2025

Investigating the Thermal Stratification Patterns of a Shallow Polymictic Boreal Shield Lake and the Effectiveness of Models in Predicting Shallow Lake Thermal Profiles

Matthew Roberts
robe4640@mylaurier.ca

INVESTIGATING THE THERMAL STRATIFICATION PATTERNS OF A SHALLOW POLYMICTIC BOREAL SHIELD LAKE AND THE EFFECTIVENESS OF MODELS IN PREDICTING SHALLOW LAKE THERMAL PROFILES

by

Matthew George Roberts

Bachelor of Science, Wilfrid Laurier University, 2020

THESIS

Submitted to the Department of Geography and Environmental Studies in partial fulfilment of the requirements for

Master of Science in Geography

Wilfrid Laurier University

Abstract

Thermal stratification is a core process of lakes which plays a critical role in shaping the ecological dynamics of lakes, influencing major processes like nutrient cycling and oxygen availability. For shallow polymictic lakes, thermal stratification patterns present some unique differences due to the complex and variable nature of these lakes.

In this thesis, I examined thermal stratification in a shallow polymictic lake in northwestern Ontario during the ice-free season of 2019 to investigate the frequency, duration, and driving factors of thermal stratification. The lake underwent 146 separate stratification events, with a median duration of 2 hours per event, with a total time spent stratified of almost one third the entire study period. The primary drivers of this thermal stratification were air temperature, wind direction, and wind speed, underscoring their significance in the thermal regime of the lake. Also observed was a brief period of hypolimnetic hypoxia prior to a phytoplankton bloom, highlighting potential influence from internal nutrient loading from sediment to have occurred under low redox conditions.

Another aspect of this investigation involved the evaluation of five one-dimensional models on their performance in reproducing the observed thermal stratification patterns in both the shallow polymictic lake and an adjacent shallow dimictic lake. While some models demonstrated success in simulating thermal stratification, particularly in the dimictic lake, they struggled to accurately capture the dynamic patterns observed in the polymictic lake. The most effective model calculated only 55 stratification events compared to the 146 that were observed

in Lake 303, suggesting the need for highly specific modeling approaches tailored to shallow polymictic lakes.

These findings exhibit the complexity of thermal stratification dynamics in shallow polymictic lakes and highlight the importance for additional field observations as well as modeling efforts to enhance our understanding of this lake type. Improving the state of modeling for these lakes is essential for accurately representing the unique nature of shallow polymictic lakes, which is crucial for effective lake management and conservation strategies.

Acknowledgements

There are truly a vast number of people who come to mind when I think about all the people who deserve my thanks. The process of grad school and specifically writing a thesis is not a simple one, and I dare say an impossible one without the support of others. I still remember my excitement in finding out I had been accepted into this program. It was in March of 2020, a very rough time during the development of the CoVid-19 pandemic where all I was allowed to do was remain indoors, a very tall order for students renting just one of several bedrooms in apartments far away from home. Still, I celebrated as best I could, a few jumps and a yell inside my cramped bedroom; the acceptance was a light in the dark, an outlook to a hopeful future.

I want to first start offering my thanks with my most important one. Never has there been a person in my life to whom I could owe so many things for so many different reasons. To my dearest, Amanda, I couldn't have done this without you. Thank you for your kind words, your assistance with keeping organized, your motivational support, and most of all your unwavering belief in me.

Many thanks also go out to Dr. Jason Venkiteswaran, my supervisor and academic mentor. Jason has helped me a great amount, like teaching me to code in R, improving my writing and communication, developing my critical thinking, as well as allowing me to solve my own problems to the best of my abilities first before seeking help.

Without a doubt I also have my family to thank. I'll start with my sweet Ila, a cat so gentle most wouldn't believe it. As my roommate throughout undergrad and grad school, she truly was there for me through it all (but it's not like she had much of a choice, lol). If only she was able to

really understand the impact she has had on my life, and any others who have been lucky to meet such a skittish creature. Thank you to my mom and dad. You two have helped me in an undefined number of ways to make so many of the things in my life even possible. No parent could do a better job in raising me and helping me achieve my goals as they have done and as they continue to do. To my sisters, thank you for shaping me into the man I am, with all the grit and patience one develops growing up with two older sisters.

A large thanks is also in order to my friends. Friendship provides so much joy and truly allows for an escape from all the stressors of life. Thank you to the day-after-day shenanigans with the science atrium crew, including Joel, Mo, Lina, Rema, and all the fellow troublemakers. Thank you to the Queen and Spadina McDonalds (not the actual restaurant) for all the fun and good times. To my friends from the sand box, you all know who you are, there is no greater comfort than knowing our 18-year friendship is untouchable.

I also need to express gratitude to the hydromicro group. As one of the few social functioning groups during the CoVid-19 pandemic, our hydromicro meetings served as both a great learning opportunity and some much needed face-to-face (virtual) interaction. Thank you to Ryan, Phil, Cat, Jenn, Raisa, and Jeremy, for being my peers and my motivators, as well as for some good times at the Gradhouse.

To my committee members, thank you for your incredible feedback and contractive criticisms. It is both an art and a skill you each possess to be able to communicate a combination of firm but positive comments on a person's hard work.

Were nearly finished, so let's shift to my thanks to Wilfrid Laurier University and Kitchener/Waterloo for being my home from 2015 until 2022. Upon first arriving in KW, I never imagined I'd grow to appreciate the city as much as I did in my years living there. Thank you to Cold Regions building for being my reliable space for Wi-Fi (but not always, Eduroam, lol) and solitude to focus on my schoolwork.

The final thank you probably often goes without saying, and maybe sometimes unconsidered, but thank you to you, the reader, for taking your time to recognize this accomplishment of mine. It truly has been a hard-fought battle to get here, and I could never ask someone to spend their time rummaging through this amalgamation of scientific jargon and ideas. I hope you enjoy, but at the very least, that you learn at least one thing which is the true measure of success by this thesis in my opinion.

Table of Contents

List of Figures

Chapter 1 Introduction

1.1 Thermal Stratification

The studying of our aquatic ecosystems forms a cornerstone of environmental science, offering insight into the intricate processes taking place within our water bodies. Of the many processes and phenomenon that exist regarding lakes, thermal stratification is one in particular which is heavily interconnected with many important topics in limnology. Thermal stratification fundamentally shapes the physical, chemical, and biological dynamics of lakes, with ties to a wide variety of things such as nutrient cycling, phytoplankton blooms, and species distribution (Orihel et al. 2015). Exploring the complexities of thermal stratification not only contributes to the understanding of lake ecosystems but also provides invaluable insights into broader environmental challenges, making it a compelling subject for scientific research as well as conservation efforts.

Thermal stratification in lakes is the natural layering of lake waters by temperature, with warmer water being less dense and floating atop colder denser water. The epilimnion is the layer which ranges from the surface to the top of the metalimnion, if any, or the hypolimnion (Wetzel 2001). Being at the surface, this layer is typically the warmest and features the majority of photosynthetic activity. The epilimnion is subject to temperature changes through its contact with the atmosphere and can be mixed due to wind. Radiation from sunlight contributes to temperature changes in the epilimnion but in some cases radiation penetrates deeper into lakes. The metalimnion is a transition layer which may be present between the epilimnion and the hypolimnion. The top of the metalimnion is often identified by the presence of the thermocline

which is the depth at which the temperature decline is steepest. Lastly, the bottom-most layer is the hypolimnion where the coldest and, in-turn, the densest waters are found.

The factors that drive mixing and changes to lake thermal profiles are wind, solar radiation, air temperature and occasionally precipitation. Wind induces mixing by generating surface waves as wind passes along the lake surface. The power and height of surface waves for shallow lakes is dependent on wind velocity, lake depth, and the fetch of the lake (Blottiere 2016). Wetzel (2001) described wave height (h) as function lacking good theoretical explanation, proportional to the square root of the fetch (x), expressed as:

$$h = 0.105\sqrt{x}$$

Fetch is the distance which wind blows unobstructed along the lake surface. With a longer fetch, waves become increasingly taller and wider, generating greater force which can induce mixing within the water column. When the wavelength of the surface wave is greater than twice the depth of the lake, the water below the surface begins to move in an elliptical orbit (Laenen and Letourneau 1996). These elliptical orbitals then influence the waters below them to move within a new smaller orbital, which is a pattern that can continue to the lake bottom, potentially influencing and resuspending the sediment (Figure 1).

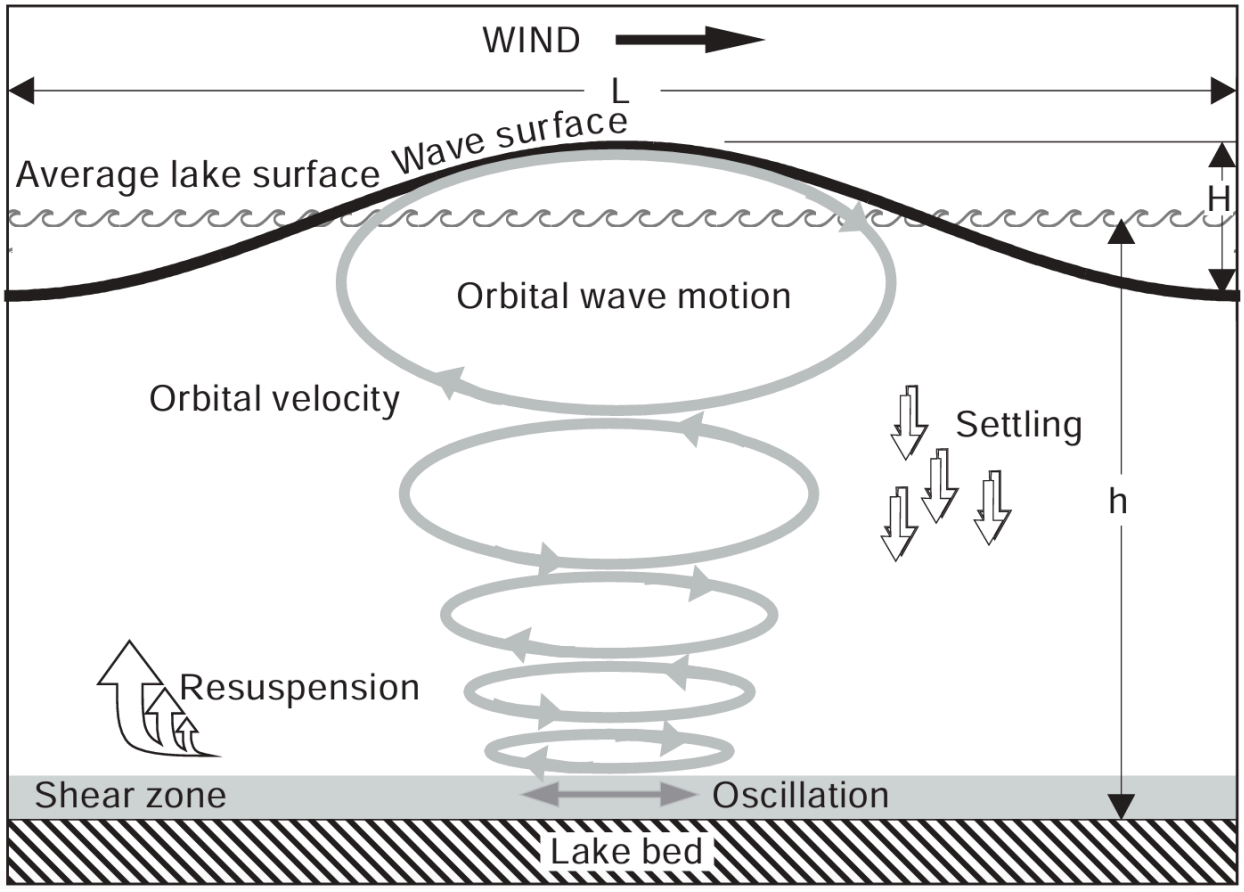

Figure 1.1. Orbital wave motion induced by wind within shallow lakes which can lead to resuspension of sediment. From Laenen and LeTourneau, U.S Geological Survey, Portland Oregon, 1996.

Solar radiation is an important driving force of lake water temperature and thermal stratification. When light energy hits the surface of the lake some is absorbed and dispersed within the upper layer of the water column while the remainder is reflected back into the atmosphere. The amount of light which is absorbed or reflected is dependent on the angle at which the light hits the water surface. The depth to which light penetrates to depends on the concentrations of suspended particles and dissolved colored compounds in the water (Wetzel 2001). Dissolved colored compounds and photosynthetic biota are two of the major controlling

factors for the absorption of light within lakes that facilitate the conversion from light into heat (Kirk 2010). As water transparency decreases with greater concentrations of dissolved colored compounds and photosynthetic biota, light cannot penetrate deeper into the water column therefore warming the upper layers of the lake (Williamson et al. 2019). Solar radiation displays a diel pattern which means that it cycles on a daily basis, peaking during the middle of the day and troughing at night. This diel pattern is particularly important with respect to thermal stratification, especially within polymictic lakes which often feature a strong relationship between stratification and solar radiation (Yang et al. 2018).

Air temperature drives lake water temperatures through consistent contact at the water-air interface which maintains a continual exchange of heat between the two media. The influence of air temperature on water varies, focusing on the gradient between the two. Air temperature also effects lakes differently within regions, but the general relationship describes the atmosphere as a heat sink with changing degrees of effect (Yu et al. 2021). Warmer air temperatures lead to less heat loss while colder air temperatures can influence a strong cooling pattern. During the spring and summer months the air temperatures typically maintain or limit the loss of heat while during the fall and winter lead to greater rates of heat loss. The loss of heat makes lakes more susceptible to wind-induced mixing. Similar to solar radiation, air temperature typically exhibits diel cycles.

Precipitation as a driver of water temperature is not as significant as the aforementioned drivers; however, its influence still deserves recognition. Precipitation can destratify lakes by direct cooling and increased inflow but publications typically only found destratification to occur

during severe storms with large amounts of rain (Znachor et al. 2008; Klug et al. 2012). Rainfall events of 10 mm or greater has been found to potentially reduce water column stratification or alter the depth of the mixed layer and thermocline (Liu et al. 2020).

Stratification and mixing are key components within lakes, with direct links to several physical and chemical interactions. Nutrient mixing is one key process with links to lake mixing, where mixing causes the nutrient rich waters of the hypolimnion upward toward the surface. Nutrient mixing benefits many organisms throughout the water column by improving access to many necessary nutrients. An example of this can be seen in the acquisition of phosphorus by some phytoplankton which cannot descend the water column (Brookes and Ganf 2001). Lake mixing also helps to replenish dissolved oxygen concentrations in the deeper layers of lakes. Oxygen primarily enters lakes at the surface and so during stratification the dissolved oxygen (DO) concentrations within the hypolimnion decrease mostly due decomposition of organic matter by bacteria (Salonen et al. 1992). DO also plays a key role in oxidation-reduction (redox) reactions, notably within the hypolimnion. Redox reactions are important for a number of processes within lakes, including the release of certain nutrients like phosphorus (P) and iron (Fe) from the sediment into the water column in a bioavailable form (Burger et al. 2005). Increased levels of P allow for the formation of harmful phytoplankton blooms, while Fe(II) availability can result in a shift toward cyanobacterial dominance within the bloom (Molot et al. 2014).

Shallow lakes (maximum depth <5 m) make up the majority of lakes worldwide (Downing et al. 2006; Meerhoff and Jeppesen 2009), therefore, research and the understanding around them is important. Surface area and depth are the main physical factors relating to lake mixing,

and as shallow lakes typically have great surface area compared to their depth, they are much more prone to water temperature changes than deep lakes (Maberly et al. 2020; Woolway et al. 2020). This susceptibility to temperature changes also makes them potentially more susceptible to climate change. The connection between climate change and thermal stratification has been an increasingly popular topic in limnology research (Coats et al. 2006; Schneider and Hook 2010; Schmid et al. 2014; Woolway et al. 2017, 2020). Thermal stratification in shallow polymictic lakes is unique in that it is typically observed as short-lived events, lasting anywhere from a few seconds to several days (Wetzel 2001; Kalff 2002; Soulignac et al. 2017). Polymictic lakes are all lakes which stratify either more than twice per year or are constantly mixing, but this variation between the amount of stratification and mixing allows for a very broad grouping of different lakes, despite many significant differences in their specific mixing regimes. The research of thermal stratification in shallow polymictic lakes suffers from a lack of observations and case studies with little information on the frequency and durations of thermal stratification (Woolway et al. 2017).

1.2 Modelling Shallow Polymictic Lakes

Modelling is an essential tool used in forecasting and management of environments which is used in the process of preventing harmful outcomes and effects. Whether purely conceptual or highly specific, models have a wide range of applications and capabilities to assist in the research and understanding of our natural ecosystems. In environmental science, models are a key component of many fields, such as emission computation (air, water, pollution), process

control, ground water, and ecosystem research (Grützner 1996). Limnology is another field which often uses models to study oxygen balances, eutrophication, and toxic substances in fresh water ecosystems (Jørgensen 1995). Models can be run based on a variety of climate scenarios to provide insight on the potential future of ecosystems. The major benefits to modelling are the capability to simulate experiments which might otherwise be expensive while also avoiding any risk associated with experimentation. Another common use case for modelling in limnology, which will be used in this thesis, is for calculating and visualizing lake thermal structures, heat exchange, and thermal stratification.

The types of models which are capable of estimating lake thermal profiles are numerous, but 1-Dimensional (1-D) models are the most popular, primarily due to their simple yet effective nature in understanding lakes as a multitude of stacked layers. Shallow polymictic lakes; however, are an underrepresented study site within environmental research (Woolway et al. 2017). Given the lack of research showcasing 1-D models and shallow polymictic lakes, the models which might best reproduce the observed thermal profiles of these lakes are not widely known. In addition, due to the transient and irregular thermal stratification patterns of shallow polymictic lakes, the effectiveness of 1-D models in mimicking stratification in these lakes has only been investigated in a small number of studies (Wilhelm and Adrian 2008; Stepanenko et al. 2013). The importance of modelling along with the aforementioned importance of understanding thermal stratification patterns highlight the need for identifying models which can accurately represent shallow polymictic lakes.

1.3 International Institute for Sustainable Development Experimental Lakes Area Lake 303 and Lake 304

The International Institute of Sustainable Development Experimental Lakes Area (IISD-ELA) is a research-based organization which specializes in the undertaking of large scale, often whole ecosystem, experiments and the collection of long-term (since 1969) meteorological, hydrological, and limnological data from boreal watersheds and natural unmanipulated lakes. IISD-ELA is approximately 55 km southeast of Kenora, Ontario in the boreal shield landscape of northwestern Ontario, Canada on the traditional land of the Anishinaabe Nation in Treaty 3 territory and the homeland of the Métis Nation. Often referred to the worlds freshwater laboratory, IISD-ELA is able to conduct whole-lake experiments and manipulations on as many as 58 small lakes and their corresponding water sheds to study a variety of topics such as algal blooms, climate change, microplastics, to name a few (Higgins et al. 2018; Mejbel et al. 2023; McIlwraith et al. 2024).The local environment consists of coniferous forests of black spruce, trembling aspen, white birch, jack pine, balsam fir, and tamarack (McCullough and Campbell 1993).

Lake 303 and Lake 304 are two neighboring headwater lakes both chosen for research in this thesis, specifically due to their shallow nature and mixing regimes. Lake 303 is a shallow (mean depth of 1.86 m and max depth of 2.79 m) polymictic lake known to mix several times annually while Lake 304 is a shallow (mean depth of 3.22 m and max depth of 7.22 m) dimictic lake. The two lakes share similar environmental characteristics and geology which allows for an effective comparison.

1.4 Thesis Objectives

The overall objective for this thesis is to contribute to the understanding of transient thermal stratification in shallow polymictic lakes. The importance of this research is highlighted by the lack of high-frequency data case studies as well as the connections it possesses to climate change and harmful phytoplankton blooms.

The following objectives will allow the accomplishment of this purpose:

1. Assess the frequency and durations of thermal stratification in Lake 303, a shallow polymictic lake (Chapter 2);
2. Identify which environmental variables best predict transient thermal stratification in Lake 303 (Chapter 2);
3. Evaluate 1-D models in their ability to calculate the observed thermal stratification for Lake 303 and Lake 304, a shallow dimictic lake (Chapter 3).

Chapter 2 Identifying Transient Stratification, its Potential Drivers, and its Connections to Episodic Hypoxia in Eutrophic Lake 303

2.1 Abstract

Thermal stratification is an important phenomenon affecting major processes and lake conditions like nutrient cycling and anoxia which can have impacts on aquatic organisms and ecosystem management. For shallow polymictic lakes, the development of thermal stratification can be more complex and variable than for deeper, stratifying lakes. Here, the occurrences of thermal stratification in a shallow polymictic lake in northwestern Ontario during the ice-free season of 2019 were investigated with the aim to understand how often shallow polymictic lakes thermally stratify, how long stratification typically lasts, and which pre-conditions most often lead to thermal stratification. Air temperature, wind speed, wind direction, precipitation, and short-wave radiation are analyzed to understand and rank their relative connections and effects due to their known influences on thermal stratification. The lake stratified 146 separate times with a median length of 2 hours, adding to a total of 1102 hours stratified out of the 3316 hours measured during the summer of 2019. Air temperature, wind direction and wind speed had the strongest influences on transient thermal stratification. The lake experienced a brief period of hypolimnetic hypoxia which is also discussed for its implications on internal nutrient loading from the sediment due to low redox conditions.

2.2 Introduction

Transient thermal stratification is a major physical process in many shallow polymictic lakes which has been historically overlooked (Holgerson et al. 2022). Early categorization of lake

thermal structures considered all polymictic lakes to be in a continuous state of mixing; however, thermal stratification has been recorded in many shallow polymictic lakes (Holgerson et al. 2016; Andersen et al. 2019). Thermal stratification is the development of a thermally induced density gradient which restricts physical and chemical mixing between layers. Thermal stratification in shallow polymictic lakes is termed transient, lasting as briefly as a few seconds to several days, and is typically associated with high solar intensity, warm weather, and low wind speeds (Imberger and Patterson 1981; Wetzel 2001; Kalff 2002). The frequency, duration, and seasonality of transient stratification in shallow polymictic lakes have important implications for a variety of physical, chemical, and biological properties including: increased surface temperature, reduced sub-surface dissolved oxygen concentrations (Loewen et al. 2007), increased nutrient (e.g. phosphorus, iron) regeneration and bioavailability (Burger et al. 2005), increased frequency and persistence of phytoplankton blooms (Yindong et al. 2021; Zhang et al. 2022), and even shifting the dominance of phytoplankton blooms towards cyanobacteria (Kanoshina et al. 2003).

Solar radiation is connected to stratification via the absorption of light at the lake's surface which transfers to thermal energy which gets stored into the upper waters and in turn increases the density gradient. Wind physically mixes lakes by producing surface waves that then create turbulence in the water column (Wüest and Lorke 2003). The effects from wind can vary based on the physical characteristics of lakes, where a greater lake fetch can proportionally allow for taller waves and therefore greater turbulence (Wetzel 2001). Air temperature acts as a consistent heat source or sink, relative to surface water temperatures, which can alter the density gradient and make lakes more or less susceptible to wind-induced mixing. Air temperature's heat sink nature comes about by regulating the loss of long-wave radiation, or heat. Precipitation also

has a notable influence to thermal stratification, where stratification and mixing can both come about after a change of the lake's surface temperatures due to the temperature of the incoming rain and/or increase of inflow to the lake. Precipitation also often occurs alongside a decrease in solar-radiation due to the associated cloud cover which in turn can act together to reduce temperature density gradients.

