Limnologica 60 (2016) 39–50

Contents lists available at ScienceDirect

Limnologica

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

How many seasons are there in a sub-tropical lake? A multivariate statistical approach to determine seasonality and its application to phytoplankton dynamics

Akeapot Srifa [a], Edward J. Phlips [b,*], John Hendrickson [c]

[a] Department of Biology, Faculty of Science, Mahasarakham University, Kantarawichai District, Mahasarakham, 44150, Thailand
[b] Program of Fisheries and Aquatic Sciences, School of Forest Resources and Conservation (SFRC), University of Florida, 7922 NW71st St. Gainesville, FL, 32653, USA
[c] St. Johns River Water Management District, 4049 Reid St, Palatka, FL 32177, USA

ARTICLE INFO

Article history:
Received 7 December 2015
Received in revised form 25 May 2016
Accepted 30 May 2016
Available online 29 June 2016

Keyword:
Cluster analysis
Principal component analysis
Seasonality
Phytoplankton
Cyanobacteria
Subtropical lake

ABSTRACT

Seasonal patterns in factors that affect primary producers are an important part of defining the structure and function of aquatic ecosystems. However, defining seasonality is often more difficult in aquatic than in terrestrial ecosystems, particularly in subtropical and tropical environments. In this study, a long-term data set for a shallow subtropical lake (Lake George, Florida, USA) was used to investigate seasonality using a range of physical, chemical and hydrological parameters. K-means cluster analysis of monthly averages among 11 selected environmental factors across 18 years suggested the overall annual pattern consists of three different seasonal clusters: a cold season (January–April), a warm season (May–August) and a flushing season (September–December). High dissolved oxygen and increased Secchi depth are key features of the cold season, while the warm season is characterized by high mean light irradiances, temperature, rainfalls, total nitrogen and phytoplankton biomass (as chlorophyll *a* level). The flushing season is characterized by high river discharge rates and high levels of dissolved nutrients and colored organic matter. Multiple response permutation procedures indicated that these seasonal cluster arrangements were significantly different than randomly permuted clusters (*A*-statistics = 0.3314, significance of delta = 0.0160, based on 1000 permutations). Results from principal component analyses supported the presence of the three seasons in the lake. Linear models explaining chlorophyll *a* levels using the 3-season system generally indicated better ratios of explained variance compared to the models without seasonal alignments, further indicating that even for sub-tropical systems defining seasons provides a better understanding of phytoplankton dynamics. The approaches used in this study provide statistically-based multivariate tools for the definition of seasonality in aquatic ecosystems. The ability to accurately define seasons is a key step in modeling the structure and dynamics of aquatic ecosystem, which is essential to the development of effective management strategies in a rapidly changing world.

1. Introduction

Over the past century, the integrity of an increasing number of aquatic ecosystems around the world has been challenged by human activities, such as cultural eutrophication and altered hydrologic regimes (Nixon, 1995; Cloern and Jassby, 2010). There is also a growing consensus that increases in greenhouse gases of human origin will lead to future shifts in climatic regimes (Edwards and Richardson, 2004; Winder and Schindler, 2004a; Winder and Schindler, 2004b; Thackeray et al., 2008). These environmental challenges have complicated the task of managing aquatic ecosystems. In order to effectively manage aquatic ecosystems, it is important to understand their structure, function and responses to changes in the environment, such as eutrophication, and shifts in climatic and hydrologic conditions. In the case of defining the responses of primary producer populations to environmental changes, these factors play particularly important roles (Smith et al., 1999; Falkowski and Oliver, 2007; Paerl and Huisman, 2008; Kratina et al., 2012) because they can impact a wide range of drivers for phytoplankton dynamics, such as nutrient availability, temperature, and hydrologic conditions (e.g. flushing rates, water residence times and vertical mixing). Statistical and modeling approaches are

* Corresponding author.
E-mail address: phlips@ufl.edu (E.J. Phlips).

http://dx.doi.org/10.1016/j.limno.2016.05.011
0075-9511/

often used to help define and even predict responses of primary producers to environmental changes, particularly as they relate to potentially ecosystem disruptive effects, such as harmful algae blooms (Braak and Verdonschot, 1995; Brink et al., 2009; Beck and Hagy, 2015; Rigosi et al., 2015).

One of the challenges in using statistical and modeling approaches to define responses of primary producers to environmental change is the ability to account for seasonal shifts in driving factors. Definitions of seasonality are often based on one or two variables, most commonly temperature and incident solar irradiance (Trenberth, 1983; Berger et al., 2010). In the case of terrestrial plant communities at higher latitudes, there is a strong link between seasonality of plant communities and temperature and photoperiod, thereby facilitating the categorization of seasons (Jolly et al., 2005). By contrast, planktonic communities often exhibit high species diversity, short generation times, and can respond quickly to a wide range of environmental conditions, thereby complicating the task of defining seasonality. This is demonstrated by the results of a recent study (Winder and Cloern, 2010) that showed a lack of clearly definable seasonal trends in phytoplankton biomass in over 30% of the time series examined across over 100 freshwater and marine ecosystems. It is widely recognized that more factors than incident irradiance and temperature can play important roles in controlling phytoplankton production, and should be included in defining seasonal categorization, such as hydrologic conditions, nutrient concentrations and the presence of light attenuating substances in the water column (Canfield et al., 1989; Agusti et al., 1990; Phlips et al., 2007; Srifa et al., 2016).

