Science of the Total Environment 648 (2019) 472–480

Contents lists available at ScienceDirect

Science of the Total Environment

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

Chlorophyll *a* predictability and relative importance of factors governing lake phytoplankton at different timescales

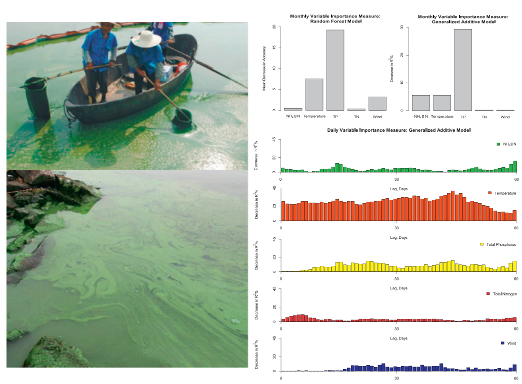
Xia Liu, Jianfeng Feng *, Yuqiu Wang *

Key Laboratory of Pollution Processes and Environmental Criteria, Ministry of Education, and Tianjin Key Laboratory of Environmental Remediation and Pollution Control, College of Environmental Science and Engineering, Nankai University, Tianjin 300071, China

HIGHLIGHTS

- Water temperature is more important predictor of daily chlorophyll *a* than nutrient.
- Nutrients are a more important predictor than water temperature at a monthly scale.
- The drivers of phytoplankton fluctuations vary at different timescales.
- Timescales have an influence on the relative role of N and P limitation in lakes.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 13 June 2018
Received in revised form 10 August 2018
Accepted 10 August 2018
Available online 11 August 2018

Editor: Sergi Sabater

Keywords:
Phytoplankton
Predictability
Random forest
Generalized additive model
Timescale

ABSTRACT

Assessing the key drivers of eutrophication in lakes and reservoirs has long been a challenge, and many studies have developed empirical models for predicting the relative importance of these drivers. However, the relative roles of various parameters might differ not only spatially (between regions or localities) but also at a temporal scale. In this study, the relative roles of total phosphorus, total nitrogen, ammonia, wind speed and water temperature were selected as potential drivers of phytoplankton biomass by using chlorophyll *a* as a proxy for biomass. A generalized additive model (GAM) and a random forest model (RF) were developed to assess the predictability of chlorophyll *a* and the relative importance of various predictors driving algal blooms at different timescales in a freshwater lake. The results showed that the daily datasets yielded better predictability than the monthly datasets. In addition, at a daily scale, water temperature was a more important predictor of chlorophyll *a* than nutrients, and the importance of phosphorus was comparable to that of nitrogen. In contrast, at a monthly scale, nutrients are more important predictors than water temperature and phosphorus is a better predictor than nitrogen. This study indicates that the drivers of phytoplankton fluctuations vary at different timescales and that timescale has an influence on the relative roles of nitrogen and phosphorus limitation in lakes, which suggests that the temporal scale should be considered when explaining phytoplankton fluctuations. Moreover, this study provides a reference for the monitoring of phytoplankton fluctuations and for understanding the mechanisms underlying phytoplankton fluctuations at different timescales.

* Corresponding authors.
E-mail addresses: fengjf@nankai.edu.cn (J. Feng), yqwang@nankai.edu.cn (Y. Wang).

https://doi.org/10.1016/j.scitotenv.2018.08.146
0048-9697/

1. Introduction

Algal blooms are a major problem in freshwater ecosystems throughout the world (Paerl and Huisman, 2008; Page et al., 2018). Predictive models can be useful for developing strategies to reduce bloom frequency and severity and for guiding actions to reduce bloom impacts (Cha et al., 2014). Identifying algal bloom drivers is essential for developing predictive models. Over the past five decades, studies have attempted to identify the drivers of undesirable or harmful algal blooms (HABs) (Cha et al., 2017). Although an excess of nutrients is always a key determinant of blooms, the relative roles of phosphorus (P) and nitrogen (N) in bloom dynamics remains hotly debated (Carpenter et al., 2016; Conley et al., 2009; Elser et al., 2009; Schindler, 1977; Schindler, 2014). Moreover, the impacts of temperature on blooms has received much attention in the past decade (Trolle et al., 2015), and these impacts will increase at an alarmingly fast rate in the future (O'Reilly et al., 2015). Some studies have found that global warming plays an important role in the global spread of phytoplankton blooms (Ma et al., 2016). The formation of HABs has been attributed, in part, to extra potential meteorological conditions (Scavia et al., 2016). For example, climate change will affect not only thermal regimes in lakes but also precipitation and thus runoff. Catchment-related fluxes in nutrients and water residence time might also have strong impacts on algal blooms and lake productivity, and changes in wind speed and thus mixing and turbulence will also contribute to these impacts. Thus, nutrients, temperature and meteorological conditions are likely to play significant roles in driving algal blooms (Rigosi et al., 2014; Scavia et al., 2016).