A thermally stratified lake can be broken up into three regions: epilimnion, metalimnion, and hypolimnion. The epilimnion is the warmest layer and is in contact with the atmosphere, while the hypolimnion is the deepest and coldest layer which lines the lakes sediments. The epilimnion and hypolimnion are separated by the metalimnion, a layer typically defined by the presence of the largest decline in water temperature with respect to depth known as the thermocline. As the epilimnion is in contact with the atmosphere, atmospheric gases like oxygen can equilibrate with the lake providing suitable oxygen conditions for aquatic biota. The epilimnion typically maintains high dissolved oxygen (DO) concentrations while the hypolimnion relies on mixing with the metalimnion and epilimnion for DO. Because of this, the hypolimnion can experience periods of low DO (typically designated at <2 mg O_2 L^{-1}) or none, which are referred to as hypoxia and anoxia, respectively.

Hypoxia and anoxia become important for biota which rely on oxygen for respiration, but also for chemical reactions and internal loading of nutrients. When DO concentrations are low, redox reactions may facilitate at the sediment-water interface which allow for the internal loading of some nutrients. High DO concentrations typically indicate high redox potential, where oxygen will often act as a oxidizing agent and prevent many other redox related reactions. Under

hypoxic or anoxic conditions, a low redox potential could allow for other redox reactions at the sediment-water interface to proceed. Some of these redox sensitive reactions are environmentally important, such as for storage and release of phosphorus and iron (Søndergaard et al. 2003).

The occurrence of phytoplankton blooms in shallow polymictic lakes poses interesting questions with respect to the sources for the nutrients which blooms require. Blooms can occur quite frequently in lakes of high trophic levels and even occasionally in lakes of lower trophic levels (Reinl et al. 2021). In many thermally stratified lakes, the development and presence of a hypoxic or anoxic hypolimnion allows for a decrease in redox potential resulting in internal loading of key nutrients for phytoplankton like soluble iron(II) and phosphorus (P) species from sediments. For shallow polymictic lakes, thermal stratification occurs very briefly, preventing the development of a hypoxic or anoxic hypolimnion which suggests that internal loading of P and iron(II) due to redox in these lakes must be limited, if at all. As redox is the main method for internal loading in many lakes (Einsele 1936; Mortimer 1942), the lack of internal loading through redox then leads to the question of where exactly the redox sensitive chemical species are acquired by phytoplankton in lakes with limited external nutrient sources. One possibility is that anoxia arises at the sediment-water interface during a brief period of thermal stratification which can support the redox conditions required for P and iron(II) release for phytoplankton that descend the water column. Despite some phytoplankton possessing a means to acquire some redox sensitive species, like cyanobacteria and iron(II) sequestration of iron(III) via siderophores, the amounts of iron(II) which are required for blooms to develop in shallow polymictic lakes indicates that siderophores are likely not the main source (Murphy et al. 1976). Internal P loading

is known to occur in some polymictic lakes which may suggest that the sediment-water interface could become anoxic for brief periods (Jensen and Andersen 1992; Ramm and Scheps 1997; Loewen et al. 2007; Bryant et al. 2010; Orihel et al. 2017). Alternatively, phosphorus has been observed in recent years to arise from internal loading from processes not related to redox, such as desorption of organic or inorganic P bound to minerals, dissolution of minerals containing phosphate, and hydrolysis or mineralization of organic matter, to name a few (Katsev et al. 2006; Joshi et al. 2015; Li et al. 2015). Compared to P, internal iron (II) loading in polymictic lakes has not been researched in as much depth but has been observed and discussed about in some previous studies (Andersen and Ring 1999; Molot et al. 2014). In addition to the possible occurrences and potential effects for episodic hypoxia and anoxia, transient stratification in polymictic lakes can influence other important phenomenon such as phytoplanktonic vertical distribution (Santos et al. 2015) and cyanobacterial biomass (Wagner and Adrian 2009).

Since shallow lakes (maximum depth <5 m) are the most abundant lake type on Earth (Downing et al. 2006; Meerhoff and Jeppesen 2009), a point could be made that their research could lead to better-informed decision making which in turn could benefit many ecosystems. Previous studies on thermal stratification and the physical and biochemical processes connected to it are numerous and relatively in-depth for deep lakes (Michalski and Lemmin 1995; Crawford and Collier 1997; Ficker et al. 2017; Liu et al. 2019), while studies of thermal stratification for shallow lakes are typically shorter-term and less conclusive in comparison despite their importance for a variety of ecosystem processes and properties, including internal nutrient generation and algal blooms (Tuan et al. 2009; Zhao et al. 2012; Yang et al. 2018). A better understanding of transient thermal stratification is increasingly important given the recent

observed trend of changes in climate including increased air temperatures, longer ice free seasons, and reduced wind speeds (Schneider and Hook 2010; Schmid et al. 2014), especially for shallow lakes which may be more prone to water temperature changes (Maberly et al. 2020; Woolway et al. 2020).

The objectives of this chapter, Objective 1 and Objective 2, are to assess the frequency and durations of thermal stratification events in a shallow polymictic boreal forest lake in Northwestern Ontario (Lake 303), and to identify which environmental variables best predict transient thermal stratification. To identify the relative importance of each of the environmental variables on thermal stratification, a classification tree, a bagged model, and a random forest model are used to identify a correlations.

2.3 Methods

2.3.1 Study Site

Lake 303 is a shallow polymictic lake found within the boreal shield landscape of the International Institute for Sustainable Development Experimental Lakes Area (IISD-ELA) in northwestern Ontario, Canada. Lake 303 is a headwater lake with a surface area of 102,012 m^2, a terrestrial catchment area of 442,000 m^2, and mean and max depths of 1.86 m and 2.79 m, respectively. The watershed is made up of mostly jack pine (*Pinus banksiana*) and black spruce (*Picea mariana*) atop thin acidic soils. The lake is found within a depression in granite bedrock. Lake 303 has been a part of three different fertilization experiments in the past. For the first experiment, the lake was fertilized with P and nitrogen (N) during the summers of 1975 and 1976

to evaluate lake's recovery after eutrophication (Levine and Schindler 1989). During the second experiment in 1984, the lake was fertilized with P and N as part of a study looking at the effectiveness of alternative methods to harvesting baitfish (Mohr 1985, Mohr 1986). The third experiment began in 2019 which only featured additions of P to observe the influence of high P loading on cyanobacterial blooms without anthropogenic N sources (Molot et al. 2021). This 2019 experiment fertilized Lake 303 weekly with 2 kg (13.3 $\mu g\ L^{-1}$) of P in the form of food grade phosphoric acid (Univar) from June 6 to October 17. A phytoplankton bloom began around August 13 as a result of fertilization. The peak of the bloom occurred on September 9. Cyanobacterial blooms were not observed in unfertilized lakes at IISD-ELA during 2019. In addition to fertilization experiments, an experiment in 2011 and 2012 involved adding rainbow trout (*Oncorhynchus mykiss*) to the lake to understand growth dynamics in wild populations.

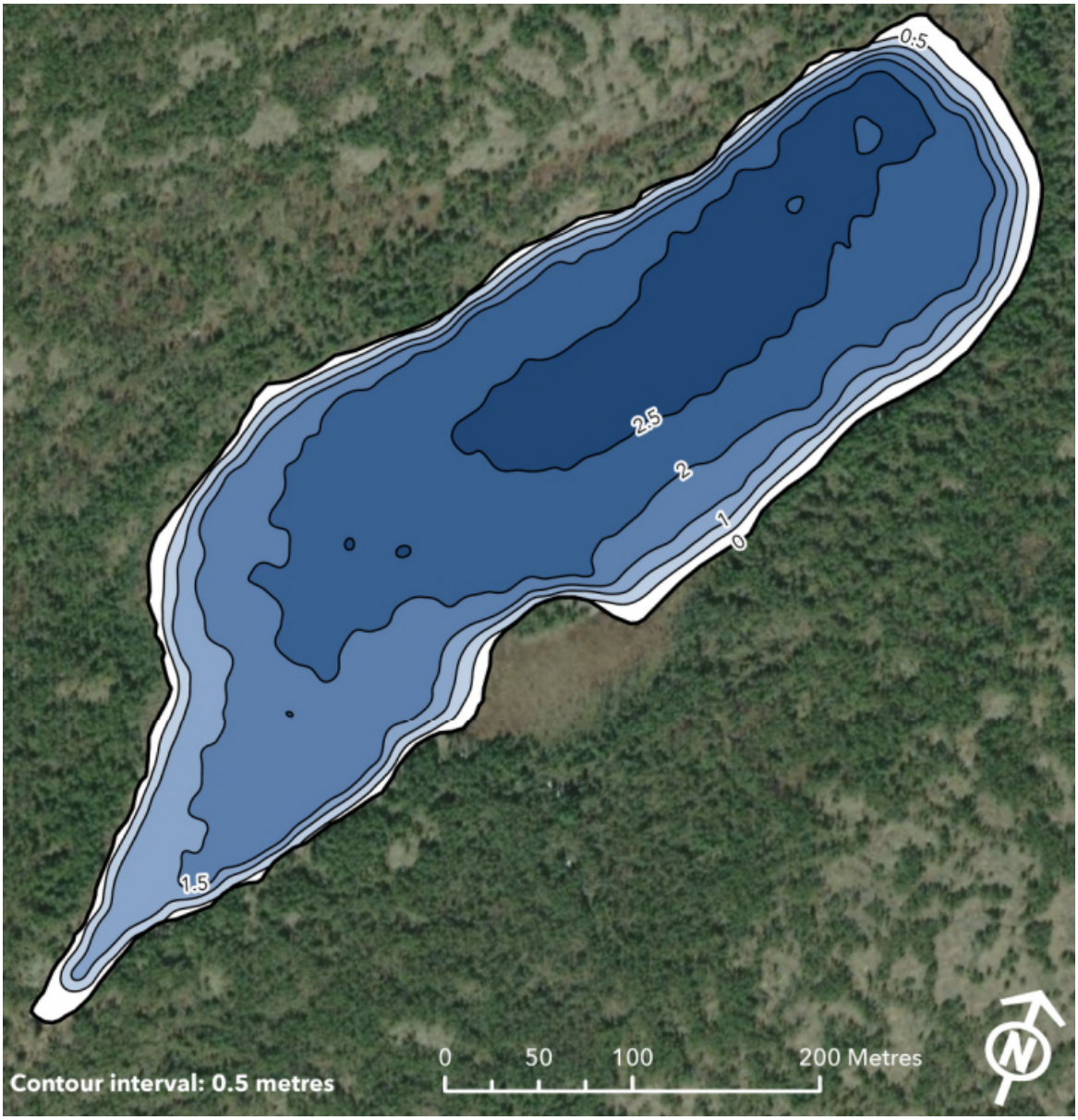

Figure 2.1. Lake 303 bathymetry. This figure was obtained and edited from IISD Experimental Lakes Area 2022.

2.3.2 Data Sources

The environmental variables used in this study, being air temperature, wind speed, wind direction, shortwave radiation, and precipitation were selected due to their known influences on stratification, along with availability of their recorded data for this site. These variables are also quite commonly available among most lake sites, allowing for similar future research to be done

for other lakes. Information on the bathymetry of Lake 303 were obtained from an online repository from IISD Experimental Lakes Area (Figure 1). HOBO U26 loggers collected DO and water temperature data at 1 m and 1.5 m. DO was collected at a resolution of 0.02 mg L^{-1} and an accuracy of ± 0.02 mg L^{-1} when below 8.00 mg L^{-1} or ± 0.05 mg L^{-1} when between 8.00-20.00 mg L^{-1}. The water temperature data at 2 m was collected by a HOBO TidbiT MX2203 temperature logger at a resolution of 0.01 °C and an accuracy of ± 0.2 °C. Both the water temperature and DO data were measured from May 20, 2019 to October 10, 2019, but due to an equipment error, data after October 6 were omitted. Data were collected every 15 minutes by the HOBO TidbiT MX2203 temperature logger and every 10 minutes by the HOBO U26 DO loggers. In order to synchronize water temperature data, only the measurements at each half hour were used.

Wind speed, wind direction, air temperature, shortwave radiation, and precipitation, were obtained from the Rawson Lake meteorological station (Station ID: 30455, Climate ID: 6036904) at IISD-ELA, located approximately 2 km from Lake 303. Each of these parameters were collected hourly. Wind speed was collected at 10 m above ground level using a RM Young anemometer every 5 s averaged for each hour to the nearest 0.1 km h^{-1} at an accuracy of ± 0.7 km h^{-1}. Wind direction was also collected by the same RM Young anemometer at a resolution of 0.1° every 5 s and averaged each hour. Air temperature was recorded with a Campbell Scientific CR3000 Micrologger built-in thermistor at a resolution of 0.01 °C and an accuracy of ± 0.3 °C. Precipitation was collected with a HyQuest Solutions TB-4 tipping bucket rain gauge at a resolution of 0.1 mm and an accuracy of ± 2%. Shortwave radiation was calculated from photosynthetically active radiation (PAR), which was measured by a Licor LI-190R quantum

sensor with an accuracy of ± 1% at a resolution of 5-10 μA per 1000 μmol s^{-1} m^{-2} and recorded by a Campbell Scientific CR1000 data logger every 5 s and averaged every 15 minutes.

2.3.3 Classifying Transient Stratification

Hutchinson (1957) defined stratification as a difference of $\geq$ 1 °C per meter depth within the water column. This simple yet robust definition allows for a clearer identification of stratification as a binary variable which was the main consideration in this process versus adopting alternative definitions such as Schmidt stability. Due to the shallow maximum depth of Lake 303, thermal stratification was interpreted as $\geq$ 0.5 °C per 0.5 m depth.

2.3.4 Statistical Analysis

All analyses were performed using R Statistical Software (v4.2.1; R Core Team 2022). The meteorologic variables were heavily skewed which required a means to better analyze each variable. To understand and identify wind speeds which may be critical to mixing, speeds were binned into quantiles based on the probability values of 0.5, 0.67, and 0.95 were used to represent half of the dataset, first standard deviation, and second standard deviation, respectively. To examine relationships between the time lengths of the stratified and mixed periods, their medians and median average deviations (MADs) were used to represent a similar statistic to the mean and standard deviation of the mean.

Classification trees were used to assess potential relationships between the environmental drivers and transient stratification. Classification trees, or classification and regression trees (CARTs) can quantify and predict the relative importance of predictor variables to an outcome variable. The outcome variable here was thermal stratification and the predictor variables were air temperature, wind speed, wind direction, shortwave radiation, and precipitation. The classification tree was produced in R using the rpart package (v4.1.19; Therneau and Atkinson 2022). The dataset was randomly partitioned to separate datapoints for the training and testing of the model. The two partitions were sized according to typical proportions for CART modelling, 70% and 30% for training and testing, respectively. The training partition is used first to develop the classification tree. The classification tree begins with all datapoints grouped within a single bin, or node, and are subsequently split using predictor variables and values until they are finally grouped into terminal nodes. At each node the datapoints may be split further into a new row with two new nodes, where the number of rows within a tree is known as the tree depth. Terminal nodes are achieved when the number of datapoints within a node are too few, referred to as the minimum node split. The terminal nodes are then classified by their datapoints based on the most popular value for the outcome variable. As terminal nodes often contain different values of the outcome variable, the terminal node's classification may incorrectly represent some of the data, known as misclassification error. Misclassification error is often the means which accuracy is assessed for trees. Next, the classification tree, created from the training partition, is tested using the remaining partition to assess the tree's accuracy on data which it was not trained on. Classification trees can further be improved for greater accuracy or interpretability by adjusting features and settings in a process

known as pruning. Pruning was done to achieve a tree with the greatest % accuracy through trial-and-error of comparing combinations of varying values for minimum node split, which is the minimum number of data points required within a node in order to split into a new row, and maximum tree depth, which is the maximum number of rows. The best % accuracy for the tree was found to result from a minimum node split of 81 and a maximum tree depth of 8.

Bootstrap aggregating (bagging), is another popular type of modelling often used in conjunction with CARTs. Bagging involves averaging an ensemble of unpruned classification trees from bootstrap samples of the dataset which can improve the accuracy from a single tree (Breiman 1996). Each tree within a bagged model is unpruned, allowing it to have many splits and great depth. As a result, the trees are less likely to overfit the dataset because the variability of each tree's structure increases, also reducing its bias. The unpruned trees are evaluated for their misclassification rates which are then averaged together often resulting in a greater accuracy than a single optimized tree. Bagging was done in R using the caret (v6.0.93; Kuhn 2022) and ipred packages (v0.9.13; Peters and Hothorn 2022). Several models were tested with varying numbers of trees to assess the impact of the number of trees to provide the lowest misclassification error. When increasing the number of trees within models, misclassification error approaches a general minimum range while the model's computational complexity continues to increase. The ipred package found that 48 trees achieved a near minimum misclassification error while still efficiently keeping the computational complexity low. While bagging does well to reduce variability in a model, other approaches such as random forests also utilize CARTs to produce large ensembles of varying CARTs to reduce model variability from both single optimized trees and bagged models.

Finally, I used a random forest modelling approach to evaluate drivers of transient thermal stratification. Random forest is a type of ensemble machine learning model which uses many randomized CARTs to reduce model variability. While both bagged and random forest models use unpruned trees, bagged model trees split based on all predictor variables while random forest trees randomly select a subset of only a few predictor variables. The random forest then provides overall results based on a majority vote among all trees within the model. This process helps to reduce overfitting by lowering bias and variability as well as improving the model generalization to new data. The random forest model used here was produced in R using the randomForest package (v4.7.1.1; Liaw and Wiener 2002). As with bagging, the size of random forest models can become computationally complex, so using the randomForest package, a forest size of 2000 was found to provide a low variability while reducing computational complexity.

The main benefit to both bagging and random forests over single classification trees is the increase in model accuracy to fitting the dataset. The major drawback to these two methods is their interpretability. CARTs provide a figure which can be easily interpreted on its own, while due to their complexity, bagged and random forest models only provide a ranking of the predictor variables based on importance within the models along with their accuracy. One way to further analyze bagged and random forest models is to use the Gini coefficient, which is a measure of inequality (Menze et al. 2009). To do this, the mean decrease of the Gini coefficient (MDG), is calculated for each predictor variable within the models. MDG quantifies the contribution of each predictor variable to the homogeneity of the trees within the models, where higher values indicate a higher importance to the outcome variable. For both the bagged and random forest

models the MDG was calculated to identify which variables had the strongest relationships with transient stratification.

In order to create a better understanding of observed hypoxic event, the R package LakeMetabolizer, was used to calculate the metabolism rates within the lake using the recorded oxygen data. The package offers the use of several metabolism models, of which the maximum likelihood (MLE) model was selected. The MLE model was used to calculate respiration (R), gross primary production (GPP), and net ecosystem production (NEP) at 1.5 m depth, the deepest depth of oxygen data in the lake. While the DO data from 1.5 m depth does not reflect true deep water respiration rates, the mean depth in Lake 303s is 1.86 m which can still allow for a useful assessment of the lakes metabolism.

2.4 Results and Discussion

2.4.1 Identifying Transient Stratification

Transient thermal stratification was a common occurrence within the lake during the study period. Transient stratification was observed 146 times between May 21 and October 6th, 2019 (Figure 2.2), ranging in duration from 1 to 106 hours. Overall, transient stratification occurred during 1102 of the 3316 hours (~33%) of the study period (Figure 2.2). Mixing periods lasted from 1 to 136 hours during the study period. Stratified period lengths had a median of 2 hours and a median absolute deviation of 3 hours while mixed period lengths had a median of 3 hours and a median absolute deviation of 4 hours. Transient stratification was observed to initiate at any time, but most likely during the morning and afternoon. Transient stratification often broke down during the night.

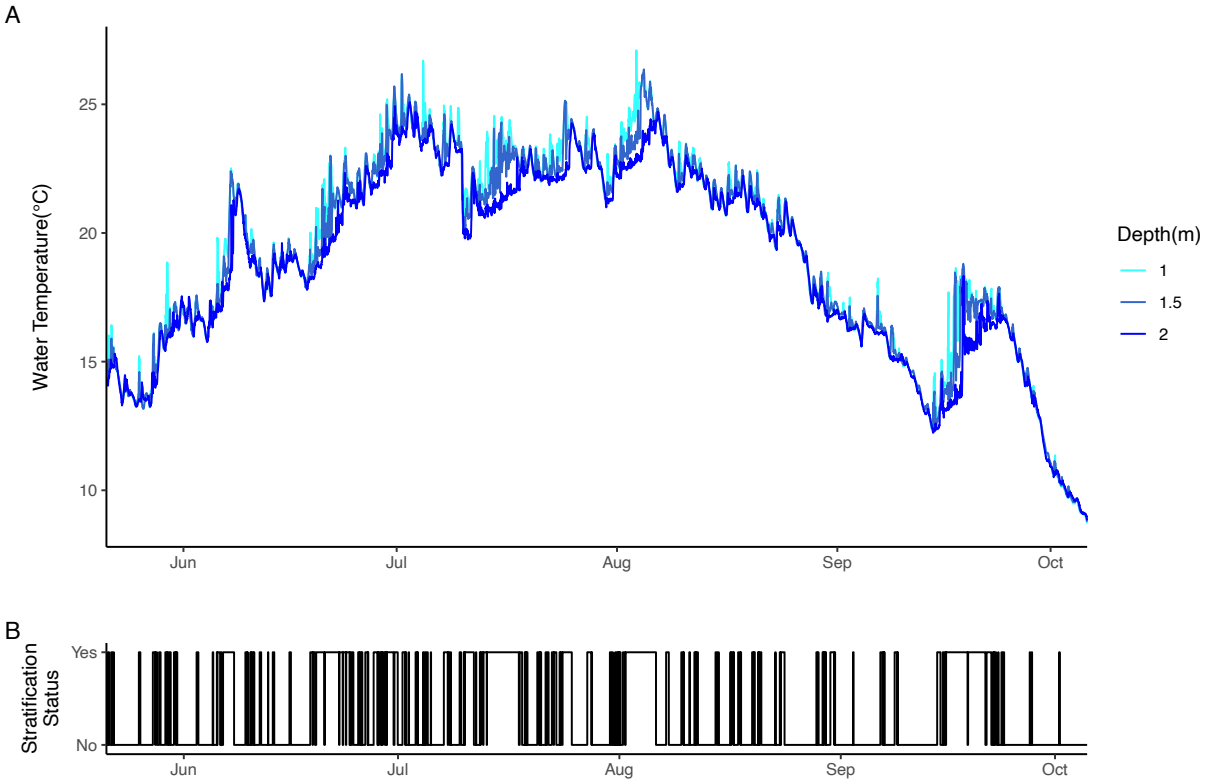

Figure 2.2. a) Water temperature measurements at 1m, 1.5m and 2m in Lake 303 from May 21 to October 6, 2019. b) Stratification status of Lake 303 based on the water temperature profile. The stratification status is indicated as either "Yes" or "No" meaning the water column was stratified or not, respectively.

Water temperatures were warmest on average during July, and the warmest water temperatures were recorded at 1 m on July 4th (26.7 °C) and August 3rd (27.1 °C) (Figure 2.2a). The highest frequency of transient stratification events occurred during July, with 42 events totaling 326 hours (Figure 2.2b). After the peak in water temperatures on August 3rd, the waters consistently cooled and stratification became less frequent until September 13th, when a short warming trend began that lasted until September 26th.