The central goal of this study was to define seasonality using multivariate statistical tools to incorporate a wide range of important physical, chemical and biological variables. Multivariate methods are recognized as powerful tools in the study of multidimensional observations in a wide range of disciplines including ecological sciences (Rencher, 2002; Everitt and Hothorn, 2011). Unlike contemporary univariate statistical analyses, many multivariate techniques do not require restrictive assumptions (Basille et al., 2013), and can be used to explore data without presumptive hypotheses (Everitt and Hothorn, 2011).

In this paper, we examine the use of two multivariate statistical approaches, i.e. K-means cluster analysis (KMCA) and principal component analysis (PCA), to define seasonality in a large subtropical lake, i.e. Lake George in Florida, USA. An eighteen-year data set for Lake George, including 17 key environmental parameters, provided an opportunity to incorporate a wide range of climatic regimes (e.g. drought and flood conditions) into the analyses. We hypothesized that the inclusion of a diverse range of parameters and conditional states into the definition of seasons would yield valuable insights into temporal variability of driving factors for phytoplankton dynamics.

2. Materials and methods

2.1. Site description

Lake George (Fig. 1) is located in Volusia County, Florida, United States. The lake is a part of the Lower St. Johns River that flows from south to north and empties into the Atlantic Ocean 173 km downstream. The lake is large (190 km^2 in surface area), shallow (mean depth of 2.8 m), subtropical (latitude 29°N) and located on a shallow-gradient basin, which results in relatively long water residence time and slow turnover rates (Phlips et al., 2007). The shallowness of the lake results in largely polymictic conditions (Brenner et al., 1990). The lake is eutrophic and regularly experiences blooms of cyanobacteria (Hendrickson et al., 2003; Phlips et al., 2007).

The primary sampling site in this study (i.e. LG) was located in the northern reach of Lake George where the lake discharges into the lower St. Johns River. The site location was chosen to represent a summation of processes within the lake before discharge downstream into lower St. Johns River Basin. The average depth at the sampling site was 2.1 m (N = 192).

2.2. Data set and water analyses

The long-term data set used in this study was obtained from the St. Johns River Water Management District (sjrwmd.com/hydrologicdata/waterquality) for the site in Lake George. The data set was comprised of water quality and physical data from September 1993–December 2010. Most samples were taken monthly, but there were periods in some winter months when water samples were taken every 2 months, and some summer months in 2009–2010 when samples were taken twice a month. In such cases mean values were calculated from over-flanking months or within a month, respectively, to balance the data set for spectral density analysis.

Water samples were taken with a 3.0 m vertical integrating sampler that evenly captured water from the surface to approximately 0.1 m above the sediment. Aliquots of water samples were analyzed for pheophytin-corrected chlorophyll *a* (Chl *a*), color, nitrate and nitrite-nitrogen (NO_x-N), ammonium nitrogen (NH_4-N), total nitrogen (TN), total phosphorus (TP), soluble reactive phosphorus (SRP), dissolved organic carbon (DOC) and silica. Dissolved inorganic nitrogen (DIN) was assumed to be the summation of NO_x-N and NH_4-N concentrations. Preparations and analyses methods for all water chemistry parameters were according to the Standard Method (American Public Health Association et al., 1992).

Water temperature (WTemp), dissolved oxygen (DO) and Secchi depth (SD) were measured on-site with a YSI or a Hydrolab multiprobe and a Secchi disk. Turbidity was determined by a LaMotte turbidimeter from water taken back to the lab.

2.3. Hydrological and meteorological data

Monthly rainfall totals for the study period were obtained from the National Climatic Data Center of the National Oceanic and Atmospheric Administration (NOAA) for the Crescent City, FL Meteorological Station (29.43333°N, 81.51667°W) (ncdc.noaa.gov/cdo-web), located approximately 10 km northeast of the sampling site. The discharge rates from the upper St. Johns River into Lake George were obtained from the United States Geological Survey river gauge monitoring site at Astor, Florida (waterdata.usgs.gov). Monthly average discharge rates were calculated from the mean daily averages. Missing monthly discharge rates from January–July 2001 were substituted by overall monthly averages across the 18-year span for the data.

2.4. Estimations for light availability

Mean light intensity in the mixed layer (I_m) was indirectly calculated from SD by using the Beer–Lambert's relationship:

$$I_m = \frac{I_0}{(K_t)(Z_m)} \left\{1 - e^{[-(K_t)(Z_m)]}\right\}$$

The mean daily surface PAR incident irradiance (I_0) was calculated from mean daily irradiance at latitude 30°N on a monthly basis (Oswald and Gotaas 1957), assuming cloudless conditions. The extinction coefficient (K_t) was estimated as 1.7 divided by the Secchi depth (SD^{-1}), and the depth of mixed layer (Z_m) was assumed to be equal to the depth of sampling site (z) due to the shallowness