CONTRIBUTION TO THE STUDY OF CYANOBACTERIAL HARMFUL ALGAL BLOOMS (CYANOHAB) BASED ON THE THRESHOLD INDEX FOR FRESHWATER LAKES

by

Kateryna Hushchyna

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at
Dalhousie University
Halifax, Nova Scotia
June 2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The worldwide phenomenon Cyanobacterial Harmful Algal Blooms (cyanoHAB) occurs with a worrisome frequency in Atlantic Canada in summer-fall periods, especially in Nova Scotia (NS) and Moncton waterbodies (NB). This work shows the dominance of cyanobacterial species associated with their potentially produced toxins in the NS and Moncton lakes and an approach to determine the thresholds of cyanoHAB. Threshold Index (TRINDEX), a combination of nutrients (absolute deviation of oxygen from 100%, phosphate, nitrate) and pigments (chlorophyll-a and phycocyanin), is proposed to detect the onset of cyanoHAB based on the independent datasets of Mattatall (2015- 2017) and Torment (2016-2018) lakes in Nova Scotia, Canada. The threshold ranges were detected by using the receiver operator curve analyses. The phycocyanin-based TRINDEX was found as a more appropriate indicator to predict cyanobacterial blooms. The research work can contribute to the monitoring and management plan for public health and freshwater resources of these provinces and globally.

LIST OF ABBREVIATIONS AND SYMBOLS USED

ACS	American Chemical Society
ADDA	3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca;4,6-dienoic acid
AGC	automatic gain control
ATP	adenosine triphosphate
AUC	area under the curve
BBML	Biofluids and Biosystems Modeling Lab, Engineering Department, Faculty of Agriculture, Dalhousie University
CE	collision energy
Chl-a	chlorophyll-a
CI	confidential interval
CN	correctly negative
CNF	correct negative fraction
CPF	correct positive fraction
CP	correctly positive
cyanoHAB	cyanobacterial harmful algal blooms
D%O	absolute deviation of dissolved oxygen from 100%
DIA	data independent acquisition
DM	direct monoclonal
dmMC-LR	microcystin with L-leucine in position 2, L-arginine in position 4 and docosahexaenoic acid in position 7
DNA	deoxyribonucleic acid
DOM	dissolved organic matter

ELISA	Enzyme-Linked Immunosorbent Assay
Fe^{2+}	iron (II)
Fe^{3+}	iron (III)
Fe-protein	iron protein
FA	formic acid
FS	full scan
GIS	geographic information system
GOGAT	glutamine oxoglutarate aminotransferase
HAB	harmful algae blooms
HDPE	high density polyethylene
HPLC	high performance liquid chromatography
HPO_4^{2-}	hydrogen phosphate
HRP	horseradish peroxidase
IN	incorrectly negative
IP	incorrectly positive
IT	time interval
J	Youden index
k	number of degrees of index
Leu1MC-LY	microcystin with L-leucine in position 1 and 2, and tyrosine in position 4
LC-HRMS	liquid chromatography-high resolutions tandem mass spectrometry
LC-MS	liquid chromatography tandem mass spectrometry
lgG	immunoglobulin G
L_i	lower limit of the *i*-parameter

LOD	limit of detection
M_i	measured parameter *i*
MeCN	acetonitrile
MC-LR	microcystin with L-leucine in position 2 and L-arginine in position 4
MC-LY	microcystin with L-leucine in position 2 and tyrosine in position 4
MC-RR	microcystin with L-arginine in positions 2 and 4
MEAN	average
ML	Mattatall lake
MoFe-protein	molybdenum-iron protein
MYC	microcystins
m/z	mass to charge ratio
n	number of variables
N_2	nitrogen as a gas
NADP	nicotinamide adenine dinucleotide phosphate
NADPH	reduced form of NADP
NB	New Brunswick
NH_3	ammonia
NF	not found
NO2	nitrite
NO3	nitrate
NS	Nova Scotia
NRC	National Research Council, Canada
OECD	The Organisation for Economic Cooperation & Development

PBS	phosphate buffered saline
PBST	phosphate buffered saline combined with tween 20
PC	phycocyanin
PETG	polyethylene terephthalate glycol
PO4	phosphate
PO_3^{3-}	phosphite ion
PVP	polyvinylpyrrolidone
Q25	first interquartile range
Q75	third interquartile range
R	coefficient of correlation
RM-ATX-a	reference material: anatoxin-a
RM-BGA	reference material: blue-green algae
RNA	ribonucleic acid
ROC	receiver operator curve
S/S	sample/standard
sp.	species
STD	standard deviation
T	threshold
T_D	threshold criterion for data
T_M	threshold criterion for model
TL	Torment lake
TMB	3,3',5,5'-Tetramethylbenzidine
TN	total nitrogen

TP	total phosphorus
TRIX	trophic index
TRINDEX	Threshold Index
U_i	upper limit of the parameter
WHO	World Health Organization
X	index to be considered

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere and deep gratitude to my supervisor Dr. Tri Nguyen-Quang for the continuous support of my master study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this master thesis. I could not have imagined having a better advisor and mentor for my study.

Besides my supervisor, I would like to thank the rest of my committee: Dr. Nancy McLean and Dr. Peter Havard for their encouragement, insightful comments, and hard questions.

I would also like to thank to the National Research Council (NRC) in Halifax who provided practical advice and assistance. The huge thanks to the Biotoxin Metrology team, including: Dr. Pearse McCarron (head of team) for his support of the project; Dr. Daniel Beach for the research advices in different steps of my toxin research from the planning experiments to the data discussion; Ms. Melanie MacArthur for the hands-on help in the laboratory analyses; Dr. Christopher O.Miles for the practical conversations toward the deeper understanding of the toxins chemistry; Ms.Cheryl Rafuse for her precious help on cyanobacterial species images and cyanobacterial physiology discussions.

I would like also to thank Dr. Ingunn A. Samdal (Toxinology Research Group, Norwegian Veterinary Institute, Oslo, Norway) for introducing the ELISA multihapten technique.

The sincere acknowledgements are addressed to the residents of Mattatall and Torment lakes especially Dave Alexander and his spouse Dawn (Lake Torment), and Blaine Odgen and Robert Maclean (Mattatall Lake) for their very active collaboration and assistance during our field trips.

I thank my Lab-mates from BBML: Kayla McLellan, Muhammad Nadeem and Dung Tri Nguyen for their involvement in the different steps during these 2 years of my research: fieldtrips, lab work, and fruitful discussions.

This research would not be achieved without the funding via the stipend from the NSERC Discovery Grant of my supervisor. To this I would express my sincere and deep gratitude. Finally, I am grateful to the financial aid from Dalhousie Faculty of Graduate Studies, which has allowed me to attend the 2d Interdisciplinary Freshwater Harmful Algal Blooms Workshop in Toronto, Ontario (April 16-18, 2018) and to present my research results. I also thank Class of 56' Graduate Scholarship in Rural Water Supply and Watershed Management from Faculty of Agriculture, Dalhousie University. Finally, my sincere acknowledges are addressed to Dr. Uday Venkatadri (ATC External Reviewer) and especially to Prof. Raymond Cote (Final Thesis Defence External Reviewer) for their precious comments, questions and suggestions to improve the quality of my work.

CHAPTER 1 INTRODUCTION

1.1 Context of the Problem and Choice of Thesis Subject

Algal blooms become a worldwide problem when marine and freshwater bodies, recreational as well as drinking water sources, suffer from a variety of blooms that could cause serious health issues for humans and animals due to their possible toxins. These toxic blooms, known as Harmful Algal Blooms (HAB) appear at a larger scale every year as a result of changes in environmental factors and climate.

HAB are blooms created by phytoplankton or algae that are unfavorable for the environment and humans either by their accumulation of toxins in the water, anoxia or huge biomass production that can cause physical issues for other habitants (Glibert et al, 2005). Among HAB family, cyanobacterial blooms (or blue-green algal blooms) in freshwater are my own research interest. The frequency and proliferation of cyanobacterial blooms have become more prominent every year due to climate change effects (Carey et al., 2004; Pearl and Paul, 2012). Not only damaging the view of lakes and stopping inhabitants from enjoying their waterbodies, HAB can also produce toxins during their life and decomposition after death. The toxins released by cyanobacteria in freshwater sources can cause many health issues and death in certain cases (Chorus and Bartram, 1999). The water could become unhealthy even dangerous for human and animal consumption, and a poisonous source for fish and invertebrates. The economical losses due to the effects of cyanobacterial HAB in freshwater in recreational and fishery industries and drinking water supply systems are very significant. Therefore, to elucidate

the bloom pattern as well as dominant factors that influence the cyanobacteria productivity and govern their behaviour nowadays is one of the main goals of the freshwater research.

The attempt to predict the bloom occurrence and proliferation under the complex context of environmental factors (nutrients, light, meteorological conditions, etc.) has led to the creation of many indices for the estimation of eutrophication. The term 'Eutrophication', from the Greek origin, meaning *Over fed,* can be applied to both marine and freshwater bodies. This is a more general definition to describe the process of enrichment of nutrients, particularly nitrogen and phosphorus, which stimulates the aquatic primary production. In this process, there are more serious manifestations leading to the visible algal blooms, algal scums, enhancing the benthic algal growth of submerged and floating macrophytes (Vollenweider, 1992). It is a well-known fact that trophic status of the waterbody can help to predict the phytoplankton growth. Less productive waters (low phytoplankton biomass) are called oligotrophic. They have clean water with low nutrients content. Mesotrophic waters characterize by medium nutrient upload and higher phytoplankton biomass. The highest amount of nutrients is found in eutrophic waterbodies. Eutrophic lakes are the most common places where HAB can develop (Oliver et al., 2012). Nevertheless, recent observations around the world also showed that oligotrophic lakes started to experience blooms as well (Anderson, et al.,2002; Sorichetti et al., 2014; Nimptsch et al., 2016) under the current climate change situation. This made the theory of eutrophication under a very serious argument that the trophic status degradation of a waterbody is not enough to be used as reason to explain the cause and consequent behaviour of cyanobacterial growth.

There exist several models and indexes that help to describe the trophic status of the waterbody and from that the phytoplankton productivity can be explained. Nevertheless, no previous work has been done to determine the threshold (onset) for the cyanoHAB occurrence.

The first goal of my research is hence focusing on the determination of the threshold for cyanobacterial bloom occurrence. Starting from the original conception of trophic index (TRIX) suggested for marine water by Vollenweider (1998), my approach will adapt it and use it as index to estimate the bloom potential occurrence in the freshwaters, as well as to evaluate the freshwater quality relating to cyanoHAB presence. The combination of nutrient factors and algal pigments can be implemented for the threshold estimation of cyanobacterial bloom in freshwater. For such concept, there is the need for a method that would allow to synthesize and insert all key data into a simple numeric expression to make information more comparable over a wide range of trophic situations, while avoiding the subjectivity in the usage of traditional trophic terminology. Moreover, this expression can be used for the prediction of bloom occurrence threshold in a combining context of nutrient and pigment present in cyanobacterial cells.

The second and third goals of my research focus on the identification of the dominant species and their released toxins; and with these research outcomes, on how they could contribute to the monitoring and management plan for water quality, public health and freshwater resources at a provincial level in NS and NB.

1.2 Research Question/Objectives and Hypotheses

Research question

The current research targets the cyanobacterial Harmful Algal Blooms (cyanoHAB) in the freshwater environment with the following research question:' *Do thresholds exist for the cyanoHAB occurrence and proliferation in freshwater environment, and how could we determine them based on common monitoring data of nutrients and environmental parameters? And how can this information be used to predict blooms?*'

Following research objectives will be carried on elucidating this research question.

Research Objectives

1. To develop a new index coined Threshold Index (or TRINDEX) for cyanoHAB based on the TRIX concept proposed for marine environment by Vollenweider et al. (1998);
2. To validate the TRINDEX by applying the real data from two Nova Scotian lakes: Mattatall Lake (Colchester and Cumberland counties) and Torment Lake (Kings County).
3. To predict the bloom based on this TRINDEX notion
4. To identify the main toxic species responsible for HAB and associating them with their released toxins (in some Nova Scotia and Eastern New Brunswick waterbodies)

Research Hypotheses

Several hypotheses have been applied to this research plan. These are:

1. The Trophic Index (TRIX) by Vollenweider et al. (1998) for marine water is completely applicable to freshwater environment.
2. Phycocyanin is one of the main parameters that can be used for prediction of cyanobacterial blooms associated with nutrients
3. Receiver Operating Curve (ROC) analyses is a useful tool for the setting of the threshold of cyanobacterial bloom appearance
4. The bloom patterns in considered field locations are generated not by one specific strain, but possibly by various dominant species (to be identified), and they could or could not release toxins depending on the environmental factors.

CHAPTER 2 LITERATURE REVIEW

2.1 Algal Blooms and Cyanobacterial Blooms

Harmful Algal Blooms (HAB) can be caused by different types of phytoplankton such as eukariotic green algae or chlorophytes, dinoflagellates, cryptophytes, chrysophytes (including diatoms) and cyanobacteria in freshwater and marine environment (Pearl et al, 2001).

The focus of my study is the freshwater cyanobacteria, as they are the main responsible phylum for HAB in fresh waterbodies.

The definition of the term 'bloom' is still under discussions. In the literature, a cyanobacterial bloom in freshwater can be considered if one of the following conditions is satisfied:

a) floating scums on the surface of the water body that can include more than 20 µg/L of chlorophyll-a and total phosphorus more than 30 µg/L (Pick, 2016);

b) high quantity of cyanobacterial cells in the water (more than 20,000 cell/mL excluding picoplankton which is 0.2-2 µm in diameter) (Pick, 2016);

c) increase of the cyanobacterial biomass with the dominance (>80%) of only one or a few species within the phytoplankton community (Humbert and Fastner, 2017)

d) dominance of cyanobacteria by biomass more than 50% in the sample (Molot et al., 2014).

e) 30±3µg/L of phycocyanin that is equal to 20,000 cells/mL and Alert 1 level by WHO (Brient et al., 2008).

Some cyanobacterial blooms found in Nova Scotia and New Brunswick are presented in Figure 2.1.

Figure 2.1: Cyanobacterial blooms in Nova Scotia and New Brunswick

Cyanobacteria are the oldest organisms on the Earth (Pearl and Otten, 2013). They are prokaryotes with a size of 2-40µm, filamentous or coccoid shapes, living in groups (colonies) or individually (Osswald et al., 2007). Cyanobacteria are photoautotrophs and can adapt to very wide ranges of temperature and nutrient. Their light harvesting mechanism, which is different from the eukaryotic algae contain phycobiliproteins allowing cyanobacteria to absorb light from a wide spectrum (Oliver et al., 2012). In the fast-changing light environment, cyanobacteria regulate a light harvesting mechanism by reducing the number of light harvesting units and turn the energy into heat. This process is called photo-adaptation (Oliver et al., 2012). There is also a photoprotective mechanism that cyanobacteria use, named an energy dissipation mechanism (Oliver et al., 2012). They have also a UV photoprotection mechanism: mycosporine-like amino acids, scytonemin that absorb UV light and help them to survive with high level of irradiance

(Pearl & Paul, 2012). Among all phytoplankton, only cyanobacteria can fix N_2 from the air in a low N environment. The ability to bind Fe^{3+} and convert it to the bioavailable form allow them cyanobacteria to get this important micronutrient from the water. Cyanobacteria have a buoyancy mechanism that allows to move in the water column using gas vesicle production to find the necessary nutrients (Carey et al., 2004). The mechanism of nutrient uptake gives cyanobacteria advantages compared to other phytoplankton when the upper water layer is nutrient depleted.
All above information indicate that cyanobacteria have many opportunities to be prevalent and bloom under different environmental conditions.

Cyanobacterial blooms have appeared more frequent in the last few decades and their duration has lasted longer in marine and freshwater environments (Pearl and Paul, 2012; EPA, 2013). The prevalence of cyanobacteria in the phytoplankton community under the climate change will follow by the increasing of cyanobacterial blooms events in the future (Carey et al., 2004). The growth rate of cyanobacteria is slower than for other eukaryotic phytoplankton, so the longer residue time is required (Chorus and Bartram, 1999). Under conditions with increasing global temperature, these conditions will be more available.

The problem of freshwater HAB has become a worrisome issue all around the globe especially for drinking water reservoirs (Asadollahfardi, 2015; Kotut et al., 2006; Liu et al., 2011; Oliver et al., 2012). The lack of ground water sources and consumption of the water directly from the lakes and ponds are real problematic issues for hot developing countries in Africa and Asia. But Canada is not an exception. As the country that has the highest amount of freshwater resources, Canada is today facing the threat of cyanoHAB. Canada has over a million lakes and ponds with more than 1000 in NS. In some part of

Canada, excess fertilization has caused the eutrophication of lakes surrounded by agricultural lands. Lakes such as Lake Champlain, Lake Ontario, Lake Erie, Lake of the Woods (LOW), and Lake Winnipeg are good examples (Pick, 2016).

There were no previous HAB surveys done to figure out the cause and origin of blooms in NS fresh waterbodies. Some surveys on water quality in general were done for areas such as Carleton river watershed (Yarmouth and Digby counties), which was monitored from 2008 after the first signs of cyanobacterial scums noticed in Fanning Lake in 2007 (Taylor, 2009). With a first survey of nine lakes (Provost, Nowlans, Hourglass, Placides, Porcupine, Parr, Ogden, Fanning, Vaughan), a bloom of *Microcystis sp.* was recorded in Nowlans Lake with a microcystin concentration of 0.3 µg/L (Taylor, 2009). The presence of *Aphanizomenon sp.* was registered in this lake as well. Fanning Lake had *Anabaena sp.* present with a high abundance. More lakes were added to the monitoring program in the Yarmouth area and from 2014, there were 14 different lakes in total including 9 in the 2008 list (Stantec Consulting, 2017). However, the data from the lakes is not consistent as the sampling occurs infrequently and isn't linked to the bloom occurrence.

In Kings County, a monitoring program consists of 13 lakes within the boundaries of the Gaspereau River watershed, with the exceptions of Lake Tupper, located within the Cornwallis Watershed; as well as Hardwood, Torment and Armstrong lakes, which are located within the LaHave River watershed. The monitoring program began from 1994, including the observation of basic parameters at one deep location for a period of 6 months from May to October annually. Unfortunately, no information about blooms and identification of species are included (Marty and Reardon, 2016).

Therefore, a bloom monitoring program around the NS province with the identification of species responsible for blooms and their released toxins is a real need.

2.2 Conditions Required for Extensive Cyanobacterial Growth

Cyanobacteria are a natural and important component of aquatic ecosystems. However, some conditions lead to their extensive growth and appearance of HAB.

According to Reynolds (1984) cyanobacteria blooms occur in rich nutrient environment. Though different cyanobacteria species have different optimal environmental conditions such as preference of light regime, temperature, amount of nutrients, and turbulence.

The following information provides the most common factors which are important for the growth of cyanobacteria and explains their importance from the physiological point of view.

2.2.1 Nutrients

2.2.1.1 Nitrogen Forms

Nitrogen is very important source for protein production. Adenosine triphosphate (ATP), nicotinamide adenine dinucleotide phosphate (NADP), deoxyribonucleic acid (DNA), RNA and amino acids contain nitrogen. The deficiency of nitrogen doesn't allow the cells to produce proteins and to process the carboxylation. Almost all phytoplankton can use ammonium, nitrate and dissolved organic nitrogen, but only cyanobacteria has an ability to fix the atmospheric nitrogen allowing them to compete for nitrogen under the nutrient depleted conditions and retain the available nitrogen in the water. Nitrogen assimilation

influences the carbon dioxide fixation rate and carbohydrate storage which in turn, influences the cell growth (Oliver et al., 2012).

Nitrogen under the form of ammonium diffuses through the cell walls (lipid and aquaporins) easily (Ritchie, 2013). Ammonia in the form of gas (NH_3) can exist in the water in high pH and can diffuse through the cell walls as well (Markou et al., 2014). However, NH_3 form is very toxic and can damage light harvesting protein complex in photosystem II in a high concentration (Markou et al., 2014). "Pump/leak" system helps to maintain the NH_3 in the cell causing the leakage of it from the cell out when ammonium uptake occurs (Ritchie, 2013). Glutamine oxoglutarate aminotransferase (GOGAT) cycle requires ATP and NADPH for the ammonium transformation to the proteins. But this energetic cost is very low (1 ATP and 1 NADPH or ferredoxin per 1 molecule of ammonium) (Oliver et al., 2012). To convert nitrite and nitrate to ammonium and further the energetic cost will be increased. The same amount of ATP is required but the number of electrons will be higher (8 electrons to transform NO2 to glutamate and 10 electrons to transform NO3 to glutamate) (García-Fernández et al., 2004). That is why ammonium is always the preferable form of nitrogen to uptake than others. There is the special mechanism that inhibits the uptake of other N forms in the presence of ammonium called "nitrogen control" process (Oliver et al., 2012).

Nitrite can be diffused through the cell wall or an active uptake is used by the cell. It is shown the toxicity effect of the nitrite on the cell if its concentration is more than 8 mM (Markou et al., 2014).

An active uptake is used by the cell to transport nitrate inside. Nitrate concentration more than 100 mM negatively influences on the growth of phytoplankton (Markou et al., 2014).

Organic forms of nitrogen can be also utilized by cyanobacteria under N-stress (Chaffin and Bridgeman, 2014). The urea degradation happens under nickel dependent urease (Oliver et al., 2012). Arginine is catabolized by a combination of the urea cycle and arginase pathway (Oliver et al., 2012). The product of both cycles is ammonium and carbon dioxide. Uptake of urea can increase the toxicity of non-N-fixing cyanobacteria (Chaffin and Bridgeman, 2014).

