.0006 \cdot q_e - 0.256$ ; LM3:  $\text{OCR} = -0.0059 \cdot \sigma_{v0} + 0.0014 \cdot q_t + 0.002 \cdot u_2 + 0.0065 \cdot \sigma'_{v0} + 0.011 \cdot Q_t + 0.00749 \cdot B_q - 0.143 \cdot \text{NEPP} + 0.0006 \cdot q_e - 1.3789$ 。

其中, LM1:  $S_u = 0.0332 \cdot \sigma_{v0} + 0.0361 \cdot q_t - 1.7512 \cdot Q_t - 38.3035 \cdot B_q + 5.0011 \cdot \text{NEPP} + 0.015 \cdot q_e + 9.0501$ ;

LM2:  $S_u = 0.0601 \cdot \sigma_{v0} - 1.431 \cdot Q_t + 67.5811 \cdot B_q + 5.801 \cdot \text{NEPP} + 0.0758 \cdot q_n + 0.027 \cdot q_e - 70.771$ 。

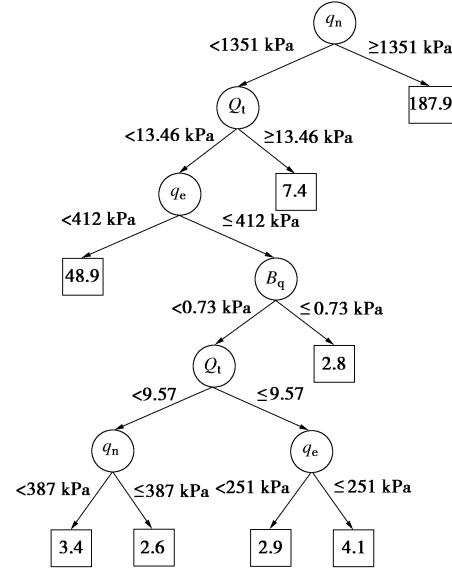


图3 回归树预测OCR

Fig. 3 Prediction of OCR of regression tree

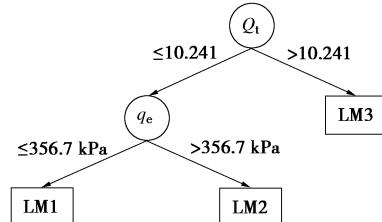


图4 模型树预测OCR

Fig. 4 Prediction of OCR of model tree

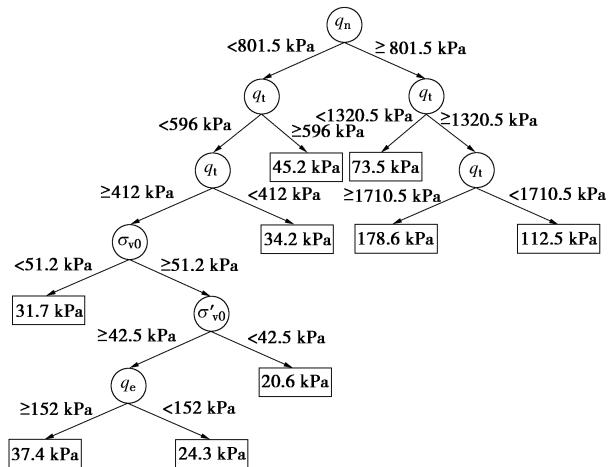


图5 回归树预测不排水抗剪强度

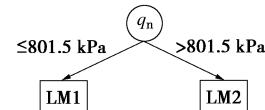
Fig. 5 Prediction of  $S_u$  of regression tree

图6 模型树预测不排水抗剪强度

Fig. 6 Prediction of  $S_u$  of mode tree

## 4 预测结果

测试之后, 将数据融合模型预测的超固结比和不排水抗剪强度值与参考的超固结比和不排水抗剪强度

值进行比较, 如图7、8所示。如前所述, 参考的OCR值由室内固结试验确定, 参考的不排水抗剪强度值由实验室三轴压缩试验确定。在每个图中, 将所有的数值曲线的相关系数( $R^2$ )、平均绝对误差(MAE), 和相对误差绝对值(RAE)作为预测值的一种误差度量。