The relative importance of environmental drivers might be site-specific (Rigosi et al., 2014). For example, the results of a linear regression model tested on >1000 U.S. lakes indicated that nutrient levels are more important predictors of algal blooms than temperature (Rigosi et al., 2014). In contrast, a study employing a Bayesian network model suggested that phytoplankton fluctuations are more sensitive to changes in water temperature than to changes in the concentration of total P in 20 globally distributed lakes (Rigosi et al., 2015). The relative roles of nutrients, temperature, mixing, hydrology, solar radiation and other drivers depend on both spatial and temporal resolutions (Blauw et al., 2018). At an inter-lake level or for intra-lake development over years, the nutrient levels are expected to override other drivers, whereas at the intra-lake level, both temperature and wind speed (and thus turbulence and nutrient mixing) could be important. Most studies on phytoplankton dynamics in both freshwater and marine sites have addressed drivers of phytoplankton responses at monthly or inter-annual scales.

The drivers of phytoplankton fluctuations can also differ among different temporal scales (Blauw et al., 2018). In marine waters, phytoplankton fluctuations at inter-annual and decadal scales are often impacted by climatic variation or changes in the eutrophication status (McQuatters-Gollop and Vermaat, 2011; Ottersen et al., 2001; Richardson and Schoeman, 2004), whereas at a seasonal scale, nutrients, temperature, solar irradiance, thermal stratification, and grazing are the major drivers of phytoplankton fluctuations and succession (Sharples et al., 2006; Sommer et al., 2012; Winder and Cloern, 2010). At even shorter timescales (e.g., monthly and daily), fluctuations in phytoplankton are largely affected by physical drivers, such as wind and turbulent mixing, which affect the spatial distribution of phytoplankton (Peter, 2005). For accurate predictions, it is important to separate the different impacts of various drivers at different temporal scales. Winder and Cloern (2010) conducted a comparative time series analysis of lakes and open oceans and found consistent differences in the relative importance of drivers between seasonal and inter-annual scales. Most studies that have investigated such temporal responses were based on marine environments (Cloern and Jassby, 2010). However, the roles of drivers at shorter timescales, i.e., monthly and daily scales, as determinants of algal biomass in freshwater lakes remain poorly understood.

The determination of a proper model structure is critical for developing a predictive model. The core of any forecasting model that used by management authorities or the community is to predict events at a relevant timescale. Many studies have focused on potential predictors of HABs without considering time-lag effects (Hollister et al., 2016; Segura et al., 2017), which results in the need for extra predictions for predictors. For example, explanatory variables (predictors) must be determined before the developed predictive model can be used to predict phytoplankton responses. Predictive models with a time lag between the drivers and responses can address this issue well because responses can be predicted without determining the predictors. There will always be some time lag between drivers and responses, and this time lag is related to either physical factors (e.g., turbulence and the spatial distribution of algal masses) or biological factors related to growth rate and population responses. Several studies have attempted to explore the development of forecasting models based on the time-lag effect of predictors (Kehoe et al., 2015; Xiao et al., 2017; Zhang et al., 2015b). For example, Zhang et al. (2015b) used an artificial neural network model to predict the water quality of the Yuqiao Reservoir (YQR) and found that this model is potentially useful for predicting eutrophication up to 2 weeks in advance.

The aim of the current study was to evaluate the predictability of phytoplankton fluctuations and the relative importance of selected key drivers among different timescales in a freshwater lake with five potential predictors (total P, total N, ammonia N, water temperature and wind speed) as drivers and chlorophyll *a* concentrations to represent the biomass of phytoplankton. Using these five predictors, we (1) developed a generalized additive model (GAM) and a random forest (RF) model based on two datasets (monthly and daily) to assess the predictability of phytoplankton fluctuations at these two timescales, (2) assessed their importance using the GAM and RF and (3) compared the differences of the relative importance at monthly and daily scales. Although these studies were performed in a single lake, we believe the insights are relevant to other comparable freshwater bodies.

2. Materials and methods

2.1. Study area and data description

The YQR (117°34′ E and 40°02′ N) is located in the north of Tianjin City in China (Fig. 1) and serves as the largest drinking water source in Tianjin (with a population of >16 million). Two rivers (Guohe River and Linhe River) enter into the YQR. The YQR receives water from an upstream reservoir (Daheiting Reservoir) via the Guohe River, and the Linhe River is a minor contributor and often runs dry. The reservoir has a watershed area of 2060 km^2, a storage capacity of 1.559 billion m^3 and a surface area of 86.8 km^2. The YQR is considered a shallow lake with a maximum depth of 12 m and an average depth of 4.7 m. The annual precipitation in the YQR Basin over the study period was approximately 750 mm/m^2. The detailed properties of the YQR are shown in Table 1.

Because the YQR is the drinking water source in Tianjin, a project aiming to protect the watershed was implemented in 2002 to reduce or eliminate the point sources of nutrients to YQR (Zhang et al., 2015a). However, the reservoir has been mesotrophic due to increased pollution in the watershed, notably from cage aquaculture in the upstream reservoir. The water quality of YQR has gradually decreased. In fact, the YQR has experienced several HABs over the past two decades.

We used the concentration of chlorophyll *a* (Chl-a, μg/L) as a proxy for phytoplankton biomass. There certainly are variations in Chl-a to carbon ratios (Jakobsen and Markager, 2016), but one should be aware that also cell specific C may vary with species and growth conditions. When comparing direct, volume-based estimates on phytoplankton with Chl-a, there is no doubt that the Chl-a's response to light is biomass-specific, but this variation does not override the positive association between Chl-a and biomass. For example, a very strong