

文章编号:1000-582X(2002)12-0141-04

基于GIS和人工神经网络技术的开采沉陷预计建模方法*

曹丽文,姜振泉

(中国矿业大学资源与地球科学学院,江苏徐州221008)

摘要:探讨了GIS支持下利用人工神经网络技术进行开采沉陷定量预测的方法和思路,完成了影响因素的选取、数据处理、开采沉陷初始预计模型的建立及验证。数据由GIS软件完成,采用BP神经网络训练方法对开采沉陷系统进行建模;进而结合验证结果,对误差进行了定性分析。研究表明,利用GIS支持下的神经网络模型对复杂的开采沉陷系统进行模拟预测,具有理论上的可行性和现实意义,说明GIS和人工神经网络技术在开采沉陷预计领域中具有广阔应用前景。

关键词:开采沉陷;预计方法;地理信息系统;人工神经网络

中图分类号:TD82

文献标识码:A

目前,常见的煤矿开采沉陷预计方法有:实测分析法、理论模拟法和影响函数法^[1],基于地下岩体赋存环境的复杂性,其中最为实用的方法是实测分析法。由于地表移动变形值与各影响因素间组成的是一个典型的非线性系统,传统的数学方法难以建立较为精确的预计模型。

煤矿开采沉陷预计属于多因素问题,可以利用地理信息系统(GIS)的功能,实现开采沉陷影响因素的输入、管理和复合分析,为建立合理的开采沉陷预计模型提供数据依据。同时,通过GIS的动态数据更新功能,可进一步建立动态的预计系统。人工神经网络(ANN—Artificial Neural Network)在处理多因素的、复杂的非线性问题时具有明显的优点,它不需要对象具有精确的数学模型,从积累的工程实例中训练、学习,建立各影响因素间的高度非线性映射关系,因此,对于非线性系统中的数据具有较高的拟合能力及预测精度;另外,人工神经网络对于残缺不全或不确定信息具有较强的容错能力^[2-3]。这些特点非常适用于研究多因素的非线性开采沉陷系统,一定程度上,可以避免传统数学方法建立模型时所遇到的尴尬。从理论上讲,将GIS与人工神经网络技术相结合应用于煤矿开采沉陷预计,具有理论上的可行性和现实意义。因此,笔者探讨了基于GIS和人工神经网络技术的开采沉陷预计建模方法,并结合实例说明这一方法的可行性。

1 GIS的特点和功能

GIS(Geographic Information System)是采集、存储、管

理、分析和描述整个或部分地球表面与空间和地理分布有关数据的空间信息系统。GIS技术最突出的优点是能够完善地把地理信息或其它数据直观地在计算机上表达出来,并能够对数据进行综合分析和多重处理。但GIS技术作为支持系统,需要与其它技术相结合,才能很好地实现评价、预测功能。

2 人工神经网络建模方法的基本原理

2.1 理论基础

假设开采沉陷影响因素组成的集合为 $X = \{x_1, x_2, \dots, x_n\}$,其中 x_1, x_2, \dots, x_n 为各影响因素;地表移动变形值用集合 $Y = \{y_1, y_2, \dots, y_m\}$ 表示,其中 y_1, y_2, \dots, y_m 为各种移动变形值。 Y 与 X 之间的关系是多因素、非线性的。将影响因素 X 作为人工神经网络的输入神经元,移动变形值 Y 作为输出神经元,通过已知样本的学习、训练,建立开采沉陷的人工神经网络模型。

2.2 BP网络学习算法^[4-5]

研究选用BP(反向传播算法——Back Propagation Algorithm)算法建模,其网络结构模型由输入层、输出层和隐含层组成,隐含层可以有一层,也可以有数层。Robert Hecht-Nielsen(1989)证明了对于任何闭区间内的一个连续函数都可以用一个隐含层的BP网络来逼近,因而,一个3层的BP网络(如图1)可以完成任意的 n 维到 m 维的映照。

BP算法的训练步骤如下:

1) 网络初始化。将各连接权值 $\{W_{ij}\}$ 和阈值 $\{\theta_j\}$ 赋予 $(-1, +1)$ 之间的随机值;

