

人工神经网络在深圳市水库富营养化评价中的应用

林高松, 黄晓英, 李娟

(深圳市环境科学研究院, 广东 深圳 518001)

摘要:对富营养化评价标准进行插值获取大量的样本,建立了基于 BP 人工神经网络的富营养化评价模型。将模型应用于评价深圳市 13 座主要水库的富营养化状况,对其成因进行分析,并提出了对策与建议。研究表明,石岩水库与深圳水库为轻度富营养化,占评价水库总数的 15.4%;西丽水库等 11 座水库为中营养,占评价水库总数的 84.6%。人工神经网络用于建立湖库富营养评价模型是适合的。

关键词:人工神经网络;富营养化评价;水库;深圳

中图分类号: X82 **文献标识码:** B **文章编号:** 1006 - 2009 (2010) 01 - 0059 - 05

Application of Artificial Neural Network Method on Eutrophication Assessment for Shenzhen Reservoirs

L N Gao-song, HUANG Xiao-ying, LI Juan

(Shenzhen Academy of Environmental Science, Shenzhen, Guangdong 518001, China)

Abstract: Abundant of training samples were gotten via interpolation of eutrophication assessment standard, and a eutrophication assessment model was established based on back propagation artificial neural networks. The model was used to assess nutritional situation of thirteen main reservoirs of Shenzhen, and its cause was analyzed, and then strategies and advices were brought forward. Research result showed that 15.4% reservoirs were slightly eutrophic including Shiyuan Reservoir and Shenzhen Reservoir, and 84.6% reservoirs were middle nutrition. Artificial neural networks without factitious weight were suitable to bright up eutrophication assessment model whose result was objective.

Key words: Artificial neural networks; Eutrophication evaluation; Reservoir; Shenzhen

富营养化是指水体接纳过量的氮、磷等营养物质,使藻类及其他水生生物异常繁殖,破坏水体的生态系统,造成水质恶化、影响水资源利用的一种现象^[1]。近年来,我国湖泊水库富营养化发展十分迅速。根据 2007 年全国 26 个重点湖库的评价结果^[2], 53.8% 的湖库处于不同程度的富营养化状态,水体富营养化已经严重威胁我国湖库水质的安全。

深圳是我国经济最发达、人口最集中的城市之一,由于无过境大江大河,几乎完全依赖于水库调蓄供水,水库水质安全直接影响到深圳的生存和发展。深圳位于亚热带地区,具有适合藻类生长的气象条件,水体富营养化威胁较大。科学评价和分析深圳市水库的富营养化状况,提出防治措施,对保

障深圳市饮用水安全和社会经济持续发展具有十分重要的意义。

富营养化评价方法主要有营养状态指数法、灰色聚类法、模糊数学法、物元分析法、随机评价法等^[3-8]。由于富营养化评价受多个因子的影响,各因子的贡献存在较大差异,且关系复杂,上述评价方法各有其适用条件和局限性,并且受到权重取值、隶属函数构造等主观因素影响。因此,迄今还没有一种公认普遍适用的评价方法^[9-10]。人工神经网络具有自学习、自组织、自适应与容错能力,可用于湖库水质富营养评价,国内众多学者也对此展

收稿日期: 2009 - 09 - 07; 修订日期: 2009 - 12 - 07

作者简介: 林高松 (1979—), 男, 广东高州人, 高级工程师, 博士, 从事水环境管理、水环境数值模拟的研究。

开了研究^[9-12]。利用神经网络建模的关键在于充足的训练样本和合理的拓扑结构,但大多数研究往往直接采用评价标准作为训练样本,过少的样本直接影响了训练效果;同时,拓扑结构太复杂,例如输出层或隐层节点数过多,导致泛化能力较差^[11]。而且,现有神经网络模型采用的分级标准或评价指标不统一,评价结果缺乏可比性和通用性^[12]。

现采用中国水利部推荐的营养状态分级标准,通过插值获得充足的训练样本,利用 BP 人工神经网络建立湖库富营养化评价模型,并对深圳市主要水库的富营养化状况进行评价,为深圳市水环境管理提供科学依据。