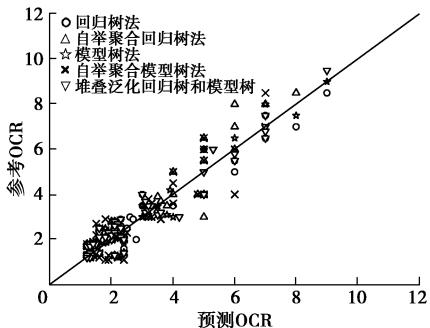


图 7 各种树法的超固结比预测值和室内试验测试值关系图  
Fig. 7 Relationship between OCRs by various tree models and laboratory test reference values

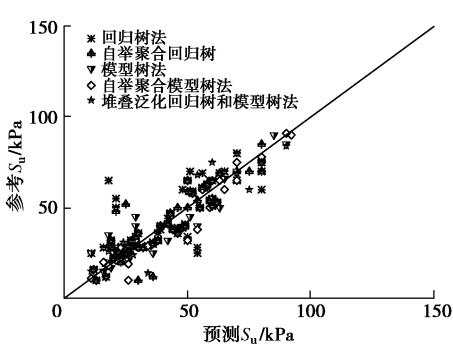


图 8 各种树法的不排水抗剪强度预测值和室内试验测试值关系图  
Fig. 8 Relationship between  $S_u$  by various tree models and laboratory test reference values

一般来说, 模型树的预测结果比回归树的预测结果更可靠, 决策级融合算法能够提高回归树和模型树的结果。基于预测值的计算性能, 不排水抗剪强度值比超固结比预测更好。如图7和表4所示, 对于所有使用的方法, OCR预测大致相当。自举聚合回归树法预测的相关系数 $R^2$ 相对较大, 自回归法和自举聚合回归法预测的平均绝对误差MAE和相对误差绝对值RAE相对较小。利用自举聚合回归树法预测的OCR相对较准确。在预测不排水抗剪强度(图8)中, 模型树比回归树表现更好。通过表5可以看出在预测不排水抗剪强度的各种预测方法中, 自举聚合模型树法的 $R^2$ 最大, MAE和RAE最小。堆叠泛化回归树和模型树法预测的要优于回归树法和自举聚合回归树法, 但差于自举聚合模型树法。

使用现有的几种计算方法(表6)估计超固结比和不排水抗剪强度(所使用的数据仍然是蔡国军等<sup>[17-18]</sup>和邹海峰等<sup>[19]</sup>通过CPTU测试得到的, 如表1所列), 并与通过室内试验得到的超固结比和不排水抗剪强度

的参考值进行比较, 如图9、10所示。在采用净锥尖阻力和超孔隙水压力的关系来预测不排水抗剪强度时 $N_{kt}$ 和 $N_{\Delta u_2}$ 的值是变化的, 因此需要在训练集中为每个类型的不排水抗剪强度试验值找到单个运行良好的 $S_{u(TC)}$ 值。对于这个数据库, 锥尖因子 $N_{kt}$ 在计算 $S_{u(TC)}$ 时取11; 对于 $N_{\Delta u_2}$ 在计算 $S_{u(TC)}$ 时取6。通过表7可以看出, 利用文献[20]计算OCR的 $R^2$ 相对较大, 平均绝对误差MAE和相对误差绝对值RAE较小即文献[20]的计算结果相对准确。同理, 利用文献[15]计算的不排水抗剪强度相对准确。将文献[20, 21]中所提到的计算OCR的传统方法结合CPTU测试数据<sup>[17-19]</sup>(表1)所得到的估计值相关系数 $R^2$ 相对数据融合方法得到的OCR预测值较低, 而平均绝对误差MAE和相对误差绝对值RAE较高; 文献[15]所提到的传统方法所得到的不排水抗剪强度估计值相关系数 $R^2$ 相对数据融合方法得到的不排水抗剪强度预测值较低, 而平均绝对误差MAE和相对误差绝对值RAE较高。总的来说, 基于树的数据融合算法在计算超固结比和不排水抗剪强度的预测值要比传统计算方法<sup>[15, 20-21]</sup>结合CPTU测试数据所得到的估计值更加准确。

表 4 各种树法计算OCR的评价参数汇总表

Table 4 Evaluation parameters for various tree models to calculate OCR

计算方法	评价系数		
	$R^2$	MAN/kPa	RAE/%
回归树法	0.91	0.4	17
自举聚合回归树法	0.92	0.4	17
模型树法	0.89	0.5	19
自举聚合模型树法	0.87	0.6	20
堆叠泛化回归树和模型树法	0.88	0.5	19

