

ANN 技术在地下水含水量预测建模中的研究与应用

宋洪伟¹ 刘继朝¹ 石建省¹ 张翼龙¹ 夏 凡² 苗青壮¹

(1. 中国地质科学院水文地质环境地质研究所, 河北 石家庄 050061; 2. 石家庄经济学院, 河北 石家庄 050031)

摘要:将人工神经网络(ANN)技术引入到地下水含水量预测工作,以华北平原和河套平原为试验场,以若干已知钻孔为验证,采用激电和电阻率测深等地面物探方法获取视电阻率 ρ_s 、视极化率 η_s 、半衰时 T_h 、衰减度 D 和偏离度 σ 等参数为输入神经元对单孔单位涌水量建立人工神经网络预测模型。同时,为消除不同地区矿化度的影响,通过实验对比引入综合参数 T^* ,改良了输入神经元的配比。最终建立以半衰时 T_h 、衰减度 D 、偏离度 σ 和综合参数 T^* 为输入神经元的含水量预测模型,进一步提高了预测精度。通过检验,发现所建立的模型对平原地区进行含水量的定量预测有着较好的效果,为含水量预测工作研究与发展带来了新理念、打开了新思路。

关键词: ANN; 涌水量; 矿化度; 偏离度; 预测模型

中图分类号: P641.7 **文献标志码:** A **文章编号:** 1000-3657(2012)04-1081-06

激发极化电测深法以其简单、快捷、效率高和成本较低等优势成为被广大水文地质工作者所认可的一种最为广泛且切实可行的水文地质勘查方法。其不受或较少受地形起伏及围岩电阻率不均匀性的影响,可充分利用其时间域(或频率域)探测方式对地下水进行间接探测^[1-2]。而人工神经网络(以下简称 ANN)技术的引入,克服了以往回归预测模型的不足,建立了电法异常参数与地下水富集程度间的量化关系,把定性的解释转变到定量的计算上来,使得激电法找水不仅限于利用其测量参数的异常值来圈定地下水赋存范围,且可以在一定程度上定量地解释出地层含水量的大小^[3-5]。传统的 ANN 含水量预测模型是以电阻率、视极化率、半衰时、衰减度等参数为输入神经元。参数的获取容易受到外界干扰而出现异常值,导致预测工作出现偏差或预测精度不够稳定,难以达到对未知水文地质单元进行精确评价的目的,推广性难遂人愿。为克服诸如矿化度等外

界因素的影响将含水层厚度和偏离度等参数引入,进一步强化模型的鲁棒性和稳定性,同时使得 ANN 模型的预测精度得到大幅提高。利用河套平原测区和华北平原测区内已知水文地质资料的水井为训练集,以通过井旁电测深方式获取前文所提的电法异常参数为输入神经元,以单孔单位涌水量为输出神经元建立了 ANN 含水量预测模型,得到了令人满意的结果^[6-10]。

1 BP 神经网络地下含水量预测模型

1.1 BP 神经网络简介

人工神经网络(简称神经网络)是由大量简单的神经元互相连接,通过模拟人的大脑神经处理信息的方式,进行信息处理和非线性转换的复杂网络系统。BP 神经网络是一种单向传播的多层前向人工神经网络,具有 3 层或 3 层以上结构,包括输入层、隐

收稿日期:2011-12-11;改回日期:2012-03-11

基金项目: 科研基金项目“太行山严重缺水地区含水结构物性特征研究”(SK201203);中国地质调查局地调项目“华北平原典型地区含水层结构调查”(G201103)和“河套平原地下水资源及其环境问题调查评价”(G09001)共同资助。

作者简介: 宋洪伟,男,硕士,地球物理探测专业;E-mail:shw2180@163.com。

含层(中间层)和输出层。此网络的中心思想是通过一边向后传播误差,一边修正误差的方法来调节网络参数,以实现或逼近所希望的输入输出映射关系。可见,BP神经网络对地下水量预测有很好的对持久性和适时预报性^[11-12]。

