Spring 5-6-2022

Enhanced Condition Assessment for Maine Lakes

Jeremy Deeds
jeremy.deeds@maine.edu

ENHANCED CONDITION ASSESSMENTS FOR MAINE LAKES

By

Jeremy Deeds

B.A. University of Maine at Farmington, 2001

M.S. Kent State University, 2003

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Ecology and Environmental Science)

The Graduate School

The University of Maine

May 2022

Advisory Committee:

Aria Amirbahman, Professor of Civil and Environmental Engineering, Co-Advisor

Stephen Norton, Professor of Earth and Climate Sciences, Co-advisor

David Courtemanch, The Nature Conservancy

Rachel Hovel, Assistant Professor of Biology

Shaleen Jain, Civil and Environmental Engineering

ENHANCED CONDITION ASSESSMENTS FOR MAINE LAKES

By

Jeremy Deeds

Dissertation Co-Advisors: Dr. Aria Amirbahman and Dr. Stephen Norton

An Abstract of the Dissertation presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Ecology and Environmental Science)
May 2022

The Influence of anthropogenic activities on lake water quality is well documented, but how those influences interact with the effects of natural features, such as watershed geology or lake morphometry, has been less explored. Further, some aspects of lake condition are influenced by factors that are not lake or watershed specific, but occur across large regions, such as weather patterns. All these factors may be interrelated in some instances, which can complicate lake condition assessments which have the purpose of determining how lakes are being affected by human activities. This dissertation investigates how lake assessments can integrate the interactions among natural features of lakes, their watersheds, and anthropogenic influences. Chapter 1 discusses the variety of factors that may affect lake condition and how those influences may confound lake condition assessments. Chapter 2 details the creation of a hydrogeomorphic lake classification, based on ecoregions and lake depth, that partitioned lakes into groups that share similarities in background water quality condition. In chapter 3, a logistic regression model is described that uses maximum depth and relative lake area beneath the epilimnion to predict which low-nutrient lakes (total phosphorus < 15 µg/L) may exhibit naturally-occurring anoxia. In chapter 4, water clarity patterns from different types of reference lakes (detailed in chapter 2) were modeled to allow for comparisons between yearly water clarity values in non-reference

lakes and a reference baseline that shifts over time. Cumulative precipitation during the lake stratification season was the primary driver of yearly differences in background lake water clarity. In chapter 5, methods were developed to measure the effect of anthropogenic shoreland disturbance on the condition of littoral habitat. Multi-metric indices based on various habitat measures were established that determine if the littoral habitat is different from a natural reference condition. Chapter 6 summarizes the research in this dissertation and offers potential foci of future lake research in Maine. The overall goal of this dissertation was to advance our collective understanding of how lakes may be variably affected by natural and anthropogenic factors, thereby allowing for better-informed lake assessments and the development of more comprehensive, achievable lake management goals. The research presented herein underscores the importance of considering the interactions of multiple cross-scale factors when evaluating lake condition, especially those related to landscape traits that influence runoff water chemistry, natural lake-specific features such as basin morphometry, large-scale weather patterns, and localized shoreland development.

DEDICATION

Dedicated to those who study, protect, or enjoy the lakes of Maine.

ACKNOWLEDGEMENTS

I am grateful for the funding support of NSF grant 1743082 (to A. Amirbahman) and the Maine Department of Environmental Protection (DEP). Much of the data used in this research were collected by more than 1,000 volunteer and professional lake monitors, largely coordinated through the Lake Stewards of Maine. I am indebted to current and previous staff at the Maine DEP who have maintained or supported a successful lake assessment program since the early 1970s, work which includes collection and management of the lake monitoring data used here. These people include Linda Bacon, Roy Bouchard, David Boucher, Jeffrey Dennis, Dr. David Courtemanch, and Matthew Scott.

Chapters 2, 3, and 4 were evaluated by several anonymous reviewers during pre-publication review processes who offered excellent comments and questions that improved the final versions. Gertrud Nürnberg provided highly insightful comments and questions for improving the final manuscript of chapter 3. Analyses in chapter 3 were enhanced by Dr. Daniel Buckley, University of Maine at Farmington, who collected high-resolution lake bathymetry data that were used in lake volume calculations. Data were supplied for chapter 5 by Kellie Merrell and Dr. Leslie Matthews of the Vermont Agency of Natural Resources, and Kirsten Hugger of the New Hampshire Department of Environmental Services. National Lake Assessment data, also used in chapter 5, was provided by the US EPA National Aquatic Resource Surveys group and I thank the field crews that collected those data in Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. Many Maine DEP staff assisted with the littoral habitat surveys used in Chapter 5, including Douglas Suitor, Joshua Noll, John McPhedran, Denise Blanchette, Karen Hahnel, Emily Zimmerman, Dr. Thomas Danielson, and several seasonal assistants. Dr. Thomas Danielson provided R script for the logistic regression plots used in chapter 3 and additional script that helped with linear discriminant model selection in chapter 5. I am grateful to Linda Bacon of Maine DEP for her support in my pursuit of this dissertation and for providing careful reviews of manuscripts that improved the final product.

I thank my advisory committee, Dr. David Courtemanch, Dr. Rachel Hovel, and Dr. Shaleen Jain for their involvement and thoughtful contributions to this process. I am especially grateful for my co-advisors, Dr. Aria Amirbahman and Dr. Stephen Norton, for the passion they have for their work, for believing that I could do this, and for providing excellent guidance every step of the way.

I am forever thankful for the support of my family in all my endeavors. My parents, Ann and John Deeds, encouraged my curiosity and instilled a sense of wonder in me about Maine's natural world from a very young age. I am extraordinarily lucky to have the love and support of my wife Shana and son Ethan, who are always up for exploring the world around us, researching interesting questions, and enjoying conversational science around our dinner table.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Lake Assessment in Maine

Lake assessments are completed in order to determine the condition of lakes, either for a single point in time or to evaluate whether lake condition is stable or changing through time. Lake assessments commonly focus on parameters related to the trophic condition and water chemistry of lakes. Trophic condition measures relate to how much biological productivity is occurring in the lake and may be used to determine if a lake is likely to experience a potentially harmful algae bloom. Trophic condition is generally assessed with water samples that are tested for chlorophyll-*a* (the pigment found in algae cells) or phosphorus (the nutrient that generally limits algae growth in northeast U.S. lakes), or with Secchi Disk Transparency (SDT, a measure of lake clarity that correlates to both chlorophyll-*a* and phosphorus) readings. Lake water chemistry is assessed to determine the status of other factors that relate to various lake dynamics, such as water temperature, dissolved oxygen, acidity, and dissolved organic carbon. Understanding how these various parameters are influenced by features of the lake's watershed, weather patterns, and local anthropogenic factors is necessary for interpreting lake data accurately and making sound lake management decisions. Regular monitoring of lake condition, throughout the yearly ice-off period and over multiple years, contributes greatly to an understanding of lake processes and how lake condition is affected by the surrounding environment.

The earliest records of lake monitoring data collected in Maine are from 1938; systematic monitoring of Maine's lakes started in earnest in 1970 by the Maine Department of Environmental Protection (DEP; ME DEP 2015). Since then, >1,000 volunteer and professional scientists have actively monitored Maine's lakes, providing data necessary for lake condition assessments. Monitoring methodology has been largely consistent during this time because of coordinated training and program management efforts by the Maine DEP and the Lake Stewards of Maine. As a result, current lake researchers now have a unique opportunity to examine decades of lake monitoring data for purposes beyond general condition assessments and trend analysis. There is potential to learn more about how Maine's lakes function in general, and to develop new tools that will inform future lake assessment efforts. By combining lake monitoring data with other datasets, such as land cover, geological information, and lake morphology, it is possible to learn more about how lakes respond to variability in their surrounding landscapes. This is especially relevant when trying to determine which aspects of lake condition may be attributable to natural watershed features (e.g., bedrock and surficial geology), anthropogenic factors (e.g., conversion of riparian forest to residential development), or lake morphology (e.g., maximum depth).

The landscape of Maine is considerably diverse both in patterns of human settlement and features of the natural landscape. The state has a land area of 91,647 km^2 and its elevation ranges from sea level to 1,608 m. Land use in the state is highly variable, ranging from relatively unpopulated forest (northwestern quadrant), to common dairy and cropland agriculture (<100 km from coast and the northeast), to urban and suburban development (largely southern and coastal). All the state was glaciated, but land currently below ~128 m above present sea level contains nutrient-rich marine-derived clay, which has implications for both agricultural activity and lake productivity. Consequently, the diversity of the Maine landscape has variable effects on lakes in the state, depending upon where the lakes are geographically located and how much human alteration has occurred in their watersheds. These interacting factors can confound interpretations of lake monitoring data, as many interacting natural and anthropogenic factors must be considered for lake data to be evaluated in the proper context.

The mission of the Lakes Assessment Section of the DEP is to monitor and assess the status of Maine lakes under the auspices of the federal clean water act (33 U.S.C §§ 1251 *et seq*) and the Maine water quality statute for lakes (MRS 38, §465-A). It is also charged with making informed assessments of lake condition and helping resource managers make sound decisions to support lake health and rehabilitation. These tasks become increasingly difficult without a clear understanding of the comparative influence between naturally-occurring phenomena and human-induced changes in lakes and their watersheds. To address these knowledge gaps, we performed four research studies: three that leveraged decades of lake monitoring data available in the Maine lakes database, and one that developed new methods to evaluate an under-studied but important component of lake condition. The overall goal of this dissertation is to advance the collective understanding of Maine lakes in general and to provide tools for more comprehensive condition assessments. These studies focus on:

1. The effects of natural features of *landscapes and watersheds* on lake condition, and the interaction between natural features and human alteration in the watershed;
2. The role of *lake morphometry* in the development of seasonal hypolimnetic anoxia;
3. Differentiating between *regional vs. local effects* on lake water clarity trends; and
4. How shoreland disturbance affects the natural condition of *littoral habitat*.

Interpreting Lake Monitoring Data: Natural vs. Anthropogenic Influences

Maine lakes occur across gradients in natural characteristics of their landscapes such as surficial and bedrock geology, soil types, topography, watershed hydrology, proximity to the ocean, atmospheric deposition, climate trends, and elevation. These natural features of lake watersheds influence the biological condition of lakes through nutrient availability, organic matter processing, primary production, and the shaping of chemical environments (Omernik 1977, Prairie and Kalff 1986, Norton et al. 1989, D'Arcy and Carignan 1997, DeVito et al. 2000, Soranno et al. 2015).

Excess phosphorus (P, the limiting nutrient in most temperate lakes; Wetzel 2001) that causes accelerated eutrophication is the principal cause of impairment that affects Maine lakes, either directly or indirectly (Maine DEP 2018). Phosphorus can enter stormwater runoff and reach lakes via many pathways including soil erosion, shoreland development, excess fertilizer application (which may also add $N0_3$ or NH_4, two other limiting nutrients), untreated storm water, and inadequately treated sewage or industrial effluent in the watershed. Various land use practices can produce variable rates of P export into lakes and their tributary waters (Dennis 1986, Petrone 2010, Soranno et al. 2015, Chen et al. 2016). Extra P in lakes generally leads to declines in water clarity and the promotion of nuisance algal growth. These detrimental effects may lead to further water quality problems such as lake anoxia, habitat degradation, diminished recreational opportunities, threatened drinking water supplies, and reductions in property values (Michael et al. 1996).

While the occurrence of excess P in Maine lakes is most commonly exacerbated by human activities, the bedrock and surficial material composition of watersheds, susceptibility of watershed soils to physical or chemical weathering, and concentration of dissolved organic carbon (DOC) play integrated roles in the P concentration and resulting trophic condition of lakes. Lake morphometry (e.g., surface area, maximum depth) can affect the dynamics of nutrient processing as a function of residence time and thermal resistance to mixing. This may be relevant to whether a lake experiences seasonal anoxia in the hypolimnion and, consequently, whether sediment P is released because of that anoxia. The effects of climate change (warmer air temperatures, more frequent and heavier precipitation events; Fernandez et al. 2020) on Maine lakes may differentially affect lakes with various morphological traits, especially as morphology relates to lake volume and residence time. The effects of climate change may also influence lake variably in different areas of Maine due to landscape-scale watershed characteristics, potentially related to the erodibility and nutrient content of watershed soils, later winter ice-on, or earlier spring ice-off. Secchi depth transparency (SDT) trends over time, which is the primary evaluation method that ME DEP uses for water quality statute attainment for Maine lakes, can reveal temporal changes in lake clarity that are not necessarily specific to an individual lake, such as change attributed to regional precipitation patterns or changes in air temperature (Schindler et al. 1996, Read and Rose 2013, Rose et al. 2017).

Lakes may be variably affected by other anthropogenic influences besides P concentration. Road salt (NaCl) is applied throughout the northeast US for winter ice management, and its use is logically more concentrated in urban areas (Kaushal et al. 2005). Runoff containing excess road salt can salinize lakes (Dugan et al. 2017), but road salt in watershed soils exchanges Ca and Mg for Na, thereby increasing the export of Ca and Mg into lakes (Sutherland et al. 2018). While Cl can be toxic to aquatic life at high concentrations (Coldsnow and Relyea 2018) and change plankton community composition at sublethal concentrations (Greco et al. 2021, Hintz et al. 2022), increases in lake Ca can create suitable habitat for infestations of exotic invasive bivalves that require higher Ca levels for shell creation (Davis et al. 2015). Application of road salts can also affect the recovery of lakes and soil from the effects of acid rain; the increase in available Na from road salts increases the export of Ca and Mg from soils, despite declines in atmospheric SO_4 deposition which also extracts Ca and Mg (Rosfjord et al. 2007).

The conversion of lake shorelands from natural vegetation to residential development is often cited as a contributing factor to cultural eutrophication of lakes because of increased land erosion and loss of nutrient retention capacity of riparian vegetation, whether development is considered in the watershed (Dennis 1986), along riparian corridors (Soranno et al. 1996), or as development density along lakeshores (Garrison and Wakeman 2000, Garrison et al. 2010). However, recent research has highlighted the detrimental effects of lake shoreland disturbance specifically on littoral habitat, specifically in the northeast US (USEPA 2009, 2016; Kaufmann et al. 2014a).

The following topics have been developed in Chapters 2, 3, 4, and 5:

Landscapes and Watersheds, and their Effect on Lake Condition

The mechanisms of watershed P mobilization and bioavailability in Maine lakes vary across geological and hydrological settings. Dillon and Kirchner (1975) found that P export from watersheds differed greatly between igneous and sedimentary geology-dominated watersheds. This was evident in both forested watersheds and agricultural watersheds, with sedimentary watersheds exporting higher and greater ranges of P concentrations. Precipitation patterns, which can vary widely across Maine (USDA 2012a), are linked to stormwater runoff that has direct association with P loading to surface waters (Soranno et al 1996). P export is also related to catchment size (Prairie and Kalff 1986) and topography (D'Arcy and Carignan 1997). The potential for lake P loading is closely related to the hydrologic setting; DeVito et al (2000) showed that lakes with close groundwater connections were likely to have the greatest increase in lake P concentration and that catchment wetlands can be effective at mitigating P export into lakes.

Reinhardt et al. (2004) found that increases in stream water total P were coincident with increased discharge and depressed pH. The total P in stream water was strongly associated with P adsorbed to particulate Al and Fe, with the Al and P entering the stream in dissolved form from the more acidic forest soils. Al and P were dissolved in the acidic soil, mobilized Al was precipitated in the stream, and then mobilized P was adsorbed to the Al. At circum-neutral pH, increases in P export from a Maine watershed with eroding agricultural soils rich in adsorbed P encountered higher pH in the stream environment where groundwater (with a pH in the low 7 range) degassed CO_2 in the stream, raising the pH to low- to mid-8s, and desorbing P from stream sediment (McDonald et al. 2019). The adsorption capacity of soil for dissolved P increases at lower pH and decreases at higher pH.

The condition of some Maine lakes is influenced by the presence of Presumpscot Formation soils in their watershed. These marine-derived sediments were deposited on now emergent land when ocean water inundated the area after deglaciation, up to what is currently 128 m above Maine's current coastline, between about ca. 15,000 - 14,000 years BP (referred to as the marine limit; Thompson and Borns 1985). These sediments presently occur in low relief areas and are easily eroded. Presumpscot soils contain the mineral apatite ($Ca_5(PO_4)_3(OH)$); apatite from any rock type or surficial material is the principal natural source of P in lakes (Norton et al. 2019). In Maine coastal regions, agricultural land and population centers commonly co-occur with the Presumpscot Formation. Furthermore, lake P concentration in Maine is positively correlated with the presence of Presumpscot soils, wetlands, agriculture, and human disturbance in watersheds (Nieratko 1992, Doolittle 2018).

While it is not feasible to conduct geochemical studies of each lake and its watershed in Maine, it is possible to focus on a few landscape attributes that influence the trophic condition of lakes because of similarities in hydrogeological settings. This approach is useful for aggregating lakes into groups that may have similar baseline water conditions, and also similar responses to anthropogenic disturbance in the watershed. This approach also facilitates the designation of minimally-disturbed reference lakes, which provide suitable comparisons for the expected conditions of different types of lakes (Stoddard et al. 2006, Herlihy et al. 2013).

Lake Morphometry and its Effect on Anoxia

Lake morphometry can affect various aspects of lake P metabolism (Taranu and Gregory-Eaves 2008), and one important connection may exist between lake morphometry and the occurrence of seasonal anoxia in the hypolimnion. Dissolved oxygen (DO) depletion can occur in the hypolimnion of lakes when the respiration rate, from the aerobic decomposition of organic matter, exceeds the rate of DO replenishment from photosynthetic activity or lake mixing. This can diminish or eliminate habitat for oxygen respiring organisms and cause the release of P from lake sediment into water via the reduction of iron hydroxide (Wetzel 2001). In addition to morphometry, DOC concentrations can influence the stability of lake stratification and associated DO; lakes with elevated DOC concentration can be more strongly stratified, which increases hypolimnetic DO depletion caused by greater isolation of bottom water, enhanced decomposition of organic matter, and heightened DOC photo-oxidation rates (Hutchinson 1957, Knoll et al. 2018).

Lake hypolimnetic anoxia (DO < 2 mg/L) is generally intensified in lakes with high algal growth and organic detrital input, both of which contribute to high nutrient availability. However, lakes in watersheds with minimal human development may exhibit seasonal anoxia because of morphometry (Molot et al. 1992), acidic DOC exported from watershed wetlands (Nürnberg 2004), or both. Many lakes in Maine experience seasonal anoxia despite minimal nutrient enrichment. Seasonal anoxia has been observed in 647 of 951 surveyed Maine lakes during peak summer stratification (1 August – 7 September), despite relatively low epilimnetic TP concentrations in most Maine lakes during this period (median = 9.0 µg/L, 25th percentile = 6.5 µg/L, 75th percentile = 13.0 µg/L; ME DEP 2015). Thus, some Maine lakes may experience hypolimnetic anoxia during peak stratification due to factors unrelated to enhanced nutrient inputs from anthropogenic sources.

The occurrence or absence of seasonal anoxia in lakes influences lake conditions year-round. For accurate interpretations of lake condition and the development of sound management objectives, it is important to determine if a lake is exhibiting anoxia because of natural features of the lake or if the anoxia is caused or exacerbated from excess nutrients originating from anthropogenic activities in the watershed.

Regional vs. Local Effects on Lake Condition

Water clarity, as measured with Secchi disk transparency (SDT), is a ubiquitously used and inexpensive water quality metric. The measurement is simple to collect, such that it can be completed with minimal expense and training, and the metric correlates to many other lake variables. In most temperate lakes, SDT correlates to trophic measures such as total P and chlorophyll-*a*, as well as suspended sediment and Colored Dissolved Organic Matter (CDOM) (Brezonik et al. 2019). SDT is highly useful for long-term lake monitoring because it provides a measure reflecting lake and watershed disturbance at the local scale, such as changes in nutrient additions, watershed erosion, and resuspension of shallow lake sediment (Soranno et al. 2015, Gunn et al. 2001). Because of the robust utility of information generated and the ability of data to be collected frequently by trained volunteers, SDT trends are the primary evaluation tool in water quality statute attainment decisions for Maine lakes (ME DEP 2015). However, SDT temporal trends can also reflect changes in factors affecting lake condition that are not specific to an individual lake, such as regional precipitation patterns or air temperature (Schindler et al. 1996, Read and Rose 2013, Rose et al. 2017).

A declining SDT temporal trend may be falsely attributed to local factors in cases where a lake is only showing a regional response to weather patterns. Alternatively, SDT trends showing stable or even decreasing water clarity may not reveal regional trends caused by positive influences on water clarity because local factors, such as watershed erosion, may overpower the regional pattern (Rose et al. 2017). As factors affecting lake water clarity interact at multiple spatial and temporal scales, the differentiation of cross-scale influences on water clarity is essential for informed interpretation of lake clarity data and any resulting management decisions.

Shoreland Development and its Effect on Littoral Habitat

National Lake Assessment (NLA) surveys evaluate the quality of littoral and lakeshore habitat by enumerating the types and coverage extents of various components of habitat features in the littoral zone (woody structure, inorganic substrate types and features, macrophyte beds, undercut banks, etc.) as well as the riparian area (layers and types of terrestrial vegetation). Poor habitat quality is generally less complex and may be missing components that are important to various types of lake-dependent biota. In the 2007 NLA, the USEPA (2009) identified that 36% of US lakes had poor lakeshore (riparian and littoral) habitat. In the Northern Appalachian ecoregion, in which Maine is located, 57% of lakes had moderate or high levels of lakeshore disturbance, and 55% of lakes had fair or poor littoral habitat condition. The 2012 NLA showed similar results as the 2007 survey, showing 51% of US lakes were in fair or poor condition for lakeshore habitat. In New England lakes, 28% and 18% of lakes were in poor condition for lake habitat complexity and littoral habitat, respectively (USEPA 2016).

These NLA survey findings showed that lakeshore disturbance, and its resulting effects on littoral habitat, is a major concern for lakes in the Northeast US (USEPA 2009, 2016; Kaufmann et al. 2014a). This identified a gap in the available lake assessment tools in this region, as systematic approaches to littoral habitat condition evaluation are not commonly practiced. This is especially relevant in Maine, where lake habitat condition is specifically addressed in a water quality statute, stating that lake "*habitat must be characterized as natural*" (MRS 38, §465-A). "Natural" is defined in statute as "*living in, or as if in, a state of nature not measurably affected by human activity*" (MRS 38, §466), but methods to measure the effect of human activity on lake littoral habitat have not been established.

Disturbance in lake watersheds and along shorelines is generally included as a contributing factor to lake eutrophication (Soranno et al. 1996, Dennis 1986, Garrison and Wakeman 2000, Garrison et al. 2010), but methods to isolate its effect on littoral habitat quality have not been incorporated into assessments of lakes in Maine. By developing methods to do so, an assessment gap may be filled and an important provision in the Maine water quality statutes may be fully addressed. With this assessment capability, lake resource managers will be better able to prioritize efforts to rehabilitate lakeshores and improve littoral habitat condition where that intervention is needed most.

Summary of Chapters

A common theme in this dissertation was the development of tools that help to inform lake assessments by enhancing the ability to differentiate between natural and anthropogenic factors that influence lakes. To investigate various aspects of lakes and their watershed that affect lake condition, the following four studies were pursued:

1) Chapter 2, *Landscapes and watersheds* (Deeds et al. 2020)*:* This study describes a Hydrogeomorphic-based lake classification, tested with mixed-effect linear models, that was used to determine which natural features of lakes and their watersheds show the greatest influence on lake condition. This classification allows for the establishment of water quality benchmarks to determine if lakes are meeting their expected thresholds for lakes in reference (minimally developed) or altered (heavily developed) watersheds based on the unique qualities inherent to each type of lake.

   <u>Needs addressed</u>: Categorizing lakes by common landscape traits that affect water quality in similar ways and defining expected data ranges based on levels of human watershed disturbance will help interpret monitoring data by providing context for lake data with respect to the important natural and anthropogenic features of their watershed.

2) Chapter 3, *Lake Morphometry* (Deeds et al. 2021a)*:* We developed a predictive logistic regression model that uses lake morphometry and stratification variables to estimate the likelihood of hypolimnetic anoxia occurring in low-nutrient (TP < 15 µg/L) Maine lakes. We tested the predictive potential of variables related to basin morphometry, positioning of lake thermal strata, epilimnetic TP, and DOC. We found that maximum lake depth and the percent of lake area below the epilimnion provided the most robust model.

   <u>Needs addressed</u>: Understanding which natural lake variables, outside of nutrient concentration, help to predict the occurrence of anoxic conditions will aid in the interpretation of lake data that show seasonal anoxia. This distinction will help prioritize management and remediation efforts by determining if hypolimnetic anoxia is a result of natural lake characteristics or human activities.

3) Chapter 4, *Regional vs. Local Effects* (Deeds et al. 2021b): We used smoothed mixed-effect generalized additive model curves derived from temporal Secchi disk transparency data collected from minimally-disturbed Maine reference lakes over time to serve as a dynamic baseline for changes in regional water clarity through time. Mann-Kendall Tau trend analyses on the residual differences between the reference baseline data and yearly Secchi data in non-reference lakes were used to assess the trends from the non-reference lakes by accounting for regional variability and focusing the assessment on local-scale changes in lake water clarity. A dynamic factor analysis revealed that cumulative precipitation during the lake stratification season explained the greatest amount of variability in reference lake water clarity in this dataset.

   <u>Needs addressed</u>: It is often challenging to determine if variations in Secchi disk clarity trends in lakes over time are due to regional background factors (e.g., weather) or localized watershed alterations (e.g., anthropogenic land development). The use of baseline trends of expected water clarity values from minimally disturbed reference lakes will facilitate better-informed interpretations of clarity trend data for individual lakes.

4) *Chapter 5, Shoreland Disturbance* (Deeds et al., to be submitted to *Lakes and Reservoir Management*): We developed an assessment methodology within the framework of the National Lake Assessment (NLA) that supports the evaluation of littoral habitat for individual lakes based on comparison to a natural reference condition. We used data from NLA littoral habitat surveys to create multi-metric indices with linear discriminant models that predict the likelihood each lake meeting the expected reference condition for littoral habitat. Linear discriminant model scores from individual sites were used to create 95% bootstrapped confidence intervals that placed lakes into categories of *natural*, *intermediate*, or *impaired* for littoral habitat.

Needs addressed: Residential development on lakeshores adversely affects littoral habitat, but there currently is not a widely used method to determine if habitat impairment is occurring in individual lakes. Maine’s water quality statute for lakes states that “habitat must be characterized as natural,” signifying a regulatory need for an assessment methodology addressing habitat condition. The ability to isolate shoreland development as a specific stressor to lakes with systematic evaluations of littoral habitat will facilitate the prioritization of shoreland condition rehabilitation on lakes that need it most.

The research included in this dissertation will help to elucidate the differences in lake condition that may be attributed to natural features of the landscape versus changes to the landscape induced by human alteration. This was done by leveraging data from hundreds of thousands of lake monitoring events that have occurred over several decades, as well as developing new assessment methods, to better understand the network of interacting factors that influence the condition of Maine lakes.

CHAPTER 2. A HYDROGEOMORPHIC AND CONDITION CLASSIFICATION FOR MAINE, USA, LAKES

Chapter Abstract

The water quality of lakes is influenced by natural landscape features, anthropogenic watershed activities, and local-scale characteristics of lake basins. Lake assessment for water quality standards is enhanced by a lake classification framework that allows for comparison of lakes of similar types and the creation of benchmark water quality values from reference lakes. Conventional lake classifications, such as those based on trophic state alone, commonly do not incorporate the natural features of the landscape or lake-specific characteristics that influence lake condition. We present a Hydrogeomorphic (HGM)-based lake classification, tested with linear mixed effects modeling, to determine which natural features of lakes and their watersheds show the greatest influence on lake condition. *A priori* classification schemes were created for model testing using combinations of various HGM features. Model strength was evaluated based on ability to predict mean lake total phosphorus and specific conductivity values. Aggregated Level IV Ecoregions stratified by two categories of maximum lake depth offered the most robust lake classification. We defined condition classes of reference, intermediate, and altered lakes based on the gradient of developed watershed area that is unique to each lake class. This approach to lake classification has implications for any region or study, as HGM variables relevant to the population of lakes of interest may be tested for efficacy in a variety of schemes to suit the goals of the classification.

Introduction

Lakes in Maine are distributed along environmental gradients that reflect differences in surficial and bedrock geology, soil types, topography, watershed hydrology, proximity to the ocean, atmospheric deposition, climate patterns, terrestrial vegetation, and elevation, all of which have a role in the export of phosphorus (P) into lakes (Omernik 1977, Prairie and Kalff 1986, Norton et al. 1989, D'Arcy and Carignan 1997, DeVito et al. 2000, Soranno et al. 2015). Gradients of these environmental factors influence the biological condition of lakes through the availability of nutrients, processing of organic matter, primary production of autotrophs, and regulation of chemical habitat in lakes.

Excess P in Maine lakes is the cause, either directly or indirectly, of impairment due to accelerated cultural eutrophication (Maine DEP 2018). P can enter runoff to lakes through many mechanisms, including soil erosion, excess fertilizer application, unmitigated storm water, and improperly treated sewage and industrial effluent. Different land uses have variable export rates of P into surface waters (Petrone 2010, Soranno et al. 2015, Chen et al. 2016). Excess P in lakes generally creates nuisance algal growth that may lead to lake anoxia and habitat degradation, undesirable recreation conditions, threatened drinking water supplies, and reduced property values (Michael et al. 1996).

Some Maine lakes are affected by the presence of marine-derived sediments (known as the Presumpscot Formation) in their watershed. These sediments were deposited on now emergent landscape when ocean water inundated the land up to 128 m above present-day sea level after deglaciation, well inland from the current coastline, between about 15,000 - 14,000 years BP (called the marine limit; Thompson and Borns 1985) (Figure 2.1). These sediments occur in low relief areas, are easily eroded and transported, and contain apatite (Ca5(PO4)3(OH)), an important source of P to lakes (Norton et al. 2019). Much of Maine’s agricultural land and population settlements are sited on the Presumpscot Formation. Mean P concentration in Maine lakes correlates positively with the presence of the Presumpscot Formation, wetland area, agriculture, and human disturbance in watersheds (Nieratko 1992).

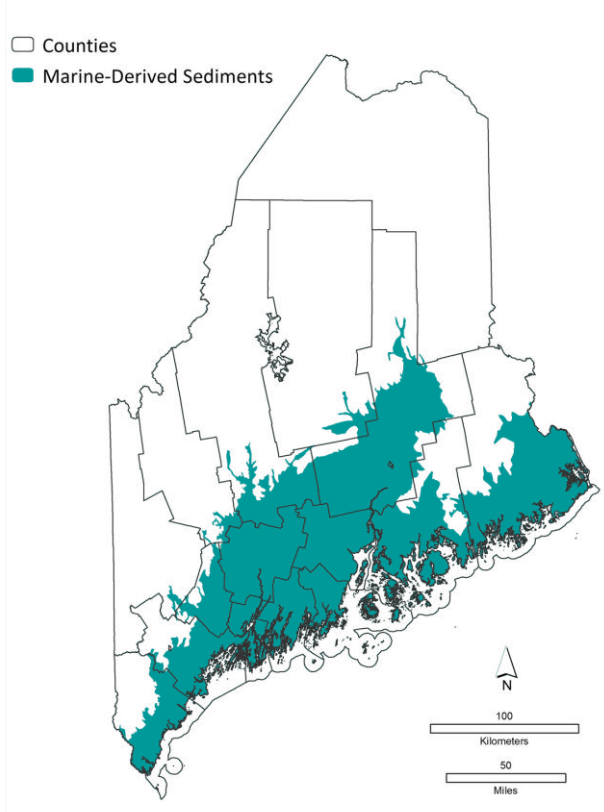


Figure 2.1. Potential location of marine-derived sediments in Maine (shaded area). The area of post-glacial marine inundation extends from modern sea level to approximately 85 m elevation.