表 5 各种树法计算不排水抗剪强度的评价参数汇总表

Table 5 Evaluation parameters for various tree models to calculate  $S_u$

计算方法	评价系数		
	$R^2$	MAN/kPa	RAE
回归树法	0.89	11	23
自举聚合回归树法	0.89	9	22
模型树法	0.93	6	18
自举聚合模型树法	0.95	6	17
堆叠泛化回归树和模型树法	0.92	7	20

表 6 现有超固结比和不排水抗剪强度计算方法

Table 6 Existing methods for OCR and  $S_u$

计算方法	字母含义
$OCR = 0.317 \left( \frac{q_t - \sigma_{v0}}{\sigma'_{v0}} \right) = 0.317 \cdot Q_t$ <sup>[20]</sup>	与文中提到相同
$OCR = \frac{0.46(q_t - u_2)}{\sigma'_{v0}}$ <sup>[21]</sup>	与文中提到相同
$S_u = \frac{q_t - \sigma_{v0}}{N_{kt}}$ <sup>[15]</sup>	$N_{kt}$ 为孔压经验系数
$S_u = \frac{\Delta u_2}{N_{\Delta u_2}}$ <sup>[16]</sup>	$N_{\Delta u_2}$ 为孔压圆锥系数

表 7 现有 OCR 和  $S_u$  计算方法的评价参数汇总表Table 7 Evaluation parameters for existing methods for OCR and  $S_u$ 

$S_u$ 计算方法	评价系数		
	$R^2$	MAN/kPa	RAE/%
OCR计算方法 <sup>[20]</sup>	0.69	0.6	25
OCR计算方法 <sup>[21]</sup>	0.63	0.67	26
$S_u$ 计算方法 <sup>[15]</sup>	0.79	10	28
$S_u$ 计算方法 <sup>[16]</sup>	0.75	12	31

## 5 讨论

### 5.1 数据融合模型的性能

基于树的数据融合算法根据孔压静力触探试验参数和相关特征预测超固结比和不排水抗剪强度方面运行的比较好, 不排水抗剪强度的预测值比超固结比的预测值更可靠。基于预测的准确性(本文定义的准确性是 $R^2$ 最大, MAE和RAE相对较小的相应的算法), 模型树和自举聚合模型树在预测不排水抗剪强度方面优于其他数据融合模型, 而回归树数模型在预测OCR方面优于其他数据融合模型。由于模型树近似于目标变量与输入变量相关的分段线性(即连续)函数, 而传统回归树近似为分段常数(即不连续)函数, 所以理论上模型树的预测结果比回归树更准确。

决策级融合算法有助于提高回归树和模型树的性能。多元回归树的自举聚合算法增加了估计函数的连续性, 从而改善了预测结果使其更加可靠。对不排水抗剪强度, 将回归树和模型树联合到一起的预测往往比单独使用回归树的预测结果好, 但比模型树的预测结果差。

由于在树的根节点选择的变量可以更好地估计训练案例的目标变量, 因此在本次回归树和模型树根的分裂变量研究中, 超固结比和不排水抗剪强度可以被假设为受根节点处的分裂变量中影响最大的变量。本文所采用的基于树的数据融合模型,  $Q_t$ 是存在大多数树预测OCR的根, 而 $q_n$  ( $q_n = q_t - \sigma_{v0}$ ) 是大多数树预测不排水抗剪强度的根<sup>[4]</sup>。 $Q_t$ 和OCR之间较好的相关性, 并且不排水抗剪强度对 $q_n$ 的依赖性符合不排水条件下的经典承载力理论, 其中 $q - \sigma_{v0} = cN_c$ ,  $\varphi = 0$ 且 $c = S_u^{[4]}$ 。

### 5.2 数据融合模型的验证

总体而言, 数据融合模型比CPTU测试方法在预测OCR方面表现更好, 但是在预测不排水抗剪强度的情况下仅比CPTU解释方法略好但可以忽略不计。在估计OCR时, 基于较高的平均绝对误差和相对误差绝对值的预测值和数值预测的相关系数值较低(图9), 数据融合模型比蔡国军等<sup>[17-18]</sup>和邹海峰等<sup>[19]</sup>所用方法的相关性要好(图10)。

