

神经网络模型在水华预警领域的应用及展望

张淑珍^{1,2}, 王培京¹, 张祎梦³

¹北京市水科学技术研究院, 北京

²北京市南水北调水质监测中心, 北京

³北京市政路桥建材集团有限公司, 北京

Email: love_shuzhen@sina.com

收稿日期: 2020年9月26日; 录用日期: 2020年10月13日; 发布日期: 2020年10月20日

摘要

从神经网络模型的特点入手, 详细介绍了神经网络模型在水华预警中的研究进展。通过梳理神经网络模型在藻类生物量预测和水华风险评估两方面的应用情况, 分析其在应用中存在的问题及优化方法, 并展望未来研究方向, 以期推动和深化水华预警模型研究。

关键词

神经网络, 水华, 预警模型, 研究进展

Application and Prospects of Neural Network Model on Algal Bloom Early-Warning

Shuzhen Zhang^{1,2}, Peijing Wang¹, Yimeng Zhang³

¹Beijing Water Science and Technology Institute, Beijing

²Water Quality Monitoring Center of Beijing South-to-North Water Diversion, Beijing

³Beijing Municipal Road & Bridge Building Material Group Co., Ltd., Beijing

Email: love_shuzhen@sina.com

Received: Sep. 26th, 2020; accepted: Oct. 13th, 2020; published: Oct. 20th, 2020

Abstract

At the beginning of this paper, the specific property of neural networks and the significance of its

application in algal bloom early-warning were introduced briefly. The research and application status of neural network models in both prediction of algal biomass and algal bloom risk assessment, as well as existing problems and improved methods in the model study were analyzed. The future research direction was prospected in order to promote and deepen the study on algae bloom early-warning model.

Keywords

Neural Networks, Algae Bloom, Early-Warning Model, Research Advances

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

神经网络模型是一种模拟大脑神经网络结构和功能的信息处理方式,具有自适应性、自组织性、容错性和不需人为确定模型参数等特征,适合解决高维非线性问题,常用于分类、识别、预测等问题[1][2]。神经网络最初是人工智能领域的一种方法,目前已经发展成为一类多学科交叉的学科领域,为各领域各专业的研究提供了新的思路和方法。

20世纪40年代M-P神经元的提出标志着神经网络模型的诞生。随着计算平台和学习算法的发展,神经网络模型由传统的生物网络简化形式的浅层神经网络逐渐向模拟人脑复杂的层次化认知特点的深层神经网络发展,到今天已经衍生出了上百种神经网络模型,通过模拟其他行业,如计算机、经济、医学、环境等,实现数据分析和利用,在数据预测、图像分类、智慧城市等领域取得非常成功的应用。目前的神经网络模型的功能大都局限在对数据的理解层面,缺乏环境感知和推理决策功能。另一方面,深层神经网络在理论上存在很多局限,加之呈指数增长的海量复杂数据,给统计学习意义下的神经网络模型的结构设计、参数选取、训练算法,以及时效性等方面都提出了新的挑战[3][4]。

水华是当水体中出现富营养状况并具备适宜的温度、光照、气候及适宜的水动力条件等有利于藻类生长或聚集的环境条件时,水体中藻类大量生长繁殖或聚集并达到一定浓度的现象[5]。浮游藻类过量增长会降低水体透明度,恶化水质(色、臭),形成缺氧,有时还产生毒素等,严重破坏水生生态系统,威胁人类的饮水安全,已成为水环境保护中亟待解决的首要问题。水华的发生具有突发性,一旦发生控制难度会加大,治理成本成倍提高,如果能够预见水华的发生,及时采取相应措施,会取得事半功倍的效果。水华预警是在水华暴发前,根据历史记录和实时监测数据对水华发展趋势、强度及风险进行分析,判断水华发生的可能性,预先进行警示,以便采取有针对性的防控措施。利用模型被认为是研究及预报水华的有效手段,其中神经网络模型是最主要的水华预警模型之一[6]。