1.2 输入输出神经元的选择

由于BP神经网络有着不需要建立精确的数学模型,就能够实现从输入空间到输出空间的非线性映射的优点,通过试验、统计的方法利用多元线性(非线性)回归模型,确定地下含水层电阻率 ρ ,极化率 η 、衰减度 D 、半衰时 T_h 为预测模型的输入神经元,单孔涌水量 Q 为预测量即输出元。通过预测出单孔涌水量在测区内的分布进行地下富水性评价。

1.3 模型结构

模型结构选择人工神经网络中使用最多、最普遍的3层BP网络结构,其中包括输入层、隐含层和输出层,还有基函数、传递函数和训练函数。

模型的基函数采用线性函数,输出 T 为输入 P 与阈值 θ 的加权和,传递函数根据模型的运行情况而定选择不同的函数。模型的结构如图1所示,输入元 $P_i(i=1,2,\dots)$ 分别对应电阻率、极化率、衰减度、半衰时的集合,输出元 T 为涌水量 Q 的集合。

模型均以矩阵形式进行输入输出如下式:

$$\text{输入 } P = \begin{bmatrix} P_{11} & P_{1n} \\ P_{21} & P_{2n} \\ P_{31} \cdots P_{3n} \\ P_{41} & P_{4n} \end{bmatrix},$$

$$\text{输出 } T = [Q_1 \quad Q_2 \cdots Q_n], (n=1,2,\dots)$$

输入神经元矩阵 P 中每一列数值为分别对应输出神经元矩阵 T 中的各列 Q (涌水量)值^[13-14]。

2 传统的 ANN 地下水含水量预测模型的应用

2.1 实验区的选取

由于模型的建立主要立足于平原区沉积地层,而内蒙古河套平原和河北衡水所在的华北平原分布范围较广,地质条件较为多样化,是很好的综合性试验场地。因此,建模训练集选定为内蒙古河套平原巴彦淖尔市杭锦后旗、五原县和河北衡水市3处水资源地共16口已知已进行抽水试验的水文孔,开展了井旁激电和电阻率测深法,对所观测到的电测深曲线进行了反演解释,计算出含水层对应的电阻率和

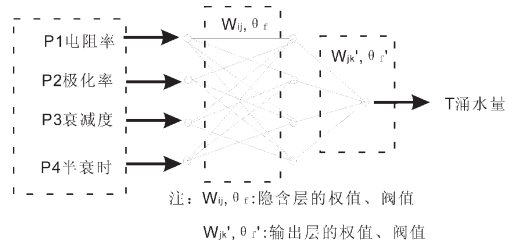


图1 神经网络模型

Fig.1 ANN model

激电参数进行统计分析。然后利用所得数据以循环验证的方式进行建模,最后利用已掌握的地处河套平原区的乌拉特前旗3眼和衡水实验区3眼已知抽水井作为测试集进行模型的推广能力测试。

2.1.1 五原实验区概况

五原地处内蒙古河套平原腹地,南临黄河,属黄河最北端,北有阴山横亘。实验区在大地构造单元上,属阴山天山纬向构造带,并受新华夏系构造的影响,形成内陆断陷盆地,为第四纪松散的地层所覆盖,沉积了较厚的湖相地层。上部是冲积、风积层,主要岩性为细砂、粉砂和砂粘土互层。砂层层理清晰,厚度10~70 m。中部为河湖交替层,主要岩性为淤泥质、粉砂与粘土互层。下部为巨厚的新老第四纪湖相沉积层,主要岩性为淤泥质砂粘土,局部土地呈盐碱化。

2.1.2 杭锦后旗实验区概况

杭锦后旗地处黄河河套冲积平原,地势平坦,西南略高,东北偏低,海拔1 040 m左右。地貌主要由冲击平原、洪积平原和河漫滩三种地形构成。地质条件与五原县类似。

2.1.3 衡水实验区

区内地层,自上而下依次为第四系、古近—新近系及古生界和元古宙地层。实验区地处河北冲积平原,地势自西南向东北缓慢倾斜,海拔高度12~30 m。地面坡降,滏阳河以东在1/8 000~1/10 000,以西为1/4000。境内河流较多,由于河流泛滥和改道,沉积物交错分布,形成许多缓岗、微斜平地 and 低洼地。