2.4.2 Potential Drivers for Transient Stratification

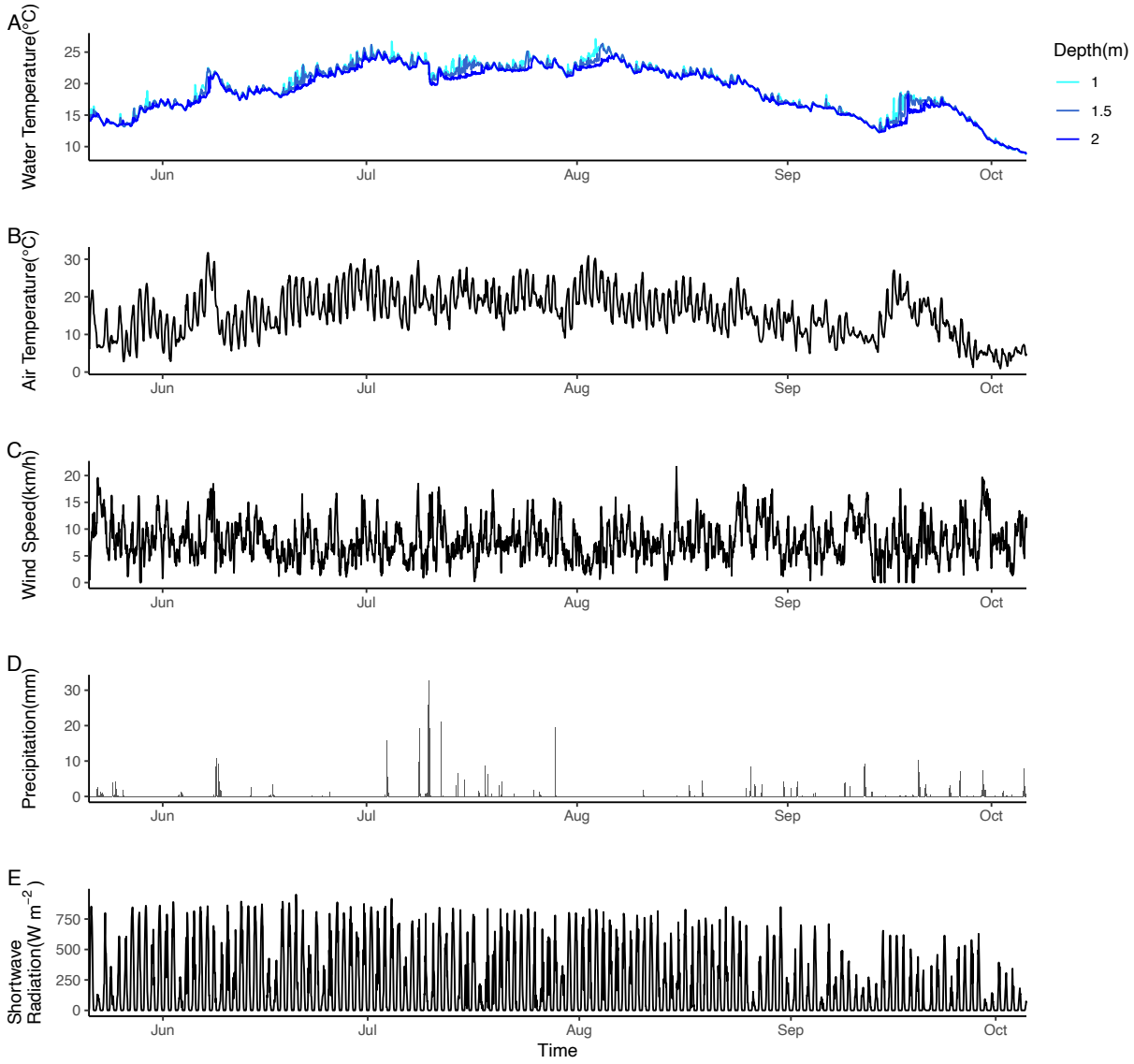

Figure 2.3. Temporal variation of a) Lake 303 water temperature b) air temperature, c) wind speed, d) precipitation, and e) shortwave radiation.

Air temperature and water temperature displayed similar seasonality, with an overall warming from May to early August and then a declining of temperatures thereafter (Figure 2.3a,b). Both air and water temperatures increased in early September before dropping further. Since most of the influence from external factors on lakes occurs at the surface, air temperature

is known to be strongly connected to water temperature. The influence of air temperature can be seen from May to early August through increasing frequency of stratification (Figure 2.2) as air temperatures rose and maintained warmer upper water temperatures of around 20.0 °C allowing for thermal stratification (Figure 2.3a). Yet when air temperatures cooled in early August, the rate of heat loss from the surface waters increased which led to a more isothermal water column which mostly continued for the remainder of the month.

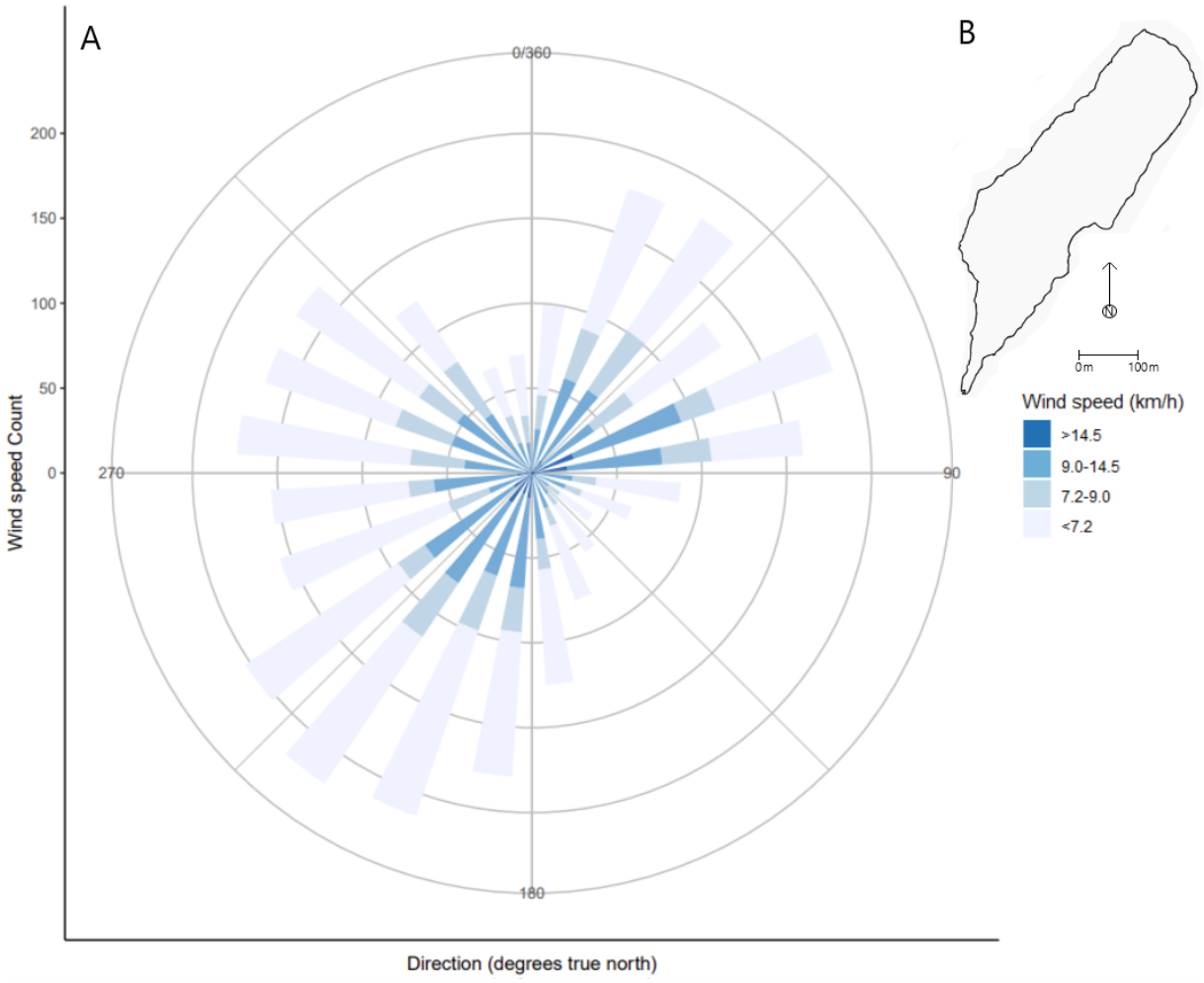

Figure 2.4. a) Polar plot of wind data including direction in degrees true north and speed binned as > 14.5 km h^{-1}, between and equal to 14.5 km h^{-1} and 9.0 km h^{-1}, between and equal to 9.0 km h^{-1} and 7.2 km h^{-1}, and less than 7.2 km h^{-1}, where 14.5 km h^{-1}, 9.0 km h^{-1}, and 7.2 km h^{-1} represent the quantile probabilities of 0.95, 0.67, and 0.5, respectively. Wind direction data was binned into 15-degree groups and centered on the midpoint of each bin. b) The shape of the surface of Lake 303 as compared to the polar plot's orientation.

The median wind speed during the study period was 7.2 ± 0.7 km h^{-1}, and wind speeds as great as 21.7 ± 0.7 km h^{-1} were measured. To understand the relative significance of wind speed due to its non-normal distribution, quantiles were calculated at probabilities of 0.5, 0.67, and 0.95, resulting in values of 7.2, 9.0, and 14.5 km h^{-1}, respectively (Figure 2.4a). Approximately 54.1% of wind directions were recorded between 30°-90° and 180°-255°, aligning closely with the fetch of Lake 303 (Figure 2.4b).

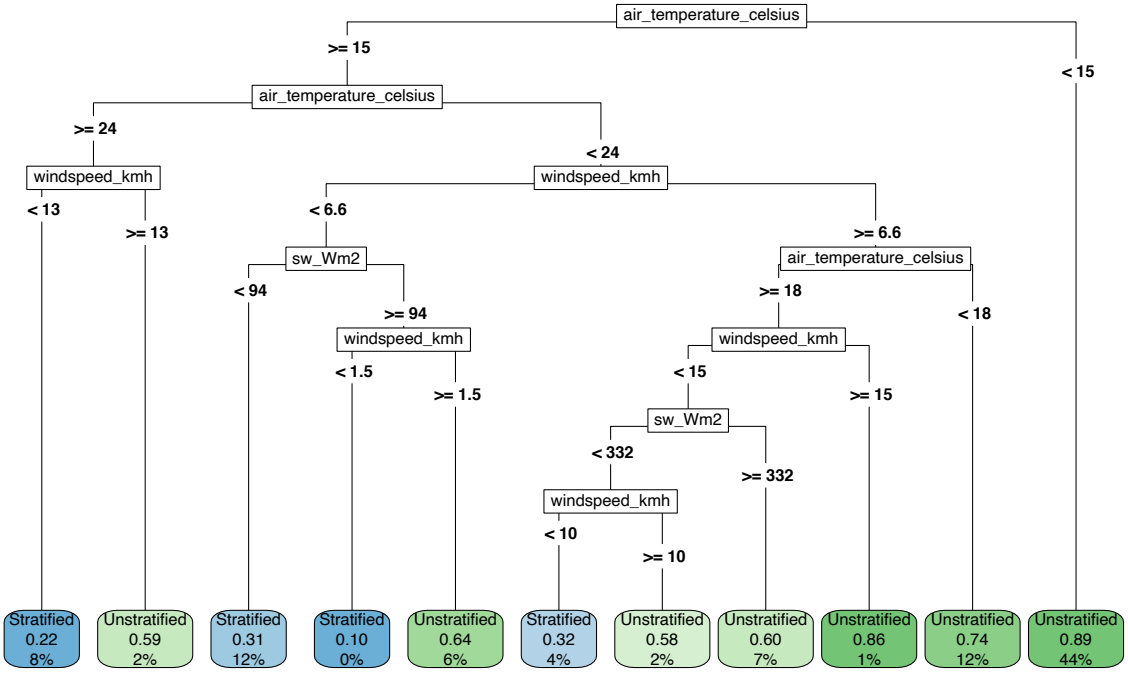

Figure 2.5. A classification tree binning the stratification status to air temperature, wind speed, and shortwave radiation (sw_Wm2) at each hour in the 2019 dataset from May 21 to October 6. Terminal nodes display the stratification status of the majority the data points within the node. Below the node class is a value for the class proportion and a value of the percentage of the entire dataset within the node. The class proportion is calculated by assigning values of 0 and 1 to the classes “Stratified” and “Unstratified”, respectively, and identifying the average of the assigned values within the node.

The classification tree (Figure 2.5) model selected air temperature to be the most effective variable in predicting transient stratification, followed by wind speed and then shortwave radiation. Precipitation and wind direction were also included in the model's calculations but their influences on stratification were not found to be as predictive. The model was found to be 78.0% accurate on the training dataset and 75.3% on the testing data. At the first model node split air temperature was used to split the data based on being above and equal to or below 15 °C. Air temperature was below 15 °C for 44% of the dataset. At this terminal node, the node class was designated as "Unstratified" with a node proportion of 0.89. Node proportion refers to the proportion of data points in a node that are classified as "Unstratified", where a node proportion value of 1 would mean all data points in the node were associated with the lake being unstratified/mixed and vice versa for a value of 0. Therefore, a node proportion of 0.89 indicates that the lake was unstratified for 89% of the data points within the node and that 11% of the data points were misclassified. The next split in the tree also used air temperature, now splitting data points based on being above and equal to or below 24 °C. The next two splits used wind speed as the predictor, where wind speeds above and equal to or below 13 km h^{-1} and 6.6 km h^{-1} when air temperatures were $\geq$ 24 °C or < 24 °C, respectively. The general pattern suggests that stratification is connected with higher air temperatures and lower wind speeds.

In pursuit of lower variability, a bagged tree model was created. The variability from bagged tree models is mostly assessed in R using a trial-and-error approach, where many bagged models are produced with differing tree numbers to understand the changes they cause in variability. Variability stabilized at 48 trees with an accuracy of 77.0% against the test data, an improvement of 1.7% points from the single tree model. The bagged model found air

temperature to be the best predictor for stratification; however, the next strongest predictors were wind direction and windspeed which differed from the single tree model (Figure 2.6a). The relative importance of each variable in the bagged model was determined using the MDG.

Next, a random forest was produced to further reduce variability. Similarly to the bagged model, the number of trees to be used in the random forest was tested to understand the effects on variability; however, as the random forests are much larger, computational efficiency was also important to consider. At 2000 trees, the random forest model's accuracy on the test data was calculated at 79.0%, an improvement of 3.7% points from the single tree model. The random forest model also found that air temperature was the greatest predictor of stratification, followed by wind direction and windspeed (Figure 2.6b). The relative importance of each variable in the random forest model was also determined using the MDG.

The importance factor of the variables was assessed by quantifying each variable's contribution to the homogeneity of the nodes, known as mean decrease gini (MDG), where larger values of MDG indicate stronger importance. With the bagged model (Figure 2.6a), air temperature had the greatest MDG value, 716.93, which was 21.7% greater than the wind direction MDG, the closest MDG value. The wind speed MDG plotted just below wind direction followed by shortwave radiation. Precipitation trails each of the variables, showing very little influence on thermal stratification. MDGs of the random forest followed a similar trend (Figure 2.6b); however, air temperature appeared as an even greater predictor for stratification proportionally in comparison with the other variables.

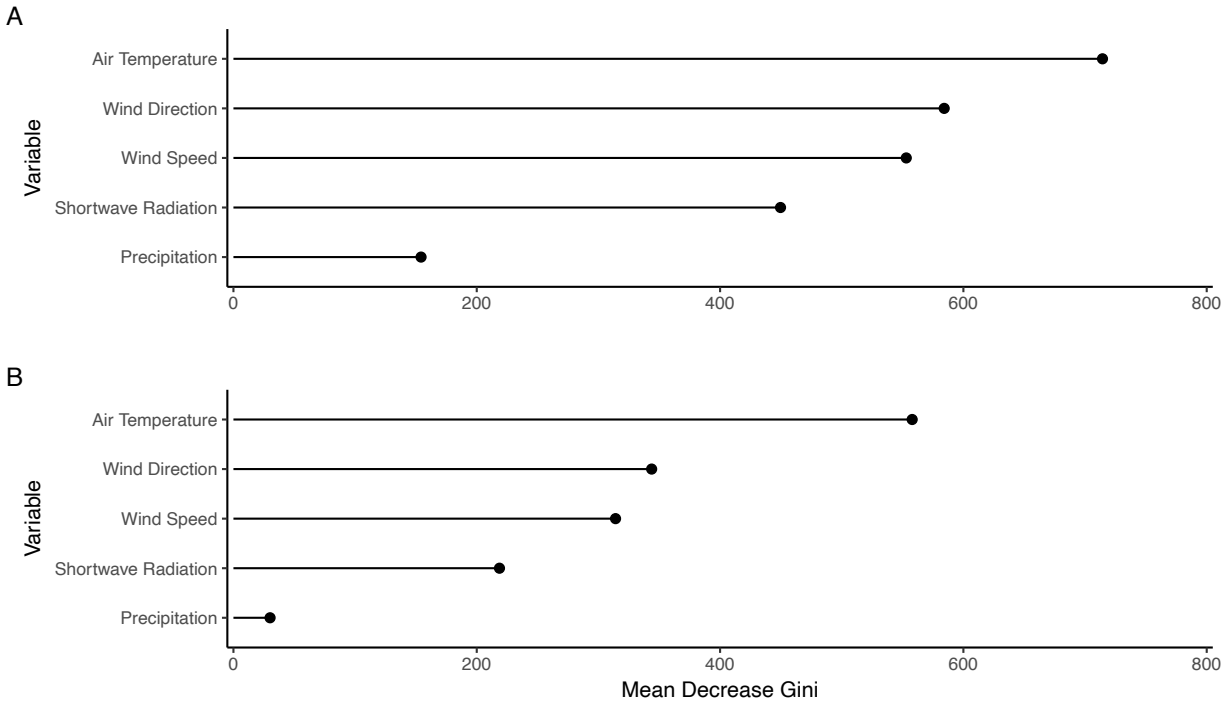

Figure 2.6. The variable importance for a) the bagged model and b) the random forest model on stratification in Lake 303 based on the mean decrease of the Gini coefficient for each variable. The two models looked at the importance of air temperature, wind direction, wind speed, shortwave radiation, and precipitation. A larger value in the mean decrease of the Gini coefficient indicates a higher importance of the variable for the models.

All three model types found air temperature to be the greatest predictor of stratification in Lake 303 over the study period. As found in the single classification tree, air temperature alone could be used to classify just under half of the data as not stratified based on a single initial split. Many studies that look at air temperature with lake thermal stratification investigate the effects of rising global air temperatures and climate change (Coats et al. 2006; Austin and Colman 2007; Schneider and Hook 2010; O'Reilly et al. 2015). Schindler et al. 1996 reported increases of 1.6 °C for both air temperature and water temperature at IISD-ELA from 1970 to 1990. Small lakes have

a high surface area to volume ratio which increases the influence of air temperature on thermal stratification more for shallow lakes than deeper lakes.

Wind speed was the second most important variable in the classification tree. The bagged and random forest models; however, found wind speed to be the third most important behind wind direction. It should be noted that wind direction and wind speed were collected at 10m above ground surface at a nearby meteorological site which could be different from the wind conditions experienced at the surface of Lake 303. Despite wind speed being commonly known as a major influence for lake mixing, surface waves in small lakes tend to be positively correlated with the lake fetch (Gorham and Boyce 1989) which can explain the importance of wind direction over speed within the two models. Surface wave height is proportional to distance, therefore, winds which propel waves along the fetch of Lake 303s oblong shape will result in larger waves, having a greater impact on mixing (Wetzel 2001). It is also possible that priority of wind direction over speed within the models suggests that the typical wind speeds that Lake 303 experiences are sufficient considering the direction of the winds.

Trends of increasing stratification duration due to decreasing wind speeds have been discovered in other parts of the world (Woolway et al. 2017). O'Reilly et al. (2003) observed reduced mixing due to decreasing wind velocities which lead to decreased nutrient upwelling to surface waters from the hypolimnion in Lake Tanganyika, Africa. Zhao et al. (2011) found that for the large shallow Lake Taihu in China, stratification was more likely when windspeeds were < 14.4 km h^{-1}.

None of the models found a significant influence from precipitation on the thermal stratification status of the lake. Research which connects precipitation to stratification is relatively scarce compared to the other parameters being explored here (Wantzen et al. 2008). Of the publications which report destratification due to precipitation, the focus is often on severe storms with a lot of rain (Znachor et al. 2008; Klug et al. 2012). The greatest daily precipitation recorded during the study period was 145.6 mm on July 9. Liu et al. (2020) found that daily rainfalls of 10 mm or greater could reduce water column stratification or alter the mixed layer depth and thermocline due to a decrease in temperature of surface waters or by increased inflow in a reservoir in eastern China.

Shortwave radiation is critical for thermal stratification via the transfer of thermal energy to the lake surface. In lakes with low clarity, or high light attenuation coefficients (K_d), light is mostly absorbed and stored in the epilimnion as thermal energy which results in increased water temperatures in the epilimnion (Mazumder and Taylor 1994; Riis and Sand-Jensen 1998; Heiskanen et al. 2015). In clear shallow lakes, light can penetrate to the sediment and warm both the water and sediment (Fang and Stefan 1998; Wetzel 2001). The transfer of energy from light to heat can therefore prevent thermal stratification if light penetration reaches the lake sediment. If phytoplankton populations proliferate and blooms develop; however, light penetration declines. Decreased light penetration allows for greater amounts of thermal energy storage in the surface waters, encouraging thermal stratification (Kumagai et al. 2000). The classification tree shows that thermal stratification arises more often from lower amounts of shortwave radiation (Figure 2.5). The two node splits using shortwave radiation both typically associate stratification with the lower values of the split, which may be due to light penetration reaching

the bottom of the lake to warm the deeper waters and sediment rather than the surface. The proportions for each of the terminal nodes following shortwave radiation splits appear to be more neutral compared to other terminal nodes in the tree, with the largest terminal node representing 12% of the dataset with a value of 0.69 indicating 69% correct classification of “Yes” data points within the node.

The collinearity of the variables to thermal stratification does bring with it a limitation regarding the lagged effects of the variables. The comparison here looks directly at the instantaneous values of each of the environmental variables compared to the stratification status of the lake, however, there are some important things to note which this method does not reflect. Firstly, the influences on the temperature of the lake waters from each of the environmental variables do not apply in the exact moment of the recorded values, that is, the heating, cooling, and mixing brought about by the environmental drivers are only recorded by the water temperature sensors in the moments afterward. The temperature sensors were at depths of 1, 1.5, and 2 m in Lake 303, therefore, the effects of air temperature, for example, would not be reflected by the water temperature data until the energy within the surface waters had transferred to the water at the sensor at 1m depth. This means that changes in thermal energy which are recorded by the sensors may be representing the influences of the environmental drivers from a previous timestep. This lag effect was not explored here due to the many complications associated with accounting for the cumulative effects caused by each of the environmental variables and their likely differing degrees of lag.

