

Statistical Regression Analysis of Water-Bloom Algae and Related Environment Factors of a Reservoir in Shenzhen City

Ying Yang, Yan Wang*, Wenhui Huang, Weiming Wang

School of Resource and Environmental Science, Wuhan University, Wuhan Hubei
Email: *wyan@whu.edu.cn

Received: Jan. 8th, 2016; accepted: Jan. 26th, 2016; published: Feb. 5th, 2016

Copyright © 2016 by authors and Hans Publishers Inc.

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

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



Open Access

Abstract

Based on five-year monitoring data of Xikeng reservoir in Shenzhen City from 2000 to 2004, data of 11 environment factors (Total phosphorus, water depth, etc.) and 3 biotic factors (algae biomass, Chl-a and blue algae biomass) were employed, and the correlations between these factors were analyzed using multi-statistic method of stepwise based on SPSS. On the basis of the analysis, considering the factors which were significantly related to blue algae biomass, a prediction model of algae biomass was developed. The results showed that algae biomass was mainly affected by water temperature, water depth and total phosphorus concentration. However, the significant factors for different sites were different, which mean the limited factors of algae growth were different in different environment. Multi-statistic method could be helpful to identify key factors but could not supply good prediction results.

Keywords

Algae Biomass, Multi-Statistic Method of Stepwise, Environmental Factor, Prediction

深圳某水库水华藻类及相关环境因子逐步回归统计分析

杨滢, 王艳*, 黄文辉, 王炜铭

作者简介: 杨滢(1994-), 女, 广东广州人, 大学本科, 地理科学专业, 研究方向是富营养化模拟。
*通讯作者。

文章引用: 杨滢, 王艳, 黄文辉, 王炜铭. 深圳某水库水华藻类及相关环境因子逐步回归统计分析[J]. 水资源研究, 2016, 5(1): 33-39. <http://dx.doi.org/10.12677/jwrr.2016.51004>

武汉大学资源与环境科学学院, 湖北 武汉
Email: *wyan@whu.edu.cn

收稿日期: 2016年1月8日; 录用日期: 2016年1月26日; 发布日期: 2016年2月5日

摘要

以深圳市茜坑水库5个监测点连续5年的监测资料为基础, 选择总磷、水深等环境理化因子, 以及藻类生物量、叶绿素a和蓝藻生物量等生物因素, 利用SPSS软件, 进行多元逐步回归分析, 分别研究与生物因素显著相关的环境理化因子。并在此基础上, 采用逐步回归统计方法, 利用与生物因素显著相关的环境因子, 建立了茜坑水库藻类生物量变化的初步预测模式。结果表明, 剔除次要环境因子后, 水库藻类生物量变化主要受水库的水深, 水温及总磷的影响, 但不同监测点的显著因子有差异, 说明藻类生长因环境不同, 限制因子也不同。逐步回归统计模型, 能够寻找影响藻类生物量的关键性因子, 为准确预测提供参考, 但尚不能对藻类爆发性生长做出精确的预测。

关键词

藻类生物量, 多元逐步回归, 环境因子, 预测

1. 引言

近年来湖泊水库的富营养化问题日益突出, 影响人们生产生活及供水安全。学者们针对富营养化机理和预测模式, 开展了大量研究工作。富营养化水质模型是研究水质动态变化的有力工具, 为富营养化研究提供便捷的方法, 为水华的防治提供有力的工具。自20世纪70年代Vollenweider提出了简单总磷模型以来, 经过众多科学家的努力, 富营养化模型已经得到了很大程度的发展。近年来, 富营养化机理性模型中发展较为完善的是生态动力学模型。该类模型涵盖了生态系统中影响藻类生物量的大部分生态学过程, 较为完善。国外已有相关模型[1][2], 如PACGAP模型, 即藻类种群生长和生产力的预测模型; PROTECH-2模型, 即浮游植物与环境因子关系模型[3]。也有部分商业软件, 如美国EPA开发的WASP中富营养模块EUTRO和CE-QUAL2。自上世纪80年代以来, 我国学者在湖泊富营养化模型构建方面也做了大量工作, 特别是在一些富营养化较为严重的湖泊, 如太湖、滇池、巢湖等, 构建了较好的富营养化生态动力学模型。

