

基于人工神经网络的河湖蓝藻水华状态评价研究

李世明,王小艺,许继平,于家斌,王立

(北京工商大学 计算机与信息工程学院,北京 100048)

摘要: 面对依旧严峻的水环境问题,水资源的高效管理成为解决该问题的关键途径。其中,水华状态评价成为重中之重,本文针对河湖蓝藻水华状态评价过程中存在的非线性和不精确性问题,构建了基于人工神经网络的蓝藻水华状态评价模型,实现了固定站点监测和遥感监测信息的有效融合。并将该模型用于太湖蓝藻水华状态评价,研究表明:评价结果与实际情况相符,验证了模型的有效性和可行性,为蓝藻水华问题深入研究提供了思路。

关键词: 人工神经网络;信息融合;蓝藻水华;水华状态评价

中图分类号:X824

文献标识码:A

文章编号:1672-643X(2017)04-0093-04

Evaluation on the algae bloom of rivers and lakes based on artificial neural network

LI Shiming, WANG Xiaoyi, XU Jiping, YU Jiabin, WANG li

(College of Computer and Information Engineering, Beijing Technology & Business University, Beijing 100048, China)

Abstract: Faced with the serious water environment problem, the efficient management of water resources is the key way to solve this problem, and the state assessment of blooms has become a top priority. Aiming at the non-linear and imprecision problems in the evaluation of algae blooms in rivers and lakes, an artificial neural network based algae bloom state evaluation model was constructed, and the effective integration of fixed site monitoring and remote sensing information was realized. The model was applied to the evaluation of algae blooms in Taihu Lake. The evaluation results were in accordance with the actual situation, which verified the validity and feasibility of the model, and provided the idea for the further study of algae bloom.

Key words: artificial neural network; information fusion; algae bloom; bloom state evaluation

1 研究背景

水体富营养化已经成为全世界普遍存在的环境问题,2017年水利部发布的最新水资源公报显示,全国处于富营养化状态的湖泊占湖泊总数的78.3%,处于中营养化状态的水库占水库总数的63.5%^[1]。对水华状态的科学评价是水环境安全评估研究的重要内容,对有效防治水污染具有重要意义。目前,在河湖水质评价研究方面,主要有层次分析法^[2-3]、主成分分析法^[4-5]、模糊数学评价法^[6]等。随着智能信息技术的发展,人工神经网络被应用于水环境评价研究中^[7],如王俭等^[8]从阈值角度出发,建立了基于人工神经网络的区域水环境承载

力评价模型;郭庆春等^[9]基于误差反向传播的前向多层神经网络,建立了用于长江水质评价模型;崔东文等^[10]在对常用的RBF、GRNN、BP、Elman 4类神经网络进行验证基础上,应用至我国24个湖泊水体富营养化评价中;宫玉琳等^[11]将粒子群优化算法和RBF神经网络相结合,建立了水质评价模型;徐康耀等^[12]提出了一种由海豚群算法与BP神经网络结合的水体质量评价方法等。目前在蓝藻水华状态评价研究方面研究较少,一些学者进行了有益探讨,如陈文健等^[13]采用脉冲耦合神经网络(PCNN)对蓝藻水域图像进行分析,对蓝藻的暴发程度进行判别;王钰^[14]利用长时间序列的卫星遥感图像数据,分析了太湖蓝藻水华时空分布特征;刘文杰^[15]在地面实测

收稿日期:2017-04-04; 修回日期:2017-05-02

基金项目:国家自然科学基金项目(51179002);北京市教委科技计划重点项目(KZ201510011011);北京市属高校创新能力提升计划项目(PXM2014-014213-000033)

作者简介:李世明(1991-),男,山东齐河人,在读硕士研究生,研究领域为河湖水质状态信息融合方法。

水体光谱特征实验的基础上,结合现有卫星数据对内陆湖泊蓝藻水华进行了评价。上述方法主要以水质监测或遥感监测信息为基础,没有对反映水质特征的多源信息进行融合分析,导致评价结果容错性差,在实际应用中存在一定的限制与不足。

本文在考虑蓝藻水华评价过程中存在的非线性、空间分布及适用性特征前提下,以站点监测和遥感监测数据为信息源,以蓝藻水华表征指标的叶绿素 a 含量作为评价指标,构建了基于人工神经网络的蓝藻水华状态评价模型,并对太湖蓝藻水华状态进行实例验证。