* 收稿日期:2002-09-18

作者简介:曹丽文(1969-),女,黑龙江佳木斯人,中国矿业大学讲师,硕士研究生,主要从事煤矿地质工程研究。

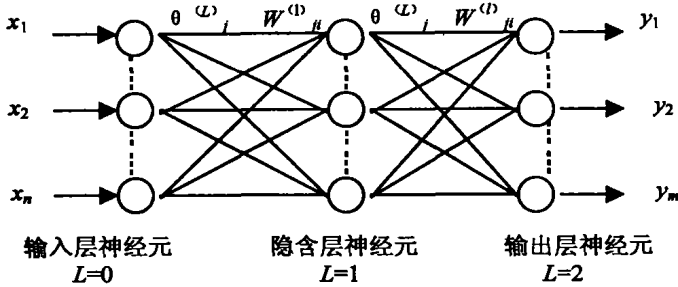


图1 3层BP网络模型结构

2) 提供训练集。输入 P 个学习样本, 即输入 $X_p = (x_1^p, x_2^p, \dots, x_n^p)^T$ 和期望输出向量 $Y_p = (y_1^p, y_2^p, \dots, y_m^p)$ 给网络, 其中 $P = 1, 2, \dots, P$;

3) 计算神经元的实际输出。对样本 P :

$$O_{pj}^{(L)} = f\left(\sum_i W_{ji}^{(L)} O_i^{(L-1)} - \theta_j^{(L)}\right)$$

式中 $O_{pj}^{(L)}$ 为第 L 层第 j 个神经元的输出; $W_{ji}^{(L)}$ 为第 $(L-1)$ 层第 i 个神经元与第 L 层第 j 个神经元的连接权值; $\theta_j^{(L)}$ 为第 L 层第 j 个神经元的阈值; f 为转换函数, $f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$;

4) 计算期望输出与实际输出的误差 E , $E = \frac{1}{2} \sum_j (y_j - Q_j)^2$;

5) 调整输入层、输出层、隐含层的连接权值 Δw 。按照 Levenberg - Marquardt 优化方法调整权值: $\Delta w = (J^T J + \mu I)^{-1} J^T e$

其中, J 为误差对权值微分的 Jacobian 矩阵, e 为误差向量, I 为单位矩阵, μ 为一个标量, 是自适应调整的;

6) 返回 2) 进行迭代, 直到误差平方和满足精度为止。

3 煤矿开采沉陷预计的 BP 网络模型

3.1 影响因素的选取和量化

3.1.1 影响因素的选取

煤矿开采过程中, 地表的移动和变形主要受控于地质和采矿两大因素, 其中地质因素包括第四系厚度、基岩厚度及岩性、层位、煤层倾角、地质构造等, 采矿因素有开采深度、开采厚度、巷道尺寸、工作面推进速度、重复采动及采煤方法和顶板管理方法等^[6]。研究中的因素选取主要考虑了以下 2 个方面的内容: 1) 基于 GIS 和人工神经网络的技术优势, 选取较其它开采沉陷预计方法多的影响因素; 2) 选取客观性强、容易进行科学量化的影响因素。

为此, 选取了煤层倾角、基岩厚度、松散层厚度、基岩岩性、相邻采区的影响、采区面积、开采深度和开采厚度、预测点与采区中心点间的位置关系、采煤方法和

顶板管理方法等作为开采沉陷建模的影响因素。这些因素中, 除了基岩岩性、相邻采区的影响、采煤方法和顶板管理方法外, 其它因素本身就是量化值, 可直接应用于 GIS 中, 从而提高了预计方法的客观性。