2.4.3 Transient Stratification and Hypolimnetic Dissolved Oxygen

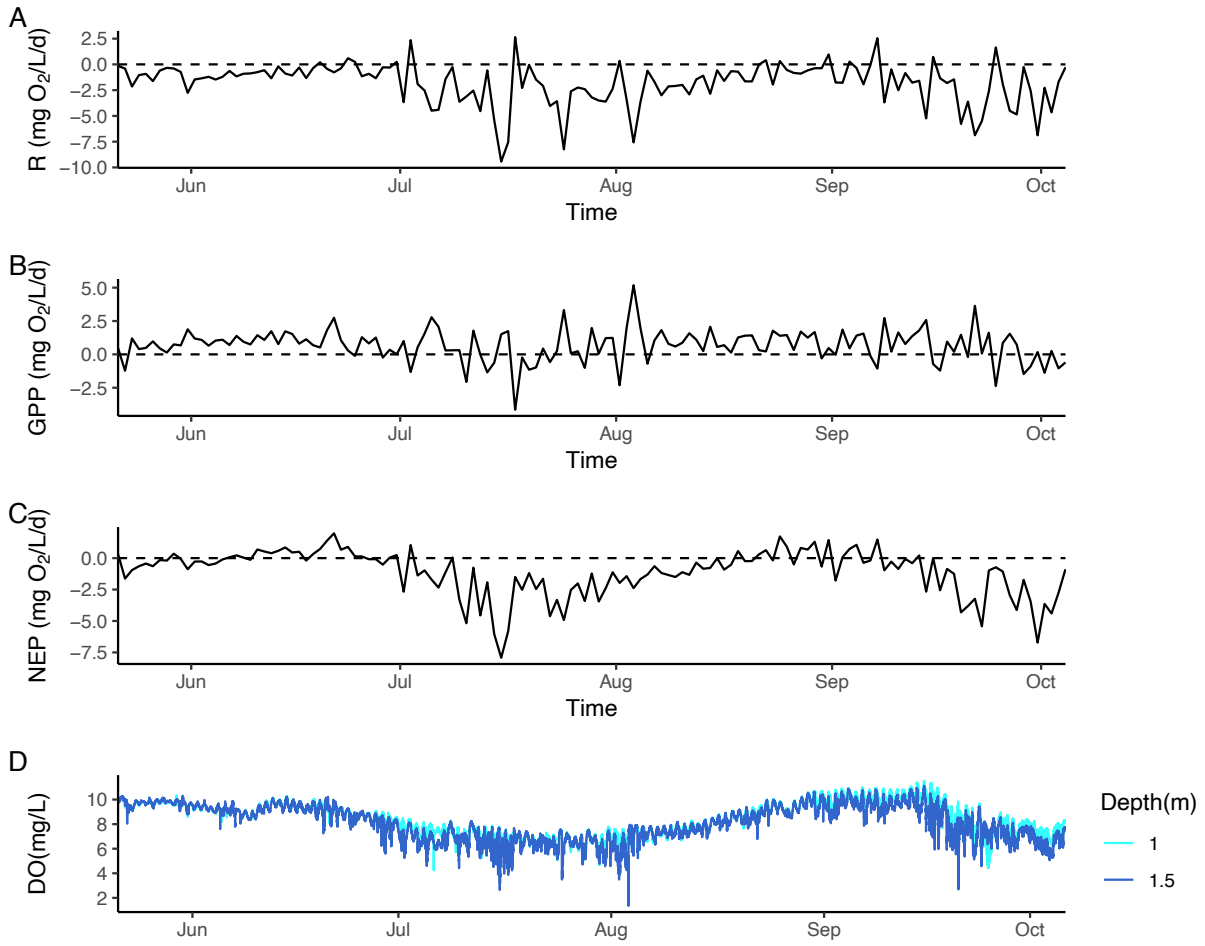

Figure 2.7. Full metabolism rates for Lake 303, with a) respiration, b) gross primary productivity, and c) net ecosystem productivity, as well as the d) DO dataset.

Throughout the study period, Lake 303 DO levels only dropped into hypoxic levels (<2 mg DO L^{-1}) on one occasion (Figure 2.7). While DO was not measured at 2m depth in Lake 303, measurements on August 3 at 1.5 m had dropped sharply to 1.35 mg L^{-1}. This decrease in DO was ephemeral, staying below 2.00 mg DO L^{-1} for only 20 minutes, before then rising back to the typical daily range near ~7.00 mg L^{-1} within an hour (Figure 2.8). The respiration rate at the same

depth on August 3 was calculated to be -7.57 mg O_2 L^{-1} d^{-1} (Figure 2.9). This drop in DO occurred just 10 days before the estimated start of the cyanobacterial bloom on August 13. Although there is limited data on the nutrients of the lake especially surrounding this hypoxic event, the timing here could suggest that the hypoxic conditions acted as a key trigger in the development of the bloom through the release of redox-sensitive nutrients from the sediment. This rapid and short-lived drop in DO at 1.5 m may also indicate an upwelling of hypoxic waters from deeper in the lake. It is possible that Lake 303 was actually hypoxic below 1.5 m (sensor depth) for an unknown period of time. Jabbari et al. (2019) captured two short-lived hypoxic events in the western basin of Lake Erie which were caused by upwelling hypolimnetic waters from the central basin of the lake.

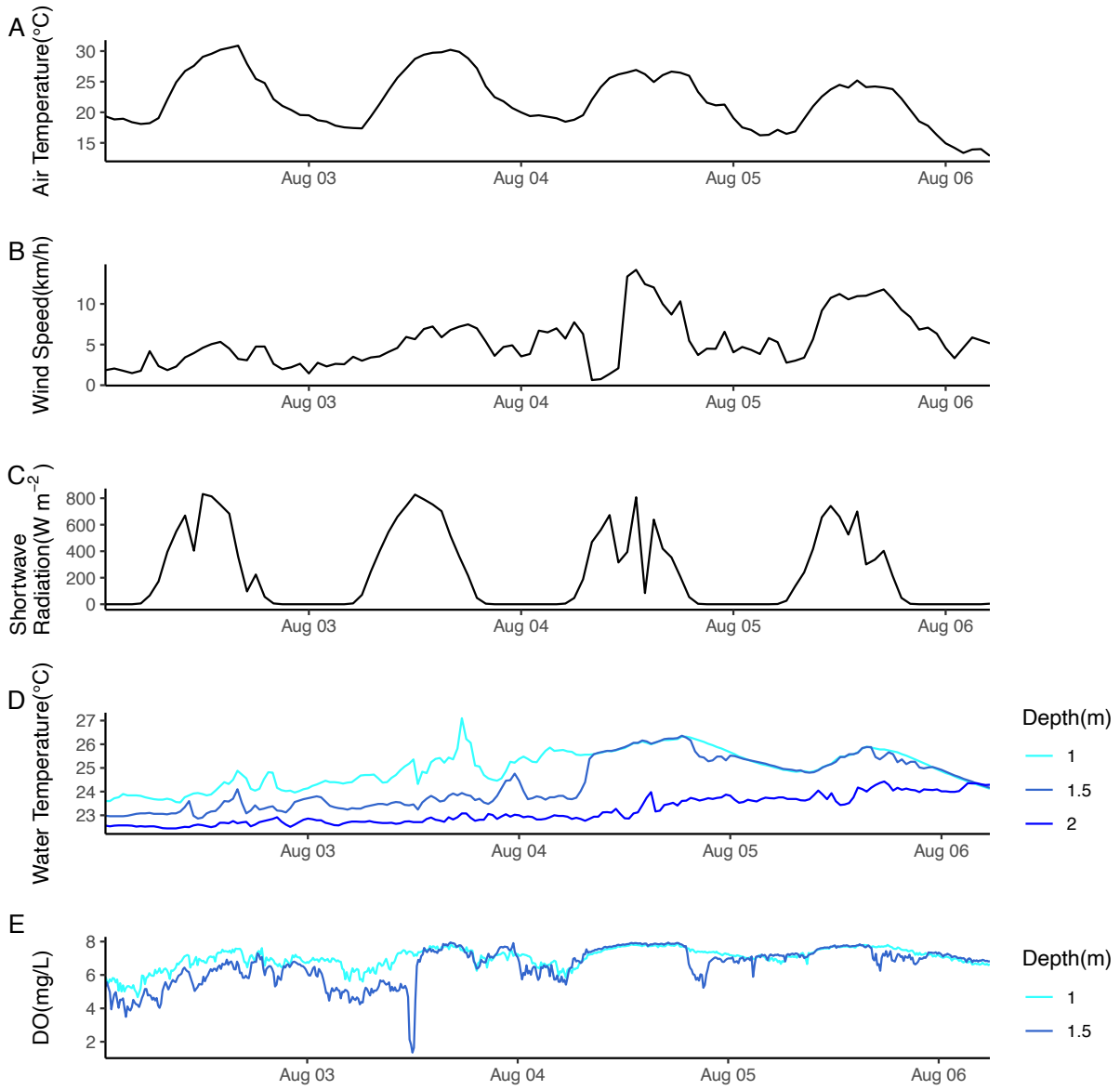

Figure 2.8. Lake 303 was stratified from August 1 at 21:00 until August 6 at 02:00. The data for a) air temperature, b) wind speed, c) shortwave radiation, d) water temperature, and e) DO are included over this stratified period. DO at 1.5m depth dropped to 1.35mg L^{-1}, the lowest value observed during the study.

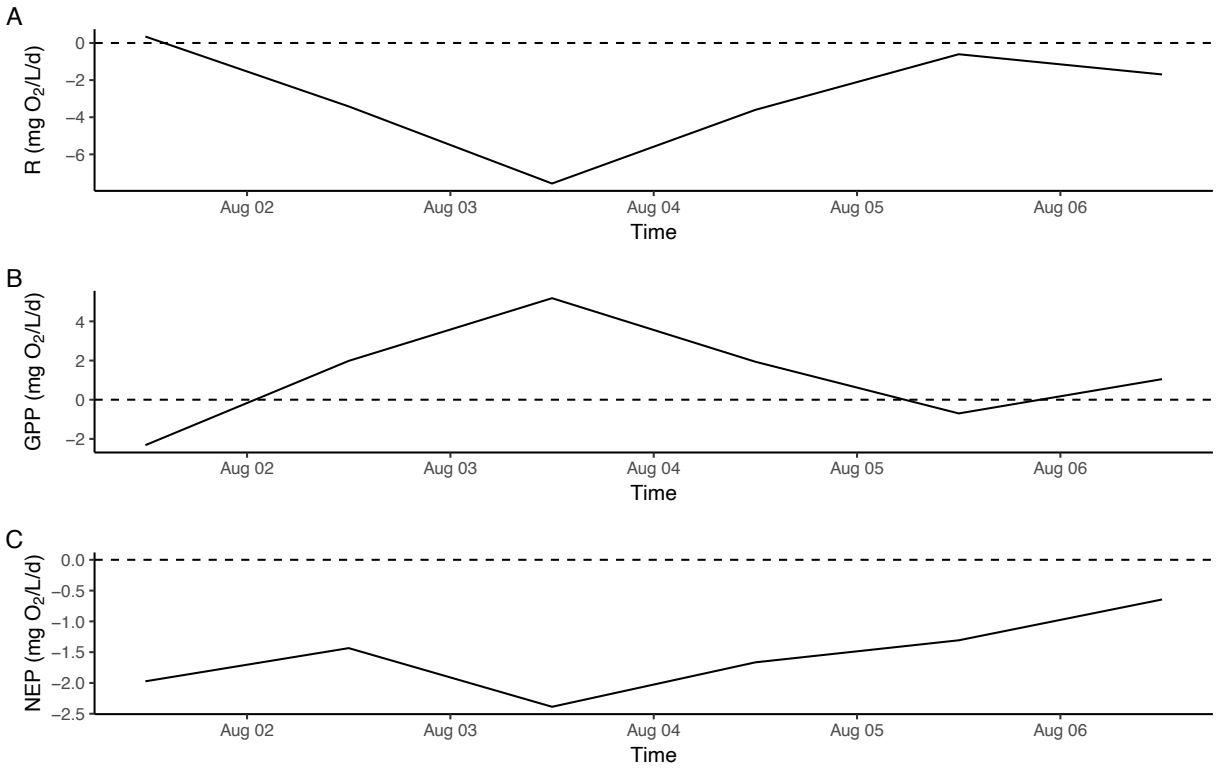

Figure 2.9. The a) respiration (R), b) gross primary productivity (GPP), and c) net ecosystem production (NEP) in Lake 303 around the time of the observed hypoxia on August 3. The dashed lines indicate 0 mg O2 L-1d-1.

Daily highs in air temperature during the hypoxic period were between 25.0-30.0 °C while lows were between 16.0-18.0 °C before dropping to approximately 13.0 °C on August 6. Increasing wind speeds in addition to this cooling of air temperatures toward the end of the event are likely the reason for destratification. Loewen et al. (2007) found that approximately 4.5 consecutive days of low wind speeds (<7 m s^{-1}) and warm air temperatures were able to induce hypoxia due to stable stratification in the shallow western basin of Lake Erie, although understanding key thresholds across ecosystems will of course depend on factors including,

morphometry, and sediment oxygen demand. Wind speed during the sampling period in 2019 for Lake 303, although much smaller than the western basin of Lake Erie, was always <7 m s^{-1}.

Apart from the biochemical implications for hypoxia, decreased oxygen levels are still potentially harmful to biota living within the lake. Rainbow trout, which have been introduced to Lake 303 on multiple occasions in the past, are known to prefer higher levels of DO ($\geq$ 7 mg DO L^{-1}) and have an estimated lower DO threshold concentration of approximately 4.2 mg DO L^{-1}. When levels fall below 4.2 mg DO L^{-1}, their survival rate beings to decline (Jiang et al. 2021). Rainbow trout were introduced into the lake first in 1969, however, none had survived, which warranted the re-introduction of the species in 2011 and 2012. DO in Lake 303 was observed to have dropped below 4.2 mg DO L^{-1} at 1.5 m or deeper on 8 occasions during the study period (Figure 2.7). Lowering levels of DO cause fish to swim into shallower waters which can lead to other issues such as predation rates or impaired health conditions (Kraus et al. 2023). It is possible that the lake's transient stratification could be responsible for inconsistent DO levels which may have been the fault of the failed initial introduction. To current knowledge, the status of rainbow trout within the lake is not known, but with the drops in DO which were observed indicate that rainbow trout may experience conditions of decreased survivability multiple times on average in the lake during the ice-free season.

2.5 Conclusion

Improving the current understanding of thermal stratification in shallow polymictic lakes can greatly assist ecosystem management to benefit prediction and prevention methods for issues such as phytoplankton blooms, anoxia, and internal nutrient loading. Lake 303 stratified 146 times and the duration of stratification ranged from as low as <1 hour to as long as >106 hours. Thermal stratification mostly occurred during June and July, aligning with warmer air temperatures. Seasonally, thermal stratification in Lake 303 was most strongly influenced by air temperature, followed by wind direction and speed. The analysis here, however, did not factor in the lagged effects of the environmental parameters. Instead, the analysis here looked purely at the state of stratification in the lake and the specific parameters at that exact moment which may not reflect the true influence of the parameters like the cumulative effect of heat which may be carried over from pervious timesteps resulting from an environmental variable.

Given evidence of increasing air temperature and water temperature at IISD-ELA (Schindler et al. 1996) and recent research suggesting that global air temperatures are rising and wind speeds are decreasing, thermal stratification could become longer and more frequent (Coats et al. 2006; Austin and Colman 2007; Schneider and Hook 2010; O'Reilly et al. 2015; Woolway et al. 2017; Maberly et al. 2020), seen through the importance highlighted here of air temperature and wind on Lake 303. Future research can look into the extent of anoxia in the hypolimnion and its pre-conditions for shallow polymictic lakes as well as the potential effects of global warming and decreased wind speeds on the duration and frequency of thermal stratification in shallow polymictic lakes. Changes in frequency and/or duration of transient

thermal stratification as a result of climate change scenarios could also present an opportunity to research how these might affect phytoplankton blooms.

Chapter 3 Using 1-D Models to Reproduce the Observed Thermal Stratification in a Shallow Polymictic Lake and a Shallow Dimictic Lake

3.1 Abstract

Given that shallow lakes are the most abundant type in the world, the importance to effectively model them is clear; however, very few studies look at the success of models on these lakes. Five models were evaluated on their performance in reproducing the observed thermal stratifications for a shallow polymictic lake and an adjacent shallow dimictic lake. Lake 303 is a shallow polymictic lake which was observed to stratify 146 times during the spring and summer of 2019 but even the model which reproduced the most stratification only reproduced 55 events. Some of the models failed to calculate any thermal stratification at all in Lake 303. Lake 304 is a shallow dimitic lake which was stratified for almost the entire spring and summer of 2019 and many of the models were able to reproduce some or most of stratification; however, the models typically calculated a much warmer lake bottom, even by as much as 10.6 °C warmer than observed. With regards to the overall temperature profiles of the two lakes the models were generally much better at reproducing the surface mixed layer temperatures than they were for the deeper layers of the lakes. The findings here suggest that shallow polymictic lakes require either highly specific models or improved parameterization for the dynamic and temporal thermal stratification patterns of these lakes.

3.2 Introduction

Modelling has grown to become an essential tool in the field of limnology, allowing for the transformation of data into terms which can unite scientists, managers, policy makers, and stakeholders to bring about important actions required for our environments (Gal et al. 2009; Schmolke et al. 2010; Trolle et al. 2012). Aquatic ecosystem models are often used to predict oxygen conditions, phytoplankton, zooplankton and fish communities, and temperature profiles (Hall and Day 1977). Models can accurately estimate lake temperature profiles using various equations and assumptions for mechanisms like mass-balancing and mechanical mixing which are especially effective for many deep stratifying lakes. Those same models; however, may not be applicable to all lake types.

With respect to temperature profiles, 1-D models have had great successes with deep and stratifying lakes which are well represented by a wide variety of capable models and publications (Perroud et al. 2009; Huang et al. 2019). Shallow and polymictic lakes; however, have seen far less specific representation in modelling. Shallow lakes are typically polymictic, either stratifying multiple times annually or are completely mixed at all times (McEnroe et al. 2013). Shallow polymictic lakes can alternate between periods of stable thermal stratification and complete mixing, often as a result of changing meteorological conditions (Soulignac et al. 2017), highlighting important factors for modelling in shallow lakes. Due to this dynamic nature of shallow polymictic lakes, models may yield many differing results due to model specific features, for example, the presence and type of parameterization of heat transport in lake sediments, the parameterization of turbulent mixing in the water column, the formulation of the absorption of

solar radiation, the parameterizations of surface sensible, latent heat, and momentum fluxes, and minimum timestep (Stepanenko et al. 2013).

Shallow polymictic lakes are not often the focus of studies or modellers when producing new models (Woolway et al. 2017). When looking at some of the most popular thermal lake 1-D models used in literature, they are often found being applied to deeper lakes than shallow polymictic lakes (Long et al. 2007; Tanentzap et al. 2007). General Lake Model (GLM), for example, is a popular model which is capable of modelling thermal stratification and yet at the time of writing there are no publications found indicating its success in reproducing thermal stratification in a shallow polymictic lake. There are some models, like FLake and Simstrat, which are known to have successfully reproduced thermal stratification for shallow polymictic lakes (Wilhelm and Adrian 2008; Stepanenko et al. 2013; Shatwell et al. 2016).

Understanding the main factors which cause shallow polymictic lakes to mix as well as how often they mix can assist with important connections to lake phenomenon such as hypolimnetic anoxia, internal nutrient loading, and phytoplankton blooms (Heaney et al. 1986). Lake mixing replenishes hypolimnetic oxygen concentrations, whereas a lack of mixing can result in anoxia which may limit aquatic species which rely on aerobic respiration to the epilimnion (Kramer 1987). Mixing may also reduce internal loading of soluble iron and phosphorus from the sediment by preventing their respective reduction-oxidation (redox) reactions through increased redox potential via the presence of dissolved oxygen (Mortimer 1942). Internal loading has been found to be a significant source of iron and phosphorus for phytoplankton communities, especially iron in the case of cyanobacteria (Molot et al. 2014). The importance of mixing on

phytoplankton blooms extends past internal loading, such as mixing leading to nutrient circulation from the hypolimnion into the epilimnion as well as changes in light attenuation which may influence the size and community composition of blooms (Hamilton et al. 2016; Northington et al. 2019). A better understanding of these factors also allows for the improvement of models to mimic and reproduce accurate simulations of these lakes.

The objective of this chapter is to evaluate 1-D model reproduction of the measured thermal stratification in shallow polymictic Lake 303 and shallow dimictic Lake 304. Five lake models are assessed on their success in reproducing the measured thermal stratification in a dimictic lake and a polymictic lake which stratifies for periods ranging from one hour to several days.

3.3 Methods

3.3.1 Study Sites

Within the boreal shield landscape of northwestern Ontario lies the International Institute of Sustainable Development Experimental Lakes Area (IISD-ELA). Lake 303 and Lake 304 are two neighboring headwater lakes at IISD-ELA with sediment-lined granitic bedrock bottoms surrounded by coniferous forests with thin soils. Lake 303 is a shallow polymictic lake with a surface area of 102,012 m^2 and mean and max depths of 1.86 m and 2.79 m, respectively (IISD Experimental Lakes Area 2022).

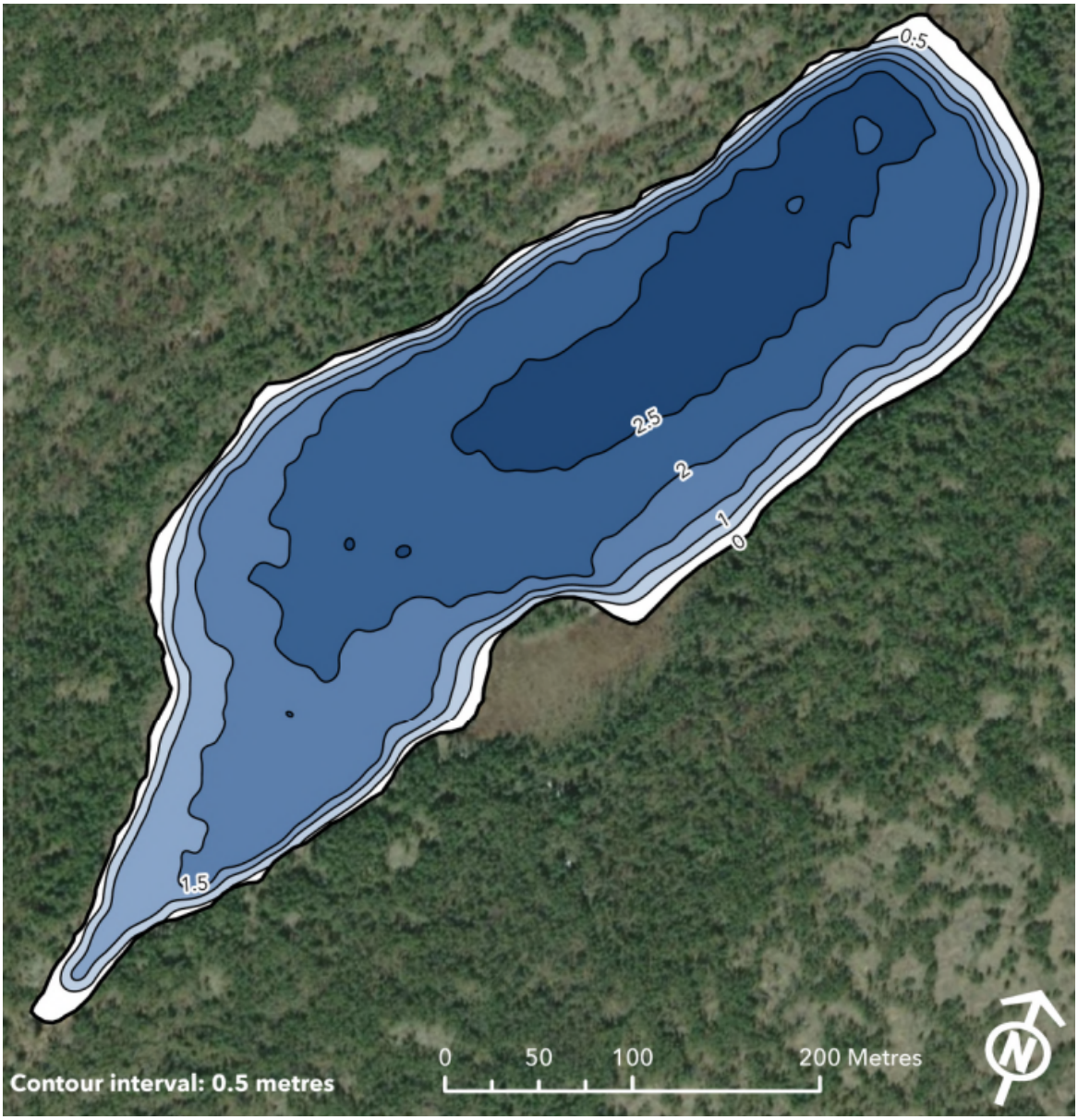

Figure 3.1. Lake 303 bathymetry. This figure was obtained and edited from IISD Experimental Lakes Area 2022.

