Accepted Manuscript

Fluorescence probes for real-time remote cyanobacteria monitoring: a review of challenges and opportunities

Edoardo Bertone, Michele A. Burford, David P. Hamilton

PII:	S0043-1354(18)30363-4
DOI:	10.1016/j.watres.2018.05.001
Reference:	WR 13764
To appear in:	*Water Research*
Received Date:	08 November 2017
Revised Date:	02 May 2018
Accepted Date:	02 May 2018

Please cite this article as: Edoardo Bertone, Michele A. Burford, David P. Hamilton, Fluorescence probes for real-time remote cyanobacteria monitoring: a review of challenges and opportunities, *Water Research* (2018), doi: 10.1016/j.watres.2018.05.001

Fluorescence probes for real-time remote cyanobacteria monitoring: a review of challenges and opportunities

Edoardo Bertone [a,b] *, Michele A. Burford[c], David P. Hamilton[c]

[a] *Griffith School of Engineering and Built Environment, Griffith University, Parklands Drive, Southport, Queensland 4222, Australia*

[b] *Cities Research Institute, Griffith University, Parklands Drive, Southport, Queensland 4222, Australia*

**Corresponding author: edoardo.bertone@griffithuni.edu.au*

[c] *Australian Rivers Institute, Griffith University, Kessels Road, Nathan, Queensland 4111, Australia*

Abstract

In recent years, there has been a widespread deployment of submersible fluorescence sensors by water utilities. They are used to measure diagnostic pigments and estimate algae and cyanobacteria abundance in near real-time. Despite being useful and promising tools, operators and decision-makers often rely on the data provided by these probes without a full understanding of their limitations. As a result, this may lead to wrong and misleading estimations which, in turn, means that researchers and technicians distrust these sensors. In this review paper, we list and discuss the main limitations of such probes, as well as identifying the effect of environmental factors on pigment production, and in turn, the conversion to cyanobacteria abundance estimation. We argue that a comprehensive calibration approach to obtain reliable readings goes well beyond manufacturers' recommendations, and should involve several context-specific

experiments. We also believe that if such a comprehensive set of experiments is conducted, the data collected from fluorescence sensors could be used in artificial intelligence modelling approaches to reliably predict, in near real-time, the presence and abundance of different cyanobacteria species. This would have significant benefits for both drinking and recreational water management, given that cyanobacterial toxicity, and taste and odour compounds production, are species-dependent.

Keywords: Artificial Intelligence; Cyanobacteria; Fluorescence; Remote sensors; Water quality; Water resources management

1 Introduction

Cyanobacteria are considered to be among the most interesting organisms for ecological and phycological studies (Beutler et al. 2003), in addition to being some of the oldest organisms on Earth. They are primary producers and some of them produce toxins and/or taste and odour compounds (Carmichael 2001, Schopf 1996, Waterbury et al. 1979). There is evidence to suggest that climate change may increase the frequency and magnitude of cyanobacterial blooms (O'Neil et al. 2012).

Submersible fluorescence sensors can rapidly provide an estimate of the biomass of algae, including cyanobacteria. Studies have shown that sensors correlate well with algal and cyanobacterial biomass measures in laboratory experiments, and thus they have been recommended as sensitive tools for real-time in-situ water management (Bastien et al. 2011, Brient et al. 2008, Gregor and

Maršálek 2004). However, there have been challenges with the calibration of these sensors, particularly in highly turbid environments, or when species composition changes rapidly, either spatially or temporally. The use of remote fluorescence instrumentation, combined with curve-fitting techniques, provides a possible means to determine the presence of different groups of algae in a sample (Hodges et al. 2017). The premise for this application, however, is that the fluorescence excitation spectra must be constant for each algal species (Beutler et al. 2002). This appears to be the case for algal groups, typically referred to by researchers studying algal fluorescence as green (chlorophytes), brown (predominantly diatoms, dinoflagellates and golden brown flagellates), or mixed algal assemblages (Beutler et al. 2002). However in the case of cyanobacteria (referred to as the blue group), the spectrum changes with environmental conditions (Beutler et al. 2003). This not only hinders the ability to differentiate different cyanobacterial species, but also undermines the reliability of such probes in estimating the total biomass of cyanobacteria. Therefore environmental and technological interferences need to be accounted for.

This review examines environmental factors and technological limitations that can affect both the cyanobacterial spectra, and accurate measurement of biomass with commonly available in-situ fluorescence probes. If these factors can be accounted for, such tools would provide a more reliable means of monitoring blooms. In addition, the use of algorithms and models derived from fluorescence sensors' data could be used to provide estimates about which cyanobacterial species are likely to be present, with links to the

79 prediction of the risk of presence of toxins, or taste and odour
80 compounds.

81 This review firstly examines how environmental factors affect the
82 pigment content of algal species. These factors complicate the
83 conversion of probe readings to cyanobacterial biomass. Next, the
84 technical constraints hindering an accurate reading, and inherent
85 interferences made in the fluorescence measurements, are examined.
86 The second part of the review assesses the feasibility of rapidly
87 estimating the relative contribution of different species to the total
88 biomass, as well as predicting toxin or taste and odour risks. Table 1
89 provides a list of the relevant literature analysed and cited based on
90 the topic discussed.

91 **Table 1 - List of relevant references based on topic**

Topic	Relevant literature cited
Environmental interferences: Chlorophyll-a and phycocyanin yields	Bryant (1982) Alpine and Cloern (1985) Everitt et al. (1990) Lee et al. (1995) Mackey et al. (1996) Henrion et al. (1997) MacColl (1998), Rapala (1998) Beutler et al. (2002) Beutler et al. (2003) Izydorczyk et al. (2005) Bryant (2006) Pemberton et al. (2007), Seppälä et al. (2007), Gregor et al. (2007) Brient et al. (2008), Randolph et al. (2008) Millie et al. (2010), Ziegmann et al. (2010), Richardson et al. (2010) Allan et al. (2011), McQuaid et al. (2011), Erickson et al. (2011) Bastien et al. (2011) NSW Office of Water (2012), Chang et al. (2012), Zamyadi et al. (2012) Bowling et al. (2013), Horváth et al. (2013) Sobiechowska-Sasim et al. (2014) Kudela et al. (2015), Korak et al. (2015), Zieger (2015) Hodges (2016b), Zamyadi et al. (2016a) Hodges et al. (2017), Steiner et al. (2017), Kong et al. (2017) McBride and Rose (in press)

Instrument limitations and opportunities	Beutler et al. (1998) Beutler et al. (2002) Beutler et al. (2003) Gregor and Maršálek (2004) Seppälä et al. (2007), Pemberton et al. (2007), Gregor et al. (2007) Brient et al. (2008) Millie et al. (2010), Ziegmann et al. (2010), Richardson et al. (2010) Pobel et al. (2011), Bastien et al. (2011), Dennis et al. (2011), Pomati et al. (2011), Hashemi et al. (2011) NSW Office of Water (2012) Bowling et al. (2013), Horváth et al. (2013), Pomati et al. (2013) Sobiechowska-Sasim et al. (2014) Hamilton et al. (2015), Zieger (2015) Zamyadi et al. (2016a), Bowling et al. (2016), Hodges (2016b), Zhou et al. (2016) McBride and Rose (in press)
Fluorescence measurement limitations	Schreiber et al. (1995) Asai et al. (2001) Gregor and Maršálek (2004) Kostoglidis et al. (2005) Gregor et al. (2007), Seppälä et al. (2007) Brient et al. (2008), Sackmann et al. (2008) Millie et al. (2010) McQuaid et al. (2011), Bastien et al. (2011), Hashemi et al. (2011) Chang et al. (2012), Zamyadi et al. (2012) Bowling et al. (2013) Korak et al. (2015), Wang et al. (2015) Zamyadi et al. (2016a), Hodges (2016b), Bowling et al. (2016) Watson et al. (2016) McBride and Rose (in press)
Estimating proportion of different algal groups, classes, species	Gieskes et al. (1988) Everitt et al. (1990) Mackey et al. (1996) Henrion et al. (1997) Pinckney et al. (2001) Izydorczyk et al. (2005) Millie et al. (2010) Bastien et al. (2011) Zieger (2015) Kpodonu et al. (2016), Zamyadi et al. (2016a) van der Linden et al. (2017)
Predicting taste and odour and toxicity risk	Hawkins et al. (2001) Vézie et al. (2002) Rohrlack et al. (2008), Henderson et al. (2008) Everson et al. (2009) Chapman (2010), Ziegmann et al. (2010) American Water Works Association (2011), Bastien et al. (2011) Everson et al. (2011) Li et al. (2012) Leloup et al. (2013) Kong et al. (2014), Wert et al. (2014) Su et al. (2015), Wang et al. (2015), Korak et al. (2015) Bertone and O'Halloran (2016), Willis et al. (2016)

	Zamyadi et al. (2016a), (Zamyadi et al. 2016b), Watson et al. (2016) Pivokonsky et al. (2016)
Relevant artificial intelligence modelling applications	Jaynes (1957) Maier and Dandy (2000), Wallace and Hamilton (2000), Singh (2000) Tanyimboh and Sheahan (2002) Embleton et al. (2003) Castelletti and Soncini-Sessa (2007), Uusitalo (2007), Hamilton et al. (2007) Fenton and Neil (2008), Lee and Wentz (2008) Johnson et al. (2010) Mosleh et al. (2012), Chen and Pollino (2012) Bertone et al. (2015), Rigosi et al. (2015) Bertone et al. (2016a), Bertone et al. (2016b)

2 Environmental interferences: Chlorophyll-a and phycocyanin yields

The functioning principle for estimating algal and, more specifically, cyanobacterial biomass using fluorometry relies on the determination of diagnostic pigments, with a focus on chlorophyll *a* (chl-a)*, b* and *c* and carotenoids, as well as phycobilins (such as phycocyanin and phycoerythrin).

