基于 ANN 与 GIS 技术的区域岩溶塌陷稳定性预测——以桂林西城区为例

胡　成[1]，陈植华[1]，陈学军[2]

（1. 中国地质大学工程学院，湖北武汉　430074；2. 桂林工学院土木系，广西桂林　541001）

摘要：岩溶地面塌陷是岩溶区常见的一种地质灾害，塌陷区域预测是进行国土规划、资源开发与灾害防治的必要工作. 由于岩溶塌陷的影响因素众多且相互作用，发展过程复杂，加之各评价因子的数值获取困难，致使长期以来塌陷区域定量预测成为一个难以解决的课题. 现行的区域预测模型不能描述塌陷形成模式的非线性特征，也难以克服评价因子权重确定过程中人为经验因素的影响. 神经网络技术的自学习、自适应与高度非线性映射特点显示了其在塌陷区域预测领域中应用的前景. 根据研究区内地面塌陷空间聚集分布的特征，提出了不同因子组合条件下塌陷发生可能性的定量化方法，结合选定的评价因子类别确定了神经网络预测模型的结构，利用 312 个塌陷点样本中的 292 个进行网络训练，余下的 20 个样本的校验结果表明该模型具有较高的可信度. 运用 GIS 技术将研究区进行评价单元划分，并获取各评价因子的取值，输入到训练好的网络中进行预测. 将各单元的输出值进行归并处理后得到研究区岩溶塌陷的稳定级分区图.

关键词：岩溶塌陷；人工神经网络；非线性；预测模型；GIS.
中图分类号：P642.254；P642.26　　**文献标识码**：A
文章编号：1000－2383(2003)05－0557－06

作者简介：胡成(1976－)，男，中国地质大学(武汉)博士研究生，主要研究方向为环境模拟评价及 GIS 技术应用. E-mail：szmnlab@cug.edu.cn

0　引言

广西桂林西城区包括琴潭开发区和西城工业区，是市重点开发的新城区. 该区碳酸盐岩分布广泛，岩溶塌陷灾害十分严重. 据地面调查统计，至 1999 年西城区 29.13 km^2 的第四系覆盖区共发生塌陷 312 个，平均塌陷密度达到 10.71 个/km^2. 严重的塌陷灾害已经成为当地经济建设和人民生命财产安全的严重威胁. 在碳酸盐岩分布地区对岩溶塌陷的发生可能性进行预测，评价岩溶环境区的盖层稳定性，是决策管理部门进行国土规划、资源开发、基础设施建设与防灾减灾的必要性工作.

目前岩溶塌陷的预测主要有以下 2 种类型：一是以评价岩溶空腔上覆松散沉积物的稳定性为任务的点状塌陷预测，预测模型大多属于建立在土体强度理论基础上的力学平衡模型[1,2]；另一种则是区域塌陷预测，其首要目标是判断出不同区域内发生岩溶塌陷可能性的相对大小. 由于岩溶塌陷的孕灾环境复杂，影响因素众多而且相互制约，加之各因素本身强烈的空间变异性以及人类活动的多变性，从而导致岩溶塌陷的区域预测、预报长期停留在定性或半定量的阶段. 虽然近年来曾有学者在岩溶塌陷区域定量化预测方面作了积极的探索[3~5]，提出了一些预测模型，在实际应用过程中取得了较明显的效果，但均未摒弃确定因素权重过程中的人为经验影响，同时也没有考虑因素耦合作用等敏感问题，因此在一定程度上降低了预测的可信度.

人工神经网络(artificial neural network，简称 ANN)作为一门新兴的系统技术，自从 20 世纪 80 年代其研究再次复苏以来，在地学界逐渐得到了日益广泛的应用并成为研究热点之一. 目前已经涉及了岩性分类、线性构造识别、遥感图像分类、地学信息提取与挖掘及资源评价预测等诸多领域，显示了强大的生命力. 但在岩溶地面塌陷预测方面仅有运

收稿日期：2003－05－19
基金项目：国家计委防灾减灾项目(计司国地[1996]12).

用 ANN 技术进行点状塌陷稳定性评价的研究[6,7]，且对网络训练目标的取值和分级问题没有作较为详细的讨论，而对于区域塌陷预测则未见相关报道. 本文以桂林西城区为例，以 GIS 的空间数据统计分析功能为操作平台，针对现行区域岩溶塌陷预测模型的不足，将 ANN 技术引入到岩溶塌陷区域评价中，作为一种新预测方法而加以初步探讨.