2.2 归一化工作

由于作为输入神经元而采集的各数据单位不一致,各种数据大小有时相差很大,会出现大样本吃小样本的情况,因而必须对输入层数据进行^[10,11]归一化处理。纵观此次数据采集的情况采取了线性函数转换,表达式如下:

$$y = (x - \text{Min}(x)) / (\text{Max}(x) - \text{Min}(x)) \quad (1)$$

式中: x 、 y 分别为转换前、后的值, $Max(x)$ 、 $Min(x)$ 分别为样本的最大值和最小值。

归一化数据见表 1 (篇幅所限略去原始数据), 由于输出形式为一元输出且数量级变化不大, 因此无需对输出神经元采取归一化处理。

2.3 隐含层的神经元个数

隐含层的神经元个数与输入输出神经元个数有关, 可参考经验公式如下:

$$\gamma = \sqrt{mn} + a \tag{2}$$

其中 γ 为隐含层神经元个数, m 、 n 分别为输入输出层神经元个数, a 为任意 1 到 10 的整数。为确保模型的稳定性和灵活性, 仅将经验公式做为参考, 而采取多种隐含层神经元个数的循环选择方式来确定其最佳个数。

2.4 预测结果

利用以上调试完成的模型, 确定传递函数采用 S 型正切函数 tansig 函数和线性函数 purelin 函数。训练函数采用 trainlm 函数, 利用全值最优化算法 Levenberg-Marquard 方法对输入数据和目标向量进行训练, 此种算法特点是训练速度快预测精度较高。对乌拉特前旗及衡水试验场 6 眼已知抽水井进行测试, 得到预测结果如表 2 所示, 可以看到, 误差结果差强人意。

3 改进之后的 ANN 地下含水量预测模型

3.1 输入神经元的分析

在模型的应用和进一步的研究工作中, 发现在不同的地质条件下, 外界因素的影响在一系列电性参数的获取工作中是不可忽视的。譬如说, 地下水的

表 1 输入神经元归一化数据
Table 1 Normalized data of input neuron

井号	电阻率 / ($\Omega \cdot m$)	半衰时/s	衰减度	极化率/%	涌水量 / m^3/h
WY04	0.879	0.579	0.840	0.605	32
WY05	0	0.263	0.640	0	48
WY06	0.423	0.165	0.670	0.402	53
WY07	0.312	0.330	0.768	0.079	60
WY09	0.665	0.611	0.600	0.582	40
HS 01	0.303	0	0.532	0.020	38
HS 02	0.223	0.234	0.469	0.587	45
HS 04	0.456	0.372	0.315	0.758	52
HS05	0.891	0.363	0.657	0.973	56
HS06	0.773	0.223	0.703	0.690	35
HJ03	0.110	0.788	0.554	0.980	55
HJ07	1	1	0.760	1	52
HJ 10	0.992	0.979	0.220	0.365	67
HJ 11	0.450	0.120	0.315	0.240	77
HJ 12	0	0.070	0.350	0.082	85
HJ 13	0.912	0.573	0.830	0.616	56

含卤度对视极化率和视电阻率的影响; 不良的接地效果对电参数的影响。因此, 经研究分析引入了偏离度和综合参数的概念, 详细分析如下。

3.1.1 偏离度参数

经大量实际工作和研究表明, 在直流电激电法

表 2 训练结果误差验证对照 (输入量均为归一化后数据)
Table 2 Comparison of training errors (all input data has been normalized)