Conversion of watershed land area from natural land cover to agricultural, urban, or residential land can adversely affect water quality from a loss of stormwater infiltration and increased erosion of watershed soils (Fraterrigo and Downing 2008, Taranu and Gregory-Eaves 2008). Agricultural areas and the associated application of P-rich fertilizers or livestock waste contribute excess P to watersheds, elevating the trophic state of lakes (Jones et al. 2001, Jones et al. 2004, Taranu and Gregory-Eaves 2008, Nielsen et al. 2012). Urban or suburban development can lead to increased trophic state in lakes from municipal or private septic sewage effluent (Muscutt and Withers 1996) and increased erosion of P-carrying sediment resulting from increased impervious surface area (Newman et al. 2006, Merugu and Seetharaman 2013). Application of road salt to mitigate winter icing can have lasting effects on the chemistry of watershed soils and lakes, resulting in increased concentration of ions in lake water

(Sutherland et al. 2018), especially Cl- which can be toxic to aquatic life (Nandana et al. 2013). Increased Cl- concentrations in lakes can also disrupt seasonal mixing regimes, leading to persistent hypolimnetic anoxia (Dupuis et al. 2019). Lakes in urban or suburban areas also may have elevated total dissolved solids, hardness, and alkalinity (Merugu and Seetharaman 2013).

The effect of natural and anthropogenic features of watersheds on lake water quality are commonly interrelated. Fraterrigo and Downing (2008) and Nielsen et al. (2012) highlight the importance of interactions among watershed land use, the nutrient transport capacity of individual watersheds, and lake water quality. Lakes that have watersheds with high capacity for nutrient transport, where nutrients are readily conveyed overland to receiving waters (Fraterrigo and Downing 2008), are influenced by nutrient inputs far from the lake more than lakes in watersheds with low capacity to transport nutrients. Lakes in watersheds with low transport capacity may only be influenced by land use conditions close to the lakeshore. Nutrient transport capacity is affected by watershed geology, soil types, and hydrologic connectivity (Fraterrigo and Downing 2008). Taranu and Gregory-Eaves (2008) demonstrated that watershed agriculture influences all lakes, and shallow lakes are more susceptible to P enrichment than deeper lakes.

There are several approaches to lake classification. Trophic-based classifications (i.e., oligotrophic, mesotrophic, eutrophic), historically used in many state programs, are useful for describing nutrient levels and the variety of biological assemblages that may exist in a group of lakes (e.g., Uttormark and Wall 1975). Biological classifications use biotic community data to identify lake types based on similar assemblages of plants or animals (e.g., Neale and Rippey 2008). Habitat-based classifications are used to define the biological potential in lakes of similar types based on lake-level habitat evaluations (e.g., Olivero-Sheldon and Anderson 2016). However, a limitation of classification schemes such as these is the necessity of obtaining field data. If data availability is limited, this reduces the number of lakes in the classification-building dataset; uncharacterized lakes may only be assigned to classification groups after new data have been collected. Another limitation of these classifications is that lake type assignments may be subject to interpretation or may change altogether if trophic status or habitat conditions evolve, allowing lakes to move between designations. While the approaches to classification discussed above may be suitable for some lake management needs, a classification used for assessments related to water quality standards must have clear, unchanging distinctions among types. For this purpose, a successful lake classification:

1. creates groups of lakes that are similar, using clearly-defined features that do not change;
2. establishes reference conditions and expected water quality ranges;
3. can be applied to lakes with little or no sampling data;
4. provides a well-defined classification scheme that promotes interpretation and communication of water quality assessments.

An alternative to habitat- or sample-based approaches is a classification that incorporates Hydrologic (e.g., landscape drainage), Geologic (e.g., soil types and parent material), and/or Morphological (e.g., depth) characteristics of lakes (Martin et al. 2011). Such an HGM classification integrates multiple lake and watershed characteristics that influence lake condition. Geographic characteristics can group lakes that have similar central tendencies in water quality data (Cheruvelil et al. 2008). Morphological variables help describe lake-to-lake variation of in-lake processes and water quality but can miss unifying features of lakes at larger spatial scales. A true HGM approach, combining lakes from areas of similar landscapes with morphological variables may offer the best classification system for water quality assessment purposes (Martin et al. 2011, Soranno et al. 2015).

An HGM lake classification framework supports lake water quality assessment and management programs by 1) assisting data interpretation; 2) simplifying management decisions by grouping similar lakes that may respond predictably to similar management approaches, and 3) establishing conditions for reference water quality values (Stoddard et al. 2006, Herlihy et al. 2013). Developing waterbody classifications and determining appropriate reference conditions increases understanding of lake and watershed processes by describing how lakes of similar type respond to changes in their watersheds.

We propose three condition classes within each HGM lake type:

1. *Reference lakes* have watershed condition and water quality that are closest to undisturbed, or "natural" condition. Water quality parameters in these lakes indicate the least disturbed condition for a lake type in a region (e.g., Stoddard et al. 2006). This condition class will help define what 'natural' water quality conditions should be for different types of lakes.
2. *Intermediate Lakes* indicate the normal range of conditions that most lakes show, after some watershed alteration. The designation of an "intermediate" lake condition will provide expected water quality parameter values for the majority of lakes within each type.
3. *Altered lakes* are in watersheds that represent the most anthropogenically disturbed for their respective classification type. Water quality parameters reflect diminished lake condition, and will inform lake assessments and management strategies by characterizing poor lake water quality conditions for different types of lakes.

There are two goals of the classification presented herein. First, to define lake types based on HGM variables that have similar natural conditions and are affected by changes in their watershed in similar ways. This is done by evaluating several *a priori* HGM classification models for their strength of classification with linear mixed effects modeling based on water quality response variables. Second, we developed a Watershed Quality Index (WQI) to define reference, intermediate, and altered watershed condition classes for all HGM lake types.

Methods

Study Area

Maine has a land area of 91,647 km^2 and approximately 2800 lakes > 4 ha. Elevation ranges from sea level to 1608 m. Six United States Geological Survey (USGS) Hydrologic Unit Code 4 (HUC4) drainage areas (mean land area approx. 14,000 km^2) fall within Maine (USGS et al. 2013; Figure A.1a). The northwestern quadrant of the state is relatively unpopulated. The northern and northwestern interior has a history of timber harvesting and re-growth. Dairy and cropland agricultural operations are largely within 100 km of the coast, and in the northeastern part of the State. These land uses were historically more extensive and are common throughout the populated portion of Maine. The state's urban centers are located primarily in southern and coastal areas, as well as along the main stems of its larger rivers.

Datasets

Lake Data - Data for Maine lakes were obtained and summarized from the Maine Lakes Database (Maine DEP 2004). The data were collected by state agency staff, certified volunteer monitors, and other collaborators from 1472 stations on 1060 lakes (representing surveys on 84% of lake surface area in Maine for lakes >4 ha). Data were screened for time period (1988-2017) and frequency (minimum 3 y of sample data per water quality parameter), resulting in 735 sampling stations in 592 lakes. Large lakes or lakes with multiple basins generally have more than one sampling station. Water quality data are based on integrated epilimnetic core samples collected during summer stratification and analyzed for trophic indicators (chlorophyll-a, Chl; total phosphorus, TP; and Secchi disk transparency, SDT), specific conductivity (SpCon), and alkalinity (Alk). All water quality data were log-transformed for statistical analyses. All statistical analyses were performed with R software, version 3.6 (R Core Team 2013).

Hydrogeomorphic (HGM) Data - Five geographical schemes were used to reflect hydrogeologic variations across Maine: 1) USGS HUC4-level areas which represent major river basins or aggregations of smaller coastal drainages (Figure A.1a); 2) level III and 3) level IV Ecoregions (Figures A.1b and c, respectively), which are areas of land with similar geology, soils, climate, land use, vegetation, and hydrology (Omernik and Griffith 2014; Omernik 1987, 1995 and 2004); 4) Maine's Biophysical regions, which are areas of land with similar vegetation and landscape features (McMahon 1990) (Figure A.1d); and 5) the marine limit (Figure 2.1). To combine similar HUC4 areas, Level IV Ecoregions, and Biophysical regions, a hierarchical cluster analysis was performed on these regionalization schemes using median values for lake TP and SpCon in each sub-region with the "cluster" package in R (Maechler et al. 2019) (Ward's method, Euclidian distance, Figure A.2). These two variables were used because they are related to mutually independent stressors for lakes: TP for nutrient enrichment and SpCon for urban and residential watershed development. Adjacent sub-regions that clustered together were grouped together (Figure A.1e-g). HUC4 areas were also combined into an alternative Ecological Drainage Unit (EDU) scheme to represent major drainage areas in Maine (Figure A.1h). The marine limit (Figure 2.1) was used to separate lakes into those with potential for Presumpscot Formation in their watersheds and those without. A final scheme was tested that separated the St. John River Drainage in northern Maine (Figure A.1h) from the rest of the state above the marine limit.

Lake surface area and maximum depth were used for morphological parameters as they are both easily measured and available for most Maine lakes. Surface area values were derived in GIS from the National Hydrography Dataset (NHD; USGS et al. 2013). Two lake size categories were established based on the median surface area for all Maine lakes > 4 ha (median = 50 ha). Maximum depth values are based on bathymetric surveys (Maine Geolibrary 2011a). Lakes were placed into two depth categories based on the mean SDT for Maine lakes (~5 m); lentic waters deeper than twice the mean SDT (10 m = the photic zone; French et al. 1982) were designated as deep lakes, waters ≤10 m deep were designated as shallow lakes.

Watershed Condition Data

Large-scale characteristics of land area related to human land use settlement were associated with each lake's direct drainage area (DDA) in GIS. Lake DDAs differ from watersheds in that they exclude upstream lakes and their DDAs. The DDAs were used in this analysis because they showed the strongest correlations between watershed condition data and lake water quality variables, stronger than total watershed area or isometric buffer areas around lake shorelines.

The Maine Land Cover Dataset (MELCD) covers all of Maine at 5 m spatial resolution, providing suitable resolution for statewide land cover data (Maine Geolibrary 2006). Impervious cover (IC) data were generated by the Maine Office of GIS with IMPERV, a raster dataset that quantifies impervious cover across Maine at 5 m resolution (Maine Geolibrary 2005). Impervious areas include buildings, roads, and parking lots. Census blocks are small geographic areas delineated by other mapped features (e.g., roads, rivers, political boundaries) by the U.S. Census Bureau that represent similar densities of population and housing units (Maine Geolibrary 2011b). The census blocks determine population statistics for specific areas. MELCD, IC, population, and housing unit density data were aggregated for individual DDAs using GIS.

To determine the effect of watershed condition on various classification schemes, the percentage of undeveloped land (TotUnd) in each DDA was calculated by subtracting developed land areas from the total area (MELCD land cover classes: urban/residential: High, Medium, and Low Intensity Development, Roads; agricultural/timber harvest: Crops, Hay, Clearcut, and Blueberry). TotUnd values were used to exclude the 25% of lakes with the most heavily developed watersheds from classification development, thereby emphasizing natural landscape factors that differentiate Maine's lakes (USEPA 2000). After excluding the most heavily developed watersheds, the number of lakes used for the models were 295 and 279 for TP and SpCon, respectively.

Classification Development

Lake classification groupings based on various combinations of HGM variables were evaluated with linear mixed effect models using the “lme4” package in R (Bates et al. 2015). We used the HGM variables as random effects on the intercept of each model (Zuur et al. 2009). Model intercepts were allowed to vary by HGM lake class. TP and SpCon were used as response variables in separate models. Overall model strength was evaluated based on Akaike Information Criterion (AIC), which provides an estimation of model strength, based on the number of model parameters and the log-likelihood of the model. The corrected AIC (AICc) accounts for model fit and complexity and is corrected for small sample sizes. A smaller AICc indicates a better-fit model (Cheruvelil et al. 2008). AICc values and other model averaging parameters were calculated with the “AICcmodavg” package (Mazerolle 2017) for each grouping scheme separately for TP and SpCon models. Analyses of variance (ANOVAs) and Tukey HSD *post-hoc* tests were performed to determine which lake types were significantly different within each classification.

After the final classification scheme was chosen, TotUnd was used to create a watershed quality index (WQI), enabling the placement of lakes in condition classes. Because the intensity and types of human disturbance vary across Maine, values of TotUnd were arcsin square-root transformed and rescaled 0-1 independently for each lake type. The arcsin square-root transformation is effective for proportional data because it accentuates differences among untransformed values as they approach 0 or 100%. This is important due to the high portion of DDAs in Maine that are largely undeveloped. This transformation enabled the WQI to accentuate differences in watershed condition unique to each lake type that influence lake water quality. Small differences in land cover percentages are particularly important concerning impervious surfaces, where slight increases in impervious area have significant effects on water quality and aquatic biota (Danielson et al. 2016). WQI thresholds to categorize condition classes were based on watershed condition, not lake water quality data, so that lake water quality value ranges indicative of the least- and most-disturbed watershed condition specific to each lake type could be determined. Three versions of condition class thresholds were tested based on 10th, 25th, 75th, and 90th percentiles of the distribution of WQI values. The low and high ends of the WQI distribution served as thresholds for Altered and Reference lakes, respectively (version one: 10th/90th percentiles, version two: 25th/75th, version three: 25th/90th). Lakes that fell between threshold values were categorized as Intermediate lakes. Two-way ANOVAs were used for TP and SpCon separately to determine which WQI threshold values differentiated condition classes best among lake types, with water quality as the response variable and lake type and condition class as predictors with interaction.

<u>Results</u>

Lake Water Quality and Watershed Condition

Trophic measures of TP, Chl, and SDT were correlated ($|R| \geq 0.5$), as were Alk and SpCon (Table 2.1). Alk and SpCon showed the strongest correlations to watershed condition variables, particularly measures of agricultural and impervious area. The metric of TotUnd showed the strongest inverse correlations to lake TP and SpCon. Therefore, TotUnd was used to evaluate classification groupings for relatedness to trophic condition and water chemistry.

Classification Scheme Selection

Models were tested for 10 geographic area schemes, then the morphological variables of depth, size and depth + size combined were added to each scheme. Models with only morphological variables were tested as well, resulting in 43 *a priori* lake classification schemes (Table 2.2). The number of lake types in each scheme ranged from two (DepthCat) to 51 (Level4 + DepthCat + SizeCat). AICc values ranged from -232 to -125 (TP) and -178 to -124 (SpCon). The top 10 models for TP and SpCon, based on AICc, are in Table 2.3. Lower AICc values mean that the water quality parameters are more similar within a single lake type than other lake types in a particular classification scheme.

Aggregated Level IV Ecoregion Groups (Figure 2.2) stratified by depth category (Eco4 + Depth) was the overall strongest classification model for TP (Table 2.3). This was followed by Eco4 + Depth + Size (ΔAICc = 2.07). The remaining TP models are comparatively unsuitable with ΔAICc values >10. The strongest SpCon models were Eco4, Eco4 + Size (ΔAICc = 1.47), and Eco4 + Depth (ΔAICc = 2.06). Eco4 + Depth was chosen as the final lake classification scheme due to high AICc ranking for both TP and SpCon. This classification was supported by significant differences among lake types for water quality values identified by ANOVA results (TP: $F_{4,290} = 25.013$, $p < 0.001$; SpCon: $F_{4,274} = 11.398$, $p < 0.001$; Figure 2.3). These "Maine Lake Regions" delineated by the Eco4 + Depth classification reflect differences in geology and land use (Table 2.4), and climate (Table 2.5) across the state.

Watershed Quality Index (WQI) and Condition Class Development

Watershed condition criteria were established for each Eco4 + Depth lake type. TotUnd had the strongest inverse correlation to TP and SpCon lake data among tested metrics (Pearson's R = -0.500 and -0.625, respectively; Table 2.1) and was consequently used for delineating watershed condition classes.

TotUnd was arcsin square root transformed and re-scaled to 0-1 for each lake type. This new value, Watershed Quality Index (WQI), is distinct from the homogenous statewide TotUnd metric. WQI distributions for each lake type were visually inspected with histograms. WQI Version 3 (25^{th}/90^{th} percentiles) was the strongest model based on F-values of the predictors and the interaction effects for both TP and SpCon (lake type*WQI v. 3; TP: $F_{8,434} = 3.95$, $p < 0.001$; SpCon: $F_{8,404} = 4.50$, $p < 0.001$) (Table A.1). Tukey HSD *post-hoc* tests showed numerous significant differences among WQI v. 3 condition classes within lake types (Figure 2.4). Threshold TP and SpCon values, serving as benchmarks for reference and altered lake condition, were assigned based on the mean values from the corresponding condition classes for each lake type (Table 2.6). Shallow and deep lakes in the Northern region were treated in one category for condition classes due to low sample size.

Table 2.1. Pearson's R correlation matrix of water quality parameters and watershed condition data. Values ≥0.5 and ≤-0.5 are in bold; n = 465 lakes.

	Parameters	Alk	SpCon	TP	Chl	SDT
Water Quality Parameters	Alkalinity	--	--	--	--	--
	Specific Conductivity	**0.729**	--	--	--	--
	Total Phosphorus	0.476	0.479	--	--	--
	Chlorophyll-*a*	0.383	0.422	**0.863**	--	--
	Secchi Disk Transparency	-0.369	-0.356	**-0.866**	**-0.770**	--
Lake Morphometry	Lake Maximum Depth	-0.076	-0.133	**-0.560**	-0.373	**0.643**
	Lake Surface Area	0.115	-0.080	-0.155	-0.078	0.199
Census Data	Estimated DDA Population	0.371	0.449	0.250	0.251	-0.131
	Estimated DDA Housing Units	0.317	0.296	0.120	0.141	-0.026
Roads	DDA road density	0.100	0.254	0.058	0.044	0.074
	DDA total road length	0.307	0.188	0.062	0.097	0.012
Land Cover Classes (MELCD and IMPERV)	High Intensity Development	0.178	0.388	0.193	0.194	-0.137
	Medium Intensity Development	0.175	0.410	0.225	0.220	-0.158
	Low Intensity Development	0.233	0.475	0.220	0.201	-0.116
	Open Space	0.240	0.327	0.167	0.144	-0.187
	Crops	**0.566**	0.492	0.432	0.390	-0.285
	Deciduous Forest	0.096	-0.014	0.030	-0.006	0.067
	Hay/Pasture	0.495	**0.576**	0.450	0.388	-0.274
	Herbaceous Vegetation	-0.109	0.075	-0.104	-0.077	0.067
	Evergreen Forest	-0.130	-0.123	-0.201	-0.100	0.094
	Mixed Vegetation	0.151	0.167	0.060	0.049	-0.048
	Scrub/Shrub Vegetation	0.031	0.097	-0.067	-0.034	-0.019
	Roads & Runways	0.076	0.378	0.053	0.059	0.061
	Forested wetlands	0.214	0.201	0.327	0.306	-0.364
	Wetlands	0.046	-0.036	0.268	0.285	-0.326
	Shore	-0.089	-0.127	-0.133	-0.091	0.106
	Bare land	0.032	0.178	0.083	0.111	-0.051
	Open Water	-0.039	-0.111	-0.144	-0.086	0.151
	Blueberry fields	-0.064	-0.069	-0.049	-0.077	0.043
	Clear-cut timber harvest area	0.129	-0.013	-0.080	-0.130	0.010
	Light/partial cut timber harvest	0.047	-0.142	-0.046	-0.096	0.003
	Heavy/partial cut timber harvest	-0.237	-0.283	-0.215	-0.232	0.157
	Forest regeneration area	0.088	-0.145	-0.046	-0.047	-0.003
	Impervious Surfaces	0.185	**0.572**	0.187	0.168	-0.057
	Hi + Med + Low development	0.253	**0.527**	0.256	0.238	-0.144
	Forest area + wetland area	-0.156	-0.079	-0.159	-0.095	0.128
	Agricultural area	**0.549**	**0.583**	0.456	0.390	-0.267
	Crops + Hay/Pasture	**0.572**	**0.610**	0.477	0.415	-0.286
	Total Undeveloped Land (TotUnd)	**-0.520**	**-0.625**	**-0.500**	-0.437	0.290

Table 2.2. Lake classification schemes tested with linear mixed effect models for TP and SpCon. Maps of the various geographic schemes listed here are in Figures 2.1, 2.2 and A1.

Model #	Model Parameters	# Lake Types	Geographic/ Morphometric Scheme
1	SizeCat	2	small lakes: < 50 ha; large lakes ≥ 50ha
2	DepthCat	2	shallow lakes: < 10 m; deep lakes ≥ 10 m
3	DepthCat + SizeCat	4	DepthCat * SizeCat
4	HUC4	6	USGS HUC4 Drainage Areas (Fig S1a)
5	HUC4 + DepthCat	12	
6	HUC4 + SizeCat	12	
7	HUC4 + DepthCat + SizeCat	24	
8	EDU	3	Ecological Drainage Units (Fig S1e)
9	EDU + DepthCat	6	
10	EDU + SizeCat	6	
11	EDU + DepthCat + SizeCat	12	
12	Biophysical	15	Biophysical Regions (Fig S1d)
13	Biophysical + DepthCat	29	
14	Biophysical + SizeCat	29	
15	Biophysical + DepthCat + SizeCat	50	
16	Level3	3	Level III Ecoregions (Fig S1b)
17	Level3 + DepthCat	6	
18	Level3 + SizeCat	6	
19	Level3 + DepthCat + SizeCat	12	
20	Level4	17	Level IV Ecoregions (Fig S1c)
21	Level4 + DepthCat	31	
22	Level4 + SizeCat	29	
23	Level4 + DepthCat + SizeCat	51	
24	Marine	2	Marine Sediment/Non-Marine Sediment (Fig 1)
25	Marine + DepthCat	4	
26	Marine + SizeCat	4	
27	Marine + DepthCat + SizeCat	8	
28	MarineSJ	3	Marine Sediment/Non-Marine Sediment + St. John River Basin
29	MarineSJ + DepthCat	6	
30	MarineSJ + SizeCat	6	
31	MarineSJ + DepthCat + SizeCat	12	
32	Eco4	3	Aggregated Level IV Ecoregions (Fig 2)
33	Eco4 + DepthCat	6	
34	Eco4 + SizeCat	6	
35	Eco4 + DepthCat + SizeCat	12	
36	Biop	4	Aggregated Biophysical Regions (Fig S1h)
37	Biop + DepthCat	8	
38	Biop + SizeCat	8	
39	Biop + DepthCat + SizeCat	16	
40	HUC4_Agg	4	Aggregated HUC4 Drainage Areas (Fig S1g)
41	HUC4_Agg + DepthCat	8	
42	HUC4_Agg + SizeCat	8	
43	HUC4_Agg + DepthCat + SizeCat	16	

Table 2.3. Linear mixed effect model results for predicting total phosphorus and specific conductivity.

Model #	Grouping Scheme	K	AICc	ΔAICc	ModelLik	AICcWt	Res.LL
Total Phosphorus							
33	Eco4 + Depth	4	-232.74	0	1.00	0.73	120.44
35	Eco4 + Depth + Size	5	-230.67	2.07	0.36	0.26	120.44
13	Biop + Depth	4	-221.35	11.39	0.00	0.00	114.74
37	Biop_Agg + Depth	4	-221.22	11.52	0.00	0.00	114.68
15	Biop + Depth + Size	5	-219.28	13.46	0.00	0.00	114.74
39	Biop_Agg + Depth + Size	5	-219.15	13.59	0.00	0.00	114.68
29	MarineSJ + Depth	4	-214.93	17.81	0.00	0.00	111.53
31	MarineSJ + Depth + Size	5	-212.86	19.88	0.00	0.00	111.53
25	Marine + Depth	4	-211.46	21.28	0.00	0.00	109.80
2	DepthCat	3	-209.80	22.94	0.00	0.00	107.94
Specific Conductivity							
32	Eco4	3	-178.73	0	1.00	0.31	92.41
34	Eco4 + Size	4	-177.26	1.47	0.48	0.15	92.70
33	Eco4 + Depth	4	-176.67	2.06	0.36	0.11	92.41
20	Level4	3	-176.27	2.46	0.29	0.09	91.18
21	Level4 + Depth	3	-176.27	2.46	0.29	0.09	91.18
22	Level4 + Size	3	-176.27	2.46	0.29	0.09	91.18
23	Level4 + Depth + Size	3	-176.27	2.46	0.29	0.09	91.18
35	Eco4 + Depth + Size	5	-175.18	3.55	0.17	0.05	92.70
28	MarineSJ	3	-160.17	18.56	0.00	0.00	83.13
29	MarineSJ + Depth	4	-159.33	19.40	0.00	0.00	83.74

Table 2.4. Land area percentages of surficial geology and land cover types in each of the three Maine Lake Regions delineated by the Eco4 + Depth classification (Figure 2.3).

Surficial Geology (500K) Type	Coastal	Interior	Northern
Beach deposits	0.0	0.0	0.0
Bedrock	1.1	5.9	1.3
Emerged beach deposits	0.0	0.0	0.0
End moraine	0.5	0.0	0.1
Eolian deposits	0.2	0.0	0.0
Eskers	1.8	1.3	0.2
Glacial outwash deposits	0.6	1.0	1.2
Glaciomoraine deposits (coarse-grained facies)	4.5	0.4	0.0
Glaciomoraine deposits (fine-grained facies)	25.3	0.8	0.0
Ice-contact glaciofluvial deposits (exclusive of eskers)	2.8	1.5	1.5
Lake-bottom deposits	0.2	0.3	0.3
Ribbed moraine	0.0	4.1	0.1
Stagnation moraine	0.0	0.0	9.7
Stream alluvium	0.9	0.6	1.6
Swamp, marsh, and bog deposits	3.6	4.6	3.6
Thin drift, undifferentiated	1.6	0.2	3.8
Till	56.9	79.1	76.4
MELCD (2004) Type			
Urban/Residential	6.52	1.29	1.25
Agriculture	8.20	1.04	5.79
Timber Harvest	8.28	15.94	15.69
Forest/Wetland	70.83	74.50	74.84
Barren/other	0.90	0.53	0.18
Water	5.27	6.71	2.24

Table 2.5. Mean area-weighted monthly climate data summarized for the three Maine Lake Regions delineated by the Eco4 + Depth classification (Figure 2.2). Data from USDA-NRCS (2012a-c).

Lake Region	Avg Precip (mm)	Avg Min Temp (°C)	Avg Max Temp (°C)
Coastal	116.8	1.2	12.3
Inland	114.6	-0.7	10.8
Northern	105.7	-1.6	9.5

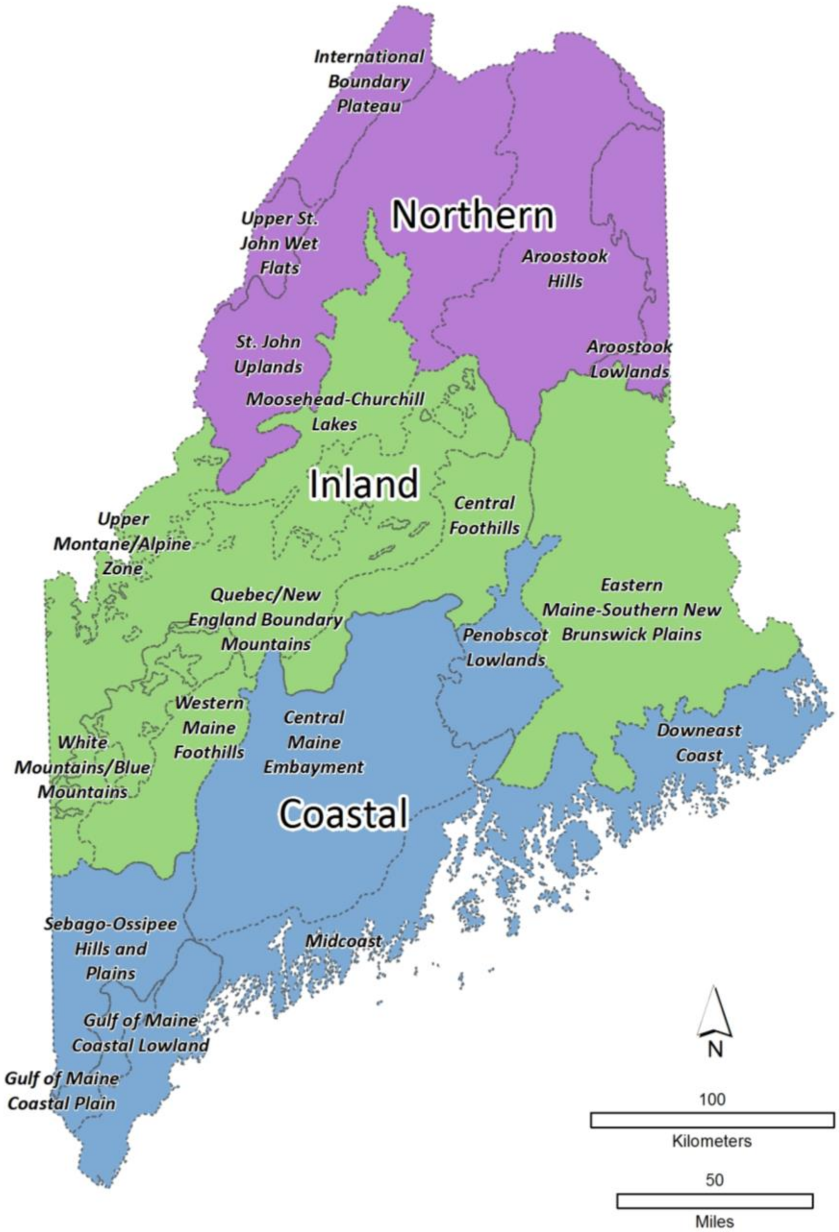


Figure 2.2. Level IV Ecoregions of Maine and the boundaries of aggregated ecoregions resulting from cluster analysis. The aggregated ecoregions serve as geographical boundaries for the Eco4 + Depth classification scheme, partitioning lakes into Coastal, Inland and Northern regions. Individual Level IV Ecoregions are labeled and demarcated with dashed lines within each of the three main regions.

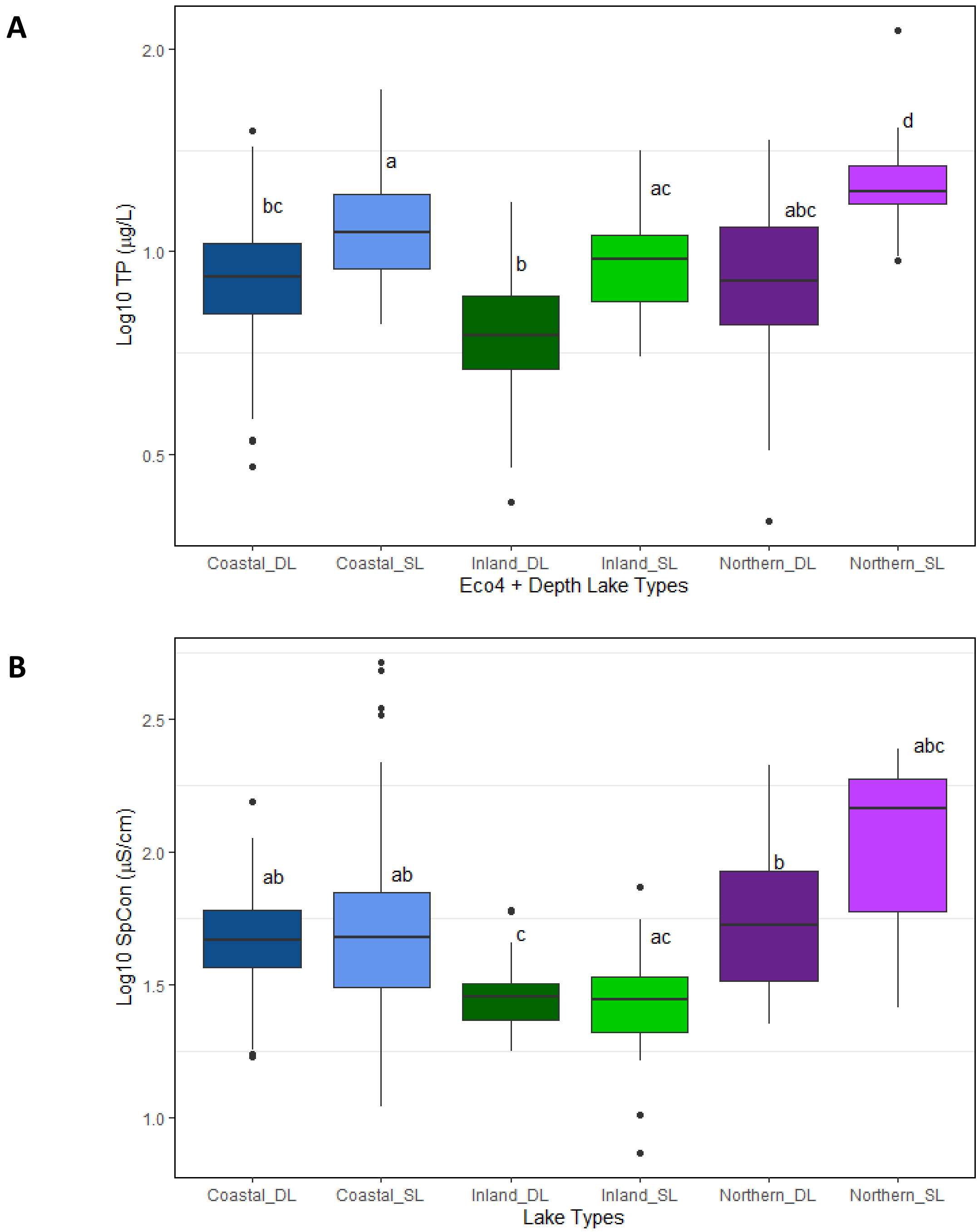


Figure 2.3. Boxplots for TP (A) and SpCond (B) for lake classes in the Eco4 + Depth scheme. DL = Deep Lakes (≥10 m maximum depth), SL = Shallow Lakes (<10 m maximum depth). Statistically different groups are noted with different lowercase letters above each box (Tukey's HSD, α = 0.05). Similar letters indicate non-significant differences.