1 BP 神经网络

ANN 是在现代神经科学研究成果的基础上，根据对人脑的组织结构、功能特征进行模仿而发展起来的一种新型信息处理系统和计算体系. 它属于高维非线性动力学系统范畴，可实现输入到输出之间的高度非线性映射，具有良好的自适应、自组织特征及较强的学习和容错能力，能通过学习人为给定的样本范例而获取知识[8]. ANN 的这些特征有助于消除或降低目前区域岩溶塌陷预测过程中不确定因素的影响，同时也预示了其在该领域中的应用前景.

BP 神经网络一般具有 3 层或 3 层以上阶层型结构，层间各个神经元实现全连接. 网络按有导师示教的方式进行监督训练学习，当一对学习模式提供给网络后，输入值从输入层经中间隐层向输出层传播，在输出层获得网络的输入响应，并遵循减小模拟值与实际值之间误差的方向，由输出层开始逐层调节各连接权值. 随着误差逆向传播修正的不断进行，网络输入响应的正确率也逐渐提高. 在达到允许的误差阈值要求后自动收敛[9]. 本文采用算法步骤为：

(1)初始化网络，为各连接权重赋值.

(2)向网络提供训练样本的输入向量与期望输出向量.

(3)选用 Sigmoid 函数为神经元的激活函数，则对于第 l 层的 j 神经元的输入值 $x_j^{(l)}$ 与输出值 $y_j^{(l)}$ 按下式进行递归计算

$$x_j^{(l)} = \sum_{i=0}^{T} w_{ji}^{(l)}(n) y_i^{(l-1)}, \tag{1}$$

$$y_j^{(l)} = 1/[1+\exp(-x_j^{(l)})]. \tag{2}$$

式中，T 为前一层(即第 $l-1$ 层)神经元的个数，$w_{ji}^{(l)}$ 为第 l 层 j 神经元与第 $l-1$ 层 i 神经元之间的连接权值. 对于网络输入层($l=0$)有：$y_j^{(0)}=x_j^{(0)}$；对网络输出层($l=L$)则有：$y_j^{(L)}=O_j^{(L)}$. $x_j^{(0)}$ 与 $y_j^{(L)}$ 分别为训练样本的输入向量与输出向量的第 j 个分量，记网络输出层 j 神经元的模拟值与实际值之间的误差 $e_j=y_j^{(L)}=O_j^{(L)}$.

(4)采用网络的均方根误差

$$E_{av} = \sqrt{\sum_{s=1}^{n}\sum_{j=1}^{m}(e_j^{(s)} - O_j^{(L)})^2/(m\times n)}$$

为判断网络迭代是否结束的判别函数. 式中 n 为训练样本的容量，m 为输出层神经元的个数，$e_j^{(s)}$ 表示第 s 个样本位于输出层 j 神经元处的模拟误差. 若 $E_{av}<\varepsilon$(预先设定的允许误差阈值)，结束学习过程，否则跳转到下一步.

(5)运用带动量项的最速下降法修正连接权值：

$$w_{ji}^{(l)}(t+1)=w_{ji}^{(l)}(t)+\eta\delta_j^{(l)}y_i^{(l-1)}+\alpha[w_{ji}^{(l)}(t)-w_{ji}^{(l)}(t-1)]. \tag{3}$$

对于输出神经元：$\delta_j^{(l)}=e_jO_j[1-O_j]$. 对于隐层神经元：

$$\delta_j^{(l)}=y_j^{(l)}(1-y_j^{(l)})\sum_k\delta_k^{(l+1)}w_{kj}^{(l+1)}.$$

式中，t 为当前迭代次数，η 为学习率，α 为惯性项系数.

(6) 转到第(3)步直至迭代结束.

根据以上算法思想编制 VC++计算程序以自动实现模拟过程.

2 区域岩溶塌陷预测的神经网络模型

2.1 评价因子的确定

评价因子的恰当选择是对岩溶地面塌陷形成本质的刻画，它不仅决定了预测神经网络模型的输入层结构，同时也影响着预测评价的准确性. 在有关的塌陷预测文献中，评价因子种类不尽相同，其个数也有较大的差别[3~7]. 事实上，岩溶塌陷是“岩溶－盖层－水”系统在各种应力作用下表现出的失稳过程，评价因子的确定应该建立在利用工程地质科学原理分析塌陷形成原因与过程的基础上. 根据实际调查资料并结合前人研究成果[10~12]，本文认为，表层岩溶带的发育强度、盖层特征及地下水的流场性质是控制研究区地面塌陷的 3 个最基本因素. 三者分别为塌陷的形成提供了发育空间、物质来源与动力条件. 其他的自然或人工因子则是通过改变此 3 个基本因素而间接地对塌陷形成施加影响，如地质构造在某种程度上决定了表层岩溶带的发育特征与规模，人工活动强度则在地下水的动态特征中予以体现等. 因此，本次研究选择了 10 个评价因子，各因子的含义及量化方法如表 1 所示.