2 蓝藻水华状态的 BP 人工神经网络评价模型

2.1 BP 人工神经网络

人工神经网络 (artificial neural network, ANN) 是理论化的人脑神经网络的数学模型,是基于模仿大脑神经网络结构和功能而建立的一种信息处理技术^[16]。在众多的神经网络模型中,BP 神经网络先进且比较成熟,适合于模式识别和信息融合分类^[17]。

BP 神经网络从结构上可以分为输入层、隐含层和输出层,各层之间权互联接,前层单元的输出不能反馈到更前层,同层之间没有连接,如图 1 所示。BP 神经网络实现了多层学习的设想,当给定网络一个输入模式时,由输入单元传到隐含层单元,经隐含层单元逐层处理后再送到输出单元,由输出层单元处理后产生一个输出模式;如果输出响应与期望输出有误差,则转入误差反向传播,将误差值沿连接通路逐层反向传送并修正各层连接权值,当各个训练模式都满足要求时,则学习结束^[18]。

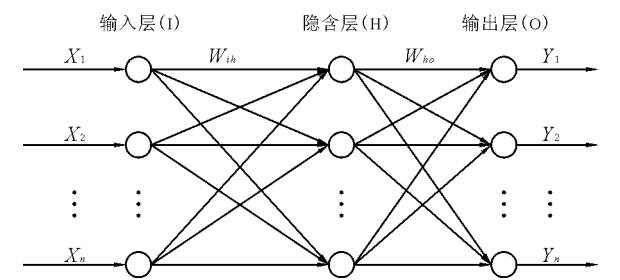


图1 BP神经网络结构

2.2 评价模型建立

根据水质固定监测站点的分布情况及蓝藻水华发生频率,选择太湖流域梅梁湾区域为对象,以太湖蓝藻水华发生风险评价指标体系作为评价依据^[19],如表 1 所示,根据叶绿素 a 含量的不同对蓝藻水华

等级进行划分,并给出了具体的水面表象描述。

表1 太湖蓝藻水华发生风险评价指标体系

水华等级	叶绿素 a 含量/ ($\text{mg} \cdot \text{L}^{-1}$)	水面现象
无水华	≤ 0.01	水面无藻类积累,水体清澈透明
轻度水华	$0.01 \sim 0.025$	水面有少量藻类积累,透明度较好
中度水华	$0.025 \sim 0.06$	水面有大面积藻类积累,透明度很低
重度水华	≥ 0.06	水面藻类积累厚度增加,透明度几乎为零

取太湖梅梁湾区域 5 个固定站点监测与遥感监测数据(如图 2 所示)。

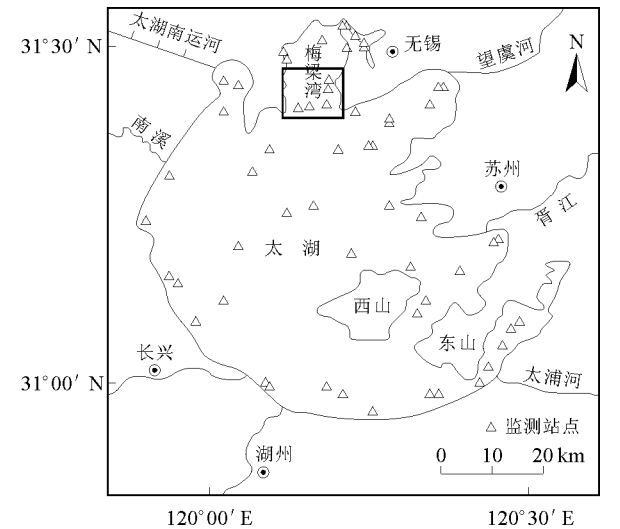


图2 固定监测点及监测区域

神经网络评价模型的输入层神经元个数为 6,分别对应 5 个固定监测站点的叶绿素 a 含量和遥感监测叶绿素 a 含量的反演数据。由于指标体系给出了对应于各类蓝藻水华等级的叶绿素 a 含量的上(下)限,因此,当叶绿素 a 含量在某个区间内时,蓝藻水华等级即属于相应类别。神经网络评价模型的输出为单输出,用 {1,2,3,4} 表示无水华、轻度水华、中度水华和重度水华状态。隐含层神经元数目根据 $\sqrt{m+n+a}$ 确定^[20],其中 m 为输入神经元数, n 为输出神经元数, a 为 1 ~ 10 的常数,经过多次测试取隐含层神经元数为 6。

