

文章编号: 1006- 7329(2000)04- 0006- 07

四种水质综合评价方法的比较^{*}

郭劲松, 龙腾锐, 霍国友, 王 红

(重庆建筑大学 城建学院, 重庆 400045)

摘要: 将人工神经网络技术运用于水质评价, 建立了地面水环境质量综合评价的BP 网络模型和Hopfield 网络模型, 并与水质综合评价方法中经典的模糊综合指数法和灰色聚类法进行了比较。结果表明,BP 网络是四种方法中的首选。

关键词: BP 模型; Hopfield 模型; 模糊指数法; 灰色聚类法

中图分类号: X 505

文献标识码: A

水质综合评价是以定量的方式直观地表征水环境质量的总体状况, 是进行水环境容量计算及实施水污染控制的重要基础。国内外的环境工作者研究开发了很多评价方法, 如单指数法、模糊综合评判法、模糊聚类法、灰色聚类法等^[1]。较常用的有模糊综合指数法和灰色聚类法。

人工神经网络是由大量神经元相互连接而成的超大规模非线性系统, 是目前国际上异常活跃的前沿研究领域之一。神经网络力图模拟人脑的一些基本特性, 如自适应性、自组织性和容错能力等来对事物进行分析判断, 已在模式识别、医疗诊断等领域取得了较好的应用效果^[2]。

本文将BP 和Hopfield 人工神经网络模型应用于水环境质量综合评价, 利用南川凤嘴江实测水质资料, 并将其与灰色聚类法和模糊综合指数法进行比较, 以期发展更具客观性的水环境质量综合评价方法。

1 四种评价模型简介

1.1 水质评价BP 网络模型

BP (Back propagation) 网络即反向传播网络, 是目前人工神经网络模式中最具代表性、应用得最广泛的一种模型。它由一个输入层、一个输出层和一个或多个隐含层组成, 各层次的神经元之间单向全互连连接, 是一种由非线性变换单元组成的前馈型网络。最简单的为三层网络。利用最陡坡降法的概念, 将输出出现的误差逐层向输入层逆向传播并“分摊”给各层单元, 以调整各单元间相应的连接权值, 使网络误差达到最小化是BP 网络的关键^[2]。BP 网络通过对有代表意义的范例的学习、训练, 就能掌握事物的本质特征, 因此, 非常适合于解决模式识别和分类问题, 这就为BP 网络应用于水质评价问题提供了理论依据。本文水质评价BP 网络模型采用三层网络, 即由一个输入层、一个输出层和一个隐含层组成。以综合水质类别作为输出层, 该层仅用一个神经元表示。实测的水质指标作为输入层, 则输入层神经元数为水质指标参数的个数。隐含层单元数 = (输入层单元数 × 输出层单元数)^{1/2}。在确定了输入层、输出层和隐含层神经元后, 即建立起了水质评价BP 网络模型。

1.2 水质评价Hopfield 网络

霍普菲尔德网络(Hopfield neural network) 是一种联想式学习网络。联想式学习网络是从问题

* 收稿日期: 2000- 03- 15

基金项目: 国家自然科学基金资助项目(59778021)

作者简介: 郭劲松(1963-), 男, 四川射洪人, 副教授, 主要从事环境科学与技术研究。

领域中取得训练范例(状态变数值),从范例中学习其内在的记忆规则,以应用于新的案例。Hopfield 网络只有一个神经元层次,各处理单元都有一个活跃值,或称为状态,用+1和-1值来表示。整个网络的状态由众多单个神经元的状态所组成。Hopfield 网络中各个神经元之间是全互连的,且连接是双向的,故网络中每个神经元的输出均反馈到同一层次的其他神经元的输入上。

当把要记忆的模式存储进 Hopfield 网络并经训练使网络稳定后,再输入有差异的模式,网络就能输出与其最相近的模式,使能量函数取得最小值。显然, Hopfield 网络适合于解决水质综合评价问题。

