Dominant factors associated with microcystins in nine midlatitude, maritime lakes

Jean Jacoby,¹* Marisa Burghdoff,² Gene Williams,² Lorraine Read,³ and F. Joan Hardy⁴

¹ College of Science & Engineering, Seattle University, Seattle, WA, USA
² Snohomish County Public Works, Everett, WA, USA
³ TerraStat Consulting Group, Snohomish, WA, USA
⁴ Washington State Department of Health, Olympia, WA, USA
* Corresponding author: jacoby@seattleu.edu

Received 9 November 2014; accepted 18 March 2015; published 22 April 2015

Abstract

The study objective was to identify factors most closely associated with the presence of cyanotoxins in 9 lakes in the Puget Sound lowlands region of western Washington, USA. Four cyanotoxins (microcystins, anatoxin-a, saxitoxin, and cylindrospermopsin), phytoplankton, and limnological parameters were monitored twice per month from June through October 2012. Microcystin (MC) was the most commonly detected cyanotoxin and was detected in every lake at least once. Nonparametric decision forests and classification trees were used to identify variables that best predicted MC categories for 2 models; (1) presence–absence of MC, and (2) MC concentrations. The best predictors of MC in concentration categories for both models were epilimnetic total nitrogen to total phosphorus (TN:TP) ratios and the abundance of potential MC-producing cyanobacteria. Model 1 showed that observations with TN:TP ratios ≤25.7 were associated with MC presence, while MC was generally absent when TN:TP ratios were >25.7 and MC-producing cyanobacteria were ≤330 cells mL⁻¹. Model 2 showed that *Microcystis* abundance >1300 cells mL⁻¹ captured moderate (>1 and ≤6 µg L⁻¹) and high (>6 µg L⁻¹) MC concentrations. Low MC concentrations (>0.05 and ≤1 µg L⁻¹) were found when TN:TP was ≤28.8 or when *Dolichospermum* abundance was >110 cells mL⁻¹. Because of their broad applicability, thresholds for these variables may be useful in evaluating public health risk in the absence of MC measurements from lakes in this and similar regions. Decision forest and classification tree models may be promising tools for lake managers to identify dominant factors and threshold limnological values associated with cyanotoxins.

Key words: classification tree, cyanobacteria, cyanotoxins, decision forest, microcystins, *Microcystis*, Puget Sound lowland maritime lakes, total nitrogen:total phosphorus ratio

Introduction

Mass occurrences or “blooms” of cyanobacteria occur in freshwater ecosystems throughout the world and create water quality problems and aesthetic nuisances. Some cyanobacteria produce toxins (cyanotoxins), which have poisoned livestock, wildlife, and pets and had adverse effects on human health (Chorus and Bartram 1999, Falconer 2005). The most commonly measured cyanotoxin worldwide is microcystin (MC), an acute hepatotoxin produced by multiple species of cyanobacteria (WHO 2003). Other cyanotoxins of increasing concern in freshwater systems include the neurotoxins anatoxin-a and saxitoxin, and cylindrospermopsin, another hepatotoxin (Chorus and Bartram 1999).

Cyanobacterial blooms are typically caused by nutrient (i.e., phosphorus and nitrogen) enrichment of aquatic systems. Other limnological factors that have been associated with cyanobacteria include high water temperature, a stable water column, low light availability, high pH, low carbon dioxide, low grazing pressure, and low total nitrogen (TN) to total phosphorus (TP)
ratios (e.g., Paerl 1988, Hyenstrand et al. 1998, Beaulieu et al. 2013). The primary factors leading to production of cyanotoxins in individual lakes are difficult to discern, however. Regional lake surveys have found MC to be associated with the same factors that promote cyanobacterial biomass, including nutrient enrichment, high water temperatures, a stable water column, low TN:TP ratios, and high pH (e.g., Kotak et al. 2000, Graham et al. 2004, Giani et al. 2005, Bigham et al. 2009, Lindon and Heiskary 2009). Other studies have shown that the most relevant of these factors vary considerably depending on lake trophic and mixing state, environmental factors, and sampling scales (e.g., Jacoby et al. 2000, Tillmanns and Pick 2011, Orihel et al. 2012, Taranu et al. 2012).

Challenges in identifying primary causes of toxic cyanobacteria reflect complex, interacting factors of environmental conditions, bloom composition, and genetic capacity of potential toxin-producing cyanobacteria species. Blooms are highly variable both spatially and temporally in the environment, making consistent tracking and monitoring difficult. Only some cyanobacterial taxa and strains within taxa are genetically capable of producing toxins, and both toxigenic and nontoxigenic strains can coexist within populations of the same species (Vezie et al. 1998, Mbedi et al. 2005). Furthermore, the same environmental factors that regulate cyanobacterial abundance (e.g., nutrients, light, temperature) also influence the composition of species and strains within cyanobacterial communities and the production of cyanotoxins.

Animal poisonings and public health concerns resulting in public access postings and lake closures due to toxic cyanobacterial blooms are well documented in western Washington and throughout the Pacific Northwest maritime region (Jacoby and Kann 2007, Hardy 2013). Thus, the ability to identify lakes most susceptible to the occurrence of cyanotoxins based on commonly collected limnological variables would be beneficial to lake managers and public health officials, not only within this region but worldwide.

Our objective was to further the understanding of the primary drivers of cyanotoxin occurrence and magnitude. Specifically, we sought to determine the key limnological parameters that would allow us to identify the lakes most likely to have toxic cyanobacterial blooms and associated risks to human health. During June through October 2012, we measured 4 cyanotoxins (MC, anatoxin-a, cylindrospermopsin, and saxitoxin), limnological parameters (TN, TP, chlorophyll a [Chl-a], temperature, Secchi disk depth), and phytoplankton species abundance in 9 midlatitude, maritime lakes.

### Study site

The 9 study lakes are located in the Puget Sound lowlands region of western Washington, USA, with 3 lakes in each of the 3 partner counties of the Seattle-Tacoma-Everett metropolitan area (Cottage, Echo, and Wilderness lakes in King County; Harts, Spanaway, and Waughop lakes in Pierce County; and Cassidy, Ketchum, and Loma lakes in Snohomish County; Fig. 1). The lakes are relatively small, ranging in size from 9 to 113 ha, and all are fairly productive, ranging in trophic status from mesotrophic to hyper-eutrophic (Table 1). Each lake has a history of toxic cyanobacterial blooms. Similar to most of the lowland lakes in this region, the study lakes have received increased nutrient loading from logging in the late 1800s and from the subsequent conversion of forested lands to urban, residential, and/or agricultural land uses. All but one of the lakes thermally stratifies during summer, with the onset of lake mixing occurring in late October. As the only polymictic lake, Waughop Lake mixes throughout summer and is one of 2 hyper-eutrophic lakes in the study.
Table 1. Morphometric characteristics and trophic status of the 9 study lakes.