表 1　区域岩溶塌陷预测评价因子选择

Table 1　Selected factors for prediction of regional cover-collapse probability

评价因子	盖层特征		岩溶表层带发育强度			地下水动态特征				
	厚度 I_1/m	成因 I_2/km²	基岩岩性 I_3/km²	空间遇洞指数 I_4	距断层距离 I_5/m	断层构造强度 I_6/m	水位平均变幅 I_7/m	水位平均变频 I_8	距强径流带距离 I_9/m	水位距基岩面距离 I_{10}/m
数据类型	比率量	命名量	命名量	比率量	比率量	比率量	比率量	比率量	比率量	比率量
量化方法		成因类型的平均塌陷密度	岩性类型的平均塌陷密度	0.5 km×0.5 km 内的溶洞数		0.5 km×0.5 km 内断层长度	地下水位变幅多年平均值	一年内水位变幅>0.5 m 次数		

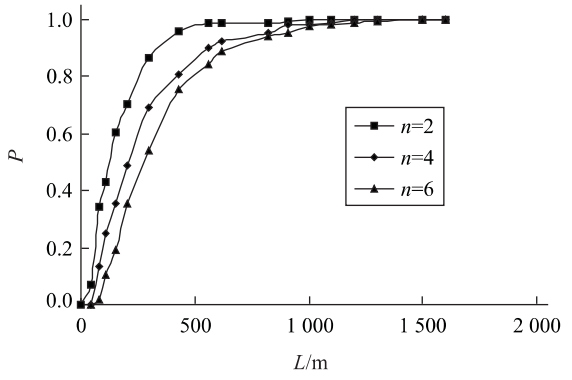

图 1　研究区塌陷点的模式分析

Fig. 1　Pattern analysis of cover-collapse plot in study area

P 代表任一塌陷点周围 L m 邻域范围内存在另外 n 个塌陷点的平均概率

2.2　评价目标值的确定

评价目标值的定量化方式与神经网络输出层的结构密切相关,是网络学习训练的必需参数之一. 其数值应能反映出不同影响因子组合条件下,岩溶地面塌陷发生的几率大小. 调查发现,研究区内塌陷点并不是毫无规律的随机分布. 利用 GIS 进行塌陷点的模式分析(pattern analysis)可以发现,在任意一塌陷点方圆 500 m 范围内存在 6 个塌陷点的可能性约为 80%(图 1),表现出了较强的空间聚集特征. 这说明塌陷点聚集区范围内的各评价因子量值是利于塌陷发生的危险组合. 不同的聚集区规模不同,评价因子量值也有差别,因而形成了不同因子组合(输入)与塌陷发生概率(输出)之间的对应关系. 据此本文选择塌陷点周围 0.5 km×0.5 km 正方形邻域内的塌陷点数作为网络训练样本的评价目标值.

2.3　预测模型结构设计

BP 神经网络通常由输入层、隐层与输出层组成. 设计网络模型的结构即是确定隐层的层数和每层神经元的个数. 输入层与输出层的神经元个数分别由评价因子的种类及评价目标值的维数决定,对于隐层的层数和神经元数目则可通过试验的途径获得. 在多次试验的基础上,确定的网络结构如下:输入层取 10 个神经元,以对应 10 个评价因子;隐层数为 2,神经元个数分别为 8 和 4;输出层设定 1 个神经元,对应于评价目标值.