井号	电阻率/ $\Omega \cdot m$	半衰时/s	衰减度/%	极化率/%	实际涌水量/ m^3/h	预测涌水量/ m^3/h
WQ02	0.879	0.689	0.760	0.635	67	63.8165
WQ03	0.333	0.433	0.670	0.675	38	35.5224
WQ05	0.110	0.126	0.730	0.202	20	17.1222
HS03	0.468	0.450	0.589	0.579	60	55.5242
HS07	0.934	0.786	0.620	0.872	83	85.8782
HS08	0.235	0.356	0.332	0.520	55	51.2894

工作中,实测的放电曲线与“理想”直线之间总会有一些偏差,所谓偏离度,系指实测放电曲线与直线方程的拟合误差,我们用均方相对误差 σ 表示,并称其为偏离度。公式如下:

$$\sigma = \frac{1}{\bar{\eta}_i} \sqrt{\frac{\sum_{i=1}^n (\eta_i + K \log t_i - B)^2}{n}} \times 100\% \quad (3)$$

式中, n 为取样点数, $\bar{\eta}_i = \frac{1}{n} \sum_{i=1}^n \eta_i$ 为观测时间段内各取样点极化率的平均值偏离度,意即偏离于“理想”直线的程度。显然, σ 值越小,实测放电曲线的“直线性”越强。通过大量的研究和资料分析,可以得知:偏离度(σ)值与供电时间、地层的湿度、孔隙液浓度及成分还有粘土含量等影响因素均有一定的物理关系。对极化率及偏离度进行统计分析如表 3 所示,可以清楚看到偏离度比极化率受到外界的影响更小,与涌水量的相关度更高,同样也使模型更加稳定。

3.1.2 综合参数—— T''

考虑到电测深反演解释过程中存在的等值现象,鉴于第四纪含水层一般呈高阻性,易产生 T 等值现象的解释误差;再者一定成分、结构的含水层对应于一定的电阻率值,但其单孔单位涌水量既与含水层成分、结构有关,又与其厚度有关。为了尽可能消除因解释所产生的误差,又能使输入神经元与预测测量有更为密切的相关性,使预测模型具有良好的推广能力,为此,将电测深反演后的含水层电阻率与其层厚度相乘作为一个输入神经元。考虑到不同的地区地下水所含矿化度的不同,因其孔隙水的导电性不同,将会导致同类富水层电阻率有较大差别。为了消除孔隙水的导电性对预测精度的影响,突出含水层有效孔隙度特征,基于阿尔奇公式,引入了综合参数,其表达式为:

$$T'' = \frac{T'}{\rho_{\text{隔}}} = \frac{\rho}{\rho_{\text{隔}}} H \quad (4)$$

表 3 电性参数与涌水量相关分析

Table 3 Analyses of electrical parameters specific capacity

神经元	σ	η
相关系数 (r)	0.970	0.856
显著性系数 (p)	0.987	0.799

式中: ρ ——含水层电阻率;
 $\rho_{\text{隔}}$ ——测点处隔水层电阻率;
 H ——含水层厚度。

从而将原基本输入神经元 ρ 和 H 组合为一个输入神经元 T'' ,使其与含水层单孔单位涌水量有更密切的相关性(表 4)。

3.2 改进后模型应用效果分析

改进后模型以半衰时 T_h 、衰减度 D 、偏离度 σ 和综合参数 T'' 为输入神经元,以单孔单位涌水量为输出神经元。同样以乌拉特前旗 3 眼及衡水 3 眼已知抽水井进行测试,得到预测收敛曲线如图 2 所示,可以看到在第 103 步时收敛达到误差目标,模型停止了工作。预测结果如表 5,可以看到预测精度得到大幅提升,证明新模型的改进工作是令人满意的。