<table>
<thead>
<tr>
<th>Lake</th>
<th>Area (ha)</th>
<th>Mean Depth (m)</th>
<th>Max Depth (m)</th>
<th>Watershed Area (ha)</th>
<th>Trophic Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassidy²</td>
<td>53</td>
<td>3.4</td>
<td>7.5</td>
<td>1072</td>
<td>Eutrophic</td>
</tr>
<tr>
<td>Cottage³</td>
<td>26</td>
<td>4.6</td>
<td>7.6</td>
<td>1731</td>
<td>Eutrophic</td>
</tr>
<tr>
<td>Echo³</td>
<td>12</td>
<td>4.3</td>
<td>9.2</td>
<td>117</td>
<td>Eutrophic</td>
</tr>
<tr>
<td>Harts⁴</td>
<td>48</td>
<td>7.9</td>
<td>15.3</td>
<td>1036</td>
<td>Eutrophic</td>
</tr>
<tr>
<td>Ketchum²</td>
<td>10</td>
<td>3.7</td>
<td>6.4</td>
<td>156</td>
<td>Hyper-eutrophic</td>
</tr>
<tr>
<td>Loma²</td>
<td>9</td>
<td>3.4</td>
<td>8.5</td>
<td>88</td>
<td>Eutrophic</td>
</tr>
<tr>
<td>Spanaway⁵</td>
<td>113</td>
<td>4.9</td>
<td>8.5</td>
<td>4403</td>
<td>Mesotrophic</td>
</tr>
<tr>
<td>Waughop⁶</td>
<td>13</td>
<td>2.1</td>
<td>2.8</td>
<td>88</td>
<td>Hyper-eutrophic</td>
</tr>
<tr>
<td>Wilderness³</td>
<td>27</td>
<td>6.4</td>
<td>11.6</td>
<td>133</td>
<td>Mesotrophic</td>
</tr>
</tbody>
</table>

¹Based on trophic state thresholds defined in Welch and Jacoby (2004)
²Snohomish County Public Works (2003)
³King County Lakes Program (2014)
⁴Washington State Department of Ecology (2001)
⁶Bortleson et al. (1976)

Methods

Limnological and cyanotoxin analysis

Samples were collected at each lake from the shoreline location that was monitored during the preceding 3 years as part of a larger study of human health effects from cyanotoxins during recreational use (Hardy 2013), as well as at a limnetic station, which better described the limnological character of the lake. Samples were collected twice per month from June through October 2012 (10 sampling dates for each lake). Shoreline samples were collected for cyanotoxin and phytoplankton analysis as surface grabs by dipping the bottle mouth-down into the water to a depth of about 0.5 m. If a cyanobacterial scum was present, a sample was also collected by skimming the water surface to capture accumulated material. Discrete samples were collected at limnetic stations using a Van Dorn water sampler at 3 depths in the water column: epilimnion 1 m; metalimnion approximately middepth; and hypolimnion 1 m from bottom. Temperature was measured at the shoreline station and at the 3 depths at the limnetic station. Secchi disk depth (transparency) was measured at the limnetic station. Phytoplankton samples were preserved in the field with Lugol’s iodine solution for subsequent microscopic taxonomic analysis and enumeration.

Phytoplankton taxonomic analysis was performed on a 1.0 mL subsample of each well-mixed lake sample using a Sedgewick-Rafter counting chamber and Leitz Laborlux K compound microscope. A transect-counting methodology was used with successive horizontal sweeps of the counting chamber under 100× magnification so that the entire 1 mL subsample volume was analyzed (i.e., 308 fields of view). Replicate subsamples of a single sample selected in each group of 50 samples were routinely analyzed as a statistical check for counting precision (e.g., to insure a sufficiently small coefficient of variation). Average counts for each phytoplankton taxon were computed from subsample results. Species were identified to genus and species where possible using standard references (e.g., Prescott 1980, Wehr and Sheath 2003). Phytoplankton densities were reported in natural units as numbers of cells, filaments, or colonies per milliliter ($mL^{-1}$).

For each sample, cell dimensions of at least 10 organisms of each taxon were computed to obtain average cell biovolume (BV) per taxon based on geometric shape. Determination of cell BV dimensions and identifications were made with a calibrated Whipple disk at 400× (high dry magnification) using a Palmer-Maloney nanoplankton chamber (0.1 mL volume) or at 1000× (oil emersion). Cell BVs were reported as $\mu m^3 mL^{-1}$.

Samples from limnetic stations were analyzed for TP, TN, and Chl-$\alpha$ in the King County Environmental Lab, Seattle, WA. TP was measured by method SM4500-P-B F, and TN was analyzed using method SM4500-N-C (APHA 2005). TN and TP were measured in raw, unfiltered water samples using a potassium persulfate digestion followed by analysis in a segmented flow autoanalyzer. Chl-$\alpha$ was measured using spectrophotometric method EPA 446.0 after filtration and a 2–24 h extraction in 90% acetone at −20 °C. Four cyanotoxins (MC, anatoxin-a, saxitoxin, and cylindrospermopsin) were analyzed in the shoreline grab.
samples and, if collected, the scum samples. Cylindrospermopsin was also measured in the metalimnetic samples from the limnetic stations. Microcystins, saxatoxin, and cylindrospermopsin concentrations were measured using antibody-based, enzyme-linked immunosorbent assay (ELISA). For MC, the ELISA plate kits from Envirotech (Portland, ME, USA) with a method detection limit (MDL) of 0.05 µg L\(^{-1}\) were used. Plate kits from Abaxis, Inc. (Warminster, PA, USA) and Beacon Analytical Systems, Inc. (Saco, ME, USA) were used for saxitoxins (MDL = 0.02 µg L\(^{-1}\)) and cylindrospermopsin (MDL = 0.05 µg L\(^{-1}\)), respectively.