以水质类别(m 个)为横轴,水质指标参数(n 个)为纵轴,构成 $m \times n$ 的 Hopfield 网络,则该网络有 $m \times n$ 个神经元。因此,水质评价 Hopfield 网络采用由 $m \times n$ 个神经元组成的网络。把五类水质标准作为要学习的标准模式,让网络“记忆”;再输入测试样本,网络经过迭代运算后就能联想起它应该归属于哪一种标准模式,该标准模式就是这个水域的综合水质类别。

1.3 水质评价模糊综合指数法

模糊综合评价是指对多种模糊因素所影响的事物或现象进行总的评价,又称模糊综合评判。所谓模糊是指边界不清晰,这种边界不清的模糊概念,是事物的一种客观属性,是事物的差异之间存在的中间过渡过程。如Ⅰ级水和Ⅱ级水的边界是无法用一个绝对的判据划分,因为它是一个连续渐变的过程^[3]。同时水环境系统是一个多因素耦合的复杂系统,各因素间关系错综复杂,表现出极大的不确定性和随机性,因此,为得到合理的水环境质量评价结果,引入模糊数学的概念是符合水质评价的客观要求的。代表性的方法有:模糊综合评判法^[4]、模糊概率法^[5]、模糊综合指数法^[6]等。

水质评价模糊综合指数法的评价步骤如下:

- 1) 拟定水质标准
- 2) 计算权重集

对于权重的确定也有很多方法,本文采用在各自超标倍数的基础上加以数学处理的方法,其计算公式为:

$$W_i = \frac{C_i/S_i}{\sum_{i=1}^n C_i/S_i} \quad (1)$$

式中, C_i 为第 i 种污染物的浓度值, S_i 为第 i 种污染物 m 个集合的算术均值。

- 3) 多指标模糊概率综合评价

若已知权重矩阵 A 及模糊关系矩阵 R , 则对于 m 个评价参数的综合评价, 其多指标模糊概率综合评价矩阵可表示为:

$$Y = A \cdot R \quad (2)$$

- 4) 计算模糊综合指数

首先构造水质标准类别向量, 根据地面水环境质量标准《GB 3838-88》, 分为 5 类, 故水质标准类别向量为:

$$S^T = (1, 2, 3, 4, 5) \quad (3)$$

模糊综合指数可表示为:

$$FCI = Y \cdot S \quad (4)$$

1.4 水质评价灰色聚类法

灰色聚类分析是建立在灰数的白化函数生成的基础上。将聚类对象对于不同的聚类指标所具有的白化数据进行归纳, 用以判断该聚类对象属于哪一类。灰数是指只知大概范围而不知其确切值的数, 显然它不是一个数, 而是一个数集, 或是一个数的区间^[3]。

水环境质量系统除了模糊性以外, 还有一种更为广泛, 内容更为深刻的特性, 这就是系统的灰色性——信息的不完全性和非确定性; 系统的模糊性和随机性是系统灰色性的两个不同方面的不

确定性。若记 $i=1,2,\dots,n$ 为聚类样本; $j=1,2,\dots,p$ 为聚类指标; $k=1,2,\dots,m$ 为灰类,即各水质分级标准,则用灰色聚类法评价水质的步骤如下:

- 1) 按聚类指标所属的灰类确定白化函数
- 2) 确定聚类权

$$\eta_{kj}=\frac{r_{kj}}{\sum_{j=1}^p r_{kj}}$$

(5)

式中, r_{kj} 为第 j 个污染指标第 k 个灰类的无量纲数; η_{kj} 为第 j 个污染指标第 k 个灰类的权值。

- 3) 求聚类系数
- 4) 按最大原则确定样本属于何类
- 5) 计算灰色聚类法评价指数 F

与模糊综合指数法相同,仍令水质标准类别向量为:

$$N^T=(1,2,3,4,5)$$

(6)

则灰色聚类法评价指数可表示为:

$$F=Y\cdot N$$

(7)

式中, $Y=\eta\cdot R$ 为灰色聚类法评判结果, R 为归一化后的单因素评判矩阵; N 为水质标准类别向量, F 为灰色聚类法综合评价指数。

2 评价实例

本文采用重庆市南川凤嘴江实测水质资料作为算例,对这四种水质综合评价方法进行比较。

2.1 水质标准和测试样本

水质标准采用《GB 3838- 88》,具体数据见表 1:

