

# 基于改进 BP 神经网络的地下水水质评价<sup>\*</sup>

## ——以新疆和田地区为例

付永锋<sup>1</sup>, 张 建<sup>2</sup>, 罗光明<sup>3</sup>, 赵基花<sup>1</sup>, 沈 冰<sup>1</sup>

(1 西安理工大学 水利水电学院, 陕西 西安 710048;

2 西安交通大学 计算机教学中心, 陕西 西安 710032;

3 新疆和田地区水文水资源勘测局, 新疆 和田 848000)

[摘 要] 在阐述神经网络基本原理的基础上, 提出了地下水水质评价的 BP 网络模型。并分别利用 BP 网络模型与传统的模糊综合评判法对新疆和田地区地下水水质进行了评价, 结果表明, BP 网络模型不仅计算简便, 而且具有较高的计算精度。

[关键词] 神经网络; BP 网络; 地下水; 水质评价; 新疆和田

[中图分类号] X824

[文献标识码] A

[文章编号] 1671-9387(2004)11-0129-04

地下水水质评价是地下水资源评价的一项重要内容, 可以根据地下水的主要成分和给定的水质标准, 分析地下水水质状况及其时空分布, 为地下水资源的开发利用、规划和管理提供科学依据。由于地下水分级标准具有模糊性和灰色性, 模糊数学和灰色系统理论方法在地下水水质评价中得到了广泛的应用<sup>[1]</sup>, 但其在处理实际问题时都需要设计若干不同的效用函数(灰色系统的白化函数、模糊数学的隶属函数等), 以及人为给定各评价指标的权重或权函数等。这些效用函数和指标权重的给定由于受人因素为因素的影响, 造成评价模式难以通用, 增加了应用的困难和人为臆断因素对结论的影响<sup>[2]</sup>。事实上, 水质评价过程既是确定待评价水体水质属于那一级国家标准水质, 而且国家水环境质量标准都是给定模式, 因此水质评价是典型的模式识别问题<sup>[3,4]</sup>。神经网络的发展为解决上述问题提供了有效的方法。

人工神经网络(Artificial Neural Network)是模拟人脑的一些基本特性, 因其具有大规模并行处理、分布式储存、自适应性、容错性等优点而广泛应用于模式识别、智能控制、知识处理和预测等众多领域<sup>[5,6]</sup>。本研究采用一种改进的 BP 网络方法, 模拟地下水水质评价因子与水质级别间的非线性关系, 对新疆和田地区地下水水质进行了评价, 为当地水资源的科学利用提供依据。

## 1 BP 网络模型

在众多的神经网络模型中, 原理最简单、应用最广泛的是基于误差反向传播算法的 BP 网络模型。它是一种多层前馈神经网络模型, 由一个输入层、一个输出层和一个或若干个隐含层组成。每一层都包含若干神经元, 神经元之间通过连接权重互连, 但同层的神经元之间没有联系。3 层 BP 网络的结构如图 1 所示。

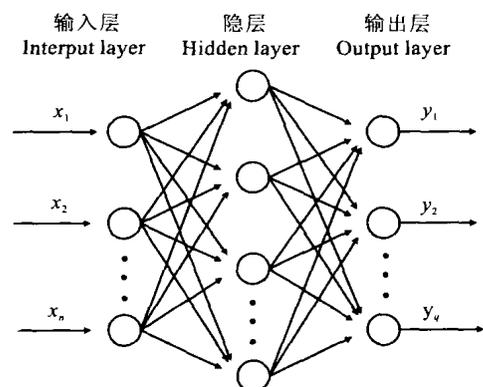


图 1 BP 神经网络结构图

Fig 1 Structure graph of BP neural network

### 1.1 网络学习规则

BP 神经网络的学习过程由信息的正向传递和误差的反向传播两部分组成。在正向传播中, 输入信息从输入层经隐含层传向输出层, 前一层的输出考

\* [收稿日期] 2003-09-12

[基金项目] 国家自然科学基金资助项目(59879022); 世界银行贷款项目(HTJ001)

[作者简介] 付永锋(1975-), 男, 山西孟县人, 在读硕士, 主要从事旱区水文及水资源研究。

虑权重、阈值后成为后一层的输入,如果在输出层得不到期望输出,则转入反向传播,将误差信号沿原来通路返回,通过修改各层神经元的权重和阈值,使得误差信号最小。将输入转换为输出相应的表达式称为处理单元的传递函数,由 2 个过程组成。首先,处理单元计算考虑权重、阈值后从所有上一层单元得到的输入

$$I_j = \sum_{i=1}^n w_{ij}x_i - \theta_j, \quad j = 1, 2, \dots, n \quad (1)$$

式中,  $I_j$  为处理单元  $j$  从前面  $n$  个神经元的输出所得到的权重的输入;  $x_i$  为第  $i$  个评价或模拟变量;  $\theta_j$  为神经元的阈值;  $w_{ij}$  为从第  $i$  个神经元到第  $j$  个神经元的连接权重。