Nitrogen fixation is the unique mechanism of the N_2 uptake. It is quite energetically costly process that require 16ATP and 8 electrons to receive 2 molecules of ammonium and 1molecule of hydrogen (Colón-López et al., 1997). Furthermore, nitrogen fixation cannot happen without available iron and molybdenum that are the structural components of the proteins (nitrogenase enzyme complex) that needs for the reduction of N_2. But at the moment, when no other form of nitrogen available in the water, cyanobacteria switch the mechanism on the N_2 fixation. Otherwise only uptake of dissolved nitrogen compounds goes. Nitrogen fixation is separated by time and space because Fe-protein and the MoFe-protein are disabled in the presence of oxygen (Corbett et al., 2006). ATP will be damaged with a contact of oxygen and nitrogenase will not be possible. Nitrogen fixation occurs at night when there is no photosynthesis and it uses the carbon storage accumulated from the daytime controlled by circadian clock (Oliver et al., 2012). Nitrogen fixation can happen at daytime inside of the special cell that called heterocyst commonly present in the filamentous cyanobacteria. Heterocyst is the cell with a thick wall contain extra glycolipid and polysaccharide layers to reduce gas diffusion and they

do not have photosynthetic harvesting mechanism. Heterocysts are not usually present in the filamentous algae if other forms of Nitrogen are available in the water (Oliver et al., 2012). They start to appear in the N-deficiency conditions. The heterocyst produces ATP only and receive glutamate from the vegetative cells. Reduction N_2 into the ammonium follows by GOGAT cycle: the reaction of ammonium with glutamate creates second amine: glutamine, that combines with 2-oxoglutorate and creates 2 molecules of glutamate; one of them will be used for the protein production and another one will be back to the glutamine cycle (Meeks and Elhai, 2002).

Cyanobacteria cannot store nitrogen in their cells. However, according to some authors there are two storage compounds – phycocyanin which is the main pigment in the antennae and cyanophycin which is non-protein, co-polymer of arginine and aspartate (Meeks and Elhai, 2002). Cyanophycin only appears under the unstable nutrient conditions before the real nutrient limiting conditions are established. It is used by cyanobacteria until other mechanisms start to work under N-stress (Oliver et al., 2012). In N-stress environment, cyanophycin granules are degraded first, followed by degradation of the phycobilisome or light harvesting mechanism of photosystem II (Oliver et al., 2012). But the degradation of phycobilisome is not a source of nitrogen but much more the down-regulation due to stress environment when the reduction of photosynthesis and re-dox reactions are followed. The cells can be in these conditions quite a long time – months. Therefore, there is no reasonable explanation to consider phycocyanin as a storage compound.

2.2.1.2 Phosphorus

Phosphorus is one of the most abundant elements inside of the cell. Its content is 0.05-3.3% of the cell (Markou et al., 2014). Phosphorus is the main element of nucleic acids (DNA, RNA), ATP, NADP, and phospholipids in the thylakoid membranes. Membranes create a barrier with two very important functions:

a) separation of electrons and protons, which is a key mechanism for ATP production;
b) provision of the proximity to transfer energy.

The main form of the available phosphorus is the mineral forms that are most abundant as hydrogen phosphate (HPO_4^{2-}) and small fraction of PO_3^{3-} in the natural waters. There are not many studies focusing on the phosphorus uptake in cyanobacteria comparing to the nitrogen (Oliver et al., 2012).

Mineral form (HPO_4^{2-}) is preferable for the uptake. The phosphorus uptake in cyanobacteria does not differ from the same process in eukaryotic algae. It does not need additional energy for the transformation. The cells can store phosphorus in the special storage as polyphosphates called a "luxury uptake" (Markou et al., 2014). This storage is enough for the several cell doublings (Oliver et al., 2012). Phosphorus affinities are higher in cyanobacteria compared to eukaryotic algae (Molot et al., 2014). The concentration of phosphorus around 0.03 mg/L in the water is enough for the sufficient growth of the cyanobacteria (Šejnohová and Maršálek, 2012). It is shown that N-fixing cyanobacteria have lower uptake rate of mineral phosphorus than non-N-fixing phytoplankton (Suttle and Harrison, 1988).

Organic phosphorus can be also used like a source of phosphorus, but it should be converted into mineral form first (Markou et al., 2014). The cell produces the extracellular alkaline phosphatase to mineralize and mobilise organic phosphorus for the active uptake (Oliver et al., 2012). These are energetically costly reactions, and the mechanism of organic phosphorus uptake switches only under the phosphorus starved conditions (Markou et al., 2014).

2.2.1.3 Iron

Iron is an essential element for cyanobacteria as they require more iron than eukaryotic phytoplankton due to the difference in the light harvesting mechanism. But in the same time, the required iron proportion for the cell growth compared with nitrogen or phosphorus is very low and can be easily found in the fresh water. Iron is the part of enzymes (ferredoxin) used for nitrogen fixation and nitrogenase (Lis et al., 2015). It is also the part of "oxygen metabolism, electron transfer, DNA, RNA and chlorophyll synthesis" (Markou et al., 2014). There is a direct transport of Fe^{2+} into the cell. (Molot et al., 2014). However, iron has a low bioavailability in natural waters. Available Fe^{2+} form is converted into Fe^{3+} under oxygen conditions forming hydroxides and oxides which are insoluble and sink rapidly. But Fe^{2+} is available on the bottom if there are anoxic conditions. So, one of the way cyanobacteria can obtain Fe^{2+} from the bottom layer and sediment pore water from not very deep lakes is its migrating ability.

In the oxygen saturated water column iron present in Fe^{3+} form is combined with dissolved organic matter (DOM). This combination creates a very strong complex that makes iron less available to uptake. Cyanobacteria as photosynthetic bacteria have a

special mechanism (siderophores) of trapping Fe^{3+} and transporting it into the cells. There are two types of siderophores: hydroxamate siderophores and catecholate siderophores. The first type is soluble in water and has less affinity to iron binding ability (Sorichetti et al., 2014). The second type is fat-soluble and has strong iron binding. It is on the surface of the cells. Special transport system let iron go to cytoplasm where reduction of Fe^{3+} and its accumulation occurs (Lis et al., 2015). Sorichetti et al. (2014) showed that the lower DOM presence in the water, the better assimilation of iron. Experimentally shown that a low concentration of iron depressed the growth rate of *Microcystis wesenbergii* and synthesis of chl-a by increasing of ATP activity (Šejnohová and Maršálek, 2012).

Some other micronutrients such as molybdenum, cobalt, and nickel are also very important for the growth of cyanobacteria (Venkiteswaran, 2019). Molybdenum and cobalt together with iron are the part of enzymes used in nitrogen fixation. Molybdenum also plays role in nitrate reduction. Nickel is a part of an enzyme called urease.

2.2.2 Water Temperature

Temperature is the most important factor that triggers a variety of processes in phytoplankton and cyanobacterial communities. The higher the temperature the more is the uptake rate of nutrients and metabolic rate in the cells. Increasing the buoyancy of cyanobacteria has also happened with a higher temperature. Higher temperature also favours the increasing of the mineralization rate of nutrients and their availability. But all organisms have an optimal temperature for their best productivity. For the phytoplankton, the optimal range of temperature is generally not very wide (several degrees). At the temperatures higher than optimal range, cells tend to degrade due to the oxidative stress

and coagulation of the protein. However, cyanobacteria are more adapted to the temperature changes and various species of cyanobacteria can live in different temperature conditions including hot springs. Freshwater species have an optimal temperature of 17-25°C but they continue to grow even when the temperature is higher (Pearl and Paul, 2012).

As the result of the global warming and climate changes, the increase of water temperature will be observed, and the cyanobacterial growth will be prevalent compared to other phytoplankton. Suttle and Harrison (1988) have shown that the competition for phosphorus uptake of diatoms decreased when the temperature went upper 17°C but increased for green algae and cyanobacteria.

As shown by Carey et al. (2004), the buoyancy of cyanobacteria increased in warmer water and decreased in temperatures less than 15°C, resulting in their settling down at the bottom. This loss of buoyancy can be caused by the carbohydrate accumulation in the cells and down regulation of photosynthesis (Thomas and Walsby, 1986). But in the real conditions we could observe floating blooms even when the ice has covered the lakes. Therefore, optimal temperature is required to accumulate cyanobacteria biomass and create bloom, but after the growth continue even if the temperature decreased.

Increasing temperature combined with low wind and high irradiance can generate the high probability for cyanobacteria growth if the nutrients concentrations are adequate.

Finally, the hydrological characteristics of the lake play a significant role in the apparition of blooms. We can find more blooms in stagnant lakes than in flowing rivers. Therefore, quiet (stagnant) and small coves could be the first place where cyanobacterial blooms can

be observed. Hydrodynamic parameters in coupling with other factors can lead blooms to develop more in the same location or to proliferate at the larger scale later.

Figure 2.2 below summarises the main factors contributing in cyanobacteria growth.

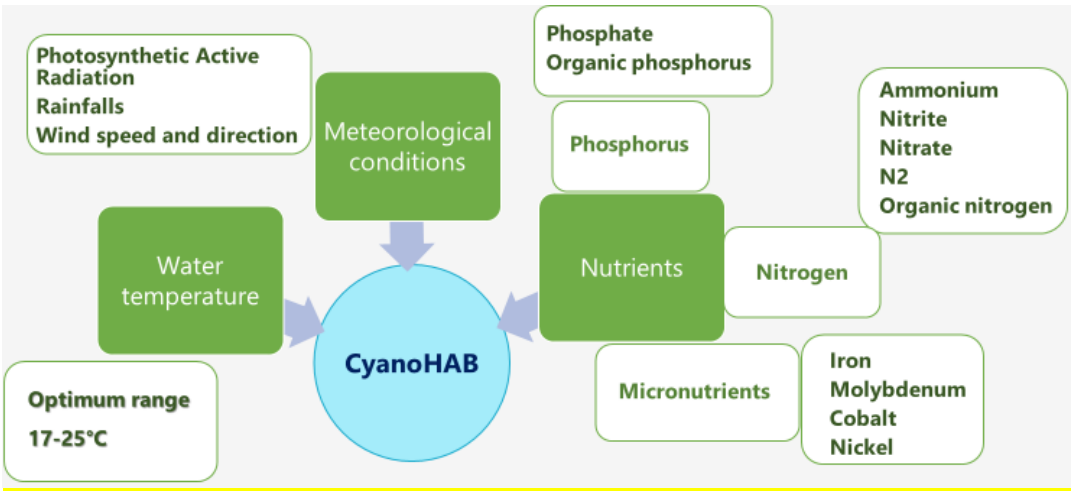


Figure 2.2: Main factors in cyanoHAB development

2.3 Damaging Consequences from CyanoHAB

The extensive growth of cyanobacteria creates many problems for waterbodies, starting such as the uncomfortable smell, anoxia and toxins released in the water. Toxins are considered to be the most serious issue. There are two main categories of toxins produced by cyanobacteria: hepatotoxins and neurotoxins (Figure 2.3). Hepatotoxins are the toxins that damage the liver by the deformation of cells and following bleeding (in case of microcystin) (Camacho, 2008) or by their involve in protein synthesis, such as accumulation of fat in liver (in case of cylindrospermopsin) (Kokocinski et al., 2017). Neurotoxins have a small ring structure and are considered to be more toxic (Camacho, 2008) damaging the nervous system leading to paralysis and respiratory failure (McLellan

and Manderville, 2017). All cyanotoxins are intracellular, that means that they can be found in the water only in case of cell lyses (Benayache et al., 2019).

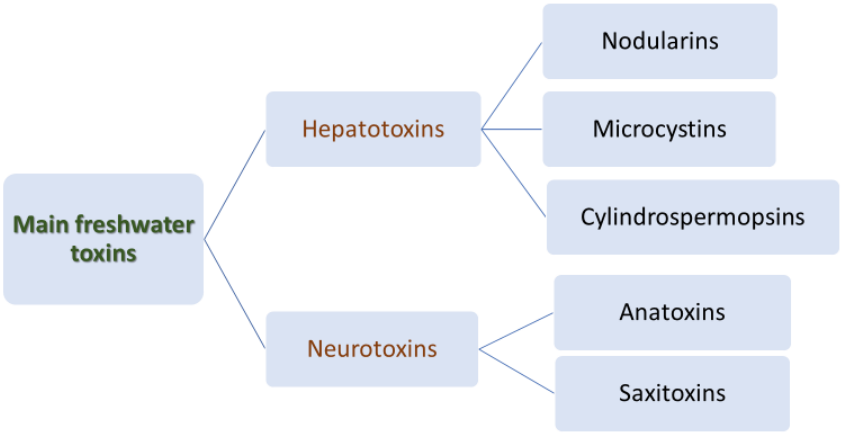


Figure 2.3: Freshwater toxins classification

Microcystin is the most common toxin that can be found during bloom episodes of cyanobacteria in freshwater. *Microcystis, Dolichospermum (Anabaena), Planktothrix, Aphanizomenon* are genus that can be associated with the microcystin production in the water. Microcystin is a cyclic heptapeptide with seven amino acids and molecular weights in the range of 890-1400 Da. The chemical properties depend on the amino acids present in the structure (Figure 2.4). It is "cyclo-(-D-Ala1-X2-D-Masp3-Z4-Adda5-D-Glu6-Mdha7-) where X and Z are variable L-amino acids, D-Masp is D-erythro-β-methylaspartic acid, Mdha is N-methyldehydroalanine and Adda is (2S,3S,8S,9S)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid" (Miles and Stirling, 2017). The most variability is happening in position 2 and 4. The more toxic variants are more hydrophilic (the example is MC-LR) (McLellan and Manderville, 2017). As a result of the structure, it can not go directly across membranes into cells and, therefore, it affects on the liver that transfer organic ions in the body (WHOb, 2003). There are around 250

different variants of microcystin which have been discovered (Spoof and Catherine, 2017). MC-LR is considered to be the most toxic (Catherine, 2017) and stable in the environment for many days. The half-life of microcystin is around 1 week in ambient conditions (WHO, 2003a). The guideline value for MC-LR in drinking water is 1.5 µg/L (Health Canada, 2014). For the recreational waterbodies, the total microcystin concentration is regulated and suggested to be less than 40 µg/L (Health Canada, 2012).

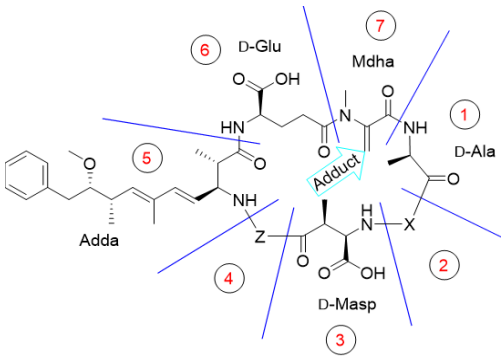


Figure 2.4: The general structure of microcystins (Miles and Stirling, 2017)

Anatoxin is the second group of toxins commonly found in the fresh water of Atlantic Canada (Hushchyna and Nguyen-Quang, 2018). *Dolichospermum flos-aqua* AKA *Anabaena flos-aquae*, *Dolichospermum planctonicum* AKA *Anabaena planctonica*, *Oscillatoria* sp., *Aphanizomenon* sp. and *Cylindrospermum* sp. produce this toxin (Chorus and Bartram, 1999). Some common analogs of anatoxins are in Figure 2.5. The molecular weight of anatoxins 166-210 Da. Anatoxin-a ($C_{10}H_{15}NO$) is the most common representative of this group. It is an alkaloid-like neurotoxin that is not stable in the environment for a long time. The half-life of this toxin is some hour under high pH and

extensive light (Smith and Sutton, 1993). The sub products of anatoxin-a are dihydro- and epoxy-anatoxins which are considered much less toxic (Bruno et al., 2017).

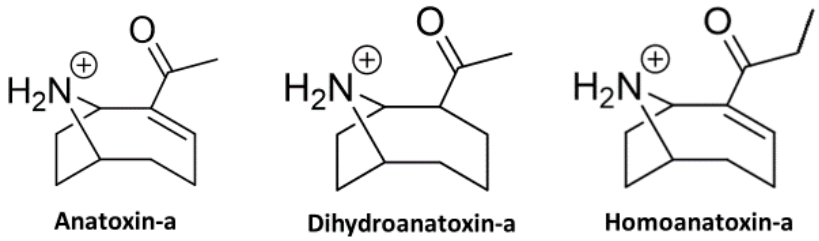


Figure 2.5: The most common analogs of anatoxins

The consequences of the toxins released during and after cyanobacterial bloom episodes are summarized in Figure 2.6. Released toxins in the water will directly influence human beings due to the consumption of water or food containing toxins. Economic losses in the fishery industry and recreational sector can occur. Water containing even small amount of toxins could be poisonous for wild and domestic animals as well as children if they swallow the water during their bath or swimming.

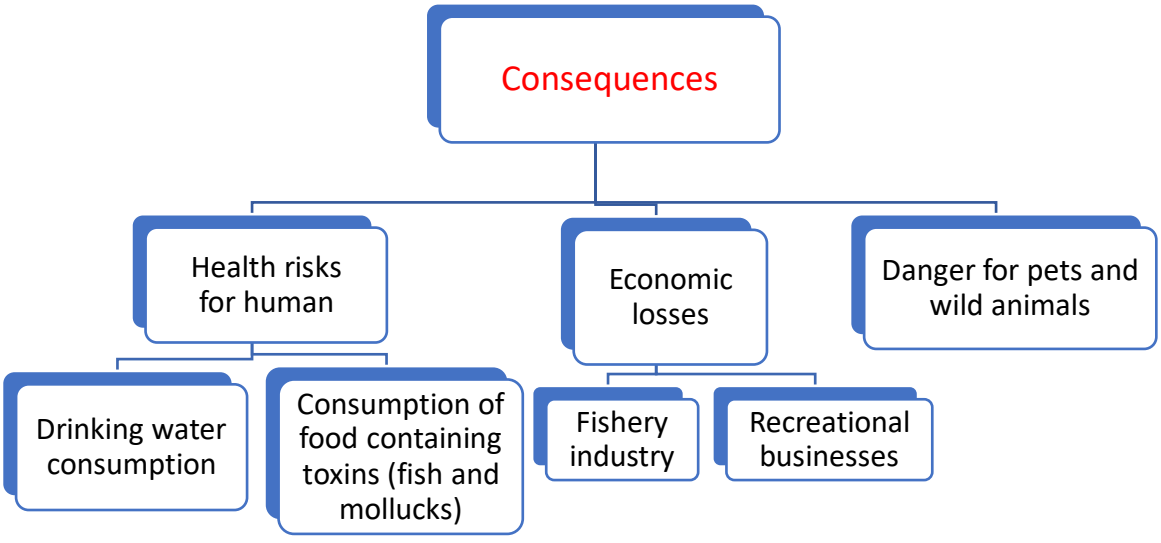


Figure 2.6: Consequences of toxins released in the waterbody

2.4 Eutrophication versus Cyanobacterial Growth

Historically bloom development was discussed and estimated based on indexes describing the eutrophication of the water. The classification of freshwater bodies according to their eutrophic status is one of the basis to estimate the productivity of the water systems. Usually, there are three categories: *oligotrophic* < *mesotrophic* < *eutrophic*. This classification is based on single factors such as dissolved oxygen (DO) (Wetzel, 2001), total phosphorus (TP) (CCME, 2004) and parameters such as chl-a, Secchi disk depth (Brylinsky, 2009; Carlson, 1977; Novotny & Olem, 1994; Serwan and Baban, 1996). The simplest index based only on one factor can be the TP level, chl-a, or cyanobacteria cell counts, suggested by the World Health Organization (CCME, 2004; WHOb, 2003) as risk factor for cyanobacterial bloom development. In fact, chl-a is not a very useful and successful parameter to detect the cyanobacterial presence in the waterbody, because it represents all green pigments generated by microalgae and cyanobacteria in the water body without any distinction. This is the reason why WHO suggested in case of the high chl-a concentrations, measure toxins concentration must be measured in order to conclude about the cyanobacterial danger. This process is certainly longer and relatively expensive. Another pigment could be effectively used to deal with cyanobacterial determination: phycocyanin (PC). This is the accessory pigment present in light harvesting antenna of the photosynthesis system known as PSII. This pigment allows us to distinguish between cyanobacterial and other phytoplankton blooms and can be used in the model combined with other important environmental factors for cyanobacterial growth. However, research on PC relating to the risk of water quality are rare (Ahn et al., 2007; Brient et al. 2008). Moreover, PC is not currently used in any official guidelines of Canada.

Other similar indexes (based one or several coupled parameters) were also used by many researchers including Brylinksy (2009) and Novotny and Olem (1994) to define a lake's trophic status. One common index in many studies is the mass ratio of total nitrogen (TN) to TP (Ndong et al., 2014). Estimation of a phosphorus load as a characteristic of trophic state was suggested for fresh waterbodies by Vollenweider (1975). Unfortunately, this index doesn't count internal load from the sediments hence it is not very accurate.

There are several indexes that have been created for inland water: Uhlmann/Verduin' Index (using carbon, nitrogen, and phosphorus), Carlson's Index (using Chl-a, TP, and Secchi disk values or transparency), Schroders Correlation Model (using a variety of nutritional and environmental factors), and the OECD Classification (using TP, average chl-a, peak chlorophyll, and transparency measured by Secchi disk) (Vollenweider, 1998). Carlson's index is commonly used by water managers, but it is not a very good tool as values of Secchi disk depends on many factors including the colour of water. It could be certainly overestimated as many lakes in Canada have a dark water colour.

Vollenweider et al. (1998) suggested an index named TRIX to determine the trophic state for waterbody of a coastal marine zone. This index was tested in different marine zones such as the Adriatic Sea, Aegean Sea, Black Sea, Tyrrhenian Sea, and north-east of the Persian Gulf. This simple combination of factors that include productivity of an ecosystem (chl-a, deviation of oxygen saturation from 100%) and nutrients (total or dissolved form of nitrogen and phosphorus) under a logarithmic format makes it easy in use. There are very few studies or even no significant works using this index for freshwater research in lakes. Only one study on landscape lakes in the North China (Chen et al., 2013) used the TRIX results from Vollenweider et al. (1998) for fresh water

sources. However, the TRIX concept, in my opinion, is reasonable to be employed for the freshwater resources as it is the combination of key biological and hydro-chemical parameters in a 'simple' relationship without specific characteristics of marine environment.

To sum up, there is still no accurate and precise indicator for the prediction of cyanobacterial growth. Moreover, no work has been done to determine the threshold that can estimate the cyanoHAB onset.

2.5 Discrimination Test versus the Threshold Definition

Discrimination tests is used for determination whether two samples are actually different (Amerine et al., 1965; Meilgaard et al., 1991; Peryam, 1958; Stone and Sidel, 1993). Discrimination tests, also known as threshold tests, are a useful method for detecting bias in decisions and for attempting to obtain the goodness of fit between model and data. In statistics, there are some different discrimination tests including Pearson Chi-Squared test, Classification tables, Hosmer-Lemeshow tests, and Receiver Operating Characteristic (ROC).