表 1 水质标准(GB 3838- 88)

水质指标	水质类别				
	I类	II类	III类	IV类	V类
溶解氧(mg/l)	饱和率 90%	6	5	3	2
BOD ₅ (mg/l)	3 以下	3	4	6	10
COD _{Mn} (mg/l)	2	4	6	8	10
非离子氨(mg/l)	0.02	0.02	0.02	0.2	0.2
大肠菌群(个/l)	100	500	10 000	50 000	100 000
挥发酚(mg/l)	0.002	0.002	0.005	0.01	0.1
氟化物(mg/l)	1.0	1.0	1.0	1.5	1.5
铬(六价)(mg/l)	0.01	0.05	0.05	0.05	0.1

测试样本采用重庆市南川凤嘴江实测水质资料,具体数据见表 2。

2.2 四种评价模型的计算方法

2.2.1 水质评价BP 网络模型算法^[7]

BP 网络计算格式分为训练(学习)和回想(评价)两个过程。训练过程在于自动自适地建立各层神经元之间互连的权值和阈值矩阵。当权值和阈值达到稳定时,即获得了人工神经网络从输入到输出的映射关系。回想过程其实质是对给定测试样本进行水质评价。

网络的训练过程由正向和反向传播组成,其过程如下:

- 1) 对节点间权值 w_{-xh} 及 w_{-hy} 和节点阈值 u_{-h} 及 u_{-y} 赋在 $(-1,1)$ 区间的非 0 均布随机初值。

表 2 测试样本集

样本编号	测 试 指 标							
	溶解氧	BOD ₅	COD _{Mn}	非离子氨	大肠菌群	挥发酚	氟化物	铬(六价)
	8 9	0 6	1 2	0 013 2	2 390	0 001	0 138	0 002
	5 0	2 8	4 6	0 023	12 000	0 001	4 399	0 002
	3 5	6 0	8 1	0 05	5 000	0 005	1 0	0 002
	6 2	4 4	5 4	0 188	3 000	0 003	0 979	0 002
	9 7	1 4	3 1	0 000 4	2 000	0 001	0 313	0 002
	10 4	1 2	1 2	0 000 5	2 380	0 001	0 054	0 002
	7 5	0 7	0 7	0 000 5	2 000	0 001	0 109	0 002
	4 9	6 3	6 5	0 000 5	5 000	0 038	0 130	0 002
	6 9	1 6	3 3	0 061 3	2 380	0 001	0 926	0 002
	10	4 0	3 0	0 02	2 380	0 003	1 0	0 04

- 2) 输入训练范例的输入向量 x , 与目标输出向量 T 。
- 3) 计算输出向量 y 。
- 4) 计算差距量 δ
- 5) 计算权值修正量 Δw 及阈值修正量 Δu 。
- 6) 更新权值 w_{-hy} 及阈值 u_{-y} 。
- 7) 重复步骤 - , 直到收敛(误差不再有明显变化) 或达到一定循环次数。
- 至此, 学习过程结束, 将训练稳定化网络用于回想, 即用于水质评价, 其过程为:
- 1) 调用经训练稳定的权值矩阵 w_{-xh} 及 w_{-hy} 和阈值向量 u_{-h} 及 u_{-y} 。
- 2) 输入测试范例的输入向量 x 。
- 3) 自动计算获得输出向量 y , 即评价结果。

2.2.2 水质评价 Hopfield 网络模型算法^[8]

- 1) 学习过程
- 先设定网络参数, 并按下式计算加权值矩阵。

$$W_{ij} = \sum_p X_i^p X_j^p$$

(8)

$$W_{ii} = 0$$

(9)

- 2) 回想(评价)过程
- (1) 设定网络参数
- (2) 调用加权值矩阵 w
- (3) 从测试范例输入初始状态变数向量 x
- (4) 计算新的状态变数向量 x

$$net[j] = \sum_i w_{ij} \cdot X_i$$

(10)

- (5) 计算李亚普诺夫能量函数

$$E = (-1/2) \sum_i \sum_j x_i w_{ij} x_j + \sum_j \theta x_j$$

(11)