3　神经网络模型参数辨识

3.1　模型参数辨识与分析

模型参数,即相邻两层中神经元的连接权重,由网络从训练样本学习中获取. 依照上述岩溶塌陷影响因子与评价目标值的量化原则,将研究区内 312 个塌陷点样本逐一取值形成输入输出模式对. 同时将塌陷点样本随机分成两部分,其中 292 个供网络训练学习,余下的 20 个用来校验模型预测精度. 由于各个输入输出项的量纲不同,在数量上存在较大差别,因此在网络学习与校验之前采用极值法进行归一化处理. 设定学习率 $\eta=0.05$,惯性项系数 $\alpha=0.6$,网络迭代停止阈值 $E_{av}\leqslant 0.02$,按如上算法进行网络训练. 经 660 371 次迭代后,样本模拟值与目标值的相关系数为 0.997,网络收敛,精度达到要求. 各神经元间的连接权值得以确定,并可据此算出各评价因子相对于输出目标值的相对影响程度(表 2). 表明研究区的地面塌陷明显地受岩溶发育程度控制,空间遇洞指数、构造强度及盖层特征是影响塌陷形成的较敏感因子,这与区内塌陷的分布特征与形成规律[10,11]是一致的.

3.2　模型预测精度校验

对参数已确定的模型利用 20 个校验样本进行测试计算发现,模拟值与目标值的最大误差为 0.153,相关系数为 0.94,误差的标准方差为 0.087. 如果按表 3 所示的标准进行区域岩溶塌陷危险性分级,易知出现 1～4 级的误判平均极差分别为 0.1,0.3,0.5 与 0.7. 校验结果显示该模型发生 1 级误判的几率为 1/20=0.05(图 2),2 级或 2 级以上误判

表 2 评价因子对于塌陷的相对影响程度

Table 2 Relative contribution of factors to cover-collapse probability

评价因子	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}
权重值/%	10.24	10.59	6.33	20.11	11.01	10.13	5.88	9.73	8.72	7.26

表 3 岩溶塌陷发生危险性分级

Table 3 Target value classes of cover-collapse probability

等级	稳定级	基本稳定级	中等稳定级	不稳定级	极不稳定级
目标归一化值区间	0.0～0.2	0.2～0.4	0.4～0.6	0.6～0.8	0.8～1.0

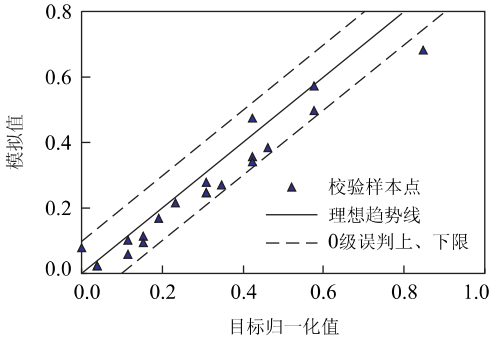

图 2 人工神经网络模型精度校验结果

Fig. 2 Validation result of ANN model precision

的可能性为 0%，具有较高的可信度，可用于区域塌陷预测.

4 区域岩溶地面塌陷预测

4.1 预测单元划分与数据准备

利用 GIS 软件的栅格处理(raster manipulation)功能将研究区内 29.13 km^2 的松散岩类覆盖区划分成 50 m×50 m 共 11 644 个正方形单元格，作为预测单元. 同时在 GIS 环境中使用不同的空间插值技术将以离散点、线形式存在的数据进行区域化，形成区域参数图而获得各预测单元所有评价因子的取值，用极值法归一化后作为神经网络预测模型的输入.

4.2 塌陷预测及可信度分析

将预测单元的评价因子归一化值输入到前述训练好的神经网络模型中进行运算，得到各预测单元相应的模拟值，并按表 3 所示的定级原则进行分类后获得研究区地面塌陷危险性等级图(图 3).

分析上述过程可知，岩溶塌陷神经网络预测的误差主要有 2 个方面：(1)各预测单元评价因子取值的准确性；(2)预测模型本身的系统误差. 前者主要受数据采样点的空间分布特征与变异程度的影响，而小面积单元格的划分将有助于减小此种误差. 后者则与网络的训练情况相关. 若忽略第一种误差影响，根据模型校验结果可估算出预测区内评价单元定级的平均准确率为 95%，发生 1 级误判的面积为 1.46 km^2. 预测结果表明西城区内稳定区、基本稳定区、中等稳定区、不稳定区与极不稳定区的面积之比为 7.35∶3.42∶1.00∶0.37∶0.18，平均塌陷密度之比为 0.11∶0.52∶1.00∶1.17∶2.71，其中大部分的塌陷点分布在中等稳定区至极不稳定区的范围内，比较符合研究区的实际情况，对该区的发展规划和灾害治理有指导意义.