Lake 304 is a dimictic lake with a surface area of 37502 m^2 and mean and max depths of 3.22 m and 7.22 m, respectively (IISD Experimental Lakes Area 2022). With Lake 304s mean depth being less than 5 m and its max depth not much above 5 m, the assumption is made here that it still behaves as a shallow lake. Both lakes have been a part of fertilization experiments in the past.

Lake 303 was fertilized in 1975 and 1984 with P and N, and in 2019 with only while Lake 304 was fertilized in 1972 with P, N, and C, in 1974 with N and C, in 1976 with N and P, and in 2019 with only P (Levine and Schindler 1989; Molot et al. 2021). Rainbow trout (*Oncorhynchus mykiss*) were introduced into both lakes in 1969; however, none survived in either lake. Afterward, rainbow trout were again re-introduced to both lakes in 2011 and 2012 again, to study growth dynamics in wild populations.

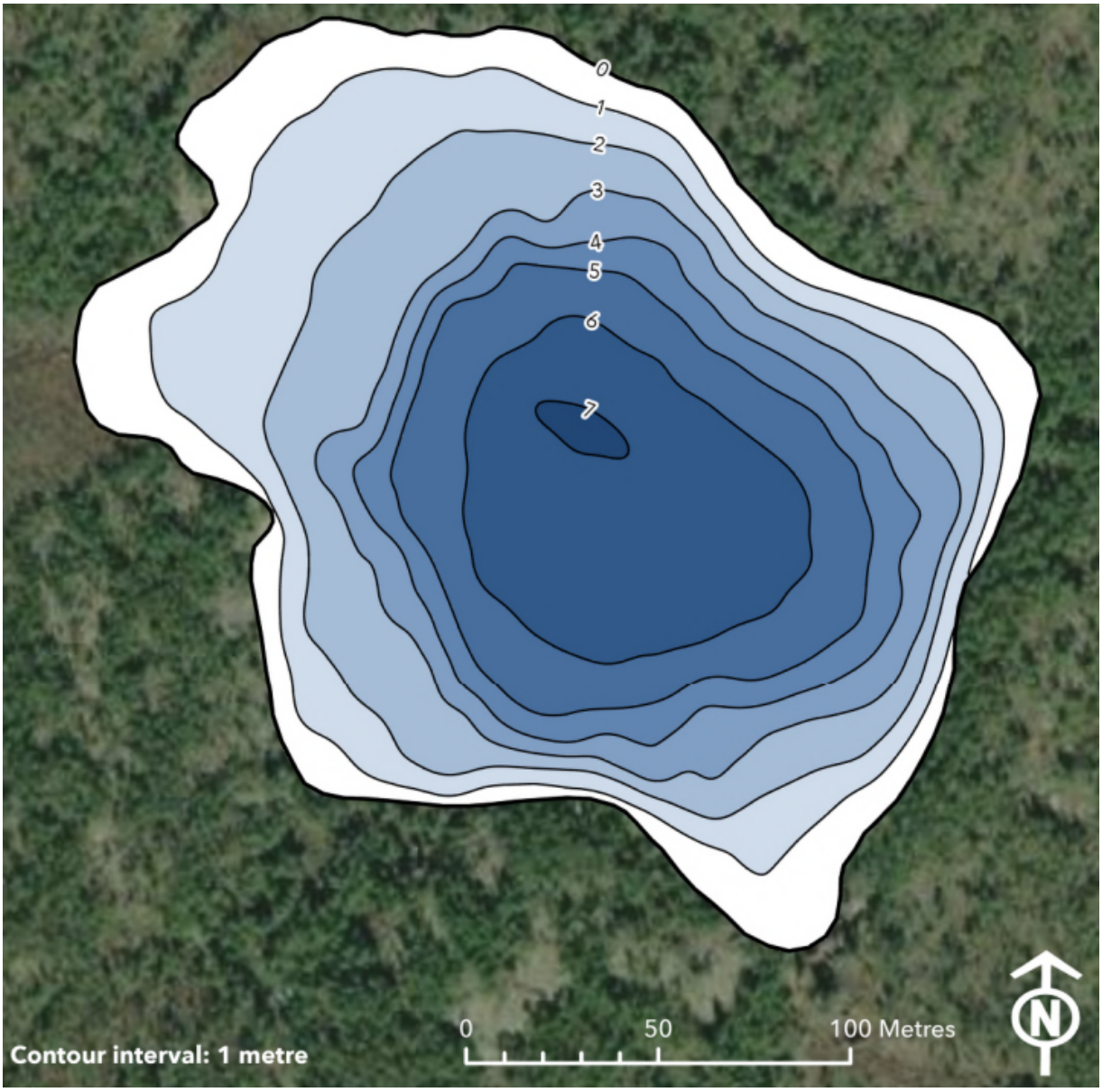

Figure 3.2. Lake 304 bathymetry. This figure was obtained and edited from IISD Experimental Lakes Area 2022.

Due to its polymictic nature, Lake 303 was selected to test the effectiveness of models in reproducing thermal stratification in a shallow polymictic lake. While Lake 304 is also a shallow lake, its deeper bathymetry and differing dimictic regime allows for a contrast from Lake 303 to

see how well the models could reproduce longer periods of stratification under the same environmental conditions.

3.3.2 Data Sources

Water temperature data was collected from May 20, 2019 to October 10, 2019 using HOBO U26 DO loggers and HOBO TidbiT MX2203 temperature loggers. Due to an equipment error, data after October 6, 2019 were omitted. HOBO U26 DO loggers collect water temperature data at a resolution of 0.02 °C and an accuracy of ± 0.2 °C. HOBO TidbiT MX2203 temperature loggers collected water temperature data at a resolution of 0.01 °C and an accuracy of ± 0.2 °C. The HOBO U26 DO logger collected data every 10 minutes while the HOBO TidbiT MX2203 temperature logger collected data every 15 minutes; however, only the data on the hour was used to synchronize water temperature and oxygen data with the environmental parameters. In Lake 303, water temperatures were taken at 1, 1.5, and 2 m depths while DO was taken at 1 and 1.5 m depths. In Lake 304 water temperatures were taken at 1, 1.5, 3, 4, and 5 m depths while DO was taken at 1 and 1.5 m depths.

Data for the wind speed, wind direction, air temperature, shortwave radiation, precipitation, and relative humidity were obtained from the Rawson Lake meteorological station (Station ID: 30455, Climate ID: 6036904) at IISD-ELA, located approximately 2 km from both Lake 303 and Lake 304. Wind speed data was collected at 10 m using a RM Young anemometer every 5 s averaged for each hour to the nearest 0.1 km h^{-1} at an accuracy of ± 0.7 km h^{-1}. Wind direction was also collected by the same RM Young anemometer at a resolution of 0.1 °C every 5 s and

averaged each hour. Air temperature was recorded with a Campbell Scientific CR3000 Micrologger built-in thermistor at a resolution of 0.01 °C and an accuracy of ± 0.3 °C. Precipitation was collected with a HyQuest Solutions TB-4 tipping bucket rain gauge at a resolution of 0.1 mm and an accuracy of ± 2%. Shortwave radiation was calculated from photosynthetically active radiation, which was measured by a Licor LI-190R Quantum Sensor with an accuracy of ± 1% and a resolution of 5-10 μA per 1000 μmol s^{-1} m^{-2} and recorded by a Campbell Scientific CR1000 data logger every 5 s and averaged every 15 minutes. The light extinction coefficient was calculated from a best fit curve of underwater photosynthetic active radiation (PAR) measurements taken each month by a Licor LI-192 Underwater Quantum Sensor with a sensitivity of ~4 μA per 1000 μmol s^{-1} m^{-2} and an absolute calibration of ± 5% in air traceable to NIST and recorded by a LI-1400 logger. Relative humidity was measured using a Vaisala HMP 45C Temperature and Relative Humidity Probe with an accuracy of ± 2% when relative humidity is between 0-90% and ± 3% when above 90%. Relative humidity was logged hourly by a Campbell Scientific CR3000 Micrologger.

3.3.3 1-Dimensional Modelling

All modelling was conducted in R (v4.2.3) through the package LakeEnsemblR (v1.2.6) which facilitates running 5 separate models from a single set of input files and produces a single file with each of the models' outputs (Moore et al. 2021). The models within LakeEnsemblR are FLake, GLM, GOTM, MyLake, and Simstrat. Each of these models are often used in a variety of ecosystems in both academic research as well as environmental management. In order to run

LakeEnsemblR for all the models it requires the basic bathymetry of the lake, a set of input water temperature data and several meteorological variables. The required meteorological variables are downwelling shortwave radiation, air temperature, wind speed at 10 m, precipitation, relative humidity, and sea level barometric pressure. LakeEnsemblR can also use data from variables like downwelling longwave radiation, cloud cover, dewpoint temperature, wind direction, and vapour pressure, but each of these can be calculated from the required variables if they are not specifically provided. Incorporated into LakeEnsemblR for both lakes was also a dataset of monthly light extinction coefficients calculated from in-situ measurements.

FLake was created by Mironov in 2008 to cater toward small-to-medium sized lakes due to an underrepresentation at the time. FLake uses a two-layer parameterization of the evolving temperature profile and the integral energy budget. FLake uses the concept of self-similarity, or assumed shape, of the temperature-depth curve to describe the structure of the thermocline. The concept of self-similarity pertains to the temperature profile, allowing for the description of the vertical temperature structure of the thermocline. This structure can be reasonably and accurately represented by a "universal" function of dimensionless depth. FLake also applies the concept of self-similarity to calculate the temperature of the thermally active upper layer of the bottom sediments. The bottom sediments in shallow lakes are known to store heat which is returned to the water column and is typically most impactful during fall and winter. For calculating the depths of the convectively-mixed layer and wind-mixed layer, FLake uses entrainment and relaxation-type equations, respectively. FLake is capable of running at hourly to daily timesteps and uses the lake-specific parameters of lake depth, water optical characteristics,

the temperature at the bottom of the thermally active layer of bottom sediments, and the depth of the thermally active layer of bottom sediments.

The General Lake Model (GLM) (Hipsey et al. 2019) was developed in 2012 by the Global Lake Ecological Observatory Network (GLEON) to support the need for a model suitable for lake types of substantial variability in morphology, hydrology, and climatic conditions. GLM uses an energy balance approach based on estimating the available turbulent energy which is calculated specifically for both surface mixing and mixing below the thermocline. Supporting both hourly and daily timesteps, GLM is capable of providing a custom number of thin layers at a fine resolution which can be particularly useful for focusing on the thermocline.

The General Ocean Turbulence Model (GOTM) was first published in Burchard et al. 1999 for use in the most important hydrodynamic and thermodynamic processes related to vertical mixing in natural waters. GOTM was formed using code from a number of effective turbulence models with various complexities, leading to its key component for solving equations regarding turbulent fluxes of quantities of momentum, salt, and heat. GOTM is capable of running at timesteps of an hour or longer.

MyLake was developed by Saloranta and Andersen 2007 to meet the need for a model more well-suited to efficient uncertainty and sensitivity analyses, something many models typically take very long periods of time to process. MyLake specifically calculates surface heat fluxes, wind stress, light attenuation and also heat flux between sediment and water. MyLake is restricted to a minimum daily timestep.

Simstrat is based on turbulence closure schemes, similar to GOTM, in which rates for vertical transport are related to turbulent kinetic energy, rather than being directly influenced by external forces (Goudsmit et al. 2002). By combining the use of a buoyancy-extended k-ε model with a seiche excitation and damping model, Simstrat is capable of predicting diffusivity below the surface mixed layer in stratified lakes where most k-ε models typically predict negligible turbulent kinetic energy. Simstrat also supports timesteps of an hour or longer.

Each of the models within LakeEnsemblR has the capability to calibrate model specific parameters and scaling factors for the input meteorological forcing which is calculated against the observed the observed water temperatures. The method used for calibration was Latin Hyper Cube (LHC) sampling, in which a near-random sample of the calibratable parameters' values were taken and tested for fit against the observed water temperature values. FLake has a custom and calibratable value for cloud cover. GLM can be calibrated for its mixing coefficient within the hypolimnion. GOTM can be calibrated for its value for minimum turbulent kinetic energy. Simstrat has a calibratable seiche value. Finally, MyLake can be calibrated for its constant for physical wind sheltering. All 5 models were calibrated individually for general wind factors and shortwave radiation. The .yaml file data with all lake specific information and the model calibrated values can be found in Appendix 5.1 and Appendix 5.2.

Both Simstrat and FLake have been used in previous studies looking at shallow lakes (Stepanenko et al. 2013, 2014). In one study they calculated the temperature difference between the surface and at 1 m of depth in the observed data and compared it to both models. FLake overestimated this temperature difference by 2.88 °C on average while Simstrat overestimated

it by only 0.07 °C on average. In another study Simstrat was reported to have produced a thermal profile most similar to the observations (Stepanenko et al. 2014). MyLake was successful in reproducing thermal stratification in Võrtsjärv, a shallow polymictic lake in Estonia with a maximum depth of 6.1 m (Woolway et al. 2017). At the time of writing GOTM and GLM are not known to have been used to investigate or successfully reproduce thermal stratification in any shallow polymictic lakes.

The performances of each model will be assessed first by comparing their ability to predict thermal stratification in general for both lakes. As it was done in Chapter 2, thermal stratification is defined as $\geq$ 1 °C per meter depth within the water column. For Lake 303, the models will be assessed for whether or not they can reproduce a similar frequency of stratification events to the observations. Finally, the models' thermal profiles will be evaluated for accuracy to the observed mean temperatures for the entire lake, the epilimnion/mixed layer, and the hypolimnion/deep layer. These accuracies will be evaluated using RMSE and difference of mean (DM).

3.4 Results and Discussion

3.4.1 Model Reproduction of Thermal Stratification

The thermal profiles calculated by each of the models for Lake 303 (Figure 3.3) and Lake 304 (Figure 3.4) exhibit different degrees of reproduction of the observed thermal profile. For Lake 303, GLM (Figure 3.3c,d) and GOTM (Figure 3.3e,f) predicted temperatures with the most short-lived stratification events, the most similar of all models to the observed thermal stratification (Figure 3.3k,l). GLM predicted temperatures yielding 35 stratification events, many

of which began and ended on or around similar dates to the observed events. GLM's longest and shortest periods of stratification were 91 hours and 1 hour, respectively. GOTM calculated water temperatures with 55 stratification events, with many also initiating and ending at similar dates to the observations; however, it's longest period of stratification was just 37 hours, compared to the observed stratification events which lasted as long as 106 hours. Both GOTM and GLM reproduced the observed general trend of a peak in stratification frequency from June 16 to July 2, as well as a low in stratification frequency during most of August. Simstrat was slightly less successful at predicting stratification, with a total of 24 periods of stratification and the longest of those lapsing 27 hours (Figure 3.3i,j). The calculated water temperatures for both FLake (Figure 3.3a,b) and MyLake (Figure 3.3g,h) did not show any thermal stratification in Lake 303 during the study period. Stepanenko et al. 2013 demonstrated the use of FLake to successfully calculate a stratified thermal profile in Großer Kossenblatter See (mean and max depths of 2 m and 5 m, respectively, however, this is approximately twice as deep as Lake 303. FLake's inability to calculate stratification in the thermal profile may be due to its calculations of the thermal profile in bottom sediments here, potentially requiring temperature data closer the sediment for greater accuracy. Stepanenko et al. 2014 found FLake to struggle with accurately calculating the development of the mixed layer, namely under weak wind conditions. The same study also reported that the thermocline in FLake was shallower than both the observations and the other models in the study, one of which was Simstrat. MyLake also, was shown to calculate a stratified thermal profile for Võrtsjärv (mean and maximum depths of 2.8 m and 6.1 m, respectively) (Woolway et al. 2017), however, in addition to being half as deep, the minimum daily timestep is

likely to be an issue for not calculating a stratified thermal profile for Lake 303 given the short durations of the observed thermal stratification events.

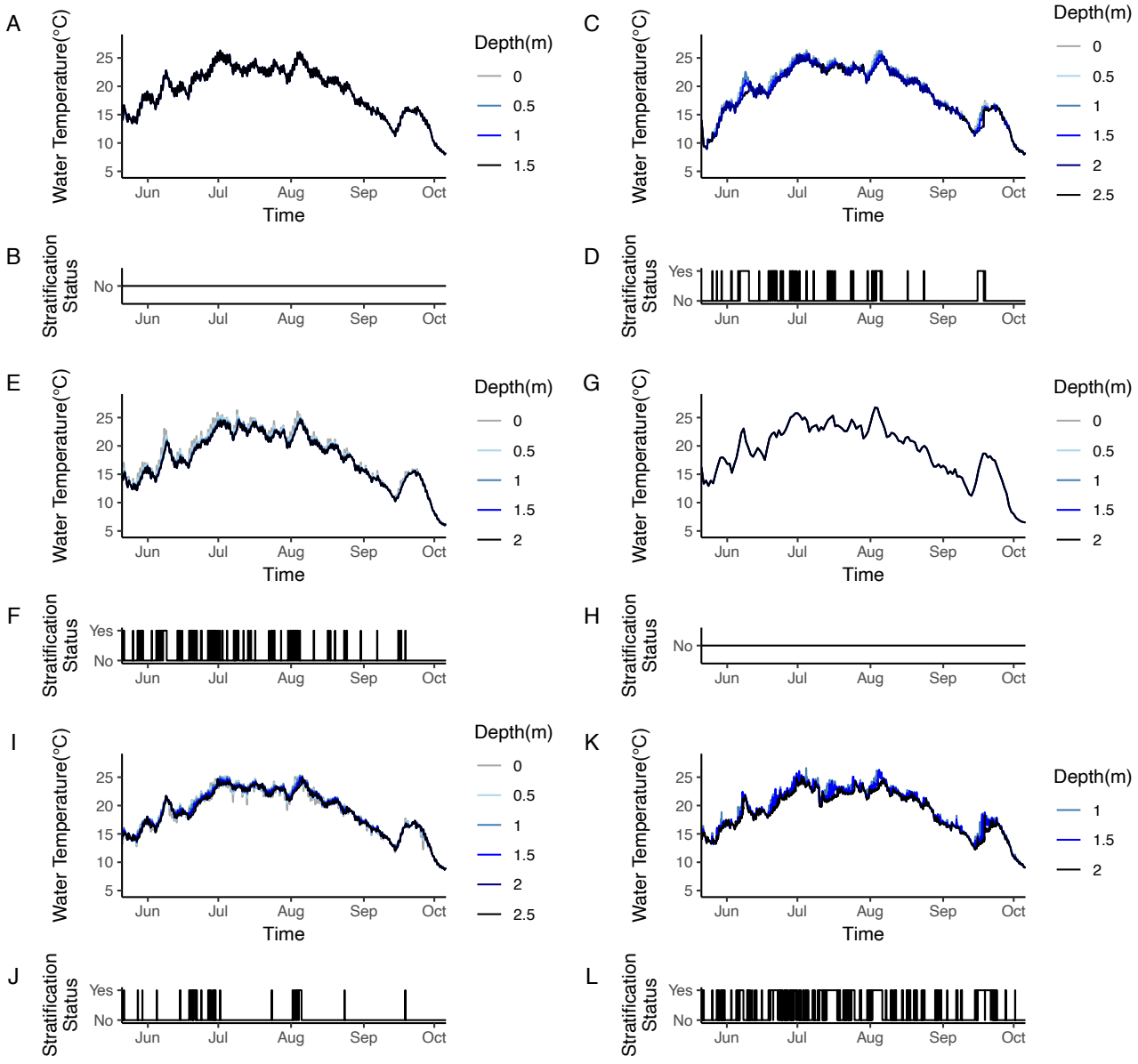

Figure 3.3. Thermal profiles and stratification statuses for Lake 303 from each of the 5 models and the observations. A) FLake thermal profile and B) stratification status, C) GLM thermal profile and D) stratification status, E) GOTM thermal profile and F) stratification status, G) MyLake thermal profile and H) stratification status, I) Simstrat thermal profile and J) stratification status, K) observed thermal profile and L) stratification status.

As Lake 304 is dimictic and more than twice as deep as Lake 303, models should generally be better equipped to reproduce its thermal profile due to its thicker mixed layer, which can withstand a greater amount of wind-induced mixing and decreases in air temperature which cause lakes to mix. GLM calculated a stratified water column for nearly the entire study period until turning over in early October, similar to the observations (Figure 3.4c,d). Simstrat reproduced most of the entire stratification; however, it calculated the summer stratification period to end on August 21, before stratifying again for approximately one week in late September (Figure 3.4i,j). FLake calculated Lake 304 to be stratified for approximately half of the study period, with brief mixing events in between before becoming isothermal after August 8th (Figure 3.4a,b). GOTM predicted 17 brief stratification periods each lasting between 1-3 hours (Figure 3.4e,f). Of the 17 brief periods, 14 of them were calculated to occur between 14:00 and 18:00, indicating a pattern of daily warming followed by nightly cooling. MyLake calculated just 7 periods of stratification which ranged from 1 day to 7 days (Figure 3.4g,h). When looking at the full thermal profiles of the models, GLM and Simstrat performed best in reproducing the colder hypolimnetic temperatures measured in Lake 304, compared to the other three models which predicted a much warmer hypolimnion.