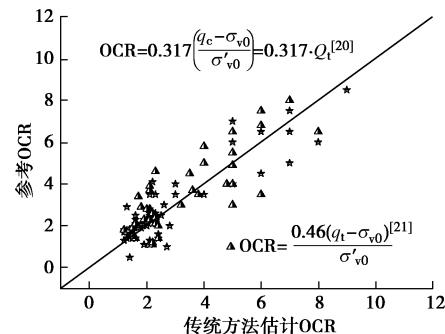
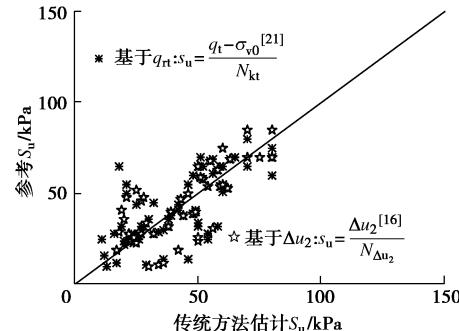


图9 现有两种方法估计OCR和室内试验测试值关系图

Fig. 9 Relationship between OCRs by two methods and laboratory test reference values

图10 现有两种方法估计 $S_u$ 和室内试验测试值关系图Fig. 10 Relationship between  $S_u$  by two methods and laboratory test reference values

## 6 结论

本文研究了基于多传感器孔压静力触探测试数据所测得的土性参数 $\sigma_{v0}$ ,  $u_0$ ,  $q_t$ ,  $u_2$ 以及附加特性, 利用特征级数据融合技术(回归树、模型树)以及决策级数据融合技术(自举聚合、堆叠泛化), 来预测超固结比和不排水抗剪强度。

(1) 数据融合模型在估计超固结比和不排水抗剪强度方面比相应的预测方法运行得更准确。由于数据融合算法具有处理异常值的能力, 因此, 它们可以很好地模拟非线性多变量问题, 包括土结构性、敏感性、矿物学和老化, 样本数据之间复杂的非线性关系。

(2) 基于树的数据融合算法能很好地预测超固结比和不排水抗剪强度。通过对多个不稳定的回归树的预测, 提高了估计函数的连续性, 从而改进了对不排水抗剪强度的预测。堆叠泛化后的回归树和模型树预测的不排水抗剪强度比回归树预测的要好, 但比模型树预测的要差。