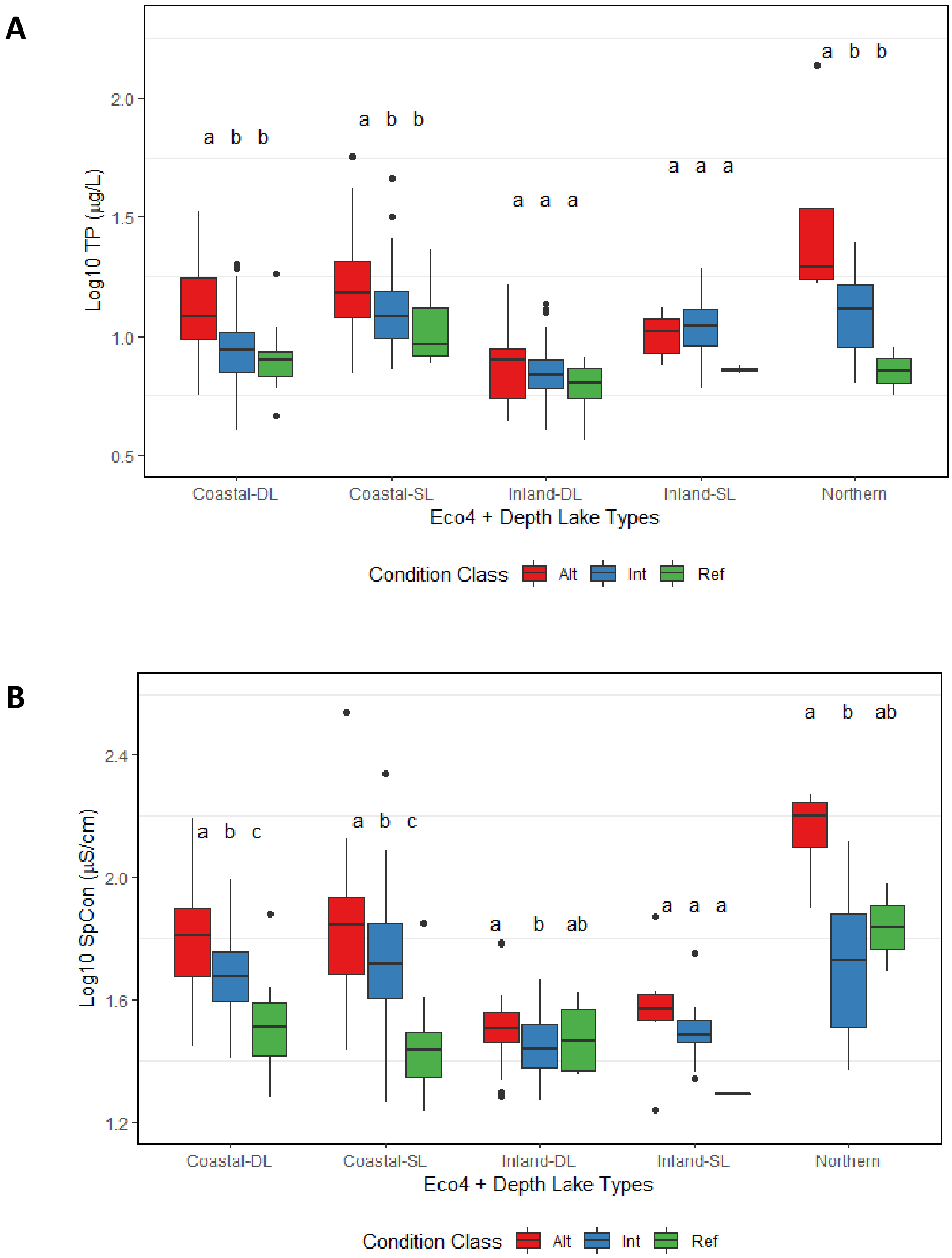


Figure 2.4. Boxplots for lake type and condition class for TP (A) and SpCon (B). DL = Deep Lakes (≥10 m maximum depth), SL = Shallow Lakes (<10 m maximum depth). Results of Tukey HSD *post-hoc* analyses for within-type condition class distinctions are in text above plots. Alt = Altered, Int = Intermediate, Ref = Reference. Deep and Shallow lakes from the Northern region were combined due to low sample size. Statistically different groups within each lake type are noted with different lowercase letters above each box (Tukey's HSD, α = 0.05). Shared letters indicate non-significant differences within lake types.

Table 2.6. Threshold values for TP and SpCond for deep and shallow lakes in reference and altered condition classes for each lake type. Values are based on mean values from reference shallow and deep lakes from each type (±SE). Water quality values for reference lakes are expected to be less than those listed in the reference columns below, and values for altered lakes are expected to be greater than those in the altered columns below. Intermediate lakes are expected to have values in between reference and altered. Deep and shallow lakes in the Northern region were combined for the condition class analysis due to small sample size. Full summary statistics from all lake types and condition classes are presented in Table A.1. *Indicates sample size of lakes with parameter data too small to calculate standard error (n<3).

Lake Type	Total Phosphorus (µg/L)		Specific Conductivity (µS/cm)	
	Reference	Altered	Reference	Altered
Coastal Deep Lakes	8.3 ± 0.7	13.4 ± 4	34.2 ± 3.2	66.3 ± 4
Coastal Shallow Lakes	12.3 ± 1.4	18.5 ± 12	30 ± 4.2	80 ± 12
Inland Deep Lakes	7.5*	10.3 ± 7.7	19.8*	40.4 ± 7.7
Inland Shallow Lakes	6.4 ± 0.5	8.2 ± 2.9	30.4 ± 2.3	37.3 ± 2.9
Northern Lakes	7.3 ± 1.7	48.4 ± 24	72.6 ± 22.9	146.6 ± 24

Discussion

The classification proposed here for Maine lakes addresses differences in natural lake condition as well as water quality changes due to anthropogenic activity in the direct watershed. Primary advantages of this classification include:

1. ease of interpretation, as the classification is based on ecoregional boundaries and one simple morphological metric - maximum lake depth;
2. lake type assignment is possible without complex or long-term monitoring data;
3. water quality gradients among lake classes are defined using conventional monitoring parameters;

4. groupings that establish reference conditions, which are necessary for most water quality agency assessments (Table 2.6);
5. reduction of statistical noise in future data analyses resulting from comparisons of lakes of dissimilar types.

The Eco4 + Depth scheme shows statistical strength and practicality in implementation for lake assessments related to water quality standards. This lake classification reflects gradients in water quality variables that respond to different stressors: TP for trophic productivity and SpCon for salt concentration. Read et al. (2015) also found that specific conductivity was related to regional and basin-specific conditions, whereas lake-specific features (e.g., basin morphometry) were most important in governing TP and other trophic variables. The effectiveness of maximum depth as a categorical variable in this classification supports Taranu and Gregory-Eaves (2008), who documented that shallow lakes are more susceptible to P enrichment than deeper lakes.

Excess watershed P is generally the result of over-application of P-rich fertilizers and transport of soil particles via erosion; but lakes can be naturally higher in P because of geological factors such as the presence of marine silt and clay (Nieratko 1992). In Maine, lake specific conductivity values may be influenced by natural factors such as geology, proximity to the marine environment (Norton et al. 1989), and anthropogenic influences such as the application of road salt (Sutherland et al. 2018). The variety of potential stressors measured by these two water quality variables accentuates the utility of having a set of reference lakes for various classes of lakes that are known to have only minimal human alteration in their watersheds (Table 2.6).

The Eco4 regions highlight landscape differences that affect lake water quality (Figure 2.2). The Coastal region corresponds with the area below the inland limit of marine sediment (Figure 2.1), where the Presumpscot Formation silt and clay contribute to the productivity of lakes. This region has the greatest exposure to marine aerosols that can lower pH and mobilize Al and base cations in watershed runoff (Wright et al. 1988, Norton et al. 1989). This sea salt effect is compounded by the application of road salt in this most heavily developed area of Maine (Table 2.4). Sutherland et al. (2018) demonstrated that road salt can have lasting effects on cations in runoff and lakes due to soil ion exchange, similar to the effect of marine-derived salts.

The Inland region is largely undeveloped, with mountainous terrain in the west and less irregular, and rolling topography in the east. The climate is cold in winter, which is meaningful for lakes because of the longer duration of winter ice cover (Hodgkins et al., 2002). Timber harvesting and recreation are the primary land uses (Table 2.4). Soils are generally acidic and coarse-loamy or loamy, although soils in some northern sections are poorly drained. Due to these landscape conditions and little human influence, lakes in the Inland region show less variation among condition classes than those in the other two regions (Figure 2.4).

The Northern Region has the only widespread calcareous bedrock in the state, resulting in lakes with naturally high alkalinity and P concentrations in the eastern part of the region. The lakes in the western portion of this region have slightly lower alkalinity, but lake alkalinity is still higher in this area of the Northern region than in the other two lake regions (Griffith et al. 2009). Timber harvesting and recreation are the most common land uses in the west, and agriculture is highly prevalent in the east. A mixture of intense agriculture and remote wilderness in this region results in the widest ranges of water quality for any of the lake types in this classification (Figures 2.4, 2.5).

The lake regions have disparate climatic conditions for both annual precipitation and temperature (Table 2.5). Temperature differences among these regions are substantial, causing variances in the length of winter ice cover and summer productivity season. The Coastal region has the warmest average minimum and maximum temperatures of the three regions. Sections of the Coastal region include the southernmost latitudes in Maine, and there is a moderating influence of the ocean on winter air temperatures. These factors combine to create a comparatively shorter winter ice-over period in this region and therefore a longer period of summer bioproductivity. The Northern region has the coldest annual average temperatures and lowest annual precipitation of the three areas (USDA-NRCS 2012*a*-*c*).

The dataset of lakes used to build the classification excluded 25% of lakes with the most heavily developed watersheds. However, the resulting lake regions reflect population distribution patterns in Maine, most notably in the Coastal region (Figure 2.2). The most densely settled areas in Maine are near the coast, which is reflected in the population density of the Coastal region (43.8 people km^{-2} per 2010 Census). Populations are more sparsely distributed in the Inland and Northern sections of the state (2.9 and 3.1 people km^{-2}, respectively). The population density that has developed in the Coastal region due to fertile land and proximity to coastal and larger riverine waters is likely compounding the influences of the natural landscape on lake water quality.

Watershed Quality Index and Condition Classes

The WQI was based on TotUnd because this metric was the strongest predictor of lake water quality condition based on linear regression models of watershed condition (Table 2.1). The TotUnd value was likely the strongest correlate for a statewide assessment because it excludes the most prevalent anthropogenically-created land cover types. The urban and residential development classes (High, Medium, and Low Intensity; Roads) address human activities in densely populated areas of the state, while Crops, Hay, Clearcut, and Blueberry land cover classes represent the agricultural and forestry practices that are most influential on water quality in the less-populated areas.

Our condition classes revealed significant differences in SpCon and TP between reference and altered classes for most lake types at α=0.05 (Figure 2.4). Coastal lakes showed significant differences for SpCon between all reference, intermediate, and altered condition classes, demonstrating the effects of high watershed development in this densely populated region of Maine, especially on deeper, larger lakes. Significant differences in water quality existed between altered and reference lakes for TP and SpCon in Coastal deep lakes, Coastal shallow lakes, Inland deep lakes, and Northern lakes, indicating that suitable reference lakes exist in these areas (Figure 2.4). A low number of Inland lakes (especially Inland shallow lakes) and generally lower concentrations of human urban or agricultural land development in this region likely led to non-significant differences in water quality values across watershed condition classes (Figure 2.4). This region has low population and little agriculture due to its mountainous terrain and less fertile soils. The Northern region shows a wide range in water quality values (Figures 2.3, 2.4; Table 2.6). This is partially due to calcareous geology and agricultural activity in the eastern portion of this area, resulting in high lake alkalinity and nutrient (N and P) concentrations. In contrast, surface waters in the more remote western portion of this area have lower alkalinity and little land use pressure other than timber harvesting (Griffith et al., 2009).

The methods to determine thresholds of watershed development to define reference and altered conditions mirror the "ambient distribution" approach proposed in Stoddard et al. (2006), except that we used the ambient watershed condition (land use) to predict water quality values indicative of most- and least-disturbed conditions, rather than using ambient nutrient concentrations to predict biological condition. Herlihy et al. (2013) defined reference lakes for the U.S. National Lake Assessment through a process of screening water quality, land use, and aerial photographs. While that approach was appropriate for that study, we did not include water quality data in our condition classification, as that would have created an assumption of the expected natural ranges of water quality parameters for the various lake types. Despite differing approaches, our WQI thresholds of ≥90th percentile for reference lakes and <25th percentile for altered lakes are similar to the results of Herlihy et al. (2013), who defined site quality based on disturbance gradients of the least disturbed 15-25% and most disturbed 20-30% of sites, respectively.

Further work to refine and validate reference and altered lake designations could help strengthen water quality threshold values for the condition classes. For example, a screening process using dated aerial photography, as used by Herlihy et al. (2013), could provide verification of our condition classes based on current watershed condition at the time of lake surveys. This may be especially relevant for reference lakes in areas of Maine facing timber harvesting pressure, which can occur quickly over large portions of watersheds and have detrimental effects on lake water quality if managed poorly (Ahtiainen 1992, Steedman 2000, Wilkerson et al. 2010).

Conclusions

The HGM classification framework in this study is based on large-scale landscape patterns, defined by aggregated Level IV Ecoregions and the maximum depth of individual lakes. Both attributes influence the condition of lakes and water quality indicators. The distinctions used in this lake classification are easily made – ecoregional boundaries are well established, and maximum lake depths are easily measured. Even though lake depth may change episodically in lakes with managed water levels, best professional judgment can place any lake into deep or shallow depth categories.

The goal of this classification was to build a framework with large-scale landscape features so that data may be applied consistently across a large region with a variable local landscape. However, it does not directly incorporate data on multiple factors that influence lakes at local scales but are much more resource-intensive to acquire at large scales, such as residence time, lake volume, sediment chemistry, susceptibility to anoxia, or human development on the immediate shoreline. The addition of such parameters may help refine future lake classifications if they can be applied at large scales.

Maine's lakes may differ from lakes in other regions in several ways. Most lakes in Maine were formed naturally by glacial processes. While many lakes are dammed to enhance and regulate water levels, there are relatively few lakes with substantial water level fluctuations. There are no Maine lakes with permitted effluent point-source discharges, except two federal fish hatcheries and one waste water treatment facility in southern Maine. Other lake classification studies may have issues that were not necessary to consider here, such as lake-level management, water scarcity issues, intensely urbanized watersheds, or N-limited rather than P-limited productivity.

While the HGM variables and water quality response parameters used in other lake classifications may differ, lake management needs and local water quality laws will dictate appropriate variables for analysis in other studies. In addition, relevant water quality response variables will differ according to classification goals. Regardless of HGM and response variables used, any variables can be tested for effectiveness in lake classification development using the approach described here.

Establishing condition classes within lake types using measures of watershed development most relevant to the study region can help to inform water quality management decisions, such as comparing water quality values to an expected reference condition. In jurisdictions with multiple lake water quality class designations, this approach may help to define numeric water quality condition criteria for each class. As environmental agencies proceed towards developing biological criteria for lakes (e.g., with aquatic macrophytes, macroinvertebrates, or sediment diatoms), the availability of condition classes that are reflective of watershed condition and supported with descriptive water quality measures will be highly important in the development of those criteria.

CHAPTER 3. PREDICTING ANOXIA IN LOW-NUTRIENT TEMPERATE LAKES

Chapter Abstract

Absence of dissolved oxygen (anoxia) in the hypolimnion of lakes can eliminate habitat for sensitive species and may induce the release of sediment-bound phosphorus. Lake anoxia generally results from decomposition of organic matter, which is exacerbated by high nutrient loads. Total Phosphorus (TP) in lakes is regulated by static aspects of the lake's watershed, but lake TP can be readily increased by human activities. In some low-nutrient lakes, basin morphometry may induce naturally-occurring anoxia. The occurrence of natural anoxia is especially important to consider in lake water quality assessments that compare observed conditions to expected reference conditions. To investigate the occurrence of natural vs. anthropogenically-influenced anoxia, we constructed a logistic regression model to calculate the probability of low-nutrient lakes (TP < 15 µg/L) developing aerial anoxic extent ≥ 10% by testing the predictive potential of variables related to basin morphometry, depths of lake thermal strata, epilimnetic TP, and dissolved organic carbon (DOC). Maximum lake depth and the proportion of lake area under the top of the metalimnion were the most important variables to predict the likelihood of hypolimnetic anoxia, which correctly predicted anoxic condition in 84% of lakes (Model 1). Adding TP as a third variable to Model 1 produced a significantly improved model (Model 2) but the prediction success rate was comparable (86%). We also present a model for lakes with limited bathymetric data, which predicts anoxia with 81% accuracy based on maximum lake depth and mean thermocline depth at peak stratification. DOC was relatively low (mean = 4.3 ± 1.5 mg/L) in the study lakes and its inclusion did not improve model performance. In Model 1, lakes with an anoxic extent ≥ 10% of lake area had significantly higher epilimnetic TP than lakes with oxic hypolimnia, regardless of prediction category or success. Our results indicate that including TP as a variable helps refine models based on morphometry alone, but lake morphometry and stratification dynamics are the most

important factors in the development of anoxic extent in low-nutrient temperate lakes. Our approach informs studies concerned with identifying key factors that influence regime shifts in a variety of ecosystems.

Introduction

Depleted dissolved oxygen (DO) concentrations in lake hypolimnia can diminish or eliminate available habitat and induce phosphorus (P) release from lake sediment into the overlying water through the reduction of oxidized iron (Fe) secondary phases (Wetzel 2001). Hypolimnetic DO declines when the respiration rate from aerobic decomposition of organic matter exceeds the oxygen replenishment rate from primary production or lake mixing. Lake anoxia (DO < 2 mg/L) is exacerbated in lakes with high nutrient concentrations; excess P promotes algal growth which adds to the detrital load and promotes anoxic conditions. In highly colored (dystrophic) lakes, elevated concentrations of dissolved organic carbon (DOC) can cause stronger thermal stratification and increased DO depletion from greater isolation of bottom waters, heightened decomposition, and increased DOC photo-oxidation rates (Hutchinson 1957, Knoll et al. 2018). Lakes free from extensive watershed development may exhibit seasonal anoxia as a result of morphometry and acidic DOC exported from watershed wetland areas (Nürnberg 2004). Lake basin morphometry may be a primary driver of both DO depletion rates and the areal extent of anoxia in some low-nutrient lakes (Molot et al. 1992). TP, which we used as a filter for lake selection, is regulated by static aspects of the lake's watershed that include bedrock and surficial geology, topography, and hydrology. However, human activity can easily alter the flux of TP from the watershed to the lake, inducing cascading and interrelated effects on lake ecology. It is necessary to determine if the drivers for lake anoxia are natural features of the system (i.e., lake basin morphometry) or due to increased nutrient input (potentially from anthropogenic sources) in order to make sound assessments of lake condition and appropriate management decisions. However, this is not always possible due to limited data availability.

Many studies have created models to predict the DO depletion rate in lake hypolimnia, in the general form of Areal Hypolimnetic Oxygen Demand (AHOD; Strom 1931). Most of these models involve a combination of measures related to nutrient load and lake morphometry (Hutchinson 1957, Reckhow 1977, Welch and Perkins 1979, Charlton 1980, Cornet and Ringler 1980, Vollenweider and Janus 1982, Reckhow and Chapra 1983, Clark et al. 2002, Rippey and McSorely 2009, Müller et al. 2012, Schwefel et al. 2018). Other models have combined morphometry and nutrient load with hypolimnetic water temperature (Cornett and Rigler 1980), mean DO at spring turnover (Livingstone and Imboden 1996), or days from onset of summer stratification (Yuan and Jones 2020). However, studies that evaluate AHOD measure the net DO consumption rate until anoxia sets in, and not the areal extent of anoxic conditions after they develop fully following peak stratification. The areal extent of anoxia is the areal portion of the lake sediment covered with anoxic water, which has implications for the extent of sediment P release and deep-water or benthic habitat suitability. DO depletion rates are not correlated with the extent or duration of anoxia in a lake, as demonstrated by Nürnberg (1995).

Reckhow (1977) developed a model that calculated the probability of anoxia based on lake morphometry (mean depth), hydrology (annual water load), and external TP load. The model predicted anoxia to occur when water load was small and external TP load was high. As a result, the model underestimates anoxia in low-nutrient lakes which may be anoxic for reasons other than high productivity (Nürnberg 1995). Molot et al. (1992) successfully predicted hypolimnetic DO concentrations at the end of summer (i.e., during peak stratification, when the thermal gradient between the epilimnion and hypolimnion is greatest) using lake morphometry, TP concentrations, and mean DO at spring turnover. They found that the lake volume to sediment area ratio was especially important in predicting the end-of-summer DO profiles in oligotrophic and oligo-mesotrophic lakes < 20 m deep. However, spring DO profiles can be difficult to collect or model, and many lake managers need a more straightforward approach to evaluate risk of natural anoxia. Additionally, neither of these models were intended to predict the expected areal extent of anoxia.

Nürnberg's (1995) Anoxic Factor (AF), defined as the ratio of temporal and spatial extent of anoxic sediment to the lake surface area, is highly effective at quantifying the extent and duration of anoxia, but requires sequential collection of DO profiles for an accurate calculation. The need for weekly to monthly DO profiles for adequate resolution (Nürnberg 2004) may limit the utility of the AF, especially in regions with many under-surveyed lakes. Modeling AF using nutrient concentration and lake morphometry is an option (Nürnberg 2004), but a simpler approach to determine expected conditions with a predictive model may be advantageous in some circumstances. This is generally the case for assessments that require comparisons to natural conditions, particularly if datasets from many remote lakes are necessary to establish a reference dataset. A simple predictive model may be a beneficial alternative where such sampling efforts are not possible.

In Maine lakes, anoxia has been observed in 647 of the 951 surveyed lakes (ME DEP, 2015) during peak stratification (defined here as 1 August – 7 September). However, most Maine lakes have relatively low epilimnetic TP concentrations during peak stratification (median = 9.0 µg/L, 25th percentile = 6.5 µg/L, 75th percentile = 13.0 µg/L). Thus, there may be many Maine lakes experiencing natural, as opposed to anthropogenically induced, hypolimnetic anoxia during peak stratification due to factors unrelated to nutrient concentration. Much limnological research focuses on impaired or eutrophic systems, while the majority of the world's lakes exist at higher latitudes (45°-75°N; Verpoorter et al. 2014) and are more likely to be nutrient limited (Abell et al. 2012). Maine lakes (43.0°- 47.3°N) are likely more similar to those of Canada and Scandinavia, which exist in similar climates, were glaciated, and many of which have largely undeveloped watersheds.

Many lake water quality assessments are predicated based on the comparison of observed condition to natural (or reference) condition (Stoddard et al. 2006, Herlihy et al. 2013). Given the importance of anoxia to overall ecosystem health and the variety of factors that influence its manifestation, it is necessary for comprehensive lake assessments to determine which lakes are predisposed to anoxia from natural conditions and which lakes are anoxic due primarily to excess nutrient concentration from anthropogenic activities. Most studies modeling DO depletion rate or the extent of anoxia in lakes combine measures of lake morphometry and nutrient concentration, which may not apply to low-nutrient lakes where anoxia may not be strongly influenced by the nutrient concentration.

Our goal was to develop a predictive model that uses morphometric and stratification variables to estimate the probability of natural anoxia occurring in low-nutrient (TP < 15 µg/L) Maine lakes. More broadly, we seek to identify the most influential abiotic variables in the expression of a key ecosystem attribute that result in abrupt shifts in ecosystem structure and function. This is especially relevant when considering the role of natural vs. anthropogenically-influenced factors that affect habitat availability in pristine environments.

Methods

Study Area and Data Sources

Maine (Figure 3.1A) has an area of 91647 km^2. Elevation ranges from sea level to 1608 m. The entire state was glaciated by the Laurentide Ice Sheet until between 15000 and 10000 years ago, south to north (Thompson and Borns 1985). The climate is north-temperate. The present Köppen climate zone is humid continental (Dfb; Beck et al. 2018), characterized by cold snowy winters, warm summers, and precipitation distributed throughout the year (Kottek et al. 2006). There are nearly 2800 lakes > 4 ha. Population density is highest in the southern and coastal areas and near the mainstems of larger rivers.

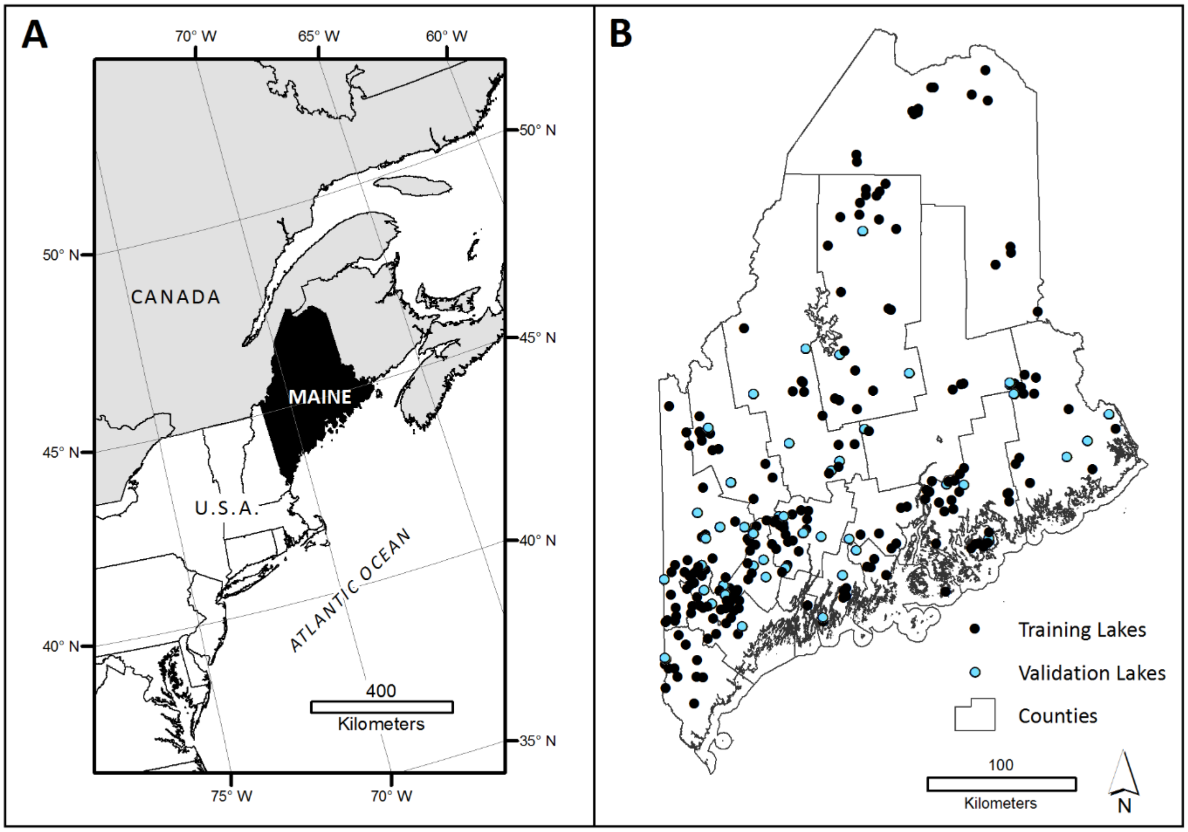


Figure 3.1. (A) Location of Maine (solid black), and (B) location of randomly selected training (n = 192, 80%) and validation (n = 43, 20%) lakes in model-building dataset.

Peak stratification temperature and DO profile data from 1989 to 2018 (n ≥ 3 y) were compiled from the deepest location of each lake. This time period captures the period of peak stratification in Maine lakes that ends before thermoclines start to erode, typically in late August to mid-September, north to south. This data screening step produced a dataset with 8070 profiles from 414 lakes (Maine Lakes Database, ME DEP 2015; Figure B.1). Profile data were typically taken from just below the water's surface, every meter through the epilimnion and metalimnion, every other meter after 15 m, and every 5 m below 25 m. If anoxia (DO < 2 mg/L) was observed, then one-meter profiles extended through the entire water column. Deepest measurements were 0.5 - 1 m above the lake bottom.

Lake chemistry data were obtained from ME DEP (2015). TP and DOC samples were collected from an integrated epilimnetic core sample during peak stratification. TP and DOC are included in our statistical analysis (if sample years between 1989-2018 were $n \geq 3$) to determine the role of nutrients and organic matter in lake anoxia (Knoll et al. 2018). While chlorophyll-*a* serves as a more direct measure of biological activity, TP was used here to facilitate comparisons to related studies that used TP as a measure of biological activity related to DO loss in lakes (Reckhow 1977, Molot et al., 1992, Clark et al. 2002, Rippey and McSorely 2009, Schewefel et al. 2017). In the current analysis, the use of TP and chlorophyll-*a* as measures of productivity were interchangeable, as each parameter produced similar model results.

Data Processing

Depths of the thermocline (horizontal plane of greatest temperature differential between adjacent layers) and the top and bottom of the metalimnion (depth interval with greatest water temperature gradient between the epilimnion and hypolimnion) were calculated for each profile in R (R Core Team 2019) using the rLakeAnalyzer package (Winslow et al. 2019). Rather than using only differences in water temperature between profile measurements to delineate stratification layer boundaries, rLakeAnalyzer calculates the position of strata based on temperature-derived water density estimates. The algorithm looks for the maximum change in water density between measurements and magnitudes of differences between adjacent depths to estimate depth values, rather than using only the difference in temperature between measured depths (Read et al. 2011). Lake profiles that did not have a developed thermocline (maximum temperature differential < 1°C below 1 m depth) were excluded from the analysis. Depths of the thermocline and the top and bottom of the metalimnion were averaged for each year in each lake, then a grand mean was calculated for the average position of each layer from each profile across all years in each lake.