Chlorophyll *a*, measured by in situ fluorometry, is a proxy for total algal biomass, as it is easy to calculate, and it is universally present in both prokaryotic and eukaryotic algae (Millie et al. 2010). Values have also been derived from water reflectance data monitored by satellites (Allan et al. 2011) and, along with other spectral data, were used to differentiate between two key cyanobacterial species (Kudela et al. 2015). However, typically chlorophyll fluorescence does not provide information on the algal community composition (Ziegmann et al. 2010). The closest that can be achieved in terms of community differentiation is at the algal class level using a range of excitation and emission fluorescence spectra, e.g., for green algae. In addition, many studies (e.g., Bowling et al. (2013), Brient et al. (2008),

Pemberton et al. (2007)), have not been able to establish a strong correlation between cyanobacterial biovolume and chl-a measured by in-situ fluorometry.

Cyanobacteria contain phycobilisomes which absorb light in a different range to chl-a, i.e., between 550 and 650 nm, expanding the range of wavelengths which may be available for photosynthesis (Bryant 2006, Korak et al. 2015). In particular, phycocyanin (PC) and phycoerythrin (PE) are two fluorescent phycobilisomes that can be used to quantify cyanobacterial biomass, with PC most commonly measured using on-line fluorescence probes. The advantage of using phycobilisomes is that: (1) they only occur in cyanobacteria, allowing differentiation from other algae; and (2) they fluoresce at higher wavelengths than dissolved organic matter (DOM), thus reducing interference from these compounds, provided quenching effects (discussed later) are accounted for. PC is generally preferred to PE because PC is produced by all cyanobacteria while PE is produced only by some species (Bryant 1982, 2006). However, PE is more prevalent in benthic cyanobacteria because they have light absorption wavelengths aligned with the spectrum for bottom waters, due to differential absorption of light through the water column, resulting in blue-green light exposure (Korak et al. 2015, MacColl 1998). Therefore PE could be useful for fluorescence probes specifically targeting benthic cyanobacterial detection.

All cyanobacterial species have a specific range of chl-a and PC contents (i.e., quota) per cell (Brient et al. 2008, Richardson et al. 2010, Seppälä et al. 2007, Sobiechowska-Sasim et al. 2014), usually proportional to the cell volume (Brient et al. 2008), and have

different ratios of pigments (Ziegmann et al. 2010). Average cell volumes can be estimated based on species by using tables, databases and calculators readily available online (e.g. (DELWP)). Such averages can be subject to variations, with previous studies finding cell volumes considerably lower than values that have been reported in such commonly used standard cell size tables, especially in summer and autumn (NSW Office of Water 2012). Each cyanobacterial pigment has an excitation and emission spectrum, hence, based on spectral signature of a sample, it is possible to identify the different phycobilins and separate them from chl-a and other background signals (Seppälä et al. 2007). A summary table of pigment ratios, based on literature data, is provided in Mackey et al. (1996). Hodges (2016a) and Hodges et al. (2017), as well as Zamyadi et al. (2016a) also summarise several studies of pigment ratios for different species. Differences in pigment content and ratios arise because each species has a different cell size and geometry (Alpine and Cloern 1985, Lee et al. 1995). For instance, filamentous species, unlike unicellular species, might form dense clumps (especially in the field), producing nonlinear responses due to the surface cells absorbing a large portion of the light (Hodges 2016a). This is not always the case, however, since other studies involving blooms of filamentous species have shown a linear correlation between cyanobacterial biomass and PC fluorescence (Seppälä et al. 2007). Adding further complexity, pigment ratios within species are affected by light regimes, nutrients, physiological status (e.g. growth stage) and strain dominance, as well as other environmental conditions (Beutler et al. 2002, Mackey et al. 1996, Seppälä et al. 2007, Zamyadi et al. 2016a, Ziegmann et al. 2010). Therefore, although

correlations between PC fluorescence and total cyanobacterial biovolume can be determined, they are usually applicable only to periods when there is a high proportion of cyanobacteria amongst the algal assemblage (e.g. Randolph et al. (2008); Brient et al. (2008), Hodges (2016a), McQuaid et al. (2011); Izydorczyk et al. (2005)). More sophisticated models are needed to account for the many potentially confounding factors that disrupt this relationship.

Studies in the laboratory using traditional spectrophotometric analysis of cultures of *Microcystis aeruginosa* and *Dolichospermum circinale* have found the highest PC cell content during the exponential growth phase (0.004 and 0.018 RFU/cell respectively), with the content reduced by 50% during stationary phase (0.002 and 0.009 RFU/cell respectively) (Chang et al. 2012). Other studies have shown that PC only dominates in the latter stages of growth (Ziegmann et al. 2010). This is consistent with Gregor et al. (2007), who determined that older cells have higher PC fluorescence. Other studies have found that the PC content is lowest when cyanobacteria growth rates are highest (Lee et al. 1995). Hence there is no agreement between studies on when PC content is likely to be higher, suggesting species and methodological differences in PC quotas. In addition, it is important to take into account the relationships between PC extraction efficiency and cyanobacteria concentrations, with previous studies showing how the PC/chl-a ratio often decreases in more diluted samples (Horváth et al. 2013).

Light history is also an important factor affecting pigment content of algae, including cyanobacteria (Erickson et al. 2011). PC content is often higher following exposure to low light (Gregor et al. 2007).

194 This is because in low light environments, cells accumulate
195 photosynthetic antenna pigments to allow them to capture more light,
196 whilst when light is not limiting, the photosynthetic pigments are
197 reduced to prevent excessive excitation energy and cell damage
198 (Seppälä et al. 2007). Under high light intensities, non-
199 photochemical quenching also occurs, leading to lower pigment
200 fluorescence (McBride and Rose in press). These findings are also in
201 line with Beutler et al. (2003), who showed that PC/chl-a ratios were
202 higher at lower light intensities. Hence, in periods of high light
203 intensity, surface cyanobacteria may produce less PC, and thus a
204 fluorometric probe reading could underestimate the cyanobacteria
205 biomass. To avoid the issue, only night readings can be considered,
206 while disregarding daytime measurements (McBride and Rose in
207 press).

208 Despite these findings, there does not seem to be agreement on the
209 effect of light intensity on PC production (Zamyadi et al. 2016a),
210 with several other studies not finding any significant effect related to
211 previous light exposure (Beutler et al. 2002, Brient et al. 2008,
212 Zamyadi et al. 2012). Other studies have measured an increase in the
213 fluorescence signal following exposure to ambient light compared to
214 dimmed ambient light (Zieger 2015). Ambient light was also shown
215 to result in higher PC yields for *Aphanizomenon* sp. (Hodges 2016a),
216 recently renamed to *Chrysosporum* sp. Differences in PC content can
217 also be found between natural and artificial light, suggesting that
218 culture studies in the laboratory may lead to inaccurate results
219 depending on the light intensities and spectra of the light source
220 (Brient et al. 2008). The effect of irradiance seems to also be species-

specific (Hodges 2016a), with the characteristic chl-a/marker pigment ratios affected by the history of light exposure (Everitt et al. 1990).

Nutrients can affect the pigment content of cells. For instance, a lack of nitrogen can lead to PC degradation in cyanobacterial cultures (Rapala 1998). Beutler et al. (2003) found that increasing phosphate and nitrate concentrations increased the PC/chl-a ratio in cyanobacteria. PC can also be released into the water during blooms, then rapidly degraded (Izydorczyk et al. 2005, Steiner et al. 2017). Differences between in-situ and laboratory experiments have also been noted, although some studies have found high similarity (Bastien et al. 2011, Brient et al. 2008, Hodges 2016a, Kong et al. 2017).

There are opportunities to estimate cyanobacteria composition from fluorescence emission spectrum patterns. Henrion et al. (1997) asserted that although the intensity of the fluorescence spectrum can change due to some of the factors outlined above, the pattern of the fluorescence spectrum does not change; thus principal component analysis of excitation-emission matrices (EEM) of surface water can be used to identify the cyanobacterial composition of the sample. Nevertheless, measuring EEM is time consuming (Ziegmann et al. 2010), compared to in-situ fluorescence probes.

3 Instrument limitations and opportunities

As outlined in Millie et al. (2010), the use of in-situ spectrometers and fluorometers has become common practice for water monitoring programs. In vivo fluorometers typically work by using a LED light

source associated with a given volume of water in front of the optical window, and measuring the emitted fluorescence from excitation of phytoplankton cells using a detection filter located perpendicularly to the excitation source. Spectrofluorometric probes, through real-time monitoring, are a useful tool to detect rapid spatiotemporal changes in cyanobacterial biomass, compared to conventional low-frequency sampling (Hamilton et al. 2015, Pobel et al. 2011). In terms of accuracy, in-vivo fluorescence is typically less precise than extracted in vitro fluorescence (McBride and Rose in press), and fluorometric and spectrofluorometric methods are more sensitive than spectrophotometric measures (Millie et al. 2010, Sobiechowska-Sasim et al. 2014). Zamyadi et al. (2016a) have reviewed the available information on in-situ fluorescence instrumentation for cyanobacterial detection.

Fluorometric probes can have a number of advantages compared to other methods, e.g. online real-time lake profiling, and can provide a good estimate of the total cyanobacterial biovolume (Bowling et al. 2016). Despite this, the readings are affected by several interferences, and so far they cannot identify specific species (Zamyadi et al. 2016a). In addition, previous studies found that probe accuracy decreases when determining cell densities, rather than biovolume, due to the presence of heterogeneous mixtures of species with different cell sizes (Bastien et al. 2011, Bowling et al. 2016). Furthermore, the probes require reliable cleaning systems to guarantee consistent levels of accuracy and reliability over time (Brient et al. 2008, NSW Office of Water 2012). Some probes have also been found to be unable to detect fluorescence below certain

lower thresholds (e.g. 2.6 µg/L (Hodges 2016a)). Some sensors also consistently underestimate cyanobacterial biomass (Bastien et al. 2011).

