Ecological Informatics 34 (2016) 129–138

Contents lists available at ScienceDirect

Ecological Informatics

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

Identifying the factors determining blooms of cyanobacteria in a set of shallow lakes

J.-P. Descy *, F. Leprieur, S. Pirlot, B. Leporcq, J. Van Wichelen, A. Peretyatko, S. Teissier, G.A. Codd, L. Triest, W. Vyverman, A. Wilmotte

ARTICLE INFO

Article history:
Received 26 October 2015
Received in revised form 19 May 2016
Accepted 23 May 2016
Available online 24 May 2016

Keywords:
Eutrophication
Modelling
Boosted regression trees
Lake management

ABSTRACT

There is a strong interest in developing a capacity to predict the occurrence of cyanobacteria blooms in lakes and to identify the measures to be taken to reduce water quality problems associated with the occurrence of potentially harmful taxa. Here we conducted a weekly to bi-weekly monitoring program on five shallow eutrophic lakes during two years, with the aim of gathering data on total cyanobacterial abundance, as estimated from marker pigments determined by HPLC analysis of phytoplankton extracts. We also determined bloom composition and measured weather and limnological variables. The most frequently identified taxa were *Aphanizomenon flos-aquae*, *Microcystis aeruginosa*, *Planktothrix agardhii* and *Anabaena* spp. We used the data base composed of a total of 306 observations and an adaptive regression trees method, the boosted regression tree (BRT), to develop predictive models of bloom occurrence and composition, based on environmental conditions. Data processing with BRT enabled the design of satisfactory prediction models of cyanobacterial abundance and of the occurrence of the main taxa. Phosphorus (total and soluble reactive phosphate), dissolved inorganic nitrogen, epilimnion temperature, photoperiod and euphotic depth were among the best predictive variables, contributing for at least 10% in the models, and their relative contribution varied in accordance with the ecological traits of the taxa considered. Meteorological factors (wind, rainfall, surface irradiance) had a significant role in species selection. Such results may contribute to designing measures for bloom management in shallow lakes.

1. Introduction

Cyanobacterial blooms, mass developments of cyanobacteria in water bodies, have become a recurrent and increasingly important phenomenon in freshwaters worldwide over recent decades (Carmichael, 2008; Codd et al., 2005a; Hudnell, 2008). The formation of such blooms in surface waters is closely linked to water eutrophication (Chorus, 2001), and presents major potential hazards for human and animal health. They interfere in various negative ways with the sustainable use of surface waters for e.g. drinking water treatment, recreation, irrigation and fisheries. Cyanobacterial blooms can present major potential hazards to human and animal health via their common ability to produce potent toxins (cyanotoxins) (Carmichael, 2008; Metcalf and Codd, 2012). The cyanotoxins are mainly released into the water column during the collapse of the blooms. Ingestion of, or contact with water containing cyanobacterial cells or cyanotoxins can cause a range of adverse health outcomes from mild to fatal health damage (Bell and Codd, 1996; Carmichael et al., 2001; Codd and Morrison, 2005b; de Figueiredo et al., 2004; Dittmann and Wiegand, 2006). The general conditions that favour bloom-forming cyanobacteria in lakes are well known (Oliver et al., 2012): as photosynthetic organisms, cyanobacteria need light, macronutrients and appropriate temperature conditions for growth. Most species which contribute to nuisance blooms are favoured by high dissolved phosphorus concentration (Reynolds, 2006), high pH (Shapiro, 1997), high temperature (Paerl and Huisman, 2008, 2009) long water retention time and, for species that can adjust their vertical position in the water column, thermal stratification (Reynolds, 2006). As a result, cyanobacteria often outcompete other phytoplankton in summer conditions in eutrophic lakes, where nuisance blooms are most likely to occur. Other biological and ecological traits may favour cyanobacterial dominance, including buoyancy regulation and a capacity to migrate vertically in the water column (Walsby, 1994), resistance to zooplankton grazing (Reynolds, 2006), including inducible defences (Van Donk et al., 2011), and efficient N_2 fixation in heterocystous species. Yet not all cyanobacteria are equivalent: they exhibit a large ecophysiological diversity, which has allowed them to develop various ecological strategies, which are summarized by their contribution to several phytoplankton functional groups developing in a wide range of conditions (Reynolds et al., 2002). Some studies have also demonstrated how biotic interactions, chiefly grazing and parasitism, may determine changes in taxa dominance (see e.g. Van Wichelen et al., 2010). Therefore, prediction and management of cyanobacterial blooms in lakes is no easy task: for instance, the contribution of climate change to the increasing incidence of 'blue-green' blooms is still debated due to complex interactions between lake characteristics and human usages, temperature

