Ecological Modelling 221 (2010) 681–688

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Exploring the influence of lake water chemistry on chlorophyll *a*: A multivariate statistical model analysis

Yong Liu[a,b], Huaicheng Guo[a,*], Pingjian Yang[a]

[a] College of Environmental Science and Engineering, Peking University, Beijing 100871, China
[b] School of Natural Resources and Environment, University of Michigan, MI 48109, USA

ARTICLE INFO

Article history:
Available online 5 April 2009

Keywords:
Absolute principal component score (APCS)
Multivariate linear regression (MLR)
Structural equation modeling (SEM)
Chlorophyll *a*
Lake Qilu

ABSTRACT

A multivariate statistical approach integrating the absolute principal components score (APCS) and multivariate linear regression (APCS-MLR), along with structural equation modeling (SEM), was used to model the influence of water chemistry variables on chlorophyll *a* (Chl *a*) in Lake Qilu, a severely polluted lake in southwestern China. Water quality was surveyed monthly from 2000 to 2005. APCS-MLR was used to identify key water chemistry variables, mine data for SEM, and predict Chl *a*. Seven principal components (PCs) were determined as eigenvalues >1, which explained 68.67% of the original variance. Four PCs were selected to predict Chl *a* using APCS-MLR. The results showed a good fit between the observed data and modeled values, with $R^2 = 0.80$. For SEM, Chl *a* and eight variables were used: NH_4-N (ammonia-nitrogen), total phosphorus (TP), Secchi disc depth (SD), cyanide (CN), arsenic (As), cadmium (Cd), fluoride (F), and temperature (T). A conceptual model was established to describe the relationships among the water chemistry variables and Chl *a*. Four latent variables were also introduced: physical factors, nutrients, toxic substances, and phytoplankton. In general, the SEM demonstrated good agreement between the sample covariance matrix of observed variables and the model-implied covariance matrix. Among the water chemistry factors, T and TP had the greatest positive influence on Chl *a*, whereas SD had the largest negative influence. These results will help researchers and decision-makers to better understand the influence of water chemistry on phytoplankton and to manage eutrophication adaptively in Lake Qilu.

1. Introduction

The awareness of lake eutrophication and aquatic ecological degradation worldwide has grown over the last decades. Primary production, especially that of phytoplankton, is used as a sensitive and accurate indicator for eutrophication assessment (Scheffer, 1998). Phytoplankton in lakes is affected by both external and internal factors, as well as the complex embedded interactions among these factors, such as watershed nutrient loading and lake nutrient concentration, light, temperature (T), water chemistry, and grazing pressure (Çamdevýren et al., 2005). Thus, decisions about managing eutrophication require a clear understanding of these interactions.

Models have proved to be useful tools in supporting management decisions about lake eutrophication (Charpa, 1997; Zou et al., 2007; Liu et al., 2008). Since the 1970s, numerous process-based and statistical models have focused on analyzing the physicochemical and ecological processes of lakes and their effects on phytoplankton (Charpa, 1997; Jørgensen and Bendoricchio, 2001). Chlorophyll *a* (Chl *a*) is widely used as a fundamental proxy for phytoplankton abundance (Çamdevýren et al., 2005). The majority of the models directly simulate interactions between phytoplankton and inorganic/organic nutrients or physiological factors, such as nitrogen (N), phosphorus (P), suspended solids, T, oxygen, and Secchi disc depth (SD) (Jørgensen and Bendoricchio, 2001; Håkanson and Boulion, 2002; Zhang et al., 2004; Elliott et al., 2006). However, few models directly factor in other water chemistry variables, such as heavy metals, despite their importance in stimulating (or limiting) phytoplankton growth (Jørgensen and Bendoricchio, 2001). This may be due to the relative lack of understanding regarding the mechanism between these water chemistry variables and phytoplankton productivity and the consequent difficulty in establishing quantitative modeling formulations. Two key issues must be addressed prior to investigating the influence of water chemistry on phytoplankton productivity: identification of the driving water chemistry variables, and quantitative expressions of their influence and relationships. Recent advances in multivariate statistical applications enable these issues to be addressed. For example, principal components analysis (PCA) can help identify the variables (Çamdevýren et al., 2005).