Issues can also arise due to the limitations of the manufacturer's machine calibration, which is typically performed using cell counts of a specific cyanobacterial species, e.g. *Microcystis aeruginosa*. Such species-specific calibration can be of limited use in any waterbody where a mix of species is present (Bowling et al. 2013, Hodges 2016a). One problem that needs to be overcome is the lack of proper laboratory-based PC determination methods, due to low PC extraction efficiency (Seppälä et al. 2007). More recently improved extraction methods have been demonstrated (Horváth et al. 2013). Calibration with different algal species with a range of biovolumes is recommended, as performed in Brient et al. (2008). Alternatively, a mixed assemblage, representative of the monitored site, can be used to relate fluorescence to a site-specific biovolume (McBride and Rose in press).

Several fluorescence devices have been developed in an attempt to characterize the species composition of algal/cyanobacterial samples (see Ziegmann et al. (2010)). Five different wavelengths were used by Beutler et al. (1998) in order to mathematically estimate the composition based on the antenna pigments of the five main groups of algae, with further studies building from those findings (Beutler et al. 2003, Beutler et al. 2002, Gregor and Maršálek 2004). Some studies have found that this approach underestimates cyanobacterial concentrations under certain conditions (Pemberton et al. 2007). Other in-situ multi-wavelength fluorometric devices have been

proposed (Beutler et al. 2002, Richardson et al. 2010). Using several wavelengths covering the excitation spectra for the main pigments present in major groups of phytoplankton can be suitable to monitor natural phytoplankton communities which have variable composition spatially and temporally (Gregor and Maršálek 2004); however, if focusing on cyanobacteria only, production of PC and other pigments is affected by several factors, such as nutrients and light, so these spectra could be expected to differ with spatial-temporal variations (Beutler et al. 2002). This implies that more work is necessary to compensate the readings for these factors and make such probes more reliable. Additionally, in-situ instrumentation to simultaneously measure the parameters causing the variability (e.g. water temperature, turbidity, etc.) would be required to correct the readings.

Fluorescence synchronous scans (SyncScans) were used by Ziegmann et al. (2010) as an alternative, faster, on-line method, to determine whether the age of cyanobacteria affects the fluorescence spectrum. The downside of using specific wavelengths is that certain cyanobacterial pigments have maxima at slightly varying wavelengths, often varying with the age of the cells (Ziegmann et al. 2010). There is, in fact, evidence that some signal overlaps can occur at wavelengths > 660 nm (Seppälä et al. 2007). Furthermore the peak wavelength of the PC emission spectrum seems to vary slightly based on species (Beutler et al. 2003, Beutler et al. 2002, Gregor et al. 2007, Seppälä et al. 2007). This is because the spectroscopic properties change based on the cellular content and abundance of phycobiliproteins, and the number and type of phycobilin chromophores (Seppälä et al. 2007).

Some probes are more effective at dealing with the issue of pigment wavelength variability than others, due to a broad wavelength bypass (e.g. 640-680 nm). The downside of this is that it can lead to greater incidence of interferences and false positives, as described below (Section 4). The selection of the optimal optical window is therefore challenging. Spectral studies with pure cultures or representative cyanobacterial species should be performed before defining the optimal optical window for a certain location and/or purchasing a particular commercial probe (Seppälä et al. 2007). This approach was used by Zhou et al. (2016) to determine the optimal wavelength for a particular cyanobacterial research application.

Flow cytometers can provide estimates of cell counts as well as morphology of different algal species, and some of these can be used in-situ (Dennis et al. 2011, Pomati et al. 2011, Pomati et al. 2013). Flow cytometers, combined with fluorescence excitation spectra, have been considered as a potential means for rapidly distinguishing between different algal groups (Hilton et al. 1989). Additionally, recently microflow cytometers have been developed (Hashemi et al. 2011), as well as prototype devices combining a microflow cytometer and a multi-channel fluorometer, to allow for distinction between cyanobacteria and other algal phyla by exciting them at different wavelengths (Zieger 2015). Another opportunity consists of modifying current spectrofluorometers for flow-through applications, but they are rarely available or used, as they are currently not sufficiently sensitive, nor robust or cheap (Seppälä et al. 2007).

4 Fluorescence measurement limitations

Given that chl-a and PC emission wavelengths are similar (685 and 650 nm respectively) for probes with a broad wavelength bypass (e.g. ± 20 nm), this might cause false detections of cyanobacterial cells (see Zamyadi et al. (2016a)). This has been deemed to be the main source of interference by some studies (Zamyadi et al. 2016a), but negligible by others (Bowling et al. 2013). Narrower bandwidths would be required; this can be achieved by using higher quality sensor materials for the sensors' glass and filters, which however are associated with higher costs (McBride and Rose in press). A simultaneous monitoring of eukaryotic algal fluorescence could help deal with the spectra overlap (Gregor et al. 2007), and the development of ad-hoc instrumentation with different excitation wavelengths could help to account for interference due to eukaryotic algae (Asai et al. 2001).

Studies using a submersible fluorescence PC probe have also shown that turbidity >50 NTU can make fluorescence in-situ PC measurements ineffective (Bowling et al. 2013). Turbidity is typically due to a combination of algae and non-algal suspended particles, hence a proper calibration model accounting for the interaction between the two variables is needed (Bowling et al. 2013). The causal agent of the turbidity is also to understand as different sources have different levels of interference with fluorescence readings. For example, in a previous study where sieved soil of a specific different grain size was gradually added to a cyanobacterial culture, small particles (≤ 0.1 mm) decreased the fluorescence reading by 12.3% at a concentration of 0.1 g/L, while large particles (0.3 mm) reduced the reading by less than one-half, to

5.9%, at the same concentration (Brient et al. 2008). In general, simultaneous measurement of PC, chl-a and turbidity is recommended in order to account for these errors (Gregor and Maršálek 2004) and allow for the development of real-time compensation models.

Cyanobacterial colonies can emit less fluorescence per cell than solitary cells because the probe excitation beam does not access cells in the inner regions of the colony, and because of scattering of light (Bowling et al. 2013, Gregor and Maršálek 2004, Hodges 2016a). Large colonies might also lead to noisy fluorescence signals and broader confidence levels (Seppälä et al. 2007). This was demonstrated with *Microcystis* colonies (Chang et al. 2012), although in other studies (McQuaid et al. 2011) it was stated that *in vivo* PC fluorescence might overestimate cyanobacterial abundance when biovolumes are high. Some studies have found that a better correlation between PC fluorescence and biomass could be achieved at higher rather than lower cyanobacterial concentrations (Gregor et al. 2007). In order to avoid saturation of the PC fluorescence signal during blooms, researchers have experimented with technological modifications to the probe; for instance Brient et al. (2008) have fitted a shutter to reduce the probe's emitted signal, and thus increased the fluorescence signal by six times without affecting linearity. However, there has been little work to specifically examine the effect of colonial cyanobacterial on fluorescence measurements (Hodges 2016a).

Species morphology may also lead to a heterogeneous distribution of cells (Hodges 2016a) causing inaccuracies in the results. Ultrasonic

devices, coupled with fluorometers, have been proposed as a mechanism to disaggregate colonies, and bring about a relatively homogeneous cell distribution (Bastien et al. 2011). Additionally, PC can be produced from lysis of cyanobacterial cells, or from cyanobacterial picoplankton not identified by microscopy (Brient et al. 2008). This may explain why fluorescence probe readings may be better correlated with species with larger cell sizes compared with smaller ones (Bowling et al. 2016). Some tools, such as microflow cytometers, are able to detect and characterize picoplankton (Hashemi et al. 2011) and thus could potentially be used in combination with conventional PC fluorescence probes to correct the readings of the latter.

A number of correction techniques have been proposed for certain fluorescence PC probes (Zamyadi et al. (2016a). They include a correction for biases due to chl-a resulting from other phytoplankton, and turbidity when estimating the abundance of a specific species (*Microcystis aeruginosa*) (Chang et al. 2012, Zamyadi et al. 2012). Additionally, corrections can be made for variations in cyanobacterial biovolume, cellular agglomeration and particle size, assuming spherical colonies and uniformly distributed cells (Chang et al. 2012). By applying these models together the error was reduced by over 90% (Zamyadi et al. 2016a). In addition, the sensor manuals for the same probe stated that (1) 1 ng L^{-1} of chl-a causes an interference of 77 cells mL^{-1} of falsely detected cyanobacteria, and (2) for each NTU of turbidity (i.e. non-fluorescent particles), the related scattering effect would be equivalent to 21 cells (Zamyadi et al. 2016a). However, this very precise specification appears to

435 oversimplify the quantification of such interferences, as these, for
436 instance, would be likely dependent on cyanobacteria species and the
437 composition relating to agents generating turbidity.

