Science of the Total Environment 409 (2011) 5353–5358

Contents lists available at SciVerse ScienceDirect

Science of the Total Environment

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

Cyanobacterial blooms: Statistical models describing risk factors for national-scale lake assessment and lake management

Laurence Carvalho [a,*], Claire A. Miller (nee Ferguson) [b], E. Marian Scott [b], Geoffrey A. Codd [c], P. Sian Davies [d], Andrew N. Tyler [e]

[a] Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK
[b] School of Mathematics and Statistics, University of Glasgow, Glasgow, G12 8QW, UK
[c] Division of Molecular Microbiology, College of Life Sciences, University of Dundee, Dundee, DD1 4HN, UK
[d] Environment Agency (EA), Waveney Works, Station Road, Haddiscoe, Great Yarmouth, NR31 9JA, UK
[e] School of Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA, UK

ARTICLE INFO

Article history:
Received 24 May 2011
Received in revised form 10 August 2011
Accepted 12 September 2011
Available online 4 October 2011

Keywords:
Algal bloom
Blue-green algae
Cyanotoxin
Phosphorus
Restoration
Water framework directive

ABSTRACT

Cyanobacterial toxins constitute one of the most high risk categories of waterborne toxic biological substances. For this reason there is a clear need to know which freshwater environments are most susceptible to the development of large populations of cyanobacteria. Phytoplankton data from 134 UK lakes were used to develop a series of Generalised Additive Models and Generalised Additive Mixed Models to describe which kinds of lakes may be susceptible to cyanobacterial blooms using widely available explanatory variables. Models were developed for log cyanobacterial biovolume. Water colour and alkalinity are significant explanatory variables and retention time and TP borderline significant (R^2-adj = 21.9%). Surprisingly, the models developed reveal that nutrient concentrations are not the primary explanatory variable; water colour and alkalinity were more important. However, given suitable environments (low colour, neutral-alkaline waters), cyanobacteria do increase with both increasing retention time and increasing TP concentrations, supporting the observations that cyanobacteria are one of the most visible symptoms of eutrophication, particularly in warm, dry summers. The models can contribute to the assessment of risks to public health, at a regional- to national level, helping target lake monitoring and management more cost-effectively at those lakes at the highest risk of breaching World Health Organisation guideline levels for cyanobacteria in recreational waters. The models also inform restoration options available for reducing cyanobacterial blooms, indicating that, in the highest risk lakes (alkaline, low colour lakes), risks can generally be lessened through management aimed at reducing nutrient loads and increasing flushing during summer.

1. Introduction

Cyanobacteria are natural inhabitants of freshwaters, where they fulfil important roles in primary production, nitrogen fixation and the cycling of matter (Howarth et al., 1988). They can, however, present hazards to the health of humans and other animals when large populations flourish to produce blooms and particularly when these accumulate on lake surfaces or along shorelines as scums. They constitute a major health hazard as they frequently produce numerous potent toxins (cyanotoxins) that can result in a range of adverse health effects from mild, e.g. skin irritations and gastrointestinal upsets, to fatal (Codd et al., 1999, 2005). Cyanotoxins constitute one of the most high risk categories of waterborne toxic biological substances. This is because not only are the health hazards which they present significant, but exposure to potentially harmful doses of the cyanotoxins can occur, with blooms and scums being a common annual feature in many lakes or reservoirs worldwide. There is, therefore, a great need for understanding where and when cyanobacterial blooms are likely to occur and to what extent. This knowledge would help target lake monitoring and management more efficiently at those lakes at highest risk of breaching World Health Organisation (WHO) and national guidelines (Chorus and Bartram, 1999; WHO, 2003, 2004).

Research over recent decades has identified a number of physical factors that favour cyanobacterial blooms, with the main focus on seasonal drivers, such as warmer temperatures, windiness and consequently the intensity of thermal stratification of the water column (Foy et al., 1976; Mischke, 2003; Reynolds, 2006). Bloom-forming cyanobacteria have been shown to be favoured by high alkalinities and associated high pH (Shapiro, 1984). It is also a widely held view that the increasing magnitude and frequency of cyanobacterial blooms is primarily related to the nutrient enrichment of freshwaters.

* Corresponding author at: Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK. Tel: +44 131 445 4343.
E-mail address: laca@ceh.ac.uk (L. Carvalho).

0048-9697/$ – see front matter. Crown Copyright © 2011 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.scitotenv.2011.09.030

Indeed there have been several studies showing empirically that bloom frequency is related to the general nutrient status of a lake (Gorham et al., 1974; Dokulil and Teubner, 2000; Downing et al., 2001; Reynolds and Petersen, 2000; Schindler et al., 2008). Supporting evidence of a relationship between nutrient enrichment and cyanobacterial abundance is largely derived from long-term studies of enrichment at a few selected individual sites, usually lowland, alkaline, eutrophic lakes, and often examining individual cyanobacterial species. There have been a few published studies examining the relative% abundance of cyanobacteria across eutrophication gradients in large datasets (Downing et al., 2001; Ptacnik et al., 2008), but a more comprehensive quantitative analysis of what factors affect actual cyanobacterial abundance across a wide range of lake types at a regional or national scale has not been carried out.