Below the anoxia threshold (< 2.0 mg DO/L), sediment P that is bound to redox-sensitive Fe(III) compounds may be released into the water column, although the presence of adequate $Al(OH)_3$ in the sediment can limit the release of P even under anoxic conditions (Kopáček et al. 2005, Lake et al. 2007, Wilson et al. 2008). The shallowest depth of anoxia that extended to the lake bottom was determined for each profile.

Bathymetric data were obtained from the Maine Office of GIS (MEGIS 2011). This dataset is a collection of data from multiple sources. Most data were collected in the mid- to late-1900s and consist of depth measurements manually collected along transects. Data resolution varies by lake surface area (A_0), ranging from 1.3 points/ha for small lakes (A_0 ~ 4 ha) to 0.1 points/ha for larger lakes (A_0 ~ 400 ha). Despite the coarseness of the data compared to the current technique of GPS-enabled sonar-captured bathymetry, the dataset includes data on such a large number of Maine lakes that it has substantial utility.

Lake volumes were calculated using a model developed in ArcMap 10.6.1 (Esri 2017) for low-resolution bathymetric data. This model most closely approximates volume calculations derived using high-resolution (average 2.5 points per ha) bathymetry from 14 Maine lakes (D. Buckley, *unpublished data*). Raster files were generated from the MEGIS bathymetric data using the Create IDW geoprocessing tool in ArcMap (settings: cell size = 10 m, power = 5, point search radius = 12). Zero-depth points were created along shorelines so as to not overweight the shallow water bathymetry points. The raster output was then clipped to the lake boundary using the Extract by Mask tool with lake polygons from the National Hydrography Dataset (USGS et al. 2013). Lake strata areas and volumes were derived in 0.1 m depth increments using the raster package in R (Hijmans 2019).

Average depths of the thermocline and metalimnion (top and bottom) were associated with the area and volume values for the corresponding stratum for each lake based on bathymetry data. These values were combined with other basic lake morphometric measurements (total surface area, total volume, maximum depth) to create 34 discrete morphological variables for testing in the logistic regression model (Table 3.1). Values for fetch (length of lake along the axis of prevailing winds – SW to NE during the maximum stratification period), MaxL (longest span of open water regardless of cardinal direction) and MaxW (longest span of open water perpendicular to MaxL) were calculated with the lakemorpho package in R (Hollister and Stachelek 2017).

Only lakes with maximum depth (z_{max}) ≥ 10 m were used in the model-building dataset, excluding shallow polymictic lakes that would likely not exhibit anoxia without substantial nutrient enrichment (median thermocline for study lakes = 6.8 m; Table 3.1). Maximum depth was used as a screening criterion over a direct measure of stratification (such as Schmidt Stability; Schmidt, 1928) to widen the applicability of the model to lakes without detailed bathymetry or historic profile data. Lakes with mean epilimnetic TP > 15 µg/L during the peak stratification period were excluded from the analysis to limit the role of nutrient enrichment in the development of hypolimnetic anoxia. This value is supported by the model created by Molot et al. (1992) for oligotrophic and oligo-mesotrophic lakes, which predicts that lakes with epilimnetic TP > 15 µg/L will have depleted hypolimnetic DO at the end of summer. The final model-building dataset consisted of 235 lakes (Figure 3.1B).

The depth of anoxia that was consistent to the bottom (thereby excluding DO minima with oxic conditions beneath) and the corresponding bathymetric stratum of each lake were used to calculate the potential maximum areal coverage of peak stratification anoxic water. Model results would likely be strengthened if anoxic extent were based on continuous profile data, so more accurate maxima may be determined, but this was not an option in this large-scale study that uses data from a large number of lakes.

Table 3.1 Variables tested in logistic regression model building analysis, with minimum, median and maximum values from the dataset (n = 235 lakes). Units are noted in description column. Descriptive statistics for variables that change seasonally are for peak stratification only.

Variable	Description	Minimum	Median	Maximum
z_{max}	Lake maximum depth (m)	10.00	17.10	96.30
$\bar{z}$	Lake average depth (volume/area) (m)	1.52	5.20	31.13
A_0	Lake total surface area (m^2)	2.48×10^4	1.93×10^6	3.05×10^8
V_0	Lake total volume (m^3)	7.86×10^4	1.03×10^7	5.50×10^9
AV_0	Lake area/volume ratio	0.03	0.19	0.66
$Therm_{\bar{x}}$	Avg. depth of peak stratification thermocline (epilimnion thickness) (m)	2.60	6.80	17.40
ThermV	Relative volume of water under average thermocline depth (% of V_0)	1.26	70.23	97.29
ThermA	Relative area of lake under average thermocline depth (% of A_0)	0.43	35.43	74.90
ThermAV	Area/Volume ratio of lake under thermocline depth	0.02	0.09	0.22
$Therm_Z$	Thickness of hypolimnion below thermocline (m)	2.20	9.80	85.60
$Therm_{Z\%}$	Thickness of hypolimnion below thermocline (% of z_{max})	16.78	60.63	88.89
$MetaT_{\bar{x}}$	Average depth of top of peak stratification metalimnion (m)	0.70	4.90	15.10
$MetaT_V$	Relative volume of water under top of metalimnion (% of V_0)	3.62	84.82	99.54
$MetaT_A$	Relative area of lake under top of metalimnion (% of A_0)	1.03	49.29	79.43
$MetaT_{AV}$	Lake Area:Volume ratio under top of metalimnion	0.03	0.10	0.29
$MetaT_Z$	Thickness of hypolimnion below top of metalimnion (m)	3.80	11.80	87.00
$MetaT_{Z\%}$	Thickness of lake below top of metalimnion (% of z_{max})	0.27	0.74	0.95
$MetaB_{\bar{x}}$	Average depth of bottom of peak stratification metalimnion (m)	3.30	9.50	18.60
$MetaB_V$	Relative volume of water under bottom of metalimnion (% of V_0)	0.18	44.89	95.22
$MetaB_A$	Relative area of lake under bottom of metalimnion (% of A_0)	0.03	17.38	71.54
$MetaB_{AV}$	Area/Volume ratio of lake under bottom of metalimnion	0.02	0.08	0.20
$MetaB_Z$	Thickness of hypolimnion below bottom of metalimnion (m)	0.60	7.20	83.70
$MetaB_{Z\%}$	Thickness of lake below bottom of metalimnion (% of z_{max})	4.20	44.17	86.92
$SS_{\bar{x}}$	Average Schmidt Stability (J/m^2)	1.20	107.60	3900.70
$\bar{z}$:z	Average depth/Maximum Depth ratio	0.11	0.29	0.61
SDR	Shoreline development ratio: (perimeter/$2\sqrt{(\pi \times A_0)}$)	1.09	2.11	10.98
SDI	Shoreline development index: (perimeter/A_0)	0.00	0.01	0.03
Cir	Comparison to perfect circle: ((perimeter/MaxL)/π)	0.36	1.25	28.04
Fetch	Length of open water along NE (45°) direction (km)	0.16	1.32	18.18
MaxL	Longest span of open water regardless of cardinal direction (km)	0.24	2.61	17.93
MaxW	Longest span of open water perpendicular to MaxL (km)	0.15	1.10	17.46
L:W	MaxL / MaxW	0.42	2.16	11.77
D_V	Volume development ratio (Hutchinson 1957): $((3 \times \bar{z})/z_{max})$	0.34	0.88	1.84
OI	Osgood Index: $\bar{z}/\sqrt{A_0}$ (m/km)	0.83	3.96	32.31
TP	Total peak stratification epilimnetic phosphorus concentration (µg/L)	2.66	7.00	14.77
DOC	Peak stratification Dissolved Organic Carbon concentration (mg/L)	1.67	4.00	10.50

Logistic Regression Model

Logistic regression produces a model that predicts the occurrence of a binary response variable (response vs. non-response). With binary data, the values of the response variables do not follow a linear trend, and errors are not normally distributed across the entire range of data (Peng et al. 2001). In logistic regression, the logit transformation on the dependent variable (*Y*) rectifies these issues by using an ordinary least squares regression. Logistic regression was used here because it does not assume a linear relationship between the binary response as do other regression techniques and some empirical models. We do not assume that the extent of lake anoxia has a linear relationship to the variables contributing to its development, especially as hypolimnetic DO approaches zero. Hypolimnetic DO depletion rate is asymptotic rather than linear over time (Yuan and Jones 2020), and it follows that factors causing the development of hypolimnetic anoxic area may be nonlinear as well. Additionally, our dataset includes many lakes (~25%) with no recorded peak stratification anoxic conditions. Because we are interested in why some lakes remain oxic throughout peak stratification, and what factors may predispose a lake to a naturally occurring anoxia, using logistic regression to predict the binary condition of oxic or anoxic conditions during peak stratification was appropriate. More details on the method are in Appendix B.1.

We focused on the extent of areal contact between anoxic water and lake sediment to investigate the role of lake morphometry in DO consumption in lake hypolimnia and set the binary response to "1" if a lake had an anoxic hypolimnion exceeding 10% of lake surface area at least once in the ≥ 3 years of profiles. This value was chosen because it created roughly the same number of responses (1, becomes anoxic at least in one year) as non-responses (0, remains oxic in all years). The 235 lakes were randomly separated into 80% training (n = 192, 112 with response) and 20% validation (n = 43, 20 with response) datasets.

Logistic regression models were created with the glm function in R (family=binomial, link=logit). Models were made for each of the 34 morphometric variables plus TP, and also in combinations of two variables for a total of 630 models. Models combining three or more variables did not increase model performance and are not presented here. Models were excluded from consideration if the variables were correlated at a threshold of $R > |0.6|$. Allison (2012) suggests that problems with multicollinearity in logistic regression occur when $R > |0.78|$, or $R^2 > 0.6$.

Model statistics were calculated with the compareGLM function of the rcompanion package in R (Mangiafico 2016). The models were ranked for performance based on the corrected Akaike Information Criterion (AICc), which provides an estimation of model strength based on log-likelihood and the number of included parameters. AICc is corrected for small sample sizes and accounts for model fit and complexity. ΔAICc provides the difference in AICc values from each model to the strongest (lowest AICc value) model. McFadden, Cox and Snell, and Nagelkereke pseudo R-squared statistics, calculated with compareGLM, were also used in model comparisons.

The final model was compared to the null model (no variables) and to the final model + TP with likelihood-ratio tests using the lmtest package in R (Zeileis and Hothorn 2002). These comparisons were made for two reasons: 1) to verify that the final model was significantly different from the null model, and 2) to see if adding TP to the final model significantly alters model performance. This would indicate whether the anoxic extent development is affected by nutrient concentration rather than occurring primarily as a function of lake morphometry specifically in low-nutrient lakes. We also present the top-ranking model that did not use volume-derived or area-derived metrics to test the application of our approach to lakes with limited bathymetric data.

Top-ranking models were evaluated on both training and validation datasets with the following additional metrics to evaluate model performance: percent correctly classified for all lakes, sensitivity (percent of anoxic lakes correctly classified), specificity (percent of oxic lakes correctly classified), and area under the receiver operating characteristic (ROC) curve, which is a measure of the relationship between false positive rates (1 – specificity) and sensitivity (Hein et al. 2012). The Area Under Curve (AUC) value measures the area of the ROC indicating correctly classified cases. An AUC approaching 1 indicates a more successful model (AUC = 1.0 indicates 100% accuracy), while an AUC value of 0.5 (50% accuracy) indicates the model predicts no better than random chance. The mean probability of response ($\pi_{\overline{x}}$) was used as the probability threshold in prediction calculations because it maximized model performance.

Diagnostic statistics were checked to verify that assumptions of logistic regression were met in all models presented here. Associations among variables and the calculated logit values were visually inspected for linearity. The standard residual error of all lakes was checked to ensure that no values were > 3, which would indicate outlier data points. Variance inflation factors (VIF) were calculated for each model to check for multicollinearity; VIF > 5 indicates problems with correlation among model variables (Fox and Monette 1992).

Results

Logistic Regression Model

The variables maximum depth (z_{max}) and relative lake area (percentage of total) below the top of the metalimnion ($MetaT_A$) accounted for the best model to predict anoxia in low-nutrient Maine lakes (AICc = 178.0; Table 3.2). This model (Model 1) was significantly different from the null model ($F = 91.0$, $p < 0.001$). Model prediction success metrics are summarized in Table 3.3. Model 1 shows that the likelihood of anoxia increases with decreasing z_{max} and increasing relative lake area below the top of the

Table 3.2. Top-ten ranking models and fit statistics from logistic regressions on training dataset (80% of dataset lakes, n = 192), plus the top-ranking model using no bathymetry-derived morphometric variables (z_{max} + $Therm_{\bar{x}}$). All models df = 191. Parameter definitions in Table 3.1. All lakes are $z_{max} \geq 10$ m and peak stratification epilimnetic TP < 15 µg/L.

Model Parameters	Rank	McFadden R^2	Cox and Snell R^2	Nagelkerke R^2	AIC	AICc	ΔAICc
z_{max} + $MetaT_A$	1	0.349	0.377	0.508	177.8	178.0	0.0
z_{max} + $MetaT_V$	2	0.340	0.370	0.498	180.2	180.4	2.4
z_{max} + ThermA	3	0.324	0.356	0.479	184.3	184.6	6.6
ThermAV + $\bar{z}$:z	4	0.324	0.356	0.479	184.3	184.6	6.6
ThermAV + Dv	5	0.323	0.355	0.478	184.6	184.8	6.8
$MetaT_V$ + $MetaT_Z$	6	0.322	0.354	0.477	184.8	185.1	7.1
$MetaT_V$ + ThermZ	7	0.316	0.349	0.470	186.3	186.5	8.5
$MetaT_V$ + $MetaB_Z$	8	0.314	0.347	0.467	186.9	187.1	9.1
z_{max} + ThermV	9	0.311	0.344	0.463	187.8	188.0	10.0
$MetaT_A$ + $MetaB_Z$	10	0.311	0.344	0.464	187.8	188.0	10.0
z_{max} + $Therm_{\bar{x}}$	29	0.289	0.325	0.438	193.5	193.3	15.6

Table 3.3. Prediction success metrics for training and validation datasets from models based on lakes with TP < 15 µg/L (Models 1 and 2), TP < 10 µg/L (Models 3 and 4), and lakes without bathymetric data (Model 5). Table Legend: TP Criterion = maximum values for total phosphorus used in lake screening; Subset = training (T, 80%) or validation (V, 20%) datasets; n = number of lakes in each subset; % Correct = percentage of lakes correctly classified with the model; Sensitivity = percent presences (anoxia present) correctly classified, Specificity = percent absences (anoxia absent) correctly classified; AUC = percent Area Under Receiver Operating Curve. Equation variables: $\log_{10}(z_{max})$ is the log of the maximum lake depth in m; $\boldsymbol{sin^{-1}}\sqrt{\mathbf{MetaT_A}}$ is the arcsin square root of the percent of lake area under the top of the metalimnion at peak stratification, log10(TP) is the log of the mean epilimnetic total phosphorus concentration in µg/L at peak stratification, and $\mathbf{log_{10}(Therm_{\bar{x}})}$ is the logarithm of the mean peak stratification thermocline depth in m.

Model Number	TP Criterion (Max. Value)	Equation	AICc	Subset	n (lakes)	% Correct	Sensitivity	Specificity	AUC
1	15	Logit = 10.028 – 10.423*$\log_{10}(z_{max})$ + 5.0207*$\sin^{-1}\sqrt{\mathrm{MetaT_A}}$	178.0	T	192	80.2	81.3	78.8	86.3
				V	43	83.7	85.0	82.6	92.4
2	15	Logit = 1.140 – 0.196*$\log_{10}(z_{max})$ + 5.953*$\sin^{-1}\sqrt{\mathrm{MetaT_A}}$ + 7.496*$\log_{10}$(TP)	159.2	T	192	80.2	82.1	77.5	89.6
				V	43	86.0	85.0	87.0	93.9
3	10	Logit = 9.171– 9.468*$\log_{10}(z_{max})$ + 4.448*$\sin^{-1}\sqrt{\mathrm{MetaT_A}}$	158.4	T	158	79.7	82.3	77.2	86.0
				V	37	83.8	90.9	73.3	89.7
4	10	Logit = -2.718 – 8.358*$\log_{10}(z_{max})$ + 4.657*$\sin^{-1}\sqrt{\mathrm{MetaT_A}}$ + 11.821*$\log_{10}$(TP)	135.9	T	158	83.5	86.1	81.0	90.1
				V	37	81.1	90.9	66.7	93.0
5	15	Logit = 13.775 – 5.969 *$\log_{10}(z_{max})$ - 6.343*$\log_{10}(\mathrm{Therm}_{\bar{x}})$	193.3	T	192	74.5	75.9	72.5	82.2
				V	43	81.4	85.0	78.3	91.5

metalimnion (Figure 3.2). All three coefficients were significant ($p < 0.001$). All model assumptions were met: the logit scale was visually observed to be linear with z_{max} and $MetaT_A$, the standard residual error was < 3 for all lakes, and there was no multicollinearity between variables (VIF < 2).

AICc values ranged from 178.0 for Model 1 to 268.7 (ΔAICc = 90.7) for the lowest-ranked of 630 tested models (ThermA + MetaTZ%). The top ten strongest models have a ΔAICc < 10, indicating that these models may offer similar results compared to the best model presented here. The strongest model that included TP with one morphometry variable ranked 11th (MetaBAV + TP, ΔAICc = 10.1). A model with TP as the only variable was ranked 300th (ΔAICc = 45.3).

Model 1 successfully predicted the correct response in 80.2% of lakes from the training dataset and 83.7% of lakes from the validation dataset. Validation dataset sensitivity, specificity, and ROC were 85.0%, 82.6%, and 92.4%, respectively (Table 3.3).

Inclusion of TP in the model

To examine the interaction of lake nutrients with morphometric variables known to be important to lake anoxia, TP was added to Model 1 as a third variable to create Model 2 (z_{max} + $MetaT_A$ + TP). Diagnostics supported the structure of Model 2: variables were linear with logit function, standard residual error < 3, and VIF <2. The AICc for Model 2 was 161.2, compared to 178.0 for Model 1 (Table 3.2), and the two models were significantly different ($F = 18.6$, $p < 0.001$). This indicates that the addition of TP to the Model 1 significantly increases model strength; however, this addition of TP did not meaningfully enhance prediction rates. The validation dataset for Model 2 had a successful prediction rate of 86.0%, sensitivity of 85.0%, specificity of 87.0% and ROC of 93.9%, values that are slightly higher but comparable to those of Model 1 (Table 3.3).

To further investigate the role of TP in these study lakes, alternative models were developed with a more stringent nutrient threshold of TP < 10 µg/L (Models 3 and 4; Table 3.3). However, despite lower AICc values, the prediction rates for this model were generally comparable to those of Models 1 and 2 which used TP < 15 µg/L (Table 3.3). Consequently, the focus here remains the TP < 15 µg/L models.

Model Without Bathymetry Data

To apply the model to lakes without bathymetric data necessary for areal and volume estimates (Table 3.1), we present the top-ranking model that uses no bathymetrically-derived measures of lake volume or area of lake strata (Model 5, Table 3.3). Model 5 predicts the likelihood of anoxia based on maximum depth of the lake (z_{max}) and the average depth of the thermocline ($\mathrm{Therm}_{\bar{x}}$) (Figure 3.3). All three coefficients were significant ($p < 0.001$) and all regression model assumptions were met. The anoxic conditions of 81.4% of lakes in the validation dataset were correctly predicted with this model, and AUC was 91.5 (Table 3.3). This was the 29th ranked model overall (ΔAICc = 13.5).

Water Chemistry and Prediction Success

Training and validation datasets were combined to compare Model 1 and 2 anoxia predictions against DOC and TP concentrations (Figures 3.4, 3.5). DOC did not show a significant difference among predicted conditions for Model 1 ($F = 0.801$, $p = 0.372$) but did for Model 2 ($F = 4.650$, $p = 0.032$). There were no significant differences in DOC for prediction correctness (Model 1: $F = 0.274$, $p = 0.601$; Model 2: $F = 0.228$, $p = 0.634$), or the interaction between predicted condition and prediction correctness (Model 1: $F = 1.889$, $p = 0.171$; Model 2: $F = 0.158$, $p = 0.691$) (Figure 3.4). Lakes that were predicted to be anoxic had significantly higher TP than lakes predicted to be oxic in both Models 1 and 2 (Model 1: $F = 32.114$, $p < 0.001$; Model 2: $F = 100.650$, $p < 0.001$). TP did not differ significantly among lakes based on prediction correctness alone for Models 1 or 2 (Model 1: $F = 0.254$, $p = 0.615$, Model 2: $F = 0.420$, $p = 0.518$). The interaction effect of predicted condition and prediction correctness was significant for TP in both Models 1 ($F = 35.782$, $p < 0.001$) and Model 2 ($F = 10.550$, $p = 0.001$). Tukey HSD *post hoc* tests for Model 1 revealed that TP concentrations in anoxic lakes were higher than in oxic lakes regardless of predicted condition (Figure 3.5A). For Model 2, lakes that were successfully predicted to be anoxic had significantly higher TP than lakes that were predicted to be oxic, regardless of prediction success. Lakes incorrectly predicted to be anoxic had TP concentrations similar to lakes incorrectly predicted to be oxic but had significantly higher TP than correctly predicted oxic lakes (Figure 3.5B).

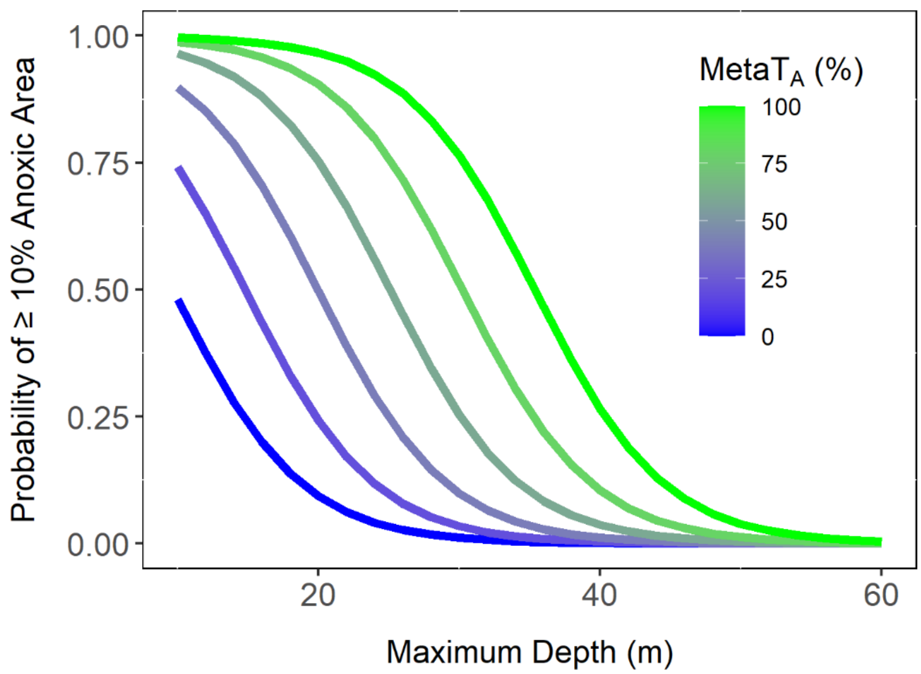


Figure 3.2. Probability of anoxia function calculated by Model 1 (Table 3.3). Maximum Depth = z_{max} in m; $MetaT_A$ = percent of the lake total area beneath the top of the metalimnion (= bottom of the epilimnion).

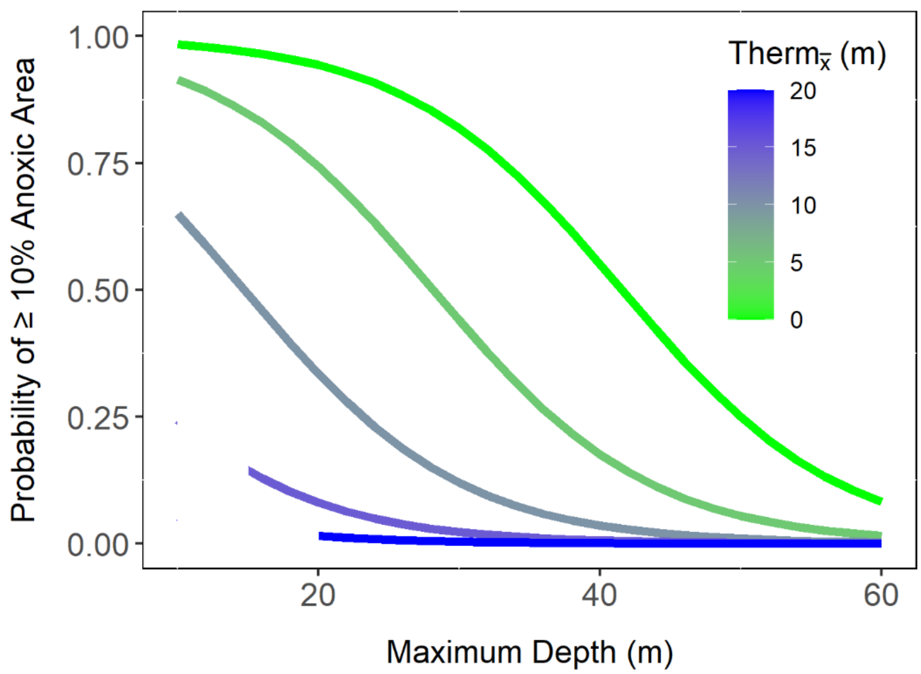


Figure 3.3. Probability of anoxia function calculated by Model 5 (Table 3). Maximum Depth = z_{max} in m; Therm$_{\bar{x}}$ = average depth of peak stratification thermocline in m.

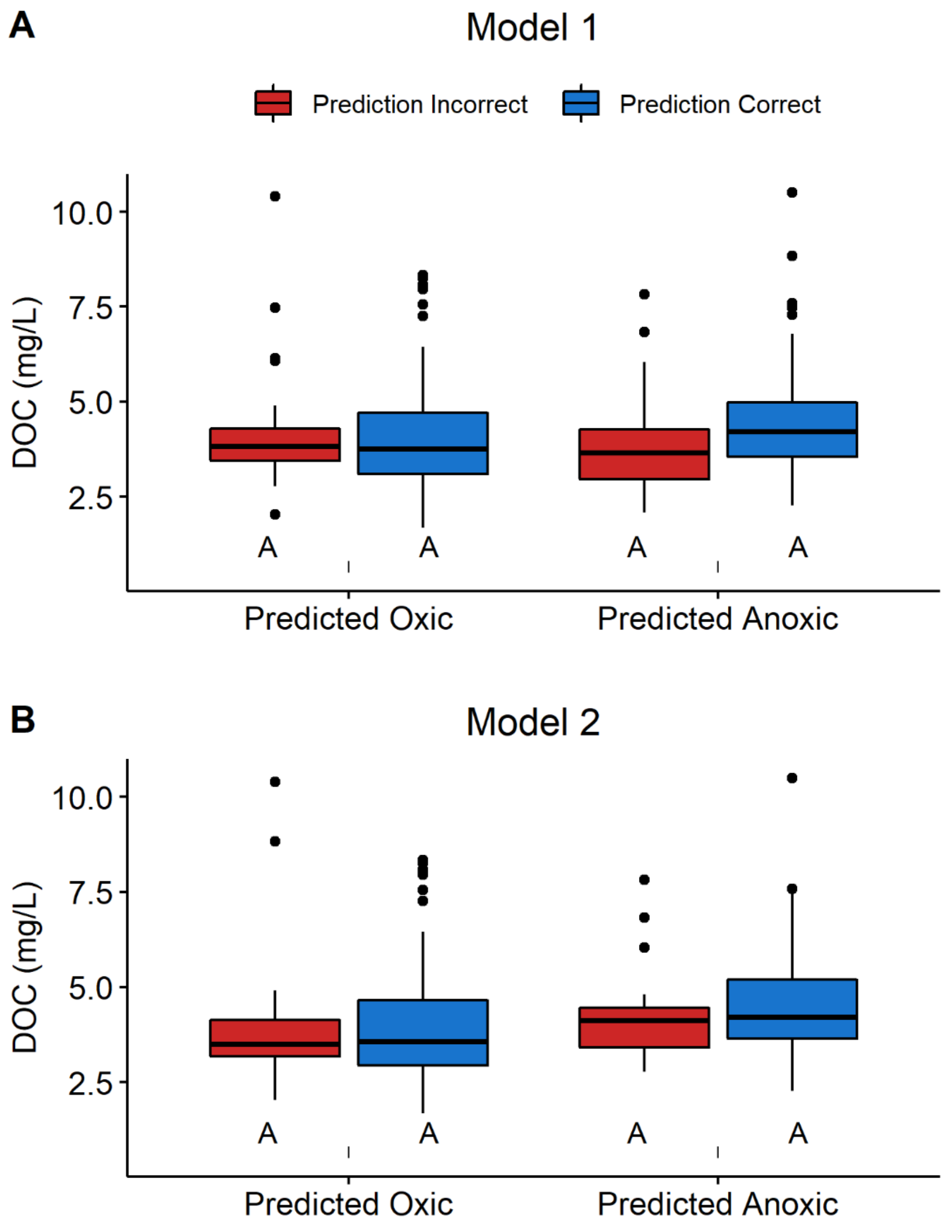


Figure 3.4. (A) Model 1 and (B) Model 2 predictions and outcomes vs mean epilimnetic DOC concentration at peak stratification. Different letters along x-axes indicate significant differences at α = 0.05.

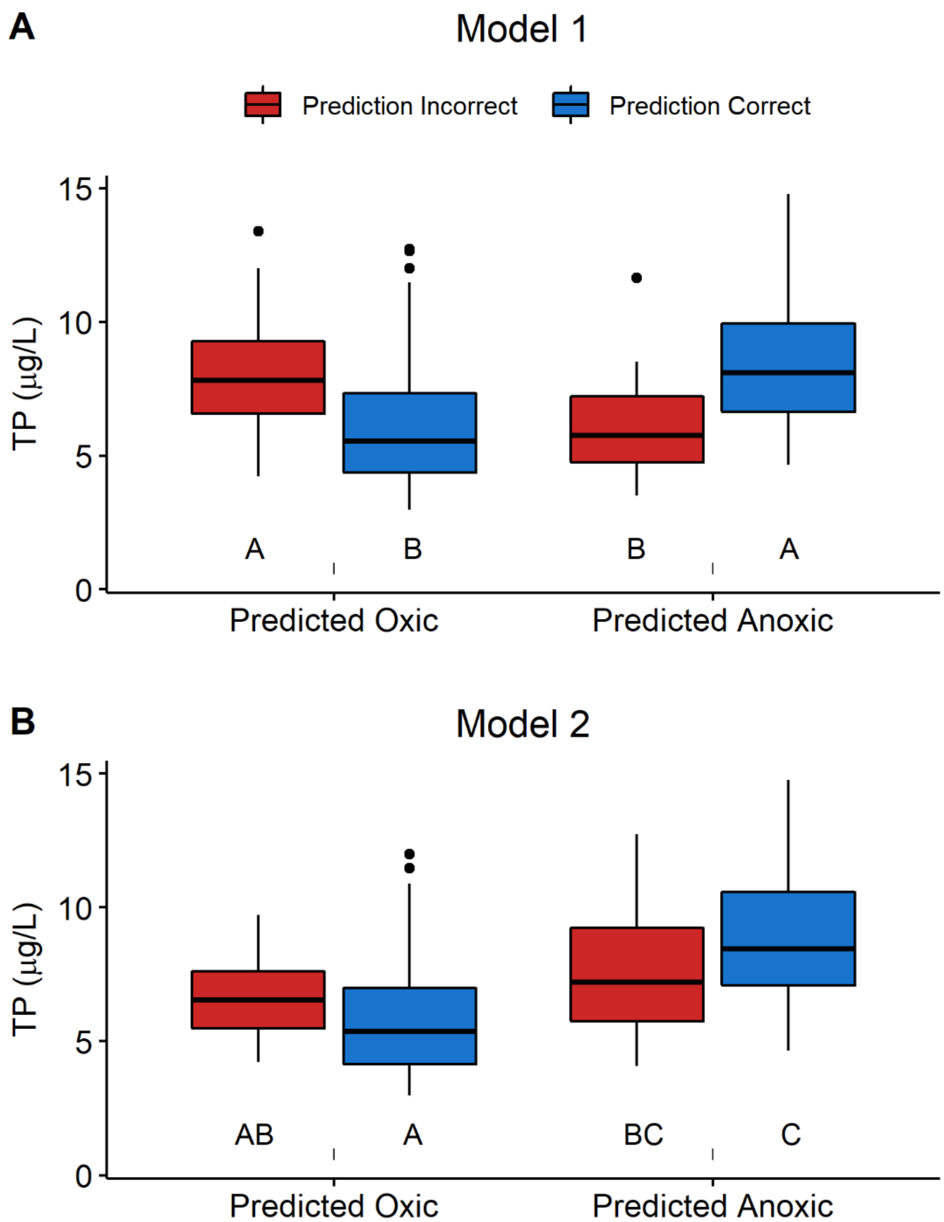


Figure 3.5. (A) Model 1 and (B) Model 2 predictions and outcomes vs mean epilimnetic TP concentration at peak stratification. Different letters along x-axes indicate significant differences at $\alpha = 0.05$.