这些生态动力学模型虽然尝试从机理上反映藻类生物量的变化, 但涉及的生态学过程复杂, 参数众多, 对数据量要求大, 并且取样频次与取样时间对模型参数测定有较大影响, 给模型的参数率定和验证造成了很大困难, 且多数主要预测藻类总生物量或叶绿素a, 专门针对蓝藻水华的模拟预测模型尚不多见。

而采用逐步多元回归方法, 分析环境因子与藻类相关关系, 方法简单易行, 在实际监测资料的基础上, 能较好的明确在特定水体中, 环境因子与富营养化藻类的关系, 为水华藻类机理模型提供较好的参数分析基础。国内学者们分别在太湖、洪泽湖、千岛湖、淀山湖、巢湖等水体[3]-[12]进行了相关研究。但由于不同水体有各自的特性, 需结合各自的情况具体分析。

深圳市茜坑水库是深圳市重要的饮用水水源地之一, 富营养化问题一直是该水库重点防治问题。为了给该水库管理部门防治富营养化提供理论支持, 本研究采用该水库连续5年的监测资料, 采用逐步回归的统计方法, 分析藻类生物量与相关环境理化因子的相关关系, 研究对藻类生物量和水华蓝藻生长影响较显著的因子, 为水库水华防治工作提供技术参考。同时筛选关键参数, 为建立该水库富营养化蓝藻预测模型提供一定的基础。

2. 研究方法

2.1. 数据来源

深圳市茜坑水库是深圳市重要的水源地之一,为全面监控水库的水质状况,该水库管理部门自2000年开始,在水库布设采样点,每月采样1~2次,分析了氨氮、总氮、总磷、溶解氧、透明度等化学指标,以及蓝绿藻及藻类总生物量等生物指标,取得了丰富的监测资料。水库共布设6个监测点,各点位置布设如图1所示。

由于监测点VI为后来设立,数据不连续,所以选用监测点I~V连续5年的监测资料进行分析。选择的理化因子包括:引水量、水深、水温、透明度、pH值、化学需氧量(COD)、溶解氧(DO)、氨氮(NH₃-N)、总氮(TN)、总磷(TP)等,其中溶解氧包括表层、4m水深处和8m水深处数据。生物因子分析包括叶绿素a、藻类生物量和蓝藻的生物量。

2.2. 统计分析

借鉴陈宇炜等[3][5][13][14]的方法,采用SPSS软件,构建回归方程,具体如下:

(1) 采用SPSS软件,对各采样点,以各理化因子为自变量,叶绿素a为因变量,选择临界值95%的显著水平($FC = FJ = 3.0$)筛选自变量[3][4],得到回归方程和复相关系数,分析得到各采样点的与叶绿素a关系最接近的环境因子。

(2) 在上述分析结果基础上,从理化因子中,剔除pH, DO, COD, SD等5项,以剩余6项理化因子为自变量,以藻类总生物量和蓝藻生物量为因变量,采用SPSS分析,分别得到两项生物因子(藻类总生物量和蓝藻生物量)的逐步回归预测方程和复相关系数等。