3.1.2 影响因素的量化

定性指标的量化方法如下: 基岩岩性以基岩的普氏硬度为量化指标; 相邻采区采动的影响包括许多因素, 为方便起见, 仅以相邻采区的面积为考虑因素, 用相邻采区影响系数(相邻采区的开采面积乘 0.000 01)表示; 对于采煤方法和顶板管理方法, 由于没有相关的量化方法, 根据采煤方法和顶板管理方法对地表移动变形影响的定性描述, 将综合机械化放顶煤开采量化为 10.0, 综机仰斜开采、全部陷落法管理顶板量化为 0.5; 预测点与采区中心点间的位置关系用 2 个参数表示。1) 以工作面长边在地表的投影为 X 轴, 以开采方向的反方向为 X 轴正向, 以采区中心点为中心, 建立直角坐标系, 观测点与 X 轴正向的夹角为观测点位置夹角; 2) 观测点与工作面在水平面的投影图上, 观测点与工作面中心点间的距离。

因此, 选取的影响因素有 11 个, 它们是观测点与倾线夹角、观测点与工作面中心点的距离、基岩厚度、第四系厚度、煤层倾角、采深与采厚之比、基岩岩性、观测点在采区的内外位置(在采区地表投影的内部时, 取 1.0, 否则取 -1.0)、开采面积、相邻采区影响系数、采煤方法和顶板管理方法。

3.2 网络结构模型的建立和训练

以兖州矿业集团兴隆庄煤矿(综合机械化倾斜分层开采、全部陷落法管理顶板方法)5306 工作面为建模区, 以 4314 工作面为模型验证区, 为研究初始预计模型对于不同井田、不同采煤方法的预测效果, 将南屯煤矿 3303 - 1 工作面(综机仰斜开采、全部陷落法管理顶板)作为另一个模型验证区, 建立 BP 神经网络预计模型并验证。

利用 GIS 的应用软件——ARC/INFO 建立 11 个影响因素的专题信息层, 生成开采沉陷综合信息层。用于 BP 网络训练建模的数据皆来自于综合信息层的 TABLES 数据文件。人工神经网络的输入层有 11 个神经元; 输出层设有 2 个神经元, 分别为下沉值和水平移动值; 根据网络的收敛速度和验证精度的训练体会, 确定隐含层神经元数目为 46。因此, 神经网络结构为 11 - 46 - 2 型。共 103 个样本, 均匀分布在 5306 工作面走向和倾向观测线上(样本的 11 种影响因素量化数据略)。

用 MATLAB 语言编制训练程序。经训练, 误差平方和为 0.000 204 387, 满足预先设定的误差平方和小于 0.000 3 的精度要求。部分网络输出结果见表 1。

表 1 5306 工作面样本实测值与网络输出值对比表

样本编号	观测值 /mm		网络输出值 /mm		相对误差 /%	
	W_0 (下沉)	U_0 (水平移动)	W (下沉)	U (水平移动)	$ W_0 - W / W_0$	$ U_0 - U / U_0$
1	-0.3	-1.0	-0.1	0.1	66.667	110.000
2	-1.3	-1.7	-2.0	-3.4	53.846	100.000
3	-15.0	-37.5	-14.6	-36.9	2.667	1.600
4	-102.7	-199.5	-100.6	-197.5	2.045	1.003
5	-218.9	-369.5	-221.5	-371.4	1.188	0.514
6	-433.2	-607.3	-431.1	-606.8	0.485	0.082
7	-744.3	-863.0	-747.7	-861.6	0.457	0.162
8	-1 167.3	-1 060.0	-1 164.9	-1 063.4	0.206	0.321
9	-1 646.0	-1 145.0	-1 644.8	-1 140.6	0.073	0.384
10	-2 124.7	-1 060.0	-2 126.7	-1 063.4	0.094	0.321
11	-2 858.8	-607.3	-2 858.7	-607.5	0.003	0.033
12	-3 037.1	-369.5	-3 042.5	-370.9	0.178	0.379
13	-3 189.3	-199.5	-3 183.7	-197.5	0.176	1.003
14	-3 252.2	-89.9	-3 255.3	-91.4	0.095	1.669
15	-3 277.2	-37.5	-3 280.2	-37.3	0.092	0.533
16	-3 287.4	-12.7	-3 284.0	-12.2	0.103	3.937
17	-3 292.0	1.0	-3 292.1	0.8	0.003	20.000
18	-3 073.1	369.5	-3 071.4	370.0	0.055	0.135
19	-2 547.7	-863.0	-2 547.1	-861.6	0.023	0.162
20	-1 846.0	1 198.6	-1 845.8	1 197.9	0.011	0.058
21	-1 267.0	1 303.0	-1 267.4	1 301.8	0.032	0.092
22	-944.0	1 198.2	-944.6	1 198.4	0.064	0.017
23	-1 167.3	1 060.0	-1 167.9	1 060.9	0.051	0.085
24	-633.0	1 000.1	-633.2	1 000.1	0.032	0.000
25	-418.0	843.1	-418.3	843.5	0.072	0.047
26	-139.0	350.0	-140.1	351.2	0.791	0.343
27	-45.0	179.5	-46.2	180.9	2.667	0.779
28	-24.0	79.0	-24.7	80.6	2.917	2.025
29	-10.0	29.9	-10.7	29.2	7.000	2.341
30	-3.3	13.2	-3.9	14.0	18.182	6.061

