

29-32

南湖富营养化的人工神经网络评价* Assesment on the Eutrophication of Water Quality of Lake Nanhu with the Artificial Neutral Network

谢宏斌

Xie Hongbin

(南宁市环保监测站 南宁 530012)

(Environmental Monitoring Station of Nanning, Nanning, 530012)

摘要 在采用权重贡献率分析法优选因子和确定因子权值的基础上,应用B-P人工神经网络对南湖水质富营养化进行综合评价,并将B-P法评价结果与用分级评分法的评价结果进行了比较,其结果与目前通用的评分法的评价结果一致。

关键词 人工神经网络 B-P算法 富营养化

中图法分类号 X 824

Abstract Based on the optimization and the determination of weight values for the factors by using the weight analysis of contribution rates, use is being made of the comprehensive assessment for the water quality eutrophication of Lake Nanhu using B-P artificial neutral network. The comparison was made for the assessment results between the methods of the B-P artificial neutral network and the classified mark. The assessment results coincide with the cases of water quality eutrophication of Lake Nanhu.

Key words artificial neutral network, B-P algorithm, eutrophication

人工神经网络是在现代神经科学研究成果基础上提出的一种数学模型,目前已在模式识别和环境科学领域中得到应用,并取得十分出色的成果^[1~4]。尽管湖泊富营养化评价的方法很多,但归根到底,评价是一个模式识别问题^[5]。本文采用人工神经网络(Artificial neural network)对南湖富营养化程度进行了综合评价,并制定了相应的南湖富营养化评价标准。

1 人工神经网络评价法

1.1 B-P网络原理简述^[2,6]

B-P网络是人工神经网络中最具代表意义和广泛应用的模型。它由输入层、隐节点层和输

1998-11-24收稿。

* 广西重点课题《南湖富营养化及防治途径的研究》的子课题。

X 824
X 524

BP算法
南湖 评价 水质评价

出层组成,相邻两层节点之间单向互联,方向由输入层到输出层。各神经元节点的输出为:

$$y_i = f\left(\sum_j w_{i,j} x_j + \theta_i\right) \quad (1)$$

式中: y_i 为第*i*个神经节点的输出; $w_{i,j}$ 为上一层*j*节点对下一层*i*节点的权值; x_j 为输入信号的第*j*个分量; θ_i 为阈值。函数一般取

$$f(x) = 1/(1 + e^{-x}) \quad (2)$$

B-P网络包括正向和反向学习过程^[6,8]。学习结束即可确定网络结构,当进行评价时,只需把数值输入网络进行正向学习,则输出结果中最大的节点就代表了模式的类别。

1.2 评价模式的建立

1.2.1 综合评价标准

影响湖泊富营养化程度的因子很多,叶绿素a(chla)、总磷(TP)、总氮(TN)、高锰酸钾指数(COD_{Mn})、生化需氧量(BOD₅)、透明度(SD)、氨氮(NH₃-N)等均是影响湖泊富营养化的重要因子。

根据有关文献^[5,9],结合南湖目前富营养化程度较高的特点,本文提出较适合南湖富营养化评价的标准,见表1。

表1 南湖富营养化的评价标准

评价等级	chla (mg/m ³)	TP (mg/L)	TN (mg/L)	COD _{Mn} (mg/L)	BOD ₅ (mg/L)	SD (cm)	NH ₃ -N (mg/L)
贫营养化	0.64	0.010	0.13	2.04	0.96	493	0.030
中营养化	3.98	0.035	0.42	4.32	2.08	176	0.102
富营养化	25.12	0.120	1.38	9.15	4.51	63	0.342
重富营养化	158.50	0.413	4.50	19.40	9.78	22	1.15
极富营养化	1000	1.415	14.64	41.14	21.25	8	3.86

1.2.2 评价因子的确定

本文选择一些起主要作用的因子来进行评价,对富营养化影响作用相对较小的因子忽略不计。首先把chla、TP、TN看作基本因子,然后根据神经网络的权重贡献率从COD_{Mn}、BOD₅、SD、NH₃-N中选出重要因子^[7]。