A more 'global' approach has been adopted to develop empirical statistical models to predict the amount of phytoplankton chlorophyll$_a$ for given concentrations of phosphorus in lake waters (OECD, 1982; Phillips et al., 2008) and also to predict natural background chlorophyll concentrations for individual lakes (Carvalho et al., 2009). The present study aims to take a similar statistical modelling approach, but to more specifically model the variability in cyanobacteria at a national level. It aims to understand the key environmental drivers which are routinely available that favour cyanobacterial abundance in lakes across the UK. We do not aim to understand the distribution and abundance of individual cyanobacterial species or functional groups as it is the risk from all potentially toxic cyanobacteria that is important. The models can be used to help identify which lakes are most susceptible to developing cyanobacterial blooms, enabling a more proactive, rather than reactive, strategy for monitoring and managing cyanobacterial health risks and other adverse impacts, including guiding restoration measures. More generally, the analysis examines whether empirical evidence from a large selection of lakes spanning a range of environmental conditions, supports nutrient concentrations as the key driver of cyanobacterial abundance in freshwater lakes.

2. Material and methods

2.1. Data

Phytoplankton data were available for summer months (July, August, September) from UK lakes during the period 2003 to 2006. Samples were either integrated tube samples from the middle of the lake, or where boat access was not possible, a sub-surface sample taken 0.3 m below the surface, using a weighted bottle and float attached to a rope and thrown from the shore near the outflow. The varying sampling method may add variability to the results (discussed later), although previous studies have shown that open water and outflow samples rarely differ markedly in terms of chlorophyll$_a$ (e.g. Bailey-Watts, 1978). Samples were preserved with Lugol's iodine solution and stored for counting for no longer than 1 year. Phytoplankton were counted using 5–10 mL Utermöhl sedimentation chambers at a range of magnifications (from ×40 to ×500) depending on cell size. In general, 400 counting units were measured across magnifications using low magnification full-chamber counts, intermediate magnification transects and high magnification fields of view. Counts and biovolume estimates of cells, colonies and filaments were made following the approach outlined by CEN (2004) and Brierley et al. (2007). A series of training workshops and ring-counts was undertaken to ensure these guidance documents were followed correctly and taxonomic identities were standardised. Sub-samples were also taken for nutrient and alkalinity measurements and analyses were carried out by accredited methods at UK environment agencies analytical chemistry laboratories. Colour was measured using absorbance at 400 nm, with measurement of a 100 Hazen standard solution at 400 nm used to convert the colour results to Hazen units.

Usually only one sample was available per month. If more than one sample was available, data were averaged by month. Some lakes were represented in the dataset by samples from different months in the same year. The 262 samples were taken from 134 lakes, with 63 lakes having more than one monthly sample in a particular year. Table 1 highlights that most of the data are for 2004 and 2005, however, there was a similar amount of data available for each month overall.

Table 2 lists the principal cyanobacterial genera considered as potentially toxin-forming in UK lakes. Taxonomy broadly followed John et al. (2002) which uses the genus name *Oscillatoria* for taxa that other floras refer to as *Planktothrix*. Records for *Snowella* and *Gomphosphaeria* were grouped as they were sometimes lumped together by counters. The most common genera known to cause toxin problems in UK freshwaters include *Microcystis*, *Aphanizomenon*, *Snowella*, *Oscillatoria* (*Planktothrix*) and *Anabaena*. The likelihood of an individual bloom of these genera containing potent toxins ranges from about 40% to at least 90% (Codd et al., 1999). The other genera were included in the analysis as they are associated with toxicity, although the toxic components are not so well characterised and they also tend to bloom less frequently.

The natural log cyanobacterial biovolume (μm^3 mL^{-1}) was taken as the response here with a small arbitrary constant of 0.001 added, before transforming, to eliminate zeros. Cyanobacterial biovolume is a direct measure of actual cyanobacterial abundance and can be, therefore, related to potential toxin concentrations (Codd et al., 2005). No a-priori subjective decisions were made in selecting explanatory variables. Data were obtained for the following widely-available explanatory variables: lake area (km^2), altitude (m above sea level), mean depth (m), alkalinity (mEq L^{-1}), colour (Pt L^{-1}), retention time (years), total phosphorus (TP) (μg L^{-1}), total nitrogen (TN) (mg L^{-1}) and chlorophyll$_a$ (μg L^{-1}). Retention times were taken from the UK lakes database (http://www.uklakes.net/) and were estimates based on lake volume and 30-year average annual total rainfall in the catchment. Table 3 includes summary statistics for all explanatory variables and for the response of cyanobacterial biovolume. It illustrates that there is a small amount of missing data for five of the variables.

2.2. Statistical methods

Due to the presence of non-linear and non-monotonic relationships between some of the explanatory variables and the cyanobacterial response being investigated, generalised additive models, GAMs (Hastie and Tibshirani, 1990; Wood, 2006) were adopted throughout, assuming normal errors. This approach, with a log cyanobacterial response, was taken since the range of values for cyanobacteria is very large going from 0 to 7.5×10^8 and adding a very small constant before transforming reflects the fact that it is unlikely that all of the zeros are truly zero counts. In these models the relationship between the response and the explanatory variables is allowed to be a smooth function instead of restricting relationships to be linear.

Since some of the lakes have measurements recorded for more than one month within the same year, generalised additive mixed

Table 1
Summary of no. of lake samples available by month and year.

Year	July	August	September
2003	0	2	2
2004	56	48	72
2005	15	24	18
2006	8	16	1
All years	79	90	93