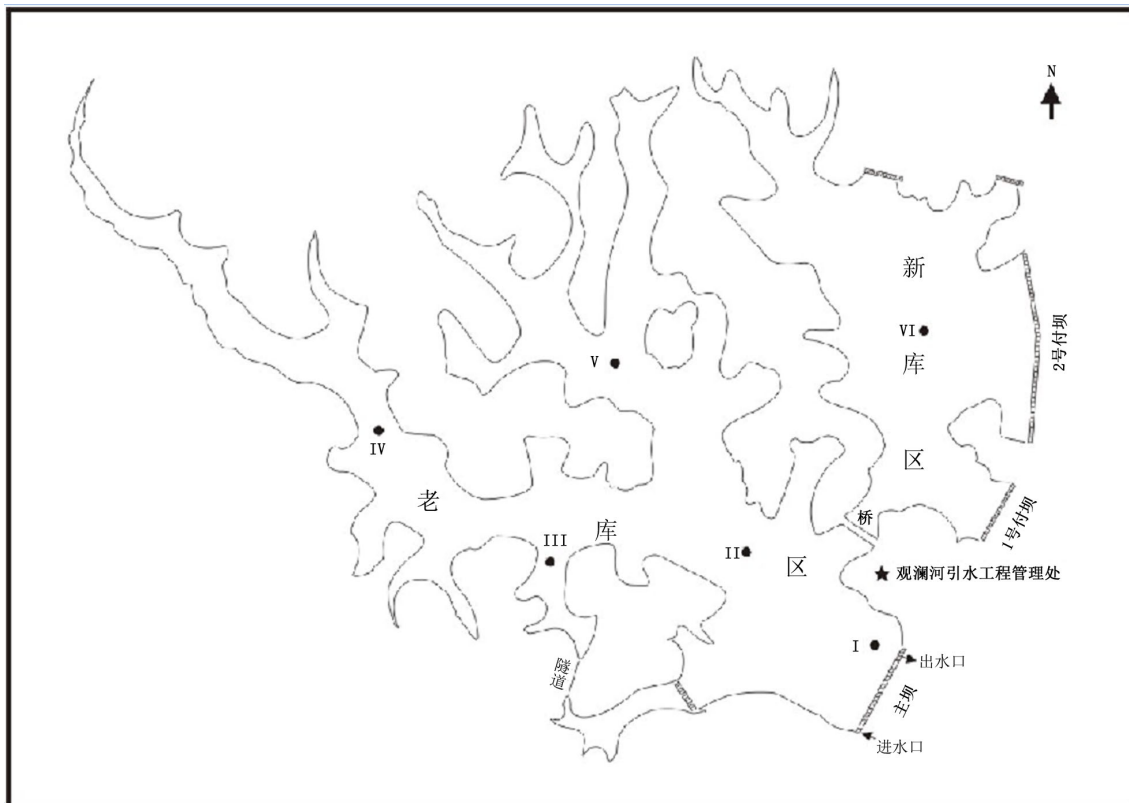


Figure 1. Location of monitoring sites
图 1. 监测点位置布设图

(3) 对 11 项理化因子进行标准化处理, 以其标准化值为自变量, 以藻类总生物量和蓝藻生物量的标准化值为因变量, 采用 SPSS 分析, 得到两项生物因子(藻类总生物量和蓝藻生物量)标准化值的逐步回归预测方程和复相关系数等。

3. 结果和讨论

3.1. 环境因子对藻类生物量的影响

本研究中涉及的环境因子均对藻类的生长有着间接或直接的影响作用。例如, 水温对藻类生长有限制作用, 氮、磷营养盐通常是藻类生长的限制性因子, 透明度和悬浮质的变化会影响水下日光辐射的变化进而影响藻类的生长[3], 同时, 透明度又会因为藻类的生长而受到影响。通过对各个因子的简单相关分析, 可以看出水库理化因子和藻类生物量的线性相关关系(见表 1)。

从上表中可知, 蓝藻为水库中的主要的藻类, 其与藻类生物量的相关性最大, 达 0.83。而在理化因子中, 藻类生物量与 COD 的相关性较大, 说明在水库中 COD 对藻类的生长有较大影响。

在本研究中, 对 DO 的分析分别设定了表层 DO、4 mDO、4 mDO 和 DO 均值这四个指标。其中, 4 mDO 与 DO 均值相关性最大, 表明了在水库中, 4 mDO 的含量最大。而 4 mDO 与藻类生物量的相关因子达 0.57, 说明 4 mDO 是又一影响藻类生长的环境因子。

营养盐的影响: 通过大量的研究, 人们一致认为, 湖泊中的氮、磷等营养物是藻类生长的主要限制性因素。在本研究中, TN 和 TP 和藻类生物量的影响因子分别为 0.41 和 0.38, 影响较为密切。

3.2. 叶绿素 a 与理化因子的逐步回归分析

在不同的水体中, 环境理化因子对藻类生长的影响程度各不相同。通过多元逐步回归分析, 可以筛选出相对重要的影响因子。藻类叶绿素 a 与理化因子的逐步回归统计结果如表 2 所示。