故神经网络的结构首先设定为:输入节点4个(代表4个因子:COD_{Mn}、BOD₅、SD、NH₃-N),输出节点4个(由于南湖各因子数值较高,故选择中、富、重富、极富营养化4个评分等级)、中间层隐节点可定为6个、学习样本的选取为: $N = 2 \times 3^4 - 2^4 = 146$ 组^[8],网络输入值由于一般限定在(0,1)之间,因此设 c_i 为第*i*个因子的输入值, b_i 为第*i*个因子的评价标准, b_{min} 为第*i*个因子的各级评价标准中的最小值,则

$$c_i = 1 - \frac{b_{min}}{b_i}, i = 1, 2, 3, 4 \quad (3)$$

假设评价结果为*j*级,则教师信号(标准输出)为

$$T_k = \begin{cases} 0.9, & k = j \\ 0.1, & k \neq j \end{cases} \quad k = 1, 2, 3, 4 \quad (4)$$

此处取0.9和0.1代替通常的1和0,是为了加快收敛速度。网络的精度用下式计算。

$$E = \frac{1}{N} \sum_{i=1}^n d_{k,i}^2 \quad (5)$$

其中 $d_{k,i} = O_{k,i} - T_{k,i}$, $k = 1, 2, 3, 4$, $O_{k,i}$ 为网络输出值。在经过6012遍的训练达到了精度要求 $E \leq 0.02$ 后停止运算,输出权值和阈值。

2
反

利用网络的权值 $w_{i,j}, v_{j,k}$ (该权值可自己设定, 然后再通过网络的反向学习逐步修正确定^[2]) 的组合可以计算每个因子对输出结果的贡献, 计算公式如下:

$$\bar{w}_i = \sum_{j=1}^n w_{i,j} \left(\sum_{k=1}^K v_{j,k} \right) \quad (6)$$

其中: \bar{w}_i 为第 i 个因子对网络输出的贡献; $w_{i,j}$ 为输入层到中间层之间的权值; $v_{j,k}$ 为中间层到输出层之间的权值。然后按公式(7) 将其归一化, 计算每个因子的权重贡献率。

$$\bar{w}'_i = \bar{w}_i / \sum_{i=1}^n \bar{w}_i \quad (7)$$

\bar{w}'_i 的最小值所对应的因子即为可剔除掉的因子, 因子剔除的限值为 $\bar{w}'_{i,\min} \leq 5\%$ 。首先用这种方法对 COD_{Mn} 、 BOD_5 、 SD 、 $\text{NH}_3\text{-N}$ 4 个因子进行筛选, 剔除 $\text{NH}_3\text{-N}$ 这个因子, 再用此法对 TP 、 TN 、 COD_{Mn} 、 BOD_5 、 SD 5 个因子进行筛选, 又剔除掉 BOD_5 。

1.3 建立评价模式

精选评价因子后, 我们把 chl_a 、 TP 、 TN 、 COD_{Mn} 、 SD 5 个因子的评分标准按 4 个评价等级, 选取 6 个隐节点建立 B-P 神经网络进行训练, 精度设置为 ≤ 0.04 即可。这样经过 1020 遍学习后, 网络学习完毕, 输出各层间的权值和阈值。

2 评价结果

2.1 现状评价结果

把南湖 1996 年~1997 年水质的监测数据平均值输入已学习好的 B-P 网络进行评价, 结果得知, 南湖富营养化状况已非常严重, 处于极富营养化状态。

2.2 近年来南湖富营养化程度变化趋势分析

把南湖 1991 年~1996 年 8 个测点的监测数据平均值输入人工神经网络的评价结果见表 2。

评价结果是利用网络输出值 O_i 与标准输出值相比而得出的。如 1991 年的输出值较接近标准输出值

表 2 南湖富营养化的评价结果 (1991 年~1996 年)