信号传递的第 2 步是由传递函数将输入  $I_j$  转换为单元的输出信号。为了模拟生物神经元的非线性特性,通常选择 S 函数(Sigmoid 函数)作为网络的传递函数,表示为

$$O_j = f(I_j) = 1.0 / (1.0 + e^{-I_j}) \quad (2)$$

式中,  $O_j$  为第  $j$  个网络传递函数。

## 1.2 BP 网络模型存在的问题及其改进

使用 3 层 BP 网络可以任意逼近任何连续函数<sup>[7]</sup>,但其主要有收敛速度慢,存在所谓“局部极小”问题,网络结构为前向结构,没有反馈连接等缺点。针对这些问题,可以进行如下改进:

1) 步长自适应。由于 BP 网络的误差曲面极不规则,如果采用固定的步长  $\lambda$ ,当  $\lambda$  较小时,在误差曲面较平坦的区域,收敛较慢;当步长  $\lambda$  较大时,又会在峡谷区域引起振荡。步长自适应算法正是针对步长的缺陷提出来的,其基本思想为:连续 2 次观测训练的误差值,如果误差下降则增大学习率,误差的反弹在一定范围内则保持步长,误差的反弹超过一定限度则减小学习率。

2) 增加动量项。这种方法是在每一次权值修正时,考虑前一次学习时的校正量而添加一惯性冲量,以滤除学习过程中的高频振荡,这样可以有效防止振荡,还能加速收敛。

## 2 地下水水质评价的 BP 模型

对于地下水水质综合评价的 BP 神经网络来说,其训练样本即为地下水水质分级标准,网络训练完成后,将网络对分级标准学习的知识和有关信息,即各层间的连接权与各个神经元的阈值保存下来,然后输入待评价样本资料,得到有关评价结论的信息,从而根据一定的规则作出有关评价结论的判断。

### 2.1 网络结构

水质评价是一个非线性关系较为复杂的问题,考虑到既满足精度要求,又要尽可能减少学习费用,本研究选用一个隐含层。设有待评价的水质监测样本  $M$  个,每个样本有  $P$  项污染指标监测值,按照国家行业规范规定,相应于  $P$  项污染指标下的水质标准浓度可将水质划分为  $C$  个等级。由于有  $P$  项水质指标参与评价,水质综合评价的结果共有  $C$  个等级,因而将输入层与输出层的神经元数分别定为  $P$  和  $C$ ,输入层的每一个神经元代表一种水质指标,而输出层的每一个神经元代表一个水质等级。

### 2.2 样本的规范化处理

为消除各指标特征之间由于量纲的不同及监测数值大小的差异对计算过程的影响,需对原始数据作规范化处理,通过实践,选用下述方法,效果较好。

$$x_i^k = \frac{x_i^k}{x_{i,\max} + x_{i,\min}}, \quad (i = 1, 2, \dots, P) \quad (3)$$

$$y_t^k = \frac{y_t^k}{y_{t,\max} + y_{t,\min}}, \quad (t = 1, 2, \dots, C) \quad (4)$$

式中,  $x_i^k$  和  $y_t^k$  分别为第  $k$  个样本的第  $i$  个输入值和第  $t$  个输出值;  $x_{i,\max}$  和  $x_{i,\min}$  分别为第  $i$  个水质指标的最大值和最小值;  $y_{t,\max}$  和  $y_{t,\min}$  分别为第  $t$  个期望输出的最大值和最小值。

### 2.3 水质的综合评价

在样本数据规范化处理后,可以将其输入到确定的网络结构中进行网络训练,当满足一定的收敛条件时就认为网络训练成功,可以进行水质评价,即将待评价的各样本资料分别输入到网络中,然后根据网络的输出信息对各样本的水质情况进行评判。

## 3 应用实例

和田地区位于新疆维吾尔自治区的南部,塔克拉玛干沙漠的南缘,区内干旱少雨,主要靠南部昆仑山融雪补给,是典型的内陆干旱气候。利用塔里木盆地灌溉与环保工程二期项目 1999 年地下水水质监测结果,对和田地区 15 个测点的地下水水样进行水质综合评价。

选择 pH 值、矿化度、总硬度、Cl<sup>-</sup> 和 SO<sub>4</sub><sup>2-</sup> 5 项水质指标进行评价,水质监测的原始资料见表 1。依据 GB/T 14848-93 将水质评价标准分为 5 个等级,这样 BP 网络模型的输入层和输出层都取 5 个神经元,试验表明,当隐层单元个数为 9 时,网络模型稳定且可获得较理想的结果,所以水质评价的 BP 网络结构可确定为 5-9-5。