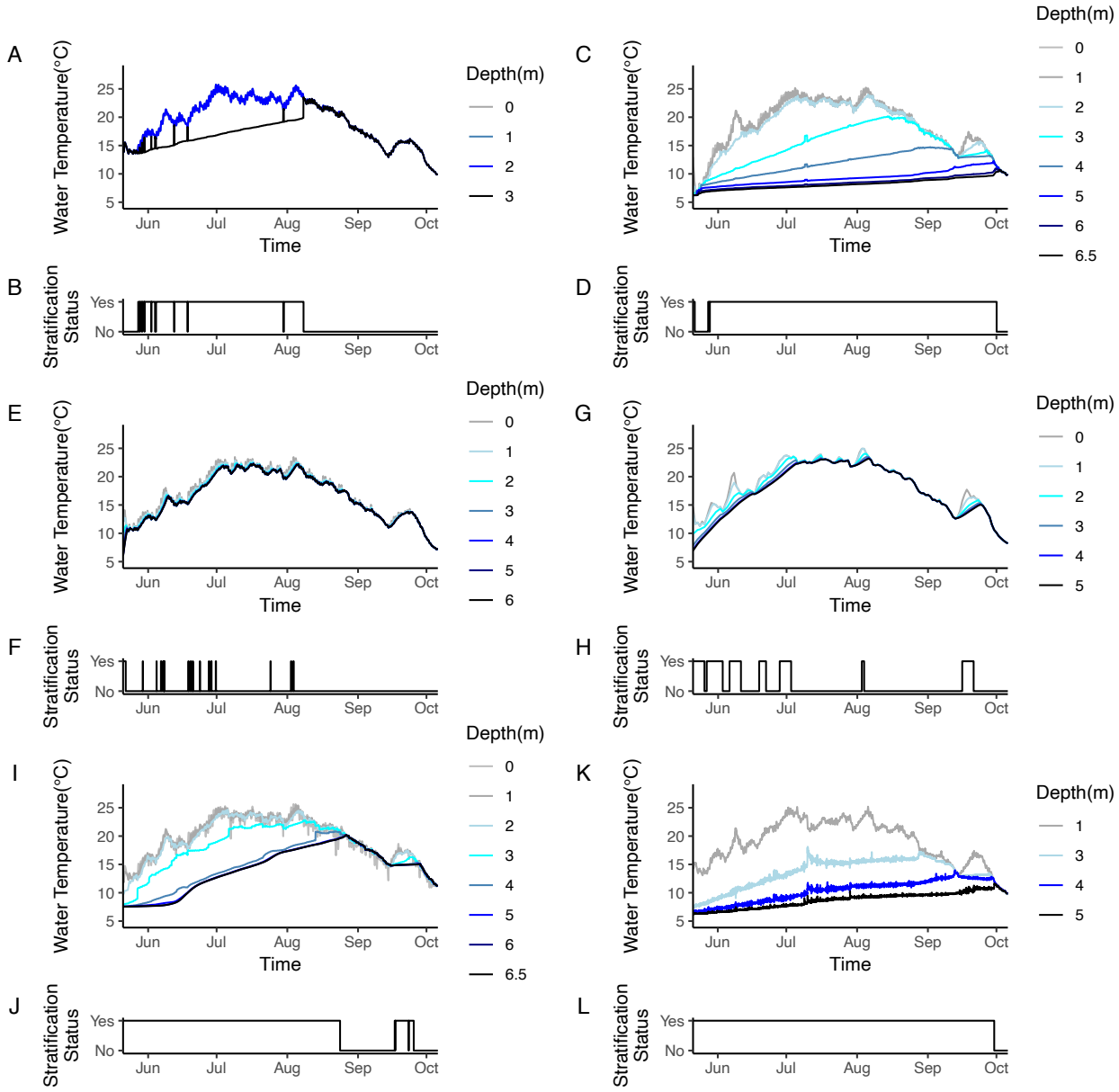

Figure 3.4. Thermal profiles and stratification statuses for Lake 304 from each of the 5 models and the observations. A) FLake thermal profile and B) stratification status, C) GLM thermal profile and D) stratification status, E) GOTM thermal profile and F) stratification status, G) MyLake thermal profile and H) stratification status, I) Simstrat thermal profile and J) stratification status, K) observed thermal profile and L) stratification status.

Mean lake temperatures calculated by each of the models were much closer to the observed mean temperatures for Lake 303 than Lake 304. During the spring and early summer

portion of the study for Lake 303, FLake, MyLake, and GLM appeared to slightly overestimate the mean lake temperature while GOTM typically underestimated it (Figure 3.5a). For the latter half of the summer and into fall, all of the models mostly underestimated the mean lake temperature, with a few exceptions of spikes in temperature calculated by MyLake throughout the study period. The difference of means (DM) for the models for Lake 303 ranged from -0.2 °C to 0.1 °C excluding GOTM which yielded -1.2 °C. GOTM and MyLake had the largest root mean square errors (RMSE), being 1.6 °C and 1.7 °C, respectively (Figure 3.6). Simstrat had a DM of -0.1 °C and the best RMSE out of all the models, which based on these suggests it calculated the most accurate mean temperature of the Lake 303. Stepanenko et al. 2014 found similar results for Simstrat in which they reported it to have produced the most accurate water temperature distribution of a shallow lake compared to FLake and several other 1-D models.

Each of the models overestimates the mean lake temperature of Lake 304 for most of the study period (Figure 3.5b). GLM plotted closest to the observed mean and had the best DM and RMSE values, 1.0 °C and 1.3 °C, respectively (Figure 3.6). For the first 6 days of the simulation GLM predicted the lake to be colder than the observed temperatures by as much as 1.4 °C before rising above the observations afterward. FLake plotted the farthest from the observed mean and had the highest DM and RMSE values, 6.2 °C and 6.8 °C, respectively.

Other studies using FLake have also shown that it can overestimate temperature in shallow lake thermal profiles, and also calculate significant positive temperature spikes (Stepanenko et al. 2014). Most models which are involved in comparisons with FLake differ by being typically turbulence or energy-balance based models, potentially highlighting issues with

how FLake calculates the surface waters and mixed layer (Stepanenko et al. 2013, 2014; Sun et al. 2020).

LakeEnsemblR has a minimum requirement for water temperature input data to have at least three depths, which means that the models here are operating with the minimum amount of observation depths for Lake 303 whereas for Lake 304 there are four observation depths. A minimum of three depths for water temperature input data potentially suggests that the greater number of observations for Lake 304 allowed for the models to more accurately identify a stronger thermal gradient than for Lake 303.

As model water temperature equations can depend on depth, the mixed layer and the layer below, referred to as the deep layer, were separately analyzed to investigate potential reasons for the differences between the models and the observations.

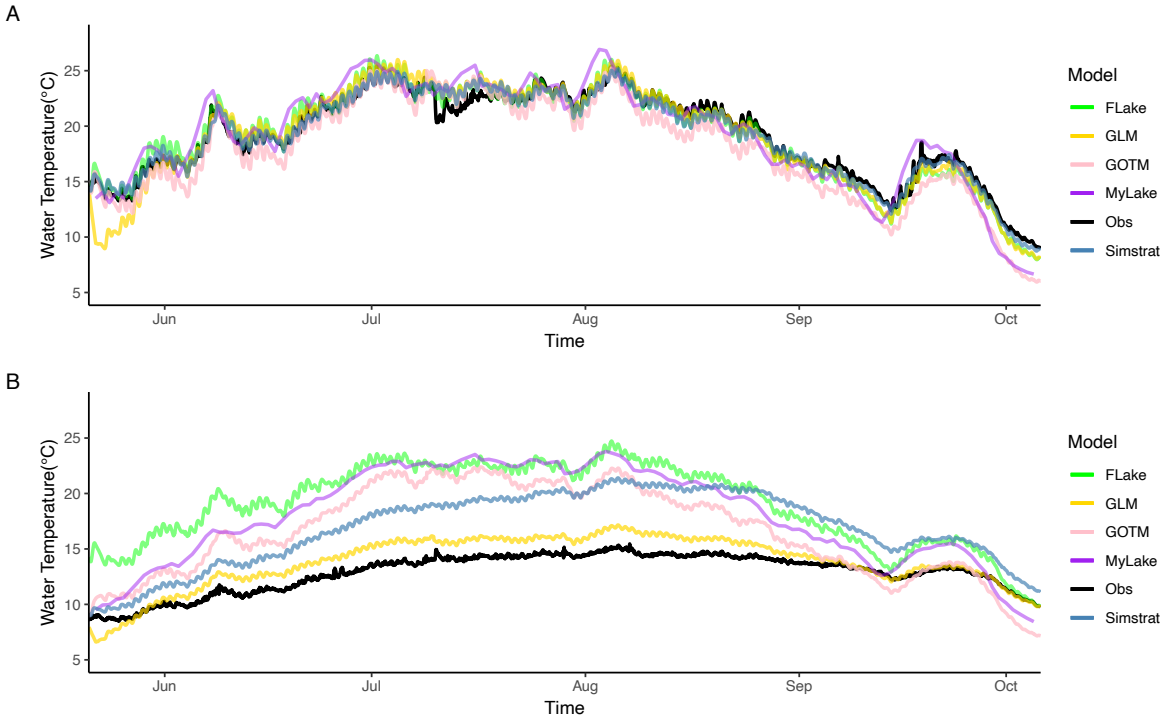

Figure 3.5. Mean temperature of the entire water column for a) Lake 303 and b) Lake 304 from each model and the observations (Obs).

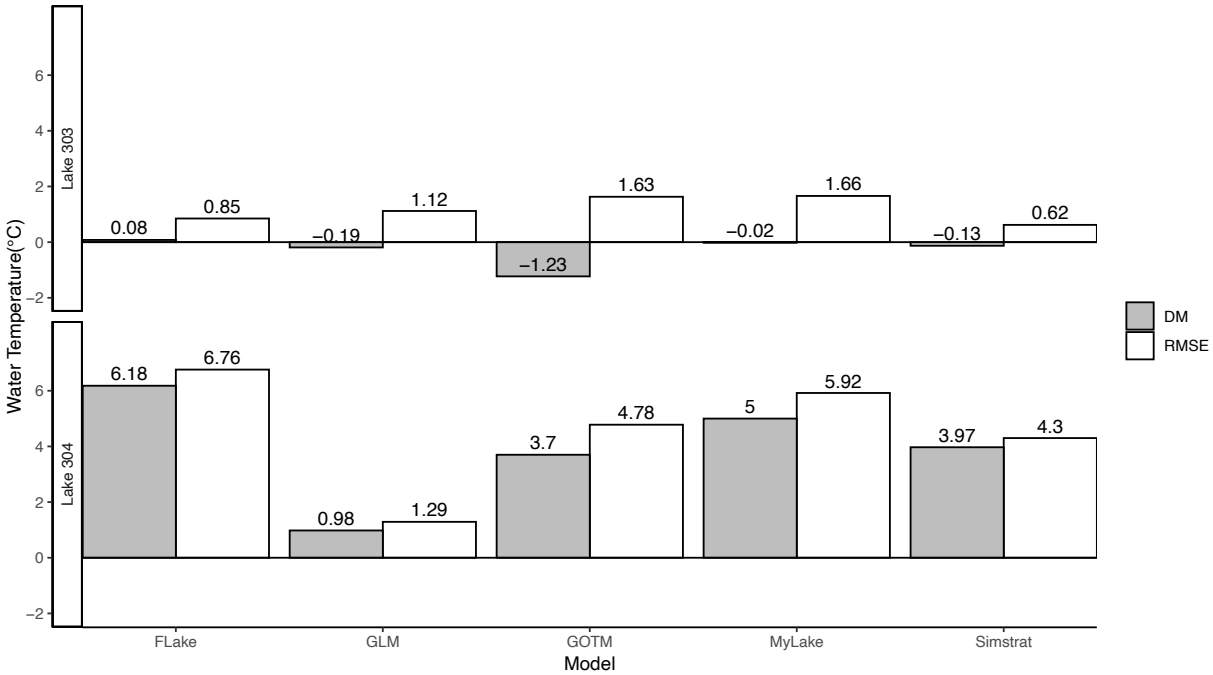

Figure 3.6. Difference of means (DM) between modeled and observed data and the RMSE of each model for the entire water columns of Lake 303 and Lake 304.

3.4.2 Temperature Reproduction in the Mixed Layer

The mixed layer is the surface layer of a lake above the thermocline where temperatures are uniform throughout. The mixed layer temperatures were calculated here as the mean of the temperatures at each of the depths above the thermocline. The thermocline depths for Lake 303 and 304 were calculated via the R package, rLakeAnalyzer, to be 1.5 m and 2.5 m, respectively, for most of the study period so these depths were selected and kept consistent for this experiment. Each of the models appeared to mimic the observations for both lakes, with a fairly congruent but offset shape in the graph (Figure 3.7). For Lake 303 (Figure 3.7a), FLake, GLM, and Simstrat reproduced the mixed layer accurately, typically deviating only during local minima and maxima where the changes in temperature were relatively rapid. GOTM typically underestimated the temperature of the mixed layer resulting in a DM value of -1.4 °C, while MyLake typically overestimated the mixed layer temperature for the first half of the study and underestimated for the latter half. Simstrat had the best RMSE value and the median DM value, while FLake had the best DM value and the second best RMSE (Figure 3.8). The mixed layer in Lake 304 (Figure 3.7b) was slightly warmer than many of the models had calculated for most of the study period. FLake and Simstrat plotted quite similarly to the observed mixed layer temperatures, with some overestimation typically toward the latter half of the study period. MyLake displayed a decent reproduction of temperature relative to the other models, after starting the simulation quite colder than the observations but then warming soon after. GLM began the simulation as the least accurate model, underestimating the mixed layer temperature in Lake 304 by 6.8 °C on May 22 before aligning similarly with the observations around early July. All the models except for FLake,

in fact, undercalculated the initial observed temperature of the mixed layer by at least 2.8 °C. FLake had the best RMSE value and tied with Simstrat for the best DM value.

GOTM mostly underestimated the temperature of the mixed layers for both lakes and was the least accurate model on overall when considering the ranks for DM and RMSE values. MyLake often predicted much stronger changes in temperature compared to the observations and produced slightly better DM and RMSE values than GOTM; however, it is difficult to compare MyLake to the other models based on calculated thermal profiles due to its daily timestep. FLake and Simstrat quite closely reproduced similar results to the observed temperatures, supported by their relatively lower DM and RMSE values in comparison to the other the models. These findings are similar to others which report FLake to be the better choice between models for recreating surface water temperatures in shallow lakes (Haddout et al. 2018). GLM also reproduced the observations quite closely, but only after warming up for the first few weeks of the simulation which likely offset the DM and RMSE values.

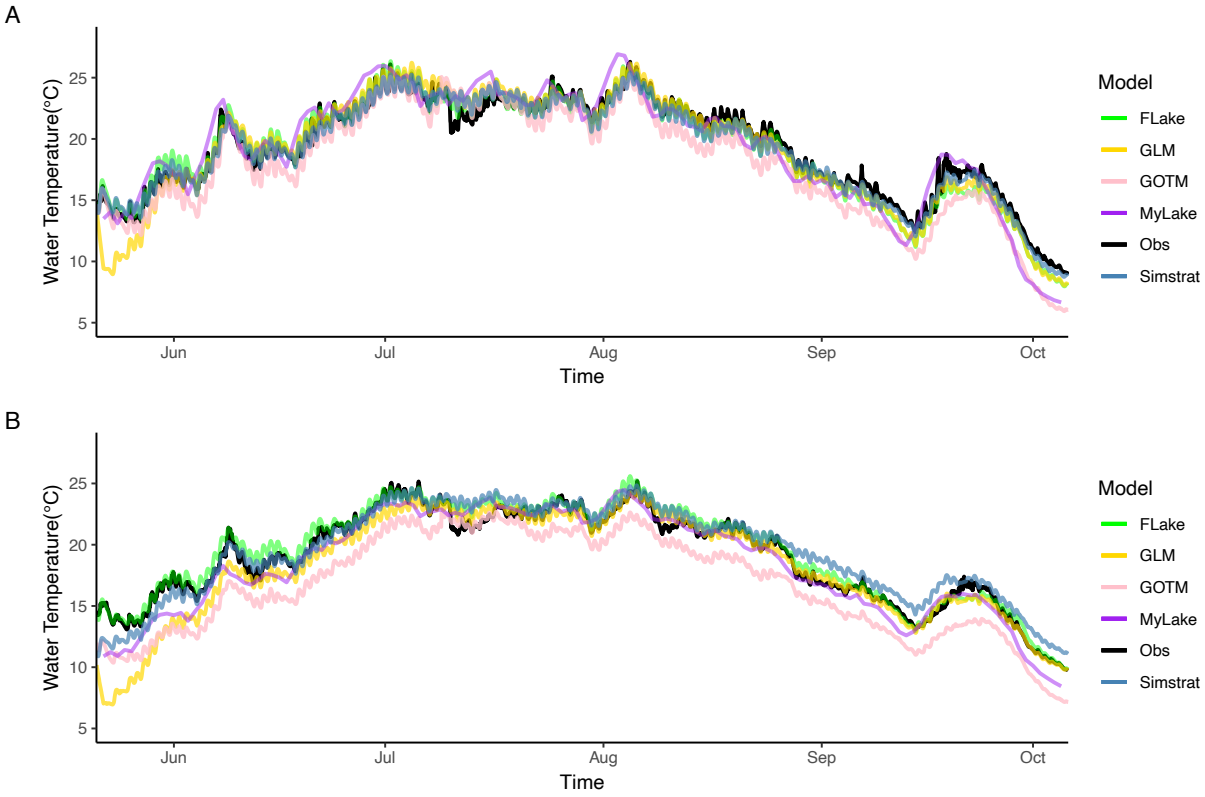

Figure 3.7. Mean temperature of the mixed layer for a) Lake 303 (0-1m) and b) Lake 304 (0-2m) from each model and the observations (Obs).

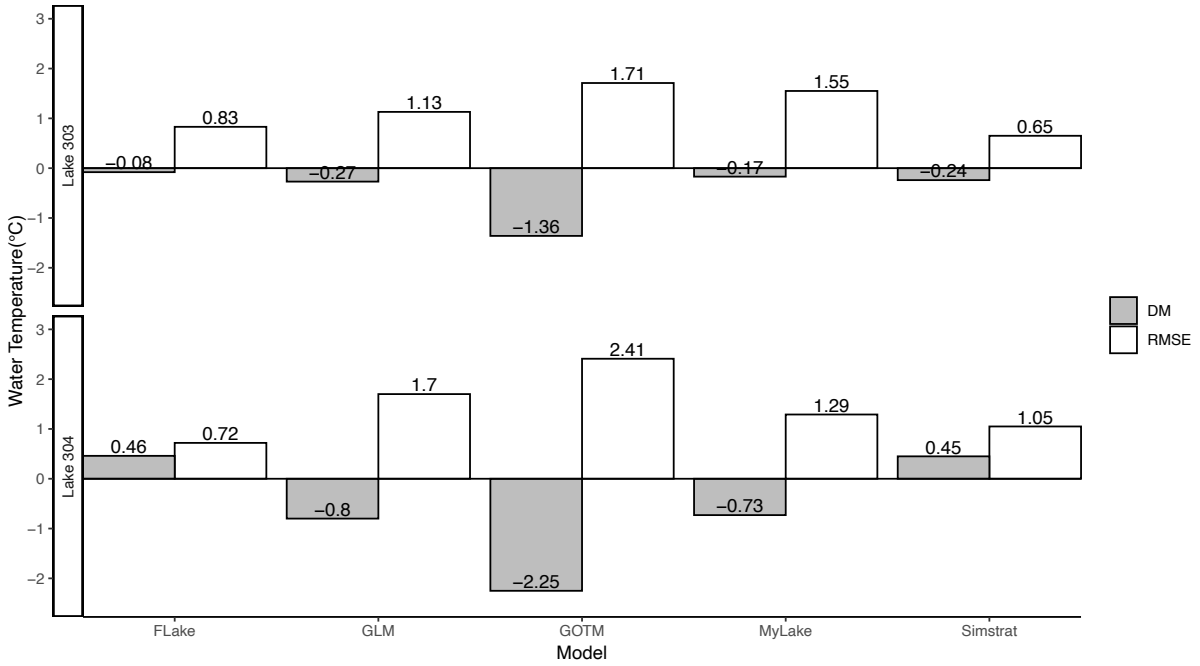

Figure 3.8. Difference of means (DM) between modeled and observed data and the RMSE for each model in the mixed layer for Lake 303 (0-1m) and Lake 304 (0-2m).

3.4.3 Temperature Reproduction in the Deep Layer

The hypolimnion temperatures were calculated as the mean of the temperatures from each of the depths from the thermocline and below. For Lake 303, the models again appear to have produced a temperature curve similar to the observations (Figure 3.9a). In the first half of the study FLake and MyLake mostly overestimated water temperatures but drew closer to the observations in the latter half. MyLake again predicted much steeper temperature changes at many points of the study compared to each of the other models and the observations. GLM reproduced similar temperatures to the observations, except for the beginning of the simulation where it underestimated by as much as 5.1 °C, which likely negatively affected its RMSE and DM values which were at the median and slightly better than the median, respectively. GOTM mostly underestimated throughout the entire study, reflected by having the worst DM value, -1.1 °C (Figure 3.10). Finally, Simstrat performed the best by reproducing a similar temperature curve and also featured the best DM and RMSE values.

As for Lake 304, each of the models except for GLM performed quite poorly compared to their mixed layer results (Figure 3.9b). For the first half of the study FLake typically overestimated temperatures by about 5-7 °C and as much as 10.6 °C on August 8. Both GOTM and MyLake predicted a steeper increase followed by a steeper decrease in the temperature curve over the study period compared to the more subtle and gradual temperature increase observed until the beginning of October. MyLake also had the worst values for DM and RMSE among all models, being 6.6 °C and 7.8 °C, respectively. Simstrat reproduced a temperature increase which

continued into late August and early September where it predicted the warmest temperatures of all the models. GLM gave the best reproduction of the Lake 304 temperature curve as well as the best DM and RMSE values, -0.2 °C and 0.4 °C, respectively.

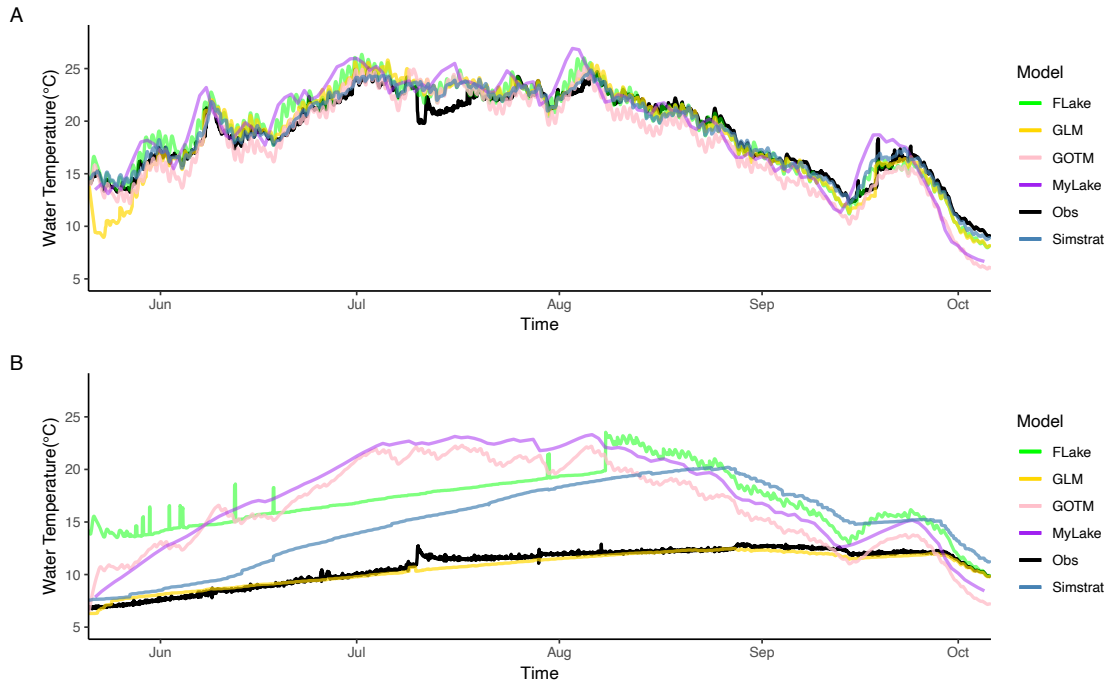

Figure 3.9. Mean temperature of the deep layer for a) Lake 303 (1.5-2.5 m) and b) Lake 304 (2.5-6.5 m) from each model and the observations (Obs).