Total MC and cylindrospermopsin concentrations (extra- and intra-cellular) in the water samples were measured after freezing, thawing, and then immediately sonicating the samples using a Vibra Cell Sonicator to ensure lysis of the cells. Saxatoxin analysis required preservation of the samples in the field (preservative provided in the ELISA kit), which also resulted in cell lysis. Anatoxin-a was measured by high-performance liquid chromatography with fluorescence detection (James et al. 1998). Whole-water samples were heat- and acid-extracted, concentrated using weak cation exchange solid phase extraction (Harada et al. 1989, James et al. 1998), reconstituted in sodium borate buffer, and derivatized with a fluorescent compound prior to injection; o-Phthaldialdehyde was added to prevent interference from primary amines (Rawn et al. 2005). The MDL for the method ranged from 0.01 to 0.25 µg L\(^{-1}\).

### Data analysis

As the most frequently detected cyanotoxin in the 9 lakes during 2012, MC was the primary focus of the statistical analysis. The low variability in concentrations of MC observed at 8 of the lakes during the study period precluded use of regression models. Instead, 2 nonparametric statistical techniques, decision forest algorithms, and classification tree models (discussed later), were used to determine interactions between variables and identify the best predictors of MC concentrations.

Two models were developed to evaluate limnological parameters for predicting MC concentrations. Model 1 was simply MC presence–absence (i.e., above or below the MDL of 0.05 µg L\(^{-1}\)). Model 2 was based on 4 MC concentration categories: Non-detected (ND ≤ 0.05 µg L\(^{-1}\)), Low (>0.05 and ≤1 µg L\(^{-1}\)), Moderate (>1 and ≤6 µg L\(^{-1}\)), and High (>6 µg L\(^{-1}\)). MC concentrations >6 µg L\(^{-1}\) were considered High because this is the guidance value established for recreational waters in Washington State (Hardy 2008). MC concentrations above the MDL and below the World Health Organization (WHO) guidance value for drinking water (1 µg L\(^{-1}\)) defined the lower and upper limits of the Low MC concentration category, respectively. The number of observations by lake within each MC concentration (µg L\(^{-1}\)) category is shown in Supplementary Table S1.

Variables used in statistical analyses included TN, TP, TN:TP ratio, and water temperature at all 3 water depths; Chl-a (epilimnetic samples only); Secchi disk depth; water column stability; and abundances and BV of total cyanobacteria, the MC-producing cyanobacteria group, and individual MC-producing cyanobacteria genera. MC-producing cyanobacteria included the following detected genera: *Doliichospermum* (formerly known as *Anabaena*), *Anaphacapsa*, *Gloeotrichia*, *Microcystis*, and *Planktothrix* (Chorus and Bartram 1999, Carey et al. 2007). Water column stability (relative thermal resistance to mixing [RTRM]) was calculated as the density difference between the surface and bottom divided by the density difference between water at 4 and 5 °C (Welch and Jacoby 2004).

Decision forests (e.g., Ho 1995, Breiman 2001) are an ensemble learning method used for classification and regression that uses randomization and bootstrapping and a recursive partitioning framework to identify the best predictors globally. Decision forests grow many classification trees, aggregating results over all the trees in the forest. Because there are many trees in the forest (i.e., good solutions are not necessarily unique), there may be multiple trees with similar accuracy of predictions using locally optimal solutions. Relationships are not constrained to a particular model form (e.g., linear), and interactions among variables are explored.

The decision forest algorithm was used to generate 10000 trees (allowing 6 variables for each tree) for predicting MC concentration categories. This algorithm allowed the computation of the relative “importance” among all variables, with a higher importance value indicative of a better predictor of MC concentration categories, aggregated over all trees in the forest. The variable importance was measured by the decrease in prediction accuracy on the observations left out of the training dataset (i.e., “out-of-the-bag” observations) using the real data versus a random reshuffling of the data. A decrease in prediction accuracy of 0.1 for a variable indicated that, averaged over all 10 000 trees in the forest, 10% of the out-of-the-bag cases were inaccurately predicted when that variable was randomly shuffled. A small decrease in prediction accuracy indicated that the observed relationship was no better than a random permutation of the data. Within a single model, the scale of importance values is relative, and the variables with the highest importance values make the greatest contribution to prediction accuracy, whereas values near zero have little predictive value.
Variables with high importance values were included in the classification tree models. The classification trees were grown using unbiased recursive binary partitioning (i.e., repeatedly splitting the data into 2 parts to isolate cases with similar outcomes) in a conditional inference framework (Hothorn et al. 2006b), which identified the most effective splits along one or more of the predictor variables using hypothesis tests of independence. The specific hypothesis test depends on the type of data being modeled (i.e., categorical, ordinal, continuous, censored, or multivariate). At each node, the predictor with the highest association of significance (i.e., the smallest p-value) for the hypothesis test of independence was selected for splitting. The objective for each split was to minimize error in predicting group membership (i.e., MC concentration categories or MC presence–absence).

All analyses were performed in R (www.r-project.org) using the party package v1.0-11, party::ctree (Hothorn et al. 2006b) and party::cforest (Hothorn et al. 2006a, Strobl et al. 2007, 2008) functions for classification trees and decision forests, respectively. Graphics were produced using ggplot2, v0.9.3.1 (Wickham 2009), in R v3.02 (R Core Team 2013) and Microsoft Excel 14.0.

### Results

#### Cyanotoxins

Microcystin was detected in all 9 lakes at least once in 2012 and on the majority of sampling dates at Waughop, Cassidy, Harts, Ketchum, Echo, and Cottage (Table 2; Fig. 2). MC concentrations in surface grab samples were generally Low (<1 µg L\(^{-1}\)) at all lakes, with the exception of Waughop, which had concentrations that exceeded 1 µg L\(^{-1}\) during 8 of 10 weeks (maximum of 69.4 µg L\(^{-1}\); Table 2).

MC concentrations in the surface scum samples were typically higher than the corresponding grab samples (Spearman rank correlation \(\rho = 0.85, p << 0.001, n = 27\); Supplementary Fig. S1). While MC concentrations measured in scums indicate worst-case concentrations that have public health implications, they did not reflect ambient levels in the lakes; therefore, the following MC analyses and results exclude results from scum samples and include only results from surface grab samples.

