Harmful Algae 31 (2014) 26–34

Contents lists available at ScienceDirect

Harmful Algae

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

# Inductive reasoning and forecasting of population dynamics of *Cylindrospermopsis raciborskii* in three sub-tropical reservoirs by evolutionary computation

Friedrich Recknagel [a,*], Philip T. Orr [b], Hongqing Cao [a]

[a] School of Earth and Environmental Sciences, University of Adelaide, SA 5005, Australia
[b] Seqwater, PO Box 16146, City East, Qld 4002, Australia

ARTICLE INFO

*Article history:*
Received 15 February 2013
Received in revised form 10 September 2013
Accepted 10 September 2013

*Keywords:*
*Cylindrospermopsis raciborskii*
Cyanobacteria blooms
Evolutionary computation
Ecological thresholds
Ecological relationships
Forecasting

ABSTRACT

Seven-day-ahead forecasting models of *Cylindrospermopsis raciborskii* in three warm-monomictic and mesotrophic reservoirs in south-east Queensland have been developed by means of water quality data from 1999 to 2010 and the hybrid evolutionary algorithm HEA. Resulting models using all measured variables as inputs as well as models using electronically measurable variables only as inputs forecasted accurately timing of overgrowth of *C. raciborskii* and matched well high and low magnitudes of observed bloom events with $0.45 \leq r^2 > 0.61$ and $0.4 \leq r^2 > 0.57$, respectively. The models also revealed relationships and thresholds triggering bloom events that provide valuable information on synergism between water quality conditions and population dynamics of *C. raciborskii*. Best performing models based on using all measured variables as inputs indicated electrical conductivity (EC) within the range of 206–280 mS m$^{-1}$ as threshold above which fast growth and high abundances of *C. raciborskii* have been observed for the three lakes. Best models based on electronically measurable variables for the Lakes Wivenhoe and Somerset indicated a water temperature (WT) range of 25.5–32.7 °C within which fast growth and high abundances of *C. raciborskii* can be expected. By contrast the model for Lake Samsonvale highlighted a turbidity (TURB) level of 4.8 NTU as indicator for mass developments of *C. raciborskii*.

Experiments with online measured water quality data of the Lake Wivenhoe from 2007 to 2010 resulted in predictive models with $0.61 \leq r^2 > 0.65$ whereby again similar levels of EC and WT have been discovered as thresholds for outgrowth of *C. raciborskii*. The highest validity of $r^2 = 0.75$ for an in situ data-based model has been achieved after considering time lags for EC by 7 days and dissolved oxygen by 1 day. These time lags have been discovered by a systematic screening of all possible combinations of time lags between 0 and 10 days for all electronically measurable variables. The so-developed model performs seven-day-ahead forecasts and is currently implemented and tested for early warning of *C. raciborskii* blooms in the Wivenhoe reservoir.



## 1. Introduction

Lakes Wivenhoe, Somerset and Samsonvale are subtropical, warm-monomictic and mesotrophic reservoirs located near Brisbane in the subtropical southeast of Queensland, Australia. Each of these reservoirs have annual recurring blooms of the potentially toxic cyanobacterium *Cylindrospermopsis raciborskii* (Woloszynska) Seenaya et Subba Raju (Orr et al., 2010). *C. raciborskii* produces cylindrospermopsins (CYN's) which are hepatotoxic alkaloid cyanotoxins that present a risk to human health (e.g. Hawkins et al., 1985) and which must be removed from raw water during water treatment. Controlling the development of *C. raciborskii* within these reservoirs is a key goal of Seqwater, the water authority responsible for the management of the reservoirs. However, *C. raciborskii* is ecologically adaptable and can form blooms under a range of light, temperature and nutrient regimes (Isvanovics et al., 2000; Sprober et al., 2003; Briand et al., 2002). It also tolerates a wide range of other environmental conditions including oligohaline waters (Caldwell, 2001) and nitrogen depleted waters through its ability to fix atmospheric nitrogen ($N_2$) (Bouvy et al., 2000; Moisander et al., 2008). Although *C. raciborskii* is being considered a tropical or sub-tropical species and recorded in tropical countries such as Brazil (e.g. Bouvy et al., 2000; Branco and Senna, 1994; DeSouza et al., 1998) it is now increasingly detected in temperate regions of Europe (Dokulil and Mayer, 1996; Padisak, 1997; Briand et al., 2002; Fastner et al.,

* Corresponding author. Tel.: +61 8 8313 3953; fax: +61 8 8303 6222.
*E-mail address:* friedrich.recknagel@adelaide.edu.au (F. Recknagel).

