

# 黄土湿陷性的智能化评价

井彦林<sup>1,2</sup>, 仵彦卿<sup>1</sup>, 崔中兴<sup>1</sup>, 宋学庆<sup>3</sup>, 李寅良<sup>2</sup>

(1. 西安理工大学 岩土所, 陕西 西安 710048; 2. 煤炭工业西安设计研究院, 陕西 西安 710054;  
3. 信息产业部 电子综合勘察研究院, 陕西 西安 710054)

**摘要:** 提出了用数据挖掘中的预测技术进行黄土湿陷性评价。根据实际工程资料建立黄土物理力学数据库, 用 BP 算法建立预测模型。实例分析表明, 预测湿陷系数所得的湿陷量计算值精度可达 89%, 这说明黄土湿陷性的智能化评价方法具有可行性和实用性。

**关键词:** 黄土湿陷性; 数据挖掘; BP 神经网络; 预测

文献标识码: A

文章编号: 1000-288X(2006)01-0053-04

中图分类号: S151.92, S155.24

## Intelligent Evaluation of Loess Collapsibility

JING Yan-lin<sup>1,2</sup>, WU Yan-qing<sup>1</sup>, CUI Zhong-xing<sup>1</sup>, SONG Xue-qing<sup>3</sup>, LI Yin-liang<sup>2</sup>

(1. Xi'an University of Technology, Xi'an 710048, Shaanxi Province, China; 2. Xi'an Design and Research Institute of Coal Industry, Xi'an 710054, Shaanxi Province, China; 3. Electronic Comprehensive Investigation and Surveying Institute of Ministry of Information Industry, Xi'an 710054, Shaanxi Province, China)

**Abstract:** This paper presents a method for assessment of loess collapsibility using the data mining technology. Coefficients of collapsibilities are predicted using the method. The database should be created based on practical engineering, and a prediction model, built with the BP neural network. The predicted loess collapse settlement is compared with the measured loess collapse settlement. Results show that prediction precision of collapse settlement is up to 89% for a specific project case. This indicates that the intelligent method of evaluating loess collapsibility is very useful in engineering.

**Keywords:** loess collapsibility; data mining; BP neural network; prediction

## 1 前言

黄土在世界上分布很广, 在北美、东欧、及我国均有分布。从厚度上来讲, 在欧、美、澳洲国家一般分布较薄, 而在我国及其邻近的哈萨克斯坦、乌兹别克斯坦等国则分布较厚, 最厚处达数百米。也正是由于分布特征以及国情的不同, 世界各国对黄土的研究重点也不尽相同。

黄土之所以为特殊性土, 主要是由于其湿陷性。目前, 国内对湿陷性的评价是采用自重湿陷系数、湿陷系数评价计算湿陷类型及湿陷等级。国内也有学者提出用本构关系等方法计算自重湿陷系数及湿陷系数、评价湿陷性。湿陷性指标的获得可通过试坑浸水试验或室内试验。由于野外浸水试验周期长, 费用高, 所以在工程中常采用室内试验测求土的湿陷性指标。进行黄土的湿陷性室内试验, 首先要在野外采取不扰动黄土试样, 在室内进行黄土固结试验, 求出湿

陷系数、自重湿陷系数等指标, 然后根据地层结构和建筑物的结构特点进行湿陷类型计算和湿陷等级计算。由于钻孔取土易被扰动, 因此对湿陷性黄土主要是人工挖探井, 人工采样。这样采取的土样质量可靠, 在 20 m 深度以内普遍采用。但当湿陷性黄土较厚时开挖较困难, 安全性不强, 文明程度不够, 且造价高, 工期长。

在工程实践中, 黄土的湿陷性指标应用非常广泛, 几乎是每个工程必须具备的试验数据。多年来我国在黄土地区进行了大量的工程实践和研究工作, 在黄土的物理力学指标方面积累有丰富的资料。干重度、重度、含水量、液隙比等指标易测、误差小, 值稳定, 而湿陷系数随试验条件、测试技术等不同而不同, 变化幅度大, 影响评价的精度。所以, 若以现代科学理论为基础, 运用先进的信息技术如数据挖掘技术, 在进行黄土的物理力学性质分析的基础之上, 掌握其变化规律, 探讨湿陷系数、与物理力学性质如干重度、