438 Other sources of interference include water temperature, with warmer
439 waters reducing fluorescence (Hodges 2016a), and the presence of
440 benthic algae (Zamyadi et al. 2012) which are reportedly difficult to
441 quantify (Watson et al. 2016). There are, however, certain probes
442 suitable for detection of benthic algae (Brient et al. 2008). In terms of
443 boundary effects, it is also important to account for a potential
444 decrease in signal next to a boundary, i.e., bottom or sides of bottles.
445 This effect has been identified in previous work (Brient et al. 2008),
446 although in other studies, sediments did not create interferences
447 (Hodges 2016a). Coloured dissolved organic matter (cDOM) with
448 similar fluorescence properties to PC and chl-a can also interfere
449 with analyses (Millie et al. 2010), greatly increasing the light
450 attenuation in the water column (Kostoglidis et al. 2005) and thus in
451 turn affecting pigment yields. The presence of humic substances can
452 also create a fluorescence output causing an over-estimation of chl-a,
453 especially when chl-a concentrations are low (Gregor and Maršálek
454 2004). Other studies have reported that interferences with DOM lead
455 to severe quenching of phycobiliprotein fluorescence (Korak et al.
456 2015), requiring modelling of nonlinear relationships between
457 concentrations and fluorescence. Wang et al. (2015) demonstrated
458 that some DOM components (specifically, protein-like compounds)
459 can be quenched by other DOM components (humic-like
460 compounds), making it difficult to quantify and account for this
461 effect. Some of the existing multi-wavelength fluorescence probes

462 have, however, incorporated diodes emitting light around the 370 nm
463 value, in order to excite, and automatically adjust for, the presence of
464 DOM (Gregor and Maršálek 2004).

465 Irradiance affects the so called fluorescence quenching and is
466 associated with a reduction of the maximum fluorescence yield. For
467 instance, the ratio between fluorescence and chl-a could present a
468 variation as high as 10-fold when in-vivo fluorescence measurements
469 are normalised to concentrations of extracted chl-a (as explained in
470 Sackmann et al. (2008) who based the variation estimate on previous
471 studies). In in-vivo experiments, fluorescence variations can occur
472 due to different mechanisms, complicating a straightforward
473 assessment (Schreiber et al. 1995). Two types of fluorescence
474 quenching exist, being (1) photochemical, and (2) non-photochemical
475 fluorescence (Sackmann et al. 2008, Schreiber et al. 1995). The latter
476 represents, in general terms, the decrease in fluorescence quantum
477 yield (Sackmann et al. 2008). The systematic reduction of
478 fluorescence signal vs chl-a ratio with high solar radiation can be
479 generally termed as daytime fluorescence quenching; such
480 fluorescence signal reduction can be as high as 80% in surface waters
481 in summer and can be present at depths greater than 50 m in coastal
482 waters (Sackmann et al. 2008). Interestingly, Sackmann et al. (2008)
483 found correlations between quenching and solar radiation very
484 similar to relationships from previous research conducted elsewhere.
485 This means that there would be potential for the development of a
486 "universal" compensation model to account for this phenomenon.

487 **5 Estimating the proportion of different cyanobacteria**
488 **species**

Techniques and algorithms need to be developed to estimate the composition and quantity of different groups of algae, or more specifically, between different cyanobacteria species in field samples (Millie et al. 2010) and as part of sediment cores (Kpodonu et al. 2016), because PC sensors per se cannot provide insight into cyanobacterial species differentiation or toxin content (Bastien et al. 2011, Izydorczyk et al. 2005). Algorithms, such as three-way principal component analysis, have previously been used to determine algal groups (Henrion et al. 1997), where five mono-species cultures were used, representative of the different main groups of phytoplankton (i.e. chrysophytes, cryptophytes, chlorophytes, diatoms and cyanobacteria). However, the analysis was based on excitation-emission matrices from data collected with laboratory instrumentation. Two decades ago, the computer program CHEMTAX was developed to estimate the relative abundance of certain algal groups based on measurement of chl-a and other carotenoid pigments using wavelengths in the visible spectrum, high performance liquid chromatography (HPLC) (Mackey et al. 1996) and factor analysis. This was re-applied in later studies (Pinckney et al. 2001). This method is more accurate than field-based methods, but time consuming, and relies on good-quality data and consistent ratios between pigments. CHEMTAX was based on previous work where phytoplankton classes were estimated from specific ratios between chl-a and other pigments, through multiple regression analysis (Everitt et al. 1990, Gieskes et al. 1988). This continues to be a widely used method throughout the world but cannot be applied to data collected remotely from sensors, as it relies on HPLC and spectrophotometric detectors.

More recently, different algal phyla have been characterised spectrally, and a prototype monitoring tool, working at different wavelengths (initially, 12 LEDs, to excite all of the most important algal pigments) was developed. It is able to distinguish the presence of cyanobacteria compared to different algal phyla, following principal component analysis (Zieger 2015). In this case, several pigments were considered in order to characterize each algal phyla, given the ability of the sensor to work at several wavelengths and excite each specific pigment. The use of this technique requires the deployment of monochromatic light to avoid the excitation of multiple pigments at the same time (Zamyadi et al. 2016a, Zieger 2015), and as a consequence this can lead to the underestimation of pigment content, especially due to the potential for variations in fluorescence peaks between different phyla or even species, as mentioned previously.

If the focus is only on cyanobacteria, a simpler approach can be developed that relies only on the PC reading, through a risk assessment/scenario analysis approach. In this case the PC reading is transformed into hypothetical cell counts of different cyanobacterial species, based on PC cell quota estimates found in the literature (van der Linden et al. 2017). Despite the fact that the method does not directly infer the species present in the sample, it deploys a "worst-case scenario" approach, providing an early warning trigger of the need for more intensive monitoring. The PC cell quota estimate still needs to take into account the background environmental conditions since, as described above, these affect pigment content.

6 Predicting taste and odour, and toxicity risk

Although the main purpose of fluorescence probe readings is to trigger monitoring and management actions based on estimated cell counts, there is potential to use them also for the estimation of taste and odour (T&O), or toxin production risk (Zamyadi et al. 2016a). For example, if the likelihood of the presence of different cyanobacteria species could be estimated, models could then be developed to link the number of cells with the toxin (or T&O) cell quota values (Zamyadi et al. 2016a). In their review, Watson et al. (2016) provided a list of the estimated geosmin and 2-Methylisoborneol (MIB) production per cell of different cyanobacterial species, also citing several studies outlining how such production is itself affected by factors such as light, temperature or nutrients (see also Rohrlack et al. (2008); Li et al. (2012), Su et al. (2015)). Geosmin production has also been linked both positively and negatively to chl-a production, as outlined in Watson et al. (2016). It is known that geosmin is produced mostly by Nostocales and Oscillatoriales species, while MIB is produced by some Oscillatoriales (Chapman 2010), although less than 3% of known cyanobacteria species have been confirmed to be able to produce MIB and/or geosmin (American Water Works Association 2011).It should also be acknowledged that T&O is also produced by a range of bacterial species, which can confound estimates based on fluorescence (e.g. Jørgensen et al. (2016)).

Previous studies linked the fluorescence intensity recorded by an in-situ PC probe to the T&O compound levels (especially MIB) caused by *Pseudanabaena* (Zamyadi et al. 2016b). Similarly, other studies linked the production of geosmin and MIB to specific cyanobacteria

571 species and noted a correlation between T&O peaks and spikes in PC
572 signal as measured by an in-situ fluorescence probe for total algae
573 (Bertone and O'Halloran 2016). These correlations would change
574 remarkably in cases where the dominant species changes or where
575 there is a mixture of different species with different T&O production
576 yields. Thus the development of a T&O predictive model based on
577 fluorescence readings should rely on a cyanobacterial species
578 prediction model, and on T&O concentration data for different
579 species; unless site-specific models are developed that are predicated
580 upon the assumption that the same dominant cyanobacterial species
581 are present at any given point in time.

582 In terms of cyanobacterial toxin content, Ziegmann et al. (2010)
583 linked the amount of toxins produced and released by a laboratory
584 culture of *Microcystis aeruginosa* to different stages of growth using
585 different maxima of a fluorescence excitation-emission matrix. They
586 found that a signal at wavelengths of 315 nm (excitation) and 396 nm
587 (emission), i.e. presumably protein-like substances, may be a useful
588 indicator of cyanobacterial toxin levels. It may be difficult, however,
589 to correctly measure the protein-like DOM due to quenching effects
590 by other DOM components (Wang et al. 2015). Other studies have
591 not been able to successfully correlate concentrations of cyanotoxins
592 and cyanobacterial biovolume to in vivo PC fluorescence (Bastien et
593 al. 2011). Additional complications in the development of a toxin
594 prediction model for a reservoir are that: (1) toxin production is
595 related to the species and in particular to the cell size (Hawkins et al.
596 2001); (2) strains vary in their toxin cell quotas, e.g.
597 *Cylindrospermopsis raciborskii* (Willis et al. 2016); (3) nutrients

598 such as nitrogen and phosphorus differentially affect the growth of
599 toxic and non-toxic strains of cyanobacteria (Vézie et al. 2002); and
600 (4) the breakdown rate of toxins can be influenced by depth-specific
601 factors such as light, water temperature, salinity, nutrients or oxygen
602 (Everson et al. 2009, Everson et al. 2011), and thus changes in depth
603 may lead to decoupling between cyanobacteria concentrations and
604 toxin concentrations.

605 Both toxins and T&O compounds are components of algogenic
606 organic matter (AOM). AOM is typically released by certain species
607 of cyanobacteria and other algae as a by-product of photosynthesis
608 (Pivokonsky et al. 2016), and its composition depends in turn on
609 growth phase, species, age, and nutrient status (Henderson et al.
610 2008, Kong et al. 2014, Leloup et al. 2013). In general, the presence
611 of AOM is difficult to quantify as it is typically much lower than the
612 background DOM (Wert et al. 2014). Hence it has been suggested
613 that variations in fluorescent DOM readings could help in
614 understanding AOM release (Korak et al. 2015), provided the
615 remaining DOM amount does not change.

616 **7 Discussion**

617 Figure 1 schematically summarises the findings of this review by
618 representing how both the measured pigment readings, using either
619 fluorescent probes or measured pigment amounts, are affected by a
620 number of environmental variables that either cause interference or
621 change the yield. Interestingly, cyanobacteria concentration and
622 species greatly affect both the actual pigment concentrations, and the
623 ability to correctly monitor them with fluorescent probes.