Anatoxin-a was detected in 6 of the 9 lakes, with the highest frequency of detection in Harts Lake (5 of 10 sampling dates), which also had the highest anatoxin-a
Table 2. Microcystin (MC), anatoxin-a (ATX), saxitoxin (SAX), and cylindrospermopsin (CYL) concentrations (µg L⁻¹) in surface (0.5 m) grab samples in the 9 study lakes June–Oct 2012. Concentrations in bold exceeded the Washington State provisional guidance values for recreational waters (i.e., 6 µg L⁻¹ MC, 1 µg L⁻¹ ATX, and 75 µg L⁻¹ SAX; Hardy 2008, 2011).

ND = Not Detected; MC MDL = 0.05 µg L⁻¹; ATX MDL = 0.01–0.25 µg L⁻¹; SAX MDL = 0.02 µg L⁻¹; CYL MDL = 0.05 µg L⁻¹

<table>
<thead>
<tr>
<th>Lake</th>
<th>Toxin</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassidy</td>
<td>MC</td>
<td>ND</td>
<td>0.09</td>
<td>0.07</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>ND</td>
<td>0.03</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Cottage</td>
<td>MC</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>0.04</td>
<td>0.12</td>
<td>0.04</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Echo</td>
<td>MC</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>ND</td>
<td>0.03</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Harts</td>
<td>MC</td>
<td>ND</td>
<td>ND</td>
<td>0.20</td>
<td>0.63</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>0.07</td>
<td>2.27</td>
<td>ND</td>
<td>0.08</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Ketchum</td>
<td>MC</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>ND</td>
<td>ND</td>
<td>0.06</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Loma</td>
<td>MC</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Spanaway</td>
<td>MC</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Waughop</td>
<td>MC</td>
<td>0.67</td>
<td>0.79</td>
<td>41.70</td>
<td>4.19</td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>0.02</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>0.03</td>
<td>0.04</td>
<td>ND</td>
<td>0.05</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Wilderness</td>
<td>MC</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>ATX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>SAX</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>CYL</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>
concentration of 2.27 µg L\(^{-1}\) (Table 2), a level that exceeded the Washington State provisional recreational guidance value of 1 µg L\(^{-1}\) (Hardy 2008).

Saxitoxin was observed in Waughop Lake in surface grab samples on 3 sampling dates (Table 2) and in scum samples on 9 of the 10 sampling dates. Concentrations increased over time, with the highest concentration of ~1 µg L\(^{-1}\) in a scum sample on 8 October 2012 (data not shown). Cylindrospermopsin was not detected in any of the shoreline or limnetic middepth samples in any lake during 2012.

**Limnological variables**

The highest TP, TN, and Chl-\(\alpha\) concentrations were measured in Ketchum, Waughop, and Harts lakes, with the lowest concentrations in Spanaway and Wilderness (Table 3; Supplementary Fig. S2a–c). Lower TP concentrations in Spanaway and Wilderness were reflected in their higher mean epilimnetic (1 m) TN:TP ratios, whereas high TP concentrations in Ketchum and Waughop resulted in the lowest TN:TP ratios of 14 and 18, respectively (Table 3; Supplementary Fig. S2d). Intermediate TN and TP concentrations were found in 4 lakes (Cassidy, Cottage, Echo, and Loma; Table 3; Supplementary Fig. S2a and b). Microcystin concentrations increased while TN:TP ratios declined in most of the study lakes through the summer (Fig. 3). Lakes with higher TN:TP ratios (Spanaway and Wilderness) tended to have lower MC concentrations, and vice versa (Cassidy, Echo, Ketchum, and Waughop).

Water transparency generally tracked nutrient and Chl-\(\alpha\) concentrations, with greater Secchi depths recorded in Spanaway and Wilderness (sampling period means of 4.7 and 4.6 m, respectively; Table 3). The other lakes exhibited low water transparencies (~1–2 m) indicative of their eutrophic state.

Surface water temperatures increased over the summer in all lakes relative to hypolimnetic temperatures, except in polymictic Waughop where surface and bottom temperatures remained relatively similar. Surface water temperatures did not vary considerably between lakes.

![Fig. 3. Epilimnetic TN:TP ratios (by mass) and microcystin (MC) concentrations (µg L\(^{-1}\), MDL = 0.05 µg L\(^{-1}\)) in near-shore grab samples (0.5 m depth) in the 9 study lakes on 10 sampling dates, June–October 2012.](image-url)
potential MC-producing cyanobacteria were observed in Harts Lake in September (predominantly *Dolichospermum*; Fig. 4). *Microcystis* was also represented in the cyanobacterial community at Cassidy and Ketchum, and to a lesser extent Echo, but was not observed in Loma, Spanaway, or Wilderness lakes, and on only one occasion in Cottage Lake (100 cells mL$^{-1}$ on 9 Sep; Fig. 4).

Although not considered to be MC-producers, *Aphanizomenon* and *Woronichinia* comprised important members of the cyanobacteria in some lakes on occasion. *Aphanizomenon* was dominant in most of the lakes (Cassidy, Cottage, Echo, Harts, Loma, Spanaway, and Waughop) during June and continued to dominate at Echo, Loma, and Spanaway during most of the summer and fall (Fig. 4). *Woronichinia* was a dominant cyanobacterium in Harts and Ketchum lakes on 2 sampling dates, and although it was frequently present in other lakes, it was neither abundant nor dominant (Fig. 4).