4 改进前后模型预测效果对比

将两种方式所建模型同样是对 6 眼抽水井的预测结果对比如表 6 所示,可以明显看出改进了输入神经元的新模型较老模型预测精度有了大幅提高,由之

表 4 综合参数与涌水量相关分析

Table 4 Analyses of comprehensive parameters and specific capacity

神经元	ρ	H	T''
相关系数 (r)	0.881	0.826	0.946
显著性系数 (p)	0.996	0.989	0.999

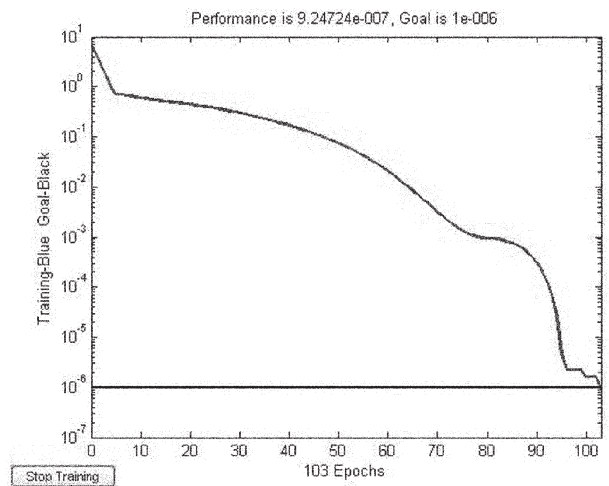


图 2 误差收敛曲线

Fig.2 Errors converged curve

表 5 新模型预测结果

Table 5 Forecast results of the new model

井号	实际涌水量 / (m ³ /h)	预测涌水量 / (m ³ /h)
WQ02		
WQ03	38	37.3230
WQ05	20	21.3823
HS03	60	56.6214
HS07	83	82.5573
HS08	55	52.3368

表 6 两种模型预测效果对比

Table 6 Comparison of the forecasting effects of the two models

井号	原模型预测误差/%	新模型预测误差/%
WQ02	4.75	1.96
WQ03	6.52	1.78
WQ05	14.39	6.91
HS03	7.46	5.63
HS07	3.47	0.53
HS08	6.75	4.84

前的误差在 3%~14% 基本降低到 0.5%~7% 之间。

5 结论

(1) BP 神经网络的中心思想是通过一边向后传播误差,一边修正误差的方法来调节网络参数,以实现或逼近所希望的输入输出映射关系。通过研究可见,BP 神经网络对地下水量预测有很好的对持久性和适时预报性。

(2) 经过以上的验证和推广性试验成功的对测区地下水的涌水量进行了量化预测,根据河套平原和华北平原水资源地内的预测检验结果表明,加入了新参数的模型具有更好的推广能力。新的模型有效抑制了矿化度因素的影响,使得推广能力得到大幅提高。

(3) 目前,课题的研究目前在相对地质地层结构较为简单的平原区的含水量预测工作中还是有较高的精度。而在基岩区,由于含水量受地质构造、岩性及汇水条件的影响较大。例如,仅通过电性参数难

以区分充填在基岩破碎带中以吸附水为主的粘土或泥页岩碎屑和重力水,因此单孔涌水量预测精度难以把控。

(4) 最后,应该指出的是,由于物探工作的多解性,仅仅利用一两种的物探方法进行预测建模仍然有欠缺。今后的发展应该更加多样化,将更多的物探方法引入,获取更多的有一定相关度的参量,使模型更加具有推广性。

参考文献 (References):