收稿日期: 2005-09-19

资助项目: 煤炭工业西安设计研究院科研基金资助

作者简介: 井彦林(1963-), 男(汉族), 博士生, 主要从事黄土力学与工程、环境岩土工程研究。E-mail: jingyanlin@sina.com。

重度、含水量、液限比等指标之间的关系,建立黄土物理力学指标的预测模型,根据黄土的一般物理力学指标间接预测出黄土的湿陷性,在实际应用中可减少取土和试验数量,提高黄土评价手段,降低投资,加快工程建设进度。

从理论上讲,大量的测试数据可为黄土力学的研究提供新的舞台,新技术的应用为黄土的研究增添新的活力。运用数据挖掘技术进行黄土研究的成果未见报道,但相关研究如探讨黄土物理力学指标同湿陷性关系的研究成果较多。具有代表性的有刘祖典(1994)用统计学理论研究黄土的湿陷性问题,探讨不同重度、含水量、液限、结构性及不同时代、不同成因类型状态下的湿陷性的变化规律。他对湿陷系数进行线性和非线性回归,提出了相应的经验公式<sup>[1]</sup>;关文章(1992)分析了天然干重度、含水量与湿陷性的关系,提出了拟合公式,给出了不同地区黄土的界限含水量和界限干密度。提出了湿陷性综合评价,即不仅是考虑自重湿陷量和总湿陷量,而是根据所有影响湿陷性的因素如渗透性、地层构成特性、附加应力、建筑物长宽比、湿陷土层厚度等,采用模糊数学的方法,用综合湿陷强度来表示评价结果<sup>[2]</sup>;胡再强(2000)通过对黄土三轴浸水试验及试验前后的扫描电镜分析,由非饱和黄土显微结构与湿陷性之间的联系,确定黄土结构与湿陷性的关系<sup>[3]</sup>;齐吉林(2001)进行原状和人工制备黄土的剪切试验,根据试验结果建立起描述土结构性的定量化指标,提出土样在不同压力作用下的结构性参数比与抗剪强度比之间呈线性关系,且不同土样的直线的截距和斜率大致相同<sup>[4]</sup>;苗天德(1990)根据黄土扫描电镜观察和土力学试验结果,提出湿陷变形系由黄土微结构失稳造成的,利用突变理论建立了湿陷性黄土的微结构失稳模型,并按照现代连续统介质力学方法推导出了湿陷性黄土的本构关系;李晓军(1996)将分形理论引入了黄土的研究,论述了黄土孔隙的分形维数,提出黄土粒度分形维数,分维数的大小代表了黄土的演化历史,分维数越大,颗粒越细,表面积越大,湿陷性越小;用孔隙比、容重、含水量、液限及显微结构特征类型建立 ANN 模型,进行湿陷类型预测<sup>[5]</sup>;滕文川(1993)根据黄土一般物理指标用模糊数学的方法研究黄土的湿陷性。这些研究所采用的思路和技术与数据挖掘还有一定距离,研究中所建模型多采用一元函数,而对湿陷性的影响是多因素的;模型中的有些参数难以测得,甚至比实测计算结果还要复杂。本文的目的是采用数据

挖掘的方法进行湿陷性研究,方法是通过实际工程数据的收集建立黄土物理力学数据库,用数据挖掘的预测技术,依据数值稳定、易于测定的物理指标对湿陷系数进行预测,达到减少试验数量、降低工程投资的目的。