### Table 3. Mean and standard deviation of limnological characteristics at the limnnetic sampling stations June–October 2012 prior to lake turnover ($n = 10$). Epilimnetic samples were collected at a water depth of 1 m. Hypolimnetic sample depths varied by lake and are shown in first row of table. Metalimnetic data are not shown.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cassidy</th>
<th>Cottage</th>
<th>Echo</th>
<th>Harts</th>
<th>Ketchum</th>
<th>Loma</th>
<th>Spanaway</th>
<th>Waughop</th>
<th>Wilderness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypo Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth (m)</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>15</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>2.5</td>
<td>7.5</td>
</tr>
<tr>
<td>TN - Epi (µg L$^{-1}$)</td>
<td>909</td>
<td>658</td>
<td>681</td>
<td>1401</td>
<td>1531</td>
<td>682</td>
<td>903</td>
<td>1752</td>
<td>386</td>
</tr>
<tr>
<td>±216</td>
<td>±202</td>
<td>±164</td>
<td>±612</td>
<td>±454</td>
<td>±83</td>
<td>±195</td>
<td>±1340</td>
<td>±132</td>
<td></td>
</tr>
<tr>
<td>TN - Hypo (µg L$^{-1}$)</td>
<td>933</td>
<td>1680</td>
<td>614</td>
<td>2149</td>
<td>2978</td>
<td>978</td>
<td>3673</td>
<td>1886</td>
<td>684</td>
</tr>
<tr>
<td>±319</td>
<td>±950</td>
<td>±208</td>
<td>±658</td>
<td>±1143</td>
<td>±94</td>
<td>±2304</td>
<td>±1215</td>
<td>±181</td>
<td></td>
</tr>
<tr>
<td>TP - Epi (µg L$^{-1}$)</td>
<td>39</td>
<td>30</td>
<td>33</td>
<td>71</td>
<td>153</td>
<td>31</td>
<td>16</td>
<td>126</td>
<td>15</td>
</tr>
<tr>
<td>±15</td>
<td>±14</td>
<td>±10</td>
<td>±30</td>
<td>±159</td>
<td>±5</td>
<td>±9</td>
<td>±154</td>
<td>±9</td>
<td></td>
</tr>
<tr>
<td>TP - Hypo (µg L$^{-1}$)</td>
<td>160</td>
<td>949</td>
<td>77</td>
<td>1335</td>
<td>2101</td>
<td>83</td>
<td>523</td>
<td>144</td>
<td>54</td>
</tr>
<tr>
<td>±114</td>
<td>±599</td>
<td>±39</td>
<td>±552</td>
<td>±853</td>
<td>±40</td>
<td>±376</td>
<td>±143</td>
<td>±36</td>
<td></td>
</tr>
<tr>
<td>TN:TP (Epi)</td>
<td>25</td>
<td>25</td>
<td>21</td>
<td>21</td>
<td>14</td>
<td>23</td>
<td>72</td>
<td>18</td>
<td>32</td>
</tr>
<tr>
<td>±5</td>
<td>±10</td>
<td>±4</td>
<td>±5</td>
<td>±6</td>
<td>±3</td>
<td>±36</td>
<td>±6</td>
<td>±17</td>
<td></td>
</tr>
<tr>
<td>Chl-α (µg L$^{-1}$)</td>
<td>21</td>
<td>14</td>
<td>11</td>
<td>23</td>
<td>49</td>
<td>5.9</td>
<td>5.9</td>
<td>28</td>
<td>4.8</td>
</tr>
<tr>
<td>±14</td>
<td>±8.2</td>
<td>±4.8</td>
<td>±28</td>
<td>±50</td>
<td>±3.9</td>
<td>±4.3</td>
<td>±28</td>
<td>±3.2</td>
<td></td>
</tr>
<tr>
<td>Secchi (m)</td>
<td>1.0</td>
<td>2.1</td>
<td>2.0</td>
<td>2.1</td>
<td>1.5</td>
<td>1.3</td>
<td>4.7</td>
<td>1.5</td>
<td>4.6</td>
</tr>
<tr>
<td>±0.2</td>
<td>±0.7</td>
<td>±0.5</td>
<td>±0.7</td>
<td>±0.8</td>
<td>±0.3</td>
<td>±2.2</td>
<td>±0.9</td>
<td>±1.2</td>
<td></td>
</tr>
<tr>
<td>Temp - Epi (°C)</td>
<td>18.4</td>
<td>19.0</td>
<td>18.8</td>
<td>20.6</td>
<td>18.5</td>
<td>17.5</td>
<td>19.8</td>
<td>19.3</td>
<td>20.2</td>
</tr>
<tr>
<td>±3.9</td>
<td>±3.8</td>
<td>±3.4</td>
<td>±4.2</td>
<td>±3.9</td>
<td>±3.9</td>
<td>±3.4</td>
<td>±4.1</td>
<td>±3.1</td>
<td></td>
</tr>
<tr>
<td>Temp - Hypo (°C)</td>
<td>11.0</td>
<td>9.9</td>
<td>11.9</td>
<td>7.5</td>
<td>9.4</td>
<td>8.3</td>
<td>12.1</td>
<td>18.6</td>
<td>13.8</td>
</tr>
<tr>
<td>±0.7</td>
<td>±1.2</td>
<td>±1.6</td>
<td>±1.2</td>
<td>±1.6</td>
<td>±0.4</td>
<td>±1.7</td>
<td>±3.6</td>
<td>±1.6</td>
<td></td>
</tr>
<tr>
<td>RTRM</td>
<td>143</td>
<td>165</td>
<td>132</td>
<td>235</td>
<td>165</td>
<td>142</td>
<td>159</td>
<td>18</td>
<td>144</td>
</tr>
<tr>
<td>±93</td>
<td>±101</td>
<td>±103</td>
<td>±97</td>
<td>±95</td>
<td>±92</td>
<td>±98</td>
<td>±29</td>
<td>±91</td>
<td></td>
</tr>
</tbody>
</table>

RTRM values increased in the 8 stratified lakes as summer stratification progressed, with mean RTRMs that ranged from ~140 to 240 during the June–October sampling period (Table 3). The RTRM in Waughop Lake was low (mean = 18), reflecting its nearly continual mixing.

**Phytoplankton**

Cyanobacteria dominated the phytoplankton community in all lakes on the majority of sampling dates throughout the sampling period (Supplementary Fig. S3). Cyanobacteria were most abundant in Cassidy, Echo, Harts, and Waughop lakes and relatively lower in Loma, Spanaway, and Wilderness. Cyanobacterial abundance and BV were highly correlated and showed similar patterns (Supplementary Fig. S4).

Cyanobacterial genera with the potential to produce MC (i.e., *Dolichospermum*, *Aphanocapsa*, *Gloeotrichia*, *Microcystis*, and *Planktothrix*) were present in all lakes on some sampling dates (Fig. 4). Highest abundances of potential MC-producing cyanobacteria were observed in Harts Lake in September (predominantly *Dolichospermum*) and in Waughop Lake on most sampling dates (predominantly *Microcystis*; Fig. 4). *Microcystis* was also represented in the cyanobacterial community at Cassidy and Ketchum, and to a lesser extent Echo, but was not observed in Loma, Spanaway, or Wilderness lakes, and on only one occasion in Cottage Lake (100 cells mL$^{-1}$ on 9 Sep; Fig. 4).