式中: x_i 为第 i 个状态变数

$$X_j = 1 \quad \text{若 } net[j] > 0$$

$$X_j = X_j \quad \text{若 } net[j] = 0$$

$$X_j = -1 \quad \text{若 } net[j] < 0$$

w_{ij} 为连接权值

θ 为连接阈值

重复(4)、(5)两步,直到收敛(能量函数不再有明显变化)或达到一定数目的学习循环。

2.2.3 水质评价模糊综合指数法算法

模糊综合指数法如 1.3 中所述,其中一些具体问题,本文作如下处理:

1) 根据国家新颁布的地面水环境质量标准《GB3838-88》,水质分为五类,即 $m = 5$, 单指标模糊概率评价矩阵 R 中的隶属度(i 指标发生 j 级水质的概率)通过线性函数来确定,按如下的公式计算:

(1) 第 1 级水, 即 $j = 1$, 其隶属函数为:

$$r_{ij} = \begin{cases} 1 & (X_i \leq S_{ij}) \\ A_{ij}(X_i - S_{i(j+1)}) & (S_{ij} < X_i < S_{i(j+1)}); A_{ij} = \frac{1}{S_{ij} - S_{i(j+1)}} \\ 0 & (X_i \geq S_{i(j+1)}) \end{cases} \quad (12)$$

(2) 第 2- ($m - 1$) 级水, 即 $j = 2, \dots, m - 1$ 时, 其隶属函数为:

$$r_{ij} = \begin{cases} 1 & (X_i = S_{ij}) \\ A_{ij}(X_i - S_{i(j-1)}) & (S_{i(j-1)} < X_i < S_{ij}); A_{ij} = \frac{1}{S_{ij} - S_{i(j-1)}} \\ A_{ij}(X_i - S_{i(j+1)}) & (S_{ij} < X_i < S_{i(j+1)}); A_{ij} = \frac{1}{S_{ij} - S_{i(j+1)}} \end{cases} \quad (13)$$

(3) 第 m 级水, 即 $j = m$ 时, 其隶属函数为:

$$r_{ij} = \begin{cases} 1 & (X_i \geq S_{ij}) \\ A_{ij}(X_i - S_{i(j-1)}) & (S_{i(j-1)} < X_i < S_{ij}) \\ 0 & (X_i \leq S_{i(j-1)}) \end{cases} \quad (14)$$

(12)、(13)、(14) 式中, X_i 为第 i 种因子的实测值; $S_{i(j-1)}$ 、 $S_{i(j+1)}$ 为第 i 种因子第 $j - 1$ 级和第 $j + 1$ 级水的标准值;

2) 模糊综合评判的某些算法如 $M(,)$ 复合算法是一种“主因素突出型”的方法。当污染因素多, 各权重较小时, 取大和取小运算将会遗失许多有用信息, 故不宜在水质综合评价中应用。本文将模糊矩阵复合运算改为一般矩阵的乘法运算, 即将算子“ $\wedge$ ”和“ $\vee$ ”改为“ $+$ ”和“ $\cdot$ ”。这样, 可将模糊矩阵 R 和权重系数 A 完全综合地在评判结果 Y 中得到反映。

2.2.4 灰色聚类法水质评价模型算法

设区域环境内有 n 个污染参数, 以 $i = 1^{\#}, 2^{\#}, \dots, n^{\#}$ 表示, 评价等级以 $j = 1, 2, 3, \dots, m$ 表示, 区域以 $k = \text{I}, \text{II}, \dots$ 表示。则根据分级标准中数值的物理意义, 规定白化函数为如下 3 种形式:

$$(1) \quad Y_{ij}(x) = \begin{cases} a_{ij}x & 0 < x < \lambda_{ij}(2) \\ Y_{ij}(s) & \lambda_{ij}(2) < x < \infty \end{cases} \quad (15)$$

$$(2) \quad Y_{ij}(x) = \begin{cases} a_{ij}x & \lambda_{ij}(1) < \lambda < \lambda_{ij}(2) \\ -a_{ij}[x - \lambda_{ik}(2)] + Y_{jk}(s) & \lambda_{ij}(2) < \lambda < \lambda_{ij}(3) \end{cases} \quad (16)$$