The ROC analysis is classified as a well-known binary discriminator test which assesses the predictive power of a binary classification system to evaluate a model in a decision-making process. This test, originally from a research domain coined "signal detection theory", was developed during World War II for the analysis of radar images. It was recognized as useful tool for interpreting medical test results and in many other fields as a method for evaluating the accuracy of analyses (Lerman et al., 2010). These techniques are now widely used in a number of fields, particularly medical research. For algal bloom

research, there were few pioneering works on the ROC analysis (Allen et al., 2008; Anderson et al., 2016; Elkadiri et al., 2016; Hamilton et al., 2009). However, these works were primarily focused on the evaluation of the predictive models coupled with satellite imaging process.

For a detailed tutorial of the use of ROC curves and related metrics, refer to Brown and Davis (2006). Some representative features of the method while respecting the nomenclature of Brown and Davis (2006) are: the decision can be correctly positive (CP), correctly negative (CN), incorrectly positive (IP) or incorrectly negative (IN) (Figure 2.7a).

If a threshold criterion for data (T_D) is set, divide the data into two groups, and then compare it with the model using the same threshold for model (T_M), we can assess model–data similarity at that threshold, effectively assessing the model ability to discriminate that threshold. A model is said to be perfect when it gives CP and CN outcomes only; the more dispersed the model/data relationship, the more IP and IN conditions will occur and the worse the model performance. If we let T_D vary with T_M, a non-parametric measure of the model's ability to simulate a given variable will be obtained, which can be compared directly with other simulated variables.

The correct negative fraction (CNF) and correct positive fraction (CPF) as defined in (1) will further assess the decision process.

$$CNF = \frac{CN}{CN+IP} \text{ and } CPF = \frac{CP}{CP+IN} \quad (1)$$

These fractions are independent of the actual numbers of positive and negative observations in the trials and express *the rate* of negative and positive observations correctly determined. A curve illustrating the model performance can then be determined by plotting CPF on the vertical axis and (1−CNF) on the horizontal axis (Figure 2.7b). CPF, or sensitivity, is the probability that case X classified correctly as above the threshold. CNF (or specificity) appearing in the term (1−CNF), which is the probability that X classified correctly as below the threshold.

The perfect model corresponds to a point in the top left hand corner of the Y axis (i.e. CNF=CPF=1), the top right (CPF=1, CNF=0) and bottom left (CPF =0 and CNF=1) of the diagram correspond to the extremes of the decision process where every trial is always deemed either positive or negative. A completely random predictor (CP=IP and CN=IN) gives a straight line CPF=1-CNF (X=Y, line of equality or random change). This can be explained by the fact that the threshold rises equal numbers of correct and incorrect positives occur. Results below this line are incorrectly predictive results.

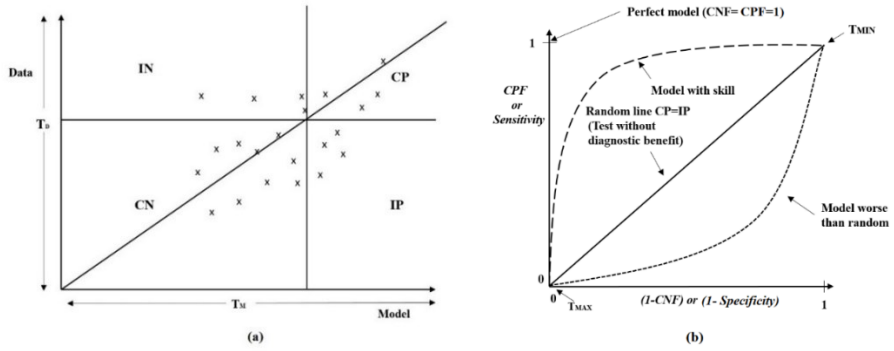


Figure 2.7: Schematic diagrams of a) discrimination analysis and b) binary discrimination skill assessment curves (Adopted from Stow et al., 2009)

Hence, ROC curves and assessments are displayed as in Figure 2.8.

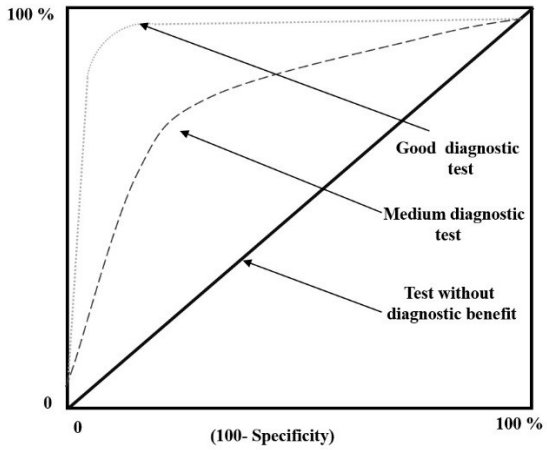


Figure 2.8: Schematic diagram of ROC analysis

Consequently, the definition of the area under the ROC curve (AUC) was introduced as a criterion to evaluate the overall performance of the discrimination test. AUC is the percentage of randomly drawn pairs for which this is true (that is, the test correctly classifies two samples in the random pair and considered as a global estimate of diagnostic accuracy). AUC may take values ranging from 0.5 (no discrimination) to 1 (perfect discrimination). A rough practical guide for evaluating the accuracy of a discrimination test with the AUC criteria described as in Table 2.1 (Carter et al., 2016).

Table 2.1: The AUC criteria to evaluate the accuracy of diagnostic test

AUC value	0.9-1.0	0.8-0.9	0.7-0.8	0.6-0.7	0.5-0.6
Evaluation	excellent (A)	good (B)	fair (C)	poor (D)	fail (F)

Another factor which can be used to estimate the effectiveness of the discrimination test is the Youden index ***J***. The Youden index ***J*** (Youden, 1950) is defined as:

$$\boldsymbol{J} = \max \{ \text{sensitivity}_c + \text{specificity}_c - 1 \} \quad (2)$$

where c ranges over all possible criterion values.

The Youden index ***J*** is commonly used to measure overall diagnostic effectiveness (Schisterman et al., 2005). Moreover, ***J*** is the maximum vertical distance or difference between the ROC curve and the diagonal or chance line. ***J*** was originally used to evaluate the test effectiveness and ranges between 0 and 1. When ***J*** values are close to 1, it indicates that the effectiveness is relatively good, while values close to 0 indicate limited effectiveness.

The method above is an easy and effective tool in the evaluation of defining of the thresholds of the binary data. Our details of calculations to apply this ROC analysis will be presented in methodology and in results sections.

CHAPTER 3 METHODOLOGY

The methodology for this research includes two categories: 1) the field sampling and lab measurements of relevant physical (temperature of water, in situ pH using YSI probe) and hydro-chemical factors (in situ dissolved oxygen DO and nutrients in the laboratory), as well as biological parameters (chl-a, PC, composition of species, toxins); and 2) mathematical model via a log formulation (TRINDEX) and Receiver Operating Curve (ROC) analyses to deal with the combining effects of the environment and complexity of collected datasets and thresholds of cyanobacterial bloom.

Data were collected during the period 2015-2018 from May to November, mainly at two locations Mattatall Lake (Colchester County, NS) and Torment Lake (Kings County, NS) for the threshold and TRINDEX study. Related to the cyanotoxin research, other locations in NS and NB were also sampled and described in the next part.

3.1 Field Locations

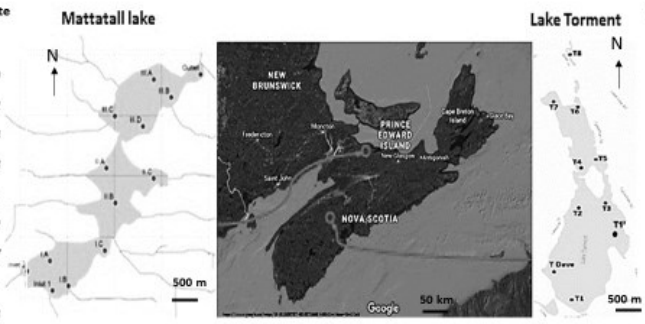
3.1.1 Mattatall Lake

Mattatall lake (ML) is approximately 5 kilometers in length, with the depth varying from approximately 1 to 10m, and an average depth of 4.5m. ML is mainly spring fed with some brooks. There is an outlet from the lake that drains into the French River. Human activities include a blueberry field and a forested area that has been clear-cut on the South-East part of the lake. With the data from two years (2015-2016), ML showed a moderately eutrophic level and contained potential toxic algal species (Hushchyna and Nguyen-Quang, 2017). The eutrophic status was explained by high level of mineral nitrogen in the beginning of summer 2015-2016 that led to a bloom of green algae

(*Mougeotia sp.)* in the middle of summer followed by a cyanobacterial bloom (*Dolichospermum planctonicum*) at late summer-autumn. Blooms followed an increase in the TP level.

Starting in June of 2015, samples were taken frequently (monthly or bi-weekly depending on weather conditions) until November. The samples were taken at locations that were determined before the trip and remained constant throughout each field trip at the surface and bottom levels (Figure 3.1). Additional samplings were done at the locations where blooms were noted. In total the number of samples for an individual parameter varied from 354 to 413. The lake covers different trophic conditions depending on the season which provides a good perspective for evaluation of new index.

Locatio n	N-Coordinate	W-Coordinate
IA	*45.683914*	*-63.483759*
IB	*45.681845*	*-63.480922*
IC	*45.684767*	*-63.475384*
IIA	*45.691568*	*-63.475148*
IIB	*45.688689*	*-63.47378*
IIC	*45.690711*	*-63.467932*
IIIA	*45.698875*	*-63.467907*
IIIB	*45.697408*	*-63.46526*
IIIC	*45.695859*	*-63.473843*
IIID	*45.695011*	*-63.469584*
Inlet 1	*45.681467*	*-63.483369*
Inlet 2	*45.683072*	*-63.487229*
Outlet	*45.699303*	*-63.460799*



Location	N-Coordinate	W-Coordinate
T1	*44.71798*	*-064.74467*
T1'	*44.72619*	*-064.73376*
T2	*44.72710*	*-064.74387*
T3	*44.72715*	*-064.73530*
T4	*44.73933*	*-064.74036*
T5	*44.74369*	*-064.73926*
T6	*44.74960*	*-064.74142*
T7	*44.75008*	*-064.74615*
T8	*44.75116*	*-064.74110*
T Dave	*44.72186*	*-064.74819*

Figure 3.1: Location of the tested lakes on the map (source: modified from Google map)

3.1.2 Torment Lake

Torment Lake (TL) is in East Dalhousie, Kings County (NS, Canada) in the upper LaHave River watershed system. There is a chain of three interconnected lakes in this system from North to South: Chain Lake which drains into Armstrong Lake, which in

turn drains into Torment Lake. The latter is a shallow lake (with the maximum depth around 21 feet) with rocky floor and low sediment accumulation on the North. In the southern section, the deeper bottom configuration has a thin sediment layer. The lake is used for residential and recreational purposes where fishery, boating and swimming are main recreational activities. It is almost surrounded by a forest with one Christmas tree farm nearby. The lake is dystrophic with a brown water, low pH, low carbonate level and high organic content (Marty and Ritchie, 2015).

Algal blooms in TL were firstly and seriously noticed in 2014. However, there was no previous study as well as a monitoring program about water quality in this watershed. During 2014-2015, some nutrients were measured but there was no data connecting to biological aspects, except for some chlorophyll-a values irregularly monitored (once a month at one location for 6 months) as a part of Kings County lake water quality monitoring program (Marty and Ritchie, 2015; Marty and Reardon, 2016).

BBML initiated its monitoring program in this lake and its data from 2015-2018 will be used for the model validation with a total number of sample equal 170.

3.1.3 Other Waterbodies

There were 10 different lakes sampled during 2015-2018 on a regular basis annually, presented in Figure 3.2 below. All these lakes will be used for the discussion of cyanobacteria species identification, responsible for blooms and their released toxins. The following paragraph describes three other main areas where waterbodies and their

watersheds have been investigated (additional to King County with Torment Lake and Colchester-Cumberland Counties with Mattatall Lake as above mentioned).

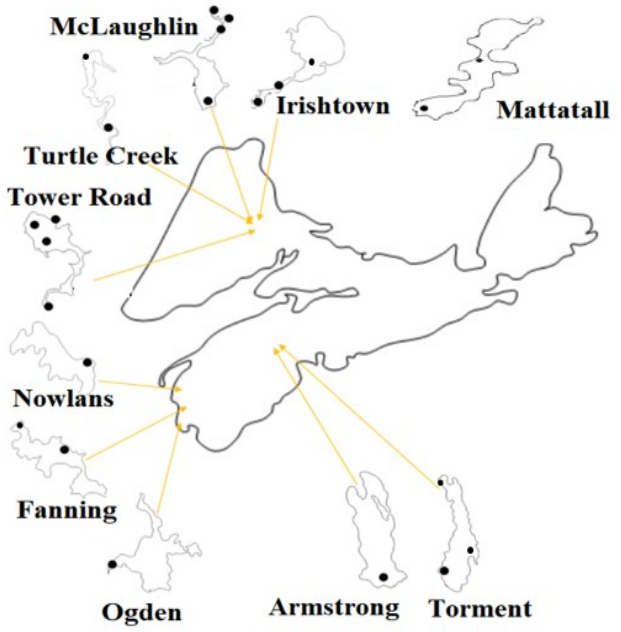


Figure 3.2: Sampling locations for taxonomy and toxins in 2015-2018

Yarmouth County is the Southeastern region of Nova Scotia, where cyanobacterial blooms became yearly events in the summer season from 2007 (Taylor, 2009). Three lakes within the Carleton river watershed were sampled during 2016-2018: Ogden, Fanning, and Vaughan (Figure 3.2). These lakes are used for recreational purpose with many cottages in the area. The lakes show interconnections between them: Ogden draining into Fanning, and then Fanning draining into Raynards; from the latter, water goes to Vaughan. Around Ogden Lake, there are residential and forestry activities. Lake Fanning has similar human activities on its watershed as found in Ogden, but there are additionally some mink farms. Vaughan Lake watershed is a mixture of residential,

forest, and mink farm activities on the Southern part of the lake (Stantec Consulting, 2017).

Digby County in NS is a neighbouring county to Yarmouth. Nowlans Lake is within the Meteghan river watershed. It has a single inlet on the East side and drains to Prime Lake on the west. The lake is very shallow with muddy shoreline. Most of the surrounded area is forest, but there are some residential areas and mink farms (Stantec Consulting, 2017).

In New Brunswick we investigated reservoirs in the Great Moncton area, including McLaughlin, Irishtown, Turtle Creek and Tower Road.

McLaughlin reservoir was used as the main drinking water supply for Moncton in the past. Nowadays it is a backup reservoir while Turtle Creek is the city Moncton's main source of drinking water. The McLaughlin reservoir is surrounded by forest, and all its territory is fenced for its protection. There are no activities on the land surrounding the reservoir and no residents live in the area (Nguyen-Quang et al., 2018).

Irishtown Reservoir, located in the heart of Moncton city, is used for recreational purposes such as kayaking, walking and other park activities. Blooms were observed in 2015 and reported by Moncton authorities earlier. No cyanobacterial bloom was detected in 2016-2017, however, there were high level of cyanobacteria detected in its water samples (Nguyen-Quang et al., 2018). CyanoHAB was registered in July 2018 (Nguyen-Quang et al., 2019).

Turtle Creek reservoir is the drinking water supply reservoir of the city. On the upper side there is a Tower Road reservoir. The latter is the new artificial waterbody (less than 5 years) that connects to Turtle Creek by dam with several gates. Before 2017 there were no

blooms identified in either reservoirs. In September 2017 a huge bloom spread over the Tower Road reservoir and it lasted approximately 1 month (Nguyen-Quang et al., 2019).

3.2 Sample Collection Method

The following is the list of equipment used in the field for data collection:

1. GPS etrex 10 (Garmin, USA) fixes the point's location;
2. Portable depth sounder (Speedtech, USA) detects the depth of the sample location;
3. YSI probe with 3 sensors (pH; dissolved oxygen (DO) in mg/L and %; conductivity and temperature) the length of the cable 4/20m (Professional Plus, Hoskin scientific LTD, USA);
4. Van Dorn 2L bottle was used for water sample collection at depth;
5. Portable sampler 6700 (ISCO, NE, USA)

Water samples for different nutrient analyses were collected in duplicate as recommended by Canadian Council of Ministers of the Environment (2011) from the surface (0.5m) and bottom (1m above the bottom) in clean HDPE bottles. If the location was too shallow (less than 3 m) the water is sampled only from the surface level. All samples were stored in cool conditions until being taken to the laboratory. In case of a bloom or colouration of water, taxonomic samples were taken and fixed with acidic Lugol's solution (Edler and Elbrächter, 2010). Related to cyanotoxin analyses, samples for microcystins were directly kept in PETG tubes while samples for anatoxin-a were preserved with a diluent at 1:10 proportion in PETG tubes to slow down the degradation of anatoxin-a which is known as a very rapid process (Chorus and Bertram, 1999). All samples for toxin analyses were immediately frozen at -20°C in the laboratory and stored until the day of measurement.

3.3 Laboratory Methods

This section will summarize four types of lab analyses that this research has involved, including 1) Nutrient analyses; 2) Pigment analyses; 3) Taxonomy identification; and 4) Cyanotoxin detection and measurements. More specifically, cyanotoxin detection and measurement were processed via three different approaches to cross-validate each other, including the direct and indirect ELISA tests (ELISA ADDA(DM) and ELISA multihapten) and LC-HRMS (ELISA stands for Enzyme-Linked Immunosorbent Assay and LC-HRMS is the acronym of liquid chromatography coupled with high resolution mass spectrometry).

Depending on the type of analyses, samples were processed at 3 different laboratories (BBML, Truro: all nutrients, pigments, direct and indirect ELISA tests; Department of Biology, University of Montreal: taxonomy identification; and NRC, Halifax: LC-MS analyses of toxins) and the following is the specific equipment used for analyses:

a) Filtration system with vacuum pump (vacuum up to 70 cm Hg, pressure up to 10 kg/cm^2)
b) Sonic Dismembrator System (Fisher Scientific, Model C-18, USA)
c) Centrifuge 5430 (Eppendorf, Germany)
d) Fluorometer Turner Designs 10AU (Hoskin scientific LTD, USA)
e) Photometer (YSI 9300) (Hoskin scientific LTD, USA)
f) Multi-scanner FC (Thermo Scientific, USA)
g) Inverted Research Microscope “Olympus IMT-2”
h) Q Exactive-HF Orbitrap mass spectrometer (ThermoFisher Scientific, Waltham, MA, USA) with an Agilent 1200 LC system (Agilent, Santa Clara, CA, USA)

i) Centrifuge Micro 6 (Fisher Scientific, USA)

j) Ultrasonic cleaner Branson 1510 (Emerson Industrial Automation, Danbury, CT, USA)

3.3.1 Nutrient Analyses

Nutrients (nitrates+nitrites and phosphates) were measured in duplicate using photometer and tablet reagent system (YSI, 2010). Results are represented in the unit of mg/L. The calibration of the photometer was processed before every measurement using photometer colour standards (Hoskin scientific LTD, USA).

The principle of photometer is based on the Beer-Lambert Law: the intensity of colour in the solution is proportional to the concentration of nutrient (Fleming and Williams, 1966). The basis of nutrients analyses (YSI, 2010) is:

a) Phosphate was analysed by the reaction between ammonium molybdate and phosphate in the acid condition. Formation of phosphomolybdic acid followed its reduction by ascorbic acid and a change of the colour of the solution to blue. The range of measurements is from 0 to 4.0 mg/L with limit of detection (LOD) equals 0.01 mg/L;

b) Nitrate+nitrite method is based on the reduction of nitrate to nitrite under the reaction with zinc-powder. The solution was filtered through 0.2 μm hydrophilic polyethersulfone membrane filters to remove all fine particles. The resulting sum of reduced nitrate and nitrite was determined by their interaction with the sulphanilic acid and creation of diazo compounds. The reaction between diazo

compound and N-(1-naphthyl)-ethylene diamine produces the red colouring of the solution. The range of measurements is 0-1.0 mg/L with LOD equals 0.01 mg/L.

3.3.2 Pigments Analyses

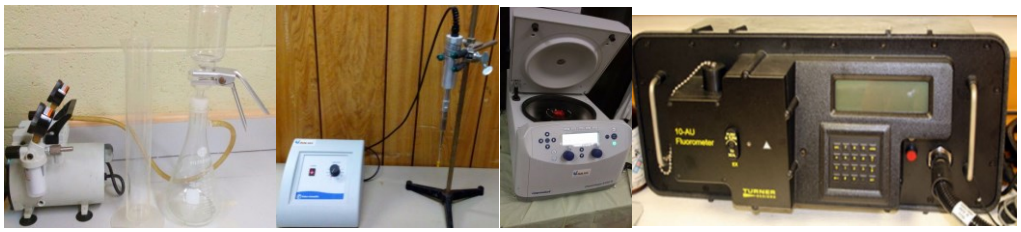
The following are different reagents have been used during pigment analyses:

a) 90% acetone: 100 mL of distilled water was added to 900 mL of acetone ACS grade (95.5%; VWR Chemicals BDH, USA). Solution was kept at room temperature.

b) phosphate buffered saline (10xPBS): 80.0 g of sodium chloride ACS grade (VWR Chemicals BDH, USA); 2.05 g of potassium chloride Lab grade (Anachemistry, Germany); 2.00 g of potassium dihydrogenphosphate ACS grade (Anachemistry, Germany) and 11.50 g of di-sodium hydrogenphosphate anhydrous ACS grade (VWR, Life Science, USA) were mixed and add to the volume of 1 L. This solution, kept in the refrigerator, was diluted 10 times prior to use.

The technique was as follows. For the detection of chl-a and PC the water was filtered through the Whatman glass microfiber filters GF/A (Whatman, GE Healthcare, Life Sciences, UK) with 1.6 µm pore diameter in green light conditions during the first hour after field trip. The volume of water to be filtered depends on the phytoplankton content and varies from some milliliters (in case of the bloom) to 300 mL (when the water being without visual presence of biomass). All filtration was done in duplicate for every type of pigment. The filters were wrapped in the foil and stored in the -80°C until the day of analyses. On the day of analyses, the samples were defrosting, filters with 5 mL of 90% acetone (for chl-a) or 5 mL of phosphate buffered saline (for PC) were sonicated for 30 sec in 40 amplitude. Additionally, 5mL of 90% acetone or PBS were added. Two steps of

centrifugation in 3500*g* at room temperature for 10 minutes and 13000*g* at 4˚C for 1.5 hours to purify the extraction were followed. The supernatant was diluted to 10 mL and analysed by Fluorometer Turner Designs 10AU (Hoskin Scientific LTD, USA). The excitation wavelength was 436nm with emission of light at 680nm for chl-a. For PC, excitation wavelength was 600nm with emission at 640nm. Different steps of the detection of chl-a and PC are represented in Figure 3.3.



a) filtration → b) sonication → c) centrifugation → d) measurements

Figure 3.3: Steps in chlorophyll-a and phycocyanin analyses

The calibration standards of both pigments provided by Beagle Bioproduct Inc. (USA). The LOD for chl-a is 0.05 µg/L; and for PC is 0.04 µg/L.