选用多组监测样本进行学习和训练,确定第一层、第二层的传递函数为 tansig,训练函数为 traingdx,学习误差为 5×10^{-4} ,学习速率为 0.01,训练次数为 10^4 ,得到神经网络的权值和阈值,如表 2 所示。

3 蓝藻水华状态评价模型验证

将太湖流域梅梁湾区域内 5 个固定站点的监测

数据分别作为一个独立信息源,将反演平均后的遥感监测数据作为一个独立信息源,测试数据见表 3。

表 2 输入层与隐含层之间的连接权值和阈值

	隐层节点 1	隐层节点 2	隐层节点 3	隐层节点 4	隐层节点 5	隐层节点 6
输入 1	0. 8144	-0. 7420	-0. 9144	0. 6407	0. 8657	0. 4827
输入 2	-1. 4950	-0. 5768	-0. 4288	-0. 0396	0. 7739	0. 5495
输入 3	-0. 3787	-0. 3567	-0. 3152	1. 2521	0. 5374	0. 9381
输入 4	0. 2633	0. 0551	0. 0005	-1. 6541	-0. 7235	-0. 4639
输入 5	0. 6972	1. 0617	-0. 1854	-0. 3089	1. 1401	0. 6436
输入 6	-0. 3793	0. 3263	1. 1599	0. 9439	-0. 8427	0. 7527
阈值 1	-1. 9398	1. 0972	0. 5024	0. 3817	1. 1489	-1. 8046
输出	-0. 6089	-0. 4961	0. 6410	-0. 2026	0. 3956	0. 2719
阈值 2	0. 2415					

表 3 测试数据集

测试组	站点	叶绿素 a 含量/(mg · L ⁻¹)
第 1 组	站点 1	0. 0423
	站点 2	0. 0354
	站点 3	0. 0366
	站点 4	0. 0406
	站点 5	0. 0443
	遥感 6	0. 0422
第 2 组	站点 1	0. 008
	站点 2	0. 0439
	站点 3	0. 0466
	站点 4	0. 0046
	站点 5	0. 0063
	遥感 6	0. 01323
第 3 组	站点 1	0. 01403
	站点 2	0. 01981
	站点 3	0. 0884
	站点 4	0. 0589
	站点 5	0. 01274
	遥感 6	0. 02314
第 4 组	站点 1	0. 0024
	站点 2	0. 093
	站点 3	0. 0969
	站点 4	0. 0344
	站点 5	0. 0538
	遥感 6	0. 0609
第 5 组	站点 1	0. 0371
	站点 2	0. 0679
	站点 3	0. 01449
	站点 4	0. 0434
	站点 5	0. 02475
	遥感 6	0. 0311

采用归一化植被指数模型对叶绿素 a 含量及其对应的遥感波段反射率比值进行回归分析,对模型参数中存在的误差采用集成学习 Bagging 算法进行修正,得到遥感反演模型为:

$$y = 0. 0338x + 0. 0013 \tag{1}$$

式中:y 为叶绿素 a 含量;x 为遥感波段反射率。

对蓝藻水华状态神经网络评价模型进行测试,结果如图 3 所示。

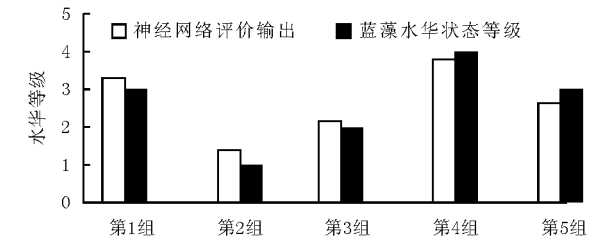


图 3 蓝藻水华状态神经网络评价模型测试结果

由图 3 可知,5 组测试数据的输出与蓝藻水华状态期望等级基本一致,其中第 1 组测试数据表示 2010 年 2 月太湖梅梁湾区域的水体处于中度水华状态,该结果与江苏省发布的太湖水环境质量状况公告相符^[21]。