$$(3) \quad Y_{ij}(x) = \begin{cases} Y_{ij}(s) & -\infty < \lambda < \lambda_{ij}(2) \\ -a_{ij}[x - \lambda_{ij}(2)] + Y_{ij}(s) & \lambda_{ij}(2) < \lambda < \lambda_{ij}(3) \end{cases} \quad (17)$$

$$i = 1^{\#}, 2^{\#}, 3^{\#}, \dots; j = 1, 2, 3, \dots$$

式中 x 为某一项污染指标值, $Y_{ij}(x)$ 为其对应的白化函数值。 $Y_{ij}(s)$ 表示聚类指标 $\lambda_{ij}(2)$ 所对应的函数值, 并规定等于 1, $\lambda_{ij}(2)$ 表示聚类指标点, a_{ij} 为斜率。

在表 2 中可看到, 大肠菌群这项指标的数值相当大, 所以需进行无量纲化处理,

$$\bar{s} = \frac{s}{10\ 000} \quad (18)$$

在这之后再行聚类权的计算。

3 评价结果及分析

将样本数据按各类评价模型计算方法进行计算, 结果如表 3:

表 3 四种评价方法对同一测试样本集的计算结果

模 型	测 试 样 本 编 号										
		1	2	3	4	5	6	7	8	9	10
BP 模型	输出结果	0.937	2.530	3.585	2.555	1.002	0.918	0.962	3.275	1.314	1.603
	水质级别	I	III	IV	III	I	I	I	III	I	II
Hopfield 模型	输出结果	/	/	/	/	/	/	/	/	/	/
	水质级别	I	III	IV	III	I	I	I	III	I	II
模糊综合指数法	输出结果	1.344	3.567	3.396	3.087	1.454	1.179	1.510	3.660	1.895	2.492
	水质级别	I	IV	III	III	I	I	II	IV	II	II
灰色聚类法	输出结果	1.224	2.353	3.184	2.714	1.447	1.175	1.334	3.256	1.678	1.316
	水质级别	I	II	III	III	I	I	I	III	II	I

注: I ~ V 级水质标准向量为 (1, 2, 3, 4, 5)。

由表 3 可知, 四种评价方法对样本 1 4 5 6 评价结果完全相同, 对样本 7 8 10 的评价结果基本相同, 其它样本结果接近。根据四种评价方法的评价过程分析可得:

人工神经网络模型摒弃了任何人为因素, 只根据水质标准本身的特点进行学习, 从而产生连接权值矩阵和阈值矩阵, 这是与传统水质评价方法的根本区别。当将测试样本输入模型后, 网络将自动自地调用权值与阈值, 进行回想或联想, 从而得到测试样本的输出结果(评价结果)。就此而言, 其输出结果与其他需人为确定权值矩阵的评价方法的输出结果相比具有更真实的客观性与唯一性。

BP 网络水质评价模型可选取任意多的水质参数进行学习, 建立不同的评价模型, 学习稳定后的网络具有相当高的精度与准确率。对于任意实测样本, 为了获得评价结果, 只需将数据作为输入向量赋给网络输入层, 则网络会自组织地给出评价结论, 因此, 学习成功后的网络具有广泛的适应性。

Hopfield 网络采用模式联想的计算格式, 用于水质综合评价形象、直观、方便。网络回想时间很短, 一般只需一到两次迭代即可完成。根据网络的特点, 它既适应于定量指标的水质参数, 也适用于定性指标的水质参数, 而且水质参数越多, 评价结果越可靠, 因此, 有其独特的个性。然而, 由于 Hopfield 网络采用模式联想, 网络未考虑水质是连续变化的这一事实, 因而不能给出指标污染物的权重, 使评价结果从理论上讲只能是接近某一类标准水质, 贴近度多少不得而知, 这是该方法的缺陷。但在对水体进行概略地分类时, 该方法的优点是十分明显的。