Concentration of chlorophyll-a and phycocyanin were calculated by the following formula:

$$C = \frac{XX \times V4\ (mL) \times V2\ (mL)}{V3\ (mL) \times V1\ (mL)} \quad (3)$$

where C – concentration of chlorophyll-a/phycocyanin, µg/L;

XX –result received from the fluorometer, µg/L;

V1 – initially filtered water, mL;

V2 –volume of extractor (90% acetone or PBS) at the first step, mL (V2=10 mL);

V3 –volume of supernatant used after centrifugation, mL;

V4 –total volume of measured sample, mL, V4=10 mL.

3.3.3 Taxonomy Identification

The identification of species was processed at the Department of Biology of University of Quebec in Montreal (UQAM), following the inverted microscope method by Lund et al. (1958). This method was modified and developed by Utermohl in 1958 (Utermohl, 1958) for the inverted microscope and saved time without sacrificing the degree of accuracy and statistical validity of the method. Depending on the density of a sample, the necessary volume was sedimented for approximately 18 hours in the counting chamber of 10mL (Hydro-Bios counting chambers). Examination under the inverted microscope (Olympus IMT-2) is followed where fifteen fields, chosen at random (Z-course), were counted at 40X magnification and at 10X magnification (cross). Identification of the species was done using identification documents (Bellinger and Sigee, 2015; Freshwater Algae of North America Ecology and Classification, 2003; Prescott, 1962) and was based on various criteria, such as: form of colonies, form of cells, terminal structures, cell contents, cell wall form and structures, presence or absence of sheaths, cellular extensions, etc.

The prevalent species which generated blooms were determined. The origin of blooms in NS and Eastern part of NB as well as types of cyanobacterial species causing HAB in freshwater is still not well known. For NS, there is not enough historical data to sketch the distribution of cyanobacterial species in the province. Therefore, data analyzed from ten

waterbodies from 2015 to 2018 was the basis for the creation of cyanobacterial species distribution map in using the GIS software.

3.3.4 Analyses of Cyanobacterial Toxins

Two main categories of freshwater toxins were analyzed from samples 2016-2018, including microcystins and anatoxins (Benayache et al., 2019). The methods used to detect toxins in the water are in Figure 3.4.

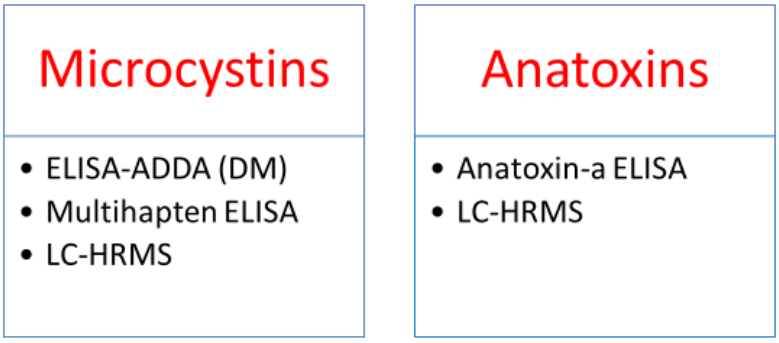


Figure 3.4: Methods of toxin detection

In the context of this thesis, for the detection of microcystins, ELISA ADDA(DM) test was used in 2016-2018. ELISA multihapten and LC-HRMS tests were processed in 2017-2018 in addition. For anatoxin analyses, anatoxin-a ELISA (Abraxis) was used in 2016-2017, and LC-HRMS was used in 2017-2018.

Total microcystins detected by 3 different methods previously cited were compared by using the Pearson coefficient of correlation (Witte and Witte, 2010) to cross-validate obtained values by different approaches.

3.3.4.1 *ELISA Kits from Abraxis (Warminster, PA, USA). Direct ELISA Test*

The method of detection of toxins is ELISA ADDA(DM) with the test kit is provided by Abraxis (Warminster, PA, USA). Three freeze-thaw cycles were applied to all considered samples before measurement to lyse the cells for detecting the total amount of toxins. Samples were filtered prior to analyses (glass fiber filters GF/C (Whatman, GE Healthcare, Life Sciences, UK) with size of pore 1.2 μm. All measurements were processed accordingly to the Abraxis procedure.

The process of direct ELISA test is summarized in Figure 3.5. It is fundamentally based on the competition of binding for toxins and enzyme conjugate (microcystin or anatoxin-a) to the anti-microcystin (or anti-anatoxin) antibodies in solution, and then binding to the goat anti-mouse antibody which was used for coating the wells. The microplate consists of 96 wells. This requires hence 38 samples for analysis in duplicate. Substrate solution contributes to the colour development and the colour intensity, representing the toxin level, is inversely proportional to the concentration of microcystin or anatoxin-a. Final measurements were read via the Multi-scanner FC at 450nm (Thermo Scientific, USA) which is shown in Figure 3.4. The quantified concentration is based on MC-LR/anatoxin-a standards and 4-parameter evaluation. LOD is 0.15 μg/L.

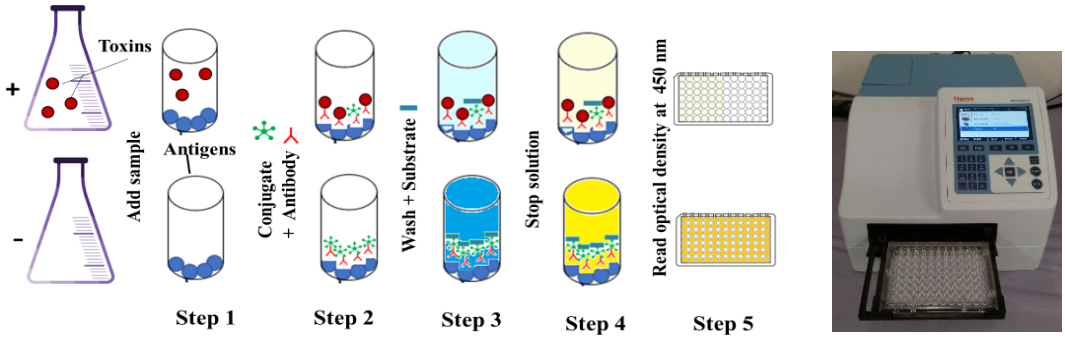


Figure 3.5: The general scheme of toxin analyses by direct competitive ELISA

The microcystins ELISA ADDA(DM) kit can be used to analyse the total amount of microcystin and nodularin in the same time. The principle is based on the recognition of the ADDA branch of the toxins which is present in nodularin as well. One of the disadvantages for this approach is that non-toxic ADDA appearing in the water after the microbial degradation of toxin can be also detected, i.e. final results could be overestimated (Harada et al., 2004).

The ELISA kit for anatoxin-a measured only anatoxin-a content in the water samples. Advantages of both methods include the easy preparation step and the use of water samples directly from the waterbody.

3.3.4.2 ELISA Multihapten (Norwegian Veterinary Institute, Norway). Indirect ELISA Test

A new approach for microcystin detection (Toxicology Research Group, Norwegian Veterinary Institute, Norway) was found and used to compare this method with ELISA ADDA(DM) and LC-HRMS. Three freeze-thaw cycles were used to lyse the cells. The reagents used in this method were as follows:

a) phosphate buffered saline (10xPBS) being prepared the same as for phycocyanin analyses (See 3.3.2. Pigments analyses). Stock solution (10xPBS) was diluted 10 times prior to use.

b) PBST: 0.5 mL of Tween 20 (Sigma Ultra, Germany) being added to 1L of PBS.

c) blocking/antibody buffer (1%PVP-PBST): 5 mL of 10% PVP stock (Veterinary Institute, Norway) being mixed with 45 mL of PBST before using.

d) sample/standard buffer: S/S buffer (10% methanol in PBS with 0.05%Tween 20): 10mL of 10xPBS stock solution, 80 mL of distilled water (Ica River Spring Water Co.Inc, ON, Canada) and 10 mL of methanol HPLC grade (VWR Chemicals BDH, USA) being mixed. 50 µL of Tween 20 was added.

e) concentrated PBST: to 10 mL of stock solution PBS (10xPBS) 50 µL of Tween 20 was added.

f) 10% sulfuric acid: 20 mL of concentrated sulfuric acid (≥95-98% ACS, FCC, VWR Chemicals BDH, USA) was added to 180 mL of distilled water.

g) K-Blue aqueous TMB substrate (Sigma Aldrich, Canada)

h) sheep anti-microcystin antibody: 30 µL of sheep antibody 80289-5b (Veterinary Institute, Norway) was added to 6 mL of 1%PVP-PBST (See step c) just before using it.

i) second antibody conjugate: 13.8 µL of rabbit anti-sheep lgG (Santa Cruz, Italy) was added to 11mL of 1%PVP-PBST (See step c)

j) 96-wells precoated immunoplates were provided by Veterinary Institute, Norway

k) standard of MC-LR (NRC; Halifax, NS, Canada) with concentration 10.01 µg/mL was used for quantification.

Lake water was filtered through glass microfiber filters GF/C (Whatman, GE Healthcare, Life Sciences, UK) with 1.2 µm pore diameter and adjusted with 100% methanol and salts (concentrated PBST (reagent (e)) to obtain 10% methanol concentration in the total volume.

The method of indirect ELISA test is based on two steps of binding of the primary antibody and labeled secondary antibody (Figure 3.6). The primary antibody is incubated with the antigen. All antibodies which were not reacted are washed away. Enzyme linked second antibody conjugate is added and incubated. All remains, which were not stick to primary antibodies, are washed away. Adding of a substrate (TMB) change the colour of the samples because of the reaction with protein (second antibody conjugate). The intensity of colour is inversely proportional to the concentration of microcystin in the solution.

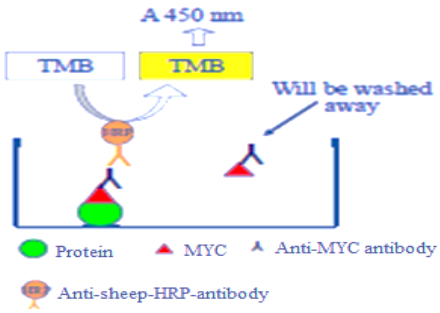


Figure 3.6: Indirect ELISA principle (Samdal at al., 2013)

The details of this test methodology could be found in Samdal et al. (2014).

3.3.4.3 LC-HRMS Analyses

LC-HRMS distinguishes the analogs of microcystin and anatoxin in the water, which can not be done by ELISA. Three freeze-thaw cycles were used to break down the cells and release toxins. The samples were diluted 1:1 with 100% methanol and sonicated in a bath for 30 seconds after that.

500 µL of all samples were filtered at 5.8rpm for 3 minutes using centrifugal filter units (0.45µm) and Centrifuge Micro 6 (Fisher Scientific, USA) and transferred to 1.5mL vials for the following LC-HRMS analyses.

Formic acid with a purity >98%, ACS grade was from EMD, Gibbstown, NJ, USA. All solvents including methanol and acetonitrile were HPLC grade from Caledon (Georgetown, ON, Canada). Deionized water was provided by a Milli-Q water purification system (Millipore Ltd., Billerica, MA, USA).

Q Exactive-HF Orbitrap mass spectrometer with a HESI-II heated electrospray ionization interface (ThermoFisher Scientific, Waltham, MA, USA) combined with an Agilent 1200 LC system (Agilent, Santa Clara, CA, USA) were used. Analyses were performed with a Symmetry Shield C18 column (3.5 µm, 150 × 2.1 mm; Waters, Milford, MA, USA) held at 40°C with mobile phases A and B of deionized water and acetonitrile with 0.1% of formic acid in both. The same equipment, mobile phases and column were used for microcystin and anatoxin detection.

Linear gradient elution for microcystin is in Table 3.1 and for anatoxin is in Table 3.2.

Table 3.1: Gradient of the mobile phases used in the microcystin procedure

Time (min)	Flow (mL/min)	A: Water with 0.1% FA (%)	B: MeCN with 0.1% FA (%)
0.00	0.3	80	20
18.00	0.3	10	90
18.10	0.5	0	100
21.00	0.5	0	100
21.10	0.3	80	20
25.00	0.3	80	20
26.00	0.3	80	20

Table 3.2: Gradient of the mobile phases used in the anatoxin procedure

Time (min)	Flow (mL/min)	Water with 0.1% FA (%)	MeCN with 0.1% FA (%)
0.00	0.2	98	2
2.00	0.2	98	2
15.10	0.2	80	20
16.00	0.3	0	100
20.10	0.3	0	100
21.00	0.2	98	2
25.00	0.2	98	2
26.00	0.2	98	2

FA – formic acid
MeCN – acetonitrile

Injection volume was 1–5 µL for microcystin and 1 µL for anatoxin.

LC-HRMS worked in positive mode collecting data of full scan and data independent acquisition (DIA).

Microcystins were detected from *m/z* 500 - 1400 using 72 *m/z* isolation windows.

Anatoxins were detected from *m/z* 100 – 500 and MS/MS data were collected for suspected ATXs with a 0.7 *m/z* isolation window.

Data processing was through Xcalibur version 2.3 (Thermo Fisher Scientific, San Jose, CA, USA). All microcystins and anatoxins in the lake samples were identified based on

masses' fragments and expected retention time in comparison with standards and the reference material RM-BGA (NRC, Halifax, NS, Canada). Quantification of microcystins and anatoxins was done using certified mix of microcystins (MC-LR: 3.07 µg/mL; dmMC-LR: 2.11 µg/mL; MC-RR: 0.34 µg/mL; nodularin: 2.02 µg/mL, Leu1MC-LY: 1.18 µg/mL) and anatoxin-a (RM-ATX-a: 4.96 µg/mL) from the National Research Council (NRC) Certified Reference Material Program (CRMP; Halifax, NS, Canada). The method LOD in full scan was in average 0.5 µg/L for microcystins and 0.2 µg/L for anatoxins.

3.4 Mathematical Model Conception

Field experiments are necessary to elucidate various factors that affect algal blooms and their proliferation. However, parameters that can be determined in field experiments are limited due to their cost and time-consuming requirements, especially as we can never measure 'all parameters' involved in the pattern for comparison. Moreover, their combining effects on the cyanobacterial growth play a significant role more than each single factor. Therefore, the mathematical model is a good approach to deal with the coupling effects of governing parameters in the bloom occurrence and proliferation. Vollenweider et al. (1998) have developed a model named TRIX to evaluate the trophic level for the marine environment. Our modification and adaptation from Vollenweider et al. (1998) concept are reasonably applied for the freshwater resources as it is the combination of main biological and hydro-chemical parameters in a relationship without specific characteristics of marine environment. Moreover, incorporation of the

discrimination test to the model in order to determine the bloom threshold was first applied.

3.4.1 Threshold Index (TRINDEX)

As most of environmental data are not normally distributed, the square root or logarithmic transformation are the best ways to approximate the random data to the normal distribution form. Logarithmic transformation showed good results in the work of Vollenweider et al. (1998). Our Threshold Index (TRINDEX) is developed in using the transformation from Vollenweider et al. (1998), precisely:

$$X = \frac{k}{n} \sum_{1}^{i=n} \left[\frac{(logM_i - logL_i)}{(logU_i - logL_{i)}} \right] \quad (4)$$

where X – index to be considered;

M_i – measured parameter *i*;

L_i – lower limit (concentration) of the parameter;

U_i – upper limit (concentration) of the parameter;

k – number of degrees of the index;

n – number of variables.

The choice of main parameters used for the threshold determination of cyanobacterial blooms was similar to the concept suggested by Vollenweider et al. (1998), that means:

a) the absolute deviation of oxygen from 100% (D%O) shows the main processes of growth of phytoplankton which can be used for the detection of bloom onset;

b) nitrogen and phosphorus were chosen in the form of nitrate (NO3) and phosphate (PO4) as the main and necessary mineral sources of nutrients for cyanobacteria growth. Moreover, these components can be easily measured on the daily basis;

c) chl-a suggested in Vollenweider et al. (1998) is not perfectly appropriate parameter for the cyanobacterial bloom detection. PC is therefore introduced into this conception as an alternative parameter to reflect the cyanobacterial presence in the phytoplankton community.

To have an appropriate range to cover different trophic conditions, we used limits of detection (LOD) as the lower limit and maximum value obtained in measurements of considered variable for the upper limit. Therefore, the difference between upper and lower limits ($logU_i - logL_i$) was defined by the difference of maximum concentration and LOD for each parameter. The number of degrees *k* was chosen to be 10 as a scaling parameter. The number *n* is 4 in TRINDEX1 when there are 4 variables (PC, D%O, NO3 and PO4 present in the formula) and *n = 5* in TRINDEX2 when there is one more variable namely pigment, chl-a, present. The formulas for TRINDEX1 and TRINDEX2 are as follows.

$$TRINDEX1 = \frac{10}{4} \times [\frac{(logPC - logL_{PC})}{(logU_{PC} - logL_{PC})} + \frac{(logD\%O - logL_{D\%O})}{(logU_{D\%O} - logL_{D\%O})} + \frac{(logPO4 - logL_{PO4})}{(logU_{PO4} - logL_{PO4})} + \frac{(logNO3 - logL_{NO3})}{(logU_{NO3} - logL_{NO3})}] \quad (5)$$

$$TRINDEX2 = \frac{10}{5} \times [\frac{(logPC - logL_{PC})}{(logU_{PC} - logL_{PC})} + \frac{(logChl-a - logL_{Chl-a})}{(logU_{Chl-a} - logL_{Chl-a})} + \frac{(logD\%O - logL_{D\%O})}{(logU_{D\%O} - logL_{D\%O})} + \frac{(logPO4 - logL_{PO4})}{(logU_{PO4} - logL_{PO4})} + \frac{(logNO3 - logL_{NO3})}{(logU_{NO3} - logL_{NO3})}] \quad (6)$$

where PC – phycocyanin, mg/L;

Chl-a – chlorophyll-a, mg/L;

D%O – absolute deviation of oxygen saturation from 100%;

PO4 – concentration of phosphate, mg/L;

NO3 – concentration of nitrate, mg/L.

Our proposed TRINDEX is used for estimation of the risk of cyanobacterial blooms based on the thresholds for bloom occurrence combined nutrient factors and productivity of the system.

3.4.2 Threshold Determination by Discrimination Test ROC

Two main steps for threshold determination in this research can be described as follows: i) statistical process for data analysis of TRINDEX; and ii) capacity of prediction via the threshold being assessed by the ROC analysis.

Precisely, the following sub-steps will be processed to evaluate the threshold using TRINDEX:

1. Calculation of TRINDEX accordingly to formula (5) to (6) for ML and hence deriving the threshold of bloom occurrence.

2. Performance of the discrimination test for 2 datasets ‘bloom’ and ‘no bloom’ by using the ROC curves and AUC (area under the curve), the confidence interval 95%, and the Youden Index *J*.

3.Validation of the discrimination test using data from TL

4. Determination the threshold for bloom

5. Estimation of the TRINDEX effectiveness

6. Application study

All dataset obtained TRINDEX values from ML was divided on two groups: 0 – no bloom, 1 – bloom. The criteria for the bloom onset was chosen based on the work of Brient et al. (2008): 30±3µg/L (or 0.030±0.003 mg/L) of PC that is equal to 20,000 cells/mL and Alert 1 level by WHO (2003b).

The bloom threshold T as the root of a *dichotomous decision* process (term used by Brown and Davis, 2006) which is based on the variable TRINDEX that will drive the outcomes of decision, as positive (yes - bloom) or negative (no - no bloom) as follows:

$$Decision(TRINDEX) = \begin{cases} + & positive, or\ bloom\ occurence\ if\ TRINDEX > T \\ - & negative, or\ no\ bloom \qquad if\ TRINDEX < T \end{cases} \tag{7}$$

Assuming that TRINDEX does indeed have some ability to adequately discriminate between positive and negative situation, there are an infinite number of possible decision thresholds. These include three following possibilities (Figure 3.7): 1) threshold T_1, calling all patterns positive with TRINDEX $\geq T_1$, would correctly identify nearly every positive pattern, although a large proportion of negative patterns would inappropriately be called positive; 2) threshold T_2 more of a balance is struck, as both positive and negative events are missed, and finally 3) T_3 most negative events are correctly identified, but a large proportion of the positive patterns are incorrectly deemed negative as well.

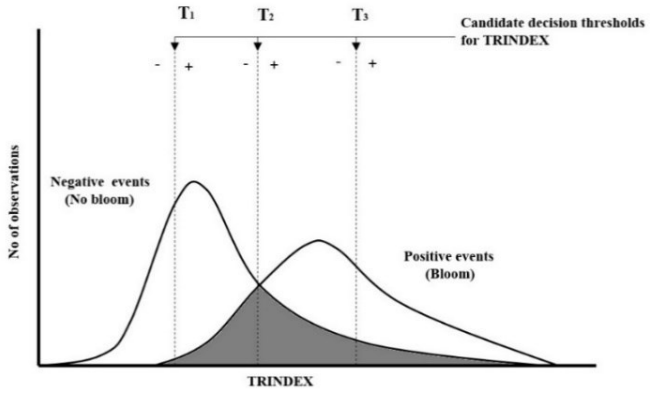


Figure 3.7: Illustration of the occurrence of positive and negative observations versus TRINDEX (adapted from Brown and David, 2006)

Four possible outcomes can result for each trial: correctly positive, correctly negative, incorrectly positive and incorrectly negative. The area measures the ***discrimination***, that is, the ability of the TRINDEX test to correctly classify those *with,* or *without the 'disease'*, i.e *bloom occurrence (yes) or no bloom (no)* respectively, as a binary variable. Therefore, we need the binary discrimination test for assessing the diagnostic accuracy (or discrimination performance) of quantitative tests throughout the whole range of their possible values, and it helps to identify the threshold T.

Graphically, the cut off point was determined by *rnorm* function in R software. The overlapping point of two distribution curves (no bloom and bloom) was considered as the cut off point.