(3) 基于树的数据融合算法在计算超固结比和不排水抗剪强度的预测值要比传统计算方法结合CPTU测试数据所得到的估计值更加准确。

## 参考文献:

- [1] 钱永兰, 杨邦杰, 雷廷武. 数据融合及其在农情遥感监测

- 中的应用与展望[J]. 农业工程学报, 2004, **20**(4): 286 - 290. (QIAN Yong-lan, YANG Bang-jie, LEI Ting-wu. Data fusion and its application prospect in agricultural condition monitoring using remote sensing[J]. Transactions of the CSAE, 2004, **20**(4): 286 - 290. (in Chinese))
- [2] 周四望, 林亚平, 聂雅琳, 等. 无线传感器网络中基于数据融合的移动代理曲线动态路由算法研究[J]. 计算机学报, 2007(6): 894 - 904. (ZHOU Si-wang, LIN Ya-ping, NIE Ya-lin, et al. A study of trajectory-based mobile agent dynamic routes algorithm for data fusion in wireless sensor networks[J]. Chinese Journal of Computers, 2007(6): 894 - 904. (in Chinese))
- [3] 刘明贵, 杨永波. 信息融合技术在边坡监测与预报系统中的应用[J]. 岩土工程学报, 2005, **27**(5): 607 - 610. (LIU Ming-gui, YANG Yong-bo. Application of information fusion in slope monitoring and prediction system Chinese[J]. Chinese Journal of Geotechnical Engineering, 2005, **27**(5): 607 - 610. (in Chinese))
- [4] GRIFFIN E P, KURUP P U. Prediction of OCR and  $s_u$  from PCPT data using tree-based data fusion techniques[J]. Journal of Geotechnical and Geoenvironmental Engineering, 2017, **143**(9): 04017045.
- [5] 陈昌彦, 王思敬, 沈小克. 边坡岩体稳定性的人工神经网络预测模型[J]. 岩土工程学报, 2001, **23**(2): 157 - 161. (CHEN Chang-yan, WANG Si-jing, SHEN Xiao-ke. Predicting models to estimate stability of rock slope based on artificial neural network[J]. Chinese Journal of Geotechnical Engineering, 2001, **23**(2): 157 - 161. (in Chinese))
- [6] 赵忠杰, 田梅. 数据融合技术在公路隧道火灾探测中的应用[J]. 西安科技大学学报, 2007, **23**(3): 435 - 438. (ZHAO Zhong-jie, TIAN Mei. Application of data fusion technology to highway tunnels fire detection[J]. Journal of Xi'an University of Science and Technology, 2007, **23**(3): 435 - 438. (in Chinese))
- [7] 刘松玉, 蔡国军, 邹海峰. 基于 CPTU 的中国实用土分类方法研究[J]. 岩土工程学报, 2013, **35**(10): 1765 - 1776. (LIU Song-yu, CAI Guo-jun, ZOU Hai-feng. Practical soil classification methods in China based on piezocone penetration tests[J]. Chinese Journal of Geotechnical Engineering, 2013, **35**(10): 1765 - 1776. (in Chinese))
- [8] SENNESET K, JANBU N, SVANO G. Strength and deformation parameters for CPT[C]// Proceedings of the 2nd European Symposium on Penetration Testing, Amsterdam, 1982: 863 - 870.
- [9] LUNNE T, CHRISTOFFERSEN H P, TJELTA T I. Engineering use of piezocone data in North Sea clays[C]// Proceedings of the 11th International Conference on Soil Mechanics and Foundation Engineering, Rotterdam, 1985.
- [10] TONG L, WANG Q, DU G, et al. Determination of undrained shear strength using piezocone penetration test in clayey soil for bridge foundation[J]. Journal of Southeast University, 2011, **27**(2): 201 - 205.
- [11] LI X, CAI G, LIU S, et al. Undrained shear strength and pore pressure changes due to prestress concrete pile installation in soft clay[J]. International Journal of Civil Engineering, 2019, **17**(2): 193 - 203.
- [12] MAYNE P W. Determination of OCR in clays by piezocone tests using cavity expansion and critical state concepts [J]. Soils and Foundations, 1991, **31**(2): 65 - 76.
- [13] CHEN B S Y, MAYNE P W. Profiling the OCR of clays by piezocone tests tests[R]. Atlanta: Georgia Institute of Technology, 1994.
- [14] 刘松玉, 蔡国军, 童立元, 等. 基于 CPTU 测试的先期固结压力确定方法试验研究[J]. 岩土工程学报, 2007, **29**(4): 490 - 495. (LIU Song-yu, CAI Guo-jun, TONG Li-yuan, et al. On preconsolidation pressure of clays from piezocone tests[J]. Chinese Journal of Geotechnical Engineering, 2007, **29**(4): 490 - 495. (in Chinese))
- [15] LUNNE T, ROBERTSON P K, POWELL J M. Cone penetration testing in geotechnical practice[M]. London: Blackie Academic & Professional, 1997: 16 - 24.
- [16] HALL D L, MCMULLEN S A H. Mathematical techniques in multi-sensor data fusion[M]. Boston: Artech House, 2004.
- [17] CAI G, PUPPALA A J, LIU S. Characterization on the correlation between shear wave velocity and piezocone tip resistance of Jiangsu clays[J]. Engineering Geology, 2014, **171**(13): 96 - 103.
- [18] CAI G, LIU S, PUPPALA A J. Comparison of CPT charts for soil classification using PCPT data: example from clay deposits in Jiangsu Province, China[J]. Engineering Geology, 2011, **121**(1/2): 89 - 96.
- [19] ZOU H, LIU S, CAI G, et al. Multivariate correlation analysis of seismic piezocone penetration (SCPTU) parameters and design properties of Jiangsu quaternary cohesive soils[J]. Engineering Geology, 2017, **228**(1/2): 11 - 38.
- [20] LIU H, MOTODA H. Feature extraction, construction and selection: a data mining perspective[M]. Amsterdam: Kluwer Academic Publishers, 1998.
- [21] CHEN B S Y, MAYNE P W. Statistical relationships between piezocone measurements and stress history of clays[J]. Canadian Geotechnical Journal, 1996, **33**(3): 488 - 498.