* Corresponding author.
E-mail address: jpdescy@gmail.com (J.-P. Descy).

http://dx.doi.org/10.1016/j.ecoinf.2016.05.003
1574-9541/

increase, water column structure, duration of stratified periods and changes in nutrient inputs from the catchment, related to agricultural practice and change in rainfall and hydrology (Carey et al., 2012; Elliott, 2012; Reichwaldt and Ghadouani, 2012; Rigosi et al., 2015).

Different approaches have been used for modelling phytoplankton blooms in freshwaters. Simulation models have so far been successful for predicting total cyanobacterial abundance in specific lakes subjected to anthropogenic pressures under different scenarios (e.g. Elliott, 2012). Other approaches have used data-based models of various kinds to achieve prediction of the probability of bloom occurrence based on various environmental variables. This applies to total cyanobacterial occurrence or abundance (e.g. Peretyatko et al., 2010; Rigosi et al., 2015) and the occurrence of particular taxa (Bobbin and Recknagel, 2001; Jeong et al., 2003; Lek et al., 2005; Recknagel et al., 1997; Recknagel et al., 2014, 2016; Zhang et al., 2012). These prediction models have mostly used various machine-learning algorithms and may be useful for determining strategies for lake management, as they help in identifying the key variables to act upon in specific lake environments in order to achieve bloom reduction or to provide an early-warning tool (Zhang et al., 2012).

Developing a capacity to predict the occurrence and abundance of particular taxa and to assess in which conditions they will arise is a critical step in bloom management, as the presence of particular cyanotoxins is primarily linked to bloom composition (Metcalf and Codd, 2012). For instance, among gas-vacuolate bloom-forming cyanobacteria, the production of microcystins is most often linked to the development of species of *Microcystis*, *Anabaena* and *Planktothrix*, but not for example, of *Aphanizomenon* or *Cylindrospermopsis* (Codd et al., 2005a; Metcalf and Codd, 2012). In this study we used a data base constituted through monitoring of phytoplankton blooms and related environmental conditions in several eutrophic shallow lakes. We applied an adaptive regression trees method, the boosted regression tree (BRT), in order to develop a model for predicting total abundance of cyanobacteria and the occurrence of the main taxa found in the studied lakes: *Aphanizomenon flos-aquae*, *Microcystis aeruginosa*, *Planktothrix agardhii* and *Anabaena* spp. We show that data processing with BRT enabled the design of satisfactory prediction models and allowed identification of the key environmental factors involved in bloom occurrence, in accordance with the ecological traits of the taxa considered.

2. Materials and methods

2.1. Study sites

Five Belgian lakes prone to cyanobacterial blooms were sampled throughout the study, carried out mostly in 2007–2008, during the B-BLOOMS2 research project (www.bblooms.be): two lakes located in Flanders (Westveld Pond and Lake Donkmeer), two lakes in the Brussels region (Ixelles Ponds, Ix1 and Ix2), and one lake in Wallonia (Lake Falemprise). Westveld Pond (mean TP 0.73 mg L^{-1}) and Lake Donkmeer (mean TP 0.33 mg L^{-1}) are situated in the neighbourhood of the city of Ghent. Westveld Pond is a small parkland, shallow pond of ca. 2000 m^2 surface area and a maximal depth of about 2 m, while Lake Donkmeer is a relatively large (ca. 86 ha), shallow (depth max 2.5 m) lake originating from peat-digging and now intensively used for recreation. Macrophytes were absent in both water bodies. The Ixelles ponds are two linked urban ponds (IxP1-upstream and IxP2-downstream) each of 1–2 ha with a mean depth of 1.5 m and mean TP concentration exceeding 0.1 mg L^{-1} before the removal of their sediment in 2010. Lake Falemprise, in the Walloon Region (Fig. 1, C) is one of the Eau d'Heure Lakes, which are used for discharge regulation of the River Sambre and for recreational activities. Lake Falemprise is a reservoir having a surface area of 47 ha, a mean depth of 2.6 m and a maximum depth of 12 m. The classic indicators of trophy, phosphorus (TP), chlorophyll a concentration (Chla) and Secchi depth (Z_S), allowed Sarmento and Descy (2008) to classify the lake as eutrophic, although the nutrient concentration was lower than in the other lakes (mean TP 0.06 mg L^{-1}). Macrophytes (mostly *Elodea nuttallii*, *Potamogeton pectinatus*, and filamentous green algae) were present in shallow littoral areas in this lake, with variable abundance between years.