For more precise detection ROC curve analyses was done using MedCalc and EXCEL softwares. All calculations of TRINDEX were rounded at 0.2 unit (eg. 0.2; 0.4; 0.6; etc.). Sensitivity (correctly positive) for every TRINDEX was calculated by assuming that the current and higher values of TRINDEX led to bloom (8). The incorrectly positive was calculated inversely in assuming that the current and higher values of TRINDEX cannot

lead to blooms (9). It is important to note that in following calculations, *sensitivity* and *specificity* were displayed in the percentage (%) instead of the fraction.

$$\textit{Correct positive (Sensitivity)} = \sum \textit{Number of TRINDEX with bloom/ Total number of bloom cases} \quad (8)$$

$$\textit{Incorrect positive (100 - Specificity)} = \sum \textit{Number of TRINDEX with no bloom /Total number of no bloom cases} \quad (9)$$

Assessment of TRINDEX discrimination by AUC and Yonden index were also performed by Medcalc.

All the same steps were done for the validation location (TL) (2015-2018) to serve as an additional way to recheck our approach.

3.4.3 A Case of TRINDEX Application

The effectiveness of the proposed TRINDEX will be discussed in using the samples collected at the same place (Dave's Cove, Lake Torment) from 2016-2019. The location is assigned TDave in Figure 3.1). TL suffers yearly from *Dolichospermum sp.* blooms. Observations for the onset and within bloom episodes were processed bi-weekly or monthly depending on the weather conditions from 2016 to 2019 (in the summer - fall periods). Water samples were taken at every time of observation analyzed and used for calculations of TRINDEX.

TRINDEX was calculated for each collection and discussed in comparison with the real observed bloom conditions. Details of calculations and results are shown in the Appendix B.

CHAPTER 4 RESULTS AND DISCUSSIONS

4.1 Distribution of Cyanobacterial Blooms and Concerned Toxins

4.1.1 Identification of Dominant Cyanobacteria Causing Blooms

The map of cyanoHAB distribution in NS and NB is presented in Figure 4.1. Our observations and study from 2015 have shown that blooms in the region were mostly associated with potentially toxic cyanobacterial species. Some other blooms were formed by green filamentous algae: *Mougeotia* sp. formed huge biomasses in the water column or *Spirogyra* sp. which is green filamentous algae was attached to rocks on the bottom at shallow places. Usually, blooms can be distinguished by scums on the water surface, increasing algal biomass or changing the colour of water. Normally both cyanobacteria and green algae could generate blooms in one place but not at the same moment. However, cyanobacteria species can be found inside of green algae filaments very often.

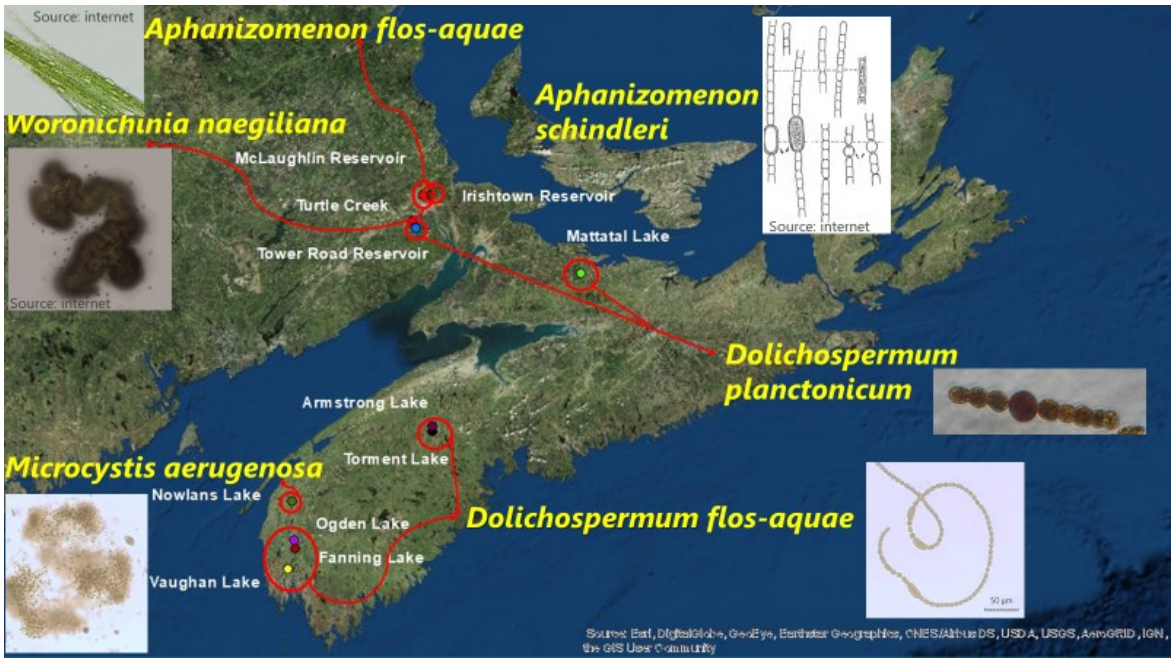


Figure 4.1: Dominant cyanobacteria species in Nova Scotia and New Brunswick

Geographically, two main directions of cyanobacterial distribution have been noted based on affected waterbodies shown in Figure 4.2:

1. First direction (South-North direction) includes Armstrong Lake and Torment Lake (King County, NS), Nowlans Lake (Digby County, NS), Ogden Lake, Fanning Lake, and Vaughan Lake (Yarmouth County, NS).

2. Second direction (East-West direction) includes Mattatall Lake (Cumberland/Colchester Counties, NS), McLaughlin reservoir, Irishtown reservoir, Tower Road reservoir and Turtle Creek reservoir (Westmorland County, NB).

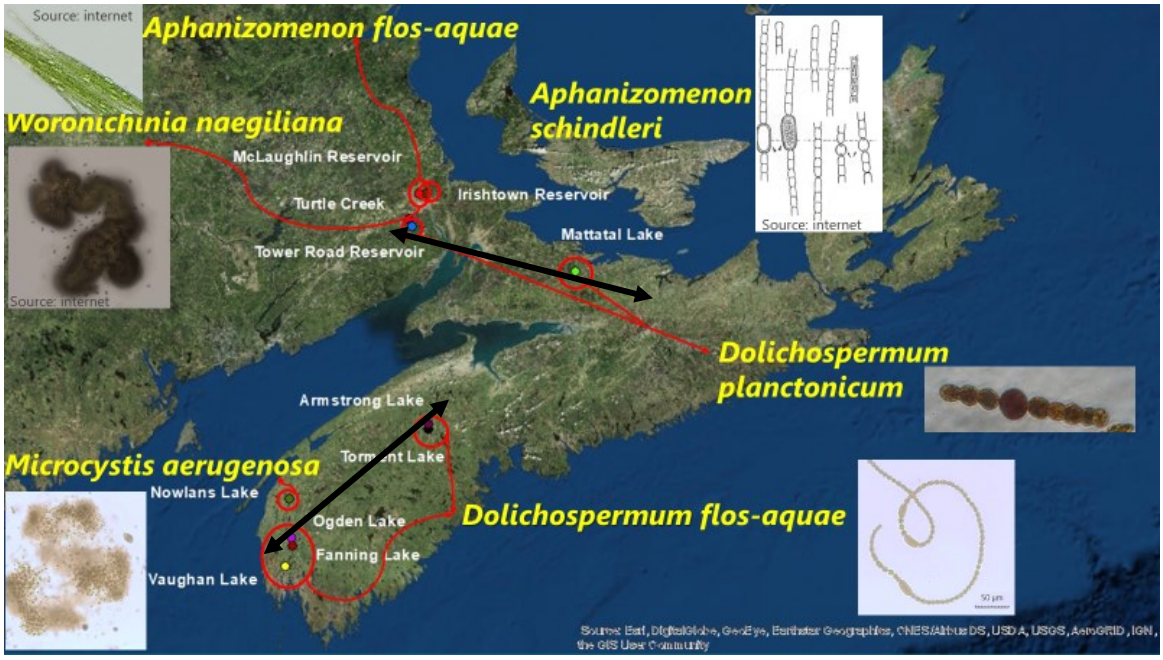


Figure 4.2: Two distribution directions of cyanobacterial development in Nova Scotia and New Brunswick waterbodies

First direction of cyanobacterial distribution: The main dominant cyanobacterial species on the South-North direction is *Dolichospermum flos-aquae* (Figure 4.3 left) which is responsible for releasing two groups of toxins (microcystin and anatoxin). These

species were found in Yarmouth County lakes (Vaughan, Fanning and Ogden) and Kings County lakes (Armstrong and Torment). The lakes have been blooming from late May to November in 2015-2018. In this area, there is also Nowlans Lake which yearly experiences blooms in summer with a high biomass of *Microcystis aeruginosa* (Figure 4.3 middle). This species is the main producer of microcystin.

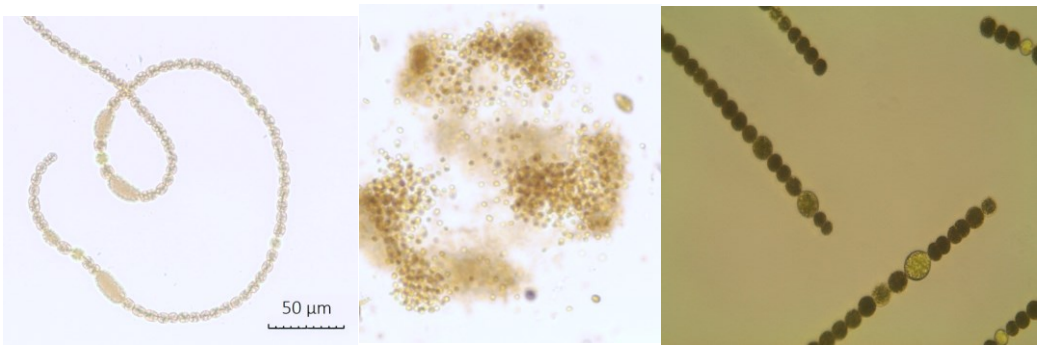


Figure 4.3: Dolichospermum flos-aquae (left), Microcystis aeruginosa (center) and Dolichospermum planctonicum (right)

Second direction of cyanobacterial distribution: *Dolichospermum planctonicum* AKA *Anabeana planctonica* (Figure 4.3 right) is the second dominant species that mostly associates with anatoxin-a production. It has been found at East-West direction: Mattatall Lake and Tower Road reservoir. Late August seems to be preferred time for this species to accumulate its biomass and bloom in September (case of Tower Road reservoir, Moncton) and lasting till December (case of Mattatall Lake).

The North area of Moncton is associated with a different cyanobacterial genus: *Aphanizomenon.* Two nearby reservoirs (McLaughlin and Irishtown in Moncton) have been found to contain species specifically from this genus: in Irishtown reservoir the most common bloom creator was *Aphanizomenon schindlery*, while in McLaughlin reservoir

Aphanizomenon flos-aquae was detected. They both could potentially release toxins during bloom episode, and this was proved by our measurements. In Mclaughlin blooms, we also found the presence of second cyanobacterial dominant species: *Woronichinia naegeliana*.

Through our observations on this direction: *Aphanizomenon schindlery* has generated bloom only in summer (July, August) while *Aphanizomenon flos-aquae* has started to bloom in August and persisted till November. Regarding *Woronichinia naegeliana,* this species has shown a behavior quite similar to *Aphanizomenon flos-aquae* in 2016-2017 (i.e bloom in August till November), but appearing earlier with small blooms in 2018.

Most of cyanobacteria creating blooms in the region are filamentous, except *Microcystis aeruginosa* and *Woronichinia naegeliana*.

The map in Figure 4.1 shows the need of monitoring cyanobacterial bloom program to improve the water quality monitoring system and management in Atlantic Canada for an early warning of HAB.

4.1.2 Toxins Following CyanoHAB

Associated with two directions of distribution of cyanobacterial species distribution previously defined, Figure 4.4 shows levels of toxins detected in considered waterbodies of NS and NB. In case that toxins were detected by more than one method (see the methodology part), the average concentration was reported in the graphs in Figure 4.4.

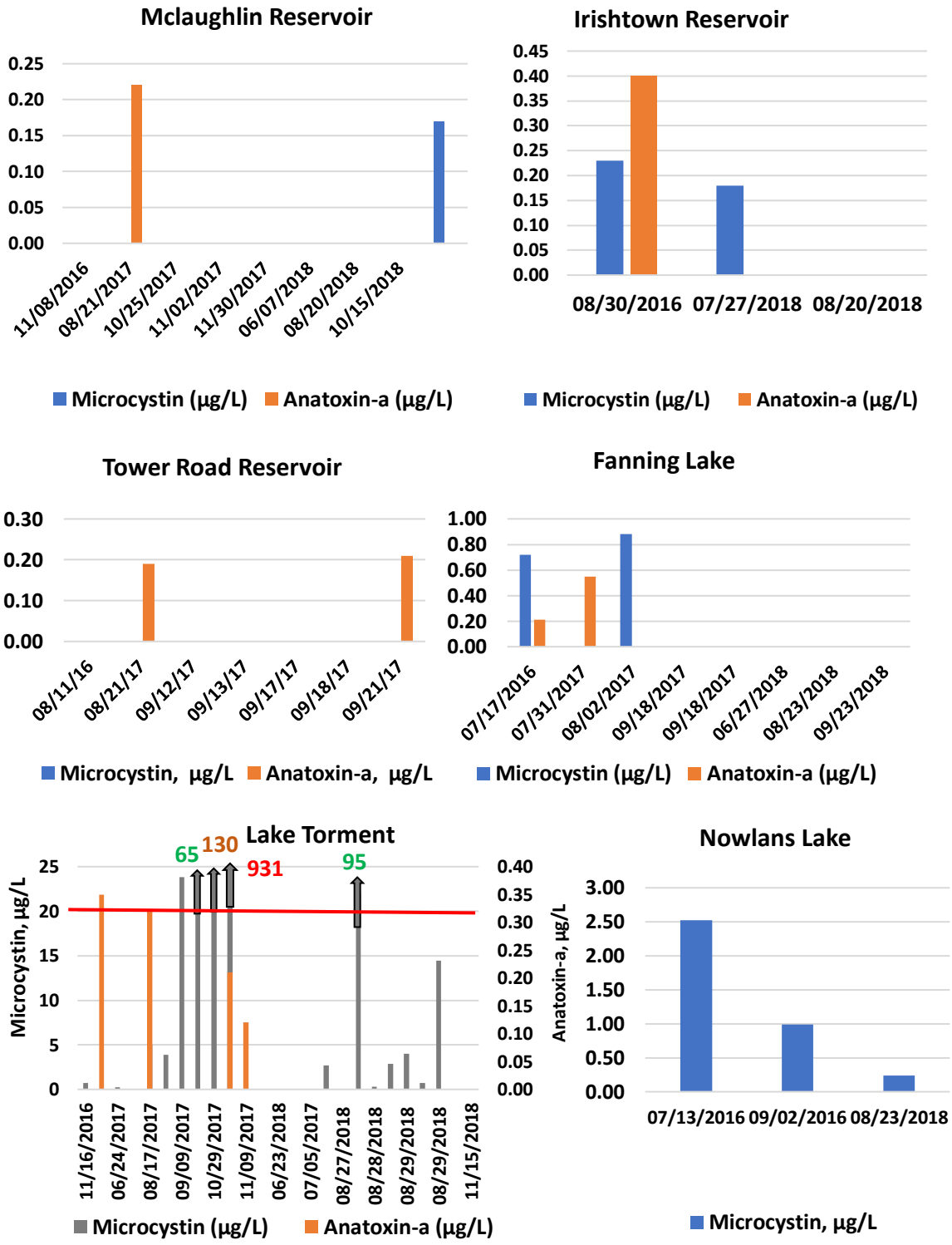


Figure 4.4: Toxins detected at monitored lakes

Starting from the South to the North, the discussion will firstly focus on watersheds in *Yarmouth County.*

Lake Ogden was the first considered in the series of lakes monitored in the Carleton river watershed. There are no lakes experienced bloom from the upstream of this lake in the same system (Stantec Consulting, 2017). Ogden Lake experienced heavy cyanobacterial blooms every summer in 2016-2018, usually in July or in August. During heavy bloom in July 2016, no toxins were found in detectable concentrations (Table 4.1). In 2018 the samples were collected after bloom period. There were no microcystin and anatoxin detected.

Table 4.1: Concentration of toxins in Ogden Lake

Location	Date	Microcystin (µg/L)				Anatoxin-(µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
O1	07/13/2016	<0.15	-	-	-	<0.15	-
	06/03/2017	-	-	NF	NF	-	NF
	08/23/2018	-	<0.05	NF	NF	-	NF

NF – not found

The second lake in the Carleton river watershed is Fanning Lake which receives the water from Ogden, and very often experiences bloom starting after Ogden. In the same time, local cyanobacterial colonies independently develop (in the eastern part of the lake), creating blooms all over the summer till the beginning of fall. Small concentrations of microcystin and anatoxin-a were found in July 2017-2018 (Table 4.2). Late summer blooms have not contained any detectable level of toxins. In 2018, no toxins were detected.

Table 4.2: Concentration of toxins in Fanning Lake

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
Outlet	07/17/2016	0.72	-	-	-	0.21	-
Jean's Place	07/31/2017	<0.15	-	-	-	0.55	-
Jean's Place	08/02/2017	0.88	-	NF	NF	<0.15	trace
Jean's Place	09/18/2017	-	-	NF	NF	-	NF
Jean's Place	06/27/2018	-	-	NF	NF	-	NF
Jean's Place	08/23/2018	<0.15	<<0.05	NF	NF	-	NF
Jean's Place	09/23/2018	-	-	NF	NF	-	NF

NF – not found

Regarding Reynard Lake, the next waterbody in the chain after Fanning, has not shown any pattern of blooms during our 2 years monitoring period. However, Lake Vaughan, the right next one had experienced cyanobacterial bloom of *Dolichospermum flos-aquae* in the end of June 2016. No more bloom episodes during our monitoring program in 2016 - 2018 were detected in this waterbody.

Moving to Digby county, Nowlans Lake has experienced *Microcystis aeruginosa,* a typical cyanoHAB during summer months every year. Microcystins were hence yearly detected in this waterbody with the maximum concentration in July 2016 (Table 4.3). It is shown that MC-LR is the main analog found in blooms of this lake.

Table 4.3: Concentration of toxins in Nowlans Lake

Location	Date	Microcystin (µg/L)			
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS	
				MC-LR	MC-LY
N1	07/13/2016	2.52	-	-	-
	09/02/2016	0.99	-	-	-
	08/23/2018	0.25	0.18	0.24	NF

NF – not found

Moving to the North direction to Kings County, some lakes in LaHave River watershed (Armstrong and Torment) were analysed. Lake Armstrong, located in the upstream of Lake Torment, had several bloom episodes during our monitoring period. In July-August (2016-2017), blooms of filamentous green algae *Mougeotia sp.* were prevalent. In June 2018, scums of cyanobacteria *Dolichospermum flos-aquae* were observed. The LC-HRMS method showed trace amounts of anatoxin-a present during bloom (Table 4.4). Microcystin was not identified in this waterbody.

Table 4.4: Concentration of toxins in Armstrong Lake

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
A3	06/23/2018	-	-	NF	NF	-	trace

NF – not found

Downstream in Torment Lake, the toxin analyses for this waterbody give the largest database among the others related to cyanotoxins in our research (Table 4.5). Toxin detection results for lake Torment show the lake has many particularities in cyanoHAB aspects and then need further research. The following are details of findings for toxin levels in Torment Lake.

Microcystins: First toxins were detected in bloom on June 18, 2017. With LC-HRMS method it was detected the presence of MC-LR and MC-LY in all samples of 2017 at Dave's Cove in Torment from September 9 to the last bloom episode on November 9. The concentration of these microcystin congeners was found increasing from September to November with a maximum value on November 9, 2017 (see Table 4.3). In parallel, detection of microcystins was cross validated by using ELISA ADDA(DM) and ELISA multihapten (see details in Table 4.5). The September 9, 2017 data for all 3 methods

shows different results as with ELISA multihapten we obtained 2 times higher results (7.7 μg/L) than with LC-HRMS (2.5 μg/L of MC-LR and 0.9 μg/L of MC-LY), while ELISA ADDA(DM) gave the values being 3 times lower (0.95 μg/L) than with LC-HRMS. The reason of this discrepancy could be explained by the different time of analyses after the sample collection: ELISA multihapten (April 2018), LC-HRMS (August 2018), ELISA ADDA(DM) (December 2018). The same tendency from ELISA multihapten is observed for samples of November 2 and 7, 2017: microcystins with this method were higher twice than the LC-HRMS results (it is important to note that November 2 and 9, 2017 samples also contained anatoxin-a as discussed below). On November 2, 2017 phytoplankton composition showed the development of *Dolichospermum flos-aquae* (abundance 3.87xE8cells/L, biomass 60761 μg/L) and *Microcystis aeruginosa* (abundance 6.56xE8 cells/L, biomass 42962 μg/L) together with the green macrophyte *Cladophora crispate* (abundance 6.1xE7 cells/L, biomass 1188779 μg/L). Both cyanobacteria continued to grow and reached the biomass 102,522 μg/L (abundance 6.5xE8 cells/L) of *Dolichospermum flos-aquae* and 107,039 μg/L (abundance 1.6xE9 cells/L) of *Microcystis aeruginosa* in the plankton community on November 9, 2017.

Analyses of 2018 data by all three methods shows close results from both LC-HRMS and ELISA multihapten approaches while ELISA ADDA (DM) gave higher concentrations of toxins. That could be explained by the degradation of toxins during the time storage even at -20°C.

We also processed the frequent interval testing of samples at the same location (Dave's cove, every 6-12 hours) on August 27 – 29, 2018. At 6pm of August 27, there was a

visible bloom with a high level of MC-LR (6.2 µg/L) registered. The sample at midnight had no toxins, that probably depends on the movements of cyanobacteria at night toward the bottom or on the prevalent wind direction. The sample of 12 hours later however gave very high amount of microcystin (MC-LR) by both methods LC-HRMS and ELISA ADDA(DM). All 3 methods showed very close results in microcystins detected. 12 hours later (August 29 at 6 am) concentration of total MC-LR was much higher than in the previous samples by LC-HRMS method. ELISA ADDA(DM) results, were only a little higher than previous. August 29 at noon, two different points in Dave's cove were analyzed:

a) in Dave's cove, the green algae *Zygnema sp., Mougeotia sp., Spirogira sp.* + *Dolichospermum flos-aquae* (abundance 3xE7 cells/L, biomass 5008 µg/L) are found developing on the bottom. MC-LR was detected by both LC-HRMS and ELISA multihapten. Results from both methods are in good fit with each other.

b) Dave's cove the surface scum grew from *Dolichospermum flos-aquae* (abundance 6xE8 cells/L, biomass 1.2xE5 µg/L), *Dolichospermum mendotae* (abundance 1xE8 cells/L, biomass 13340 µg/L), and *Zygnema sp.* (abundance 2xE7 cells/L, biomass 3xE5 µg/L). Total MC-LR was found lower by ELISA ADDA(DM) than by LC-HRMS.