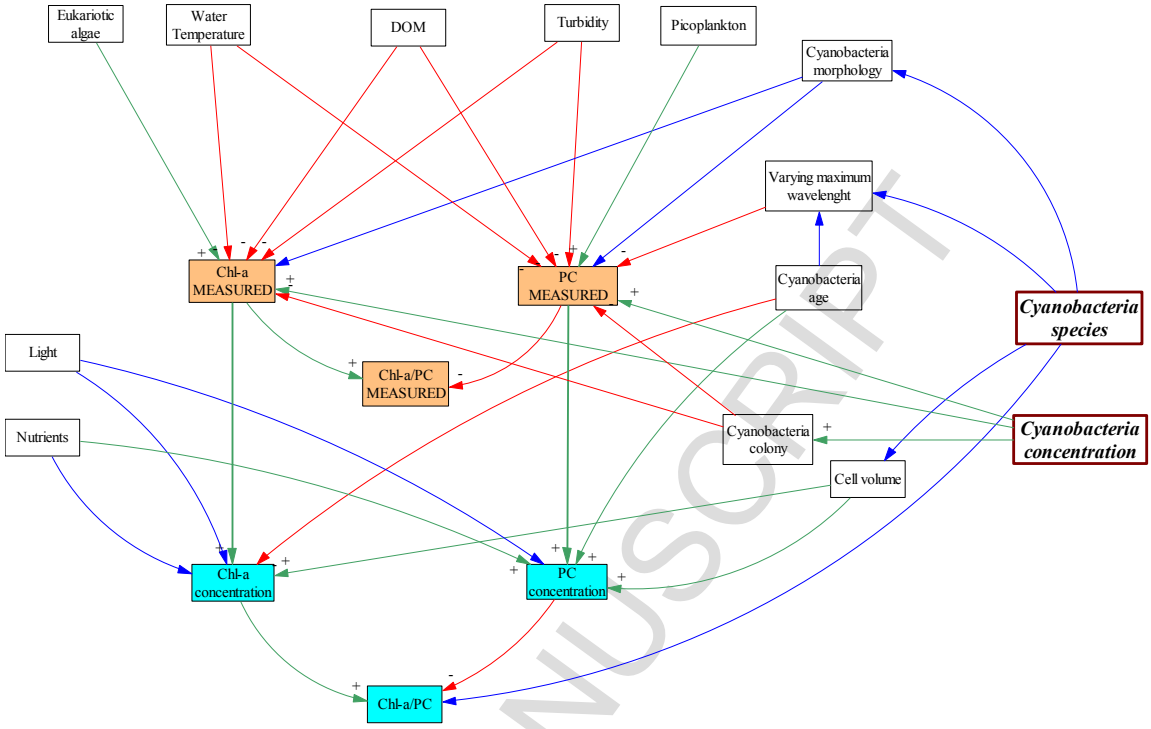


Figure 1 – A summary of parameters affecting chl-a and PC concentrations and their measurements. Green arrows indicate that the input positively affects the output (i.e., an increase in the input value implies an increase in the output value), red arrows imply that the input negatively affects the output (i.e., an increase in the input value implies a decrease in the output value), and blue arrows represent an unspecified link (i.e., there is insufficient, or contrasting, evidence in the literature regarding the way the input affects the output).

In-situ PC fluorometry has been recommended as a cyanobacterial bloom management tool in certain regions of the world, e.g. Australian locations (NSW Office of Water 2012). Importantly, several factors can affect the accuracy and reliability of in-situ

640 fluorescence probes, and a large body of research work has been
641 conducted globally to quantify the effects of a number of these
642 parameters. However, no study has yet been conducted which
643 simultaneously quantifies the effects of most, if not all, of these
644 variables to enable to correct measured chlorophyll and PC
645 fluorescence values, and in turn quantify abundance and species of
646 cyanobacteria. If such a comprehensive dataset can be developed,
647 there would be potential to develop a predictive model providing the
648 probability of having different cyanobacteria species in a reservoir,
649 based on corrected readings of commercially available in-situ
650 fluorescent PC and chl-a probes. To increase accuracy, such a model
651 could also rely on other regularly monitored data (e.g. water
652 temperature, nutrients, time of the year) and site-specific historical
653 data of cyanobacterial species prevalence under different
654 environmental conditions. A large amount of research work has been
655 undertaken around the world to predict cyanobacteria blooms (which
656 is out of the scope of this review – for information see e.g. Oliver et
657 al. (2012)), and thus this knowledge and historical site-specific data
658 could be incorporated in the proposed model.

659 In recent decades, artificial intelligence has been used in several
660 applications, especially in the water industry, to solve complex
661 problems. Modelling techniques, such as artificial neural networks,
662 have been extensively applied and have enabled the optimization of
663 water resource management (Maier and Dandy 2000), while a variety
664 of data-driven or hybrid models have been developed to leverage
665 remote sensors and historical data by achieving water treatment

666 operation optimisation and cost savings (Bertone et al. 2015, Bertone
667 et al. 2016b).

668 Pattern recognition is another technique which would also potentially
669 suit the proposed research goal, as it could enable the decomposition
670 of a "signal" (i.e. total chl-a and PC measurements of a field sample)
671 into chl-a and PC contributions from different species. An early
672 application of pattern recognition for classification of different
673 phytoplankton species (including cyanobacteria) is represented by the
674 work of Embleton et al. (2003), with the models able to classify the
675 four different predominant species of a Northern Ireland lake, based
676 on microscopy images of the samples. A very similar modelling
677 attempt was performed later by Mosleh et al. (2012). Similarly,
678 CHEMTAX represents an early attempt at algal class level
679 classification; but it does not rely on in-situ fluorescence probes or
680 fully account for interferences. Based on this review, it is argued that
681 artificial intelligence approaches combined with a well-designed set
682 of experiments and a large amount of site-specific historical data,
683 would allow the development of a similar model, but that may further
684 allow classification of cyanobacteria species based on in-situ data
685 only (Figure 2).

686

687 Bayesian Networks (Fenton and Neil 2008) are also commonly used
688 for risk assessment and analysis and have been also widely deployed
689 in the water resource management area (Bertone et al. 2016a,
690 Castelletti and Soncini-Sessa 2007). There are applications in the
691 cyanobacterial bloom management field (Johnson et al. 2010).

692 Bayesian Networks represent a modelling candidate for this research
693 problem due to their ability to deal with missing data and uncertainty,
694 through the use of conditional probabilities and the integration of
695 experts' opinion into the network (Chen and Pollino 2012, Uusitalo
696 2007). Outputs from other models that can provide insights on the
697 mechanisms causing and governing blooms (e.g. Wallace and
698 Hamilton (2000)) can also be incorporated. Although Bayesian
699 Networks have been applied to determine the probability of
700 cyanobacterial blooms (Hamilton et al. 2007, Rigosi et al. 2015),
701 they have not been developed in order to predict the likelihood of the
702 presence of different, distinct algal groups or cyanobacteria species
703 based on sensor data.

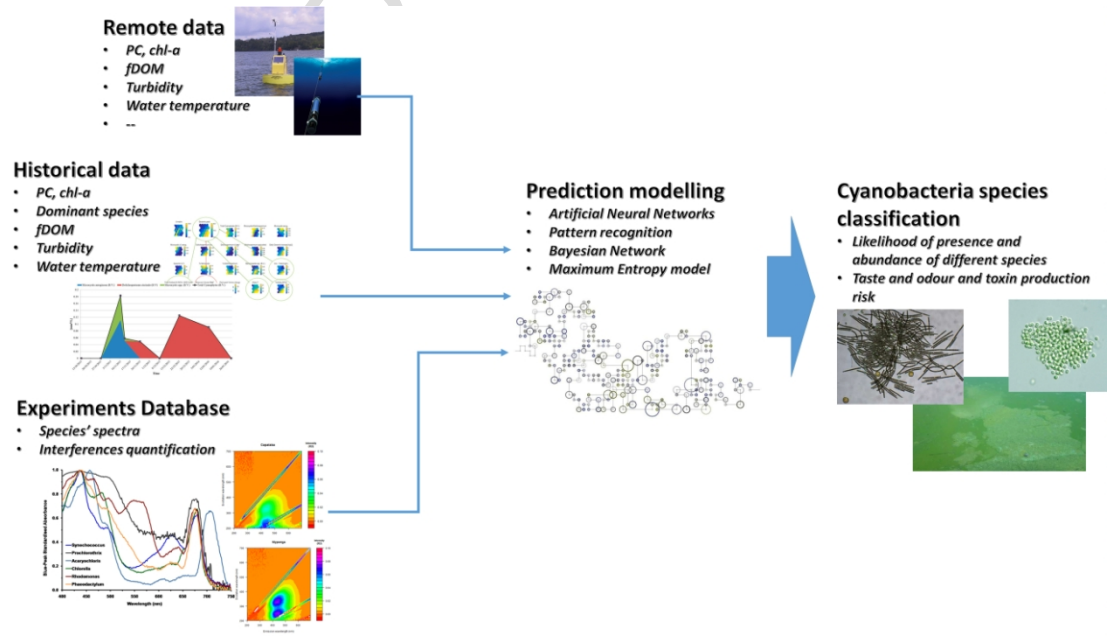
704 The entropy theory, specifically maximum entropy modelling, could
705 also be used for this purpose. In simple terms, the principle of
706 maximum entropy states that, among a number of potential
707 combinations of modelling outputs satisfying all a system's
708 constraints, the best solution is the one maximising the Shannon
709 entropy (Jaynes 1957). In terms of cyanobacterial species
710 classification, it can be seen that a number of solutions (e.g.
711 probability and numbers of different species) would exist that satisfy
712 the same set of constraints (e.g. sum of species-specific PC signals
713 must equal the total PC signal as per probe reading). Hence,
714 maximum entropy modelling would facilitate the identification of the
715 true solution. Entropy theory has been applied to several
716 environmental and water resource problems (Singh 2000). For
717 instance, Tanyimboh and Sheahan (2002) used maximum entropy
718 principles to identify the best water distribution system design in

719 order to optimise the balance of cost and reliability. Another example
720 is given by Lee and Wentz (2008), who used a Bayesian Maximum
721 Entropy approach to assimilate soft data and better predict spatial
722 water use variability. However, there is limited evidence of
723 applications in the water quality field, and in particular for
724 cyanobacteria prediction or management.