2. 神经网络模型在水华预警中的应用现状

2.1. 神经网络模型在水华预警领域应用的适用性

水华是由浮游藻类生长特性、水体理化特性、生物、气象及水文等因子共同作用所致,各因子之间关系复杂,存在随机性、不确定性和非线性特征,目前对于其发生的临界条件和机理不完全清楚[7][8][9]。浮游植物与营养盐相关模型主要研究在不同营养物条件下浮游植物的生长情况,适用于只有一种限制性营养物质的均匀水体,而实际水体情况要复杂得多,其预测精度较低。生态动力学模型是用来模拟生态

系统受物理、化学、生物因素影响后的动态变化,模拟结果接近真实环境,但该类模型结构复杂,参数繁多,需要大量数据支持,参数率定十分困难。多元线性回归、自回归移动平均模型等传统统计模型一般采用线性关系进行简化,当水华暴发的限制性因素发生变化时预测效果往往不佳。因此,确定性的机理模型以及传统的统计模型在水华预警中的应用均存在一些困难与问题。

自 20 世纪 80 年代以来,神经网络、遗传算法、粒子群算法和蚁群算法等试图通过自然生态机制来求解复杂优化问题的新型智能算法相继被提出,为环境建模提供一种新的思路[10] [11]。其中神经网络能够有效地解决复杂的非线性问题,在环境模型优化中有广泛的应用。对于水华这类具有高度非线性和突变性的复杂生态活动,基于神经网络模型模拟预测浮游植物的生长及繁殖的水华建模预测,能将水华发生时间与程度预测到一个合理的范围内,模拟藻类动态变化过程,为水华预警提供了一种有效途径。基于误差反向传播算法(Error Back Propagation Training)的多层前馈神经网络(简称 BP 神经网络)是最常用的水华预警模型,由输入层、隐含层和输出层组成,隐含层可以有若干层,每一层的神经元只接受前一层神经元的输出。输入层的神经元节点获取输入样本数据的基本信息,根据随机产生的输入权值和阈值将信息传递给隐含层的神经元节点;隐含层神经元通过激活函数对传入隐含层的数据进行处理;在隐含层通过激活函数处理的数据最终输出到输出层。神经网络水华预警模型的构建步骤为:

1) 筛选模型变量。利用文献调研、数据分析、实验模拟等方法选择影响水华发生的最关键因子作为模型的输入变量。模型的输出变量一般选择表征藻类生物量的指标,如叶绿素 a 浓度或藻密度等。

2) 确定网络结构。根据输入变量、输出变量确定模型输入层和输出层的神经元节点数,进而确定隐含层的层数和神经元节点数。

3) 模型训练与优化。通过输入、输出参数历史数据构成的训样本和学习算法,寻找各层神经元之间连接的权值和阈值等参数的最优组合,对数据间的非线性因果映射关系进行感知与学习,即完成模型的训练。

4) 实现预测。依据未来影响水华发生的关键因子的实测值,利用训练好的模型来定量预测水华预警指标的变化。

从模型构建步骤不难看出,神经网络模型的性能受模型输入变量、网络结构、学习算法以及训练样本的影响。选取适合的输入变量、收集足够的历史数据是建模的前提,设计合理的网络结构、采取适合的学习算法是建模的基础。

2.2. 在藻类生物量预测方面的应用

藻类生物量通常以藻密度、叶绿素 a 浓度等指标表征,能够在一定程度上反映水华发生的强度[12] [13]。预测藻类生物量是水华预警研究中最重要的一部分内容。French 和 Recknagel [14]最早将神经网络用于藻类生物量的预测。该研究以 pH、营养盐等水质指标及浮游动物生物量为模型输入变量,特定藻类的生物量为输出变量,通过误差反向传播(BP)算法训练模型,构建了一个三层的 BP 神经网络模型,实现了对德国 Saldenbach 水库 7 种常见优势藻种的藻密度及其随时间变化情况的模拟。该模型基本上能够反映藻类种群生物量的变化趋势,如高峰和低谷,但对藻类生物量大小的预测精度还有待进一步提高。此后,国内外学者纷纷构建了各类神经网络模型,成功实现了对水库[15] [16]、湖泊[17] [18]、河流[19]、海岸线[20] [21]等不同环境条件水体的藻类生物量预测。