No blooms were observed in the cove after that until October 4th. Analyses of total toxins in October gave the values of MC-LR (14.74 µg/L) and MC-LY (7.38 µg/L) by LC-HRMS. ELISA ADDA(DM) results also confirmed the MC-LR value (12.34 µg/L) but MC-LY cannot be detected. November sample showed the absence of microcystins.

Anatoxins: Only anatoxin-a with small amount was detected in 2017. With Anatoxin ELISA kit, we detected anatoxin-a in samples of June and August 2017 and with LC-

HRMS, small concentrations of anatoxin-a were found in November samples. All the samples were analyzed by LC-HRMS in 2018, but no detectable toxins were found.

Table 4.5: Concentration of toxins in Torment Lake

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
TDave	11/16/2016	< 0.15	-	-	-	< 0.15	-
Tdave	06/18/2017	0.74	-	-	-	0.35	-
T1'bottom	06/24/2017	< 0.15	-	-	-	< 0.15	-
T7 bottom	06/24/2017	0.24	-	-	-	< 0.15	-
Tdave	08/17/2017	-	-	-	-	0.32	-
Tdave	08/29/2017	> 5.00	-	-	-	< 0.15	-
Tdave	09/09/2017	0.94	7.70	2.15	0.90	-	NF
Tdave	09/16/2017	>25.0	-	20.2	2.40	-	trace
Tdave	10/29/2017	-	-	57.8	7.90	-	NF
Tdave	11/02/2017	-	178	72.1	9.00	-	0.21
Tdave	11/09/2017	-	1102	684	76.6	-	0.12
Tdave	06/21/2018	-	-	NF	NF	-	NF
Tdave	06/23/2018	-	-	NF	NF	-	NF
Tdave	06/27/2018	-	-	NF	NF	-	NF
Tdave	07/05/2017	-	-	NF	NF	-	NF
	08/17/2018	-	-	NF	NF	-	NF
Tdave	08/27/2018 6pm	1.35	0.58	6.20	NF	-	NF
	08/28/2018 1am	-	-	NF	NF	-	NF
Tdave	08/28/2018 noon	>100	-	91.0	NF	-	NF
Tdave	08/28/2018 6pm	0.36	0.26	0.27	NF	-	NF
Tdave	08/29/2018 6am	0.69	-	5.05	NF	-	NF
Tdave surface	08/29/2018 noon	2.14	3.13	6.65	NF	-	NF
Tdave bottom	08/29/2018 noon	0.67	0.82	0.75	NF	-	NF
Tdave	10/04/2018	12.34	-	14.74	1.83	-	NF
Tdave	11/15/2018	<0.15	-	NF	NF	-	NF

NF – not found

Mattatall Lake (Cumberland and Colchester Counties) has mainly contained cyanobacterial blooms of *Dolichospermum planctonicum* (syn. *Anabaena planctonica),* which have been extremely prominent in 2014-2016. However, no blooms were detected after 2016 due to by our hypothesis that modifications of flow streams and outlet were done by the lake residents and shutting down the forest clear-cutting activity. Two 2016

samples showed the absence of both anatoxin-a and microcystin (Table 4.6). This can be explained by the long storage time of samples (1 year) prior to analyses.

Table 4.6: Concentration of toxins in Mattatall Lake

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
Inlet	08/16/2016	≤0.15	-	-	-	≤0.15	-
Blain	10/04/2016	≤0.15	-	-	-	≤0.15	-

Moving to the west direction, the drinking water for Great Moncton, Tower Road reservoir, shows the same prevalent cyanobacterial species of *Dolichospermum planctonicum* with bloom patterns. Anatoxin-a was detected firstly in the water sample from this reservoir in August 2017 before visual blooms have developed. However, no toxins were found during bloom episodes of September 13-15, 2017. The bloom degradation (the end of bloom period) was accompanied by the anatoxin-a detection again in the waterbody. Microcystins, however, were not found in any water samples of this reservoir in 2017. In 2018, we did not observe any visible blooms in this reservoir, and neither toxins were detected. That could be connected to weather conditions which were quite different from 2017.

Moving to the North of TRR, there is the first and current drinking water reservoir of the city: Turtle Creek. Turtle Creek (TC) had an episode of filamentous algae development 3 km from the Tower. This area is stagnant and shallow (1m deep). Filamentous green algae *Mougeotia sp.* (abundance – 3xE7 cells/L, biomass -1xE6 µg/L) were 'permanently' found here and they are responsible for blooms appearing in this location. Also, *Microcystis aeruginosa* was found in small amount (abundance - 5xE6 cells/L,

biomass -330 µg/L). The latter can be probably the cause of intracellular toxins (MC-LR) found by our all three methods for the samples taken from this location, precisely, 0.6µg/L (LC-HRMS) and 0.5 µg/L (ELISA multihapten) being close to each other. Concentration of microcystins in ELISA ADDA(DM) method was lower (0.16 µg/L).

Table 4.7: Concentration of toxins in Tower Road reservoir

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
TRR1	11/08/2016	≤0.15	-	-	-	<0.15	-
TRR1	08/21/2017	≤0.15	-	-	-	0.19	-
TRR1	09/12/2017	≤0.15	-	-	-	<0.15	-
Spillway		≤0.15	-	-	-	<0.15	-
TRR1	09/13/2017	≤0.15	-	-	-	<0.15	-
TRR1	09/17/2017	≤0.15	-	-	-	0.16	-
TRR4	09/18/2017	≤0.15	-	-	-	<0.15	-
TRR1	09/21/2017	≤0.15	-	-	-	0.21	-
TRR2		≤0.15	-	-	-	<0.15	-
TRR3		≤0.15	-	-	-	0.21	-
TRR4		≤0.15	-	-	-	<0.15	-
TRR5		≤0.15	-	-	-	0.20	-
TRR6		≤0.15	-	-	-	<0.15	-
TRR7		≤0.15	-	-	-	0.46	-
TRRInletL		≤0.15	-	-	-	-	-
TRRInletR		≤0.15	-	-	-	-	-
Spillway		≤0.15	-	-	-	0.41	-
Spillway algae		≤0.15	-	-	-	0.26	-
Spillway out		≤0.15	-	-	-	0.28	-
By-pass	10/10/2017	-	-	NF	NF	-	NF
TRR1	04/23/2018	-	-	NF	NF	-	NF
TRR7	06/19/2018	-	-	NF	NF	-	NF
TRR7/TRR8	08/14/2018	-	<0.05	NF	NF	-	NF

NF – not found

Table 4.8: Concentration of toxins in Turtle Creek reservoir

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
Old Dam	11/08/2016	<0.15	-	-	-	<0.15	-
TC1	08/14/2018	0.16	0.46	0.48	NF	-	NF

NF – not found

McLaughlin reservoir experienced mixed blooms of *Aphanizomenon flos-aquae* and *Woronichinia naegeliana* from August to November. These blooms developed in the spillway and north parts of the reservoir near the road. In 2016-2017 late fall blooms didn't release toxins in the water. However, August 2017 showed the presence of small amount of anatoxin-a near the road. In 2018 the early small bloom of *Woronichinia naegeliana* (in the beginning of June 2018) was detected at M8 location near the road. That was the first returning early bloom episode in this reservoir since 2016. However, toxins were not detected by using LC-HRMS method. A bloom in this reservoir was found again near the road (outlet location) in August 2017/18. In the end of August 2017, the bloom was formed by *Dolichospermum spiroides* (biomass 50693 µg/L) and *Aphanizomenon flos-aquae* (biomass 49870 µg/L). The algal accumulation at surface in the inlet area was captured on August 20, 2018 and showed the thin film of surface algae was generated by three different cyanobacterial species: *Dolichospermum planctonicum*, *Aphanizomenon flos-aquae and Woronichinia naegeliana,* with an abundance 1.2xE9 cells/L and biomass 175,025 µg/L. It is also noted that this bloom started in the presence of huge filamentous green algae bloom of *Mougeotia sp*. in the same area. However, no toxins were detected by all 3 methods.

In 2018, there was a mixed bloom of *Aphanizomenon flos-aquae* mostly with a presence of *Woronichinia naegeliana* developed only at spillway part. LC-HRMS method did not detect any toxin, but ELISA ADDA(DM) showed low concentration of microcystins (0.17 µg/L).

Table 4.9: Concentration of toxins in McLaughlin reservoir

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
M13	11/08/2016	≤0.15	-	-	-	<0.15	-
M8	08/21/2017	≤0.15	-	-	-	0.22	-
M1/M8	10/25/2017	-	-	NF	NF	-	NF
M1	11/02/2017	-	-	NF	NF	-	NF
M1/M8	11/30/2017	-	-	NF	NF	-	NF
M8	06/07/2018	-	-	NF	NF	-	NF
M7/M8/M8'	08/20/2018	<0.15	<0.05	NF	NF	-	NF
M1	10/15/2018	0.17	-	NF	NF	-	NF

NF – not found

Multiples cases of toxins released were registered in Irishtown National Park reservoir in the downtown of Moncton city. In 2016 small concentrations of both kinds of toxins were detected in the end of August by ELISA approach (Table 4.10). There was no visible bloom observed but the biomass of *Aphanizomenon schindleri* was detected high (13,363 µg/L). In 2017, there was no bloom and no measurements of toxins. In 2018 the huge bloom of *Aphanizomenon schindleri* developed in the end of July in this reservoir near the park pedestrian bridge. No microcystins by LC-HRMS were recorded, but ELISA multihapten and ELISA ADDA(DM) methods detected small concentrations of microcystins. One sample from a near location to the bloom was analysed in 1 month after bloom episode. Cyanobacteria *Aphanizomenon schindleri* abundance was still high (abundance 2xE8 cells/L), but visually the bloom was dispersed. No toxins were detected at that moment.

Table 4.10: Concentration of toxins in Irishtown reservoir

Location	Date	Microcystin (µg/L)				Anatoxin-a (µg/L)	
		ELISA ADDA(DM)	ELISA multihapten	LC-HRMS		Anatoxin-a ELISA	LC-HRMS
				MC-LR	MC-LY		
I4	08/30/2016	0.23	-	-	-	0.40	-
I2	07/27/2018	0.19	0.17	NF	NF	-	NF
I1	08/20/2018	<0.15	<0.05	NF	NF	-	NF

NF – not found

To conclude about the cyanotoxin session:

1) it could be said that mostly microcystins (with two congeners LR and LY) were found in many considered lakes associated with cyanobacterial bloom. Anatoxin-a with small concentration was detected after bloom episodes in certain waterbodies including Fanning, Torment, Irishtown and Tower Road Reservoir. The importance of an early detection and warning system related to HAB and toxins is therefore indispensable for NS and Great Moncton, NB.

2) Statistical analyses of three different methods of detection for cyanotoxins as previously mentioned were conducted. In using Pearson coefficient of correlation, it is shown that there exists a mutual relationship between all three methods which is highly positive, i.e Pearson coefficient of correlations greater 0.9 with the maximum value obtained for ELISA multihapten and LC-HRMS (Table 4.11).

Table 4.11: Pearson coefficient of correlation between methods

Between methods	R	Number of samples
LC-HRMS-ELISA ADDA(DM)	0.943	32
LC-HRMS-ELISA multihapten	0.986	21
ELISA ADDA(DM)-ELISA multihapten	0.923	21

R – Pearson's coefficient of correlation

However, there are also some discrepancies in the results between 3 methods due to following reasons:

a) the storage time of samples after collection from the field: the longer samples accumulated in freezer the lower the results. The optimum time for sample storage should be less than one month (Zaffiro et al., 2016).

b) the difference in limits of detection (LOD): we noticed that LC-HRMS could not detect very low concentrations because its LOD was several times higher than the ones in ELISA methods. This makes ELISA methods more sensitive.

4.2 Application of TRINDEX to Freshwater Environment (Mattatall Lake)

4.2.1 Results Obtained from TRINDEX Application

Our model is developed using ML data from 2015-2017. Basic statistical parameters are summarised in Table 4.12. The sample size differs due to extreme maximum values elimination. It is noticed that there is the fluctuation of data between no bloom and bloom periods.

Table 4.12: Statistical data for Mattatall Lake

	Absolute deviation of oxygen from 100%	**Chlorophyll-a**	**Phycocyanin**	**Phosphate**	**Nitrate**
Number of samples	310	292	310	310	310
Maximum value	100	1.59	1.80	1.52	1.70
Minimum value	0.01	5.E-5	4.E-5	0.01	0.01
Average	13.9	0.02	0.08	0.05	0.20
Standard deviation	15.5	0.09	0.54	0.13	0.18

**All data in* mg/L *except deviation of oxygen which is dimensionless number*
*** The minimum value is the detection limit of the method for each variable*

Appendix A shows a correlation between parameters for reference. It is important to note that Pearson's coefficient of correlation represents only the linear correlation between

parameters, and this is not the case of our model where the non-linearity is dominant. This non-linearity between our chosen parameters can be seen via correlation of each pair of parameters for both lakes in Appendix A.

The experimental data related to water parameters for both lakes (Mattatall and Torment) are not normally distributed. The use of log transformation therefore is one approach to convert them into the *'normal distribution'* and TRINDEX can be then processed. Data used for the determining lower and upper limits are in Table 4.13.

Based on Table 4.13 and formulas (5, 6) in the previous section, TRINDEX will be described as follows.

$$TRINDEX1 = 2.5 \times [\frac{logPC+4.4}{4.7} + \frac{logD\%O+2}{4.0} + \frac{logPO4+2}{2.2} + \frac{logNO3+2}{2.2}] \quad (10)$$

$$TRINDEX2 = 2 \times [\frac{logPC+4.4}{4.7} + \frac{logCh+4.3}{4.5} + \frac{logD\%O+2}{4.0} + \frac{logPO4+2}{2.2} + \frac{logNO3+2}{2.2}] \quad (11)$$

Table 4.13: Limits and ranges - Mattatall Lake data

Limits and ranges	L_i	U_i	$LogL_i$	$LogU_i$	$LogU_i$-$LogL_i$
Phycocyanin, mg/L	4.E-5	1.855	-4.4	0.3	4.7
Chlorophyll-a, mg/L	5.E-5	1.692	-4.3	0.2	4.5
D%O	0.01	1000	-2.0	2.0	4.0
Phosphate, mg/L	0.01	1.52	-2.0	0.2	2.2
Nitrate, mg/L	0.01	1.72	-2.0	0.2	2.2

D%O – absolute deviation of oxygen from 100%
L_i – lower limit (concentration) of the parameter: LOD
U_i – upper limit (concentration) of the parameter: maximum observed concentration

The following Table 4.14 shows some statistical comparison for both TRINDEXes. It is observed that two TRINDEXes are quite close. Some discrepancies can be explained by the unequal contribution of each parameter during bloom scenarios.

Table 4.14: Statistical comparison of TRINDEX1 and TRINDEX2

	TRINDEX1	TRINDEX2
Sample size	309	292
Maximum	7.70	7.40
Minimum	1.58	1.88
Average	4.32	4.42
Standard deviation	1.15	1.09

4.2.2 Receiver Operator Curve Analyses and Threshold Determination

It was noticed that points with hypoxia conditions on the bottom increased TRINDEX values without any signs of cyanobacterial bloom. Therefore, these data were removed from the computations. In the end our dataset for TRINDEX1 was 266 and for TRINDEX2: 249. The distinct scenario for both bloom and no bloom conditions for TRINDEX1 and TRINDEX2 using *rnorm* in R software is graphically represented in Figure 4.5.

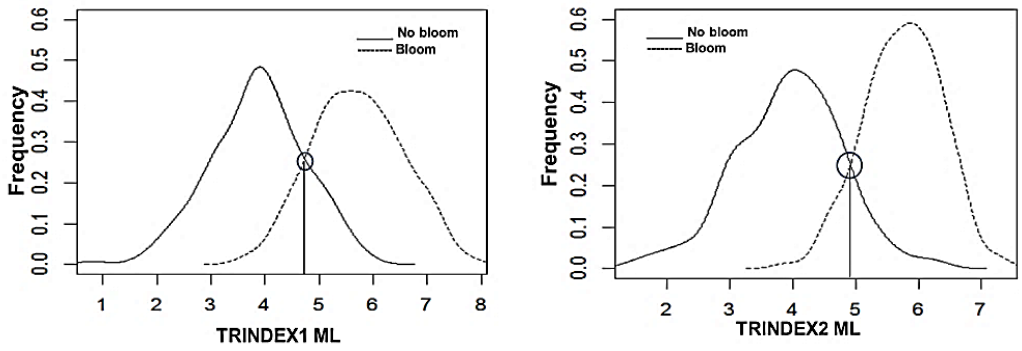


Figure 4.5: Distribution of no bloom and bloom cases TRINDEX1 and TRINDEX2, ML

The normal distribution of no bloom and bloom cases for both TRINDEXes shows approximately the same value 5 as the possible cut-off point (Figure 4.5). However, this cut-off point should be validated by ROC curve analysis to precise the threshold value for

cyanobacterial bloom, as the graphical determination could not be accurate enough, hence leading eventually to the wrong estimation of bloom scenarios.

The ROC curve analyses were processed through Table 4.15 (with field observation data) and Figure 4.6. Calculations for the correctly positive and incorrectly positive are the following.

There were 266 values of TRINDEX1 from ML during 2015-2017 used in ROC analyses, among them 74 cases with bloom and 192 cases without bloom (Table 4.15). Single cases of bloom were detected when TRINDEX1 started from value 4 (Table 4.15). It is observed that the higher TRINDEX1 (greater than 5.0), the more frequent bloom cases are recorded than no bloom cases. Therefore, the TRINDEX1 range from 4.0 to 5.0 is the marginal situation, where there is possibly no visible bloom sign but just only small disturbances of the environmental conditions (leading to a higher TRINDEX1) which could trigger the cyanobacterial bloom. When TRINDEX1 = 6.2, the bloom cases were the maximum (14 cases).

There were 249 calculated values of TRINDEX2 (Table 4.15) with 75 bloom cases and 174 no bloom cases. The lowest TRINDEX2 when bloom was observed is 4.4, and from TRINDEX2 = 5.2 the number of bloom cases occur more often than no bloom cases. The maximum number of no bloom cases were noticed when TRINDEX2 = 4.0 and the maximum bloom cases were when TRINDEX2 = 6.2. From above analyses, the proposition of transition range for TRINDEX2 is from 4.4 to 5.2 and the threshold value for bloom occurrence suggested is 5.2.

Table 4.15: Data for ROC curve of Mattatall Lake

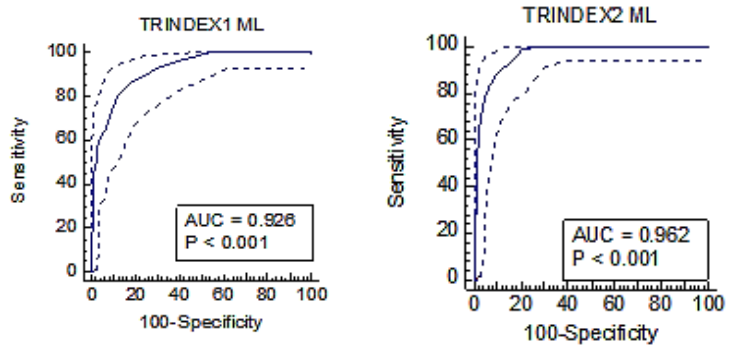
TRINDEX	TRINDEX1					TRINDEX2				
	Bloom	No bloom	Total	Sensitivity	(100-Specificity)	Bloom	No bloom	Total	Sensitivity	(100-Specificity)
1.60		1	1	100	100					
1.80		1	1	100	99		1	1	100	100
2.00		5	5	100	99		1	1	100	99
2.20		3	3	100	96		3	3	100	99
2.40		3	3	100	95		4	4	100	97
2.60		4	4	100	93		4	4	100	95
2.80		12	12	100	91		4	4	100	93
3.00		12	12	100	85		18	18	100	90
3.20		5	5	100	79		6	6	100	80
3.40		6	6	100	76		10	10	100	76
3.60		17	17	100	73		18	18	100	71
3.80		21	21	100	64		17	17	100	60
4.00	2	21	23	100	53		26	26	100	51
4.20	3	24	27	97	42		13	13	100	36
4.40	3	13	16	93	30	1	13	14	100	28
4.60	2	10	12	89	23	5	11	16	99	21
4.80	4	11	15	87	18	2	7	9	92	14
5.00	5	6	11	81	12	5	7	12	89	10
5.20	7	4	11	75	9	4	3	7	83	6
5.40	5	7	12	65	7	7	3	10	77	5
5.60	8	2	10	59	3	10	2	12	68	3
5.80	3	3	6	48	2	8	1	9	55	2
6.00	4		4	44	1	5		5	44	1
6.20	14		14	39	1	15	1	16	37	1
6.40	5	1	6	20	1	7	1	8	17	1
6.60	6		6	13	0	2		2	8	0
6.80						1		1	5	0
7.00	2		2	5	0	1		1	4	0
7.20	1		1	3	0	1		1	3	0
7.40						1		1	1	0
Total	74	192	266			75	174	249		

From the ROC curves (Figure 4.6), the appropriate threshold for bloom onset can be chosen. It should have the maximum sensitivity and at the same time the minimum incorrect positive cases. Two axes of our ROC curve are determined by *the sensitivity* 100% (defined as the probability of correct positive results) and (*100% - specificity*) (the

probability of incorrect positive results) are calculated. Precisely, the *incorrect positive* cases show TRINDEX are high but there are no blooms occurring, while the *incorrect negative ones* show the opposite scenario: TRINDEX are low but blooms are observed.

The ROC curves (Figure 4.6) can show the cut off points (which are close to the left upper corner to represent the combination of the highest sensitivity and lowest incorrect positive). For TRINDEX1 the best combination of high sensitivity (81%) and low incorrect positive (12%) is equivalent to the point 5.0 in Table 4.15. So, all results of TRINDEX1 equal to or greater than 5.0 have to be considered as resulting in cyanobacterial blooms (with 19% of correctly positive results coming into transition phase).