Discussion

This study shows that the existence of anoxic conditions may be successfully predicted in low-nutrient temperate lakes ≥10 m deep based on morphometry and stratification dynamics alone, without inclusion of a measure of nutrient concentration. While the addition of lake TP concentration to Model 1 slightly increased chances of successful prediction (Model 2), this does not appear necessary for effective identification of lakes that are susceptible to naturally-occurring anoxia. The poor performance of the TP-only model (ΔAICc = 45.3) underscores the greater importance of lake morphometry in the development of anoxia in these low-nutrient lakes. Our approach was effective at identifying principal variables that predict a critical ecological response.

Our diagnostic models predict that the likelihood of hypolimnetic anoxia increases with decreasing z_{max} and an increasing $MetaT_A$, which represents the percent of the total lake area below the epilimnetic mixing zone during the period of peak stratification (Figure 3.2). Maximum depth was correlated to lake volume ($R^2 = 0.516$, $p < 0.001$), which is important because deeper, more voluminous dimictic lakes hold a larger reserve of DO and can better mitigate its loss in the hypolimnion before the autumnal mixing event replenishes DO to the entire water column. Deeper lakes also have sustained colder hypolimnetic water through the summer months because of greater thermal stability, higher DO acquired during spring overturn due to increased solubility, and decreased sediment oxygen demand in colder water. Further, z_{max} was correlated to other lake measures, such as surface area ($R^2 = 0.327$, $p <$ 0.001), TP ($R^2 = 0.221$, $p < 0.001$) and DOC ($R^2 = 0.029$, $p = 0.007$),); all of which affect DO demand in lakes. $MetaT_A$ relates to DO depletion by providing a measure of the areal extent of the lake sediment area in contact with water below the epilimnion. In lakes with large shallow areas where much of the lake sediment is in contact with epilimnetic waters (small $MetaT_A$), DO loss from decomposition of sediment organic matter may be largely replenished from surface mixing in the epilimnion. In contrast, if this shallow area of the lake is proportionally small (large $MetaT_A$), the larger isolated portion of the lake will undergo DO consumption without epilimnetic replenishment.

We had a similar strong prediction success rate in validation data (n = 43 lakes) for both Models 1 and 2 (84% and 86%, respectively; Table 3.3). However, there was a significant difference in performance between Models 1 and 2, as indicated by the results of likelihood-ratio tests and the difference in their respective AICc values. The significantly higher epilimnetic TP in Model 1 lakes that were anoxic versus oxic, regardless of prediction category or correctness (Figure 3.5A), indicates that Model 1 results are associated with lake TP even though individual model parameters were not strongly correlated with TP alone (z_{max}: $R^2 = 0.22$; $MetaT_A$: $R^2 = 0.02$). This result attributes Model 1 prediction failures to differences in lake TP concentration. Not surprisingly, Model 2 (z_{max} + $MetaT_A$ + TP) presented significant differences in TP among lakes based on actual and predicted anoxic condition. Lakes that were predicted to be anoxic (correctly or incorrectly) had higher TP than correctly predicted oxic lakes, but anoxic lakes that were incorrectly predicted to be oxic had similar TP to oxic lakes regardless of prediction success (Figure 3.5B). Lakes that were incorrectly predicted in Model 2 did not have significantly different TP. These results indicate that including TP helps to refine the model based on morphometry alone, and that nutrient concentration still influences the development of anoxic extent in low-nutrient lakes. Incorrect predictions in Model 2 could not be explained by patterns in spatial distribution of lakes (Appendix S1: Figures B2, B3), bedrock geology (Appendix S1: Figure B.4), natural features of the lake or landscape, or anthropogenic condition of watersheds. It is possible that these incorrect predictions are related to temporal factors, such as timing of spring turnover or onset of summer stratification (Livingstone and Imboden 1996, Yuan and Jones 2020), which were not included in this study but may inform future analyses with this dataset.

Incorrect predictions in Model 2 may also be associated with the narrow ranges of lake depth and epilimnetic TP used to constrain the study dataset. By reducing the variability in these parameters, we are introducing the possibility of missing descriptive relationships in shallower, more nutrient rich lakes. This may be especially relevant with our low TP criterion, as samples with low TP concentrations may have been inaccurately evaluated because of analytical challenges with low detection limits and increased possibility of contamination.

Lakes that exhibit anoxia despite being predicted to be oxic based on morphometry alone (Model 1), or morphometry and TP concentrations (Model 2), may develop a more extensive anoxic area than correctly-predicted oxic lakes with similar morphometry. It is possible that these lakes are particularly vulnerable to anoxia from comparatively small increases in nutrient concentration. This group of lakes deserves closer study, as well as identification in lake management investigations regarding the creation of lake-specific thresholds for shoreline development and nutrient load assimilation (Nürnberg 1997).

In attempts to further minimize the role of TP in predicting anoxic extent, thereby accentuating the role of lake morphometry, we created alternative versions of Models 1 and 2 using a more restrictive criterion of including only lakes with < 10 µg/L TP (Models 3 and 4; Table 3.3). These models had lower AICc values but overall similar prediction success rates to Models 1 and 2. This suggests that there may not be great differences in the association of morphometry, or morphometry and TP, with the development of anoxic extent among lakes with either TP < 10 or TP < 15 µg/L. The lower AICc values may be due to lower sample size of lakes with TP < 10 µg/L compared to TP < 15 µg/L lakes; regardless, this change in model structure did not translate into substantially better prediction rates.

Model 5 may be used in lakes where detailed bathymetry data are not available (Table 3.3). This model predicts that the likelihood of hypolimnetic anoxia increases with decreasing maximum depth and decreasing depth of the average thermocline ($Therm_{\bar{x}}$) (Figure 3.3). While $Therm_{\bar{x}}$ is correlated strongly with the average depth to the top of the metalimnion ($MetaT_{\bar{x}}$; R^2 = 0.94), it is weakly associated with the relative lake area under the top of the metalimnion ($MetaT_A$; R^2 = -0.12), the parameter used in Models 1-4. In the current dataset, except for associations with metalimnion, $Therm_{\bar{x}}$ was most strongly correlated to total lake volume (V_0; $R^2 = 0.55$). Therefore, $Therm_{\bar{x}}$ likely works in this model partially because it is serving as an estimate of V_0. As discussed above, deeper, higher volume lakes have larger reserves of DO and a greater thermal stability and are better able to mitigate DO loss during stratification. Despite correlating to V_0, $Therm_{\bar{x}}$ was not strongly correlated to z_{max} ($R^2 = 0.31$), so no collinearity problems were identified with Model 5. Prediction success metrics were similar but slightly lower for Model 5 compared to Model 1 (81.4% vs. 83.7% correct in validation datasets; 0.92 AUC for both). There may be utility in applying this model to lakes where only a single profile is available, but prediction accuracy would likely be increased if long-term average thermocline depth was used instead.

Bathymetric data detailed enough for calculating the morphometric variables used in this study may be unavailable in many situations, possibly making Model 5 of more widespread applicability than the other models presented in this study. However, the availability of models developed with detailed stratigraphic morphometry data (Models 1 - 4) may enhance the validity of non-bathymetric models (Model 5), because comparisons may be made to models developed from lakes with more detailed data.

DOC was not an important predictor variable in any iteration of these models, suggesting that DOC variation at the range observed in this dataset (mean = 4.3 ± 1.5 mg/L; Table 3.1) does not play an important role in the areal extent of hypolimnetic anoxia in deep, low-nutrient temperate lakes. This supports the findings of Nürnberg (1995), who determined that DOC accounted for only a minor role in AF and that lake morphometry and biological productivity were the primary drivers controlling the extent of lake anoxia. However, other studies of different lake types have found that increases in DOC have led to increased anoxia in a shallow eutrophic temperate lake (z_{max} = 2.9 m, Brothers et al. 2014) and a small humic boreal lake (z_{max} = 12 m, Couture et al. 2015). Model 2 lakes predicted to be anoxic had significantly higher DOC than lakes predicted to be oxic, but no other differences in DOC were observed among prediction categories in Models 1 or 2 (Figure 3.4). Model 2 lakes predicted to be anoxic were also higher in TP (Figure 3.5B), suggesting that the higher DOC concentration in these lakes is due to increased biological productivity (Nürnberg and Shaw 1999). DOC may be a latent variable in the development of anoxia in these lakes, since DOC concentration can affect light attenuation and thermocline positioning (Fee et al. 1996).

TP was used in this study as a proxy for the biological activity that may contribute to loss of DO through the decomposition of organic matter. Using chlorophyll-*a*, a direct measure of lake biological productivity, produced similar model results. This was not surprising as the close relationship between TP availability and chlorophyll-*a* concentration is well documented and serves as the basis of management activities in most temperate lakes. However, the TP-chlorophyll-*a* relationship, while robust in large datasets (and strongest during peak stratification in most Maine lakes), may be weaker at smaller scales. Yuan and Jones (2020) found that chlorophyll-*a* was more successfully predicted by TP for individual lakes when TP was partitioned into dissolved P, P bound to suspended sediments, and P bound in phytoplankton cells. Future studies of lakes with a wider range of DOC than lakes in the present study, and consideration of other variables such as suspended solids, may help to further quantify the interactions of TP, chlorophyll-*a*, and DOC in the anoxic extent of lakes.

We tested several lake morphometric variables that relate to potential for wind-driven mixing events during the stratification period (SDR, Fetch, MaxL, MaxW, L:W, OI; Table 3.1). None of these variables appeared in any of the top 25 ranked models, which was surprising as AF is correlated with OI (Nürnberg 1995), and OI helps to explain some of the variance between TP and AF (Nürnberg 2019). Average depth was also not an important variable for predicting areal extent of anoxia, but it has been a component of models in other studies linking shallower average depth with increased probability of anoxia (Reckhow 1977) and higher DO consumption rates (Cornet and Ringler 1980, Rippey and McSorley 2009). These unexpected results may be related to the constrictions placed on the z_{max} and TP in our dataset. We purposefully excluded lakes that were shallow enough to likely experience only ephemeral stratification. While lakes with z_{max} < 10 m may develop ephemeral anoxia during short stratification periods, frequent mixing and wind-driven disturbance of sediment may also contribute to elevated trophic state in these lakes. Alternatively, z_{max} may have been more important than other morphometric variables due to several lakes in the dataset with large surface area, shallow average depth, and small deep holes that are susceptible to local anoxia. Therefore, the use of z_{max} may account for this situation and produce a better model for these lakes. Additionally, we excluded lakes with TP ≥ 15 µg/L, because lakes above that threshold may have anoxic condition governed by nutrient concentration rather than basin morphometry (Molot et al. 1992). By excluding shallower lakes and only including lakes with lower TP, our models found that different morphometric variables more successfully predict the occurrence of anoxia in low- nutrient, deeper lakes. While these constraints on lake morphometry and condition may limit model applicability in some areas where smaller and nutrient-rich lakes are the norm, it may apply well in northern latitudes where low-nutrient lakes are more common (Abell et al. 2012).

Our models can help identify low-nutrient lakes that may be especially sensitive to development of natural anoxic conditions. Combining the likelihood of natural anoxia with Al:Fe sediment chemistry data (Kopáček et al. 2005) can provide a valuable assessment of how vulnerable a lake is to internal release of sediment-bound P, informing lake management with loading allowances and restoration targets. There is potential for guiding management decisions for high nutrient lakes as well, as these models may help identify situations where eutrophication is being exacerbated by natural basin morphometry, or lakes that will still experience anoxia even if excesses in nutrients and organic matter in water and sediments are reduced. These models can potentially support decisions necessary in long-term restoration plans with respect to artificial hypolimnetic aeration or chemical treatments for the precipitation of P.

Our results have clear implications for fisheries management as well. The model response could be re-calibrated with a hypoxia threshold of 5.0 mg/L DO, a minimum value required to support most coldwater fisheries such as salmonids (Davis 1975). This adaptation of the model would address the probability of particular amounts of habitat loss for coldwater fish due to natural hypoxia in specific lakes. In a fisheries management version, it may be of interest to focus on volumetric DO loss measures in order to quantify the potential loss of available three-dimensional habitat, rather than the two-dimensional area of sediment-water interface as we did here.

All of the top-ranking models (Table 3.2) included variables related to maximum lake depth and area or volume metrics related to the position of stratification layers, rather than whole-lake measures. This emphasizes the importance of thermal stratification of lake layers in the development of anoxia in lakes. Consequently, annual variability in the positioning of these layers likely has implications for year-to-year fluctuations in anoxic condition.

There are many opportunities for further work to extend this model and refine its approach for a variety of study questions. For example, we did not account for temporal weather variability or the climate gradient from costal to northern Maine, which may explain the limited importance of variables such as fetch or OI in our results. We also did not account for the annual variability in DOC concentration or quality in lakes and watershed runoff that can occur due to photochemical degradation of DOC (Porcal et al. 2010). Water temperature metrics were not included as potential model variables, although temperature (and temperature gradients between stratification layers) likely relate to the extent of hypolimnetic anoxia. Water temperature may be especially important if this analysis were expanded to larger geographic areas, spanned multiple climate regions, or incorporated forecasts of future climate conditions.

In light of research addressing the role of lake morphometry in stratification changes brought about by climate change, our results suggest that it may be possible to identify which lakes are more susceptible to increasing anoxic extent due to a warming climate. Both Schindler (2001) and Kraemer et al. (2015) found lakes developed deeper thermoclines (among other modifications) in response to variables attributed to climate change in surveys of global and boreal lakes, respectively. Schindler (2001) attributed the deepening thermoclines to reduced runoff which increased residence time and lowered allochthonous DOC inputs, thereby increasing solar radiation penetration. According to our models, this scenario would reduce the likelihood of anoxia, as the area below the epilimnion would decrease. However, Robertson and Ragotzkie (1990) and Snucins and Gunn (2000) found that thermoclines were shallower in warmer years due to faster onset of spring stratification and rapid heating of surface waters. Maine lakes have been experiencing increasingly earlier ice-outs since the mid-20th century (Hodgkins et al. 2002, Ellis and Greene 2019). Given the low mean DOC concentrations of lakes in our study, it is likely that these lakes are more likely to experience increasingly shallower thermoclines as temperatures warm and ice-off periods lengthen. This is supported by preliminary analyses of long-term temperature data for Maine lakes (ME DEP, *unpublished data*). In this situation, these deep, low DOC lakes will be increasingly susceptible to anoxia as the percentage of area beneath the epilimnion increases. Lakes with z_{max} 10-25 m may be the most susceptible to shallower thermoclines, as these lakes show the greatest differences in probability of anoxia attributed to $MetaT_A$; deeper lakes (>25 m) may be more resilient to the development of anoxic conditions from an expanding $MetaT_A$ (Figure 3.2). This may be especially critical as low-nutrient lakes become rarer across the landscape (Stoddard et al. 2016). Models such as these, which enhance our understanding about the ecological dynamics of pristine systems and habitats, will help inform management objectives intent on preserving them, including on the terrestrial landscape.

In summary, our results suggest that morphometry and stratification dynamics may be the most important variables in predicting the development of anoxia in low-nutrient deep lakes. The models presented here identify a way to evaluate the potential for natural anoxia without the need to incorporate direct measures of nutrient condition that may require extensive and costly sampling in some lakes. These diagnostic models provide lake managers with simple yet robust tools to evaluate the potential for natural anoxic conditions in either single lakes (e.g., drinking water supplies) or many lakes across a region (e.g., regional studies, agency water quality assessments), and compare observed results to an expected reference condition. This approach does not replace the pioneering work on quantifying the extent of lake anoxia (e.g., Nürnberg 1995, 2004) or predicting DO profiles (Molot et al. 1992) but addresses a gap in the literature by developing a method to quantify the probability of natural anoxia occurring in low-nutrient lakes. This was an especially critical gap for lake research in areas similar to Maine: glaciated, north-temperate climate, with primarily smaller, low-nutrient lakes. The methods used to build the models presented here helped identify which variables are effective at predicting an important ecosystem attribute and will help to inform future assessments regarding observed versus expected conditions with respect to ecological regime shifts.

CHAPTER 4. SHIFTING BASELINES AND CROSS-SCALE DRIVERS OF LAKE WATER CLARITY: APPLICATIONS FOR LAKE ASSESSMENT

Chapter Abstract

Temporal Secchi depth trends are used in lake assessment to evaluate lake condition and possible shifts in trophic state. For accurate lake assessments, it is important to differentiate regional trends from lake-specific trends, but this can be confounded by interacting factors. We present a divergent trend analysis which uses temporal patterns of Secchi depth water clarity from 1999-2018 for five different types of reference lakes from minimally disturbed watersheds to create dynamic baselines against which we evaluate Secchi depth trends from non-reference lakes in Maine, USA. We used mixed-effect generalized additive models to generate smoothed curves of the expected baseline Secchi depth for each reference lake type to account for the nonlinear dynamics of lake condition through time. The majority of non-reference lakes (74%) showed no difference between measured trend (actual Secchi depth data) and divergent trend (residual Secchi depth from baseline trends). The most common difference in lakes with inconsistent trend test results showed stability in measured trends but apparent declining trends in divergent Secchi depth clarity. We used a Dynamic Factor Analysis (DFA) model to help interpret the variation and shifts observed in baseline reference lake trends. The best DFA model identified two common trends in water clarity among lake types and precipitation during the primary stratification season as the most informative covariable. Because precipitation amount and intensity are expected to increase according to predictive climate models for the Northeast US, our results suggest that baseline lake clarity in Maine will decrease with climate change.

Introduction

Water clarity, as quantified with Secchi depth, is perhaps the most widely used lake water quality metric due to its simplicity of measurement, low cost, and correlation to many lake variables. In most temperate lakes, Secchi depth correlates to trophic measures of total phosphorus and chlorophyll-*a*, in addition to suspended sediment and Colored Dissolved Organic Matter (CDOM) (Brezonik et al. 2019). Secchi depth is highly informative for long-term lake monitoring as it provides a measure of ecosystem disturbances at the local scale, including nutrient additions, watershed erosion, and resuspension of lake sediment (Gunn et al. 2001). However, temporal Secchi depth trends can also correlate to large-scale factors that influence water clarity which are not specific to an individual lake, such as regional precipitation or air temperature (Schindler et al. 1996, Read and Rose 2013, Rose et al. 2017).

The identification of water quality changes caused by localized anthropogenic activities is the basis of the need for lake assessments. Local influences on lake clarity often include increased watershed erosion or excess nutrient loading, both of which are generally related to anthropogenic activities (Soranno et al. 2015). Water quality assessments for lakes are generally predicated on the comparison of the observed condition to a baseline condition defined from minimally disturbed reference lakes (Stoddard et al. 2006, Herlihy et al. 2013). Many factors influence lake water clarity at multiple scales, so it is critical to differentiate the nonlinear dynamics in water clarity trends due to regional effects from those due to local factors.

Secchi depth trends may be falsely interpreted as declining water quality due to local factors when the lake is only responding to a regional factor such as weather patterns. Alternatively, lakes with stable Secchi depth may not indicate the influence of a positive regional trend because local effects, such as watershed erosion and nutrient enrichment, may counter or overwhelm the regional pattern (Rose et al. 2017). Therefore, the time period of interest and concurrent weather factors become relevant to proper assessment of lake Secchi depth trends. The literature reveals discrepancies for Maine lake Secchi depth trends, with studies finding different results for clarity trends during slightly different but overlapping time periods. McCullough et al. (2013) found that water clarity (using satellite-inferred Secchi depth) declined in Maine lakes during the period of 1995-2010. However, Canfield et al. (2016) reported that satellite-inferred lake clarity was stable for 1990-2010, but lake clarity based on Secchi depth field measurements increased from 1976-2013. Neither McCullough et al. (2013) nor Canfield et al. (2016) attempted to separately analyze lakes in minimally-disturbed watersheds to differentiate influencing variables on water clarity patterns between regional and local factors.

The degree to which lake clarity is sensitive to precipitation patterns varies with many attributes of lakes and their watersheds, such as watershed land use (Rose et al. 2017, McCullough et al. 2019). For example, agricultural land generally exports more nutrients, particularly phosphorous and nitrogen, to lakes in runoff than other land use types (Vaithiyanathan and Correll, 1992, Carpenter et al. 1998). Sensitivity to precipitation is an important factor for water clarity in reference lakes, such as those included here, with largely undeveloped watersheds. Water clarity appears to be less sensitive to precipitation changes in lakes with higher catchment to lake area ratios or higher percent wetland area in the watershed, and more sensitive if the lake has a lower watershed forested area, greater maximum depth, lower trophic state, or clarity controlled by CDOM rather than algal abundance (Rose et al. 2017, McCullough et al. 2019). Due to the variable nature of lake response to climate patterns and the predicted changes to climate in the northeastern United States (Fernandez et al. 2020), it is necessary to further our understanding of how climate change will affect baseline conditions in various lake types.

We established temporal patterns of Secchi depth water clarity in five different types of reference lakes from minimally disturbed watersheds (Deeds et al. 2020) to establish a baseline against which we evaluate Secchi depth trends from non-reference lakes in Maine, USA. We define “measured trend” as the trend based on actual Secchi depth measurements, and “divergent trend” as the trend based on the residual differences between the *measured* values and the *expected* Secchi depth values based on the reference lake baseline trends within the corresponding lake type. This approach enhances conventional assessments of lake trends by creating nonlinear, dynamic baselines of expected lake condition during the same time period. These dynamic baselines help to disentangle the effects of widespread regional factors and localized anthropogenic impacts on Secchi depth trends. It is important to understand why temporal shifts are occurring in reference lakes, especially if shifts in temporal patterns are due to yearly weather variation which is predicted to change in the future. To better understand how and why baseline lake Secchi depth trends shift over time, we applied a Dynamic Factor Analysis to trends observed in the five types of reference lakes to detect shared patterns in Secchi depth trends and evaluate the influence of regional climate covariables among different lake types.

Methods

Study Area

Maine has a land area of 91,647 km^2 and approximately 5,000 lakes > 1 ha. Elevation ranges from sea level to 1,608 m. The state was completely deglaciated by about 10,000 years ago. Most lakes have natural outlets although many are depth- and area-enhanced with outlet dams. The northern and northwestern portions of the state have a history of timber harvesting and re-growth, and periodic serious pest invasions with massive dieback. The northwestern quadrant of the state is relatively unpopulated. Dairy and cropland agricultural operations are largely within 100 km of the coast, which largely coincides with nutrient-rich and easily erodible marine clay deposits (Deeds et al. 2020). Agriculture is also prevalent in the northeastern part of the state. Human population is most dense in southern Maine and near the Atlantic coast. Development near lakeshores is primarily residential, both summer and year-round.

Deeds et al. (2020) identified five unique hydrogeomorphic lake types (henceforth, “lake types”) in three regions of Maine based on Omernik Level IV ecoregional groupings (Omernik 1987) and maximum lake depth: Coastal Deep Lakes, Coastal Shallow Lakes, Inland Deep Lakes, Inland Shallow Lakes and Northern Lakes; Figure 4.1). Northern Lakes is a grouping of both deep and shallow lakes in this study to accommodate low sample size in this under-surveyed region of Maine. The threshold between deep and shallow lakes was defined as maximum depth = 10 m, which is the approximate mean depth of the photic zone in our dataset of Maine lakes (approximated as twice mean Secchi depth; French et al. 1982). Anthropogenic watershed pressures vary in form and intensity across Maine, so a Watershed Quality Index was developed that identified reference, intermediate, and altered condition categories of lake watersheds that were specific to the scales of watershed disturbance observed in each lake type (Deeds et al. 2020). This was done by calculating percentiles of total watershed disturbance (urban, residential, agricultural, and harvested forest land cover) that occur in the direct watersheds of lakes in each type. Reference lakes have ≥ 90^{th} percentile *least* disturbed Watershed Quality Index values for direct watersheds, altered lakes are ≤ 25^{th} percentile *most* disturbed, and intermediate lakes have Watershed Quality Index values between the 25^{th} and 90^{th} percentiles. Secchi depth trends from the reference lakes of each type were used to create baseline trends for the calculation of residual values (i.e., difference from baseline) of Secchi depth trends in non-reference lakes.

Divergent Trend Analysis

Secchi depth readings have been collected in Maine lakes by state agency staff, professional researchers, and community scientists using standardized protocols since 1971, coordinated by the Lake Stewards of Maine's Volunteer Lake Monitoring Program. Trained monitors are frequently re-certified for data collection and recording procedures. Secchi depth measurements are taken with a viewing scope to reduce glare and to reduce reading variability occurring from various weather conditions. Data are quality controlled and maintained in a database housed at Maine Department of Environmental Protection (ME DEP 2015).

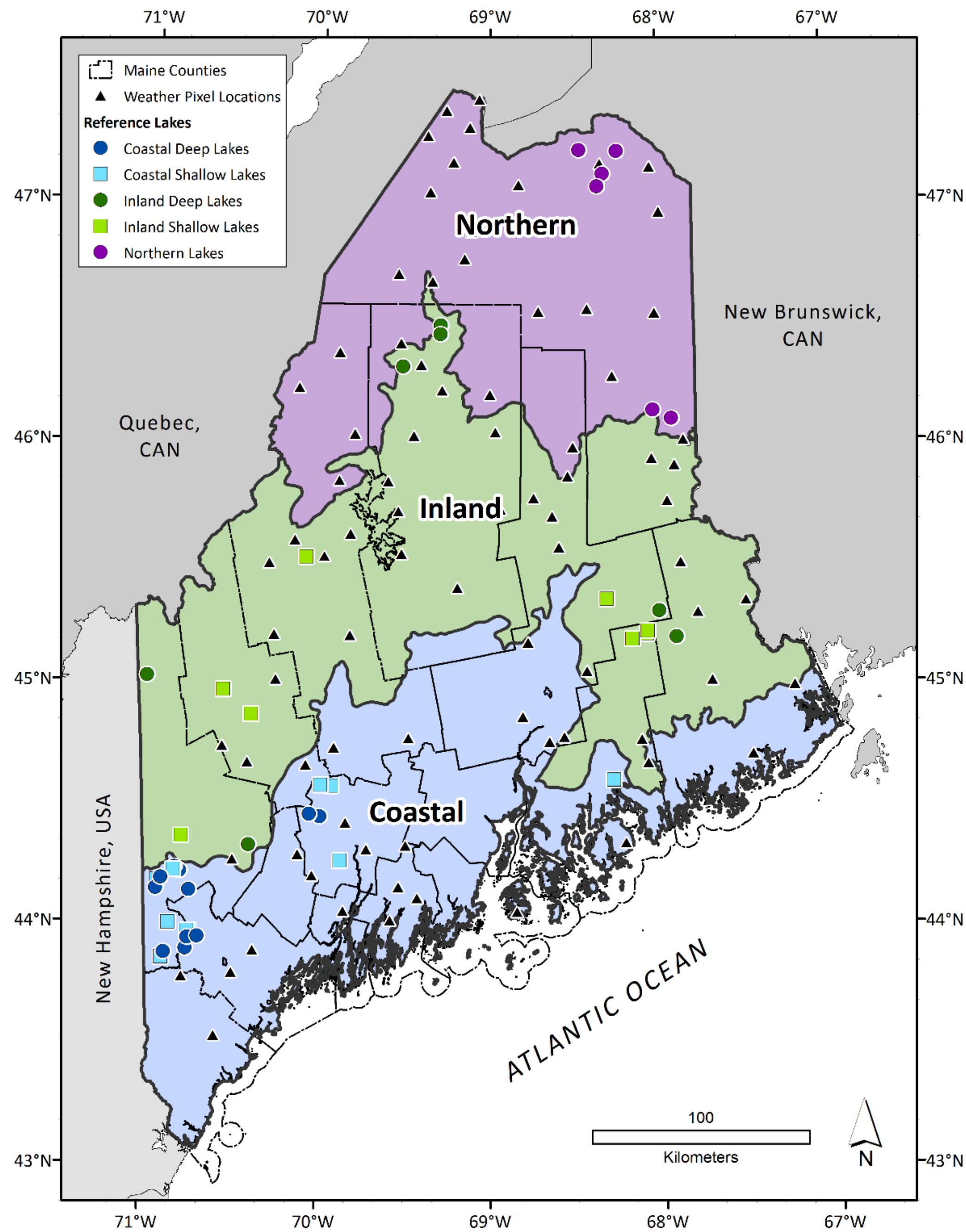


Figure 4.1. Map of Maine, USA, showing boundaries of three lake regions (Deeds et al. 2020), locations of reference lakes of five types, and random locations used to assemble weather data from Daymet's single-pixel interpolated weather data extraction tool (Thornton et al. 2020) for the Dynamic Factor Analysis.

Secchi depth data were summarized for each lake in each year of a 20-year study period (1999-2018) by calculating the mean values during the approximate period of peak lake stratification in Maine (1 August – 7 September). Only lakes with ≥ 10 yr of Secchi depth data with no ≥ 5 yr gaps in data record were included in the analysis. These criteria were relaxed for Northern Lakes, where long-term datasets are not as common, to ≥ 7 yr of data with no gaps ≥ 10 years. These data restrictions yielded a dataset with 310 lakes. Mann-Kendall Tau trend analysis was used to detect changes in Secchi depth over time using the Kendall package in R (McLeod 2015; R Core Team 2021). This non-parametric test determines whether median values of Y (Secchi depth, in this case) will increase or decrease with further increases in X (years) (Helsel et al. 2020). The Mann-Kendall Tau algorithm computes the positive or negative difference between each value (starting with the second value in a series) and all subsequent values. A value of 1, 0 or -1 is assigned to each position, and the sum of these integers (*S*) is computed. The test statistic Tau (τ) is calculated as

$$\tau = \frac{S}{\left(\frac{n(n-1)}{2}\right)} \tag{1}$$

and varies from -1 to 1. Tau is analogous to the Pearson's correlation coefficient in regression analysis, with larger negative values indicating stronger negative trends and larger positive values indicating stronger positive trends. If S and τ are significantly different from zero, then a statistically significant trend is identified (Helsel et al. 2020). Maine DEP currently designates lakes with Secchi depth trends with $\tau \geq |0.5|$ over ≥ 10 years as exhibiting a changing trophic condition (ME DEP 2015).

Reference lakes (Deeds et al. 2020) were used to establish reference lake trends that reflect non-anthropogenic regional variation in Secchi depth values across Maine (Table C1). We required that there be a minimum of three lakes in each year (from the subset of lakes that met the data inclusivity restrictions above) to calculate the yearly reference baseline trends. Two lake types, Inland Shallow Lakes and Northern Lakes, had insufficient sample size of reference lakes that met these data density criteria, so intermediate-disturbance lakes (based on the Watershed Quality Index, Deeds et al. 2020) were added to supplement the reference datasets for these lake types. Inland Shallow and Northern Lakes also had incomplete time series during the study period (some years with < 3 lakes with Secchi depth data), so Inland Shallow Lakes are represented from 2004-2018 and Northern Lakes are represented from 1999-2013. Forty of the 310 study lakes were designated as reference, leaving 270 non-reference lakes for comparison to baseline trends.