1 基于人工神经网络的富营养化评价模型

中国水利部在《中国水资源公报》中采用线性插值评分法 (linear interpolation scoring method, 简称 SCO 法) 评价全国湖库的营养状况,即根据表 1 的营养状态分级标准得到单个指标的营养状态分值,然后等权平均得到综合营养状态分值,以此确定营养状态级别^[13]。但是,SCO 法默认了所有指标的权重相等,实际上没有考虑各评价指标对营养状态的不同贡献,这种人为权重赋值方式会导致评价结果存在主观性,而人工神经网络可避免该缺陷。

表 1 地表水富营养化评价标准

mg/m³

Table 1 Eutrophication assessment standard of surface water

mg/m³

营养状态级别	营养状态分值	叶绿素 a	TP	TN	I _{Mn}	透明度
贫营养	0	0	0	0	0	37
	10	0.5	1	20	0.15	10
	20	1	4	50	0.4	5
中营养	30	2	10	100	1	3
	40	4	25	300	2	1.5
	50	10	50	500	4	1
富营养	60	26	100	1 000	8	0.5
	70	64	200	2 000	10	0.4
	80	160	600	6 000	25	0.3
	90	400	900	9 000	40	0.2
	100	1 000	1 300	16 000	60	0.12

单位为 mg/L; 单位为 m。

人工神经网络是基于生物学的神经网络原理而建立的一种信息处理系统,具有极强的非线性映射能力。现采用最为常用的前馈型人工神经网络,训练方法为误差反向传播 (Back Propagation, BP) 算法,简称 BP 神经网络^[14]。BP 神经网络具有 3 层或 3 层以上的神经元,通过邻层神经元之间的连接实现信息的传递。如果输入节点数为 n , 输出节点数为 m , 则神经网络是从 n 维欧氏空间到 m 维欧氏空间的映射, 3 层结构的 BP 神经网络, 见图 1。

BP 神经网络在学习过程中由正向传播和反向传播两部分组成,正向传播是数据由输入层经隐层处理传向输出层;反向传播是误差信号从输出层向输入层传播并沿途调整各层连接权重和各层神经元的阈值,以使误差信号不断减小,通常采用 Sigmoid 函数作为神经元的激发函数。Sigmoid 函

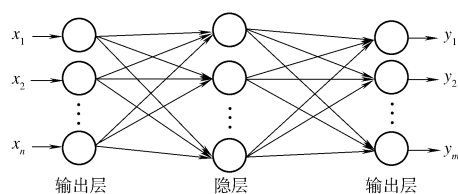


图 1 BP神经网络

Fig 1 Sketch map of Back Propagation artificial neural network

数为:

$$f(x) = \frac{1}{1 + e^{-x}}$$

误差函数定义为:

$$Err(w) = \frac{1}{N} \sum_{i=1}^N F(x_i, w) - y_i^2$$

式中: x_i 为输入数据; w 为神经网络的权重; y_i

为期望输出; $F(x_i, w)$ 为实际输出; N 为训练数据个数。

神经网络的设计主要包括 3 个方面的内容: 网络拓扑结构的设计, 包括隐层数、隐层神经元节点数; 神经网络主要参数的确定, 包括动量项系数、学习率、学习次数的确定; 训练样本的组织, 即训练样本的选择和处理^[15-16]。现就训练样本、拓扑结构和 BP 神经网络的主要参数对建立的模型进行介绍。

1.1 训练样本

生成足够多符合湖库富营养化评价标准的训练样本、检验样本和测试样本是应用 BP 神经网络建立鲁棒性和泛化能力好的模型的关键^[11]。其中, 训练样本用于训练神经网络的权重; 检验样本用于实时动态监控训练过程, 判断训练过程是否“训练不足”或“过训练”; 测试样本不参与训练过程, 主要用于判断已建立的神经网络的泛化能力。现以表 1 的富营养化评价标准为指标, 各指标的分级区间内均匀插值生成 100 个样本, 共获得 1 000 个样本。按区间均匀分布原则抽取 800 个作为训练样本 (即每个区间中随机抽取 80 个), 检验样本和测试样本各抽取 100 个。对训练样本的输出数据添加服从均值为 0、标准差为 0.1 的正态随机噪声, 以防止训练样本过于平滑, 同时也是提高泛化能力的一种有效方法^[12]。