从表 1 的数据可以看出,神经网络模型输出结果与实际观测值十分吻合,说明经过训练后,网络最终确定的各层神经元之间的连接权值和各神经元的阈值是合理的。这样,确定了开采沉陷的 BP 神经网络初始模型(稳定的网络结构、连接权值和阈值)。

3.3 网络模型的验证与结果分析

3.3.1 网络模型的验证

将兴隆庄煤矿 4314 工作面和南屯煤矿 3303-1 工作面的未知观测点作为输入向量,赋予 BP 网络模型,

根据模型的连接权值和阈值,按照公式

$$O_{ij}^{(L)} = f(\sum_i W_{ji}^{(L)} O_i^{(L-1)} - \theta_j^{(L)}),$$

由计算机计算网络输出值(网络预测结果略)。为方便比较,将两工作面的实测值与网络输出结果绘制成图,见图 2、图 3、图 4 和图 5。(以下图中,“—”为实际观测值,“★”为网络输出值,纵轴 U 为开采沉陷水平移动值,横轴 L 为水平投影面上观测点与工作面中心点间的距离,单位:m)

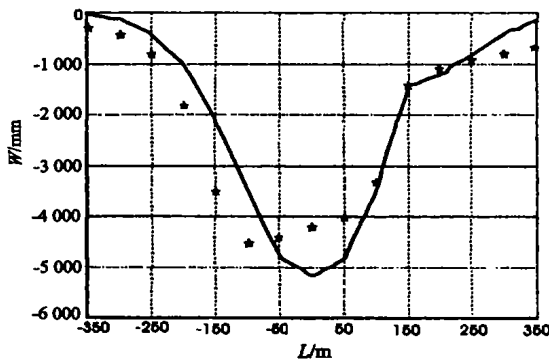


图 2 4314 工作面开采沉陷下沉值对比

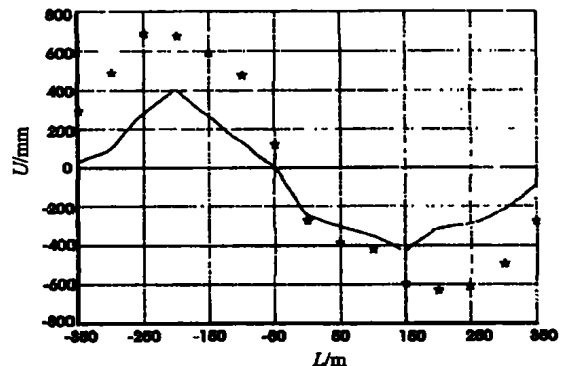


图 3 4314 工作面开采沉陷水平移动值对比

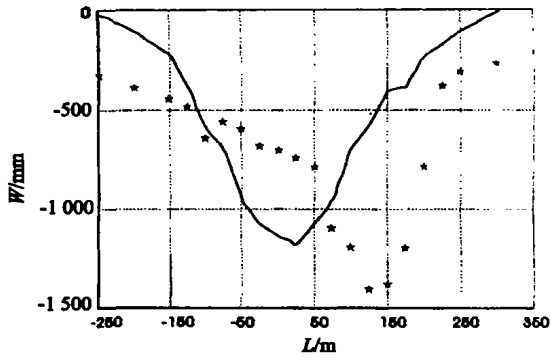


图4 3303-1工作面开采沉陷下沉值对比

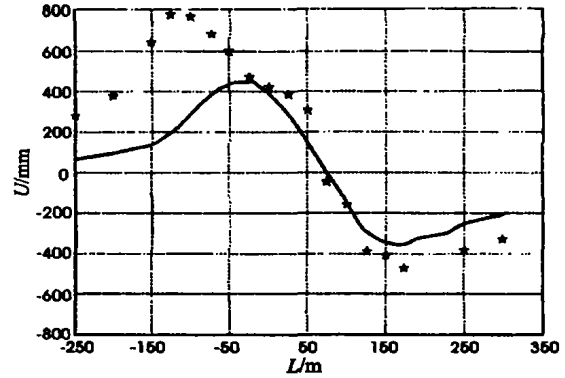


图5 3303-1工作面开采沉陷水平移动值对比

由以上对比图可以看出,网络输出结果基本反映了地表移动变形的变化趋势,说明GIS与人工神经网络技术相结合可以作为煤矿开采沉陷预测的一种行之有效的办法。

3.3.2 验证结果分析

2个工作面的验证结果基本反映了开采沉陷的下沉规律,但总体误差较大。经分析,产生误差的主要原因如下:

1) 部分影响因素有待于用更为严谨的科学方法进行量化,如相邻采区的影响、采煤方法和顶板管理方法等。

2) 样本的选取和隐层单元数的确定问题。本次研究的样本仅来自于兴隆庄煤矿的5306工作面,因此,样本的代表性不高;虽然考虑了样本在工作面上的均匀分布性,但由于观测资料有限,只把分布在走向和倾向两个方向的观测点作为样本使用。样本数量有限、均匀性不佳,影响了模型的代表性;另外,本次研究中的隐层单元数是根据多次训练实验确定的,缺乏理论依据。

与建模区处于同一井田的4314工作面验证精度略高于3303-1工作面,说明除了训练样本的均匀分

布特性外,样本的代表性对预测精度也有重要影响。

4 结论

将GIS和人工神经网络技术相结合应用于煤矿开采沉陷预计研究是可行的。该方法在一定程度上丰富了开采沉陷预计理论。

参考文献:

- [1] 吴佩,葛家新,王玲丁,等.开采沉陷预计一体化方法[M].徐州:中国矿业大学出版社,1998.
- [2] 张立明.人工神经网络的模型及其应用[M].上海:复旦大学出版社,1993.
- [3] 周翔,朱学愚.神经网络在岩溶水资源评价中的应用[J].中国岩溶,1999,18(4):337-341.