## 2 数据挖掘简介

数据挖掘(data mining, DM)是 20 世纪 90 年代中期发展起来的一门高新技术。它是从大量的不完全的、有噪声的、模糊的、随机的实际数据中提取隐含在其中的、不为人们所知的、但又有潜在应用价值的信息和知识的过程。它是知识发现(knowledge discovery in database, KDD)中的关键环节。数据挖掘的主要挖掘对象可以是结构化的关系型数据(仓库),也可以是半结构化数据或网络数据;任务主要包括特征规则挖掘、辨识规则挖掘、关联(分为布尔关联和量化关联两种)规则挖掘、分类规则挖掘、数据聚类、预测、趋势性规则挖掘、公式发现、可视化、数据纠正等;算法有归纳学习法、决策树、遗传算法、神经网络、统计分析、模糊数学、关联分析、粗糙集、规则发现、公式发现和关联顺序等;数据挖掘的结果输出有图形、报告、逻辑公式等多种形式。数据挖掘的过程包括数据预处理、方法选择、建立模型、挖掘实施、规则抽取、知识解释等环节<sup>[6]</sup>。

## 3 数据库的建立

本文的研究思路是首先根据实际工程的黄土固结试验等结果建立数据库,然后进行预测模型的建立,用预测模型对实际工程的湿陷性进行评价,以验证模型的预测效果,最后对模型做出评价。本文直接使用采样数据建立数据库。数据的采样方法直接关系到挖掘的质量。数据挖掘中采样方法有随机采样、分层采样和窗口采样,本文采用分层采样法,即根据分布的不均衡控制采样的频率。当数据分布密度较大时,采样频率低,反之采样频率高。采样范围选定在陕西省关中中部地区,采样原则是要有广泛性和代表性,样本尽量覆盖多种地貌,要求样本全部为探井人工采取,均来自已完成的实际工程。这些数据是工程设计中已采用过的数据,均已经过严格分析,甚至测量不确定度分析,所以质量较为可靠。以此原则,共收集黄土固结试验数据 1 257 组,每组试验数据包括取土深度、含水量、密度、比重、液限、塑限、压缩系数、湿陷系数等指标,指标来源及场地特性见表 1。

表 1 训练样本场地特征

工程名称	所处地貌	湿陷性 土层厚 度/m	土样 数量 /组	工程名称	所处地貌	湿陷性 土层厚 度/m	土样 数量 /组
景观 360	渭河三级阶地	15.0	163	黄陵总机厂	沮水河一级阶地	4.0	31
白水煤矿铁路	黄土塬	8.0	43	黄陵一号井水池	黄土塬	12.0	32
西安矿院宿舍楼	渭河三级阶地	12.0	18	长武商贸大厦	渭河二级阶地	18.0	37
王家岭爆破库	洪积扇	8.5	25	长安县三水厂	黄土斜坡	15.0	59
铁一院科研楼	黄土塬	10.0	37	彩虹主厂房	渭河一级阶地	7.0	74
蒲白马村矿浴室	黄土塬	12.0	30	彩虹配料厂房	渭河一级阶地	8.0	14
蒲白电厂	黄土塬	15.0	137	彬县下沟煤矿	泾河一级阶地	11.0	179
联合油气站	黄土塬	20.0	273	彩虹住宅三期	渭河二级阶地	10.0	105

## 4 工程实例

### 4.1 场地概况

拟建的陕西澄城董东煤矿工业场地, 位于陕西省澄城县城东北 4.3 km。距离训练样本所在场地最近者为蒲白矿务局矸石电厂, 两者相距约 70 km。董东煤矿地貌为渭北黄土高原, 地形较平坦, 北高南低, 地面标高 722.15~725.54 m, 最大高差 3.39 m。场地内主要建(构)筑物有井架、绞车房、原煤仓、机修车间、消防材料库、35 kV 变电所、日用消防水池、单身宿舍、办公楼、休息室等。场地地层表层为耕土, 其下依次为第四系全新统残积黑垆土( $Q_4^{2cl}$ )、上更新统风积黄土( $Q_3^{col}$ )、残积古土壤( $Q_3^d$ )、中更新统风积黄土( $Q_2^{col}$ ), 共 6 层黄土, 2 层古土壤。常规压力下, 湿陷性土层厚近 20 m。自重湿陷量的计算值为 386~940 mm, 场地为自重湿陷类型。