由表 2 可以看出, 各个监测点筛选出的重要因子各不相同。

Table 1. Relationship of algae biomass and related environmental factors

表 1. 水库一些理化因子和藻类生物量的线性相关关系

	藻类生物量	水深	透明度	表温	表层 DO	4 mDO	8 mDO	pH	NH ₃ -N	T-N	T-P	COD _{Mn}	BLUE	green	diatom
藻类生物量	1														
水深	0.51	1													
透明度	-0.05	0.52	1												
表温	0.34	0.36	-0.39	1											
表层 DO	0.27	0.15	0.06	0.1	1										
4 mDO	-0.57	-0.53	0.2	-0.81	-0.12	1									
8 mDO	-0.36	-0.41	0.31	-0.86	0.05	0.91	1								
Ph	0.41	0.14	-0.12	0.32	0.42	-0.18	-0.14	1							
NH ₃ -N	0.2	0.26	-0.44	0.74	0.3	-0.59	-0.6	0.35	1						
T-N	0.41	0.46	-0.25	0.67	0.22	-0.69	-0.65	0.29	0.88	1					
T-P	0.38	0.35	-0.32	0.42	0.12	-0.5	-0.46	0.1	0.64	0.65	1				
COD _{Mn}	0.80	0.49	-0.24	0.71	0.09	-0.78	-0.66	0.4	0.59	0.72	0.55	1			
蓝藻	0.83	0.43	0.17	0.26	0.01	-0.46	-0.33	0.38	-0.14	0.14	0.07	0.63	1		
绿藻	0.48	-0.14	-0.34	0.06	0.24	-0.01	0.07	0.21	0.15	0.05	0.34	0.36	0.26	1	
硅藻	0.58	0.57	-0.13	0.32	0.36	-0.51	-0.32	0.2	0.56	0.64	0.51	0.58	0.12	0.06	1

Table 2. Statistical regression results of Chl-a and related environmental factors

表 2. 叶绿素 a 与理化因子的逐步回归统计

监测点	入选变量	逐步回归方程	复相关系数	综合 F 值
1	COD _{Mn}	$Chla = -0.014 + 0.012 \times COD - 0.002 \times \text{引水量} - 0.007 \times NH3-N + 0.005 \times TN$	0.938	29.225
	引水量			
	NH3-N			
	T-N			
2	COD _{Mn}	$Chla = -0.04 + 0.15 \times COD + 0.004 \times DO \text{ 均值}$	0.934	61.021
	DO 均值			
3	表温	$Chla = 0.077 - 0.002 \times \text{表温}$	0.54	5.751
4	COD _{Mn}	$Chla = -0.007 + 0.15 \times COD - 0.042 \times \text{透明度}$	0.966	63.641
	透明度			
5	COD _{Mn}	$Chla = -0.01 + 0.005 \times COD + 0.004 \times \text{表层 DO} + 0.003 \times \text{水深} - 0.34 \times \text{透明度}$	0.927	24.545
	表层 DO			
	水深			
	透明度			

P1 点共有 4 个因子入选，其中 COD 和 TN 为正相关，引水量和 NH3-N 为负相关。

P2 点有 COD 和 DO 均值这两个因子入选，且均为正相关。

P3 最为特殊，只有表温这个因子影响较为显著。其复相关系数和综合 F 值均最低，分别为 0.54 和 5.751。

P4 有 COD 和透明度这两个因子入选。其复相关系数和综合 F 值均最高，分别为 0.966 和 63.641。

而 P5 有 COD、表层 DO、水深和透明度这四个因子入选，其中透明度为负相关，其他因子为正相关。

P3 点为引水入口附近监测点，受引水影响较大，P1、P2、P4 和 P5 均有 COD 这个因子入选，而且均为正相关。表明，各个理化因子中，COD 对叶绿素 a 的影响最为显著。

3.3. 藻类生物量和蓝藻生物量的逐步回归分析

上述分析结果是假设所有理化因子与叶绿素 a 之间是单向的线性相关，忽略了藻类生长或者大规模爆发，也会影响水体理化因子，如：藻类等光合作用能产氧气，影响水体溶解氧的变化，藻类生长过程中产生的化学物质，也会影响水体中 PH 的变化，这些环境因子可以说是以藻类为“因”而形成的“果”因素，难以作为自变量用于预测藻类生长[3] [4]。

因此，要预测藻类生物量尤其是蓝藻生物量的变化，需要剔除溶解氧、pH、SD 等因素，用藻类总生物量和蓝藻生物量分别作为因变量，计算逐步回归方程(见表 3)。

由表 3 可见，监测点 4 无法建立回归方程，监测点 1, 2, 3, 5 的回归方程分别与水深、TP，表温有关。各个方程的复相关系数相差不大，表明，水深、温度和营养物质对藻类生物量的影响相当。