1568-9883/$ – see front matter 
http://dx.doi.org/10.1016/j.hal.2013.09.004

**Table 1**
Summary of water quality data of the three reservoirs used in this study.

| | Wivenhoe | | Somerset | Samsonvale |
|---|---|---|---|---|
| | Historical data 1999–2009 | On-line data 2007–2010 | Historical data 1999–2009 | Historical data 1999–2009 |
| | Mean ± SD | | | |
| Electrical conductivity (EC) $\mu$S cm$^{-1}$ | 354.90 ± 72.64 | 348.75 ± 72.1 | 235.75 ± 45.35 | 226.94 ± 41.62 |
| Turbidity (TURB) NTU | 12.92 ± 56.76 | 2.71 ± 3.59 | 2.32 ± 2.25 | 2.72 ± 1.4 |
| Water temperature (WT) °C | 22.66 ± 3.70 | 22.51 ± 3.61 | 22.69 ± 3.98 | 22.47 ± 3.64 |
| Dissolved oxygen (DO) mg L$^{-1}$ | 8.67 ± 1.25 | 7.81 ± 1.61 | 7.55 ± 1.82 | 7.96 ± 1.4 |
| pH | 8.21 ± 0.32 | 8.15 ± 0.54 | 7.97 ± 0.46 | 7.95 ± 0.34 |
| Silica ($SiO_2$) mg L$^{-1}$ | 3.24 ± 2.69 | | 4.25 ± 2.13 | 2.27 ± 1.78 |
| Total nitrogen (TN) mg L$^{-1}$ | 0.48 ± 0.078 | | 0.59 ± 0.15 | 0.56 ± 0.09 |
| Total phosphorus (TP) mg L$^{-1}$ | 0.019 ± 0.012 | | 0.028 ± 0.017 | 0.018 ± 0.007 |
| Chlorophyll_a (Chl_*a*) mg L$^{-1}$ | 7.82 ± 4.54 | 6.41 ± 3.58 | 10.66 ± 7.49 | 11.6 ± 5.23 |
| *Cylindrospermopsis* cells/mL | 10,108 ± 20,698 | 6098 ± 12,466 | 10,259 ± 250,489 | 14,853 ± 19,110 |

2007), New Zealand (Wood and Stirling, 2003), Canada (Hamilton et al., 2005) and the U.S. (Chapman and Schelske, 1997; Calandrino and Perl, 2011).

The drivers of *C. raciborskii* blooms are highly complex and the synergism between *C. raciborskii* bloom development and physical-chemical water quality conditions is still poorly understood. Novel techniques for inductive reasoning and forecasting from historical data can reveal environmental conditions and thresholds that have triggered blooms of *C. raciborskii* in the past, and synthesise short-term forecasting models for up to 14 days ahead. These models can give early warning for timely operational control and risk management of cyanobacteria bloom events.

Artificial neural networks (ANN) and evolutionary computation (EC) have proved to be powerful techniques for inductive reasoning and forecasting of highly complex limnological data (e.g. Recknagel et al., 1997, 2013). Even though both ANN and EC achieve similar good multivariate forecasting, only EC performs mathematically explicit syntheses and representation of underlying models. The hybrid evolutionary algorithm HEA (Cao et al., 2006, 2013) has been specifically designed and customised for inducing predictive and explanatory models from complex ecological data.

This study applied HEA to 11 years of water quality data collected from sites close to the dam wall of the three reservoirs and developed forecasting models to predict *C. raciborskii* cell concentrations up to 2 weeks in advance. The models also reveal underlying ecological relationships that trigger mass developments of *C. raciborskii*. Since the three lakes are similar in climate, circulation patterns and eutrophication levels, collective properties of *C. raciborskii* in the three lakes have been identified.

## 2. Materials and methods

### 2.1. The Lakes Wivenhoe, Somerset and Samsonvale

Eleven years of water quality data collected from Lakes Wivenhoe, Somerset and Samsonvale in south-east Queensland, Australia between 1999 and 2010 (see Table 1) were used for this study. The reservoirs are located to the north and west of Brisbane (Fig. 1) and supply drinking water for about 1.5 million people. Lake Samsonvale (also known as North Pine Dam) (27°15′S, 152°55′E) is the smallest reservoir with a surface area of 21.8 km$^2$ at full drinking water supply volume of 215,000 mL and has a bubble plume destratifier which is activated during the summer months to reduce stratification and help control *Cylindrospermopsis raciborskii* cell concentrations. Lake Somerset (27°7′S, 152° 33′E) is smaller with a surface area of 42.1 km$^2$ at full supply volume of 380,000 mL and is located upstream of the third, largest and most spatially diverse reservoir, Lake Wivenhoe (27°24′S, 152°36′E) (Orr et al., 2010). Lake Wivenhoe has a surface area of 107.5 km$^2$ and full supply volume of 1,165,000 mL. The catchments for these reservoirs are typically unprotected, with more than 50% dominated by cattle grazing pasture, and approximately 20% being natural vegetation (Orr et al., 2010).

Table 1 summarises the water quality data measured in the three reservoirs over an 11 year periods that have been utilised for modelling. Since the intervals for measuring the historical data ranged between weekly and biweekly, and sampling dates differed for physical, chemical and biological variables, the data have been interpolated to suit daily time steps for forecasting. For simplicity, we used linear interpolation and in order to

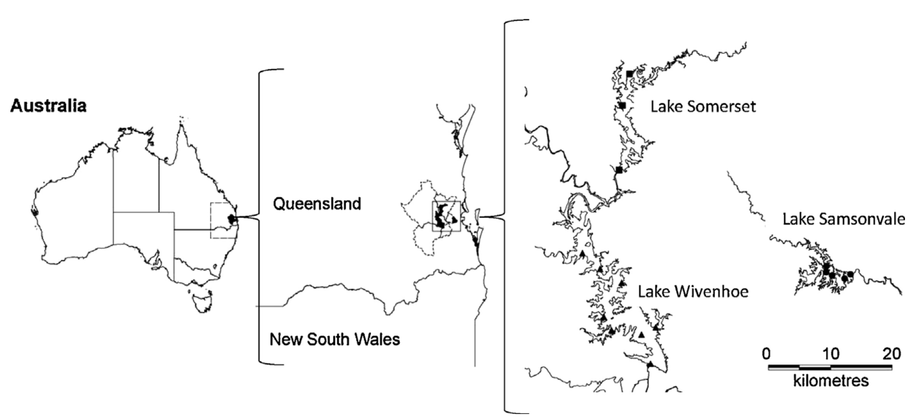


**Fig. 1.** The locations of the three reservoirs within Australia.