5 结论与建议

(1)BP 神经网络预测模型能够从大量塌陷样本的反复学习中主动调整自身的连接权值，实现系统参数的优化，从而获得不同评价因子组合到塌陷发生概率之间非线性映射关系，降低了预测过程中人为经验因素的影响，并可通过样本校验测试模型的预测精度. 这是其他塌陷预测模型难以具备的优点. 实践结果证明将神经网络技术应用于区域岩溶塌陷的预测是可行的，具有较高的可信度. 它为该问题的解决提供了一条新途径，同时也显示了 ANN 技术在处理复杂系统预测问题中的强大潜力.

(2)网络结构与学习参数是决定网络是否收敛及收敛速度的最重要因素. 通过多次试验的结果对比确定一个满意解是行之有效的解决方法.

(3)评价因子的选择及预测单元的因子取值关系到模型预测结果的准确性. 评价因子种类过多势必增加输入的噪音，不利于网络模型的判断；过少则不足以刻画岩溶塌陷受多因素共同影响的特征. 评价因子的确定应建立在充分理解塌陷形成过程的基础上. GIS 技术的运用提高了预测单元各评价因子数值获取的效率，缩短了预测评价的周期.

(4)由于资料的限制，本次研究中 2 个定性评价因子采用了相应的平均塌陷密度予以近似量化，这

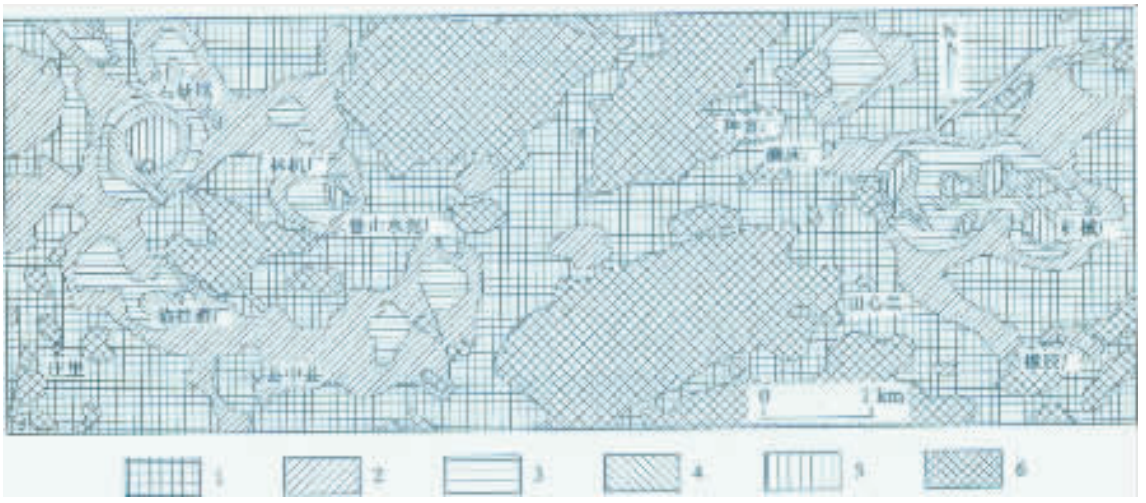

图 3　岩溶地面塌陷预测分区

Fig. 3　Zones of cover-collapse probability prediction

1. 稳定区;2. 基本稳定区;3. 中等稳定区;4. 不稳定区;5. 极不稳定区;6. 非评价区(基岩裸露区)

在一定程度上影响了预测精度. 若有勘查资料的支持,建议分别采用盖层的平均剪切强度与岩石的溶蚀率来代替.

参考文献:

[1] Goodings D J, Abdulla W A. Stability charts for predicting sinkholes in weakly cemented sand over karst limestone [J]. Engineering Geology, 2002, 65: 179－184.

[2] 陈星,黄润秋. 岩溶塌陷的地质概化模型[J]. 水文地质与工程地质,2002,(6): 30－34.
CHEN X, HUANG R Q. Geological conceptive models of karst collapse [J]. Hydrogeology and Engineering Geology, 2002, (6): 30－34.

[3] 张发旺,贾秀梅,赵华. 灰色统计方法及其在岩溶塌陷预测分析中的应用[J]. 河北地质学院学报,1996, 19(2): 144－150.
ZHANG F W, JIA X M, ZHAO H. Application of gray system statistic method in the field of karst collapse prediction analysis [J]. Journal of Hebei College of Geology, 1996, 19(2): 144－150.