4 结 论

本文建立了基于神经网络的蓝藻水华状态评价模型,该模型在太湖梅梁湾湖域的蓝藻水华状态评价中结果准确且与事实相符。证明了该模型适用于水华状态评价,能够解决状态评价过程中的非线性和不精确性问题,实现对反映水质特征的站点监测信息与遥感监测信息的有效融合。验证了该模型具有良好的稳定性、容错性和可靠性,为有效防治蓝藻

水华污染提供决策依据,对深入研究蓝藻水华问题提供有效途径。

参考文献:

- [1] 中华人民共和国水利部. 2015 年中国水资源公报[EB/OL]. [2017-01-19]. http://www.mwr.gov.cn/zwzc/hygb/szygb/qgszygb/201701/t20170119_785887.html.
- [2] 李新, 石建屏, 曹洪. 基于指标体系和层次分析法的洱海流域水环境承载力动态研究[J]. 环境科学学报, 2011, 31(6):1338-1344.
- [3] 杨钧婷, 陈秋波, 彭黎旭. 基于联合层次分析法和“模加和”法的海南水环境承载力评价[J]. 南水北调与水利科技, 2013, 11(3):47-51.
- [4] 周广峰, 刘欣. 主成分分析法在水环境质量评价中的应用进展[J]. 环境科学导刊, 2011, 30(1):75-78.
- [5] 万金保, 曾海燕, 朱邦辉. 主成分分析法在乐安河水质评价中的应用[J]. 中国给水排水, 2009, 25(16):104-108.
- [6] 凌敏华, 左其亨. 水质评价的模糊数学方法及其应用研究[J]. 人民黄河, 2006, 28(1):34-36.
- [7] 王栋, 曹升乐. 人工神经网络在水文水资源水环境系统中的应用研究进展[J]. 水利水电技术, 1999, 30(12):4-7.
- [8] 王俭, 孙铁珩, 李培军, 等. 基于人工神经网络的区域水环境承载力评价模型及其应用[J]. 生态学杂志, 2007, 26(1):139-144.
- [9] 郭庆春, 何振芳, 李力. 基于人工神经网络的长江水质评价模型[J]. 信息技术, 2013(8):54-56.
- [10] 崔东文. MATLAB 神经网络在湖库富营养化程度评价中的应用——以全国 24 个湖泊富营养化程度评价为例[J]. 环境研究与监测, 2012, 1(25):42-48.
- [11] 宫玉琳, 大化, 张静鹤. 基于粒子群优化算法和 RBF 神经网络的水质评价方法研究[J]. 长春理工大学学报(自然科学版), 2014, 37(5):141-145.
- [12] 徐康耀, 葛考, 赵建强, 等. 基于 DPA-BP 神经网络的地下水水质综合评价[J]. 节水灌溉, 2015(9):66-69.
- [13] 陈文健, 莫利, 杨建华. 基于脉冲耦合神经网络的太湖蓝藻水华程度分析[J]. 福建电脑, 2008, 24(8):97-97+105.
- [14] 王牲. 基于 MODIS 数据的太湖蓝藻水华时空分布特征研究[D]. 南京:中国科学院南京地理与湖泊研究所, 2009.
- [15] 刘文杰. 内陆湖泊蓝藻水华的遥感监测与评价研究[D]. 北京:中国地质大学, 2013.
- [16] 丛爽. 面向 Matlab 工具箱的神经网络理论与应用[M]. 合肥:中国科学技术大学出版社, 2009:55-57.
- [17] 杨丽花, 佟连军. 基于 BP 神经网络模型的松花江流域(吉林省段)水环境承载力研究[J]. 干旱区资源与环境, 2013, 27(9):135-140.
- [18] 胡守仁. 神经网络导论[M]. 北京:国防科技大学出版社, 1993:129-130.
- [19] 李颖, 施择, 张榆霞, 等. 关于用藻密度对蓝藻水华程度进行分级评价的方法和运用[J]. 环境与可持续发展, 2014, 39(2):67-68.
- [20] 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, 1993:134-135.
- [21] 江苏省环境保护厅. 2010 年 2 月江苏省水环境质量状况[EB/OL]. [2010-03-17]. http://www.jshb.gov.cn/jshbw/hbzl/szyb/201005/t20100527_155397.html.