Although not considered to be MC-producers, *Aphanizomenon* and *Woronichinia* comprised important members of the cyanobacteria in some lakes on occasion. *Aphanizomenon* was dominant in most of the lakes (Cassidy, Cottage, Echo, Harts, Loma, Spanaway, and Waughop) during June and continued to dominate at Echo, Loma, and Spanaway during most of the summer and fall (Fig. 4). *Woronichinia* was a dominant cyanobacterium in Harts and Ketchum lakes on 2 sampling dates, and although it was frequently present in other lakes, it was neither abundant nor dominant (Fig. 4).
Microcystis was present during approximately half of the detected MC observations in the study lakes (notably Cassidy, Echo, Harts, Ketchum, and Waughop; Table 2; Fig. 4). The only detected MC in Spanaway Lake was on 23 September 2012 when there was a mixed cyanobacterial assemblage of Dolichospermum and Aphanizomenon. MC was detected in September and October in Lake Loma, even though MC-producers were not observed (Aphanizomenon dominated). MC in Cottage Lake was also associated with cyanobacteria other than Microcystis. During the 6 occasions (Aug–Oct) when MC was detected in Cottage Lake, Dolichospermum was the only potential MC-producing cyanobacterium present in the samples.

**Decision forest and classification tree models**

The decision forest algorithm and classification tree models yielded similar results using either cyanobacterial abundance or BV as predictors of MC. Because cyanobacterial abundance is a commonly used metric to assess risk due to cyanotoxins, those model results are described below. The decision forests and classification trees based on cyanobacterial BV are included in the Supplementary Material (Supplementary Fig. S5a–b, S6a–b).

**Model 1: presence–absence of MC**

In Model 1, variables with the highest importance values were Woronichinia abundance, epilimnetic TN:TP ratios, and MC-producing cyanobacterial abundance; the remaining variables barely improved the prediction accuracy relative to a random reshuffling of the data (Fig. 5).

Using the same set of variables included in the decision forest, the classification tree analysis indicated that the presence–absence of MC in the 9 study lakes was largely predicted by epilimnetic TN:TP ratios and the abundance of potential MC-producing cyanobacteria (Fig. 6). Specifically, the most effective classification tree showed that observations with TN:TP ratios ≤25.7 were associated with MC presence with 73% accuracy (45 of
62 cases with low TN:TP ratios had MC present; Node 2, Fig. 6). Observations with TN:TP ratios >25.7 and with MC-producing cyanobacteria ≤330 cells mL$^{-1}$ were associated with MC absence with 95% accuracy (21 of 22 cases; Node 4, Fig. 6). Observations with TN:TP ratios >25.7 and MC-producing cyanobacteria >330 cells mL$^{-1}$ were associated with MC presence (6 of 6 cases; Node 5, Fig. 6).

The overall misclassification error of this classification tree model was 20% [(17+1)/90]. The model more accurately predicted MC presence than MC absence with misclassification errors of 2% (1 of 52 cases) and 45% (17 of 38 cases), respectively.

Model 2: four MC concentration categories

For this model, variables identified by the decision forest algorithm as the best predictors of MC concentration categories included Woronichinia and Microcystis abundance, epilimnetic TN:TP ratios, and abundance of potential MC-producing cyanobacteria (Fig. 7). Using the same set of variables included in the decision forest to grow classification trees produced results that included Aphanocapsa as a predictor; however, Aphanocapsa was only present in Waughop Lake at relatively Low abundance (300–15,400 cells mL$^{-1}$) during August–October and in Wilderness on 2 occasions at especially low levels (80 and 100 cells mL$^{-1}$). Aphanocapsa appeared as a primary splitting variable in the classification trees only because it was largely unique to Waughop, the only lake with High MC concentrations. It was therefore excluded from the classification tree analysis because it was not a generalizable predictor of MC.

![Figure 5. Model 1: Predictor variables of microcystin (MC) presence–absence in the 9 study lakes ranked by their importance values as measured by the mean decrease in accuracy in the decision forest.](image)

![Figure 6. Results of the classification tree analysis for Model 1: microcystin absence–presence.](image)
A classification tree model that excluded *Aphanocapsa* resulted in a simple model (Fig. 8) similar to that predicted for MC presence–absence (Fig. 6). *Microcystis* abundance provided the first split in the data; specifically, Microcystis abundance >1300 cells mL$^{-1}$ captured all of the Moderate and High MC observations and few of the observations in the lower concentration categories (Node 7, Fig. 8). When Microcystis was ≤1300 cells mL$^{-1}$, the TN:TP ratio was the next classifying variable, with ratios ≤28.8 predicting Low MC concentrations with 63% accuracy (Node 3). For TN:TP ratios >28.8, Dolichospermum abundance was the final classifying variable, with cell counts of ≤110 cells mL$^{-1}$ predicting ND observations of MC with 100% accuracy (13 of 13 cases, Node 5). Higher Dolichospermum abundance was associated with a mixture of ND and Low MC observations (4 and 3 cases, respectively; Node 6, Fig. 8).

The overall misclassification error for this classification tree model was 34%. The model most accurately predicted the Moderate and Low categories with misclassification error rates of 0% and 12% (0 of 6 and 5 of 41 cases, respectively). The ND category was not readily distinguishable from the Low category, which led to a high misclassification rate of 66% (25 of 38 cases). The High category (present only at Waughop) was 100% misclassified if it was considered as a category separate from Moderate; however, if Moderate and High MC levels were considered a single category (i.e., >1 µg L$^{-1}$), then this combined category had 0% misclassification error.

For both MC prediction models, primary variables used to split the data in the classification trees matched the

![Classification Tree Diagram](image-url)

**Fig. 7.** Model 2: Predictor variables of microcystin concentration category (not detected [ND], Low, Moderate, High) in the 9 study lakes ranked by their importance values as measured by the mean decrease in accuracy in the decision forest.

![Classification Tree Diagram](image-url)

**Fig. 8.** Results of the classification tree analysis for Model 2: Four microcystin (MC) categories (ND, Low, Moderate, High).
variables of highest importance from decision forest results (Fig. 5 and 7), with the exception of *Woronichinia* abundance. The absence of *Woronichinia* in the classification models indicated that this variable split the data in a similar manner as other closely related variables (i.e., MC-producing cyanobacteria in Model 1 and the combination of *Microcystis* abundance and TN:TP ratios in Model 2). Furthermore, the inclusion of *Woronichinia* in the trees resulted in higher misclassification errors; thus, the most effective classification trees did not use it to split the data.