[4] 陈学军,陈植华,陈先华,等. 桂林西城区岩溶塌陷模糊层次综合评价[J]. 桂林工学院学报,2000, 20(1): 112－116.
CHEN X J, CHEN Z H, CHEN X H, et al. Fuzzy-hierarchy prediction of karst collapse in the west area of Guilin city [J]. Journal of Guilin Institute of Technology, 2000, 20(1): 112－116.

[5] Kaufmann O, Quinif Y. Geohazard map of cover-collapse sinkholes in the "Torunaisis" area, southern Belgium [J]. Engineering Geology, 2002, 65: 117－124.

[6] 贺玉龙,杨立中,黄涛. 人工神经网络在岩溶塌陷预测中的应用研究[J]. 中国地质灾害与防治学报, 1999, 10(4): 86－90.
HE Y L, YANG L Z, HUANG T. Research on application of artificial neural network in the prediction of karst collapse [J]. Chinese Journal of Geological Hazard and Control, 1999, 10(4): 86－90.

[7] 包惠明,胡长顺. 岩溶地面塌陷神经网络预测[J]. 工程地质学报,2002, 10(3): 299－304.
BAO H M, HU C S. Neural network prediction of karstic ground collapse [J]. Journal of Engineering Geology, 2002, 10(3): 299－304.

[8] 陈守余,周梅春. 人工神经网络模拟实现与应用[M]. 武汉:中国地质大学出版社,2000. 1－36.
CHEN S Y, ZHOU M C. Realization and application of ANN models [M]. Wuhan: China University of Geosciences Press, 2000. 1－36.

[9] 阎平凡,张长水. 人工神经网络与模拟进化计算[M]. 北京:清华大学出版社,2000. 1－32.
YAN P F, ZHANG C S. ANN and simulating calculations of evolution [M]. Beijing: Tsinghua University Press, 2000. 1－32.

[10] 朱寿赠,周健红,陈学军. 桂林市西城区岩溶塌陷形成条件及主要影响因素[J]. 桂林工学院学报,2000, 20(2): 100－105.
ZHU S Z, ZHOU J H, CHEN X J. Analysis of forming conditions and main influential factors of karst collapse in west urban district, Guilin city [J]. Journal of Guilin Institute of Technology, 2000, 20(2): 100－

105.

[11] Chen X J, Chen Z H, Hu C. The dominant factors and spatial-temporal distribution of karst collapse in the west area of Guilin city [A]. Proceedings of international symposium on hydrogeology and the environment [C]. Beijing: China Environmental Science Press, 2000. 502—507.

[12] Cooley T. Geological and geotechnical context of cover collapse and subsidence in mid-continent US clay-mantled karst [J]. Environmental Geology, 2002, 42: 469—475.

ANN- and GIS-Based Regional Prediction of Cover-Collapse Probability: A Case Study in West Part of Guilin City

HU Cheng[1], CHEN Zhi-hua[1], CHEN Xue-jun[2]

(1. *Faculty of Engineering, China University of Geosciences, Wuhan* 430074, *China*; 2. *School of Civil Engineering, Guilin Institute of Technology, Guilin* 541001, *China*)

Abstract: Cover-collapse is one of frequent geological hazards in karst zone. Predicting the probability of cover-collapse occurrence is a requisite task for territorial planning, resource exploiting and hazard harnessing. Regional quantitative prediction of cover-collapse probability has become very intricate problem for the following reasons: (1) interactions between many influencing factors; (2) a sophisticated developing process; (3) difficulty associated with the value acquisition of factors. Some recent prediction models can not display the nonlinear characteristics of collapse development pattern, nor can they eliminate the impact of empiricism during the course of weights allocation. Three major characteristics of artifical neural network (ANN) technology, i. e. self-learning, self-adapting and nonlinear mapping, indicate a powerful application potential in collapse prediction field. This paper reports the methodology of developing an ANN model to predict cover-collapse occurrence probability. An approach was established to measure the relative probability of the collapse corresponding to certain factor combinations, and some impact factors were specified. Consequently, the structure of ANN prediction model was created. 292 stochastic collapse samples from the sample aggregate, of which the size was 312, were used to train the ANN model. The testing results of the 20 remaining samples show that this model has a good precision. Evaluation grids division and their value acquisition of every factor were accomplished with the aid of GIS software tools. The unstable probability of each grid was calculated through the trained ANN model, which enables us to delineate the different stable zones in the study area.

Key words: cover-collapse; artifical neural network (ANN); nonlinear; prediction model; GIS.