725 In conclusion, all the proposed modelling approaches are valuable
726 options in achieving the goal of data-driven cyanobacterial species
727 classification prediction; the optimal model can be selected based on
728 features of available data, or from a direct comparison of prediction
729 accuracy over a test set of data. Regardless of the model choice it is
730 evident that modelling options exist to potentially achieve such
731 research goals.

732

733



734

735 Figure 2 – Graphical conceptual representation of a cyanobacterial
736 species classification model

8 Conclusions

In this review paper we identified both impediments and opportunities for using online remote fluorescence probes for accurate and reliable monitoring of cyanobacteria in reservoirs. The technology has become cheaper and more accessible over time, but we have highlighted a number of environmental and technological impediments affecting the reliability of probe readings. We argue that unless these interferences and variability in pigment production are accounted for in a comprehensive calibration model, many of these probes provide unreliable information about cyanobacterial densities or biovolume, and little or no information about species.

On the other hand, research has shown that it is possible, in most cases, to quantify how those variables and processes affect the estimation of cyanobacterial species and abundance. Starting from a specific location, it is possible to conduct experiments to calibrate the readings to most of the conditions at that location. In addition, there is potential to use such experimental outputs, combined with literature data, as inputs to a complex predictive model. This would potentially combine pattern recognition algorithms and Bayesian Networks, to provide the likelihood of occurrence and biomass estimates of different cyanobacterial species, based on real-time fluorescence probe readings and other remotely collected data. Based on the predicted species, the risk of toxins, or taste and odour compound concentrations could also be estimated, thus assisting water resource managers and operators to be more effective and proactively manage cyanobacteria-related issues related to recreational and drinking water.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Allan, M.G., Hamilton, D.P., Hicks, B.J. and Brabyn, L. (2011) Landsat remote sensing of chlorophyll a concentrations in central North Island lakes of New Zealand. International Journal of Remote Sensing 32(7), 2037-2055.

Alpine, A.E. and Cloern, J.E. (1985) Differences in in vivo fluorescence yield between three phytoplankton size classes. Journal of Plankton Research 7(3), 381-390.

American Water Works Association (2011) Algae: source to treatment, American Water Works Association.

Asai, R., Horiguchi, Y., Yoshida, A., McNiven, S., Tahira, P., Ikebukuro, K., Uchiyama, S., Masuda, Y. and Karube, I. (2001) Detection of phycobilin pigments and their seasonal change in Lake Kasumigaura using a sensitive in situ fluorometric sensor. Analytical Letters 34(14), 2521-2533.

Bastien, C., Cardin, R., Veilleux, É., Deblois, C., Warren, A. and Laurion, I. (2011) Performance evaluation of phycocyanin probes for the monitoring of cyanobacteria. Journal of Environmental Monitoring 13(1), 110-118.

Bertone, E. and O'Halloran, K. (2016) Analysis and Modelling of Taste and Odour Events in a Shallow Subtropical Reservoir. Environments 3(3), 22.

Bertone, E., Sahin, O., Richards, R. and Roiko, A. (2016a) Extreme events, water quality and health: A participatory Bayesian risk assessment tool for managers of reservoirs. Journal of Cleaner Production 135, 657-667.

Bertone, E., Stewart, R.A., Zhang, H., Bartkow, M. and Hacker, C. (2015) An autonomous decision support system for manganese forecasting in subtropical water reservoirs. Environmental Modelling & Software 73, 133-147.

Bertone, E., Stewart, R.A., Zhang, H. and O'Halloran, K. (2016b) Hybrid water treatment cost prediction model for raw water intake optimization. Environmental Modelling & Software 75, 230-242.

Beutler, M., Wiltshire, K., Meyer, B. and Moldaenke, C. (1998) Differentiation of spectral groups of algae with computer-based analysis of fluorescence excitation spectra. Vom Wasser 91, 61-74.

Beutler, M., Wiltshire, K.H., Arp, M., Kruse, J., Reineke, C., Moldaenke, C. and Hansen, U.P. (2003) A reduced model of the fluorescence from the cyanobacterial photosynthetic apparatus designed for the in situ detection of cyanobacteria. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1604(1), 33-46.

808 Beutler, M., Wiltshire, K.H., Meyer, B., Moldaenke, C., Lüring, C.,
809 Meyerhöfer, M., Hansen, U.-P. and Dau, H. (2002) A fluorometric
810 method for the differentiation of algal populations in vivo and in situ.
811 Photosynthesis Research 72(1), 39-53.

812 Bowling, L., Ryan, D., Holliday, J. and Honeyman, G. (2013)
813 Evaluation of in situ fluorometry to determine cyanobacterial
814 abundance in the Murray and Lower Darling Rivers, Australia. River
815 Research and Applications 29(8), 1059-1071.

816 Bowling, L.C., Zamyadi, A. and Henderson, R.K. (2016) Assessment
817 of in situ fluorometry to measure cyanobacterial presence in water
818 bodies with diverse cyanobacterial populations. Water Research 105,
819 22-33.

820 Brient, L., Lengronne, M., Bertrand, E., Rolland, D., Sipel, A.,
821 Steinmann, D., Baudin, I., Legeas, M., Le Rouzic, B. and Bormans,
822 M. (2008) A phycocyanin probe as a tool for monitoring
823 cyanobacteria in freshwater bodies. Journal of Environmental
824 Monitoring 10(2), 248-255.

825 Bryant, D.A. (1982) Phycoerythrocyanin and phycoerythrin:
826 properties and occurrence in cyanobacteria. Microbiology 128(4),
827 835-844.

828 Bryant, D.A. (2006) The molecular biology of cyanobacteria,
829 Springer Science & Business Media.

830 Carmichael, W.W. (2001) Health effects of toxin-producing
831 cyanobacteria:"The CyanoHABs". Human and ecological risk
832 assessment: An International Journal 7(5), 1393-1407.

833 Castelletti, A. and Soncini-Sessa, R. (2007) Bayesian Networks and
834 participatory modelling in water resource management.
835 Environmental Modelling & Software 22(8), 1075-1088.

836 Chang, D.-W., Hobson, P., Burch, M. and Lin, T.-F. (2012)
837 Measurement of cyanobacteria using in-vivo fluoroscopy – Effect of
838 cyanobacterial species, pigments, and colonies. Water Research
839 46(16), 5037-5048.

840 Chapman, A.D. (2010) Algae: Source to Treatment. American Water
841 Works Association (ed), pp. 125-142.

842 Chen, S. and Pollino, C. (2012) Good practice in Bayesian network
843 modelling. Environmental Modelling & Software 37, 134-145.

844 DELWP Biovolume Calculator Sheet.
845 https://www.water.vic.gov.au/__data/assets/excel_doc/0027/65592/B
846 IOVOLUME-CALCULATOR.XLSX, Environmental Health Unit,
847 Department of Human Services, Victoria.

848 Dennis, M., Landman, M., Wood, S.A. and Hamilton, D. (2011)
849 Application of flow cytometry for examining phytoplankton
850 succession in two eutrophic lakes. Water Science and Technology
851 64(4), 999-1008.

852 Embleton, K., Gibson, C. and Heaney, S. (2003) Automated counting
853 of phytoplankton by pattern recognition: a comparison with a manual
854 counting method. Journal of Plankton Research 25(6), 669-681.

Erickson, J.S., Hashemi, N., Sullivan, J.M., Weidemann, A.D. and Ligler, F.S. (2011) In situ phytoplankton analysis: there's plenty of room at the bottom. Analytical chemistry 84(2), 839-850.

Everitt, D., Wright, S., Volkman, J., Thomas, D. and Lindstrom, E. (1990) Phytoplankton community compositions in the western equatorial Pacific determined from chlorophyll and carotenoid pigment distributions. Deep Sea Research Part A. Oceanographic Research Papers 37(6), 975-997.

Everson, S., Fabbro, L., Kinnear, S., Eaglesham, G. and Wright, P. (2009) Distribution of the cyanobacterial toxins cylindrospermopsin and deoxycylindrospermopsin in a stratified lake in north-eastern New South Wales, Australia. Marine and Freshwater Research 60(1), 25-33.

Everson, S., Fabbro, L., Kinnear, S. and Wright, P. (2011) Extreme differences in akinete, heterocyte and cylindrospermopsin concentrations with depth in a successive bloom involving Aphanizomenon ovalisporum (Forti) and Cylindrospermopsis raciborskii (Woloszynska) Seenaya and Subba Raju. Harmful Algae 10(3), 265-276.

Fenton, N. and Neil, M. (2008) Risk assessment and decision analysis with Bayesian Networks, CRC Press, Taylor & Francis Group, Boca Raton, FL, USA.

Gieskes, W., Kraay, G., Nontji, A., Setiapermana, D. and Sutomo (1988) Monsoonal alternation of a mixed and a layered structure in the phytoplankton of the euphotic zone of the Banda Sea (Indonesia): a mathematical analysis of algal pigment fingerprints. Netherlands Journal of Sea Research 22(2), 123-137.