早期的研究往往选取水华发生机理研究中较关注的影响因子,如营养盐、水温、太阳辐射及浮游动物数量等因子作为模型输入变量,显然该方法受当前理论认识的局限,网络输入变量选择任意性较大。为此,许多学者引入了输入变量筛选方法。Zeng 等[22]基于 BP 神经网络的北京六海水华预警模型研究中,将影响水华发生的各因子与叶绿素 a 浓度做回归分析,选择相关系数较大指标作为模型输入变量。夏杰

等[23]利用主成分分析法对影响浮游生物总量的理化因子进行主要影响因子分析,并以确定的主要理化因子作为网络的输入参数。结果表明基于主成分分析的 BP 神经网络模型对浮游生物总量的预测结果与真实值的拟合度高达 0.99 以上。桑文璐等[24]在三峡水库香溪河的叶绿素 a 浓度预测研究中,首先通过聚类分析将形似或相关的指标进行分类,然后根据灰色关联度理论,以叶绿素 a 浓度为母因素,各富营养化因素为子因素,计算两者的关联度,依据分类结果和关联度排序选择富营养化主控因子,即为模型输入变量。回归分析、主成分分析、聚类分析、关联度分析等方法的引入一定程度上克服了神经网络结构设计上的盲目性,为输入变量的筛选提供了更可靠的方法。但是,输入变量的筛选方法多样,至今没有统一标准,且选出的输入变量也不同。因此,所建神经网络模型泛化能力差,只能在小范围内应用。

2.3. 在水华风险评估方面的应用

水华风险是指水华发生强度及其可能性。对于尚未发生水华的水体提前进行水华风险评估,确定其暴发水华的可能性和风险性,可有效减少水华的发生及其带来的影响。李杰等[25]使用抚仙湖历史水环境和藻类生物量数据作为训练样本,训练得到成熟的神经网络模型。随后将国家地表水质标准中的 I 类、II 类、III 类、IV 类、V 类水质标准中总氮、总磷、高锰酸盐指数、氨氮参数输入训练成熟的神经网络模型,得到不同水质类别下浮游植物生物量,并以此将水华风险等级分为初级、中级和强烈,用来评估抚仙湖水华风险。王巍巍等[26]依据藻密度大小划分了 4 个水华发生级别,并基于 BP 神经网络和广义回归网络实现华南地区 S 水库水华风险评估。

由于水华发生机理和临界特征尚不明确,如何表征水华风险及其风险等级尚未有定论,并且水华是否暴发、暴发强度和范围具有很大的不确定性,因此神经网络模型常常与模糊理论结合用于水华风险的评估。郑剑锋等[27]基于内集-外集模型提出了水华风险概率,划分了水华预警等级,并基于粗糙集约简-径向基函数(RBF)神经网络模型对天津清静湖水华预警等级进行了预测,预测精度达 85.7%。张艳会等[28]采用自组织特征映射神经网络、模糊风险评价法及 GIS 地学统计分析法对全太湖水华发生风险进行评价和区划,揭示了太湖水华灾害风险的空间分布规律。

2.4. 神经网络水华预警模型的优化

2.4.1. 敏感性分析

尽管神经网络模型被广泛应用于水华预警研究,但神经网络模型存在着黑箱特性,无法直接分析模型输入变量与输出变量的响应机制。为此,许多学者使用敏感性分析方法来分析输入变量对输出的影响,从而确定哪些环境变量对浮游藻类生长、水华发生发展起主要作用。张克鑫等[29]运用 BP 神经网络模型预测宁波大学校内池塘颤藻生物量,并通过基于扰动的敏感性分析发现,pH 的变化对池塘中颤藻生物量的影响较为显著,说明 pH 是该水体藻类水华的主要驱动因素。敏感性分析还可用于判断所建模型是否合理,弥补神经网络模型机理不明的问题。Park 等[15]为预测韩国 Juam 水库和 Yeongsan 水库叶绿素 a 浓度,分别建立了神经网络模型和支持向量机模型,应用 LH-OAT (Latin-Hypercube One-factor-At-a-Time) 方法对模型输入变量进行了敏感性分析和排序,与藻类生长影响因素比的分析后发现所构建神经网络模型没有支持向量机模型合理。