- [4] 贺玉龙,杨立中,黄涛.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):86-76.
- [5] 施鸿宝.神经网络及其应用[M].西安:西安交通大学出版社,1993.
- [6] 何国清,杨伦,凌虞娣,等.矿山开采沉陷学[M].徐州:中国矿业大学出版社,1991.

Prediction Model of Exploitation Sink Using Ann and GIS

CAO Li-wen, JIANG Zhen-quan

(College of Mineral Resource and Geoscience, China University of Mining and Technology, Xuzhou 221008, China)

Abstract: A new idea is put forward on researches of prediction method of coal mining subsidence supported by GIS, and a new method is given about quantitative prediction of mining subsidence by means of GIS and ANN (Artificial Neural Network). This paper has completed a lot of work concerning choosing factor, data processing, establishing and validating preliminary ANN prediction model. Further more, data processing is carried out by GIS software, and the BP training method is used for modeling the exploitation sink system. Subsequently, the error is qualitatively analyzed with considering the result of verification. Researches show that the ANN prediction model supported by GIS has theoretical feasibility and realistic significance in predicting complex exploitation sink system, and GIS and ANN possess wide application prospects in the prediction of exploitation sink.

Key words: exploitation sink; prediction method; GIS; ANN (Artificial Neural Network)

(责任编辑 姚 飞)