时间	输出结果				精度 E	评价结果
	O_1	O_2	O_3	O_4		
1991 年	0.200 0	0.181 5	0.703 4	0.317 6	0.023 50	重富营养化
1992 年	0.191 6	0.183 4	0.699 3	0.141 3	0.013 92	重富营养化
1993 年	0.199 4	0.185 8	0.749 8	0.266 8	0.007 19	重富营养化
1994 年	0.200 8	0.191 3	0.191 1	0.815 0	0.000 38	极富营养化
1995 年	0.200 5	0.189 4	0.300 4	0.710 3	0.018 25	极富营养化
1996 年	0.200 6	0.191 3	0.235 8	0.771 8	0.002 15	极富营养化

(0.1, 0.1, 0.9, 0.1), 最大节点为 O_3 , 所以评价结果为第 3 级, 即重富营养化^[2]。从评价结果看, 1994 年、1995 年、1996 年明显比 1991 年、1992 年、1993 年的富营养化程度高。从神经网络原理可知, 输出值与标准值的差距 (即精度) 越小则在该等级内的营养化程度就越高。可见 1991 年的富营养化程度比 1993 年要小, 在重富营养化阶段呈递增趋势。而进入极富营养化阶段后, 南湖富营养化程度略有波动, 其变化表现为 1994 年较大, 1995 年小, 1996 年又变大。

从单项评价因子的数值来分析, 1991 年~1993 年各项因子的数值均为逐步增长, 全湖的富营养化程度也在增强。但 1994 年~1996 年, 由于 TP 和 TN 的数值异常变大, 其他各项因子的数值反而在一定程度上受到了抑制, 故 1994 年~1996 年的富营养化程度出现了波动, 并不是一直增强。1994 年以前, 南湖属于营养化的响应阶段, 1994 年以后则处于非响应阶段的异常营养状态。进入该阶段后, 考虑到 TP 和 TN 已变得异常, 要评价南湖的营养化状态, 去掉这

两项因子再评价更适合些。故采用 chl_a 、 COD_{Mn} 、 BOD_5 、 SD 这4项因子进行评价, 结果如表3所示。

由表3可知, 除去 TP 和 TN 这两项因子后, 南湖仍处于极富营养化阶段, 但由精度可看出富营养化的程度则相对于保留 TP 和 TN 时要小, 且富营养化的程度

变化表现为1994年较大, 1995年略小, 1996年最大, 呈逐年振荡上升态势。

利用分级评分法对南湖1996年~1997年全年水质进行评价, 结果也为极富营养化, 与神经网络的评价结果一致。

3 结语

应用 B-P 人工神经网络进行水质富营养化评价, 能客观地反映水质富营养化状况; B-P 人工神经网络还可通过对权重贡献率的分析, 提取关键性神经元节点, 从而实现对因子的优选; 应用 B-P 人工神经网络对富营养化标准样本一旦训练完毕, 只需通过简单的加法和乘法运算, 就可以对任何湖泊水质富营养化进行评价, 因此该方法适用性强, 通用性好。

参考文献

- 1 李祚泳. 基于 B-P 网络的水质营养状态评价模型及其效果检验. 环境科学学报, 1995, 15 (2): 186~191.
- 2 李祚泳. B-P 网络用于环境质量分类研究. 环境科学, 1994, 15 (5): 75~77.
- 3 李祚泳, 盛祖康. 人工神经网络用于大气环境质量评价与排序. 四川环境, 1994, 13 (3): 45~47.
- 4 李祚泳, 邓新民. 综合环境质量的 B-P 网络二级评价. 环境科学研究, 1995, 8 (3): 32~35.
- 5 李祚泳. 基于指数标度法的大气污染物标准浓度限值. 中国环境科学, 1994, 14 (3): 196~199.
- 6 李祚泳. 环境监测优化布点的人工神经网络模型. 中国环境科学, 1997, 17 (1): 27~29.
- 7 Li Zuoyong, Zhang Huijun, Deng Xinmin. A new assessment method of eutrophication of lake based on AHP-PCA. China Environmental Science, 1994, 5 (1): 37~42.
- 8 刘首文等. 人工神经网络在湖泊富营养化评价中的应用研究. 上海环境科学, 1996, 15 (1): 11~14.
- 9 金相灿等. 中国湖泊富营养化. 北京: 中国环境科学出版社, 1990.

(责任编辑: 黎贞崇 邓大玉)