**Discussion**

Similar to other studies throughout the world (e.g., Chorus and Bartram 1999), we found that MC was the most prevalent cyanotoxin in the 9 Puget Sound lowland lakes during June–October 2012. Microcystin was also most frequently detected in the larger 30-lake study during the previous 3 years (Hardy 2013). This larger study concluded that MC was the most important cyanotoxin in the region from a public health perspective. Based on these findings, limnological parameters associated with MC production became the primary focus of analysis in the 9-lake study.

The principal conclusion of our analysis is that the best predictors of MC presence and MC concentrations were epilimnetic TN:TP ratios and the abundance of potential MC-producing cyanobacteria, especially *Microcystis* and *Dolichospermum*. Specifically, Model 1 (developed to evaluate MC presence–absence) showed that observations with lower TN:TP ratios were generally associated with MC presence, while MC was mostly absent when TN:TP ratios were higher and MC-producing cyanobacterial abundance was low (Fig. 6). Model 2 (focused on MC concentration categories) showed that even fairly low *Microcystis* abundance was associated with all of the Moderate and High MC concentrations. Low MC concentrations were found when TN:TP was low or when *Dolichospermum* was present (Fig. 8). The relevance of our results to TN:TP ratios, MC-producing cyanobacteria, and other limnological factors and the use of nonparametric decision forest analysis and classification tree models are discussed in the following sections.

**TN:TP ratios and MC**

The association between low TN:TP ratios and MC found in this study reinforces the findings of numerous other studies around the world. Low TN:TP ratios have long been recognized to favor cyanobacterial dominance in lake phytoplankton (e.g., Smith 1983, Pick and Lean 1987). Smith showed that the percentage of cyanobacteria in the phytoplankton of 17 lakes worldwide was low (<10%) when TN:TP >29 and increased dramatically when ratios fell below 29. Commonly noted rationale for low TN:TP ratios favoring cyanobacteria include their lower cellular N:P requirement than other phytoplankton as well as the ability of some cyanobacteria to fix N.

More recently, relationships between TN:TP ratios and MC concentrations have been found in regional and global datasets (Graham et al. 2004, Orihel et al. 2012, Harris et al. 2014). These studies have shown that MC concentrations increase as TN:TP ratios decrease, which is a logical extension of the relationship between cyanobacteria in general and their relative requirement for N and P as described in earlier work. Orihel et al. (2012) found that high MC occurred under low TN:TP ratios (<23) in high-nutrient Canadian lakes, although the Scott et al. (2013) reanalysis of this same dataset indicated that High MC occurred at intermediate TN:TP (12–23) and elevated nutrient concentrations. The global dataset analyzed by Harris et al. (2014) indicated that High MC concentrations (>20 µg L⁻¹) generally occur at TN:TP <30, which is consistent with the threshold identified by Smith (1983) for cyanobacterial dominance and with our findings of measurable MC at TN:TP ratios of ≤25.7 for Model 1 and ≤28.8 for Model 2.

Although there seems to be a TN:TP threshold (≤25.7) associated with MC detection in the 9 study lakes, the relationship may not be causal. It is well established that, as eutrophication proceeds, lake TN:TP ratios tend to decrease, reflecting low ratios in wastewater sources and increased internal P cycling, including cyanobacterial uptake of P in sediments and subsequent translocation to the water column (e.g., Downing and McCauley 1992, Xie et al. 2003). As noted by others (e.g., Orihel et al. 2012, Scott et al. 2013), the apparent association between TN:TP ratios and MC concentrations may be due to covariation between low TN:TP and eutrophy (with increased cyanobacterial dominance) rather than due to a causal mechanism. Low TN:TP ratios (with sufficient N and P) such as those observed in Waughop and Ketchum lakes, however, may select for toxigenic populations of *Microcystis* (Otten et al. 2012). The role of TN:TP as a driver or a covariant of MC concentration remains a question unanswered by this study.

Our findings are consistent with other studies where MC concentrations increased with nutrient enrichment (Kotak and Zurawell 2007, Dolman et al. 2012, Orihel et al. 2012). The highest MC concentrations were found in the 2 hyper-eutrophic lakes replete with N and P (Ketchum and Waughop) and the lowest MC in the 2 mesotrophic lakes (Spanaway and Wilderness).

© International Society of Limnology 2015

DOI: 10.5268/IW-5.2.808
**Microcystin-producing cyanobacteria**

Our results confirm other lake studies that have shown that the abundance of potentially toxigenic cyanobacteria genera is correlated with MC concentrations (e.g., Giani et al. 2005, Dolman et al. 2012). Ketchum and Waughop lakes, which had the highest MC concentrations, also had higher abundance and BV of potential MC-producing cyanobacteria, especially *Microcystis*, than the other 7 study lakes. Even relatively low densities of *Microcystis* yielded high potential for Moderate and High concentrations of MC in our study. *Microcystis* was found in nearly half (44%) of the samples where MC was detected and in 91% of observations in which MC was >1 µg L\(^{-1}\). In the absence of *Microcystis*, MC concentrations <1 µg L\(^{-1}\) were associated with *Dolichospermum* spp.

It is noteworthy that low densities or even absence of MC-producing cyanobacteria did not preclude MC presence in this study. As demonstrated by Lake Loma, measurable MC was found when potential MC-producing cyanobacteria were not observed. These species may have been present in the lake but at such low densities that they could have been missed during collection or analysis of the samples. It is also possible that MC may have persisted in the water column following release from cyanobacterial cells, as has been observed in other lake systems (Lahti et al. 1997, Graham et al. 2012, Zastepa et al. 2014).

Although *Woronichinia* had high importance values in the decision forest algorithms (Fig. 5 and 7), this genus was not identified as a strong predictor of MC in the classification tree models (Fig. 6 and 8). The ability of *Woronichinia* to produce MC has not been confirmed to date, although it may have that potential (Bober et al. 2011). Genetic analysis of *Woronichinia* from Waughop Lake supports its classification as a non-MC producing species in this study (J. Dyble-Bressie, NOAA Northwest Fisheries Science Center, 7 October 2011, unpubl data). *Woronichinia*’s covariance with species that are potential MC producers would likely explain its relatively high position in the decision forest rankings because *Woronichinia* was observed in most of the study lakes and found at High abundance in several.