PCA is used to explain the variance of a large set of intercorrelated variables, which are transformed into a smaller set of independent principal components (PCs) (Singh et al., 2005). PCs

* Corresponding author. Tel.: +86 10 62751921; fax: +86 10 62751921.
E-mail addresses: lyong@umich.edu (Y. Liu), hcguo@pku.edu.cn (H. Guo).

0304-3800/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2009.03.010

provide the most meaningful information about the targeted data set. PCA has been widely applied to pattern recognition to reduce the dimensions of multivariate problems with a minimum loss of original information (Perkins and Underwood, 2000). Recently, absolute principal components score (APCS) analysis and multivariate linear regression (MLR) have been integrated to create a popular receptor model (APCS-MLR) that has been used in water-quality assessment, source apportionment, and Chl *a* prediction (Çamdevýren et al., 2005; Singh et al., 2005; Zhou et al., 2007). We applied APCS-MLR in this study to identify the driving water chemistry variables, reduce the number of variables for further analysis, and predict lake Chl *a* concentration. However, despite the ability of PCA to reduce variables to higher order PCs, it has little flexibility to specify model structure, such as interactions among variables (Arhonditsis et al., 2006). Hence, we used structural equation modeling (SEM), another multivariate statistical method, based on the APCS-MLR results to quantitatively describe the influence of the key water chemistry variables on Chl *a*.

A powerful method to explore the relationship among correlated variables, SEM incorporates ideas from regression, path analysis, and factor analysis. Wright (1934) first introduced the method in biological population research, and it has been expanded to a range of research areas, including social sciences, psychology, chemistry, and biology (Kline, 1998; Hung et al., 2007; Kashy et al., 2008). In SEM, latent (unobserved) variables are involved in the analysis, and multiple and overlapping regressions can be statistically tested. SEM provides an effective way to develop and test a hypothesized underlying mechanism among variables (La Peyre et al., 2001). In contrast to traditional multivariate regression, such as MLR, SEM has several advantages: indirect measurement of latent variables, which can reduce the dimensionality of data and also help to analyze the direct and indirect effects of variables; incorporation of uncertainties in measurement errors or lack of validity of the observed variables; and provision of a simultaneous solution to a set of multiple regression relationships, thus helping to test the hypothesis (Malaeb et al., 2000; La Peyre et al., 2001; Arhonditsis et al., 2006; Hung et al., 2007). Despite its wide use in the social sciences, SEM applications in the environmental sciences and ecology are relatively limited, but increasing (Buncher et al., 1991; Liu et al., 1997; Arhonditsis et al., 2006; Bayard and Jolly, 2007). For example, Malaeb et al. (2000) used SEM to examine the hypothesized interdependencies of four latent variables in estuaries: sediment contamination, natural variability, biodiversity, and growth potential. Eleven environmental variables were involved in SEM to test the conceptual model that explored the pattern of effects of the first three latent variables on growth potential by recreating the matrix of observed correlations. The data fit well, accounting for 81% of biodiversity and 69% of growth potential variability. In a Lake Washington and Neuse River estuary case study, Reckhow et al. (2005) applied SEM to relate the water pollutant concentration of each criterion to the probability of compliance with the designated use. Their SEM results could provide decision-makers with a practical assessment of the risks of violating various water-quality criteria. Arhonditsis et al. (2006) used SEM to explore the ecological patterns of two lakes, eutrophic Lake Mendota and mesotrophic Lake Washington, in terms of how abiotic conditions and biological interactions affected phytoplankton dynamics during summer stratification.