对蓝藻生物量与环境因子进行回归分析，发现，只有监测点 P2 能建立回归方程，入选变量为 TP，回归方程为： $BLUE = -0.016 + 6.187 \times TP$ ，复相关系数为 0.482，综合 F 值为 6.353。

3.4. 标准化回归分析

以监测点 P2 为例，对藻类生物量和各环境因子进行标准化处理，标准化后的逐步回归方程以 TP 为显著因子，回归方程为：

Table 3. Statistical regression results of algae biomass and related environmental factors**表 3.** 藻类生物量与相关环境因子的逐步回归统计结果

点号	入选变量	逐步回归方程	复相关系数	综合 F 值
1	水深	$PHY = -0.051 + 0.170 \times \text{水深}$	0.571	10.141
2	TP	$PHY = 0.987 + 7.814 \times TP$	0.586	10.998
3	表温	$PHY = 5.239 - 0.145 \times \text{表温}$	0.501	6.381
5	水深	$PHY = 0.524 + 0.233 \times \text{水深}$	0.469	5.911

$$\ln(PHY + 1) = 0.659 + 2.518 \times \ln(TP + 1)$$

复相关系数为 0.566, 与标准化前相近。

同样以监测点 P2 为例, 对蓝藻生物量和各环境因子进行标准化处理, 标准化后的逐步回归方程以水深为显著因子, 回归方程为:

$$\ln(PHY + 1) = -1.078 + 0.616 \times \ln(\text{Depth} + 1)$$

其中, 复相关系数为 0.469, 显著性较差。与未标准化回归分析相比, 未标准化回归方程显著因子为 TP, 更符合实际情况。

4. 结论

对藻类生物量与环境理化因子的相关性回归分析, 结果表明, COD 的影响最为显著, 但化学需氧量受藻类生长的影响而变化, 目前无证据表明 COD 影响藻类生长, 因而难以用于预测藻类生长情况。剔除 COD, DO 等次要环境因子外, 藻类生物量主要受水深, 水温及 TP 的影响。但不同监测点的显著因子不同。表明, 藻类生长因环境不同, 限制因子也不同。为了建立蓝藻预测模型, 对各个因子及藻类生物量进行标准化处理, 标准化后的回归方程相对未标准化方程, 相关性并无显著改善。

综上所述, 逐步回归统计模型, 能够寻找影响藻类生物量的关键性因子, 为准确预测提供参考, 但尚不能对藻类爆发性生长做出精确的预测, 要对该水库蓝藻水华进行较准确的模拟预测, 还需在分析显著因子的基础上, 结合藻类生长机理, 建立预测模型。

基金项目

国家基础科学人才培养基金《武汉大学地理科学理科基地》科研能力训练项目(J1103409); 中央高校基本科研业务费专项资金武汉大学自主科研项目(121066)。

参考文献 (References)

- [1] ANGELINI, R., PETRERE JR., M. A model for the plankton system of the Broa reservoir, Sao Carlos, Brazil. *Ecological Modelling*, 2000, 126(2-3): 131-137. [http://dx.doi.org/10.1016/S0304-3800\(00\)00261-1](http://dx.doi.org/10.1016/S0304-3800(00)00261-1)
- [2] MENSHTUKIN, V. V., ASTRAKHANTSEV, G. P., YEGOROVA, N. B., RUKHOVETS, L. A., SIMO, T. L. and PETROVAC, N. A. Mathematical modeling of the evolution and current conditions of the Ladoga Lake ecosystem. *Ecological Modelling*, 1998, 107(1): 1-24. [http://dx.doi.org/10.1016/S0304-3800\(97\)00184-1](http://dx.doi.org/10.1016/S0304-3800(97)00184-1)
- [3] 陈宇炜, 秦伯强, 高锡云. 太湖梅梁湾藻类及相关环境因子逐步回归统计和蓝藻水华的初步预测[J]. *湖泊科学*, 2001, 13(1): 63-71.
CHEN Yuwei, QIN Boqiang and GAO Xiyun. Prediction of blue-green algae bloom using stepwise multiple regression between algae & related environmental factors in Meiliang Bay, Lake Taihu. *Journal of Lake Science*, 2001, 13(1): 63-71. <http://dx.doi.org/10.18307/20010110>
- [4] 王兆群, 张宁红, 张咏. 洪泽湖藻类与环境因子逐步回归统计和蓝藻水华初步预测[J]. *中国环境监测*, 2012, 28(4): 17-20.
WANG Zhaoqun, ZHANG Ninghong and ZHANG Yong. Prediction of blue-green algae bloom using stepwise multiple regression between algae and related environmental factors in Hongze Lake. *Environmental Monitoring in China*, 2012, 28(4): 17-20.