此外,基于神经网络模型和敏感性分析可以考察水华发生时常见影响因素的范围,为水华预警提供支撑。王重阳等[30]基于 BP 神经网络模型,采用全样本敏感性分析法,在多个环境因素共同影响下,分别分析了不同理化因子与模型输出的影响,依据全样本敏感性系数最高的范围提出了水华爆发时不同理化因子的范围。

2.4.2. 混合模型

神经网络模型需要大量的历史数据来训练,给实际应用带来了一定的困难,同时它的学习收敛速度

还有待提高。因此,一些学者运用混合模型的方法,将神经网络与其他方法相结合,取长补短,继而获得更好的应用效果。目前这方面工作有神经网络与模糊逻辑、遗传算法、小波分析、混沌、粗集理论和灰色系统等的融合。Xi 等[16]借助浮标式荧光探头的叶绿素 a 浓度或藻密度监测数据,建立了单一输入参数的小波神经网络模型。该模型能够有效、准确的处理非平稳数据,比单独神经网络模型和传统的统计模型预测结果更准确。吴巧媚等[31]构建了结合小波变换与 BP 神经网络的水华预测模型对北海叶绿素 a 进行预测,小波变换用于对叶绿素 a 信号进行降噪处理。与 BP 神经网络模型对比,该混合模型能反映叶绿素 a 的变化趋势,且预测精度更高。王德喜等[32]建立了基于灰色-BP 神经网络的水华智能预测模型,该混合模型兼具灰色理论模型良好的拟合和和外推特性,以及神经网络非线性预测优势,适合于小样本条件、长期、高精度的藻类水华预测。

3. 结论

近年来,神经网络模型在水华预警研究中取得了很大发展,其成功应用于各类水体藻类生物量预测和水华风险评估。神经网络模型结构设计存在随意性,具有黑箱特性,且需要大量数据训练模型。为了克服这些缺陷,在模型输入变量选取、敏感性分析和混合模型等方面对模型进行了优化。

4. 展望

结合当前神经网络在水华预警中的研究现状及存在问题,神经网络水华预警模型亟待在以下方面强化:

1) 随着实时在线监测技术、遥感监测技术等监测技术的发展,快速、大范围、动态的水华监测越来越普及,进而产生了更丰富、更密集、更庞大的数据[33][34][35]。杨斌等[36]为提高 RBF 神经网络水华预测模型的预测精度,将遥感监测信息作为预测因子之一。赵文喜等[17]以海河干流段典型断面的水质在线监测及气象站高频、实时数据为基础构建了基于 BP 神经网络的叶绿素 a 日变化量预测模型。石野等[37]将基于灰色-神经网络的水华预测模型嵌入到水质无线远程监测系统中,实现对湖库水华的智能预测。因此,在线监测、遥感监测、物联网等高新技术与神经网络水华预警模型的融合取得了较好的应用效果,已引起了极大的关注。

2) 水华发生的复杂性决定了任何一个模型很难精准模拟水华发生。加强对水华发生影响因素的分析、筛选,将有助于进一步提高模型对复杂输入的处理效率和预测精度,仍是未来应用的研究重点之一。同时,神经网络模型是黑箱模型、机理不明,将神经网络模型与机理模型相结合的水华预警研究思路将是一个可行的方向。