As APCS-MLR reduces the number of variables, and SEM provides a quantitative description of their relationships, the integration of APCS-MLR and SEM should be effective at exploring the influence of water chemistry on Chl *a* in lakes. In this study, we developed an integrated multivariate statistical approach using APCS-MLR and SEM to model the relationship between water chemistry and Chl *a*, including physical factors, nutrients, and toxic substances. Our case study was Lake Qilu, a heavily polluted and eutrophicated lake in southwestern China. The model results will help researchers and local decision-makers to understand the mechanism of Chl *a* and control eutrophication in an effective manner.

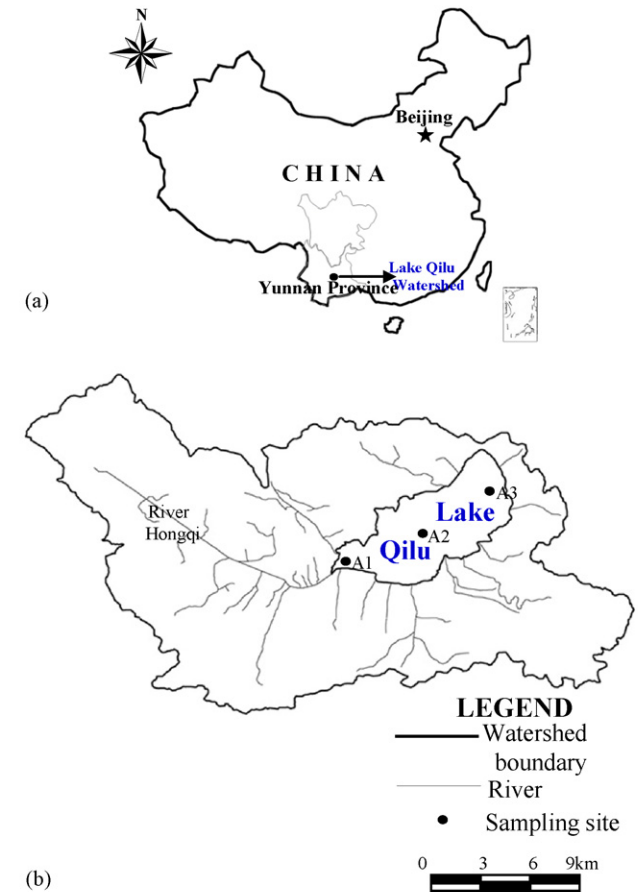

Fig. 1. Lake Qilu and the watershed.

2. Materials and methods

2.1. Study area and data sources

Lake Qilu is an altiplano lake located on the Yunnan-Guizhou Plateau (altitude: 1797.65 m) of southwestern China, in Yuxi City, Yunnan Province. The lake surface area is about 37.26 km^2, and its volume is approximately 1.68×10^8 m^3 (Fig. 1). The watershed, at latitude 24°4′ to 24°14′N and longitude 102°33′ to 102°52′E, covers about 354.2 km^2 (Liu et al., 2007a). The maximum depth of the shallow lake is 6.8 m, and the average depth is 4.0 m. The long water residence time of about 560 days renders the lake more vulnerable to pollution. Lake Qilu is severely eutrophicated due to nutrient loading from industrial pollution, intensive agricultural activities, and recent rapid urbanization, according to the China Ministry of Environmental Protection (MEP, 2007). Lake water quality has not met the targeted goal of 8 μg l^{-1} Chl *a* set by the local environmental protection bureau (EPB). Wastewater from industrial sources and urbanized areas, together with agricultural nonpoint contaminants, are the main pollution sources. As one of the four most severely polluted lakes in Yunnan Province, Lake Qilu is under government mandate to develop an effective water-quality management plan and reduce major eutrophication by 2015. Thus, the mechanisms by which water chemistry influences phytoplankton, measured as Chl *a*, must be investigated to aid local decision-makers in reducing pollutants and managing eutrophication.