Gregor, J. and Maršálek, B. (2004) Freshwater phytoplankton quantification by chlorophyll a: a comparative study of in vitro, in vivo and in situ methods. Water Research 38(3), 517-522.

Gregor, J., Maršálek, B. and Šípková, H. (2007) Detection and estimation of potentially toxic cyanobacteria in raw water at the drinking water treatment plant by in vivo fluorescence method. Water Research 41(1), 228-234.

Hamilton, D.P., Carey, C.C., Arvola, L., Arzberger, P., Brewer, C., Cole, J.J., Gaiser, E., Hanson, P.C., Ibelings, B.W. and Jennings, E. (2015) A Global Lake Ecological Observatory Network (GLEON) for synthesising high-frequency sensor data for validation of deterministic ecological models. Inland Waters 5(1), 49-56.

Hamilton, G.S., Fielding, F., Chiffings, A.W., Hart, B.T., Johnstone, R.W. and Mengersen, K. (2007) Investigating the use of a Bayesian network to model the risk of Lyngbya majuscula bloom initiation in Deception Bay, Queensland, Australia. Human and Ecological Risk Assessment 13(6), 1271-1287.

Hashemi, N., Erickson, J.S., Golden, J.P., Jackson, K.M. and Ligler, F.S. (2011) Microflow Cytometer for optical analysis of phytoplankton. Biosensors and Bioelectronics 26(11), 4263-4269.

Hawkins, P.R., Putt, E., Falconer, I. and Humpage, A. (2001) Phenotypical variation in a toxic strain of the phytoplankter,

904 Cylindrospermopsis raciborskii (Nostocales, Cyanophyceae) during
905 batch culture. Environmental Toxicology 16(6), 460-467.

906 Henderson, R.K., Baker, A., Parsons, S.A. and Jefferson, B. (2008)
907 Characterisation of algogenic organic matter extracted from
908 cyanobacteria, green algae and diatoms. Water Research 42(13),
909 3435-3445.

910 Henrion, R., Henrion, G., Böhme, M. and Behrendt, H. (1997) Three-
911 way principal components analysis for fluorescence spectroscopic
912 classification of algae species. Fresenius' Journal of Analytical
913 Chemistry 357(5), 522-526.

914 Hilton, J., Rigg, E. and Jaworski, G. (1989) Algal identification using
915 in vivo fluorescence spectra. Journal of Plankton Research 11(1), 65-
916 74.

917 Hodges, C. (2016a) A validation study of phycocyanin sensors for
918 monitoring cyanobacteria in cultures and field samples. Masters of
919 Science (Research) thesis, The University of Waikato.

920 Hodges, C.M. (2016b) A validation study of phycocyanin sensors for
921 monitoring cyanobacteria in cultures and field samples, University of
922 Waikato.

923 Hodges, C.M., Wood, S.A., Puddick, J., McBride, C.G. and
924 Hamilton, D.P. (2017) Sensor manufacturer, temperature, and
925 cyanobacteria morphology affect phycocyanin fluorescence
926 measurements. Environmental Science and Pollution Research, 1-10.

927 Horváth, H., Kovács, A.W., Riddick, C. and Présing, M. (2013)
928 Extraction methods for phycocyanin determination in freshwater
929 filamentous cyanobacteria and their application in a shallow lake.
930 European journal of phycology 48(3), 278-286.

931 Izydorczyk, K., Tarczynska, M., Jurczak, T., Mrowczynski, J. and
932 Zalewski, M. (2005) Measurement of phycocyanin fluorescenceas an
933 online early warning system for cyanobacteria in reservoir intake
934 water. Environmental Toxicology 20(4), 425-430.

935 Jaynes, E.T. (1957) Information theory and statistical mechanics.
936 Physical review 106(4), 620.

937 Johnson, S., Fielding, F., Hamilton, G. and Mengersen, K. (2010) An
938 Integrated Bayesian Network approach to Lyngbya majuscula bloom
939 initiation. Marine Environmental Research 69(1), 27-37.

940 Jørgensen, N.O., Podduturi, R. and Burford, M.A. (2016) Relations
941 between abundance of potential geosmin-and 2-MIB-producing
942 organisms and concentrations of these compounds in water from
943 three Australian reservoirs. Journal of Water Supply: Research and
944 Technology-Aqua 65(6), 504-513.

945 Kong, Y., Lou, I., Zhang, Y., Lou, C.U. and Mok, K.M. (2017)
946 Advances in Monitoring and Modelling Algal Blooms in Freshwater
947 Reservoirs, pp. 45-67, Springer.

948 Kong, Y., Zhu, L., Zou, P., Qi, J., Yang, Q., Song, L. and Xu, X.
949 (2014) Isolation and characterization of dissolved organic matter
950 fractions from antialgal products of Microcystis aeruginosa.
951 Environmental Science and Pollution Research 21(5), 3946-3954.

952 Korak, J.A., Wert, E.C. and Rosario-Ortiz, F.L. (2015) Evaluating
953 fluorescence spectroscopy as a tool to characterize cyanobacteria
954 intracellular organic matter upon simulated release and oxidation in
955 natural water. Water Research 68, 432-443.

956 Kostoglidis, A., Pattiaratchi, C.B. and Hamilton, D.P. (2005) CDOM
957 and its contribution to the underwater light climate of a shallow,
958 microtidal estuary in south-western Australia. Estuarine, Coastal and
959 Shelf Science 63(4), 469-477.

960 Kpodonu, A.T.N., Hamilton, D.P., Hartland, A., Laughlin, D.C. and
961 Lusk, C.H. (2016) Coupled use of sediment phosphorus speciation
962 and pigment composition to infer phytoplankton phenology over 700
963 years in a deep oligotrophic lake. Biogeochemistry 129(1-2), 181-
964 196.

965 Kudela, R.M., Palacios, S.L., Austerberry, D.C., Accorsi, E.K.,
966 Guild, L.S. and Torres-Perez, J. (2015) Application of hyperspectral
967 remote sensing to cyanobacterial blooms in inland waters. Remote
968 Sensing of Environment 167, 196-205.

969 Lee, S.J. and Wentz, E.A. (2008) Applying Bayesian Maximum
970 Entropy to extrapolating local-scale water consumption in Maricopa
971 County, Arizona. Water Resources Research 44(1).

972 Lee, T.-y., Tsuzuki, M., Takeuchi, T., Yokoyama, K. and Karube, I.
973 (1995) Quantitative determination of cyanobacteria in mixed
974 phytoplankton assemblages by an in vivo fluorimetric method.
975 Analytica Chimica Acta 302(1), 81-87.

976 Leloup, M., Nicolau, R., Pallier, V., Yéprémian, C. and Feuillade-
977 Cathalifaud, G. (2013) Organic matter produced by algae and
978 cyanobacteria: quantitative and qualitative characterization. Journal
979 of Environmental Sciences 25(6), 1089-1097.

980 Li, Z., Hobson, P., An, W., Burch, M.D., House, J. and Yang, M.
981 (2012) Earthy odor compounds production and loss in three
982 cyanobacterial cultures. Water Research 46(16), 5165-5173.

983 MacColl, R. (1998) Cyanobacterial phycobilisomes. Journal of
984 structural biology 124(2), 311-334.

985 Mackey, M., Mackey, D., Higgins, H. and Wright, S. (1996)
986 CHEMTAX—a program for estimating class abundances from
987 chemical markers: application to HPLC measurements of
988 phytoplankton. Marine Ecology Progress Series, 265-283.

989 Maier, H.R. and Dandy, G.C. (2000) Neural networks for the
990 prediction and forecasting of water resources variables: a review of
991 modelling issues and applications. Environmental Modelling &
992 Software 15(1), 101-124.

993 McBride, C. and Rose, K. (in press) Automated high-frequency
994 monitoring and research. In "Lake Restoration Handbook: A New
995 Zealand Perspective". Hamilton DP, Collier, K., Quinn JM and C,
996 H.-W. (eds), Springer.

997 McQuaid, N., Zamyadi, A., Prévost, M., Bird, D. and Dorner, S.
998 (2011) Use of in vivo phycocyanin fluorescence to monitor potential
999 microcystin-producing cyanobacterial biovolume in a drinking water
1000 source. Journal of Environmental Monitoring 13(2), 455-463.

1001 Millie, D.F., Pigg, R.J., Fahnenstiel, G.L. and Carrick, H.J. (2010)
1002 Algae: Source to Treatment. American Water Works Association
1003 (ed), pp. 93-114.

1004 Mosleh, M.A., Manssor, H., Malek, S., Milow, P. and Salleh, A.
1005 (2012) A preliminary study on automated freshwater algae
1006 recognition and classification system, p. S25, BioMed Central.

1007 NSW Office of Water (2012) Evaluation of a YSI fluorometer to
1008 determine cyanobacterial presence in the Murray and lower Darling
1009 Rivers, p. 65, New South Wales Department of Primary Industries, a
1010 division of NSW Department of Trade and Investment, Regional
1011 Infrastructure and Services.

1012 O'Neil, J.M., Davis, T.W., Burford, M.A. and Gobler, C.J. (2012)
1013 The rise of harmful cyanobacteria blooms: The potential roles of
1014 eutrophication and climate change. Harmful Algae 14, 313-334.

1015 Oliver, R.L., Hamilton, D.P., Brookes, J.D. and Ganf, G.G. (2012)
1016 Ecology of cyanobacteria II, pp. 155-194, Springer.

1017 Pemberton, K.L., Smith, R.E., Silsbe, G.M., Howell, T. and Watson,
1018 S.B. (2007) Controls on phytoplankton physiology in Lake Ontario
1019 during the late summer: evidence from new fluorescence methods.
1020 Canadian Journal of Fisheries and Aquatic Sciences 64(1), 58-73.