- [5] 张雅燕, 吴志旭, 朱淑君. 千岛湖藻类及相关环境因子多元线性回归和鱼腥藻预测模型的建立[J]. 中国环境监测, 2002, 18(3): 37-41.
ZHANG Yayan, WU Zhixu and ZHU Xujun. Model prediction of anabaena bloom using backward multiple regression between algae & related environmental factors in Thousand-island Lake. Environmental Monitoring in China, 2002, 18(3): 37-41.
- [6] 张大玲. 淀山湖富营养化分析与水质预测预警[D]: [硕士学位论文]. 上海: 东华大学, 2008.
Zhang Daling. The analysis on eutrophication and prediction advance warning on water quality in Dianshan Lake. Shanghai: Donghua University, 2008.
- [7] 姜霞, 王书航, 钟立香, 金相灿, 孙世群. 巢湖藻类生物量季节性变化特征[J]. 环境科学, 2010, 31(9): 2051-2062.
JIANG Xia, WANG Shuhang, ZHONG Lixiang, JIN Xiangcan and SUN Shiqun. Seasonal variation characteristics of algae biomass in Chaohu Lake. Environmental Science, 2010, 31(9): 2051-2062.
- [8] 韩新芹, 叶麟, 徐耀阳, 蔡庆华. 香溪河库湾春季叶绿素 a 浓度动态及其影响因子分析[J]. 水生生物学报, 2006, 30(1): 89-94.
HAN Xinqin, YE Lin, XU Yaoyang and CAI Qinghua. Analysis of factors impacted on algal growth during period of spring algal bloom in Xiangxi River Bay. Journal of China Three Gorge University (Natural Science), 2006, 30(1): 89-94.
- [9] 郑利, 谢平, 林匡飞, 王松波, 杨华, 杨洪. 武汉莲花湖微囊藻毒素含量变化特征及其影响因素的研究[J]. 农业环境科学学报, 2004, 23(6): 1053-1057.
ZHENG Li, XIE Ping, LIN Kuangfei, WANG Songbo, YANG Hua and YANG Hong. Variation characteristics of microcystin concentration and influencing factors in Lake Lianhuahu (Wuhan). Journal of Agro-Environment Science, 2004, 23(6): 1053-1057.
- [10] 刘晓慧. 西流湖水体营养状况评价及藻类环境因子多元分析[D]: [硕士学位论文]. 郑州: 郑州大学, 2005.
Liu Xiaohui. Assessment of water eutrophication and multiple analysis of algae and environmental factors. Zhengzhou: Zhengzhou University, 2005.
- [11] 梁俐, 邓云, 郑美芳, 魏希. 基于 CE-QUAL-W2 模型的龙川江支库富营养化预测[J]. 长江流域资源与环境, 2014, 23(Z1): 103-111.
LIANG Li, DENG Yun, ZHENG Meifang and WEI Xi. Prediction of eutrophication in the Longchuan River based on CE-QUAL-W2 model, resources and environmental and environment in the Yangtze Basin, 2014, 23(Z1): 103-111.
- [12] ZHOU X. J., WEI, X., LUO, G. Y., ZHOU, X. B. and Li, X. Y. Analysis of eutrophication status and chla prediction models for Da'ning River. Agricultural Science & Technology, 2012, 13(3): 608-612, 616.
- [13] 卢崇飞, 高惠璇, 叶文虎. 环境数理统计学应用及程序[M]. 北京: 高等教育出版社, 1988.
LU Chongfei, GAO Huixuan and YE Wenhui. Application and program of environmental mathematical statistics. Beijing: Higher Education Press, 1988.
- [14] ANDERSON, C. R., SAPIANO, M. R. P., PRASAD, M. B. K., *et al.* Predicting potentially toxigenic *Pseudo-nitzschia* blooms in the Chesapeake Bay. Journal of Marine Systems, 2010, 83(3): 127-140. <http://dx.doi.org/10.1016/j.jmarsys.2010.04.003>