基金项目

水体污染控制与治理科技重大专项 2018 年项目(2018ZX07101005)。

参考文献

- [1] 袁冰清, 陆悦斌, 张杰. 神经网络与深度学习基础[J]. 数字通信世界, 2018(5): 32-33.
- [2] 徐学良. 人工神经网络的发展及现状[J]. 微电子学, 2017, 47(2): 239-242.
- [3] 焦李成, 杨淑媛, 刘芳, 等. 神经网络七十年:回顾与展望[J]. 计算机学报, 2016, 39(8): 1697-1716.
- [4] 赵崇文. 人工神经网络综述[J]. 山西电子技术, 2020(3): 94-96.
- [5] 郑建军, 钟成华, 邓春光. 试论水华的定义[J]. 水资源保护, 2006(5): 45-47.
- [6] 徐宪根, 周游, 黄骏, 等. 蓝藻水华发生机理及预警研究进展[J]. 污染防治技术, 2017, 30(5): 22-24+54.
- [7] Lee, T.A., Rollwagen-Bollens, G. and Bollens, S.M. (2015) The Influence of Water Quality Variables on Cyanobacterial Blooms and Phytoplankton Community Composition in a Shallow Temperate Lake. *Environmental Monitoring and Assessment*, **187**, 315. <https://doi.org/10.1007/s10661-015-4550-2>