TRINDEX2 (Figure 4.6 right) has the best combination of high sensitivity (83%) and low incorrect positive (6%), equivalent to the value 5.2 in Table 4.15. Both ROC curves for TRINDEX1 and TRINDEX2 have the AUC value in the range 0.9-1.0 ($p < 0.001$) so can be evaluated as excellent (see Table 2.1). Therefore, there are two cut-off points: 5 can be considered as threshold for TRINDEX1 and 5.2 as threshold for TRINDEX2.



___ Receiver operating characteristic (ROC) curve
---- 95% confidence interval
AUC – area under the curve

Figure 4.6: ROC curves for TRINDEX1 and TRINDEX2, ML

4.2.3 Independent Verification of Received Results (Torment Lake)

The independent verification was done based on TL data. Blooms appeared 'randomly' at the surface from June to November, increasing in frequency from 2014 to 2018, when our monitoring program was established. The problem of bloom was caused by *Dolichospermum flos-aquae* which can produce microcystins, anatoxins and saxitoxins. The species that created cyanobacteria blooms in TL are different from ML, but from the same genera. Therefore, TL can be a good validation site for the model.

The normal distribution approach shows that the two cut-off points are: 4.6 for both TRINDEXes (Figure 4.7).

The same analyses as for ML were processed for TL to validate these threshold values for both TRINDEXes (Table 4.16 and Figure 4.7).

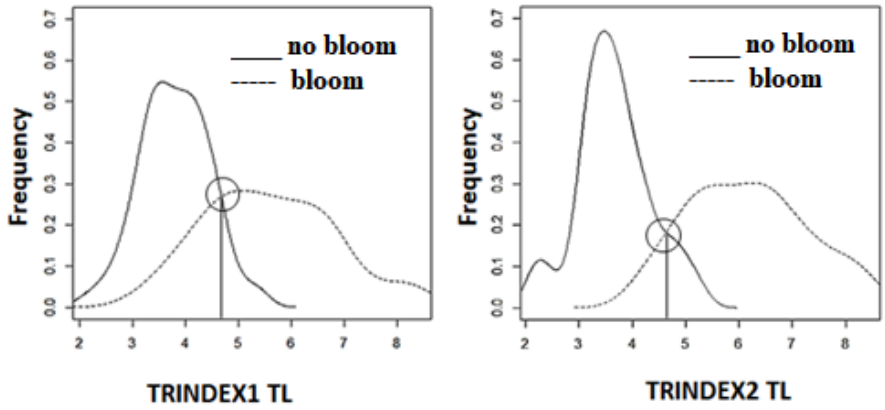


Figure 4.7: Distribution of no bloom and bloom cases TRINDEX1 and TRINDEX2 (Torment Lake)

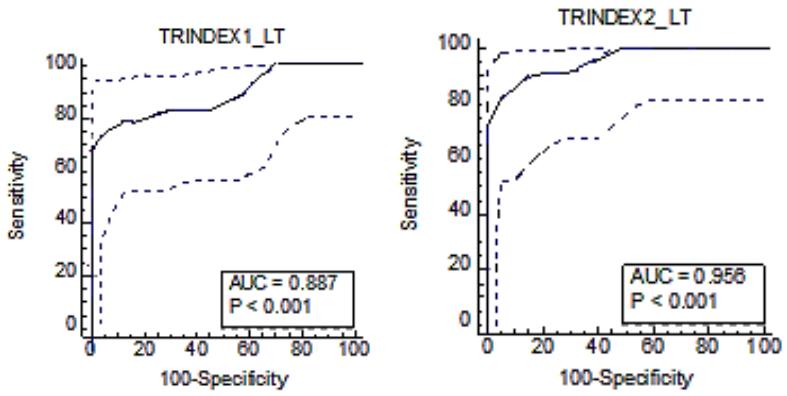
Table 4.16: Data for ROC curve of Torment Lake

TRINDEX	TRINDEX1					TRINDEX2				
	Bloom	No bloom	Total	Sensitivity	(100-Specificity)	Bloom	No bloom	Total	Sensitivity .	(100-Specificity)
1.60		1	1	100	1					
2.00		2	2	100	99					
2.20		2	2	100	98		2	2	100	1
2.40		4	4	100	97		2	2	100	99
2.60		8	8	100	94		4	4	100	97
2.80		8	8	100	88		6	6	100	95
3.00		5	5	100	83		13	13	100	90
3.20		12	12	100	80		4	4	100	82
3.40	1	13	14	100	71		16	16	100	79
3.60	2	9	11	96	62		8	8	100	68
3.80	1	15	16	88	56		24	24	100	62
4.00		25	25	83	46	2	24	26	100	46
4.20	1	16	17	83	28.8		16	16	92	29
4.40		12	12	79	18	1	12	13	92	19
4.60		5	5	79	10	2	10	12	88	10
4.80	2	6	8	79	6		2	2	79	3
5.00	1	1	2	71	2		2	2	79	2
5.20	2	2	4	67	1					
5.40	3		3	58	0	1		1	79.2	1
5.60	3		3	46	0	3	1	4	75	1
5.80						2		2	63	0
6.00						4		4	54	0
6.20	1		1	33	0	1		1	38	0
6.60	1		1	29	0					
6.80	4		4	25	0	1		1	33	0
7.00						1		1	29	0
7.40	1		1	8	0	3		3	25	0
7.60						1		1	13	0
7.80	1		1	4	0	1		1	8	0
8.20						1		1	4	0
Grand Total	**24**	**146**	**170**			**24**	**146**	**170**		

Analyses of 170 values of TRINDEX1and TRINDEX2 from TL (2015-2018) were computed. Among them 24 cases were bloom and 146 cases were without bloom (Table

4.16). Single cases of bloom were detected when TRINDEX1 started from value 3.4 (Table 4.16). It is observed that the higher TRINDEX1 (greater than 5.2), only bloom cases are recorded, and the maximum bloom cases (4 cases) happened when TRINDEX1 = 6.8. At TRINDEX1 equal 5.0-5.2 the number of bloom and no bloom cases were the same. Therefore, it can be said that the TRINDEX1 range from 3.4 to 4.8 is the marginal situation.

The lowest TRINDEX2 showing bloom is 4.0, and from TRINDEX2 = 5.2 the number of bloom cases occur more often than no bloom cases. The maximum number of no bloom cases were noticed when TRINDEX2 = 3.8-4.0 and the maximum bloom cases were when TRINDEX2 = 6.0. From above analyses, the proposition of transition range for TRINDEX2 is from 4.0 to 5.2 and the threshold value for bloom occurrence suggested is 5.2.



___ Reciver operatinf characteristic (ROC) curve
---- 95% confidence interval
AUC – area under the curve

Figure 4.8: ROC curves for TRINDEX1 and TRINDEX2, TL

ROC curves showed that TRINDEX1 (Figure 4.8 left side) has the best combination of high sensitivity (79%) and low incorrect positive (6%), equivalent to the value 4.8 while

TRINDEX2 (Figure 4.8 right side) has the best one at sensitivity (79%) and incorrect positive (2%), equivalent to the value 5.0.

4.2.4 Discussions on the TRINDEX Results from Two Independent Lakes

To sum up the analysed information above, with ML, it is observed that TRINDEX1 has a range of values from 4.0 to 5.0, considered as the transition phase, i.e. potential for a bloom occurrence. The value 5.0 can be considered as a threshold. With TL, transition range for TRINDEX1 is 3.4-4.8 and 4.8 is a threshold. TRINDEX2 has transition phase with values 4.4-5.2 (ML) and 3.8-5.2(TL); the threshold is 5.2 for both lakes.

Tables 4.17a, 4.17b and 4.17c also show calculations for TL data in comparison with ML calculations.

Table 4.17a: Statistical analyses of TRINDEX1 and 2 for both lakes (by using MedCalc)

	TRINDEX1		**TRINDEX2**	
	Mattatall	**Torment**	**Mattatall**	**Torment**
Sample size	266	170	249	170
Positive group [a]	74 (27.82%)	24 (14.12%)	74 (29.72%)	24 (14.12%)
Negative group [b]	192 (72.18%)	146 (85.88%)	175 (70.28%)	146 (85.88%)

[a] *results = 1 (having bloom)*
[b] *results = 0 (no bloom)*

Table 4.17b: Area under the ROC curve (AUC) for data of the two lakes

	TRINDEX1		TRINDEX2	
	Mattatall	**Torment**	**Mattatall**	**Torment**
Area under the ROC curve	0.926	0.887	0.961	0.956
Standard Error [a]	0.0162	0.0465	0.0106	0.0226
95% Confidence interval [b]	0.887 to 0.954	0.830 to 0.930	0.929 to 0.981	0.914 to 0.982
z-score	26.237	8.324	43.505	20.172
Significance level *p* (Area=0.05)	< 0.0001	<0.0001	< 0.0001	<0.0001

[a] *DeLong et al., 1988 (the method recommended by MedCalc to calculate standard error and CI95%)*
[b] *Binomial exact*

Table 4.17c: Youden index for data from two lakes

	TRINDEX1		TRINDEX2	
	Mattatall	**Torment**	**Mattatall**	**Torment**
Youden index *J*	0.69	0.73	0.78	0.78

From these results, we can see that ML data with AUC (TRINDEX1) = 0.926 shows that the discrimination test is excellent (95%CI: 0.887 to 0.954; $p < 0.0001$).

The AUC (TRINDEX2) = 0.961 shows that the discrimination test is excellent (95%CI: 0.929 to 0.981; $p < 0.0001$) as well. A test has (at least some) discriminatory power if the 95% confidence interval of AUC does not include 0.50 (Carter et al. 2016).

The results confirm the good fit of our threshold 5.0 (TRINDEX1) and 5.2 (TRINDEX2) for ML as the AUC = 0.926 and 0.961, that is, the discrimination test was excellent (Table 2.1).

Tables 4.17a,b,c show comparison between LT and ML data in term of threshold and ROC curve analyses. Although the sample size of two lakes is different, number of samples (170 cases for analyses for LT compared to 266 values of ML) is enough for a

statistical significance (95% confidential interval, maximum error estimated around 0.2 and standard deviation 1.17).

An area of more than 0.9 (TRINDEX1: 0.926 and TRINDEX2: 0.961 for ML) implies that in a hypothetical experiment in which we randomly select pairs of positive cases (bloom) an incorrect negative result is low comparable to that of incorrect positive result. With the environmental factors that can affect the lake system, the random excitation can cause change of stability around the equilibrium point and beyond this equilibrium point, blooms occur, i.e., instability scenario will cause the HAB.

Considering TRINDEX1 and 2 as 'biomarker' for bloom, J is significant in our tests: 0.69 for ML TRINDEX1 and 0.78 for ML TRINDEX 2 (Table 4.17c). The results are similar for TL.

The significant level represented by p-value is also needed to be discussed herein, as it stands for the probability that the observed sample AUC is found when the true (population) AUC is 0.5. When p is small ($p < 0.05$) it can be concluded that AUC is significantly different from 0.5. There is therefore evidence that the laboratory test does have an ability to distinguish between the two groups 'bloom and no bloom'.

A brief comparison of TRINDEX1 and TRINDEX2 is made for two independent lakes (Mattatall and Torment) shown in Table 4.18.

Table 4.18: Comparison of statistical characteristics TRINDEX1 and TRINDEX2 for both lakes

	TRINDEX1		TRINDEX2	
	Mattatall	**Torment**	**Mattatall**	**Torment**
Sample size	266	170	249	170
Average	4.39	3.94	4.45	4.10
Standard deviation	1.17	1.03	1.16	1.08
Maximum	7.60	7.82	7.40	8.11
Minimum	1.17	1.70	1.80	2.23

From Table 4.18, it is noticed that the range of TRINDEX1 and TRINDEX2 is slighly different between two lakes. That could be due to the presence of different species and PC content as well as nutrients in two lakes. As TRINDEX2 combines both pigments PC and chl-a, it seems inaccurate for the prediction of cyanobacterial bloom thresholds due to the increase of chl-a by other phytoplankton rather than just cyanobacteria, hence increasing TRINDEX2 higher than the real threshold. Therefore, TRINDEX1 based only on PC seems the better indicator to estimate the threshold for cyanobacterial bloom than TRINDEX2. The lowest value of TRINDEX1 when blooms appear in 2 lakes was chosen for the transition phase and the cut-off point was chosen is the theshold for bloom onset.

4.3 Thresholds Suggested via TRINDEX and Application

TRINDEX1 is our main recommendation for the bloom occurrence indicator. In taking the results from two lakes into account, the final scale of thresholds can be proposed as follows:

- When TRINDEX1 < 3.4: no bloom should happen;

- When TRINDEX1 is between 3.4-4.8: there will be a high risk of cyanobacterial bloom development; this range is the transition phase.
- When TRINDEX1 > 4.8: cyanobacterial blooms are likely occurred.

I would highlight some important points from this work as below:

1) The original and modified TRIX with only chl-a cannot tell about the existence of cyanobacteria in the waterbody, but our suggested TRINDEX1 is able to do that. It is necessary to remind that TRINDEX1 is not an indicator like TRIX for trophic level of the lake, but indicator for the cyanobacterial bloom onset.
2) Because the chl-a is a common pigment for all kinds of micro-plants and algae, and it varies significantly under different environmental conditions, therefore PC should play the main role in the indicator of cyanobacterial bloom threshold. From this point of view, the use of PC in TRINDEX1 leads to the possible prediction of bloom occurrence.
3) The range called 'transition phase' should be understood as a 'risk' possibility or potential bloom, i.e. that can lead to unsteady situation and blooming immediately depending on many other factors, such as light, temperature, microelements etc.
4) It is noticed that TRINDEX1 can be very high in the summer at bottom of the stratified lakes. As the oxygen is quite low at the bottom, it could completely increase the values of TRINDEX1 without any cyanobacterial existence. It is suitable to assume that the TRINDEX1 can be used in the epilimnion layer in case of stratification or all the depth of lake if there is no stratification.

5) Temperature is a very important parameter for cyanobacterial growth. However, this factor was not considered in our TRINDEX model due to the fact that different potentially toxic cyanobacterial species grow with various temperature ranges. It is assumed that TRINDEX1 could be used when temperature is greater than 15ºC which is the lowest temperature observed in Atlantic Canada for cyanobacterial blooms to develop.

6) Wind factors and light intensities were also not considered yet in TRINDEX due to their complexity and the lack of data. However, these parameters will be certainly taken into account in future research.

Finally, the scheme in Figure 4.9 is suggested as a summarized management tool for bloom onset prediction based on TRINDEX.

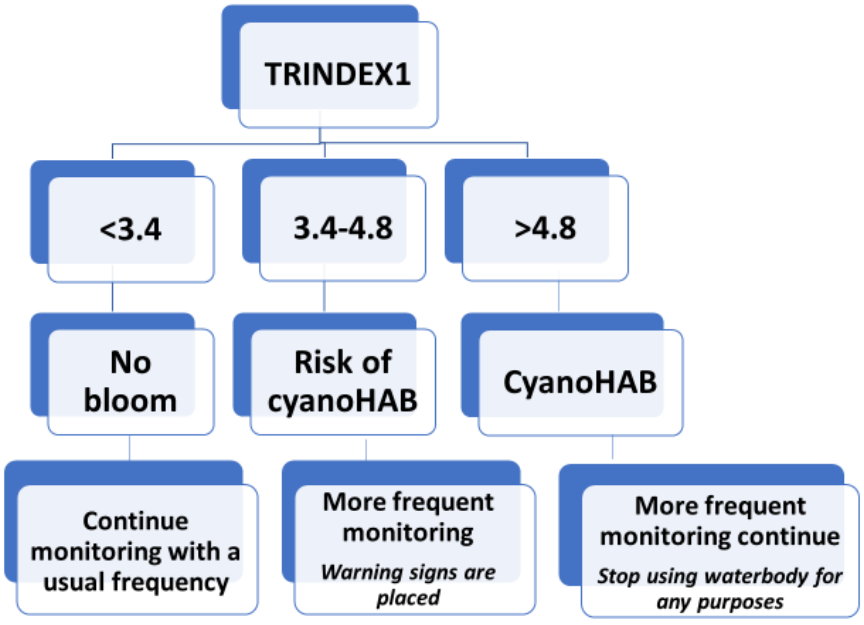


Figure 4.9: TRINDEX application for prediction of bloom onset

Figure 4.9 shows three scenarios of risk that could lead to a management decision for waterbody when dealing with bloom issue. 1) When no visible bloom conditions and TRINDEX<3.4, the monitoring plan for waterbody should follow its established routine; 2) However, if no visible bloom on the surface, but TRINDEX of the lake goes between 3.4-4.8, the risk for a cyanoHAB growth increases. A more frequent sampling plan with all nutrient parameters, plus taxonomy and toxins should be put in place. Also, the early warning signs could be placed at the affected areas in order to inform residents about the algal growth concerns. 3) In case of TRINDEX calculated greater than 4.8, the risk of blooming issues is supposed very high, and blooms even could be observed (on the surface) or not clearly observed (in case of their dissipation in the water column). In this last scenario, any activities for people and pets must be absolutely restricted from concerned waterbody. The monitoring plan should be more intensive during the bloom episodes.

CHAPTER 5 CONCLUSIONS

Work done and Innovation of the research

- Via the dataset of over 10 waterbodies in Nova Scotia and Moncton area of New Brunswick, two main distribution directions of cyanoHAB development in the region are discovered and monitored. Main species causing blooms in the region are well identified including *Dolichospermum flos-aqua, Dolichospermum planctonicum, Microcystis aeruginosa* and *Aphanizomenon sp*.
- With these dominant species for toxic and potential toxic blooms, associated cyanotoxins are also identified. The main variant MC-LR of microcystins, among more than 200 known congeners around the world, was detected in many of the waterbodies cited above. The second congener MC-LY was found additionally in Lake Torment only in fall 2017-2018. For anatoxin family, Anatoxin-a (only one analog of anatoxins) was detected in 2017 by LC-HRMS at certain waterbodies in our study program. However, only small concentrations were detected in some samples.
- The use of three methods for microcystin analyses is a pioneering work and this thesis can contribute to cyanotoxin research, not only in the Canadian context but also globally. The comparison and cross-validation of these three methods were well confirmed by Pearson's coefficients of correlation. The thesis results show hence the importance of a long-term monitoring plan for cyanobacterial blooms to improve the water management system in Atlantic Canada. An accurate detection and estimation for the HAB consequences in freshwater are totally necessary in this region.
- The suggestion of TRINDEX and its applications in freshwater can be considered as innovative technique for bloom prediction indicator. The proposed TRINDEX indicator can show the variation of cyanobacterial dominance in phytoplankton. Moreover, TRINDEX can orient towards the determination of threshold for HAB occurrence. Different combinations of nutrients and pigments were introduced to

identify the best combination of factors that will influence on HAB before the formation of cyanoHAB bloom occurs.

- It is also the first time the ROC analysis for discrimination tests is introduced in algal research to determine the threshold ranges. From there, a practical and easy applicable scheme to evaluate the bloom onset based on TRIDEX is suggested for water management and monitoring.

Significance of the work and contributions to scientific knowledge

- Prediction of cyanobacteria growth as well as cyanoHAB occurrence and proliferation will bring benefits to Nova Scotia, Atlantic Canada and the entire Canada where cyanobacterial blooms become yearly a very common and worrisome issue, as they are generated quite often by potentially toxic species. The outcomes of this work can significantly contribute to development of a monitoring plan and management of water quality, public health as well as to the lake management. Scientists and engineers from the water or algae-related fields, as well as ecologists can also get benefits from this research.
- This work is the first done in the identification of cyanoHAB in NS and Eastern NB. Two dominant directions of cyanobacterial distribution were discovered. The cyanotoxins associated with these species are also detected. This step contributed significantly in the limnology knowledge, enrich the dataset of algal and environmental research for two concerned provinces, as well as to open many research avenues (see below).
- This work is also the first one in systematic research on cyanoHABS and cyanobacterial toxins in both NS and eastern NB.
- TRINDEX is the first indicator suggested for bloom onset prediction of cyaoHAB in the literature of algal research. This is not a single parameter index, but a combination of many involving factors in algal blooming. The proposed TRINDEX indicator significantly contributed not only to the study of cyanoHAB for detection of the bloom onset, but also to the

mathematical approaches and modeling stage with ROC analysis. The latter can be concluded as a state-of-art approach for the discrimination of binary data. The ROC discrimination test could be extended to many other environmental studies to deal with threshold issues.

Future perspectives

The thesis has opened many new avenues for the research on HAB in Atlantic Canada as well as in the entire Canada where numerous waterbodies were reported with invading harmful algal blooms last several years. Some of these avenues can be suggested as follows.

- Our proposed TRINDEX should be tested with other lakes to confirm its ability to predict the beginning of cyanoHAB.
- TRINDEX should be assessed with other factors such as wind and light intensity
- Threshold determination needs to be more studied in the future in combining all of stressors for the onset of blooms.
- More lakes and watersheds in NS and NB should be included in the general plan of study to support two distributions of cyanobacterial species or/and discover the new directions.
- Determination of mutual relationships between toxin concentration and dominant cyanobacterial species detected in different seasons under different environmental conditions is a real need for water management and monitoring, leading to the development of bloom risk management program in Atlantic provinces.

BIBLIOGRAPHY

Ahn C.Y., Joung S.H., Yoon S.K., and Oh H.M. (2007). Alternative alert system for cyanobacterial bloom, using phycocyanin as a level determinant. Journal of Microbiology, 45(2): 98-104.

Allen J.I., Smyth T.J., Siddorn J.R., and Holt M. (2008). How well can we forecast high biomass algal bloom events in a eutrophic coastal sea? Harmful algae 8: 70–76.

Amerine M.A., Pangborn R.M., and Roessler E.B. (1965). Principles of Sensory Evaluation. Academic, New York.

Anderson D.M., Glibert P.M. and Burkholder J.M. (2002). Harmful Algal Blooms and Eutrophication: Nutrient Sources, Composition, and Consequences. Estuaries 25 (4b): 704–726.

Asadollahfardi G. (2015). Water Quality management. Assessment and Interpretation. Springer.

Bellinger E.G., and Sigee D.C. (2015). Freshwater Algae: Identification, Enumeration and Use As Bioindicators – 2e. John Wiley & Sons, Ltd.

Benayache N.-Y., Nguyen-Quang T., Hushchyna K., McLellan K., Afri-Mehennaoui F.-Z. and Bouaïcha N. (April 3rd, 2019). An Overview of Cyanobacteria Harmful Algal Bloom (CyanoHAB) Issues in Freshwater Ecosystems [Online First], IntechOpen, DOI: 10.5772/intechopen.84155.