### 4.2 预测模型及效果检验

BP 神经网络虽有很多缺点, 其数学基础还没有完全建立起来, 但其高度非线性的映射能力使其应用非常之广, 本文采用该方法建模。1943 年, 最早的 MP 神经网络用于二值分类, 所以传递函数是阶跃函数; 20 世纪 80 年代, 神经网络尤其是 BP 网络有了很大的发展, 出现了 Sigmoid 和 Tansig 等传递函数, 但这两类函数是由阶跃函数演化来的, 还主要是用于分类。湿陷系数是一个连续型数值, 其预测实际上是一个回归问题。以 Tansig 为例, 在进行回归计算时, 当网络的某层输出接近于 -1 或 +1, 网络处于饱和状态, 调整能力较差, 使网络迭代次数增加, 收敛性减弱, 训练时间延长。数据变换是数据挖掘的预处理阶段非常重要的一环。使用 BP 算法, 常用的数据变换是归一化处理。在进行连续函数的预测时, 当数值落在饱和段, 网络收敛速度减慢(见图 1)。如果要使网络训练速度加快, 应将各层输出限制在非饱和区, 要

做到这点可通过调整权值或阈值等方法来实现。最简单的办法是使数值减小, 因为在非负范围内, 该函数为增函数, 但数值过小会使误差增大。经试算, 将输入映射在 [0, 0.18] 范围内, 其网络精度和收敛速度都令人比较满意。因此, 本文对数据标准化后, 原变量范围在 [0, 0.18] 之间, 而没有采取常用的归一化处理。建模时传递函数采用 Tansig 函数, 采用 3 层网络; 输入层为 8 个节点, 隐含层为 20 个节点, 输出层为 1 个节点, 即湿陷系数。网络误差为 0.01。预测变量结合文献[1]、[2]对黄土湿陷性影响因素的分析, 选定为取土深度、含水量、比重、密度、液限、塑限、压缩系数  $a_{1-2}$  及自重应力等 8 个指标。自重应力不是实测指标, 故需对数据库增加这一新的属性, 其数值据取土深度及其上覆土的密度计算得出。

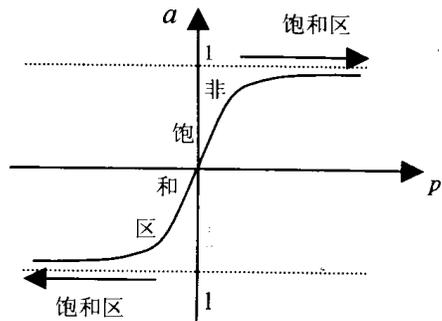


图 1 Tansig 函数曲线

建模时, 首先用数据库中的 1 257 组数据对网络进行训练, 训练完成后即得到网络权值, 以此建立预测模型。用所建立的预测模型对陕西澄城董东煤矿工业场地 9 个探井土样的湿陷系数进行预测, 然后用预测的湿陷系数进行黄土的湿陷量计算, 计算结果见表 2。试验值所得湿陷量计算值为 704~1 082 mm, 预测湿陷系数所得湿陷量计算值为 789~956 mm, 预测精度为 81.7%~99.2%, 平均为 89.3%, 预测计算的湿陷等级同试验值完全一致, 均为 Ⅴ级。

表 2 预测与试验值结果对比

建筑物名称	勘探点 编号	湿陷土层起 始深度/m	试验值		预测值		
			湿陷量/mm	湿陷等级	湿陷量/mm	准确率/%	湿陷等级
变电所	1	2.49~ 16.50	920	Ⅴ	789	85.8	Ⅴ
	4	2.18~ 16.50	842	Ⅴ	849	99.2	Ⅴ
日用消防水池	5	1.84~ 16.50	960	Ⅴ	838	87.3	Ⅴ
	9	3.36~ 16.50	704	Ⅴ	796	86.9	Ⅴ
单身宿舍	10	2.09~ 16.50	984	Ⅴ	871	88.5	Ⅴ
	12	2.02~ 16.50	901	Ⅴ	847	94.0	Ⅴ
	16	1.50~ 16.50	1082	Ⅴ	956	88.4	Ⅴ
休息室	21	2.70~ 16.50	1008	Ⅴ	824	81.7	Ⅴ
消防材料库	34	2.06~ 16.50	930	Ⅴ	854	91.8	Ⅴ