1.2 网络拓扑结构

神经网络的隐层及其节点越多, 训练能力越强, 但其泛化能力越差。为了确保训练后的神经网络

模型的泛化能力和预测能力, 应在满足精度要求的前提下尽可能采用紧凑的神经网络结构。因此, 现采用 3 层结构的 BP 神经网络, 即一个输入层、一个隐层和一个输出层。由表 1 可知, 富营养化评价的指标有 5 个, 评价结果为 0~100 之间的数值, 因此输入层有 5 个节点, 输出层有 1 个节点。隐层的节点个数仍然没有较好的方法事先确定, 通常是在训练过程中逐渐增加其数目, 反复训练, 直到精度达到要求为止^[15]。

1.3 网络的主要参数与模型的验证

根据上述方法获得的 800 组训练样本和 100 组检验样本对 BP 神经网络进行训练, 剩余 100 组测试样本用于验证神经网络的拟合效果。隐层节点数和训练次数在训练过程中通过实际检验后确定, 一般而言, 如果网络训练误差下降极其缓慢, 甚至不再下降, 则可停止训练。在主要参数中, 学习率取 0.04, 动量项系数取 0.9。经检验, 隐层节点数取 4 是最紧凑的, 并且误差达到较低水平, 而训练次数达到 20 000 后误差下降不明显, 因此取 20 000 即可, 由此得到训练好的人工神经网络。

由于剩余的 100 组测试样本未参与训练过程, 因此能够真实地反映神经网络的泛化能力和预测能力。采用测试样本对训练后的神经网络的拟合效果进行验证, 结果表明期望输出与实际输出的误差较小, 平均误差仅 0.6%。因此, 神经网络训练效果良好, 能够用于富营养化评价。部分测试样本的期望值与神经网络预测值的对比结果见图 2。

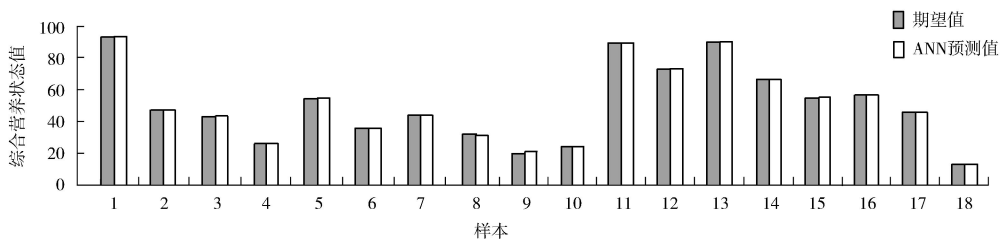


图 2 BP 神经网络预测值与期望值对比

Fig 2 Comparison of anticipation value and prediction value of BP artificial neural networks

2 深圳市水库富营养化评价

2.1 水库富营养化评价结果

评价对象包括了深圳市 13 座位于饮用水水源保护区内的水库: 深圳水库、西丽水库、梅林水库、铁岗水库、石岩水库、罗田水库、清林径水库、赤坳

水库、松子坑水库、径心水库、铜锣径水库、枫木浪水库和三洲田水库, 其供水量超过深圳市供水总量的 90%, 基本代表了深圳市饮用水源的整体环境质量。分析和掌握上述水库的富营养化状况, 对于保护水源、保障人民身体健康和加强饮用水源地管

理具有重要的现实意义。

以 2007 年深圳市水库水质的年平均值为例，

采用已建立的人工神经网络模型对水库的富营养化状况进行评价，结果见表 2。

表 2 深圳市主要水库富营养化评价结果

Table 2 Eutrophication assessment results of main reservoirs of Shenzhen

水库名称	叶绿素 a	TP	TN	I_{Mn}	透明度	营养状态值	评价结果
深圳水库	8.9	57.8	1494.5	1.3	0.5	50.3	轻度富营养
西丽水库	11.3	45.7	1098.7	1.3	0.7	48.3	中营养
梅林水库	3.8	23.9	513.3	1.4	0.9	43.0	中营养
铁岗水库	23.9	38.9	945.0	2.7	0.6	49.8	中营养
石岩水库	40.4	43.2	1428.7	2.9	0.6	51.5	轻度富营养
罗田水库	13.6	39.6	1125.0	3.4	0.6	49.2	中营养
清林径水库	2.6	14.1	746.6	2.2	1.1	40.5	中营养
赤坳水库	2.9	16.3	760.2	2.2	1.1	41.3	中营养
松子坑水库	3.3	18.2	812.5	1.7	1.0	41.3	中营养
径心水库	1.8	16.0	731.7	2.0	1.3	39.9	中营养
铜锣径水库	2.2	19.2	619.5	1.7	1.2	40.5	中营养
枫木浪水库	2.1	17.4	641.8	1.8	1.2	40.2	中营养
三洲田水库	2.2	16.9	675.2	1.8	1.3	39.7	中营养