表 1 新疆和田地区地下水水质监测资料

Table 1 The data of groundwater quality

mg/L

测点编号 No.	pH	矿化度 Mineralization degree	Cl <sup>-</sup>	SO <sub>4</sub> <sup>2-</sup>	总硬度 Total hardness
39	7.5	1 010	207	194	461
40	7.7	384	72.7	58.8	184
41	7.5	1 350	318	284	584
42	7.4	2 420	668	359	894
43	7.4	1 220	299	210	502
44	7.6	2 010	509	388	649
45	7.6	3 430	1 040	629	1 037
46	7.8	1 780	536	343	457
47	7.6	4 180	1 400	517	645
48	7.9	1 060	279	210	445
49	7.8	919	257	229	412
50	7.4	1 190	253	253	510
51	7.5	1 290	299	151	494
52	7.6	1 190	276	135	490
53	7.6	1 080	193	214	470

对于地下水水质综合评价的 BP 神经网络来说, 其训练样本即为地下水水质分级标准。依据 GB/T 14848-93, 地下水分类标准的 IV 类与 V 类标准的界值是同一数值, 该标准规定小于等于该值为 IV 类水, 大于该值为 V 类水。而水环境质量标准的划分一般都是指一个浓度区间。为了符合评价的要求, 按照文献 [8] 提出的方法来确定分级代表值: I 类水的标

准界值作为 I 类水的分级代表值, II 类水的分级代表值为 I 类水和 II 类水标准界值的中值, 其余依次类推, 将 V 类水 (IV 类) 的界值作为 V 类水的分级代表值。这样, 由国家地下水质量评价标准及上述分级原则, 可以得到本研究评价因子的类别分级代表值, 也就是 BP 网络的训练样本, 见表 2。

表 2 地下水水质评价标准类别分级代表值和网络期望输出及训练结果

Table 2 Water quality standards and the expected outputs and learned results of ANN

级别 Grade	pH	矿化度/ (mg · L <sup>-1</sup> ) Mineralization degree	Cl <sup>-</sup> / (mg · L <sup>-1</sup> )	SO <sub>4</sub> <sup>2-</sup> / (mg · L <sup>-1</sup> )	总硬度/ (mg · L <sup>-1</sup> ) Total hardness	期望输出值 Expected outputs	网络训练结果 Learned results
I	7.0	300	50	50	150	1 0 0 0 0	0.991 0.012 0.000 0.000 0.000
II	7.8	400	100	100	225	0 1 0 0 0	0.010 0.986 0.012 0.000 0.000
III	8.2	750	200	200	375	0 0 1 0 0	0.000 0.011 0.986 0.011 0.000
IV	8.6	1 500	300	300	500	0 0 0 1 0	0.000 0.000 0.012 0.987 0.011
V	9.0	2 000	350	350	550	0 0 0 0 1	0.000 0.000 0.000 0.007 0.993

注: pH 值以 7.0 为 I 类, 随着增加和减少, 超出区间 [5.5, 9.0] 时为 V 类水, 表中未列出 pH < 7.0 时的分级标准。

Note: When pH is 7, the water quality is grade I, the quality grade would increase or decrease with the change of pH. And when pH goes beyond [5.5, 9.0], the water quality would be grade V. The classification criterions were not included in this table when pH < 7.0

取动量因子  $\eta = 0.3$ , 初始学习率  $\alpha(1) = 0.03$ , 随机赋予网络的初始权值和阈值, 控制误差取为 0.0001。利用 BP 网络模型进行训练, 表 2 给出了对应于每一类水质时网络的期望输出值与网络训练结果。当网络收敛并达到所需精度时, BP 网络模型就建立了水质评价标准与评价级别之间的复杂对应关系, 这时可以利用网络进行水质综合评价。将和田地区地下水水质资料输入已训练好的网络模型中, 按照其输出值与评价级别的贴近度进行判断, 以决定

其归属的水质等级。为验证改进 BP 网络模型水质评价结果的可靠性, 本研究同时运用模糊综合评判法<sup>[9]</sup>对水质进行了评价, 2 种方法的评价结果见表 3。从表 3 可以看出, 水质评价的 BP 网络模型评价结果与模糊综合评判法的评价结果基本一致。可见, BP 神经网络模型用于水质评价是可行的; 相比之下, BP 网络由于其各个神经元输出值差别明显, 可以更明确的判定水质级别, 因而评价更客观。