2.2. Methods for data collection

The sampling of the lakes was carried out using a common protocol, on a weekly basis from spring to autumn in 2007 and 2008. Integrated samples of the whole water column were collected in the very shallow lakes, whereas discrete samples were taken every metre from surface to bottom (0–4 m) in Lake Falemprise. For the latter, data from previous studies in 2002–2004 were included. Limnological profiles were acquired using a YSI 6600 or Hydrolab DS5 multiprobe. Weather data in Lake Falemprise were collected by an in situ Davis Vantage Pro 6150 C meteo station, and irradiance was recorded with a surface LICOR sensor LI-190 SB connected to a data logger. For the other lakes, data were obtained from the closest station of the RMI (Royal Meteorological Institute). Zooplankton was collected from each lake, using buckets or a Schindler-Patalas plankton trap, and enumerated using a dissecting microscope for the larger forms (small and large cladocerans, adult cyclopoid and calanoid copepods) and an inverted microscope for the smaller forms (copepod nauplii, rotifers). Physical and chemical measurements were based on standard techniques of water analysis.

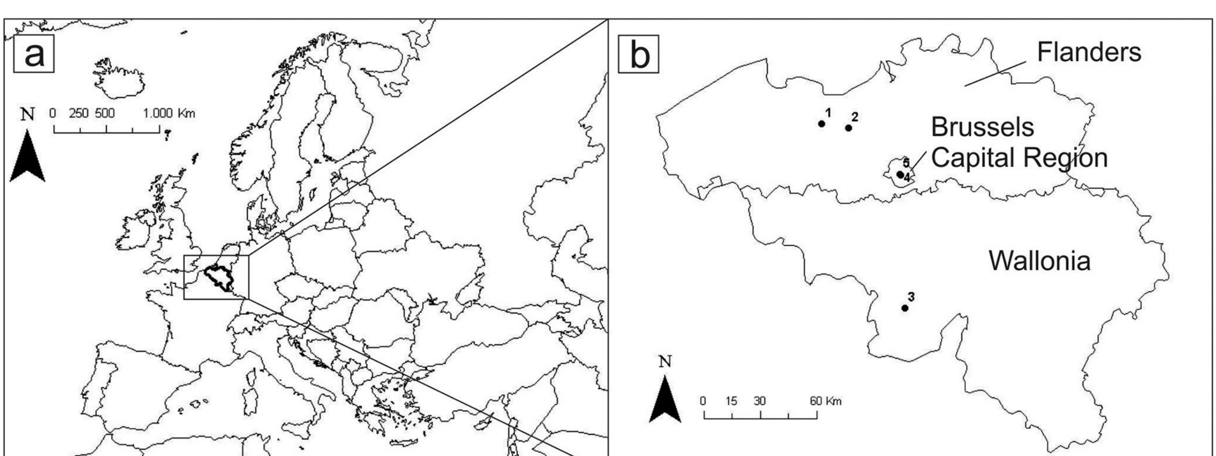

Fig. 1. Map of the studied sites. a: Location of Belgium in Europe. b: Location of the studied water bodies in the Belgian regions: 1. Westveld Pond; 2. Lake Donkmeer; 3. Lake Falemprise; 4, 5. Ixelles Ponds.