单位为 mg/L; 单位为 m。

根据人工神经网络模型的评价结果，在 13 座评价水库中，营养状态分值由高至低依次为：石岩水库 > 深圳水库 > 铁岗水库 > 罗田水库 > 西丽水库 > 梅林水库 > 赤坳水库 > 松子坑水库 > 清林径水库 > 铜锣径水库 > 枫木浪水库 > 径心水库 > 三洲田水库。其中，石岩水库与深圳水库为轻度富营养化，占评价水库总数的 15.4%；西丽水库等 11 个水库为中营养，占评价水库总数的 84.6%。铁岗水库、罗田水库和西丽水库虽然为中营养，但营养状态值均超过 48，已接近轻度富营养化水平。

2.2 原因分析

尽管深圳市的水库以中营养为主，但部分水库已达到或接近轻度富营养化水平。其中，石岩水库、铁岗水库、罗田水库和西丽水库的叶绿素 a 均超过 10 mg/m^3 ，深圳水库也达到 8.9 mg/m^3 ，暴发水华的风险相对较大。

(1) 深圳市水库属于调蓄型水库，约 70% 水源来自东江。东江水源的氮、磷质量浓度偏高，而且随着东江上游地区经济发展，水质存在下降迹象。东江水源对深圳水库水质影响极大，来水中的氮、磷是水库营养盐的重要来源。

(2) 由于历史原因，深圳市水库流域开发强度普遍较高，建成区面积比重大，不可避免的对水库水质造成影响。近年来，虽然深圳水库、石岩水库

等相继开展了污水截排工程，但截污系统不完善，仍有部分生活污水、工业废水及面源污染负荷进入水库，对水库水质造成了较大影响。

(3) 深圳地处亚热带，阳光充沛，温度偏高，一般 3 月—11 月水温均超过 20°C ，具有适合藻类生长的气象条件。

受东江来水及水库周边污染影响，深圳市水库营养盐较为充足，并且具有适合藻类繁殖的气象条件，富营养化威胁较大。另一方面，由于深圳市水库库容较小，而用水量较大，水体交换速度相对较快，一定程度上起了抑制藻类繁殖、防止水华暴发的作用。

2.3 对策与建议

(1) 加强东江流域的区域合作，建立适合的行政协调机制，并通过法律、经济补偿等方式，引导东江流域的合理开发，共同保护和改善东江水源水质，减少来水带来的营养物质。

(2) 加快完善各水库的污水截排工程，加强环境管理，依法清除饮用水源保护区内的排污口，减少入库污染负荷。

(3) 严格控制饮用水源保护区内的土地利用开发强度，加强水库流域生态补偿机制的研究，以适当方式迁移或安置水源保护区内的居民和企业，减轻乃至消除人类活动对水库水质的影响。

(4)制定水华暴发应急预案,一旦发生事故,各部门能够及时响应,将危害降至最低。

3 结论

(1)神经网络具有自学习、自组织、自适应与容错能力,只要有充足的训练样本,并且采用紧凑的网络结构,完全适用于建立湖库富营养评价模型,并且避免了人为权重赋值的缺陷,具有较强的客观性。

(2)在深圳市 13 座水库中,石岩水库与深圳水库 2 座水库为轻度富营养化,占评价水库总数的 15.4%;其余 11 座水库为中营养,占评价水库总数的 84.6%。铁岗水库、罗田水库和西丽水库虽然为中营养,但营养状态值均超过 48,已接近轻度富营养化水平。

(3)深圳市水库富营养化的威胁较大,主要原因是东江来水氨磷含量偏高,同时水库流域开发强度大、截污系统不完善,大量污染物进入水库,导致水库营养盐充足,适当的气象条件也为藻类生长提供了基础。因此,建议加强东江流域的区域合作以改善来水水质,并且完善污水截排工程、控制水库流域的土地开发利用强度,从而减少入库营养物质,防止水库富营养化和水华暴发。