表 3 BP 神经网络与模糊综合评判法的水质综合评价结果比较

Table 3 Results comparison between BP algorithm and fuzzy evaluation method

编号 No.	BP 神经网络输出结果 Results of BP neural network	水质级别 Grade	模糊综合评判法评价结果 Results of fuzzy evaluation	水质级别 Grade
39	0.000 0.000 0.915 0.405 0.000	III	0.125 0.209 0.348 0.318 0.000	III
40	0.724 0.355 0.000 0.000 0.000	I	0.748 0.252 0.000 0.000 0.000	I
41	0.000 0.000 0.035 0.996 0.004	IV	0.086 0.144 0.068 0.297 0.404	V
42	0.000 0.000 0.000 0.001 0.999	V	0.107 0.107 0.000 0.000 0.786	V
43	0.000 0.000 0.299 0.990 0.000	IV	0.134 0.134 0.279 0.436 0.016	IV
44	0.000 0.000 0.000 0.001 0.999	V	0.067 0.202 0.000 0.000 0.731	V
45	0.000 0.000 0.000 0.001 0.999	V	0.038 0.114 0.000 0.000 0.847	V
46	0.000 0.000 0.000 0.005 0.996	V	0.000 0.196 0.089 0.170 0.546	V
47	0.000 0.000 0.000 0.001 0.999	V	0.029 0.088 0.000 0.000 0.882	V
48	0.000 0.000 0.722 0.907 0.000	IV	0.000 0.263 0.340 0.397 0.000	IV
49	0.000 0.000 0.880 0.661 0.000	III	0.000 0.324 0.383 0.292 0.000	III
50	0.000 0.000 0.307 0.989 0.000	IV	0.156 0.156 0.203 0.388 0.097	IV
51	0.000 0.000 0.442 0.974 0.000	IV	0.124 0.207 0.138 0.531 0.000	IV
52	0.000 0.000 0.764 0.840 0.000	IV	0.087 0.261 0.185 0.466 0.000	IV
53	0.000 0.000 0.852 0.656 0.000	III	0.082 0.246 0.348 0.323 0.000	III

## 4 结 论

BP 神经网络用于地下水水质评价, 只需以各评价参数的分级标准为基础生成训练样本, 用训练后得到的网络模型可以对需评价的地下水水质样本进行评价, 网络模型一旦训练好, 就可以推广应用于该

类环境问题的评价, 因此 BP 网络模型具有很好的通用性。由于人工神经网络方法具有高度非线性函数映射功能, 使地下水水质评价结果的精度大大提高, 与传统水质评价方法相比, 神经网络模型因其大量参数均由学习所得, 因而评价结果更客观、合理。

### [参考文献]

- [1] 沈迅伟, 袁春伟, 王世和. 基于模糊聚类分析的水污染评价[J]. 南京化工大学学报, 1998, 20(增刊): 30- 33
- [2] 黄胜伟, 董曼玲. 自适应变步长 BP 神经网络在水质评价中的应用[J]. 水利学报, 2002, (10): 119- 123
- [3] 万幼川, 谢鸿宇, 吴振斌, 等. GIS 与人工神经网络在水质评价中的应用[J]. 武汉大学学报(工学版), 2003, 36(3): 7- 12
- [4] 杨志英. BP 神经网络在水质评价中的应用[J]. 中国农村水利水电, 2001, (9): 27- 29
- [5] 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
- [6] 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- 10
- [7] 刘曙光, 刘明远, 何 钺, 等. 一种基于同伦 BP 网络记忆模糊规则的参数自调整 ANN-PI 控制及应用[J]. 系统工程与电子技术, 1997, (1): 45- 51
- [8] 夏 军. 区域水环境及生态环境质量评价——多级关联评估理论的应用[M]. 武汉: 武汉水利电力大学出版社, 1999 83- 85
- [9] 付永锋, 沈 冰, 关东海, 等. 模糊综合分析法和和田地下水水质中的应用[J]. 地下水, 2003, (3): 147- 149

## Application of BP network to groundwater quality evaluation ——Case study of Hotan sub-project area Xinjiang

FU Yong-feng<sup>1</sup>, ZHANG Jian<sup>2</sup>, LUO Guang-ming<sup>3</sup>, ZHAO Ji-hua<sup>1</sup>, SHEN Bin<sup>1</sup>

(1 Institute of Water Resources and Hydroelectric Engineering, Xi'an University of Technology, Xi'an, Shaanxi 710048, China;

2 Teaching Center of Computer, Xi'an Jiaotong University, Xi'an, Shaanxi 710032, China;

3 Hotan Survey Bureau of Hydrology and Water Resources, Hotan, Xinjiang 84800, China)

**Abstract:** Based on the explanation of rationale of artificial neural network, the BP Network model for groundwater quality evaluation is established. Then the model and the traditional fuzzy synthetic appraisal method are used to evaluate the groundwater quality in Hotan Prefecture of Xinjiang Autonomous Region. The case study shows that the model possesses the advantage of high accuracy and good practicability and adaptability.

**Key words:** artificial neural network; BP neural network; groundwater; water quality evaluation; Hotan Xinjiang