(1) 人工神经网络是一种多元非线性分类方法, 本文首次将之用于黄土的湿陷性研究及应用。验证结果证明了这种方法的有效性和实用性。

(2) 预测变量不但考虑了黄土本身的物理力学性质, 还通过自重应力考虑了其应力状态。预测模型在规划和初勘阶段可直接用于工程, 在施工图设计阶段此预测方法可将工程中缺失的湿陷性指标补充完整, 提高湿陷性评价的可靠性。

(3) 用数据挖掘的方法对湿陷系数进行预测是一种有效的方法, 本文所采用的训练样本主要来自陕西关中, 实例表明这种方法是可行的, 下一步有必要在更广的范围内开展此项研究。

## [ 参 考 文 献 ]

- [1] 刘祖典. 影响黄土湿陷系数因素的分析[J]. 工程勘察, 1994(5): 6—10.
- [2] 关文章. 湿陷性黄土工程性能新篇[M]. 西安: 西安交通大学出版社, 1992. 36—53.
- [3] 胡再强, 沈珠江, 谢定义. 非饱和黄土的显微结构与湿陷性[J]. 水利水运科学研究, 2000(2): 68—71.
- [4] 齐吉林. 土的结构性及其定量化参数的研究[D]. 西安: 西安理工大学出版社, 1999.
- [5] 李晓军. 黄土的特性分析及黄土湿陷性的预测[D]. 西安: 西安科技大学出版社, 1996.
- [6] 刘同明. 数据挖掘技术及其应用[M]. 北京: 国防工业出版社, 2001. 223—225.

(上接第 37 页)

例如缩小还草与还林的补助年限的差距, 让农民在还生态林、经济林、草之间做出更加多样的选择。在一些暂不具有还林草条件的地区, 甚至可以将退耕与还林草分步实施。发达国家在农产品过剩而对林产品需求增长时, 将“多余”的耕地向林地的转变通常是间接的, 即先把边缘地区的耕地变为永久性牧场, 接着才有部分牧场变为林地<sup>[5]</sup>。还要改革现有的林业政策。目前的退耕还林政策限制了农户的林业经营权, 规定生态林的比例为 80%, 而经济林的比例为 20%。并且, 生态林是禁止采伐的, 用材林的采伐也必须得到主管部门的严格审批。难以保证农户投资林业的激励, 也使得退耕还林无法保持其持续性, 因而需要做出必要的调整。这需要政府充分利用市场机制, 通过优惠的补贴政策和税收政策, 使农户在退

耕还林中确实得到实惠, 形成还林的激励, 从而形成国家(宏观)目标与农户(微观)目标的激励相容<sup>[6]</sup>。

## [ 参 考 文 献 ]

- [1] 田均良, 刘国彬. 黄土高原退耕还林工程中现存问题及有关建议[J]. 水土保持通报, 2004, 24(1): 63—78.
- [2] 王乃斌, 沈洪泉, 赵存兴, 等. 黄土高原地区资源与环境遥感调查数据集[M]. 西安: 西安地图出版社, 1991.
- [3] 徐晋涛, 陶然, 徐志刚. 退耕还林: 成本有效性、结构调整效应与经济可持续性[J]. 经济学季刊, 2004(4): 8.
- [4] 赵曦. 西部地区退耕还林试点工程问题与对策研究[J]. 农业经济问题, 2000(12): 2—5.
- [5] 宋乃平, 张凤荣, 李国旗, 等. 西北地区植被重建的生态学基础[J]. 水土保持学报, 2003, 17(5): 1—4.
- [6] 蒋海. 中国退耕还林的微观投资激励与政策的持续性[J]. 中国农村经济, 2003(8): 30—36.