Reference lake time series were transformed into smoothed curves based on mixed-effed generalized additive models (GAMs) using the mgcv package in R (Wood 2021), with Secchi depth as the response variable, year as a fixed effect, and individual lake as a random intercept to account for differences in mean clarity among lakes. GAM models were used here because of their effectiveness at approximating values in nonlinear data. GAMs extend generalized linear models by replacing linear terms of Generalize Linear Models with smoothing functions called splines (Shadish et al. 2014). GAMs fit non-linear data with cubic splines, which are turning points in the data. The smoothing factor k indicates the number of cubic splines in the model, and therefore, the smoothness of the model fit; a lower k results in a smoother surface. If k = number of observations (years of data here), then the GAM curve intersects each observation point. To standardize the degree of smoothing across lake types with different number of observations (years), we restricted k by dividing the number of years of data for each lake by 3, as this calculation provided nonlinear model curves that best approximated Secchi depth variability while avoiding overfitting (Fisher et al 2017).

For individual lakes, the annual mean of measured Secchi depth values was subtracted from the yearly GAM-derived value of its respective reference lake type to obtain yearly residual values based on the baseline trend. These time series based on residual values were then analyzed with Mann-Kendall trend tests to produce the "divergent" trend result. Trend test results indicate if a lake is increasing in clarity (significantly positive τ), decreasing in clarity (significantly negative τ), or if clarity is stable or too variable to detect a trend (τ not significantly different from zero). For each lake type, the results of the divergent trend test results were compared to the results of the trend tests based on measured values.

Dynamic Factor Analysis

Dynamic Factor Analysis (DFA) is fully described in Zuur et al. (2003). In brief, it is a multivariate analysis used to estimate underlying common trends and important covariates among multiple time series. DFA models separate time series in terms of a combination of common trends and factor loadings (discussed below), explanatory variables (covariates), a level parameter (constant), and a noise component (unexplained error). Different covariables may be added to the model to test for improved performance. DFA was applied to the reference lake Secchi depth data to determine if observed patterns in baseline trends are associated with underlying common patterns or identifiable covariates.

The DFA model equation may be distilled to matrix notation, in a form similar to linear regression:

$$y_t = Z\alpha_t + c + Dx_t + e_t \qquad (2)$$

Where y_t is an $N \times 1$ vector of the variable y for N time series at time t, α_t is the value of the common trends at time t, c is the level parameter which is a $N \times 1$ vector of constants that allows each linear combination of common trends to move up or down, and e_t is a $N \times 1$ vector of noise component (unexplained error) at time t. The matrix Z consists of the number of time series (N) × the number of common trends (M) and contains the factor loadings which determine the linear combinations of common trends. Factor loading values are compared to determine which response variables are most strongly related to each common trend. Explanatory values may be added as a vector x of L covariates at time t; D is an $N \times L$ matrix containing regression coefficients. Note that α represents hypothetical variables (common trends) that cannot be explained by measured explanatory variables (x). Therefore, the intent of DFA is to create a model with a reasonable fit and a minimum M. A higher M will increase model fit but will also cause more parameters to have to be estimated, which may complicate interpretation. Common trends across all time series account for autocorrelation and reduce dimensions among the time series by representing undescribed processes that share similarity among time series. Factor loading values are calculated for each time series to measure the magnitude of the association of each time series with each common trend.

All DFA analyses were completed in R using the MARSS package (Holmes et al. 2020). The five time series used in the DFA were composed of standardized z-scores [(value – mean)/standard deviation] of yearly reference lake mean Secchi depth measurements, with one time series for each lake type. Z-scores are used rather than absolute measurements in DFA so that the analysis is focused on year-to-year differences within and among time series rather than the magnitude of differences in data among time series.

Akaike's corrected Information Criterion (AICc) was used as a measure of DFA model goodness-of-fit because it may be calculated for different types of variance, any value of *M*, or when there are relatively few observed data points (Zuur et al. 2003). Models with differences between AICc values (ΔAICc) < 2 are highly similar, and models with ΔAICc in the 2-7 range may be highly competitive (Burnham and Anderson 2002). Evidence ratios were calculated to aid in model comparisons using the R package qpcR (Speiss 2018). This ratio indicates the strength of evidence that the top scoring model is the best; for example, an evidence ratio of 10 indicates that evidence for the top model being the best one is 10 times stronger than the model in question. DFA model fit for individual lake types was evaluated with the ratio of the residual sum of squares (based on differences between the data and modeled estimates) and the observation sum of squares (based on differences among observations to the overall mean) ($SS_{residuals}/SS_{observations} = SS_{ratio}$). High values of SS_{ratio} indicate that a particular time series, or a portion of that time series, are not fitted well with the chosen DFA model. Ratio values closer to zero indicate a better model fit (Zuur et al. 2003).

After the best-performing model structure was selected, different covariates were tested to examine if climate variables were associated with identified trends. To generate covariates, weather data were obtained from Daymet (https://daymet.ornl.gov) using the single pixel extraction tool. Daymet provides interpolated estimates of daily weather parameters in a 1 km^2 grid for continental North America, PR, and HI (Thornton et al. 2020). The single-pixel extraction tool calculates daily weather data for a given GPS coordinate based on the closest 1 km^2 grid cell. To summarize weather data both across Maine and within the three lake regions, we generated random points across each of the lake regions, with the number of points standardized by land area (approximately 1 point per 10,000 km^2: Coastal = 22, Inland = 35 and Northern = 27 points; Figure 4.1). Mean monthly values during the study period were calculated for all points to generate monthly statewide weather parameters, and the process was repeated for points within each region so that we could investigate differences in weather patterns among regions. Average air temperature, total precipitation, percent of days with any precipitation, and percent of days with intense precipitation (defined as ≥ 10 mm/d; Griffiths and Bradley 2007) were summarized by monthly means. Monthly climate metrics were aggregated into various time periods in each year: winter (Dec – Feb), spring (Mar – May), summer (July – Aug), winter and spring (Dec – May), the approximate lake stratification season in Maine (April – August), and water year (previous October – September, inclusive). Covariate models were ranked by model strength with ΔAICc and Evidence Ratios. Sum of Squares metrics, discussed above, were used to evaluate model improvement within each lake type by comparing SS_{ratio} values from the non-covariate models to the SS_{ratio} in covariate models.

Results

Divergent Trend Analysis

Inland Deep and Inland Shallow reference lakes showed significantly improving water clarity from 1999-2018 (Table 4.1, Figure 4.2A). Trends for Coastal Deep, Coastal Shallow and Northern

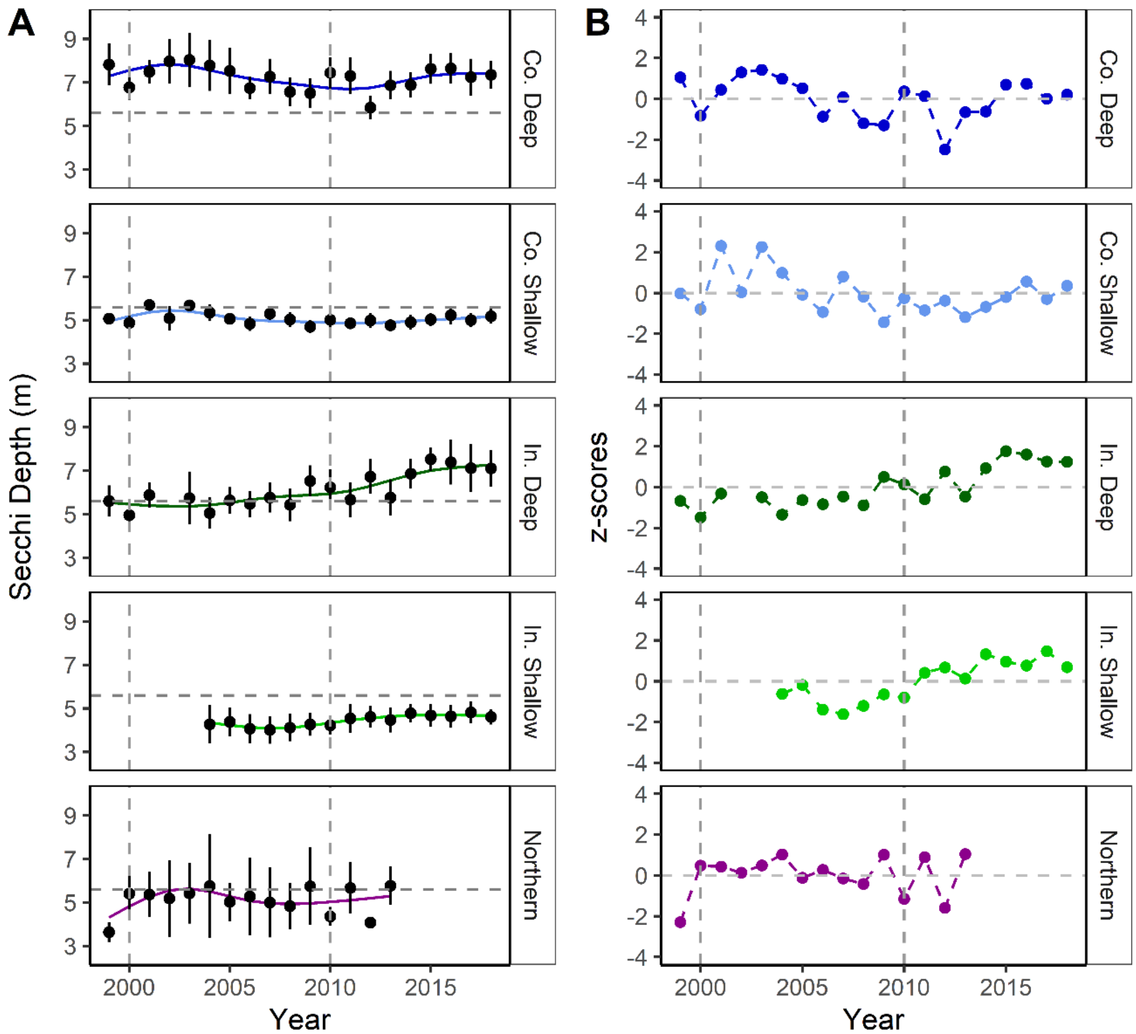


Figure 4.2. (A) Reference lake mean Secchi depth values (± mean se) with predicted GAM smoothing trend curves (solid lines). Dashed line at Secchi depth = 5.6 m represents the overall mean Secchi depth for reference lakes in this study. (B) Time series of z-scores [(value – mean)/standard deviation] for reference lake mean Secchi depth measurements for each lake type used in the Dynamic Factor Analysis. Y axis is unitless. 0 = no difference from mean of each lake type. Vertical dashed lines at 2000 and 2010 are for visual reference. Co. = Coastal, In. = Inland.

Table 4.1. Mann-Kendall trend tests results for mean Secchi depth values (measured) and GAM-predicted Secchi depth values for five reference lake types from 1999-2018. * indicates significance at α = 0.05.

Reference Lake Type	Measured Secchi depth (Yearly means)		GAM-predicted Secchi depth	
	τ	***p***	***τ***	***p***
Coastal Deep	-0.189	0.256	-0.242	0.144
Coastal Shallow	-0.211	0.206	-0.263	0.112
Inland Deep	0.591	<0.001*	0.895	<0.001*
Inland Shallow	0.600	0.002*	0.657	0.001*
Northern	0.010	1.000	0.086	0.692

reference lakes showed non-significant (stable or variable) measured Secchi depth trends. Trend tests based on GAM-derived smoothed Secchi depth values showed the same results as measured value trends in reference lakes. Coastal Deep Lakes had the highest overall mean Secchi depth (6.84 m) while Inland Shallow Lakes had the lowest overall mean Secchi depth (4.25 m) (Table C2).

The differences in Mann-Kendall results between measured and divergent trend tests for non-reference lakes varied in six different ways, including no difference between tests (Table 4.2). Example progressions for the five observed changes from actual trend to residual trend are displayed in Figure 4.3. For most lakes, results were consistent between measured and divergent trend tests (n = 200 out of 270 non-reference lakes, or 74%). Twenty-six lakes (10% of total) show stable Secchi depth in measured trends but declining Secchi depth in divergent trends (measured: stable to divergent: negative; Figure 4.3E-F). All but one of these lakes were Inland Deep Lakes. The third most common result was a change from a measured trend of stable clarity to a divergent trend of increasing clarity (measured: stable to divergent: positive; n = 22, 8%; Figures 4.3F-G). All but one of these lakes were Coastal Deep Lakes. The other two potential result changes, either negative or positive measured trends with stable divergent

Table 4.2. Comparison of trend test outcomes between measured and divergent trend tests for non-reference lakes. Positive = significantly increasing Secchi depth, Negative = significantly decreasing Secchi depth, Stable = Secchi depth trend had non-significant result, No Difference Between Tests = matching outcome between measured and divergent trend tests. α = 0.05. Total values for all lake types combined are in bottom row in italics.

Lake Type	Total number of non-reference lakes	Measured: Negative – Divergent: Stable	Measured: Positive – Divergent: Stable	Measured: Stable – Divergent: Negative	Measured: Stable – Divergent: Positive	No Difference Between Tests	Percent of lakes matching reference trend
Coastal Deep	113	6	1	0	21	85	29%
Coastal Shallow	77	1	3	0	0	73	5%
Inland Deep	54	0	5	25	0	24	56%
Inland Shallow	12	0	2	0	0	10	17%
Northern	14	1	3	1	1	8	43%
Totals	*270*	*8*	*14*	*26*	*22*	*200*	*26%*

trends, were represented by a small number of lakes (n = 8 and 14, or 4% and 5%, respectively; Figures 4.3A-B, 4.3C-D).

Of the five lake types, Inland Deep had the largest proportion of lakes with a difference in test results between measured and divergent trends (56%, Table 4.2). The most common difference in Inland Deep Lakes was measured: stable to divergent: negative (Figure 4.3E-F). Northern Lakes had different trend test results in 43% of lakes, with at least one lake representing all four of the change categories. Coastal Deep Lakes showed different results in 29% of lakes, and the most common change was measured: stable to divergent: positive (Figure 4.3G-H). Only two of 12 Inland Shallow Lakes showed a change in trend test result, and both were measured: positive to divergent: stable (Figure 4.3C-D). Coastal Shallow Lakes had the lowest percentage of lakes exhibiting a change in trend test result (5%).

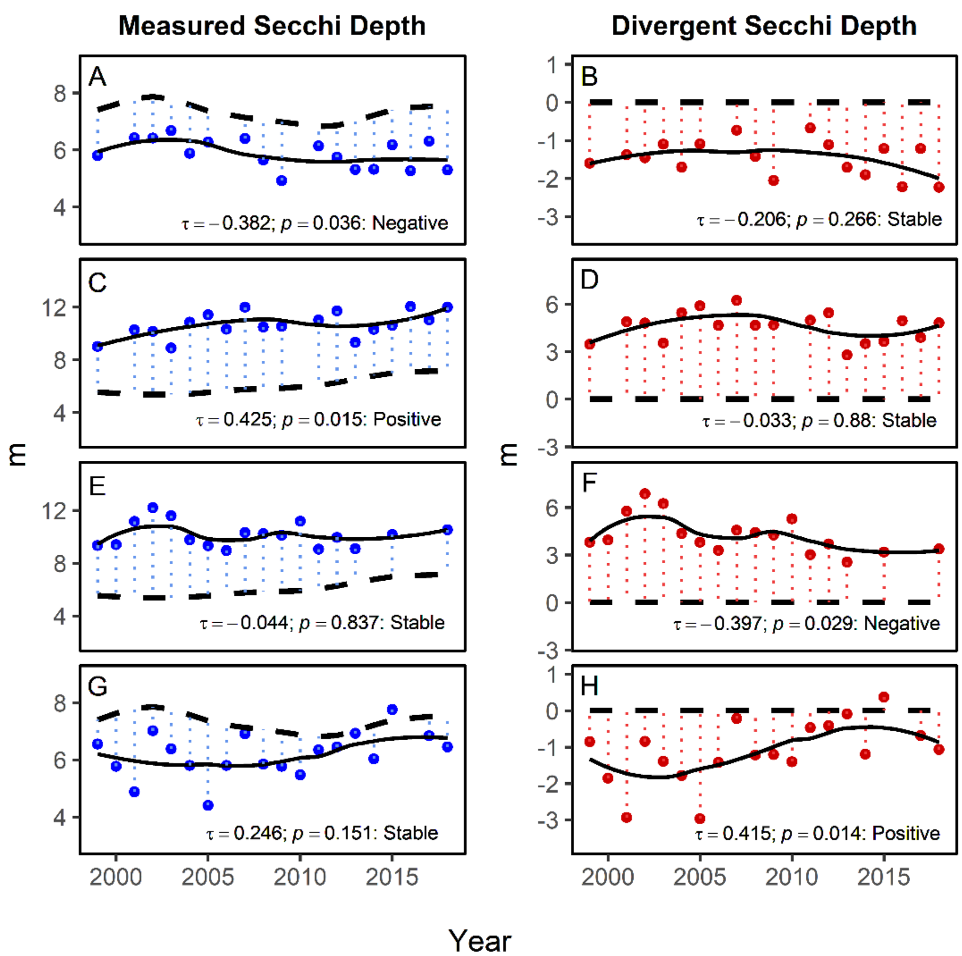


Figure 4.3. Examples of measured and divergent trend analyses from individual non-reference lakes for the four observed changes in trend test result. Measured Secchi depth trend plots (left) show the mean yearly Secchi depth readings in meters (dots), LOESS moving-average curve based on Secchi depth readings (solid lines), reference GAM curve for the respective lake types (heavy dashed lines), and the residual distances of the measured Secchi depth readings from the expected reference condition (light dashed vertical lines). Divergent Secchi depth plots (right) show the same data as the measured plots (left), but with the reference GAM line flattened at zero and the Secchi depth yearly means shown as residual distances in m from the reference GAM. The τ and p values for the trend tests are displayed in each plot with test result (positive = significantly increasing trend in water clarity, negative = significantly declining trend, and stable = non-significant trend. The four changed outcomes are represented by rows: (A-B) measured: negative to divergent: stable; (C-D) measured: positive to divergent: stable; (E-F) measured: stable to divergent: negative; and (G-H) measured: stable to divergent: positive. Reference trends used in each plot: A, B, G, H: Inland Deep Lakes; C-F: Coastal Deep Lakes.

Dynamic Factor Analysis

Standardized time series of Secchi depth z-scores showed variation among lake types, but all types displayed an apparent increase in Secchi depth from approximately 2011 to 2018 (except Northern Lakes, where data were absent) (Figure 4.2B). The selected DFA model (AIC = 244.18, ΔAICc = 3.44, evidence ratio = 5.59; Model 4, Table C3) had diagonal and equal variance and two common trends, indicating that variance was shared among lake types and there was no year-to-year correlation among types. This model was chosen because it had the lowest AICc scores and evidence ratio among models with two common trends. Three models had lower AICc scores and one common trend, but common trends in these models showed associations with only two lake types (Inland Deep and Inland Shallow). Model 4 identified a second common trend that provided an association with two additional lake types (Coastal Deep and Coastal Shallow) and lower (stronger) values of SS_{ratio} across multiple lake types. DFA Models 2-6 all had ΔAICc < 7, so these may all be considered similar enough for consideration as an appropriate model (Burnham and Anderson 2002). Model 4 was chosen as the best model among these competitive models. It presented associations with two more lake types than Models 1-3, and SS_{ratio} values indicated better model fit within each lake type.

The common trends associated with Model 4 indicated two separate patterns in Secchi depth trends among the five lake types. Common trend 1 starts with an increase from 1999 until 2003, after which it steadily decreases until 2012, then gradually increases again until 2018 (Figure 4.4). Common trend 2 is stable from 1999 until the late 2008, and then has a sharp increase until it remains relatively stable from 2015 through 2018. Factor loading values indicate the strength and direction of association with this trend for each individual lake type. Common trend 1 is associated most strongly with Coastal Deep and Coastal Shallow Lakes. Common trend 2 is most strongly associated with Inland Deep and Inland Shallow Lakes. Northern Lakes were not strongly associated with either trend (factor loadings for both trends < 0.2).

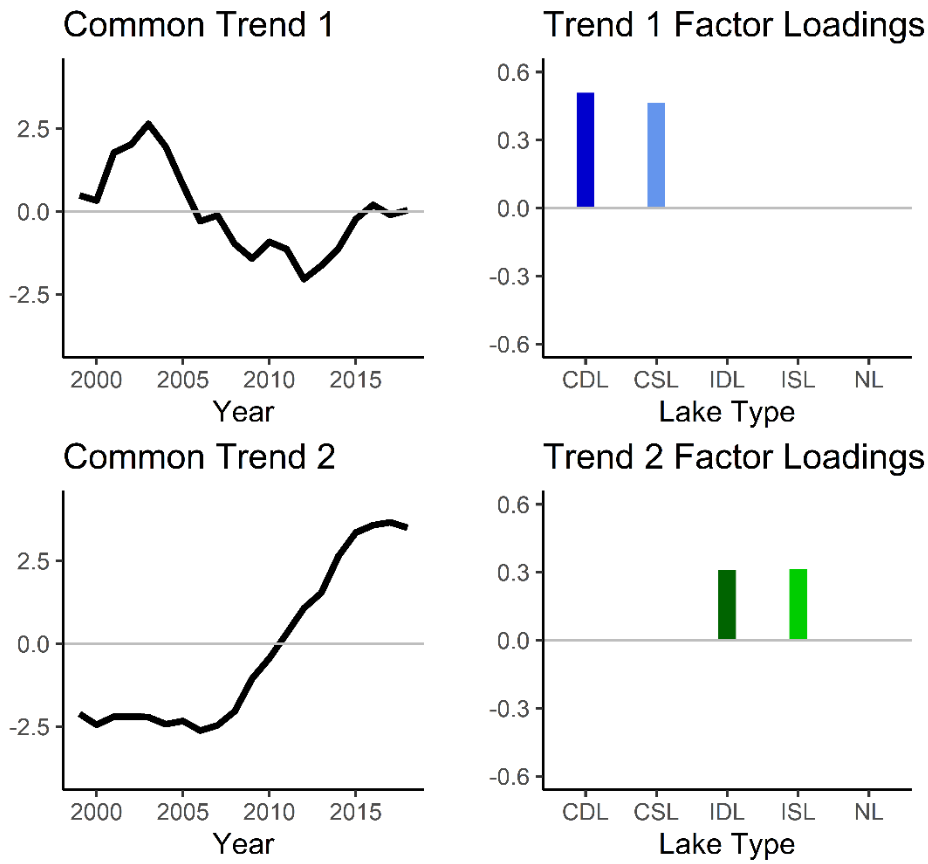


Figure 4.4. The two common trends and factor loading values for DFA Model 4. Lake types with factor loading values < 0.20 have been excluded from each common trend. Y axes are unitless. CDL = Coastal Deep Lakes, CSL = Coastal Shallow Lakes, IDL = Inland Deep Lakes, ISL = Inland Shallow Lakes, NL = Northern Lakes.

SS_{ratio} ($SS_{residuals}/SS_{observations}$) values evaluate the DFA model fit for each lake type (Table 4.3). Northern Lakes have the highest SS_{ratio} value for Model 4 (42.96), indicating the weakest fit among the five lake types. Coastal Deep and Coastal Shallow Lakes have moderately low SS_{ratio} values (0.99 and 1.25), and the Inland Deep and Inland Shallow types show the lowest SS_{ratio} values and best overall Model 4 fit (0.42 and 0.35, respectively). Model 4 fits for each lake type are shown in Figure 4.5A along with observed Secchi depth data (z-scores) and associated common trends.

Annual weather data were z-scored and applied to Model 4 as covariates (x_t in Equation 2) to see if regional weather patterns might improve DFA model fit. Mean monthly statewide precipitation for April – August (hereafter, “Precipitation”), which is the approximate stratification period for most dimictic lakes in Maine; Figure 4.6) was chosen as the best covariate model (AIC = 225.33, ΔAICc = 0; Table C4). Precipitation was chosen as the best covariate because it had strong performance based on its AICc score and produced the greatest improvements in model fit across all lake types as measured by SS_{ratio} (Table 4.3). The evidence ratio suggests that Model 4 + Precipitation is over 460 times more likely to be the best model over Model 4 with no covariates (ΔAICc = 12.27; Table C4). Combinations of paired precipitation and air temperature covariates were also tested (e.g., Precipitation + Degree Days) but none exceeded model performance of Precipitation as the sole covariate and are not presented here. Sixteen covariate models showed improved AICc scores over the no covariate model (Model 4). Two other covariate models had ΔAICc < 7, which indicates that two other climate metrics (percent of days with precipitation during the stratification period and number of degree days during the water year) may also improve model fit for Secchi depth trends in this DFA (Table C4).

There was improvement in model performance in all lake types with the addition of Precipitation as a covariable to Model 4 (Table 4.3). The greatest difference in model performance for a single lake type was for Northern Lakes (ΔSS_{ratio} = 41.67, or 188% improvement), which had the overall weakest fit in Model 4 with no covariates. Coastal Shallow Lakes had the next largest increase in model

Table 4.3. Sum of Squares (SS) model comparisons between DFA Model 4 (no covariates) and Model 4 + Precipitation covariate (mean monthly April – August precipitation) for observations of the fitted model and model residuals. SS_{ratio} ($SS_{residuals}/SS_{observations}$) is a model fit statistic, with lower SS_{ratio} values indicating a better fit model for that lake type. ΔSS_{ratio} is the mathematical difference between the SS_{ratio} values of Model 4 and Model 4 + Precipitation; positive values indicate increased model performance (Zuur et al. 2003). ΔSS_{ratio} percent difference is the percent difference between the two SS ratios and presents a normalized metric for model change with the addition of the covariate within each lake type. SS_{res} = SS of the residuals (based on differences among data and modeled estimates), SS_{obs} = SS of the observations (based on differences among observations and the overall mean).

Lake Type	Model 4			Model 4 + Precipitation			ΔSS_{ratio}	ΔSS_{ratio} Percent Difference
	SS_{res}	SS_{obs}	SS_{ratio}	SS_{res}	SS_{obs}	SS_{ratio}		
Coastal Deep	7.05	7.10	0.99	6.83	9.43	0.72	0.27	31%
Coastal Shallow	8.32	6.63	1.25	5.75	11.28	0.51	0.74	84%
Inland Deep	4.30	10.36	0.42	3.65	11.45	0.32	0.10	26%
Inland Shallow	2.93	8.41	0.35	1.85	9.17	0.20	0.15	53%
Northern	13.45	0.31	42.96	6.78	5.24	1.29	41.67	188%

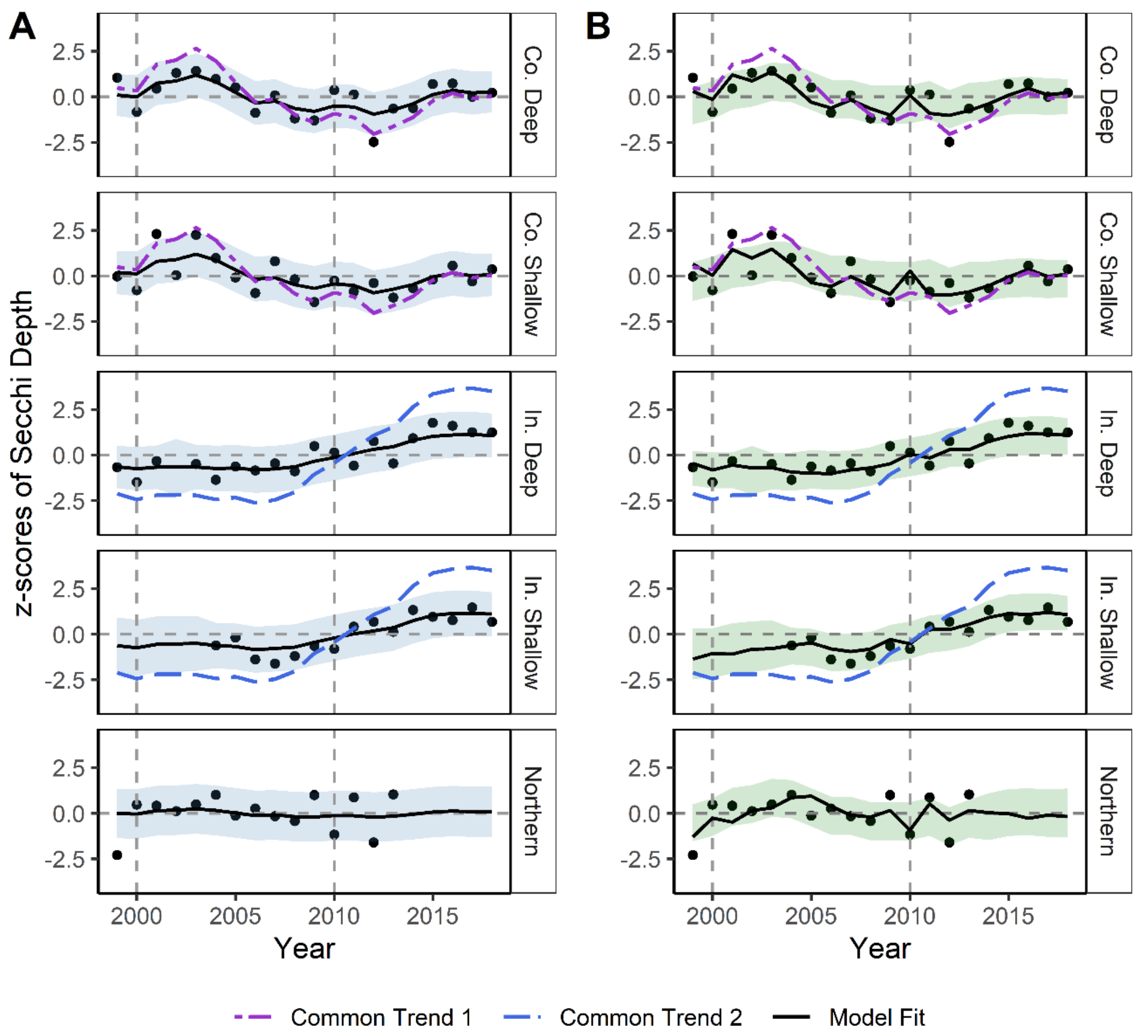


Figure 4.5. (A) DFA Model 4 (no covariables) for each lake type and (B) Model 4 + Precipitation fit results for each lake type. Points are observed Secchi depth values (z-scores), black lines represent DFA model fits, and purple and blue dashed lines represent common trends 1 and 2, respectively. The shaded area is the 95% confidence interval of the model fit line for each lake type. Each lake type is shown with its most strongly related common trend, except for Northern Lakes which had no strong common trend association (factor loadings < 0.2). Vertical dashed lines at 2000 and 2010 are for visual reference. Co. = Coastal, In. = Inland. Y axes are unitless.

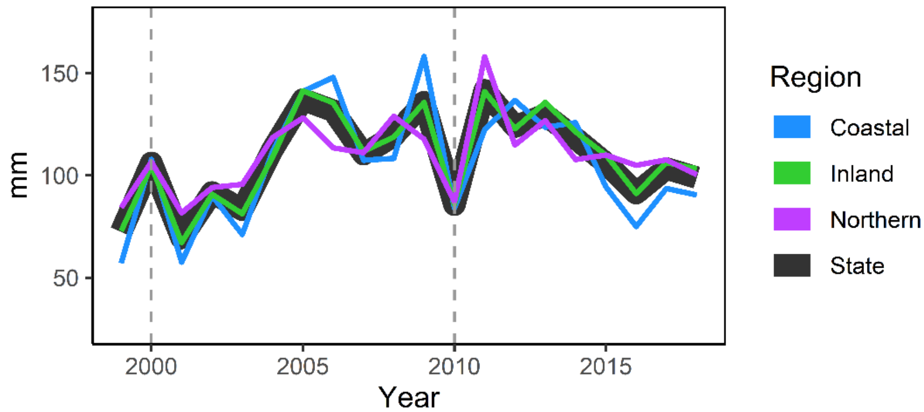


Figure 4.6. Precipitation (mean monthly precipitation during the stratification season, Apr-Aug, in mm) for the three lake regions and the Maine statewide average from 1999-2018. Data are from Daymet (Thornton et al. 2020). Vertical dashed lines at 2000 and 2010 are for visual reference.

fit (ΔSS_{ratio} = 0.74, 84%). The remaining three lake types showed moderate improvement in model fit (Table 4.3). The increase in model responsiveness with the addition of Precipitation is seen graphically by comparing the model fits (solid black lines) in Figures 4.5A (Model 4 alone) and 4.5B (Model 4 + Precipitation).