The abundance of potential MC-producing cyanobacteria associated with measurable MC in this study was much lower than the risk guidelines published by WHO (Chorus and Bartram 1999, WHO 2003) and others (e.g., Stone and Bress 2007) for recreational waters. In addition, *Microcystis* abundance thresholds were found to be especially low (1300 cells mL\(^{-1}\)) for predicting MC concentrations >1 µg L\(^{-1}\) in our study. The WHO guidelines use thresholds of 20000 cyanobacteria cells mL\(^{-1}\) for a relatively low probability of adverse health effects risk level and 100000 cells mL\(^{-1}\) for moderate risk, with high risk indicated by the presence of cyanobacterial scums (WHO 2003). In the State of Oregon, USA, a lake is posted when the total toxigenic cyanobacterial densities are >100 000 cells mL\(^{-1}\) or when *Microcystis* or *Planktothrix* exceed 40 000 cells mL\(^{-1}\) (Stone and Bress 2007).

Using the above thresholds would have greatly underestimated risk from MC to recreational users of the lakes in our study; therefore, we believe that risk levels based on abundance of potentially toxigenic cyanobacteria should be carefully evaluated using regional or even lake-specific data that link cyanotoxin concentrations to cyanobacterial species abundance.

**Other limnological factors and MC**

Other limnological factors, such as temperature and water column stability, were expected to be key drivers in this study. Surface water temperatures and thermal stability increased through the summer, and MC concentrations generally increased as summer progressed (Supplementary Fig. S1). This trend was also discernible in the larger 30-lake study during 2009–2011, with the highest MC concentrations measured in late summer and fall before lake mixing (Hardy 2013); however, water temperature and RTRM were not identified by the decision forests as strong predictors of MC in the 9-lake dataset. While temperature and RTRM did not explain across-lake variation in MC concentrations in this study, these factors seem to be positively correlated with cyanobacterial abundance and MC concentrations within individual lakes. Tillmanns and Pick (2011) also highlighted the influence of both temporal and spatial scale in evaluations of environmental drivers of MC variability. They found that factors previously associated with MC at a larger spatial scale in among-lake studies (i.e., TP, TN, and light attenuation) did not explain temporal MC variation within one lake at any sampling scale, indicating that trends in multi-lake studies may not hold for individual lakes.

There is growing concern that cyanobacterial blooms will increase in frequency and magnitude due to increased temperature and water column stability resulting from global climate change (Paerl and Paul 2011, Kosten et al. 2012). As noted in many studies, however, the response of cyanobacteria to temperature is not straightforward and is influenced by nutrients and physical properties of lakes (i.e., water column stability in stratified and nonstratified lakes). These interactions complicate evaluation of effects of changing climate on the proliferation of harmful cyanobacteria. Literature analysis by Tararu et al. (2012) supports the premise that temperature promotes cyanobacteria once critical nutrient levels are met and indicates that stratified and nonstratified lakes would be expected to respond differently to temperature and nutrient increases.
Waughop Lake was the only nonstratified lake in the study and also had higher MC concentrations (and cyanobacteria) than the other 8 study lakes. The cyanobacterial community in Waughop Lake is a notoriously consistent, year-round producer of MC as well as other cyanotoxins (saxitoxin, anatoxin, and cylindrospermopsin; Hardy 2013), and its high cyanotoxin potential is likely related to the lake’s hypereutrophic and polymictic state.

Decision forest and classification tree models

Decision forests and classification tree models proved to be useful tools for identifying the best individual or combinations of variables associated with MC concentrations. Examining the consistencies between the trees and the forests indicated which variables were dependably strong predictors across the entire dataset (i.e., TN:TP ratios and *Microcystis* abundance). Observing where differences arose between the trees and the forests indicated which variables were specific to subsets of the data (e.g., *Aphanocapsa* associated with High MC concentrations in Lake Waughop in Model 2), or where an individual variable could be replaced with combinations of variables for more accurate predictions (i.e., *Woronichinia* in both models). Decision forests and classification tree models were also useful in identifying abundance thresholds for potential MC-producing cyanobacteria associated with MC concentration categories. Because the abundance of potential MC-producing cyanobacteria is often used to estimate risk to human health, this approach has potential to be broadly applicable for use in assessing cyanotoxins and their associated risk in lakes in other regions of the world.

Conclusion

The findings of this study support the importance of nutrient relationships and relative cyanobacterial abundance in understanding MC dynamics in freshwater lakes. Low TN:TP ratios were the most dominant factor associated with MC presence in the study lakes. Cyanobacterial abundance, specifically *Microcystis*, was the principal driver of High and Moderate MC concentrations, even at lower densities than expected based on current risk guidelines (Chorus and Bartram 1999, WHO 2003, Stone and Bress 2007). These factors may be helpful to lake managers in using existing data to identify other midlatitude lakes at risk of MC episodes and the associated human and animal health concerns. Analysis of lake datasets using decision forest and classification tree models to identify dominant factors and threshold limnological values associated with the presence of cyanotoxins is a promising tool for effective lake management.

Acknowledgements

This study was funded through a Harmful Algae Bloom-Related Illness Surveillance System (HABISS) (RFA #EH08-801) cooperative agreement with the US Centers for Disease Control and Prevention. Lorraine Backer, Epidemiologist, was the CDC Project Manager for HABISS. Phytoplankton enumerations were conducted by Maribeth Gibbons, Water Environmental Services, Inc. Other partners in the Washington State HABISS project include Kathy Hamel and Lizbeth Seebacher, Washington Department of Ecology; Ray Hanowell and Lindsay Tuttle, Tacoma Pierce County Health Department; Sally Abella, Debra Bouchard, and Beth LeDoux of King County Department of Natural Resources; and Gabriella Hannach and Francis Sweeney of King County Environmental Laboratory. We also thank the many volunteers who collected samples from the lakes in King and Pierce counties. Two anonymous reviewers provided constructive comments that improved earlier drafts of this manuscript.

References


Dominant factors associated with microcystins in nine midlatitude, maritime lakes


King County Lakes Program. 2014. King County small lakes information and data. Seattle (WA): King County Water and Land Resources Division [cited 26 June 2014]. Available from: http://green2.kingcounty.gov/SmallLakes/default.aspx


**Supplementary Material**

Supplementary Material is available for download via the Inland Waters website, https://www.fba.org.uk/journals/index.php/IW:

- Supplementary Figures S1–6; Supplementary Table S1.