- [8] 黄鹏, 田腾飞, 张文安, 等. 换水周期对直流河道藻类生长的影响[J]. 环境工程, 2018, 36(7): 22-26.
- [9] Trombetta, T., Vidussi, F., Mas, S., *et al.* (2019) Water Temperature Drives Phytoplankton Blooms in Coastal Waters. *PLoS ONE*, **14**, e214933. <https://doi.org/10.1371/journal.pone.0214933>
- [10] 黄佳聪, 高俊峰. 智能算法及其在环境预警中的应用[J]. 环境监控与预警, 2010, 2(3): 5-8+17.
- [11] 易仲强. 智能算法在湖库富营养化预测中的应用研究综述[J]. 水电能源科学, 2010, 28(8): 33-36+26.
- [12] Boyer, J.N., Kelble, C.R., Ortner, P.B., *et al.* (2009) Phytoplankton Bloom Status: Chlorophyll a Biomass as an Indicator of Water Quality Condition in the Southern Estuaries of Florida, USA. *Ecological Indicators*, **9**, S56-S67. <https://doi.org/10.1016/j.ecolind.2008.11.013>
- [13] 杜胜蓝, 黄岁樑, 臧常娟, 等. 浮游植物现存量表征指标间相关性研究 II: 叶绿素 a 与藻密度[J]. 水资源与水工程学报, 2011, 22(2): 44-49.
- [14] French, M. and Recknagel, F. (1994) Modeling of Algal Blooms in Freshwaters Using Artificial Neural Networks. *Computational Mechanics*, **2**, 87-94.
- [15] Park, Y., Cho, K.H., Park, J., *et al.* (2015) Development of Early-Warning Protocol for Predicting Chlorophyll-a Concentration Using Machine Learning Models in Freshwater and Estuarine Reservoirs, Korea. *Science of the Total Environment*, **502**, 31-41. <https://doi.org/10.1016/j.scitotenv.2014.09.005>
- [16] Xiao, X., He, J., Huang, H., *et al.* (2017) A Novel Single-Parameter Approach for Forecasting Algal Blooms. *Water Research*, **108**, 222-231. <https://doi.org/10.1016/j.watres.2016.10.076>
- [17] Wang, Z., Huang, K., Zhou, P., *et al.* (2010) A Hybrid Neural Network Model for Cyanobacteria Bloom in Dianchi Lake. *Procedia Environmental Sciences*, **2**, 67-75. <https://doi.org/10.1016/j.proenv.2010.10.010>
- [18] Wang, H., Zhu, R., Zhang, J., *et al.* (2018) A Novel and Convenient Method for Early Warning of Algal Cell Density by Chlorophyll Fluorescence Parameters and Its Application in a Highland Lake. *Frontiers in Plant Science*, **9**, 869. <https://doi.org/10.3389/fpls.2018.00869>
- [19] 赵文喜, 周滨, 刘红磊, 等. 基于 BP 神经网络的海河干流叶绿素浓度短时预测研究[J]. 水利水电技术, 2017, 48(11): 134-140.
- [20] Velo-Suarez, L. and Gutierrez-Estrada, J.C. (2007) Artificial Neural Network Approaches to One-Step Weekly Prediction of *Dinophysis acuminata* Blooms in Huelva (Western Andalusia, Spain). *Harmful Algae*, **6**, 361-371. <https://doi.org/10.1016/j.hal.2006.11.002>
- [21] Melesse, A.M., Krishnaswamy, J. and Zhang, K. (2008) Modeling Coastal Eutrophication at Florida Bay Using Neural Networks. *Journal of Coastal Research*, **2**, 190-196. <https://doi.org/10.2112/06-0646.1>
- [22] Zeng, W., Song, Q., Liu, H., *et al.* (2010) Research on ANN-Based Pre-Warning Water Bloom Model of LiuHai Lake in Beijing. *Procedia Environmental Sciences*, **2**, 625-635. <https://doi.org/10.1016/j.proenv.2010.10.070>
- [23] 夏杰, 吴文青, 许海洋. 基于主成分分析的 BP 神经网络在水华预测中的应用[J]. 世界生态学, 2018, 7(2): 53-60.
- [24] 桑文璐, 纪道斌, 朱士江, 等. 基于 WNN 和 SVM 模型的香溪河 Chl-a 浓度预测[J]. 环境科学与技术, 2018, 41(S2): 95-99.
- [25] 李杰, 张玲芬. 抚仙湖藻类水华暴发风险的神经网络模型预测[J]. 环境科学导刊, 2015, 34(6): 4-7.
- [26] 王巍巍, 武延坤, 朱佳. 基于神经网络的水库藻类预警模型研究[J]. 给水排水, 2015, 51(11): 103-105.
- [27] 郑剑锋, 焦继东, 孙力平. 基于神经网络的城市内湖水华预警综合建模方法研究[J]. 中国环境科学, 2017, 37(5): 1872-1878.
- [28] 张艳会, 李伟峰, 陈求稳. 太湖蓝藻水华发生风险区划[J]. 湖泊科学, 2015, 27(6): 1133-1139.
- [29] 张克鑫, 陆开宏, 朱津永, 等. 基于 BP 神经网络的藻类水华预测模型研究[J]. 中国环境监测, 2012, 28(3): 53-57.
- [30] 王重阳, 马菁华, 吴睿钰, 等. 基于 BP 神经网络的水华预测模型及其敏感性分析[J]. 信息记录材料, 2018, 19(11): 81-83.
- [31] 吴巧媚, 刘载文, 王小艺, 等. 小波神经网络在北京河湖水华预测中的应用[J]. 计算机工程与应用, 2010, 46(12): 233-235.
- [32] 王德喜, 王晓凯, 王小艺. 基于灰色 - 神经网络的城市湖库水华预测研究[J]. 测试技术学报, 2013, 27(4): 349-353.
- [33] 熊晶, 黄丹, 沈帆, 等. 湖北省水华预警监测体系的探讨[J]. 环境科学与技术, 2015, 38(S1): 312-315.
- [34] 郁建桥, 吕学研. 太湖遥感可测性蓝藻水华发生条件分析[J]. 环境科学与技术, 2015, 38(6): 93-98.
- [35] 尹美琳, 张嘉琪, 路春晖, 等. 蓝藻实时监测技术综述[J]. 价值工程, 2019, 38(22): 208-209.
- [36] 杨斌, 许继平, 王凌斌, 等. 湖泊藻类水华神经网络预测研究[J]. 工业控制计算机, 2014, 27(4): 52-53.
- [37] 石野, 王晓凯, 张春梅. 湖库水质无线远程监测与水华预测方法研究[J]. 测试技术学报, 2019, 33(3): 216-222.