1021 Pinckney, J.L., Richardson, T.L., Millie, D.F. and Paerl, H.W. (2001)
1022 Application of photopigment biomarkers for quantifying microalgal
1023 community composition and in situ growth rates. Organic
1024 Geochemistry 32(4), 585-595.

1025 Pivokonsky, M., Naceradska, J., Kopecka, I., Baresova, M.,
1026 Jefferson, B., Li, X. and Henderson, R. (2016) The impact of
1027 algogenic organic matter on water treatment plant operation and
1028 water quality: A review. Critical Reviews in Environmental Science
1029 and Technology 46(4), 291-335.

1030 Pobel, D., Robin, J. and Humbert, J.-F. (2011) Influence of sampling
1031 strategies on the monitoring of cyanobacteria in shallow lakes:
1032 Lessons from a case study in France. Water Research 45(3), 1005-
1033 1014.

1034 Pomati, F., Jokela, J., Simona, M., Veronesi, M. and Ibelings, B.W.
1035 (2011) An automated platform for phytoplankton ecology and aquatic
1036 ecosystem monitoring. Environmental science & technology 45(22),
1037 9658-9665.

1038 Pomati, F., Kraft, N.J., Posch, T., Eugster, B., Jokela, J. and Ibelings,
1039 B.W. (2013) Individual cell based traits obtained by scanning flow-
1040 cytometry show selection by biotic and abiotic environmental factors
1041 during a phytoplankton spring bloom. PloS one 8(8), e71677.

1042 Randolph, K., Wilson, J., Tedesco, L., Li, L., Pascual, D.L. and
1043 Soyeux, E. (2008) Hyperspectral remote sensing of cyanobacteria in
1044 turbid productive water using optically active pigments, chlorophyll a
1045 and phycocyanin. Remote Sensing of Environment 112(11), 4009-
1046 4019.

1047 Rapala, J. (1998) Toxin production by freshwater cyanobacteria:
1048 effects of environmental factors, Biocenter Viikki Universitatis
1049 Helsingiensis.

1050 Richardson, T.L., Lawrenz, E., Pinckney, J.L., Guajardo, R.C.,
1051 Walker, E.A., Paerl, H.W. and MacIntyre, H.L. (2010) Spectral
1052 fluorometric characterization of phytoplankton community
1053 composition using the Algae Online Analyser®. Water Research
1054 44(8), 2461-2472.

1055 Rigosi, A., Hanson, P., Hamilton, D.P., Hipsey, M., Rusak, J.A.,
1056 Bois, J., Sparber, K., Chorus, I., Watkinson, A.J. and Qin, B. (2015)
1057 Determining the probability of cyanobacterial blooms: the
1058 application of Bayesian networks in multiple lake systems.
1059 Ecological applications 25(1), 186-199.

1060 Rohrlack, T., Edvardsen, B., Skulberg, R., Halstvedt, C.B., Utkilen,
1061 H.C., Ptacnik, R. and Skulberg, O.M. (2008) Oligopeptide
1062 chemotypes of the toxic freshwater cyanobacterium Planktothrix can
1063 form sub-populations with dissimilar ecological traits. Limnology
1064 and Oceanography 53(4), 1279-1293.

1065 Sackmann, B., Perry, M. and Eriksen, C. (2008) Seaglider
1066 observations of variability in daytime fluorescence quenching of
1067 chlorophyll-a in Northeastern Pacific coastal waters. Biogeosciences
1068 Discussions 5(4), 2839-2865.

1069 Schopf, J.W. (1996) Cyanobacteria: pioneers of the early Earth.
1070 NOVA HEDWIGIA BEIHEFT 112, 13-32.

1071 Schreiber, U., Endo, T., Mi, H. and Asada, K. (1995) Quenching
1072 Analysis of Chlorophyll Fluorescence by the Saturation Pulse
1073 Method: Particular Aspects Relating to the Study of Eukaryotic
1074 Algae and Cyanobacteria. Plant and Cell Physiology 36(5), 873-882.

1075 Seppälä, J., Ylöstalo, P., Kaitala, S., Hällfors, S., Raateoja, M. and
1076 Maunula, P. (2007) Ship-of-opportunity based phycocyanin
1077 fluorescence monitoring of the filamentous cyanobacteria bloom
1078 dynamics in the Baltic Sea. Estuarine, Coastal and Shelf Science
1079 73(3), 489-500.

1080 Singh, V.P. (2000) The entropy theory as a tool for modeling and
1081 decision-making in environmental and water resources.

1082 Sobiechowska-Sasim, M., Stoń-Egiert, J. and Kosakowska, A. (2014)
1083 Quantitative analysis of extracted phycobilin pigments in
1084 cyanobacteria—an assessment of spectrophotometric and
1085 spectrofluorometric methods. Journal of Applied Phycology 26(5),
1086 2065-2074.

1087 Steiner, K., Wood, S.A., Puddick, J., Hawes, I., Dietrich, D.R. and
1088 D.P, H. (2017) A Comparison of bacterial community structure,
1089 activity and microcystins associated with formation and breakdown
1090 of a cyanobacterial scum. Microbial Ecology

1091 Su, M., Yu, J., Zhang, J., Chen, H., An, W., Vogt, R.D., Andersen,
1092 T., Jia, D., Wang, J. and Yang, M. (2015) MIB-producing
1093 cyanobacteria (Planktothrix sp.) in a drinking water reservoir:
1094 Distribution and odor producing potential. Water Research 68, 444-
1095 453.

1096 Tanyimboh, T. and Sheahan, C. (2002) A maximum entropy based
1097 approach to the layout optimization of water distribution systems.
1098 Civil Engineering and Environmental Systems 19(3), 223-253.

1099 Uusitalo, L. (2007) Advantages and challenges of Bayesian networks
1100 in environmental modelling. Ecological Modelling 203, 312-318.

1101 van der Linden, L., Courtois, S., Faye, G., Steinmann, D. and
1102 Newcombe, G. (2017) Post-processing of phycocyanin probe outputs
1103 for better reservoir and water treatment outcomes, Shanghai, China.

1104 Vézie, C., Rapala, J., Vaitomaa, J., Seitsonen, J. and Sivonen, K.
1105 (2002) Effect of nitrogen and phosphorus on growth of toxic and
1106 nontoxic Microcystis strains and on intracellular microcystin
1107 concentrations. Microbial ecology 43(4), 443-454.

1108 Wallace, B.B. and Hamilton, D.P. (2000) Simulation of water-bloom
1109 formation in the cyanobacterium Microcystis aeruginosa. Journal of
1110 Plankton Research 22(6), 1127-1138.

1111 Wang, Z., Cao, J. and Meng, F. (2015) Interactions between protein-
1112 like and humic-like components in dissolved organic matter revealed
1113 by fluorescence quenching. Water Research 68, 404-413.

1114 Waterbury, J.B., Watson, S.W., Guillard, R.R. and Brand, L.E.
1115 (1979) Widespread occurrence of a unicellular, marine, planktonic,
1116 cyanobacterium [10]. Nature 277(5694), 293-294.

1117 Watson, S.B., Monis, P., Baker, P. and Giglio, S. (2016)
1118 Biochemistry and genetics of taste-and odor-producing
1119 cyanobacteria. Harmful Algae 54, 112-127.

1120 Wert, E.C., Korak, J.A., Trenholm, R.A. and Rosario-Ortiz, F.L.
1121 (2014) Effect of oxidant exposure on the release of intracellular
1122 microcystin, MIB, and geosmin from three cyanobacteria species.
1123 Water Research 52, 251-259.

1124 Willis, A., Chuang, A.W., Woodhouse, J.N., Neilan, B.A. and
1125 Burford, M.A. (2016) Intraspecific variation in growth, morphology
1126 and toxin quotas for the cyanobacterium, Cylindrospermopsis
1127 raciborskii. Toxicon 119, 307-310.

1128 Zamyadi, A., Choo, F., Newcombe, G., Stuetz, R. and Henderson,
1129 R.K. (2016a) A review of monitoring technologies for real-time
1130 management of cyanobacteria: Recent advances and future direction.
1131 TrAC Trends in Analytical Chemistry 85, 83-96.

1132 Zamyadi, A., Henderson, R.K., Stuetz, R., Newcombe, G., Newtown,
1133 K. and Gladman, B. (2016b) Cyanobacterial management in full-
1134 scale water treatment and recycling processes: reactive dosing
1135 following intensive monitoring. Environmental Science: Water
1136 Research & Technology 2(2), 362-375.

1137 Zamyadi, A., McQuaid, N., Dorner, S., Bird, D.F., Burch, M., Baker,
1138 P., Hobson, P. and Prevost, M. (2012) Cyanobacterial detection using
1139 in vivo fluorescence probes: managing interferences for improved
1140 decision-making. Journal-American Water Works Association
1141 104(8), E466-E479.

1142 Zhou, Y., Jeppesen, E., Zhang, Y., Shi, K., Liu, X. and Zhu, G.
1143 (2016) Dissolved organic matter fluorescence at wavelength 275/342
1144 nm as a key indicator for detection of point-source contamination in a
1145 large Chinese drinking water lake. Chemosphere 144, 503-509.

1146 Zieger, S. (2015) Optical detection module of algae species used as
1147 early warning system, Graz University of Technology, Graz.

1148 Ziegmann, M., Abert, M., Müller, M. and Frimmel, F.H. (2010) Use
1149 of fluorescence fingerprints for the estimation of bloom formation
1150 and toxin production of Microcystis aeruginosa. Water Research
1151 44(1), 195-204.

1152

Highlights

- Opportunities, methods and caveats for fluorescence monitoring of cyanobacteria
- Artificial intelligence could exploit sensors data
- Potential for real-time prediction of relative abundance of cyanobacteria species