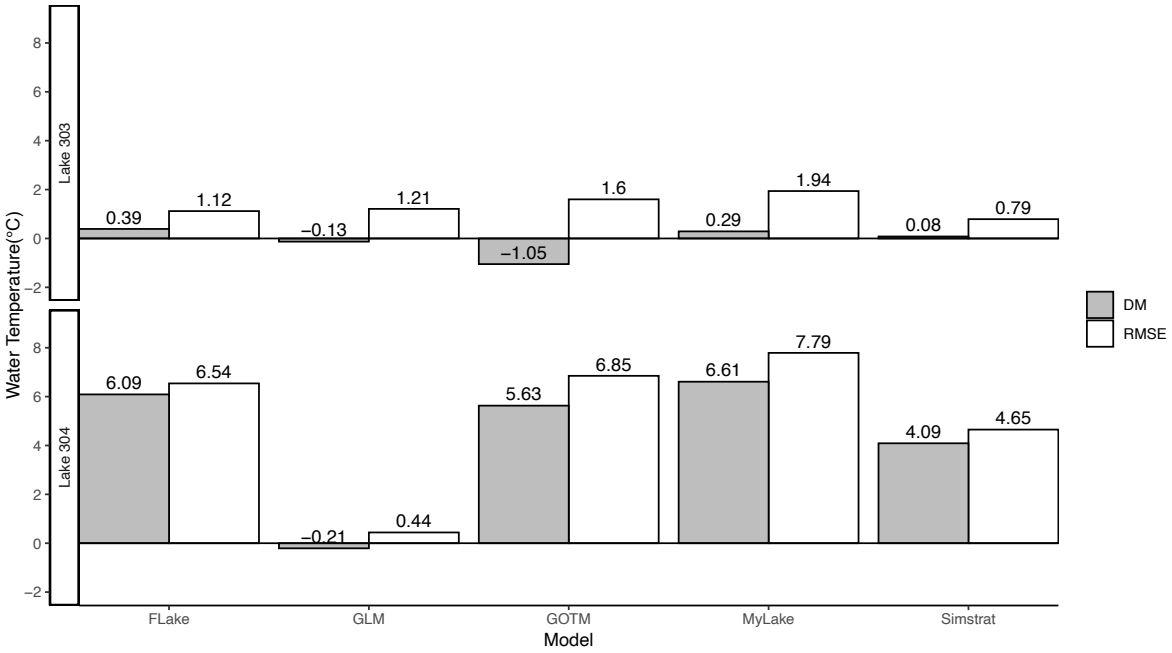

Figure 3.10. Difference of means (DM) between modeled and observed data and the RMSE for each model in the deep layer for Lake 303 (1.5-2.5 m) and Lake 304 (2.5-6.5 m).

To investigate differences between the models in calculating stratification, the mixed and deep layers were separated to narrow a search for notable factors. In Lake 303, both the DM and RMSE values for FLake and MyLake were worse for the deep layer than the mixed layer, suggesting that these models have a stronger ability to calculate temperature at the lake surface than deeper depths. FLake has been shown to calculate temperature gradients near the surface as great as 12 °C m^{-1} in a shallow lake in one study (Stepanenko et al. 2013). FLake has also been previously observed to perform the poorest for shallow lakes in terms of recreating the complete thermal structure compared to other 1-D models, such as Simstrat (Sun et al. 2020). GOTM again had the worst DM and RMSE values; however, GOTM was the only model that exhibited better values for both DM and RMSE in the deep layer than the mixed layer which may be attributed to the success of its turbulent mixing calculations. Simstrat and GLM both had better DM values

when comparing the deep layer to the mixed layer. In Lake 304, every model except GLM had worse DM and RMSE results for the deep layer than the mixed layer. These results suggest that these models typically reproduce the surface and the mixed layer of Lake 303 and Lake 304 quite effectively but stray in their calculations of the thermocline and below.

3.4 Conclusion

Thermal stratification observed in Lake 303 was best reproduced by the models GLM and GOTM, while Simstrat was also somewhat successful. Even still, these three models only reproduced a portion of the frequency of stratification as well as the overall time the lake was stratified. The success of these models could indicate that modelling the entire water column of shallow polymictic lakes requires a focus on turbulence and energy balance calculations. Although GOTM reproduced the highest frequency of thermal stratification as well as relatively similar timings of events, it also had the worst DM and RMSE values for both the mixed and deep layers. GLM was also not as effective relative to the other models in replicating the entire thermal profile of the lake. In this study FLake and MyLake were the only models to specifically consider heat flux between sediment and water which could explain their calculated warmer lake bottom and lack of thermal stratification. For Lake 304, all models did generally well to predict stratification; however, all models underestimated the observed stratification. For Lake 304 GLM was the best model in reproducing the observed stratification which, combined with its success for Lake 303, may prove it to be a good choice for future studies looking to model stratification in shallow lakes. Overall, each of the models appeared to calculate similar trends of warming and cooling for Lake 303 and the upper layer of Lake 304; however, the magnitude of change during shifts in temperature were typically greater in the models than the observations, with sudden

changes in temperature being stronger in the models than were observed. The mixed layer was typically reproduced better than the deep layer for both lakes which could suggest that each of the models is correct in understanding the interactions at the surface but are not well tuned for the deeper depths of these two shallow lakes. For research looking toward modelling the surface water or mixed layer temperatures, the best results might be found using FLake or Simstrat.

Chapter 4 Conclusions and Recommendations

4.1 Transient Thermal Stratification in Lake 303

There are very few case studies on shallow polymictic lakes focusing on their thermal profiles and thermal stratification patterns. As part of Objective 1, The experiments within this thesis successfully investigated Lake 303, a shallow polymictic lake, using hourly temperature data to observe how often and how brief the lake was stratified during the ice-free months of 2019. The frequency and durations of transient thermal stratification within Lake 303 were successfully observed and investigated, with 146 stratification events totaling 1102 of the 3316 hours of the study period from May to October in 2019. Stratification periods ranged from less than one hour to as long as 106 hours with a median duration of 2 hours.

In line with Objective 2, the classification tree, bagged, and random forest models allowed for a clear comparison between the chosen environmental factors and transient thermal stratification in Lake 303. The environmental factors most associated with observed stratification were air temperature, wind direction, and wind speed, in that order. Short-wave radiation and precipitation were not as closely connected to stratification. The classification tree allowed for a simple interpretation of how effective air temperature was as an indicator for stratification, classifying as much as 44% of the dataset within the first node as not stratified based on when temperatures were < 15 °C. The bagged and random forest models agreed with the results of the classification tree in terms of the rank of predictor variables, only differing in their improved rank for wind direction. The lack of connection between short-wave radiation and thermal stratification was not suspected and warrants further research. The only connection regarding short-wave radiation found in the classification tree was actually suggesting that higher levels

were more often associated with mixing than stratification. One possibility for this may have been that the clarity of the water may have allowed for the short-wave radiation to penetrate into the deeper layers of the lake instead of inducing thermal stratification by warming only at the surface.

On August 3 a hypoxic event occurred in Lake 303 which saw DO levels at 1.5 m depth drop to 1.35 mg L^{-1} and remain below 2 mg L^{-1} for 20 minutes before rising above hypoxic levels. This event took place approximately 10 days before the estimated start of a cyanobacterial bloom in the lake which is an interesting timing regarding the possibility for this event to have allowed for redox conditions which may have promoted internal loading of Fe and P. As the data on the lake nutrient levels is limited, conclusions can only be speculated about from this observation.

4.2 Reproducing Transient Thermal Stratification in Models

The final objective of this thesis, Objective 3, was to evaluate the ability of FLake, GOTM, GLM, MyLake, and Simstrat, several commonly used 1-D models, to mimic the observed thermal stratification patterns in both Lake 303 and Lake 304. Lake 303 is a shallow discontinuous polymictic lake, as evidenced by the 146 stratification events observed in the 2019 study period, therefore, the ability of the models to calculate similar stratification was the main objective of this experiment. Lake 304 was also investigated to allow for a comparison with the model's ability to calculate a dimictic thermal regime for another shallow lake under the same environmental drivers. For Lake 303, GOTM calculated 55 thermal stratification events, the most out of all the

models; however, this number is much smaller than the 146 observed events. The next best model was GLM, which calculated 35 stratification events.

In terms of the durations of stratification events, GLM calculated events as short as 1 hour and as long as 91 hours while GOTM calculated events ranging from only 1 to 37 hours, compared to the observed range of 1 to 106 hours. Simstrat trailed GOTM and GLM in the frequency of stratification, with just 24 stratification events, the longest of which lapsed just 27 hours. The remaining two models, FLake and MyLake, did not calculate any stratification within Lake 303. For Lake 304, GLM was the most successful in calculating the observed thermal stratification as well as the overall thermal profile. The success of GLM and GOTM seems to suggest that models which focus on turbulence and energy balance equations may be most effective when modeling the thermal profiles of shallow polymictic lakes. FLake and MyLake were the only models which considered the heat flux at the sediment-water interface, which may be where these models strayed in their calculations for Lake 303. Overall, models were much closer to correctly calculating the temperature for the upper layers of the lakes than the deeper layers.

4.3 Future Research Suggestions

Throughout this thesis, several areas have been identified which stand to benefit from additional research. There is a need for more case studies involving high resolution temperature data for shallow polymictic lakes to better understand their thermal stratification patterns. These case studies would provide more insight on the frequency and duration of thermal stratification in more shallow polymictic lakes. Related to this, research can look into the impacts of transient

thermal stratification and any connections it may have to physical or chemical interactions relating to concepts such as redox and internal loading or phytoplankton blooms. Studies which apply models to shallow polymictic lakes within other regions besides northern Ontario could also be beneficial to identify regional differences and the influences of different climates.

Climate change is an ever-growing research topic (Harper et al. 2021), and so its connections to thermal stratification in shallow polymictic lakes regarding duration and frequency are also of importance. Many studies look into climate change and thermal stratification in lakes as a general group but specific focuses on polymictic lakes could outline issues and findings unique to these lakes.

Research is required to pursue an improved model with the capability of effectively predicting the thermal profiles and stratification patterns of shallow polymictic lakes. Models are an invaluable tool which are a core part of research, but first there needs to be a more accurate and reliable model for shallow polymictic lakes. The investigation in Chapter 3 found that turbulence models as well as energy balance models provided the best overall results while models using equations for heat flux at the sediment-water interface included too much warmth in the hypolimnion. If a new model were to be created specifically for shallow polymictic lakes, using models like GLM, GOTM, and Simstrat as a starting point may prove useful.

Appendix

5.1 LakeEnsemblR L303 .yaml File With Model Parameterizations

```
location:
  name: Lake 303                    # name of the lake
  latitude: 49.7                    # latitude [degrees North; min=-90.0; max=90.0]
  longitude: -93.7                  # longitude [degrees East; min=-360.0; max=360.0]
  elevation: 400                    # elevation of lake surface above sea level [m]
  depth: 2.79                       # maximum water depth [m; min=0.0]
  hypsograph: L303_bathymetry.csv   # hypsograph [csv file]
  init_depth: 2.79                  # initial height of lake surface relative to the bottom [m]
time:
  start: 2019-05-21 00:00:00        # start date and time [yyyy-mm-dd HH:MM:SS]
  stop: 2019-10-06 00:00:00         # stop date and time [yyyy-mm-dd HH:MM:SS]
  time_step: 3600.0                 # time step for integration [s; min=0.0]
config_files:
  GOTM: GOTM/gotm.yaml              # GOTM config file [yaml file]
  GLM: GLM/glm3.nml                 # GLM config file [nml file]
  Simstrat: Simstrat/simstrat.par   # Simstrat config file [json-format file]
  FLake: FLake/flake.nml            # FLake config file [nml file]
  MyLake: MyLake/mylake.Rdata       # MyLake config file [Rdata file]
observations:
  temperature:
    file: L303_wtemp.csv            # file with observed water temperature profiles, with column
      headers according to LakeEnsemblR vocabulary [csv file; if none use NULL or leave empty]
  ice_height:
    file: NULL                      # file with observed ice height, with column headers according to
      LakeEnsemblR vocabulary [csv file; if none use NULL or leave empty]
  water_level:
    file: NULL                      # file with observed water level in meter above bottom of the lake
input:
  init_temp_profile:
    file: NULL                      # initial temperature profile [csv file; if none use NULL or leave
      empty; if empty/NULL, the observations file will be used]
  meteo:
    file: L303_meteo.csv            # file with meteorological forcing data, with column headers
      according to LakeEnsemblR vocabulary [csv file]
  light:
    Kw: L303_kw.csv                 # light extinction coefficient [m-1 or csv file]
  ice:
    use: true                       # turn on ice models? [true/false]
inflows:
```

```
  use: false                        # use in- and outflows? [true/false]
  file: LakeEnsemblR_inflow_standard_round.csv # file with inflow data, with column headers
      according to LakeEnsemblR vocabulary [csv file; must be provided if inflows -> use is true]
  number_inflows: 0                 # number of inflows in the inflow file
outflows:
  use: false                        # use outflows? [true/false]
  file: LakeEnsemblR_inflow_standard_round.csv # file with outflow data, with column headers
      according to LakeEnsemblR vocabulary [csv file; must be provided if outflows -> use is true]
  number_outflows: 0                # number of outflows in the outflow file
  outflow_lvl: -1                   # height of the outflow above the ground. If the outflow is a
      surface outflow use "-1". If there are more than one outflow in the outflow file, this must
      be a list with one value per outflow
output:
  file: L303_ensemble_output        # name of output file, excluding extension
  format: netcdf                    # format [text/netcdf]
  depths: 0.5                       # depths to extract output [m]
  compression: 4                    # set to an integer between 1 (least compression) and 9 (most
      compression), this enables compression for the variable as it is written to the file
  time_unit: hour                   # time unit [second, hour, day]
  time_step: 1                      # number of time units between output [min=1]
  time_method: mean                 # treatment of time dimension [point=instantaneous, mean,
      integrated; only used in GOTM]
  variables:
    - temp
    - ice_height
    - w_level
scaling_factors:                    # scaling factors to apply to meteorological input, either for all
      models or model-specific. If not specified, no scaling is applied. If both "all" and model-
      specific are specified for a certain model, only the model-specific scaling is applied.
  Simstrat:
    wind_speed: 0.82083
    swr: 1.0089
  GOTM:
    wind_speed: 0.78828
    swr: 1.0995
  GLM:
    wind_speed: 1.0058
    swr: 0.99544
  FLake:
    wind_speed: 0.76657
    swr: 0.98684
  MyLake:
    wind_speed: 0.87374
    swr: 0.90042
```

```
model_parameters:                    # Parameters and scaling factors based on calibration in
    Moore et al. (2020), LakeEnsemblR: An R package that facilitates ensemble modelling of
    lakes.
  FLake:                           # FLake specific parameters
    fetch_lk: 700                  # Typical wind fetch [m]
    c_relax_C: 0.016399
  GLM:                             # GLM specific parameters
    bsn_len: 700                   # Length of the lake basin, at crest height [m; default=NULL]
    bsn_wid: 200                   # Width of the lake basin, at crest height [m; default=NULL]
    mixing/coef_mix_hyp: 1.8832
  GOTM:                            # GOTM specific parameters
    k_min: 2.568e-05               # minimum turbulent kinetic energy [m^2/s^2; min=0.0;
      default=1.00000000E-10]
  Simstrat:                        # Simstrat specific parameters
    a_seiche: 0.003208
  MyLake:                          # MyLake specific parameters
    Phys.par/C_shelter: 0.22126
calibration:                       # calibration section
  met:                             # Meteo scaling parameter
    wind_speed:                    # Wind speed scaling
      lower: 0.75                  # lower bound for wind speed scaling
      upper: 1.25                  # upper bound for wind speed scaling
      initial: 1                   # initial value for wind speed scaling
      log: false                   # log transform scaling factor
    swr:                           # shortwave radiation scaling
      lower: 0.9                   # lower bound for shortwave radiation scaling
      upper: 1.1                   # upper bound for shortwave radiation scaling
      initial: 1                   # initial value for shortwave radiation scaling
      log: false                   # log transform scaling factor
  Simstrat:                        # Simstrat specific parameters
    a_seiche:
      lower: 0.0008                # lower bound for parameter
      upper: 0.0036                # upper bound for parameter
      initial: 0.001               # initial value for parameter
      log: false                   # log transform scaling factor
  MyLake:                          # MyLake specific parameters
   Phys.par/C_shelter:
     lower: 0.05                   # lower bound for parameter
     upper: 0.35                   # upper bound for parameter
     initial: 0.15                 # initial value for parameter
     log: false                    # log transform scaling factor
  GOTM:                            # GOTM specific parameters
    turb_param/k_min:
      lower: 1E-6                  # lower bound for parameter
```

```
    upper: 4E-5                  # upper bound for parameter
    initial: 1E-5                # initial value for parameter
    log: true
  GLM:                           # GLM specific parameters
    mixing/coef_mix_hyp:
    lower: 0.1                   # lower bound for parameter
    upper: 2                     # upper bound for parameter
    initial: 1                   # initial value for parameter
    log: false                   # log transform scaling factor
  FLake:                         # FLake specific parameters
    c_relax_C:
    lower: 0.00003               # lower bound for parameter
    upper: 0.3                   # upper bound for parameter
    initial: 0.003               # initial value for parameter
    log: true                    # log transform scaling factor
```

5.2 LakeEnsemblR L304 .yaml File with Model Parameterizations

```
location:
  name: Lake 304                 # name of the lake
  latitude: 49.7                 # latitude [degrees North; min=-90.0; max=90.0]
  longitude: -93.7               # longitude [degrees East; min=-360.0; max=360.0]
  elevation: 400                 # elevation of lake surface above sea level [m]
  depth: 7.2                     # maximum water depth [m; min=0.0]
  hypsograph: L304_bathymetry.csv          # hypsograph [csv file]
  init_depth: 7.2                # initial height of lake surface relative to the bottom [m]
time:
  start: 2019-05-21 00:00:00             # start date and time [yyyy-mm-dd HH:MM:SS]
  stop: 2019-10-06 00:00:00              # stop date and time [yyyy-mm-dd HH:MM:SS]
  time_step: 3600.0              # time step for integration [s; min=0.0]
config_files:
  GOTM: GOTM/gotm.yaml                   # GOTM config file [yaml file]
  GLM: GLM/glm3.nml              # GLM config file [nml file]
  Simstrat: Simstrat/simstrat.par        # Simstrat config file [json-format file]
  FLake: FLake/flake.nml         # FLake config file [nml file]
  MyLake: MyLake/mylake.Rdata            # MyLake config file [Rdata file]
observations:
  temperature:
    file: L304_wtemp.csv         # file with observed water temperature profiles, with column
headers according to LakeEnsemblR vocabulary [csv file; if none use NULL or leave empty]
  ice_height:
```

```
    file: NULL                      # file with observed ice height, with column headers according to LakeEnsemblR vocabulary [csv file; if none use NULL or leave empty]
  water_level:
    file: NULL                      # file with observed water level in meter above bottom of the lake
input:
  init_temp_profile:
    file: NULL                      # initial temperature profile [csv file; if none use NULL or leave empty; if empty/NULL, the observations file will be used]
  meteo:
    file: L303_meteo.csv            # file with meteorological forcing data, with column headers according to LakeEnsemblR vocabulary [csv file]
  light:
    Kw: L304_kw.csv                 # light extinction coefficient [m-1 or csv file]
  ice:
    use: true                       # turn on ice models? [true/false]
inflows:
  use: false                        # use in- and outflows? [true/false]
  file: LakeEnsemblR_inflow_standard_round.csv # file with inflow data, with column headers according to LakeEnsemblR vocabulary [csv file; must be provided if inflows -> use is true]
  number_inflows: 0                 # number of inflows in the inflow file
outflows:
  use: false                        # use outflows? [true/false]
  file: LakeEnsemblR_inflow_standard_round.csv # file with outflow data, with column headers according to LakeEnsemblR vocabulary [csv file; must be provided if outflows -> use is true]
  number_outflows: 0                # number of outflows in the outflow file
  outflow_lvl: -1                   # height of the outflow above the ground. If the outflow is a surface outflow use "-1". If there are more than one outflow in the outflow file, this must be a list with one value per outflow
output:
  file: L304_ensemble_output        # name of output file, excluding extension
  format: netcdf                    # format [text/netcdf]
  depths: 0.5                       # depths to extract output [m]
  compression: 4                    # set to an integer between 1 (least compression) and 9 (most compression), this enables compression for the variable as it is written to the file
  time_unit: hour                   # time unit [second, hour, day]
  time_step: 1                      # number of time units between output [min=1]
  time_method: mean                 # treatment of time dimension [point=instantaneous, mean, integrated; only used in GOTM]
  variables:
    - temp
    - ice_height
    - w_level
```

```
scaling_factors:                    # scaling factors to apply to meteorological input, either for all models or model-specific. If not specified, no scaling is applied. If both "all" and model-specific are specified for a certain model, only the model-specific scaling is applied.
  Simstrat:
    wind_speed: 0.7506
    swr: 0.96035
  GOTM:
    wind_speed: 1.2353
    swr: 0.95362
  GLM:
    wind_speed: 0.98485
    swr: 0.90393
  FLake:
    wind_speed: 1.0645
    swr: 0.99304
  MyLake:
    wind_speed: 1.2398
    swr: 0.90844
model_parameters:                    # Parameters and scaling factors based on calibration in Moore et al. (2020), LakeEnsemblR: An R package that facilitates ensemble modelling of lakes.
  FLake:                           # FLake specific parameters
    fetch_lk: 275                  # Typical wind fetch [m]
    c_relax_C: 0.00029845
  GLM:                             # GLM specific parameters
    bsn_len: 275                   # Length of the lake basin, at crest height [m; default=NULL]
    bsn_wid: 225                   # Width of the lake basin, at crest height [m; default=NULL]
    mixing/coef_mix_hyp: 0.22968
  GOTM:                            # GOTM specific parameters
    k_min: 1.0747e-06              # minimum turbulent kinetic energy [m^2/s^2; min=0.0; default=1.00000000E-10]
  Simstrat:                        # Simstrat specific parameters
    a_seiche: 0.00084626
  MyLake:                          # MyLake specific parameters
    Phys.par/C_shelter: 0.074804
calibration:                       # calibration section
  met:                             # Meteo scaling parameter
    wind_speed:                    # Wind speed scaling
      lower: 0.75                  # lower bound for wind speed scaling
      upper: 1.25                  # upper bound for wind speed scaling
      initial: 1                   # initial value for wind speed scaling
      log: false                   # log transform scaling factor
    swr:                           # shortwave radiation scaling
      lower: 0.9                   # lower bound for shortwave radiation scaling
      upper: 1.1                   # upper bound for shortwave radiation scaling
```

```
    initial: 1                  # initial value for shortwave radiation scaling
    log: false                  # log transform scaling factor
Simstrat:                       # Simstrat specific parameters
  a_seiche:
    lower: 0.0008               # lower bound for parameter
    upper: 0.0036               # upper bound for parameter
    initial: 0.001              # initial value for parameter
    log: false                  # log transform scaling factor
MyLake:                         # MyLake specific parameters
 Phys.par/C_shelter:
   lower: 0.05                  # lower bound for parameter
   upper: 0.35                  # upper bound for parameter
   initial: 0.15                # initial value for parameter
   log: false                   # log transform scaling factor
GOTM:                           # GOTM specific parameters
  turb_param/k_min:
    lower: 1E-6                 # lower bound for parameter
    upper: 4E-5                 # upper bound for parameter
    initial: 1E-5               # initial value for parameter
    log: true
GLM:                            # GLM specific parameters
  mixing/coef_mix_hyp:
    lower: 0.1                  # lower bound for parameter
    upper: 2                    # upper bound for parameter
    initial: 1                  # initial value for parameter
    log: false                  # log transform scaling factor
FLake:                          # FLake specific parameters
  c_relax_C:
    lower: 0.00003              # lower bound for parameter
    upper: 0.3                  # upper bound for parameter
    initial: 0.003              # initial value for parameter
    log: true                   # log transform scaling factor
```

References

Andersen F, Ring P (1999) Comparison of phosphorus release from littoral and profundal sediments in a shallow, eutrophic lake. Biogeochemistry 127:175–183. https://doi.org/10.1023/A:1017027818233

Andersen MR, Kragh T, Martinsen KT, et al (2019) The carbon pump supports high primary production in a shallow lake. Aquat Sci 81:0. https://doi.org/10.1007/s00027-019-0622-7

Austin JA, Colman SM (2007) Lake Superior summer water temperatures are increasing more rapidly than regional temperatures: A positive ice-albedo feedback. Geophys Res Lett 34:1–5. https://doi.org/10.1029/2006GL029021

Blottiere L (2016) The effects of wind-induced mixing on the structure and functioning of shallow freshwater lakes in a context of global change

Breiman L (1996) Bagging Predictors. Mach Learn 24:123–140. https://doi.org/10.1007/BF00058655

Brookes JD, Ganf GG (2001) Variations in the buoyancy response of microcystis aeruginosa to nitrogen, phosphorus and light. J Plankton Res 23:1399–1411. https://doi.org/10.1093/plankt/23.12.1399

Bryant LD, Lorrai C, McGinnis DF, et al (2010) Variable sediment oxygen uptake in response to dynamic forcing. Limnol Oceanogr 55:950–964. https://doi.org/10.4319/lo.2010.55.2.0950

Burchard H, Bolding K, Villarreal MR (1999) GOTM, a general ocean turbulence model. Theory, implementation and test cases. Eur Com Rep EUR 18745 103 pp.