[参考文献]

- [1] 金湘灿. 中国湖泊水库环境调查研究 (1980—1985) [M]. 北京: 中国环境科学出版社, 1990.
- [2] 国家环境保护总局. 2007 年中国环境状况公报 [OL]. [2009-09-23]. <http://www.mep.gov.cn/plan/zkgh>
- [3] 李凡修, 辛焰, 陈武. 集对分析用于湖泊富营养化评价研究 [J]. 重庆环境科学, 2000, 22(6): 10-11.
- [4] 林衍, 顾恒岳, 周富春. 湖泊水质富营养化评价的物元分析法 [J]. 环境工程, 1996, 14(5): 44-49.
- [5] 王明翠, 刘雪芹, 张建辉. 湖泊富营养化评价方法及分级标准 [J]. 中国环境监测, 2002, 18(5): 47-49.
- [6] 谢平, 李德, 陈广才, 等. 基于贝叶斯公式的湖泊富营养化随机评价方法及其验证 [J]. 长江流域资源与环境, 2005, 14(2): 224-228.
- [7] 朱庆峰, 廖秀丽, 陈新庚, 等. 用灰色聚类法对荔湾湖水质富营养化程度的评价 [J]. 中国环境监测, 2004, 20(2): 47-50.
- [8] 郑晓红, 汪琴. 淀山湖水质状况及富营养化评价 [J]. 环境监测管理与技术, 2009, 21(2): 68-70.
- [9] 胡明星, 郭玲香, 郭达志. 湖泊水质富营养化评价的多准则神经网络法 [J]. 上海环境科学, 1998, 17(4): 14-16.
- [10] 张晓丹, 黄朝君, 徐茵, 等. 基于 RBF 神经网络的湖库水质富营养化程度评价模型 [J]. 环境工程, 2007, 25(2): 80-82.
- [11] 楼文高. 湖库富营养化人工神经网络评价模型 [J]. 水产学报, 2001, 25(5): 474-477.
- [12] 邓大鹏, 刘刚, 李学德, 等. 基于神经网络简单集成的湖库富营养化综合评价模型 [J]. 生态学报, 2007, 27(2): 725-731.
- [13] 中华人民共和国水利部. 2007 年中国水资源质量年报 [OL]. [2009-09-23]. <http://www.mwr.gov.cn/xygb/szyzlnb/index.aspx>
- [14] ROBERT H N. Theory of the back-propagation neural networks Proceedings of the International Joint Conference on Neural Networks [C]. San Diego, CA: HNC, Inc., 1989: 593-604.
- [15] 苑希民, 李鸿雁, 刘树坤, 等. 神经网络和遗传算法在水科学领域的应用 [M]. 北京: 中国水利水电出版社, 2002.
- [16] 秦传玉, 赵勇胜, 张伟红, 等. 基于 BP 神经网络的齐齐哈尔地区地下水水质评价 [J]. 环境监测管理与技术, 2007, 19(2): 15-18.

· 简讯 ·

美国有毒化学品排放减少 6%

中国化工报消息 美国环保局上周发布年度综合分析时称,美国 2008 年的有毒化学品排放量较上年下降了 6%。该局称,统计表明,2008 年美国制造业、公用设施及其他设备共向大气、土地及水中排放有毒化学品 38.6 亿磅,这一排放量较 2007 年下降了 6%。其中,改善最明显的是向大气中排放的化学品质量,下降了 14%;排向大地的排放量基本没变化;而进入水系的化学品质量则增加了 3%。

美国环保局称这是其首次在同一年度对当年收集到的数据进行年度分析。该局还称,全美国只有很少一部分工厂达到当前的联邦排放要求。2008 年该局共收到 21 000 家工厂共计 650 种化学品的排放数据,较上年减少了 5%。美国环保局表示,上报数据的工厂减少可能是一些厂由于经济衰退已关闭,该局计划对此进行进一步的调查。

2008 年,全美排放的包括铅、二恶英和水银等危险物质总量有所减少,但多氯联苯 (PCBs) 的排放量却增加了 121%。美国自 1979 年就禁止生产多氯联苯,而去年排放量却出现大幅增长,对此美国环保局认为是由于制造商为停止这类物质的使用,将其置于危险废物处置设备中所致。

摘自 www.jshh.gov.cn 2010-01-04