Water clarity in both Coastal lake types was significantly correlated with statewide Precipitation (Deep: R^2 = 0.343, p = 0.007; Shallow: R^2 = 0.454, p = 0.001), while lake types in the other regions were not significantly correlated (Inland Deep: R^2 = 0.016, p = 0.608; Inland Shallow: R^2 = 0.035, p = 0.502; Northern: R^2 = 0.153, p = 0.149). No differences in Precipitation were identified among regions (F = 0.163, p = 0.921; Figure 4.6), and results of correlation tests among lake types and precipitation estimates from their respective regions were not different from the correlations with statewide precipitation data.

Discussion

Divergent Trend Analysis

Mean water clarity increased for reference Inland Deep and Inland Shallow Lake types during the study period (Table 4.1). The other three lake types did not show significant changes in water clarity during the same time. The approximate period of 2011-2018 shows increases in GAM-derived clarity trends for all reference lakes except for Northern Lakes, where these data are absent (Figure 4.2). This seven-year period of increasing clarity in Coastal and Inland lake types coincides with a period of steady decrease in the amount of precipitation during the stratification period (Figure 4.6).

Comparisons of measured Secchi depth trends to divergent Secchi depth trends based on reference lakes of various types show that lakes most commonly exhibited no difference or one of four differences in trend test result (Table 4.2, Figure 4.3). The majority of lakes (74%) showed matching trend test results, indicating that the non-reference lake Secchi depth trend was similar to the baseline reference lake trend, and therefore the lake reflected regional baseline Secchi depth patterns. The most common difference in trend test result was where measured trends were stable and divergent trends were declining, meaning that a lake's Secchi depth measured trend is stable according to actual Secchi depth measurements, but the divergent trend analysis shows an apparent decline in water clarity because Secchi depth did not increase as expected according to reference lake trends (Figure 4.3E-F). This difference in trend test result indicates that there may be local factors influencing lake clarity more strongly than regional factors. This result was mostly associated with Inland Deep Lakes, which had a significantly increasing measured Secchi depth trend in reference lakes (Table 4.1). The next most common change in trend test result was measured: stable to divergent: positive, indicating again that the lake trend was identified as stable with measured Secchi depth data, but a period of declining water clarity in the baseline trend created an apparent increase in water clarity for the non-reference lake (Figure 4.3G-H). This occurrence may be due to a periodic decline in baseline water clarity observed in reference lakes, or possibly changes in local conditions that contributed to increased lake clarity in the non-reference lake. This result was most common in Coastal Deep Lakes, which showed a period of declining water clarity from about 2003-2012, as reflected in Common Trend 1. The third most common change in trend test result was measured: positive to divergent: stable, which suggests that measured lake clarity increases may be attributed to the "background" signal that was observed in reference lakes (Figure 4.C-D). In this instance, the trend in non-reference lake clarity is consistent with the regional baseline trend observed in reference lakes, suggesting that local drivers such as land use are not influencing lake clarity. The final difference observed in trend results, measured: negative to divergent: stable, only occurred in a small number of lakes (3%). This change indicates that a lake showed a trend of declining water clarity, but the baseline trend showed a similar decline during the same period. This is analogous to the measured: positive to divergent: stable result, but in the opposite direction. There are two other possible results that were not encountered in our study: measured: positive to divergent: negative and measured: negative to divergent: positive. Both of these results would indicate that water clarity in the non-reference lake is trending in the same direction of the reference baseline but that the trend is not as strong.

The strongly increasing Secchi depth trend in Inland Deep reference lakes has the potential to create a large difference in trend between the expected baseline condition seen in reference lakes versus the observed conditions in non-reference lakes, because for a non-reference lake to be considered stable in the divergent analysis, it would also have to exhibit a similarly steep increase in Secchi depth measurements over time. Consequently, the trend analyses for Inland Deep Lakes had the largest proportion of non-reference lakes show apparent reductions in clarity (positive to stable or stable to negative) in the divergent trends (56%; Table 4.2). Even lakes with moderately increasing water clarity may appear to show stable or apparent declining clarity when compared against a strong reference trend (Figure 4.3C-D). Inland Shallow reference lakes also showed a significantly increasing water clarity trend, but only 2 of 12 non-reference Inland Shallow Lakes (17%) had differing divergent trend results (Table 4.2). It is not possible to draw strong conclusions with this low sample size, and the results are likely confounded by the inclusion of non-reference lakes in calculation of the baseline trend (see *Methods*), but it seems that the non-reference Inland Shallow Lakes are generally following the regional baseline pattern observed for this lake type.

Coastal lake types had the largest number of sampled lakes in this study, but also the lowest proportion of lakes exhibiting differences in divergent trend test result (apart from the under-sampled Inland Shallow Lakes). This indicates clarity trends in non-reference Coastal Deep and Coastal Shallow Lakes are generally consistent with the baseline clarity trends observed in Coastal reference lakes. Possible explanations for this may be that reference and non-reference watersheds for both Coastal lake types are more similar than in other lake types. The approach used here to define reference lakes used the gradient of watershed disturbance unique to each lake type (Deeds et al. 2020). Lakes in the Coastal region may experience greater pressures from watershed development and recreation than other Maine lake types due to their high water clarity and proximity to population centers. This is especially relevant for the Coastal Deep reference lakes, which are larger and have the highest water clarity of all Maine reference lake types. Regardless of differential development pressures, since the reference lakes of both Coastal lake types showed non-significant (stable) trends in water clarity over the study period (Table 4.1), we infer that most non-reference lakes in this region also have stable water clarity.

Northern Lakes also had non-significant reference lake trend (Table 4.1), perhaps due to high variability in water clarity rather than stability (Figure 4.2). Additionally, low sample size and lack of true reference lakes with long-term data likely affected the results for this lake type. Higher amounts of watershed development, especially agriculture, occur in Northern Lakes compared to the other four reference lake types (Table C2), as non-reference lakes were used to generate the Northern Lake GAM reference trend. Lakes in watersheds with prevalent agriculture can experience increased effects of precipitation (Rose et al. 2017), as stormwater runoff from agricultural land generally has higher nutrient concentrations than runoff from other land cover types (Carpenter et al. 1998).

Dynamic Factor Analysis

To advance our understanding of nonlinear dynamics and shifting baselines of Secchi depth in Maine reference lakes, we employed a Dynamic Factor Analysis to investigate possible commonalities in Secchi depth trends among the various types of Maine lakes. We applied climatic variables as covariates to examine the associations of precipitation and air temperature patterns with Secchi depth trends.

The DFA model that best fit our reference lake data had two common trends that showed associations with Coastal Deep and Coastal Shallow Lakes (Common trend 1) and Inland Deep and Inland Shallow Lakes (Common trend 2). Northern Lakes were not strongly associated with either common trend, indicating that neither common trend reflects the temporal patterns of water clarity observed in these lakes. This may be due to the high variability in water clarity observed in the Northern Lakes, which do not reflect true “reference” conditions observed in the minimally disturbed watersheds of reference lakes of the other four types. The high variability in Northern Lake clarity data was reflected in the especially poor DFA model fit for these lakes, as measured by the SS_{ratio} (Table 4.3).

April through August mean monthly precipitation was the strongest covariable in the DFA covariate models (Table C3). Many other studies have also demonstrated the effect of precipitation on lake water clarity (Schindler et al. 1996, Read and Rose 2013, Rose et al. 2017, McCullough et al. 2019). Six covariate models had $\Delta AICc < 7$, indicating that many alternative weather variables may help inform clarity trends (Burnham and Anderson 2002), including precipitation intensity, frequency of precipitation, and number of degree days. Precipitation can decrease water clarity in lakes through the erosion and continued suspension of soil particles that transport nutrients and CDOM from the watershed into the water. Elevated concentrations of nutrients can stimulate algal growth and additional CDOM in lake water can reduce light penetration, both of which will reduce water clarity

(Schindler et al. 1996, Rose et al. 2017). This association may be especially relevant over longer time periods, as fluctuations in precipitation during the lake stratification period over multiple decades has been shown to influence lake water color and subsequent chlorophyll-*a* response (Carpenter and Pace 2018). It is surprising, however, that mean monthly cumulative precipitation showed a stronger association with Secchi depth trends than measures of precipitation intensity, which can be an indication of weather events that induce greater amounts of watershed soil erosion. It may be that the large spatial scale of this study produces a stronger association to regional precipitation amounts than precipitation intensity because intense rainfall events, typically convective storms, are generally more localized and may not be recorded by remote weather stations. Exploring this distinction further may be especially relevant in Maine, where an increasing frequency of intense storms has already been documented (Fernandez et al. 2015). Studies focused on single lakes, or groups of lakes in close proximity, with local weather stations may show stronger associations with precipitation intensity rather than cumulative totals (e.g., Coats 2010).

ΔSS_{ratio} (Table 4.3) may be considered as a measure of sensitivity to precipitation, as this metric evaluates change in DFA model fit with the addition of precipitation as a covariable. The DFA non-covariate model had a poor fit for Northern Lakes especially; however, this lake type had the largest improvement in model fit with the addition of the Precipitation covariate (Table 4.3). The sensitivity of precipitation in these Northern Lakes may be due to the high amounts of watershed agriculture in these lakes compared to the other lake types. Rose et al. (2017) found that watershed agriculture, where present, was most important in explaining lake water clarity variability between dry and wet years. Runoff in the watersheds of the Northern Lakes used as reference lakes here likely carries higher concentrations of nutrients than runoff in the watersheds of other lake types considered in this study.

The improvements in DFA model fit were greater for the two Coastal lake types than for the two Inland lake types with the addition of the Precipitation covariate, as measured by ΔSS_{ratio} (Table 4.3). We expected that the two deep lake types, Coastal and Inland, would have the greatest response to a precipitation covariate after the findings of other studies that water clarity in deep, unproductive lakes are more sensitive to precipitation (Rose et al. 2017, McCollough et al. 2019). However, sensitivity to precipitation seemed to be related more strongly to lake region here than to other characteristics of lakes or their watersheds. The shallow Coastal and Inland lake types in our study actually showed greater sensitivity to precipitation than the Coastal and Inland deep lakes, as measured by ΔSS_{ratio} percent differences (Table 4.3). It is possible that these shallower lakes are more sensitive to precipitation than deep lakes in the same region because of the greater proportions of the lake volume that is replaced with stormwater during precipitation events. Further, the shallow lakes are more likely to be polymictic and experience periodic resuspension of lake sediment and nutrients throughout the stratification season, particularly in association with storms (Stockwell et al. 2020).

Variables such as maximum depth, catchment area: lake area ratio, or residence time did not explain the patterns in sensitivity to precipitation among reference lakes (Table C2). Instead, water clarity in both Coastal lake types was significantly correlated with statewide precipitation whereas lake types in the other regions were not. We were not able to detect stronger correlations between regional precipitation estimates and the water clarity values among lake types and regions, and we also did not identify a difference in precipitation amounts among our lake regions. Further research exploring the connections between weather patterns and lake condition may help to explain the different dynamics we observed among lake regions.

The pattern of decreasing precipitation during the stratification period from 2011-2018 (Figure 4.6) appears to correspond with a period of increasing Secchi depth in all lake types (Figure 4.2). Our covariate model shows that including precipitation as a covariate improves DFA model fit for all lake types (Table 4.3). Precipitation amounts in Maine have oscillated over time, but the annual statewide average precipitation has increased 15% from 1895 to 2018, and 30% from 1960 - 2018 (Fernandez et al.

2020). High-intensity weather events and total amount of precipitation has been predicted to increase in frequency according to climate models for the northeast US and Maine (Diffenbaugh et al. 2005, Mallakpour and Villarini 2017, Fernandez et al. 2020). Considering these model projections, our results indicate that baseline Secchi depth values in Maine lakes could shift toward lower water clarity, especially if those increases occur during the stratification period. This is of particular relevance for water quality assessments that evaluate lake condition based on the expected condition of minimally-disturbed reference lakes. Understanding how precipitation patterns will influence the expected lake condition and Secchi depth trends in the future will allow lake researchers to adapt to shifting baselines of reference lake condition.

Seven of 16 covariate models that improved model fit over Model 4 included an air temperature variable (Table C4). Ambient air temperatures can affect almost all hydrological, chemical, physical, and biological properties of lake systems, including the watershed. However, lakes may respond heterogeneously to widespread increases in air temperatures (O'Reilly et al. 2015). The response of lakes to variation in air temperatures can be dependent on interactions among air temperature, maximum depth, and water clarity (Snucins and Gunn 2000, Rose et al. 2016). Clearer lakes have greater light penetration, making them more susceptible to increases in water temperature, both in the mixing zone and in the hypolimnion, as well as inducing deeper thermoclines (Schindler 2001, Read and Rose 2013). Alternatively, warm years may induce shallower summer thermocline depths due to more rapid warming of surface waters and onset of stratification earlier in spring (Robertson and Ragotzkie 1990, Snucins and Gunn 2000). Thus, clear lakes, such as the set of reference lakes in this study (all types have mean Secchi depth > 4 m), will be variably susceptible to the effects of climate warming based on future precipitation patterns and any changes in water clarity (Read and Rose 2013, McCollough et al. 2019). The connections among lake clarity, precipitation, and air temperature in a changing climate have implications for shifting baselines of expected lake clarity trends and should continue to be the focus of future studies.

Lake-specific variables beyond maximum depth categories were not incorporated in this DFA, but other factors certainly have an effect on the weather sensitivity related to water clarity. For example, an important variable that might be considered in local-scale studies is the lake-specific potential for sediment P release or retention, depending on the sediment aluminum (Al) ratios with iron (Fe) and P, as determined by sequential extractions of sediment (Psenner and Pucsko 1988, Kopáček et al. 2005, Lake et al. 2007). Irreversible sequestration of PO_4 by sediment Al diminishes sediment release of P during anoxia that could cause higher algal productivity, leading to reduction of water clarity. Common trends identified through the DFA may have captured some latent factors (Zuur et al. 2003) that incorporate similarities among lakes or that were associated with lake types.

Our analysis was possible because of long-term datasets, which are invaluable for enhancing our understanding of how ecological systems change and react to their environments over time. We show that long-term data from minimally-disturbed reference sites can improve ecological assessments by separating site-specific effects from regional shifts in baseline conditions. We were able to account for nonlinear dynamics in lake condition over time by developing predictive smoothing curves to measure shifts in baseline water clarity through time. Without long term-datasets, it would be impossible to account for these nonlinear dynamics, which represent interactions among regional weather patterns and local watershed activity. Annual data collection from reference lakes, which is not always a funding priority for agencies (but should be), is important for research examining trends in water quality and interactions between lakes and their watersheds.

This work has established a connection between weather patterns and the temporal shifts in baseline lake conditions that may be expected to vary during climate change. By enhancing our understanding of the relationship between precipitation and water clarity over time, this research may aid in the interpretation and prediction of the effects of climate change on water clarity trends in pristine lakes in Maine and elsewhere.

CHAPTER 5. ASSESSMENT INDICES OF LITTORAL HABITAT CONDITION FOR LAKES IN MAINE AND NEW ENGLAND, USA

Chapter Abstract

Littoral habitat is critical for lake biota but it is adversely affected by residential shoreland development through the loss of riparian and nearshore vegetation and reduced structural complexity. However, there currently exists no assessment methodology for evaluating littoral habitat condition at the single-lake scale in the northeast US. We addressed this assessment need by creating multi-metric indices of littoral habitat condition that focus on residential development as the stressor. We did this by calculating a collection of littoral metrics based on the NLA (National Lake Assessment) Physical Habitat (PHAB) survey field observations, the literature, and others created for this study. Metrics were used to build Linear Discriminant Analysis (LDA) models to find the best combination of littoral habitat measures to predict site classification based on measures of shoreland disturbance. Lake PHAB survey data were used from NLA surveys as well as state-level surveys completed in Maine, New Hampshire, and Vermont. We partitioned data into six groups of lakes: two groups in a Maine-only dataset (Deep and Shallow lakes, n = 102), and four groups of New England regional lakes (Deep-Large, Deep-Small, Shallow-Large, Shallow-Small, n = 361). The two Maine LDA models showed prediction success rates in model validation datasets >90%, while the four Regional Models ranged from 80.8% to 85.4%. We used 95% bootstrapped confidence intervals based on LDA scores from each site on a lake to make assessment designations of natural (meeting reference quality), impaired (not meeting reference quality), or intermediate (existing between natural and impaired) littoral habitat condition for each lake. Regional Deep-Large lakes had the largest proportion of natural lakes (47%), while Regional Shallow-Small lakes had the highest proportion of impaired lakes (47%) among the six lake groups. Our results show that efficacious single-lake littoral habitat assessments may be completed within the framework of

NLA PHAB methodology, but that confidence in assessment results, and therefore better-informed management decisions, can be improved with finer-scale observation data.

Introduction

Suitable littoral habitat is essential for lake biota, but residential development along lakeshores has detrimental effects on the structure and function of these habitats (Kaufmann et al. 2014a). Physical habitat in the littoral areas of lakes is commonly evaluated in three major components: coarse woody habitat (CWH; submerged tree debris), littoral macrophytes, and substrate composition (e.g., percentages of boulders, cobble, gravel, sand, silt, and clay-sized particles). CWH provides structure for large and small fish, and colonization surfaces for bacteria, surficial diatoms, and macroinvertebrates (Everett and Ruiz 1993, Roth et al. 2007, Lawson et al. 2011, Twardochleb et al. 2016, Dustin and Vondracek 2017). CWH may be diminished in littoral areas where vegetation has been removed from nearby riparian areas, which reduces recruitment of new woody structure (Christensen et al. 1996, Jennings et al. 2003, Francis and Schindler 2006, Marburg et al. 2006, Dustin and Vondracek 2017, Chhor et al. 2020). Deliberate CWH removal may occur in littoral areas intended for recreational swimming or boating (Lepore 2013). Loss of CWH has cascading effects on prey availability, growth rates, reproductive success, and mortality across trophic levels in lake communities (Everett and Ruiz 1993, Schindler et al. 2000, Sass et al. 2006, Helmus and Sass 2008, Brauns et al. 2011).

Macrophyte communities are an important biological component of lake littoral areas, but also provide habitat structure for other lake biota. Littoral macrophytes can influence spawning, refuge, and feeding in fish communities (Crowder and Cooper 1982, Savino and Stein 1982) and provide important nursery grounds for young fish (Hayse and Wissing 1996, Weaver et al. 1997). Macrophyte beds help to keep lake water clear through sediment retention, nutrient uptake, and absorption of wave energy (Scheffer et al. 1993). Alterations in macrophyte community composition have been associated with human land use patterns along lake shores, as measured by declines in community composition metrics

(Hatzenbeler et al. 2004, Mikulyuk et al. 2011, Beck et al. 2013) or changes in relative abundances of functional groups (Radomski and Goeman 2001, Alahuhta et al. 2014, Dustin and Vondracek 2017, Chhor et al. 2020). The effects of shoreland development on littoral macrophyte communities may be more pronounced in deeper lakes with lower watershed development, indicating cross-scale linkages among littoral communities, shoreland development, lake morphology, and land use in the larger watershed (Beck et al. 2013). Declines in faunal species that depend on macrophytes for structural habitat have been related to altered macrophyte communities associated with riparian disturbance (Butler and DeMaynadier 2008).

Littoral substrate composition may be altered in areas where lake shoreland has been anthropogenically developed. This is especially relevant in cobble and gravel habitats, which provide important interstitial spaces for macroinvertebrates and incubation areas for developing fish eggs. Erosion from the shoreland area can cover the littoral zone with fine sediments (e.g., sand or silt), restricting water circulation in these spaces and eliminating this habitat for oxygen-sensitive species (Merrell et al. 2009, Horne 2020, Ostendorp et al. 2020).

The 2007 National Lakes Assessment (NLA) found that 36% of the U.S. lakes had poor lakeshore habitat (i.e., little structural complexity and potentially missing components), and lakes with poor lakeshore habitat were three times more likely to be in poor biological health (USEPA 2009). In the Northern Appalachian Ecoregion, where New England is located (Figure 5.1A), 57% of lakes had moderate or high levels of lakeshore disturbance. Additionally, 55% of Northern Appalachian lakes had fair or poor shallow water habitat. For comparison, only 20% of lakes were in fair or poor condition with respect to phosphorus concentration, which is commonly regarded as the primary stressor to most northeastern US lakes. The 2012 NLA supported the findings of the 2007 NLA, showing that 51% of the nation's lakes had lakeshore habitat classified as 'most disturbed' or 'moderately disturbed' (USEPA 2016). Reduced riparian vegetation cover was identified as a primary

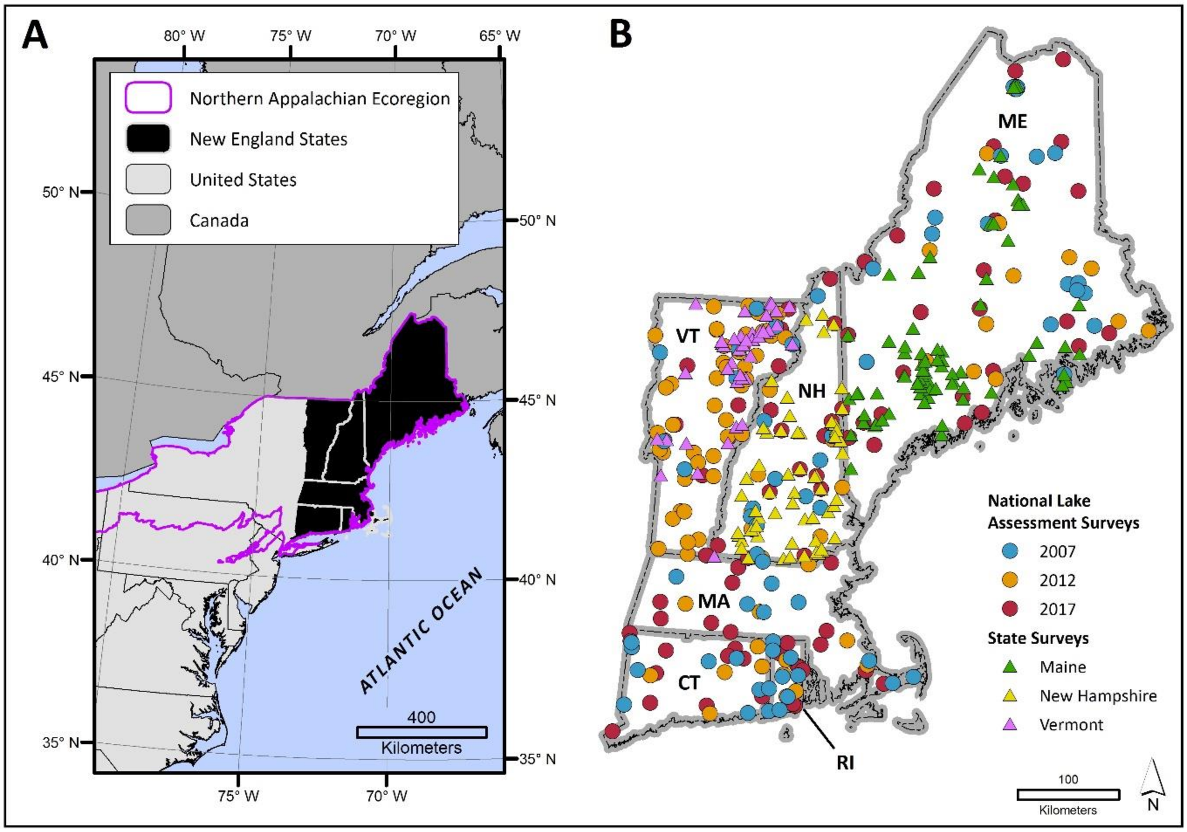


Figure 5.1. A) Locations of the six US states comprising the New England region and the Northern Appalachian ecoregion used in the National Lakes Assessment. B) Lakes included in this study to create littoral habitat assessment models. Lakes surveyed as part of the National Lakes Assessment are circles; lakes surveyed as part of state-level surveys are triangles. CT = Connecticut, ME = Maine, MA = Massachusetts, NH = New Hampshire, RI = Rhode Island, VT = Vermont.

stressor to 21% of Northern Appalachian lakes, behind only phosphorus (31%) and nitrogen (22%). For New England specifically, the 2012 NLA results indicated that 28% and 18% of lakes were in the 'most disturbed condition' for lake habitat complexity (a combination of riparian and littoral habitats) and shallow water habitat, respectively (USEPA 2016).

Shoreland disturbance is generally included as a contributing factor to lake eutrophication (Soranno et al. 1996, Dennis 1986, Garrison and Wakeman 2000, Garrison et al. 2010), but systematic procedures to isolate its effect on littoral habitat have not been largely incorporated into assessments of individual lakes. The purpose of the NLA is to evaluate lakes across the US, not to serve as a tool for evaluation of single lakes. However, single-lake evaluation is a primary objective of most state environmental agencies and other lake-focused resource managers. NLA research determined that lakeshore disturbance is a major concern for lakes in the Northeast US because of its adverse effects on littoral habitat (USEPA 2009 and 2016, Kaufmann et al. 2014a). These findings identified a gap in assessment needs in this region, especially in Maine where lake habitat condition is specifically addressed in the water quality statute for lakes (MRS 38, §465-A). Recognizing this need, we sought to develop an assessment methodology within the framework of the NLA that would support the evaluation of littoral habitat based on comparisons to a natural, minimally disturbed reference condition (Stoddard et al. 2006). We achieved this by slightly modifying the NLA physical habitat survey approach and recalculating a collection of potential metrics to develop indices of littoral habitat condition. Our goals were to 1) numerically describe the range of littoral habitat conditions in natural (reference quality) settings in New England lakes, 2) isolate the effects of shoreland development as a specific stressor to littoral habitat, and 3) develop a systematic assessment method that objectively determines if littoral habitat is measurably different, or not, from the expected reference condition of individual lakes.

Methods

Physical Habitat Surveys

Physical habitat (PHAB) survey protocols and lake selection methods for the NLA are described in Kaufmann et al. (2014a, 2014b) and USEPA (2017). In brief, the percent coverage of a variety of littoral and riparian habitat characteristics (Table 5.1) are estimated by trained field crews at 10 randomly placed, equidistantly spaced shoreland stations (Figure 5.2). Each station consists of a 10 m x 15 m littoral plot (linear longer edge along shore) and a 15 m x 15 m riparian plot. Cover classes are used to estimate percent coverage of various habitat components. NLA cover classes include absent (0%), sparse (0-10%), moderate (10-40%), heavy (40-75%), and very heavy (>75%). Littoral habitat coverage estimates consist of three functional groups of aquatic macrophytes (submerged, floating-leaved, and emergent), fish cover types (large and small woody habitat, total macrophytes, live trees, overhanging riparian vegetation, ledges, boulders, and human structures), and substrate composition. Multiple layers of the vegetative canopy in the riparian plot are estimated, from ground cover to trees ≥ 5 m tall. Indicators of human influence within and adjacent to PHAB plots are recorded and enumerated as any of 13 different types of influences observed within (one point) or adjacent to (0.5 points) sites. NLA surveys occur once every five years, with the first three assessments completed in 2007, 2012, and 2017.

Lake Selection

NLA survey lakes were selected by the US EPA with a spatially balanced probability sampling design from a set of US waterbodies with surface area ≥4 ha (2007 NLA) or ≥1 ha (2012 NLA and 2017 NLA) and ≥1 m maximum depth (Figure 5.1B; USEPA 2017). Additional reference lakes were hand-picked by the EPA to represent lakes in minimally-disturbed watersheds within each ecoregion.

Table 5.1. Physical Habitat (PHAB) metrics for littoral, riparian, and human influence field observations collected in the National Lake Assessment surveys. All observations were enumerated with percent cover classes except for human influence metrics, which were scored either a one (inside plot), 0.5 (adjacent to plot), or 0 (not observed near or in plot).

Littoral habitat structure:	
SUBMGT	Submergent plant cover
EMRGT	Emergent plant cover
FLOAT	Floating-leaved plant cover
FC_LWH	Large woody habitat (>30 cm diameter) cover
FC_SWH	Small woody habitat (<30 cm diameter) cover
FC_LIVETR	Live tree cover
FC_OVERH	Overhanging vegetation cover
FC_LEDGE	Ledges cover
FC_BOULD	Boulder cover
FC_HUMAN	Human structure littoral habitat cover
SUB_BEDR	Percent bedrock cover
SUB_BOULD	Percent boulder cover
SUB_COB	Percent cobble cover
SUB_GRAV	Percent gravel cover
SUB_SAND	Percent sand cover
SUB_SCM	Percent silt, clay, or muck cover
SUB_WOOD	Percent wood cover
SUB_ORG	Percent organic detritus cover
Riparian vegetation/condition:	
C_BIG	Large tree canopy cover (> 5m tall and > 30 cm DBH)
C_SM	Small tree canopy cover in riparian plot (> 5m tall and < 30 cm DBH)
UN_WDY	Large shrub and sapling cover (0.5 m - 5 m tall)
GC_WDY	Ground woody vegetation cover (< 0.5 m tall)
GC_INUN	Percent riparian plot cover inundated with water
GC_NONW	Ground non-woody vegetation cover
Human Influence (HI) (in or adjacent to PHAB plot):	
HI_Buildings	Buildings
HI_ParkBeach	Recreational park facilities or human-made beach
HI_Comm	Commercial facilities
HI_DocksBoats	Docks or boats
HI_RowCrop	Row crop agriculture
HI_Fields	Agricultural pasture or hay fields
HI_Orchard	Orchards
HI_Lawn	Lawns
HI_Wall	Walls or dikes
HI_Trash	Trash or landfill
HI_Roads	Roads
HI_Powerlines	Powerlines
HI_LowImpTr	Low-impact trails or other minimal sign of human influence

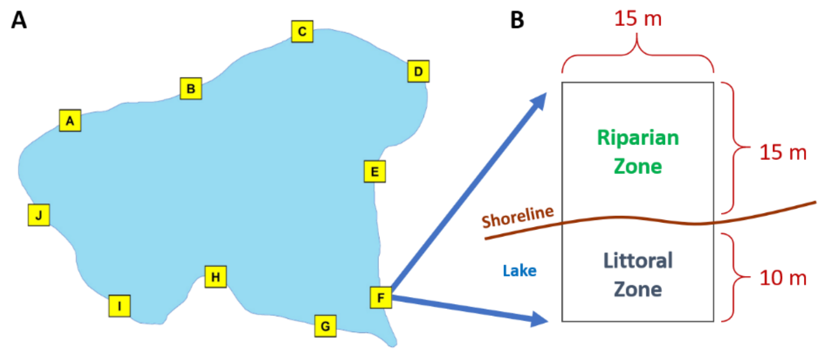


Figure 5.2. A) NLA Physical Habitat survey design of 10 equidistant, randomly placed points around a lake (A-J). B) Schematic of riparian and littoral survey plots that are established at each of the 10 stations.