- [1] 李金铭. 地电场与电法勘探[M]. 北京:地质出版社,2005:208-215.
Li Jinming. Electric Field and Electrical Prospecting [M]. Beijing: Geological Publishing House, 2005:208-215 (in Chinese)
- [2] 刘国辉, 张献民, 贾学民. 地下水资源电法勘探新技术[M]. 北京:地震出版社, 2007:60-64.
Liu Guohui, Zhang Xianmin, Jia Xuemin. New Technology of Groundwater Resources Electrical Prospecting [M]. Beijing: Seismological Press, 2007:60-64(in Chinese).
- [3] 凌成鹏, 孙亚军, 杨兰和, 等. 基于 BP 神经网络的孔隙充水矿井涌水量预测 [J]. 水文地质工程地质. 2007, (5):55-58.
Ling Chengpeng, Sun Yajun, Yang Lanhe, et al. Prediction of inrush water of mine pore water yield based on BP artificial neural network[J]. Hydrogeology and Engineering Geology, 2007 (5):(in Chinese with English abstract).
- [4] Dong YZ, Ma YK et al. Pattern recognition of the characteristics of AE source using neural network [J]. Proc. Of 14th WCNDT, New Delhi: 1996.
- [5] R.O.Hansen.Feature recognition from Potential fields using neural networks [J]. Geophysical, 1997(3).
- [6] 李金铭. 激电找水的新参数——偏离度 [J]. 勘察科学技术, 1993, (6):52-56.
Li Jinming. The New Parameters of IP for Ground Water Investigation—Deviation [J]. Investigation Technology, 1993(6): 52-56. (in Chinese with English abstract).
- [7] McDonald.G, Harbaugh A W. A modular three - dimensional finite - difference groundwater flow model[M]. On - trio: Waterloo Hydro geologic Inc, 1998.
- [8] 杨进, 孟海东, 傅良魁. 激电找水数据解释方法与系统 [J]. 物探与化探, 1998, 22(3):204-210.
Yang Jin, Meng Haidong, Fu Liangkui. Interpretation method and system of IP data processing in groundwater exploration[J]. Geophysical and Geochemical Exploration, 1998, 22(3):204-210 (in Chinese with English abstract).
- [9] Khanna. Function of Neural Networks[J]. Addhun-Wesley, 1990.
- [10] Scanlon. Uncertainties in estimating water fluxes and residence times using environmental tracers in an arid unsaturated zone [J]. Water Resource. Res.,2000,36:395-409.
- [11] 葛哲学, 孙志强. 神经网络理论与 MATLAB2007 实现 [M]. 北

- 京:电子工业出版社, 2005:58-78.
- Ge Zhexue, Sun Zhiqiang. Theory and the Realization of MATLAB7 Neural Network [M]. Beijing:Electronic Industry Press, 2005:58-78(in Chinese).
- [12] Bear. Hydraulics of Groundwater [M]. New York:McGraw-Hill, 1979.
- [13] 韩万林, 张幼蒂, 李梁, 等. 地震多参数 BP 神经网络预测煤层厚度 [J]. 煤田地质与勘探, 2001, 29(4):53-54.
- Han Wanlin, Zhang Youdi, Li Liang, et al. Earthquake parameters of BP ANN to predict seam thickness [J]. Coal Geology and Exploration, 2001, 29(4):53-54(in Chinese with English abstract).
- [14] Hayashi Y, Buckley Czogala E. Fuzzy neural networks with fuzzy signals and weights [J]. IEEE Transom Neural Networks, 1992, 3 (5):673-678.

The study and application of ANN to the Modeling of underground water content forecast

SONG Hong-wei, SHI Jian-sheng, LIU Ji-chao,
ZHANG Yi-long, XIA Fan, MIAO Qing-zhuang

(1. Institute of Hydrogeology and Environmental Geology, CAGS, Shijiazhuang 050061, Hebei, China; 2. Shijiazhuang University of Economics, Shijiazhuang 050031, Hebei, China)

Abstract: This paper has introduced the technology of artificial network into the modeling of underground water content forecast. Hetao plain and Huabei plain were chosen as the testing ground with a number of known local agro-wells as the verification sites. Induced Polarization (IP) and resistivity sounding and other surface geophysical methods were used to construct the artificial neural network (ANN) model based on such parameters as apparent resistivity, polarization rate, half-life, decay rate and rate of deviation in the relevance. Then, the comprehensive parameter was added to improve the inputting neurons. Finally, the quantitative prediction model of the water content was established. The results of mean-variance test show that this technique has a good effect in the plain area. The study has provided a new concept and a new idea for the forecasting work in hydrogeological exploration.

Key words: ANN; water inflow; mineralization; deviation; forecast model

About the first author: SONG Hong-wei, male, born in 1983, master, assistant researcher, majors in geophysical exploration; E-mail: shw2180@163.com.