灰色聚类法中白化函数的建立与模糊数学中隶属函数的建立颇为相似, 都注意到水质等级界限的模糊性, 结果表示也相同。两种方法最显著的不同点在于权重的计算方法上。在模糊综合指数法中计算权重的含义是, 根据各污染物的超标情况进行加权, 超标越多, 加权越大。在不同的水质级别之间, 各污染物的权重应该是不同的, 但根据这一原则计算权重, 各污染物的权重都是相同的。灰色聚类法认为各污染物的危害性反映在变化幅度不同的分级标准中, 根据不同的水质类别确定相应级别各污染物的权重, 这一点避免了模糊数学方法只用一个权重划分水质级别的不合理之处, 但这种权重计算方法显然完全没有考虑超标污染物对水质的特别损害, 更有着其明显的缺陷。由此可以看出: 用模糊综合评判法易使评价结果偏重, 而用灰色聚类法评价水质又容易偏轻, 从表 3 中样本 1、2、3、4、5、6、7、8、9、10 的实际评价结果来看也是如此。模糊综合评判法和灰色聚类法计算过程都相当复杂, 编写程序可使运算速度加快, 但由于隶属度值和白化函数的计算均要求 s_{ij} 是单调递增的, 即

$S_{i1} < S_{i2} < S_{i3} < S_{i4} < S_{i5}$ 。若每次水质评价时选取的指标不完全一样,则编写的程序需作相应调整后方可使用,增加了评价工作的难度。另一方面,若要提高程序的通用性,则必须增加数据、指标的前处理工作。因此该类方法的运算程序很难具有通用性。

4 结 论

综合上述分析可得以下主要结论:

1) BP 网络、Hopfield 网络、模糊综合指数法、灰色聚类法这四种水质评价方法都属于水质综合评价模型,对水环境质量评价的结论也是基本一致的。它们推导严谨,有较充分的理论依据,应用于水质评价均有其合理性,适用于作深入的水环境质量分析。

2) 人工神经网络模型模拟人脑思维,具有很强的自组织、自学习、自适应能力,用于水质评价有以下更突出的优点:权值的获得是网络对水质标准学习的结果,摒弃了人为确定权值的主观影响,使评价结果更具客观性;一个学习成功的网络应用于任意一个实例样本,只需将样本的参数作为输入向量赋予网络,即可得到评价结果输出。计算简便,运算程序的实用性和通用性更强。这一点是模糊综合指数法和灰色聚类法做不到的,它们必须根据不同的实例选取参数和调整权值。

3) 综合比较四种方法,BP 网络水质综合评价模型是更佳的方法,应是这四种方法的首选。

参考文献:

- [1] 郭劲松,等. 水资源水质评价方法分析与进展 [J] 重庆环境科学, 1999, 15(6): 1~ 5
- [2] 胡守仁,余少波,等. 神经网络导论 [M] 北京: 国防科技大学出版社, 1993
- [3] 侯克夏. 环境系统工程 [M] 北京: 北京理工大学出版社, 1992
- [4] 陈飞星. 一种新的水质评价模式 [J] 中国环境科学, 1986, 6(6): 54~ 58
- [5] 马建华,等. 水质评价的模糊概率综合评价法 [J] 水文, 1994, (3): 21~ 24
- [6] 邓峰. 水质评价的一种新方法——模糊综合指数法 [J] 中国环境科学, 1991, 11(1): 63~ 66
- [7] 郭劲松,等. 水环境质量的隶属度 BP 网络决策模型 [J] 中国给水排水, 2000, 16(3): 11~ 15
- [8] 沈清,胡德文,等. 神经网络应用技术 [M] 北京: 国防科技大学出版社, 1993

A Comparison of Four Methods of Water Quality Assessment

GUO Jing-song, LONG Teng-rui, HUO Guo-you, Wang Hong

(Faculty of Urban Construction, Chongqing Jianzhu University, Chongqing 400045, China)

Abstract In this paper, a back-propagation network, a Hopfield network, Fuzzy comprehensive index method and classification of gray are compared in application for water quality assessment. The comparison result shows that the BP network is the best one of these four methods.

Keywords BP-network; Hopfield-network; Fuzzy comprehensive index method; classification of gray