Burger DF, Hamilton DP, Pilditch CA, et al (2005) Sediment phosphorus release during stratification in polymictic Lake Rotorua, New Zealand. SIL Proceedings, 1922-2010 29:811–814. https://doi.org/10.1080/03680770.2005.11902791

Coats R, Perez-Losada J, Schladow G, et al (2006) The warming of Lake Tahoe. Clim Change 76:121–148. https://doi.org/10.1007/s10584-005-9006-1

Crawford GB, Collier RW (1997) Observations of a deep-mixing event in Crater Lake, Oregon. Limnol Oceanogr 42:299–306. https://doi.org/10.4319/lo.1997.42.2.0299

Downing JA, Prairie YT, Cole JJ, et al (2006) Abundance and Size Distribution of Lakes, Ponds and Impoundments. Limnol Oceanogr 51:469–478. https://doi.org/10.1016/B978-012370626-3.00025-9

Einsele W (1936) Über die beziehungen des eisenkreislaufs zum phosphatkreislauf im eutrophen see. Arch für Hydrobiol 664–686

Fang X, Stefan G (1998) Temperature variability in lake sediments. Water Resour Res 34:717–729

Ficker H, Luger M, Gassner H (2017) From dimictic to monomictic: Empirical evidence of thermal regime transitions in three deep alpine lakes in Austria induced by climate change. Freshw Biol 62:1335–1345. https://doi.org/10.1111/fwb.12946

Gal G, Hipsey MR, Parparov A, et al (2009) Implementation of ecological modeling as an effective management and investigation tool: Lake Kinneret as a case study. Ecol Modell 220:1697–1718. https://doi.org/10.1016/j.ecolmodel.2009.04.010

Gorham E, Boyce FM (1989) Influence of Lake Surface Area and Depth Upon Thermal Stratification and the Depth of the Summer Thermocline. J Great Lakes Res 15:233–245. https://doi.org/10.1016/S0380-1330(89)71479-9

Goudsmit G-H, Burchard H, Peeters F, Wüest A (2002) Application of k-ϵ turbulence models to enclosed basins: The role of internal seiches. J Geophys Res Ocean 107:23-1-23–13. https://doi.org/10.1029/2001jc000954

Grützner R (1996) Environmental modeling and simulation — applications and future requirements. Environ Softw Syst 113–122. https://doi.org/10.1007/978-0-387-34951-0_9

Haddout S, Priya KL, Boko M, Azidane H (2018) Comparison of one-dimensional (1-D) column lake models prediction for surface water temperature in eight selected Moroccan lakes. ISH J Hydraul Eng 24:317–329. https://doi.org/10.1080/09715010.2017.1376294

Hall CA., Day JJ. (1977) Ecosystem Modelling in Theory and Practise: An Introduction with Case Histories. John Wiley Sons

Hamilton DP, Salmaso N, Paerl HW (2016) Mitigating harmful cyanobacterial blooms: strategies for control of nitrogen and phosphorus loads. Aquat Ecol 50:351–366. https://doi.org/10.1007/s10452-016-9594-z

Harper SL, Cunsolo A, Babujee A, et al (2021) Trends and gaps in climate change and health research in North America. Environ Res 199:111205. https://doi.org/10.1016/j.envres.2021.111205

Heaney SI, Smyly WJP, Talling JF (1986) Interactions of Physical, Chemical and Biological Processes in Depth and Time within a Productive English Lake during Summer Stratification. Int Rev der gesamten Hydrobiol und Hydrogr 71:441–494. https://doi.org/10.1002/iroh.19860710402

Heiskanen J., Mammarella I, Ojala A, et al (2015) Effects of water clarity on lake stratification and lake-atmosphere heat exchange. J Geophys Res Ocean 120:7412–4828. https://doi.org/10.1002/2014JD022938

Higgins SN, Paterson MJ, Hecky RE, et al (2018) Biological Nitrogen Fixation Prevents the

Response of a Eutrophic Lake to Reduced Loading of Nitrogen: Evidence from a 46-Year Whole-Lake Experiment. Ecosystems 21:1088–1100. https://doi.org/10.1007/s10021-017-0204-2

Hipsey MR, Bruce LC, Boon C, et al (2019) A General Lake Model (GLM 3.0) for linking with high-frequency sensor data from the Global Lake Ecological Observatory Network (GLEON). Geosci Model Dev 12:473–523. https://doi.org/10.5194/gmd-12-473-2019

Holgerson MA, Richardson DC, Roith J, et al (2022) Classifying Mixing Regimes in Ponds and Shallow Lakes. Water Resour Res 58:1–18. https://doi.org/10.1029/2022WR032522

Holgerson MA, Zappa CJ, Raymond PA (2016) Substantial overnight reaeration by convective cooling discovered in pond ecosystems. Geophys Res Lett 43:8044–8051. https://doi.org/10.1002/2016GL070206

Huang A, Lazhu, Wang J, et al (2019) Evaluating and Improving the Performance of Three 1-D Lake Models in a Large Deep Lake of the Central Tibetan Plateau. J Geophys Res Atmos 124:3143–3167. https://doi.org/10.1029/2018JD029610

Hutchinson GE (1957) A Treatise on Limnology, Volume I: Geography, Physics, and Chemistry, 1st edn. Wiley

IISD Experimental Lakes Area (2022) IISD Experimental Lakes Area: Bathymetry Data Package, 1968-2022

Imberger J, Patterson JC (1981) A dynamic reservoir simulation model - DYRESM: 5. ACADEMIC PRESS, INC.

Jabbari A, Ackerman JD, Boegman L, Zhao Y (2019) Episodic hypoxia in the western basin of Lake Erie. Limnol Oceanogr 64:2220–2236. https://doi.org/10.1002/lno.11180

Jensen HS, Andersen FO (1992) Importance of temperature, nitrate, and pH for phosphate release from aerobic sediments of four shallow, eutrophic lakes. Limnol Oceanogr 37:577–589. https://doi.org/10.4319/lo.1992.37.3.0577

Jiang X, Dong S, Liu R, et al (2021) Effects of temperature, dissolved oxygen, and their interaction on the growth performance and condition of rainbow trout (Oncorhynchus mykiss). J Therm Biol 98:102928. https://doi.org/10.1016/j.jtherbio.2021.102928

Jørgensen SE (1995) State of the art of ecological modelling in limnology. Ecol Modell 78:101–115. https://doi.org/10.1016/0304-3800(94)00120-7

Joshi SR, Kukkadapu RK, Burdige DJ, et al (2015) Organic matter remineralization predominates phosphorus cycling in the mid-bay sediments in the chesapeake bay. Environ Sci Technol 49:5887–5896. https://doi.org/10.1021/es5059617

Kalff J (2002) Limnology - Inland Water Ecosystems. Prentice Hall 1:592

Kanoshina I, Lips U, Leppänen JM (2003) The influence of weather conditions (temperature and wind) on cyanobacterial bloom development in the Gulf of Finland (Baltic Sea). Harmful Algae 2:29–41. https://doi.org/10.1016/S1568-9883(02)00085-9

Katsev S, Tsandev I, L'Heureux I, Rancourt DG (2006) Factors controlling long-term phosphorus efflux from lake sediments: Exploratory reactive-transport modeling. Chem Geol 234:127–147. https://doi.org/10.1016/j.chemgeo.2006.05.001

Kirk JT (2010) Light and Photosynthesis in Aquatic Ecosystems. Cambridge Univ Press 3:. https://doi.org/https://doi.org/10.1017/CBO9781139168212

Klug JL, Richardson DC, Ewing HA, et al (2012) Ecosystem effects of a tropical cyclone on a network of lakes in northeastern North America. Environ Sci Technol 46:11693–11701. https://doi.org/10.1021/es302063v

Kramer DL (1987) Dissolved oxygen and fish behavior. Environ Biol Fishes 18:81–92. https://doi.org/10.1007/BF00002597

Kraus RT, Cook HA, Faust MD, et al (2023) Habitat selection of a migratory freshwater fish in response to seasonal hypoxia as revealed by acoustic telemetry. J Great Lakes Res 49:1004–1014. https://doi.org/10.1016/j.glr.2023.01.004

Kuhn M (2022) caret: Classification and Regression Training. https://cran.r-project.org/package=caret

Kumagai M, Nakano SI, Jiao C, et al (2000) Effect of cyanobacterial blooms on thermal stratification. Limnology 1:191–195. https://doi.org/10.1007/s102010070006

Laenen A, Letourneau AP (1996) Upper Klamath Basin Nutrient-Loading Study — Estimate of Wind-Induced Resuspension of Bed Sediment During Periods of Low Lake Elevation. US Geol Surv 1–12

Levine SN, Schindler DW (1989) Phosphorus, nitrogen, and carbon dynamics of experimental Lake 303 during recovery from eutrophication. Can J Fish Aquat Sci 46:2–10. https://doi.org/10.1139/f89-001

Li W, Joshi SR, Hou G, et al (2015) Characterizing phosphorus speciation of Chesapeake Bay sediments using chemical extraction, 31P NMR, and X-ray absorption fine structure spectroscopy. Environ Sci Technol 49:203–211. https://doi.org/10.1021/es504648d

Liaw A, Wiener M (2002) The R Journal: Classification and regression by randomForest. R J 2:18–22

Liu M, Zhang Y, Shi K, et al (2020) Effects of rainfall on thermal stratification and dissolved oxygen

in a deep drinking water reservoir. Hydrol Process 34:3387–3399. https://doi.org/10.1002/hyp.13826

Liu M, Zhang Y, Shi K, et al (2019) Thermal stratification dynamics in a large and deep subtropical reservoir revealed by high-frequency buoy data. Sci Total Environ 651:614–624. https://doi.org/10.1016/j.scitotenv.2018.09.215

Loewen MR, Ackerman JD, Hamblin PF (2007) Environmental implications of stratification and turbulent mixing in a shallow lake basin. Can J Fish Aquat Sci 64:43–57. https://doi.org/10.1139/F06-165

Long Z, Perrie W, Gyakum J, et al (2007) Northern lake impacts on local seasonal climate. J Hydrometeorol 8:881–896. https://doi.org/10.1175/JHM591.1

Maberly SC, O'Donnell RA, Woolway RI, et al (2020) Global lake thermal regions shift under climate change. Nat Commun 11:1–9. https://doi.org/10.1038/s41467-020-15108-z

Mazumder A, Taylor WD (1994) Thermal structure of lakes varying in size and water clarity. Limnol Oceanogr 39:968–976. https://doi.org/10.4319/lo.1994.39.4.0968

McCullough GK, Campbell P (1993) Lake variation and climatic change study: ELA lakes 1986-1990 II. Watershed geography and lake morphology. Can Tech Rep Fish Aquat Sci 1898 iv:29

McEnroe NA, Buttle JM, Marsalek J, et al (2013) Thermal and chemical stratification of urban ponds: Are they "completely mixed reactors"? Urban Ecosyst 16:327–339. https://doi.org/10.1007/s11252-012-0258-z

McIlwraith HK, Dias M, Orihel DM, et al (2024) A Multicompartment Assessment of Microplastic Contamination in Semi-remote Boreal Lakes. Environ Toxicol Chem 43:999–1011. https://doi.org/10.1002/etc.5832

Meerhoff M, Jeppesen E (2009) Shallow Lakes and Ponds. Encycl Inl Waters 645–655. https://doi.org/10.1016/B978-012370626-3.00041-7

Mejbel HS, Irwin CL, Dodsworth W, et al (2023) Long-term cyanobacterial dynamics from lake sediment DNA in relation to experimental eutrophication, acidification and climate change. Freshw Biol 68:1875–1893. https://doi.org/10.1111/fwb.14074

Menze BH, Kelm BM, Masuch R, et al (2009) A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. BMC Bioinformatics 10:1–16. https://doi.org/10.1186/1471-2105-10-213

Michalski J, Lemmin U (1995) Dynamics of vertical mixing in the hypolimnion of a deep lake: Lake Geneva. Limnol Oceanogr 40:809–816. https://doi.org/10.4319/lo.1995.40.4.0809

Mironov D (2008) Parameterization of Lakes in Numerical Weather Prediction. Description of a

Lake Model

Molot LA, Higgins SN, Schiff SL, et al (2021) Phosphorus-only fertilization rapidly initiates large nitrogen-fixing cyanobacteria blooms in two oligotrophic lakes. Environ Res Lett 16:. https://doi.org/10.1088/1748-9326/ac0564

Molot LA, Watson SB, Creed IF, et al (2014) A novel model for cyanobacteria bloom formation: The critical role of anoxia and ferrous iron. Freshw Biol 59:1323–1340. https://doi.org/10.1111/fwb.12334

Moore TN, Mesman JP, Ladwig R, et al (2021) LakeEnsemblR: An R package that facilitates ensemble modelling of lakes. Environ Model Softw 143:105101. https://doi.org/10.1016/j.envsoft.2021.105101

Mortimer CH (1942) The Exchange of Dissolved Substances between Mud and Water in Lakes. J Ecol 30:147. https://doi.org/10.2307/2256691

Murphy TP, Lean DRS, Nalewajko C (1976) Blue-green algae: Their excretion of iron-selective chelators enables them to dominate other algae. Science (80-) 192:900–902. https://doi.org/10.1126/science.818707

Northington RM, Saros JE, Burpee BT, McCue J (2019) Changes in mixing depth reduce phytoplankton biomass in an Arctic lake: Results from a whole-lake experiment. Arctic, Antarct Alp Res 51:533–548. https://doi.org/10.1080/15230430.2019.1692412

O'Reilly CM, Alin SR, Piisnier PD, et al (2003) Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa. Nature 424:766–768. https://doi.org/10.1038/nature01833

O'Reilly CM, Rowley RJ, Schneider P, et al (2015) Rapid and highly variable warming of lake surface waters around the globe. Geophysical Research Letters, 42: 1-9. Geophys Res Lett 1–9. https://doi.org/10.1002/2015GL066235.Received

Orihel DM, Baulch HM, Casson NJ, et al (2017) Internal phosphorus loading in canadian fresh waters: A critical review and data analysis. Can J Fish Aquat Sci 74:2005–2029. https://doi.org/10.1139/cjfas-2016-0500

Orihel DM, Schindler DW, Ballard NC, et al (2015) The "nutrient pump:" Iron-poor sediments fuel low nitrogen-to-phosphorus ratios and cyanobacterial blooms in polymictic lakes. Limnol Oceanogr 60:856–871. https://doi.org/10.1002/lno.10076

Perroud M, Goyette S, Martynov A, et al (2009) Simulation of multiannual thermal profiles in deep Lake Geneva: A comparison of one-dimensional lake models. Limnol Oceanogr 54:1574–1594. https://doi.org/10.4319/lo.2009.54.5.1574

Peters A, Hothorn T (2022) ipred: Improved Predictors. https://cran.r-project.org/package=ipred

R Core Team (2022) R: A language and environment for statistical computing. https://www.r-project.org/

Ramm K, Scheps V (1997) Phosphorus balance of a polytrophic shallow lake with the consideration of phosphorus release. Hydrobiologia 342–343:43–53. https://doi.org/10.1007/978-94-011-5648-6_5

Reinl KL, Brookes JD, Carey CC, et al (2021) Cyanobacterial blooms in oligotrophic lakes: Shifting the high-nutrient paradigm. Freshw Biol 66:1846–1859. https://doi.org/10.1111/fwb.13791

Riis T, Sand-Jensen K (1998) Development of vegetation and environmental conditions in an oligotrophic Danish lake over 40 years. Freshw Biol 40:123–134. https://doi.org/10.1046/j.1365-2427.1998.00338.x

Salonen K, Kankaala P, Tulonen T, et al (1992) Planktonic food chains of a highly humic lake. Hydrobiologia 229:143–157. https://doi.org/10.1007/bf00006997

Saloranta TM, Andersen T (2007) MyLake-A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. Ecol Modell 207:45–60. https://doi.org/10.1016/j.ecolmodel.2007.03.018

Santos R, Saggio A, Silva T, et al (2015) Short-term thermal stratification and partial overturning events in a warm polymictic reservoir: effects on distribution of phytoplankton community. Brazilian J Biol 75:19–29. https://doi.org/10.1590/1519-6984.05313

Schindler DW, Bayley SE, Parker BR, et al (1996) The effects of climatic warming on the properties of boreal lakes and streams at the Experimental Lakes Area, northwestern Ontario. Limnol Oceanogr 41:1004–1017. https://doi.org/10.4319/lo.1996.41.5.1004

Schmid M, Hunziker S, Wüest A (2014) Lake surface temperatures in a changing climate: A global sensitivity analysis. Clim Change 124:301–315. https://doi.org/10.1007/s10584-014-1087-2

Schmolke A, Thorbek P, DeAngelis DL, Grimm V (2010) Ecological models supporting environmental decision making: A strategy for the future. Trends Ecol Evol 25:479–486. https://doi.org/10.1016/j.tree.2010.05.001

Schneider P, Hook SJ (2010) Space observations of inland water bodies show rapid surface warming since 1985. Geophys Res Lett 37:1–5. https://doi.org/10.1029/2010GL045059

Shatwell T, Adrian R, Kirillin G (2016) Planktonic events may cause polymictic-dimictic regime shifts in temperate lakes. Sci Rep 6:1–14. https://doi.org/10.1038/srep24361

Søndergaard M, Jensen J, Jeppesen E (2003) Role of sediment and internal loading of phosphorus in shallow lakes. Hydrobiologia 506:135–145

Soulignac F, Vinçon-Leite B, Lemaire BJ, et al (2017) Performance Assessment of a 3D

Hydrodynamic Model Using High Temporal Resolution Measurements in a Shallow Urban Lake. Environ Model Assess 22:309–322. https://doi.org/10.1007/s10666-017-9548-4

Stepanenko V, Jöhnk KD, Machulskaya E, et al (2014) Simulation of surface energy fluxes and stratification of a small boreal lake by a set of one-dimensional models. Tellus, Ser A Dyn Meteorol Oceanogr 66:. https://doi.org/10.3402/tellusa.v66.21389

Stepanenko VM, Martynov A, Jöhnk KD, et al (2013) A one-dimensional model intercomparison study of thermal regime of a shallow, turbid midlatitude lake. Geosci Model Dev 6:1337–1352. https://doi.org/10.5194/gmd-6-1337-2013

Sun L, Liang XZ, Ling T, et al (2020) Improving a Multilevel Turbulence Closure Model for a Shallow Lake in Comparison With Other 1-D Models. J Adv Model Earth Syst 12:1–23. https://doi.org/10.1029/2019MS001971

Tanentzap AJ, Hamilton DP, Yan ND (2007) Calibrating the Dynamic Reservoir Simulation Model (DYRESM) and filling required data gaps for one-dimensional thermal profile predictions in a boreal lake. Limnol Oceanogr Methods 5:484–494. https://doi.org/10.4319/lom.2007.5.484

Therneau T, Atkinson B (2022) rpart: Recursive Partitioning and Regressiong Trees. https://cran.r-project.org/package=rpart

Trolle D, Hamilton DP, Hipsey MR, et al (2012) A community-based framework for aquatic ecosystem models. Hydrobiologia 683:25–34. https://doi.org/10.1007/s10750-011-0957-0

Tuan N V., Hamagami K, Mori K, Hirai Y (2009) Mixing by wind-induced flow and thermal convection in a small, shallow and stratified lake. Paddy Water Environ 7:83–93. https://doi.org/10.1007/s10333-009-0158-x

Wagner C, Adrian R (2009) Cyanobacteria dominance: Quantifying the effects of climate change. Limnol Oceanogr 54:2460–2468. https://doi.org/10.4319/lo.2009.54.6_part_2.2460

Wantzen KM, Junk WJ, Rothhaupt KO (2008) An extension of the floodpulse concept (FPC) for lakes. Hydrobiologia 613:151–170. https://doi.org/10.1007/s10750-008-9480-3

Wetzel RG (2001) Limnology - Lake and River Ecosystems. Acad Press 1:1006

Wilhelm S, Adrian R (2008) Impact of summer warming on the thermal characteristics of a polymictic lake and consequences for oxygen, nutrients and phytoplankton. Freshw Biol 53:226–237. https://doi.org/10.1111/j.1365-2427.2007.01887.x

Williamson CE, Neale PJ, Hylander S, et al (2019) The interactive effects of stratospheric ozone depletion, UV radiation, and climate change on aquatic ecosystems. Photochem Photobiol Sci 18:717–746. https://doi.org/10.1039/C8PP90062K

Woolway RI, Kraemer BM, Lenters JD, et al (2020) Global lake responses to climate change. Nat Rev Earth Environ 1:388–403. https://doi.org/10.1038/s43017-020-0067-5

Woolway RI, Meinson P, Nõges P, et al (2017) Atmospheric stilling leads to prolonged thermal stratification in a large shallow polymictic lake. Clim Change 141:759–773. https://doi.org/10.1007/s10584-017-1909-0

Wüest A, Lorke A (2003) Small-scale hydrodynamics in lakes. Annu Rev Fluid Mech 35:373–412. https://doi.org/10.1146/annurev.fluid.35.101101.161220

Yang Y, Wang Y, Zhang Z, et al (2018) Diurnal and Seasonal Variations of Thermal Stratification and Vertical Mixing in a Shallow Fresh Water Lake. J Meteorol Res 32:219–232. https://doi.org/10.1007/s13351-018-7099-5

Yindong T, Xiwen X, Miao Q, et al (2021) Lake warming intensifies the seasonal pattern of internal nutrient cycling in the eutrophic lake and potential impacts on algal blooms. Water Res 188:116570. https://doi.org/10.1016/j.watres.2020.116570

Yu SJ, Ryu IG, Park MJ, Im JK (2021) Long-term relationship between air and water temperatures in Lake Paldang, South Korea. Environ Eng Res 26:1–12. https://doi.org/10.4491/eer.2020.177

Zhang M, Zhang Y, Deng J, et al (2022) High-resolution temporal detection of cyanobacterial blooms in a deep and oligotrophic lake by high-frequency buoy data. Environ Res 203:111848. https://doi.org/10.1016/j.envres.2021.111848

Zhao L, Zhu G, Chen Y, et al (2011) Thermal stratification and its influence factors in a large-sized and shallow Lake Taihu. Adv Water Sci 22:844–860

Zhao Q, Sun J, Zhu G (2012) Simulation and exploration of the mechanisms underlying the spatiotemporal distribution of surface mixed layer depth in a large shallow lake. Adv Atmos Sci 29:1360–1373. https://doi.org/10.1007/s00376-012-1262-1

Znachor P, Zapomělová E, Řeháková K, et al (2008) The effect of extreme rainfall on summer succession and vertical distribution of phytoplankton in a lacustrine part of a eutrophic reservoir. Aquat Sci 70:77–86. https://doi.org/10.1007/s00027-007-7033-x