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

Brown C.D. and Davis H.T. (2006). Receiving operating characteristics curves and related decision measures: A tutorial. Chemometrics and Intelligent Laboratory Systems 80: 24-38.

Bruno M., Ploux O., Metcalf J. S., Mejean A., Pawlik-Skowronska B., and Furey A. (2017). Anatoxin-a, Homoanatoxin-a, and Natural Analogues. In: Meriluoto J., Spoof L., and Codd G.A. (eds), Handbook on Cyanobacterial Monitoring and Cyanotoxin Analysis, John Wiley & Sons, Ltd, Chichester, UK: 138–147.

Brylinksy M. (2009). Lake Utopia Water Quality Assessment. Final report, Dept. Environment of New Brunswick, Canada.

Camacho F.A. (2008). Macroalgal and Cyanobacterial Chemical Defenses in Freshwater Communities. In Amsler C. D. (ed). Algal Chemical Ecology. Springer-Verlag Berlin Heidelberg: 105-116.

Canadian Council of Ministers of the Environment (CCME) (2004). Canadian water quality guidelines for the protection of aquatic life: Phosphorus: Canadian Guidance Framework for the Management of Freshwater Systems. In: Canadian environmental quality guidelines, 2004, Canadian Council of Ministers of the Environment, Winnipeg.

Canadian Council of Ministers of the Environment (2011). Protocols manual for water quality sampling in Canada. PN 1461 ISBN 978-1-896997- 7-0 PDF.

Carter J.V., Pan J., Rai S.N., and Galandiuk S. (2016) ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery, 59 (6): 1638–1645.

Carey C.C., Ibelings B.W., Hoffmann E.P., Hamilton D.P., and Brookes J.D. (2004). Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. Water research 46: 1394-1407.

Carlson R.E. (1977). A trophic state index for lakes. Limnology and Oceanography 11: 361-369.

Catherine A., Bernard C., Spoof L., and Bruno M. (2017). Microcystins and Nodularins. In: Meriluoto J., Spoof L., and Codd G.A. (eds), Handbook on Cyanobacterial Monitoring and Cyanotoxin Analysis, John Wiley & Sons, Ltd, Chichester, UK: 109–126.

Chaffin J.D. and Bridgeman T.B. (2014). Organic and inorganic nitrogen utilization by nitrogen-stressed cyanobacteria during bloom conditions. Journal of Applied Phycology 26: 299–309.

Chen Y., Niu Z. and Zhang H. (2013). Eutrophication assessment and management methodology of multiple pollution sources of a landscape lake in North China. Environmental Science and Pollution Research vol.20: 3877–3889.

Chorus I. and Bartram J. (1999). Toxic Cyanobacteria: A guide to their public health consequences, monitoring and management. WHO (World Health Organization), E & FN Spon. London.

Colón-López M., Sherman D.M., and Sherman L.A. (1997). Transcriptional and translational regulation of nitrogenase in light-dark- and continuous-light-grown cultures of the unicellular cyanobacterium *Cyanothece* sp. Strain ATCC 51142. Journal of Bacteriology 179: 4319–4327.

Corbett M.C, Hu Y.L., Fay A.W., Ribbe M.W., Hedman B., and Hodgson K.O. (2006). Structural insights into a proteinbound iron-molybdenum cofactor precursor. Proceedings of the National Academy of Sciences USA 103:1238-1243.

DeLong E.R., DeLong D.M., and Clarke-Pearson D.L. (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837-845.

Edler L. and M. Elbrächter (2010). The Utermöhl method for quantitative phytoplankton analysis. Microscopic and Molecular Methods for Quantitative Phytoplankton Analysis. Manuals and guides. Editors B.Karlson, C. Cusack and E. Bresnan, UNESKO: 13-20.

Elkadiri R., Manche C., Sultan M., Al-Dousari A., Uddin S., Chouinard K., and Abotalib A.Z. (2016). Development of a Coupled spatiotemporal Algal Bloom Model for Coastal Areas: A Remote Sensing and Data Mining-Based Approach. IEEE Journal of selected topics in applied Earth observations and remote sensing 9 (11): 5159-5171.

EPA (2013). Impacts of Climate Change on the Occurrence of Harmful Algal Blooms.Office of Water, EPA 820-S-13-001.

Fleming I. and Williams D.H. (1966). Spectroscopic methods in organic chemistry. McGraw-Hill Publishing Company Limited.

Freshwater Algae of North America Ecology and Classification (2003). Edited by Wehr J.D. and Sheath R.G. Amsterdam, Boston: Academic Press.

García-Fernández J.M., Tandeau de Marsac N., and Diez J. (2004). Streamlined regulation and gene loss as adaptive mechanisms in *Prochlorococcus* for optimized nitrogen utilization in oligotrophic environments. Microbiology and Molecular Biology Reviews 68:630–638.

Glibert P. M., Anderson D. M., Gentien P., Graneli E., and Seller K.G. (2005). The global, complex phenomena of Harmful Algal Blooms. Oceanography Vol.18 (2): 130-141.

Hamilton G., McVinish R., and Mengersen K. (2009). Bayesian model averaging for harmful algal bloom prediction Ecological Applications 19 (7): 1805–1814.

Harada K., Imanishi S., Kato H., Mizuno M., Ito E., and Tsuji K.K. (2004). Isolation of Adda from microcystin-LR by microbial degradation. Toxicon, 44 (1): 107-109. DOI: 10.1016/j.toxicon.2004.04.003.

Health Canada (2012). Guidelines for Canadian Recreational Water Quality, Third Edition. Water, Air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario. (Catalogue No H129-15/2012E).

Health Canada (2014). Guidelines for Canadian Drinking Water Quality—Summary Table. Water and Air Quality Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Humbert J.-F. and Fastner J. (2017). Ecology of Cyanobacteria. In Meriluoto J., Spoof L., and Codd G.A. (eds) Handbook of Cyanobacterial Monitoring and Cyanotoxin Analysis, First Edition. John Wiley & Sons, Ltd, Chichester, UK: 10-18.

Hushchyna, K. and Nguyen-Quang T. (2017). Using the modified Redfield ratio to estimate harmful algal blooms. Environmental Problems 2 (2): 101-108.

Hushchyna K. and Nguyen-Quang T. (2018). Research and monitoring directions for Harmful Algal Blooms (HAB) in the Nova Scotia (NS) and New Brunswick (NB) freshwater. An abstract. 2nd Interdisciplinary Freshwater Harmful Algal Blooms Workshop, April 16th-18th, Toronto, ON, Canada.

Kotut K., Ballot A., and Krienitz L. (2006). Toxic cyanobacteria and their toxins in standing waters of Kenya: implications for water resource use. Journal of Water Health 4(2): 233-45.

Lerman D. C., Tetreault A., Hovanetz A., Bellaci E., Miller J., Karp H., Mahmood A., Strobel M., Mullen S., Keyl A., and Toupard A. (2010). Applying signal-detection theory to the study of observer accuracy and bias in behavioral assessment. Journal of Applied Behavior Analysis 43(2): 195–213. http://doi.org/10.1901/jaba.2010.43-195

Lis, H., Kranzler, C., Keren, N. and Shaked, Y. (2015). A Comparative Study of Iron Uptake Rates and Mechanisms amongst Marine and Fresh Water Cyanobacteria: Prevalence of Reductive Iron Uptake. Life 5: 841-860.

Liu Y., Chen W., Li D., Huang Z., Shen Y., and Liu Y. (2011). Cyanobacteria-/cyanotoxin-contaminations and eutrophication status before Wuxi drinking water crisis in Lake Taihu, China. Journal of Environmental Science (China) 23(4): 575-81.

Lund J.W.G., Kipling C., and Le Cren E.D. (1958). The Inverted Microscope Method of Estimating Algal Numbers and the Statistical Basis of Estimations by Counting. Freshwater Biological Association, The Ferry House, Ambleside Westmorland, England: 144-177.

Markou G., Vandamme D., and Muylaert K. (2014). Microalgal and cyanobacterial cultivation: The supply of nutrients. Review. Water research 65:186-202.

Marty J. and Ritchie M. (2015) Kings County lake monitoring, 2014 season. Report. Municipality of the County of Kings, Nova Scotia, Canada.

Marty J. and Reardon M., (2016). King county lake monitoring program 2015 season. King County Lake Monitoring Program 2016 Season. Report. Municipality of the County of Kings.

McLellan N.L. and Manderville R.A. (2017). Toxic mechanisms of microcystins in mammals. Toxicology Research 6: 391–405.

Meeks J.C. and Elhai J. (2002). Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiology and Molecular Biology Reviews 66: 94–121.

Meilgaard M., Civille C.V., and Carr B.T. (1991). Sensory Evaluation Techniques. CRC, Boca Raton, FL.

Miles C.O. and Stirling D. (2017). Microcystin mass calculator, version 15 from 01 May 2017. http://www.toxinology.no/Downloads.aspx. DOI: 10.13140/RG.2.2.27959.11688 · License: CC0.

Molot, L.A., Watson, S.B., Creed, I.F., Trick, C.G., McCabe, S.K., Verschoor, M.J., Sorichetti, R.J., Powe, C., Venkiteswaran, J.J., and Schiff, S.L. (2014). A novel model for cyanobacteria bloom formation: the critical role of anoxia and ferrous iron. Freshwater Biology 59: 1323–1340.

Ndong M., Bird D., Nguyen-Quang T., Boutray M., Zamyadi A., Vincon-Leite B., Lemaire J. B., Prevost M., and Dorner S. (2014). Estimating the risk of cyanobacterial occurrence using an index integrating meteorological factors: Application to drinking water production. Water Research. Vol.56: 98–108.

Nguyen-Quang T., Fraser H., Hushchyna K., McLellan K., and Sabir, Q. (2018). Continuation of the systematic study (2nd year – 2017) of toxic algal blooms in McLaughlin and Irishtown Reservoirs to add to the water quality program of the City of Moncton- New Brunswick. Technical Report submitted to Government of New Brunswick, Environmental Trust Fund, Environmental and Local Government, Canada.

Nguyen-Quang T., Fraser H., Hushchyna K., and McLellan K. (2019). Toxic algae in the Moncton city drinking water resources: an urgent systematic investigation plan for species identification, toxicity determination, and prediction of bloom proliferation. Technical Report submitted to Government of New Brunswick, Environmental Trust Fund, Environmental and Local Government, Canada.

Nimptsch J., Woelfl S., Osorio S., Valenzuela J., Moreira C., Ramos V., Castelo-Branco R., Nuno Leao P. and Vasconcelos V. (2016). First record of toxins associated with cyanobacterial blooms in oligotrophic North Patagonian lakes of Chile—a genomic approach. International Review of Hydrobiology 101: 57-68.

Novotny V. and Olem H. (1994). Water Quality: Prevention, Identification, and Management of Diffuse Pollution. Van Nostrand Reinhold.

Oliver R.L., Hamilton D.P., Brookes J.D. and Ganf G.G. (2012). In Whitton B. A. (ed.) Physiology, Blooms and Prediction of Planktonic Cyanobacteria. Ecology of Cyanobacteria II. Their Diversity in Space and Time. Springer: 155-195.

Osswald J., Rellán S., Gago A., and Vasconcelos V. (2007). Toxicology and detection methods of the alkaloid neurotoxin produced by cyanobacteria, anatoxin-a. Environment International 33: 1070–1089.

Pearl H.W., Fulton R. S., Moisander P.H. and Dyble J. (2001). Harmful freshwater algae blooms with anemphasis on Cyanobacteria. Review article. The scientific world, 1: 76-113.

Pearl H. W. and Paul V. J. (2012). Climate change: Links to global expansion of harmful cyanobacteria. Water research 46: 1349-1363.

Pearl H.W. and Otten T.G. (2013). Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls. Environmental Microbiology. DOI 10.1007/s00248-012-0159-y.

Peryam D.R. (1958). Sensory difference tests. Journal of Food Technology, 12: 231–236.

Pick F.R. (2016). Blooming algae: a Canadian perspective on the rise of toxic cyanobacteria. Canadian Journal of Fisheries and Aquatic Sciences 73: 1149–1158.

Prescott G.W (1962). Algae of the western great lakes area. WM.C.Brown CO.INC, Dubuque, Iowa, USA, 977p.

Reynolds C.S. (1984). The ecology of freshwater phytoplankton. Cambridge University Press, Cambridge, London, New York, New Rochelle, Melbourne, Sydney.

Ritchie R.J. (2013). The Ammonia Transport, Retention and Futile Cycling Problem in Cyanobacteria. Microbial Ecology 65: 180–196.

Samdal I.A., Lovberg K., Ballot A., and Miles C.O. (2013). Microcystin and nodularin ELISA - A new approach to developing generic antibodies. Poster at Marine and Freshwater Toxins Analysis, Fourth Joint Symposium and AOAC Task Force Meeting Baiona, Spain May 5-9.

Samdal I.A., Ballot A., Lovberg K.E., and Miles C.O. (2014). Multihapten approach leading to a sensitive ELISA with broad cross-reactivity to microcystins and nodularin. Environmental science & technology, 48: 8035-8043.

Schisterman E.F, Perkins N.J., Liu A., and Bondell H. (2005). Optimal cut-point and its Corresponding Youden Index to discriminate individuals using pooled blood samples. Epidemiology 16 (1): 73-81.

Šejnohová L. and Maršálek B. (2012). Microcystis. In Whitton B. A. (ed.) Ecology of Cyanobacteria II. Their Diversity in Space and Time editing. Springer: 195-228.

Serwan M. and Baban J. (1996). Trophic classification and ecosystem checking of lakes using remotely sensed information. Hydrological Sciences Journal 41(6): 939-957.

Smith C. and Sutton A. (1993). The Persistence of Anatoxin-a in Reservoir Water, Report no. FR0427. Foundation for Water Research, UK.

Sorichetti R.J., Creed I.F. and Trick C.G. (2014). The influence of iron, siderophores and refractory DOM on cyanobacterial biomass in oligotrophic lakes. Freshwater Biology 59: 1423–1436.

Spoof L. and Catherine A. (2017) Appendix 3: Tables of microcystins and nodularins. In: Meriluoto J., Spoof L., and Codd G.A. (eds), Handbook on Cyanobacterial Monitoring and Cyanotoxin Analysis, John Wiley & Sons, Ltd, Chichester, UK: 526–537. https://doi.org/10.1002/9781119068761.app3.

Stantec Consulting LTD (2017) Results of 2017 water quality survey of eleven lakes in Yarmouth and Digby Counties, final report for Carleton River Watershed Area Water Quality Steering Committee, Darmouth, NS, Canada.

Stone H., and Sidel J.L. (1993). Sensory Evaluation Practices, 2d ed. Academic, San Diego, CA.

Stow C.A., Jolliff J., Mc Gillicuddy D.J., Doney S.C., Allen J.I., Friedrichs M.A.M., Rose K.A., and Wallhead P. (2009). Skill assessment for coupled biological/physical models of marine systems. Journal of Marine Systems 76: 4-15.

Suttle C.A. and Harrison P.J. (1988). Ammonium and phosphate uptake rates, N:P supply ratios, and evidence for N and P limitation in some oligotrophic lakes. Limnology and Oceanography 33(2): 186-202.

Taylor D. (2009). A water quality survey of nine lakes in the Carleton river watershed area. Yarmouth County, NS. Water & Wastewater Branch Nova Scotia Environment.

Thomas R.H. and Walsby A.E. (1986). The effect of temperature on recovery of buoyancy by Microcystis. Journal of General Microbiology 132: 1665-1672.

Utermohl H. (1958). Zur Ver vollkommung der quantitativen phytoplankton-methodik. Mitteilung Internationale Vereinigung Fuer Theoretische unde Amgewandte Limnologie, 9.

Venkiteswaran J. (2019). Constant nutrient loads yield changing annual cyanobacteria bloom dynamics. An abstract. 3nd Interdisciplinary Freshwater Harmful Algal Blooms Workshop, April 24th-26th, Toronto, ON, Canada.

Vollenweider R.A. (1975). Input-output model with special reference to the phosphorus loading concept in limnology. Schweiz.Z.Hydrolo vol.37: 53-58.

Vollenweider R. A., Rinaldi A. and Montanari G. (1992). Eutrophication, structure and dynamics of a marine coastal system: results of ten years monitoring along the Emilia-Romagna coast (Northwest Adriatic Sea), in Marine Coastal Eutrophication. Science of the Total Environment: 63-106.

Vollenweider R.A., Giovanardi F., Montanari G., and Rinaldi A. (1998). Characterization of the trophic conditions of marine coastal waters, with special reference to the NW Adriatic Sea. Proposal for a trophic scale, turbidity and generalized water quality index. Environmetrics, Vol. 9: 329–357.

Wetzel R. G. (2001). Limnology. Lake and River Ecosystems. 3rd ed. Academic Press.

Witte R.S. and Witte J.S. (2010). Statistics, 9th ed., John Wiley&Sons, Inc.

World Health Organization (WHO) (2003a). Guidelines for drinking-water quality. 2nd ed. Addendum to Vol. 2. Health criteria and other supporting information., Geneva, 1998.

World Health Organization (WHO) (2003b). Guidelines for safe recreational water environments. Vol. 1, Coastal and fresh waters.

Youden W.J. (1950). An index for rating diagnostic tests. Cancer 3: 32-35.

YSI (2010). YSI 9300 and 9500 direct-read photometer. User manual, YSI Incorporated.

Zaffiro A., Rosenblum L., Wendelken S.C. (2016). Method 546: Determination of Total Microcystins and Nodularins in Drinking Water and Ambient Water by Adda Enzyme-Linked Immunosorbent Assay. EPA, USA.

APPENDIX A CORRELATION MATRICES BETWEEN PARAMETERS

Table A1: Correlation matrix of chosen parameters in Mattatall Lake

	D%O	**Chl-a**	**PC**	**PO4**
R (Chl-a)	0.022			
p-value	0.711			
R (PC)	-0.003	*0.952*		
p-value	0.962	0.000		
R (PO4)	*0.345*	-0.005	-0.014	
p-value	0.000	0.928	0.812	
R (NO3)	-0.096	-0.052	-0.098	-0.015
p-value	0.092	0.373	0.086	0.793

Table A2: Correlation matrix of chosen parameters in Torment Lake

	D%O	**Chl-a**	**PC**	**PO4**
R (Chl-a)	0.017			
p-value	0.827			
R (PC)	-0.013	*0.351*		
p-value	0.862	0.000		
R (PO4)	-0.104	*0.202*	0.042	
p-value	0.178	0.008	0.587	
R (NO3)	-0.142	*0.445*	*0.163*	-0.023
p-value	0.064	0.000	0.034	0.768

D%O – deviation of oxygen from 100%
Chl-a – chlorophyll-a
PC – phycocyanin
NO3 – nitrate
PO4 – phosphate
R – Pearson's coefficient of correlation – Italic number: correlation should be considered

APPENDIX B APPLICATION OF TRINDEX

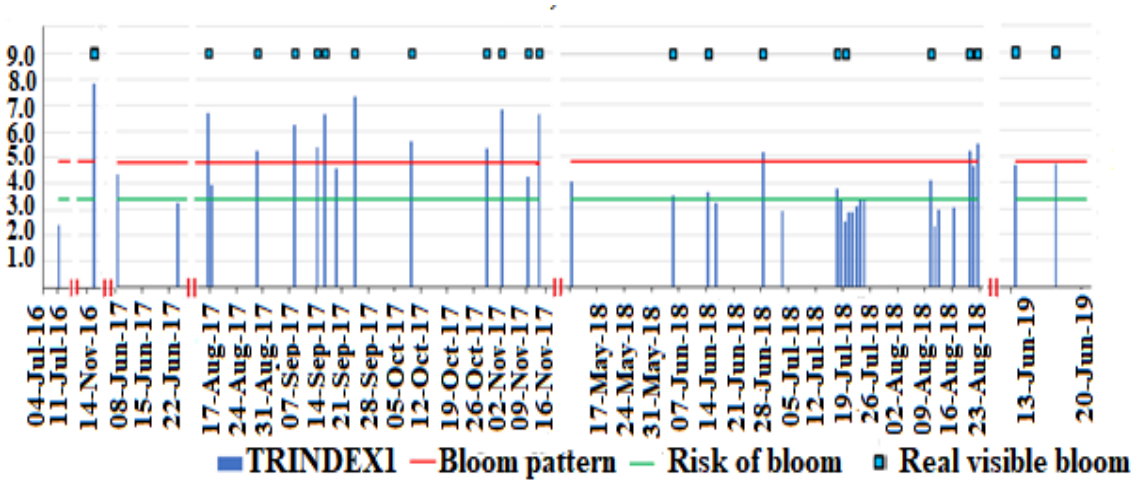


Figure B1: Routine observations of bloom onsets at the same location, TDave

Figure B1 summarizes observations and shows TRINDEX1 thresholds: for transitional phase beginning (green line) and for bloom onset (red line). The transitional phase is defined in between these two lines. These forty-one samples from regular observations during summers 2016-2019 at TDave are compared with the real visible blooms: there were 22 cases in total with stable blooms or just bloom onset. Among them thirteen cases with TRINDEX1 greater than 4.8 (stable bloom cases) and nine cases with TRINDEX1 greater 3.4 but less than 4.8 (transitional cases).

The evaluation of the model performance was done via Accuracy, Precision, Recall and F1 Score metrics. Precisely, among these 40 observations, we have:

- Four incorrect positive cases (10%)
- Twenty-two true positive (TP) cases
- One incorrect negative (FN) case: 5%
- Thirteen true negative (TN) cases:

The accuracy of the model is defined by $\frac{TP+TN}{TP+FP+FN+TN}$ and corresponds to 87.5%.

The recall or true positive rate (TPR) is $\frac{TP}{TP+FN}$, which is 95.7% in our case. This ratio defines how many real positive cases the model can predict.

The precision of the model is given by the ratio $\frac{TP}{TP+FP}$ which equals to 84.6%. Precision can show how accurate the model predicts actual positive events.

Finally, the F_1 score is used to measure the balance between precision and recall and not focussing on actual negatives because the large number of true negative cases can influence on the accuracy of our predictions:

$$F_1 = \frac{2*(Recall*Precision)}{Recall+Precision} = \frac{2*(0.957*0.846)}{0.957+0.846} = 0.898$$

In summary, the TRINDEX1 model which is applied to the real observation data from TL for 3 summers is 87.5% accurate, with a recall 0.95 (which is excellent as far above 0.5), a precision of 84.6%, and a F_1 score near 0.9 (which is also very good as F_1 defined in the range from 0 (bad test) to 1 (excellent test)).