Many of the EPA-selected lakes for NLA in New England (Figure 5.1) are quite small due to the high prevalence of small waterbodies in this region (e.g., 70% of Maine lakes are ≤ 10 ha). While small lakes are ecologically important, they do not receive the same intensity of shoreland development pressure as larger lakes in New England. To better represent larger lakes and their greater development stress, additional state-level surveys were completed in Maine, New Hampshire, and Vermont on hand-picked lakes that reflect a broader gradient of lake size and shoreland development based on desktop reconnaissance with aerial imagery and local knowledge regarding the shoreland condition of various lakes. Lakes included in this analysis were surveyed between 2007 and 2019. Survey data from these hand-picked lakes were combined with NLA PHAB survey data from these states as well as Connecticut, Massachusetts, and Rhode Island for creation of New England Regional habitat assessment models (Figure 5.1B).

We investigated the efficacy of littoral habitat condition indices for Maine lakes, as well as for lakes in all of New England, including Maine lakes (hereafter, “Regional”). Departures from standard NLA methodology occurred in state-level surveys in two instances: Maine split the moderate cover class (10-40%) into two groups: 10-25% and 25-40% because of the large number of observations that occurred within 10-40% coverage during early pilot surveys. The Maine-only analysis used all six cover classes collected in those surveys. Vermont recorded absolute percentage values for all observations. Maine and Vermont state-level survey data were both included in Regional analyses, but data from those states were reduced to the five cover classes used in NLA so that those data would harmonize with the remainder of the Regional surveys that were collected with standard NLA methodology (five cover classes).

Vermont crews did not collect full substrate composition (but did record percent sand coverage of littoral plots) in state-level surveys, but the number of state surveys they completed made a substantial contribution to the Regional dataset. Consequently, Regional models were calculated without metrics requiring complete substrate composition.

Non-metric multidimensional scaling was used to test patterns of littoral habitat metrics across gradients of lake trophic condition, depth, surface area, ecoregion, and state to investigate various groupings of lakes that may reflect gradients in natural habitat condition (*natural* is defined here as not being measurably different from the conditions measured in minimally-disturbed reference sites). The most informative environmental gradients identified were lake depth and surface area (Figure S1). Therefore, habitat condition models were developed for two Maine lake groups (Deep and Shallow) and four Regional lake groups (Deep-Large, Deep-Small, Shallow-Large, and Shallow-Small). Sub-groupings of deep and shallow lakes in the smaller Maine-only dataset based on lake surface area had decreased model performance so were not used here. The cut-off value for deep and shallow lakes was 10 m maximum depth, which is limnologically relevant in a classification of Maine lakes, as it is associated with many other lake variables including trophic condition (Deeds et al. 2020). Other studies have also determined that lake depth has an association with the habitat and biological communities in littoral areas (Beck et al. 2013, Lewin et al. 2014). The criteria designating large lakes were ≥ 81 ha (200 ac) for Regional deep lakes and ≥ 12 ha (30 ac) for Regional shallow lakes, as these surface areas roughly bisected the Regional depth-stratified lake datasets. The Maine-only analysis included 102 lakes; 361 lakes were included in the Regional analysis.

Metric Calculations

Raw PHAB field observations, from NLA methods (Table 5.1), were used to calculate a variety of metrics for potential inclusion in linear discriminate models (Table 5.2). Characteristics of riparian and littoral habitats, and human influence metrics used in the NLA, were calculated following Kaufmann et al. (2014a). Additional metrics were calculated according to Miler et al. (2014) and new ones were created for this study. All metrics were scaled 0-1 based on the minimum and maximum values of each metric observed within each lake group. In contrast to the NLA approach, which groups all site data for each lake (A-J) and evaluates that lake as a single site, we maintained separate site scores so that each lake could be evaluated with scores from 10 sites each. This was to account for lakes with differing levels of shoreland development along different sections of the shore. Whole-lake metrics were included in the collection of candidate lake metrics to capture potentially important lake-wide measures. Because PHAB observations are based on cover class bins, we used the mid-point of each bin in relevant metric calculations.

Determining Reference vs. Highly Developed Shoreland sites

We adapted the lakeshore human disturbance index used in NLA (RDis_IX, Kaufmann et al. 2014b) to a metric that focuses on residential development (e.g., buildings or lawns) as the primary anthropogenic stressor on lakeshores (RDis_Site; Table 5.2), rather than a focus on agriculture as the primary human influence stressor on lakeshores. Agricultural land use in the watershed can be the most significant stressor on the entire lake in some situations, but typically littoral habitat condition is controlled by other human activities. We adapted this metric to better represent residential development as the primary stressor on New England lakeshores, and to address the impact that it can have on littoral habitat (Christensen et al. 1996, Francis and Schindler 2006, Marburg et al. 2006, Merrell

Table 5.2. Physical Habitat (PHAB) calculated metrics for littoral, riparian, and human influence.

Name	Source[1]	Description
Littoral habitat structure:		
TOTMAC	N	Total macrophyte cover: (SUBMGT + EMRGT + FLOAT)
FLTEMG	N	FLOAT + EMRGT plant cover
SUBFLT	N	SUBMGT + FLOAT leaf plant cover
FC_NAT	N	Sum of all fish cover metrics except FC_Human
FC_ALL	N	Sum of all fish cover metrics
SUB_VAR	N	Littoral substrate site variety (number of substrate types at site)
SUB_PSIZE	N	Littoral Substrate Mean Particle Size: Cover-weighted $\log_{10}$ mean diam (mm); estimated from substrate cover type percentages
SMNTCVR	N	Total of some natural fish cover types (FC_SWH, FC_LIVETR, FC_OVERH, FC_LEDGE, FC_BOULD)
LITCVR_Q	N	Shallow water habitat index: (SMNATCVR / 1.5 + FC_LWH/0.2875 + FLTEMRG / 1.515) / 3
LITCVR_B	N	Alternative shallow water habitat index: (FC_NAT + (FC_LWH / 0.2875)) / 2
LITCVR_C	N	Alternative shallow water habitat index: (FC_NAT + (FC_LWH / 0.2875)) + ((EMRGT + FLOAT) / 1.515) / 3
LK_SUB_VAR	N	Littoral Substrate Variety (Mean number of Substrate Types per site in lake)
LK_LITCVRQ	N	Lake Mean for LITCVRQ across all sites
LK_LITCVRB	N	Lake Mean for LITCVRB across all sites
LK_SMNTCVR	N	Lake Mean for SMNTCVR across all sites
TWH	C	Total Woody Habitat (FC_LIVETR + FC_OVERH + FC_LWH + FC_SWH)
TLWH	C	Total Littoral Woody Habitat (FC_LIVETR + FC_LWH + FC_SWH)
FINES	C	Total substrate fines (SUB_SCM + SUB_ORG)
ROCKY	C	Total rocky substrate (SUB_BOULD + SUB_COB + SUB_GRAV)
TOTFC	N	All fish cover metrics
HCI	C	Habitat Complexity Index: (FC_LWH + FC_LIVETR + FC_LEDGE + FC_BOULD)
VCI	C	Vegetation Community Index: (TOTMAC) / (1 - (SUB_BEDR + SUB_BOULD))
SHANNON	M	Shannon-Wiener habitat diversity of: TOTMAC, FC_LWH, FC_SWH, FC_LIVETR, FC_OVERH, FC_LEDGE, FC_BOULD
SIMPSON	M	Simpson habitat diversity of: TOTMAC, FC_LWH, FC_SWH, FC_LIVETR, FC_OVERH, FC_LEDGE, FC_BOULD
EVENNESS	M	Evenness of habitat features: TOTMAC, FC_LWH, FC_SWH, FC_LIVETR, FC_OVERH, FC_LEDGE, FC_BOULD
Riparian vegetation/condition:		
rvpCAN	C	Presence or absence of lower-level vegetation canopy (C_SM or UN_WDY)
rviCanopy	N	Total canopy cover: (C_SM + UN_WDY)
rviWoody	N	Total cover of riparian vegetative layers: (C_SM + UN_WDY + GC_WDY + GC_INUN)
rviwoody2	N*	Natural Riparian Cover Complexity: Total cover of riparian vegetative layers (C_SM + UN_WDY + GC_WDY + GC_INUN + GC_NONW + rvpCan)
RVeg_Site	N*	Site-level Riparian Vegetation Condition: (((rviWoody2/max(rviWoody2)) + (GC_INUN /max(GC_INUN)))/2)
Human Influence Metrics:		
HI_LB	C	Site Human Influence score for lawn and buildings: (HI_Lawn + HI_Buildings)
HI_NONLB	C	Site Human Influence score for all HI metrics except HI_Lawn and HI_Buildings
HI_ANY	N	Proportion of stations around lake with any human influence within plots (# of sites with Human Influence score > 0 / Total number of sites)
RDis_Site	N*	Lakeshore Human Disturbance Index – Site level: (1 - (1/(1+(HI_NONLB) + (5*HI_LB))) + HI_ANY)/2
RipScore	C	Riparian Condition Score: (1+RVeg_Site) / (1+RDis_Site)

[1]N = NLA (Kaufmann et al. 2014a); M = Miler et al. 2014; C = current study.
*metrics were modified from their original formulae.

et al. 2009, Brauns et al. 2011, Dustin and Vondracek 2017). We also adapted the NLA Riparian Vegetation Complexity Index (RVegQ, Kaufmann et al. 2014b) to be a more appropriate measure of natural riparian vegetation at New England lakeshores. This was done by including the cover and complexity of all layers of riparian vegetation and natural ground cover (e.g., duff vs. bare ground) to create the metric RVeg_Site. RDis_Site and RVeg_Site were combined into a single metric, RipScore (Table 5.2), to place all sites along a gradient of most natural (minimal human influence, high RipScore) to most disturbed (greatest human influence, low RipScore).

Sites were designated as "reference" that scored above the 80th percentile of RipScore values, and "developed" if they fell below the 20th percentile of RipScore values as these thresholds provided the best separation of lakeshore condition. Sites designated as either reference or developed were used to create separate datasets for building linear discriminant models to predict site condition based on littoral habitat metrics. The model-building datasets had an equal number of sites designated as either "reference" or "developed". The model-building datasets were randomly partitioned into training and validation sub-sets (80% and 20% of sites, respectively). The training sub-set was used to build and train the candidate models, and the validation sub-set was to test the effectiveness of the candidate models with new data. Overall model effectiveness was evaluated based on the percent of sites correctly classified (as reference or developed) by the model in both the training and validation sub-sets of data.

Linear Discriminant Modeling

Linear discriminant analysis (LDA) models were built for the six groups of lakes discussed above, using the MASS package in R (Venables and Ripley 2002). LDA effectively identifies the best combination of metrics for classifying new data with multi-metric indices (Danielson et al. 2012). Combinations of metrics were tested in as many as 3,000,000 iterations per model (e.g., 10 metrics in a model x 25 potential metrics) for each lake. Models were chosen with combinations of metrics that best predicted (discriminated) the correct affiliation of sites between the *a priori* reference and developed site categories and had no correlated metrics (Variance Inflation Factors < 5; Fox and Monette 1992).

The final step of our model selection process was a sensitivity analysis which measured the change in model effectiveness when a particular metric is absent from that model. To do this, we ran successive LDA models within each lake group that were missing one of the metrics from the full model. We enumerated the importance of each metric within each model by calculating the difference in the percent of correctly identified sites from training and validation datasets between full *n*-metric models and *n-1* metric models, creating the variables ΔTrain% and ΔValid% for each metric, respectively. ΔTrain% and ΔValid% represent the change in the percent correct from each category (new value – original value), so negative values indicate declines in model performance. Metrics which did not decrease model performance (ΔTrain% and ΔValid% ≥ 0) when absent were omitted from final models. We also used the results of the sensitivity analyses to evaluate the importance of each metric within the final version of each model, because metrics with lower ΔTrain% and ΔValid% values (i.e., more negative) had a greater effect on model success.

LDA models were applied to all sites from each lake to place lakes in habitat condition assessment categories. LDA scores are the linear combination of the products of metric values and their model coefficients. Here, LDA scores were calculated for each of the 10 sites. Higher (positive) LDA scores indicated that the site has habitat closer to undisturbed reference conditions, and lower (negative) values indicate habitat conditions were closer to those found in disturbed sites. LDA scores near zero designate intermediate habitat condition. Bootstrapped 95% confidence intervals (CIs) were calculated from the 10 LDA scores for each lake with the Boot package in R (10k resample with replacement; Canty and Ripley 2021). The accelerated bias-corrected (BCA) method was used because it is non-parametric and corrects for skewness and bias in the bootstrap distribution. It is considered the most consistently effective technique for calculating CIs (Puth et al. 2015). The CIs based on the 10 LDA scores for each lake were used to determine if the lake had littoral habitat more indicative of natural or developed shoreland conditions. This approach of using CIs is effective at capturing the variation and overall condition of the sites based on these scores and their distance from zero. CIs that do not cross zero indicate that the data are significantly different from zero. If the CI's lower bound for a lake was > 0, the lake was designated as *natural,* meaning the habitat is reflective of reference-quality conditions; if the CI upper bound was < 0, it was designated as *impaired,* meaning that the habitat is measurably different from the expected reference condition; and if the CI range contained 0, the lake was assigned *intermediate* status, indicating that the habitat condition exists between *natural* and *impaired*.

Results

The top LDA models for the Maine-Deep and Maine-Shallow PHAB data both had over 90% of sites correctly classified in validation datasets, with 90.5% of the Maine-Deep and 90.6% of the Maine-Shallow sites correctly classified (Table 5.3). The percent classified correctly in the validation datasets from the Regional lake groupings ranged from 85.4% (Regional Deep-Large and Deep-Shallow) to 80.8% of sites (Regional Shallow-Large). The Maine-only models for deep and shallow lakes showed better model performance by 5.1% and 9.4% each when compared to the average percentage of correctly categorized sites in the validation data from their deep and shallow Regional model counterparts (average of Regional Deep and Shallow lakes combined). LDA models with Maine-only data but with five cover classes instead of six (10-25% and 25-40% cover classes combined into the 10-40% cover class per NLA methodology) had lower model performance. The top-performing model for Maine-Deep lake datasets with five cover classes had 83.5% correct in the training dataset and 88.1% correct in the validation data, representing declines in model performance of 3.3% and 2.4%, respectively. The best Maine-Shallow model with five cover classes had 75.3% and 75.8 % correct in training and validation datasets, respectively, representing declines in respective model performance of 2.5% and 14.8%.

Table 5.3. Metrics used and performance statistics for Linear Discriminant Analysis models for each lake type.

Metric	Maine Deep	Maine Shallow	Regional Deep - Large	Regional Deep - Small	Regional Shallow - Large	Regional Shallow - Small
EMRGT		X	X		X	
FLOAT	X		X			X
FLTEMG						
SUB_SAND	X	X	X	X	X	
SUB_SCM		X				
LK_SUB_VAR				X		X
SUB_VAR	X					
SUB_PSIZE	X					
EVENNESS		X	X			
FC_SMH		X				
FC_LWH	X					X
FC_NAT						
FINES				X		X
HCI		X	X		X	X
LITCVRB		X			X	
LK_LITCVRB			X	X		
LK_LITCVRQ	X				X	X
LK_SMNTCVR	X	X	X	X	X	
SHANNON	X			X	X	X
SIMPSON						
SMNTCVR				X		
TWH			X	X		X
TOTFC			X		X	X
TLWH				X		X
Depth Criteria	≥ 10 m	< 10 m	≥ 10 m	≥ 10 m	< 10 m	< 10 m
Size Criteria	--	--	≥ 200 ac	20-200 ac	≥ 30 ac	< 30 ac
Number of Lakes[1]	56	46	64	58	103	125
Training % Correct	86.8	77.8	80.1	73.4	75.1	75.5
Validation % Correct	90.5	90.6	85.4	85.4	80.8	81.6

[1]*Lakes in the Maine-Deep and Maine-Shallow groups were also used in the Regional models.*

The metrics included in the top models varied across lake groupings, but several were common among multiple lake groups (Table 5.3). SUB_SAND and LK_SMNTCVR were the most common, occurring in five models each. HCI and SHANNON occurred in four models and five metrics (EMRGT, FLOAT, LK_LITCVRQ, TOTFC, TWH) were common among three models. Most metrics decreased with riparian disturbance (lower *RipScore*) with the most noticeable exception of SUB_SAND, which increased with shoreland disturbance (Figure 5.3).

Sensitivity analyses showed that one model performed best with 10 metrics (Regional Shallow Small), two with nine metrics (both Regional Deep models), and the rest with eight metrics (Table 5.4). The absence of remaining metrics would have decreased model performance as measured by ΔTrain% and ΔValid%. Some metrics were not a factor in either the validation or training datasets, having a value of zero for ΔTrain% or ΔValid but not both. These metrics were left in the models because a decline in either value indicates a decline in overall model performance.

Whole-lake metrics (LK_SMNTCVR, LK_LITCVRQ, LK_LITCVRB) were among the most important metrics (lowest ΔTrain% and ΔValid% values) in all lake models (Table 5.4). SUB_SAND was also among the most important metrics in all models (mean SUB_SAND for all models: ΔTrain% = -2.4; ΔValid% = -6.4) but was relatively less important in Maine shallow lakes (sixth most important of eight metrics; ΔTrain% = -2.6; ΔValid% = -3.1). Macrophytes were important in the Regional Deep lakes (RDL: FLOAT ΔTrain% = -1.8, ΔValid% = -4.2; RDS: EMRGT ΔTrain% = -1.8, ΔValid% = -4.1) but were not as important in the Maine Deep model (FLOAT ΔTrain% = -1.1, ΔValid% = -0.0). Substrate variability (SUB_VAR) was important in the Maine deep model but not the Regional deep lake models. Emergent macrophtyes were important in the Regional Shallow-Large lake model (EMRGT ΔTrain% = -1.8, ΔValid% = -4.1), but macrophyte metrics were less important in the other two shallow lake models.

LDA Scores were calculated for each site according to corresponding model coefficient values (Table 5.4). Mean LDA scores (the mean score from all sites on a lake) ranged from -1.820 to 2.428 ($\bar{x}$ = 0.144 ±

0.020 SE) in Maine Deep lakes (Figure 5.4A) and from -2.514 to 2.514 ($\bar{x}$ = 0.263 ± 0.019 SE) in Maine Shallow lakes (Figure 5.4B). Regional deep lake scores ranged from -1.826 to 2.455 ($\bar{x}$ = 0.100 ± 0.014 SE) and -2.048 to 2.324 ($\bar{x}$ = 0.080 ± 0.016 SE) in large and small lakes, respectively (Figures 5.5A, 5.5B). Regional shallow lakes LDA scores ranged from -2.140 to 2.011 in large lakes ($\bar{x}$ = 0.080 ± 0.008 SE) and -1.866 to 3.108 in small lakes ($\bar{x}$ = -0.029 ± 0.008 SE) (Figures 5.6A, 5.6B). All individual lake assessment scores are in Table D1. Upper and lower bounds of 95% bootstrapped confidence intervals based on LDA scores for each lake were used to place lakes in habitat condition assessment categories (Figs 4-6; Table D1).

Lake groups had generally similar proportions of *natural*, *intermediate*, and *impaired* lakes (Figure 5.7). The Regional Shallow-Small lake model was an exception to this with 47% (n = 59) of lakes designated as *impaired*, which was the highest impairment percentage among all lake groups. Regional Shallow-Small lakes also had the lowest percentage of *natural* lakes (30%, n = 37). Maine-Shallow Lakes had only 26% *impaired* lakes (n = 12), the lowest percent *impaired* of all groups. Regional Deep-Small had the highest percentage of *natural* lakes (41%, n = 24).

RDis_Site (human influence) values were highest in Maine-Deep and Regional Deep-Large lakes (although Regional Deep-Large RDis_Site scores were only significantly higher than those of two other lake groups at $\alpha = 0.05$), indicating the most intense human influence occurred within these two lake types (Figure 5.8A). Maine shallow lakes had the highest RVeg_Site values (Riparian vegetation condition) among the six lake types, suggesting this class of lakes had the most natural riparian vegetation condition (Figure 5.8B). RipScore values were more evenly distributed across lake types, but Regional deep lakes were lower than most (i.e., most disturbed; Figure 5.8C). LDA scores across all lake groups were similarly associated with RDis_Site (negatively) and RipScore (positively) (Figures 5.8D, 5.8F). The associations among LDA scores and RVeg_Site scores were slightly more variable but was positive across five lake groups, with Maine Shallow lakes being the exception (Figure 5.8E).

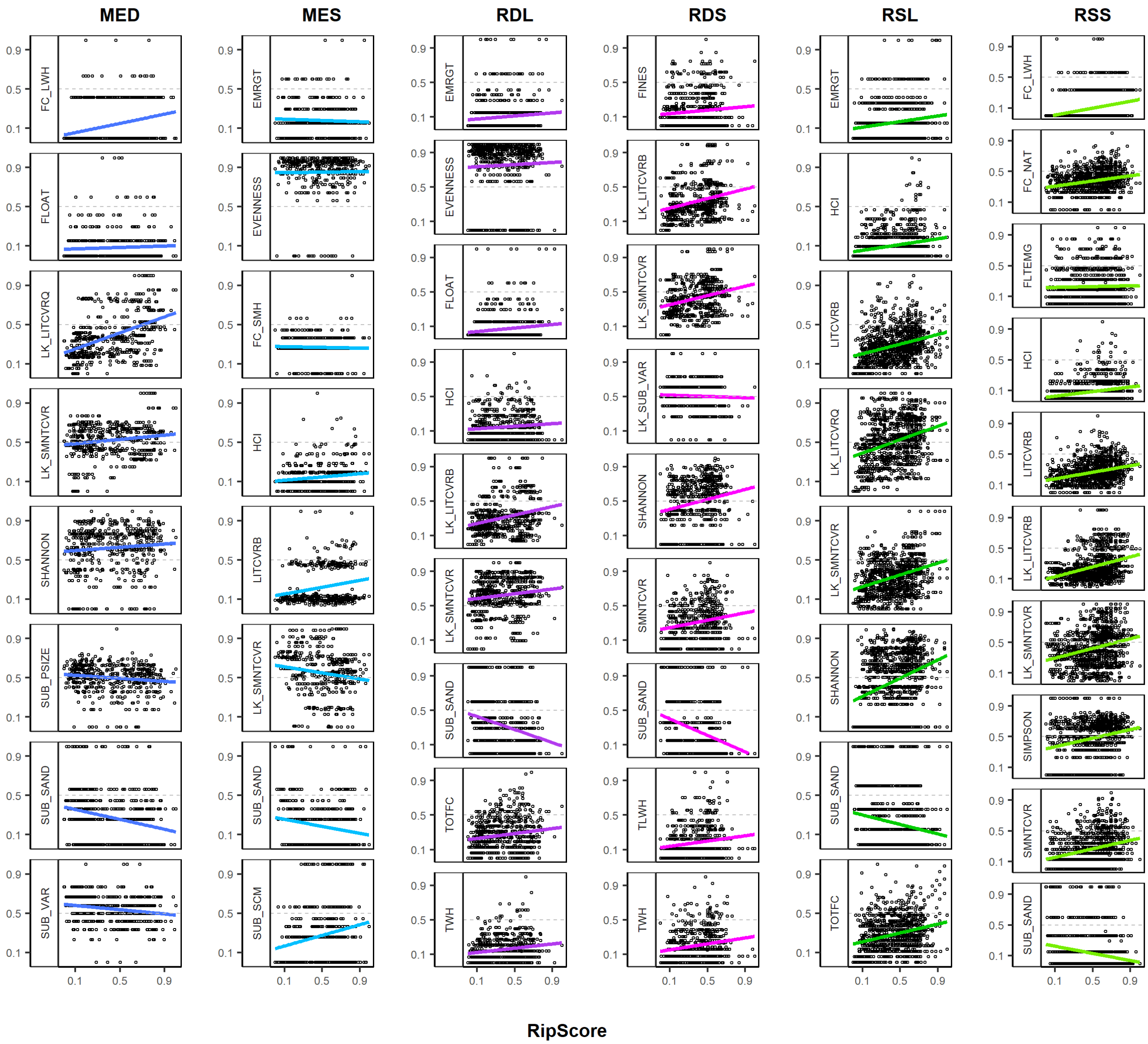


Figure 5.3. Scatterplots of RipScore vs. final model metrics from all six LDA models, with linear regression lines. RipScore was used to evaluate site condition; a higher value indicates a more natural shoreland condition. All metrics have been scaled 0-1 based on the minimum and maximum values within each lake group. MED = Maine Deep, MES = Maine Shallow; RDL = Regional Deep Large, RDS = Regional Deep Small, RSL = Regional Shallow Large, RSS = Regional Shallow Small.

Table 5.4. Model coefficients (LD1) and sensitivity values for Training (ΔTrain%) and Validation (ΔValid%) datasets, which is the change in percent correct in a new model missing only that metric for six Linear Discriminant Analysis models. Metrics are sorted by ΔValid%, which represents the importance of that metric to correct site classification in the validation dataset.

Metric	LD1	ΔTrain%	ΔValid%	Metric	LD1	ΔTrain%	ΔValid%
Maine Deep (MED)				**Maine Shallow (MES)**			
SUB_SAND	-1.333	-2.7	-4.8	HCI	3.916	-4.6	-15.6
SUB_PSIZE	1.423	-2.2	-4.8	LK_SMNTCVR	-3.997	-6.6	-12.5
SUB_VAR	-5.431	-2.2	-4.8	SUB_SCM	1.382	-1.3	-12.5
FC_LWH	0.762	-1.1	-4.8	EVENNESS	2.327	-3.3	-6.2
LK_LITCVRQ	4.049	-8.8	-2.4	LITCVRB	1.667	-3.9	-3.1
FLOAT	0.695	-1.1	0.0	SUB_SAND	-0.647	-2.6	-3.1
LK_SMNTCVR	-0.072	-1.1	0.0	EMRGT	-0.944	-1.3	-3.1
SHANNON	0.657	-1.1	0.0	FC_SMH	-0.758	-1.3	0.0
Regional Deep-Large (RDL)				**Regional Shallow-Large (RSL)**			
SUB_SAND	1.657	-1.8	-4.2	LK_SMNTCVR	3.984	-3.0	-5.5
LK_LITCVRB	3.410	-7.7	-4.2	SUB_SAND	-0.995	-0.6	-5.5
LK_SMNTCVR	0.587	-4.1	-4.2	EMRGT	2.693	-1.8	-4.1
TOTFC	-3.742	-3.2	-4.2	LITCVRB	3.279	-1.2	-4.1
FLOAT	1.657	-1.8	-4.2	TOTFC	-2.589	-1.2	-2.7
EMRGT	2.349	-4.1	-2.1	LK_LITCVRQ	-1.474	-0.6	-2.7
TWH	5.556	-3.2	-2.1	HCI	-0.681	-0.9	-1.3
HCI	1.017	-2.7	-2.1	SHANNON	1.367	0.3	-1.3
EVENNESS	-0.301	-2.3	0.0				
Regional Deep-Small (RDS)				**Regional Shallow-Small (RSS)**			
LK_LITCVRB	2.502	-4.3	-7.4	LK_LITCVRB	3.889	-3.3	-3.9
LK_SMNTCVR	2.633	-2.4	-7.4	HCI	-2.063	0.7	-3.9
SUB_SAND	-1.074	-0.5	-7.4	SUB_SAND	-0.854	-1.4	-3.0
TWH	2.439	-1.9	-2.5	SMNTCVR	4.159	0.0	-3.0
SHANNON	0.705	-0.9	-2.5	FC_LWH	3.035	-0.7	-2.0
LK_SUB_VAR	-1.911	-1.9	0.0	SIMPSON	0.722	0.3	-2.0
SMNTCVR	-1.406	-1.4	0.0	LK_SMNTCVR	-0.994	0.5	-2.0
TLWH	-1.591	-1.4	0.0	LITCVRB	-3.166	-0.9	-1.0
FINES	-0.416	-0.5	0.0	FC_NAT	0.593	0.0	-1.0
				FLTEMG	0.627	-0.7	0.0

Applying models to new survey data (example for Maine-Deep):

LDA Score = 0.695*FLOAT + 0.762*FC_LWH -1.333*SUB_SAND - 5.431*SUB_VAR + 1.423*SUB_PSIZE + 0.657*SHANNON - 0.072*LK_SMNTCVR + 4.049*LK_LITCVRQ

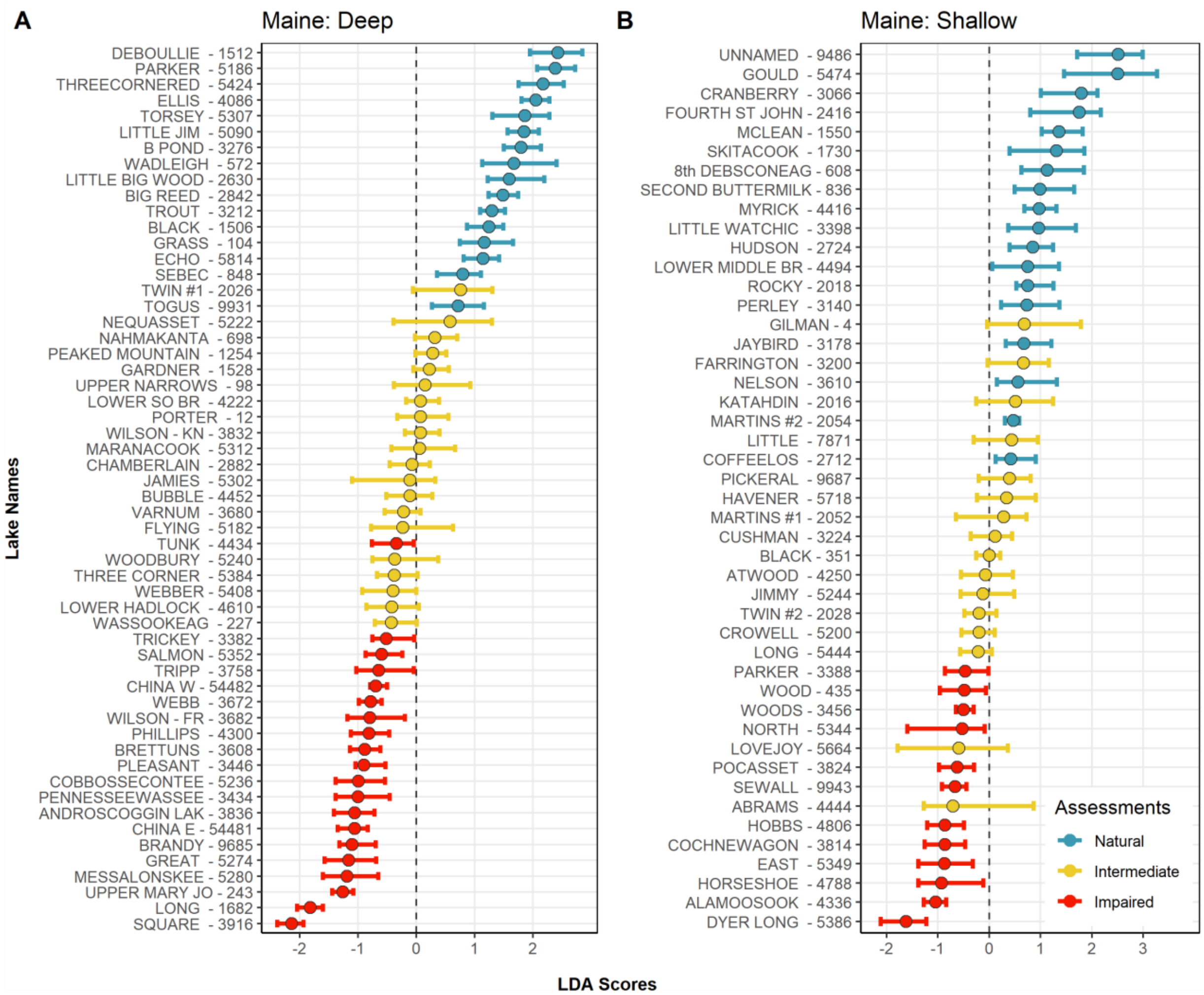


Figure 5.4. Mean LDA scores and bootstrapped 95% confidence intervals (CIs) from models for Maine lakes with habitat condition assessment results. The horizontal bars show the range of the CIs. See Table 5.2 for depth and size criteria.

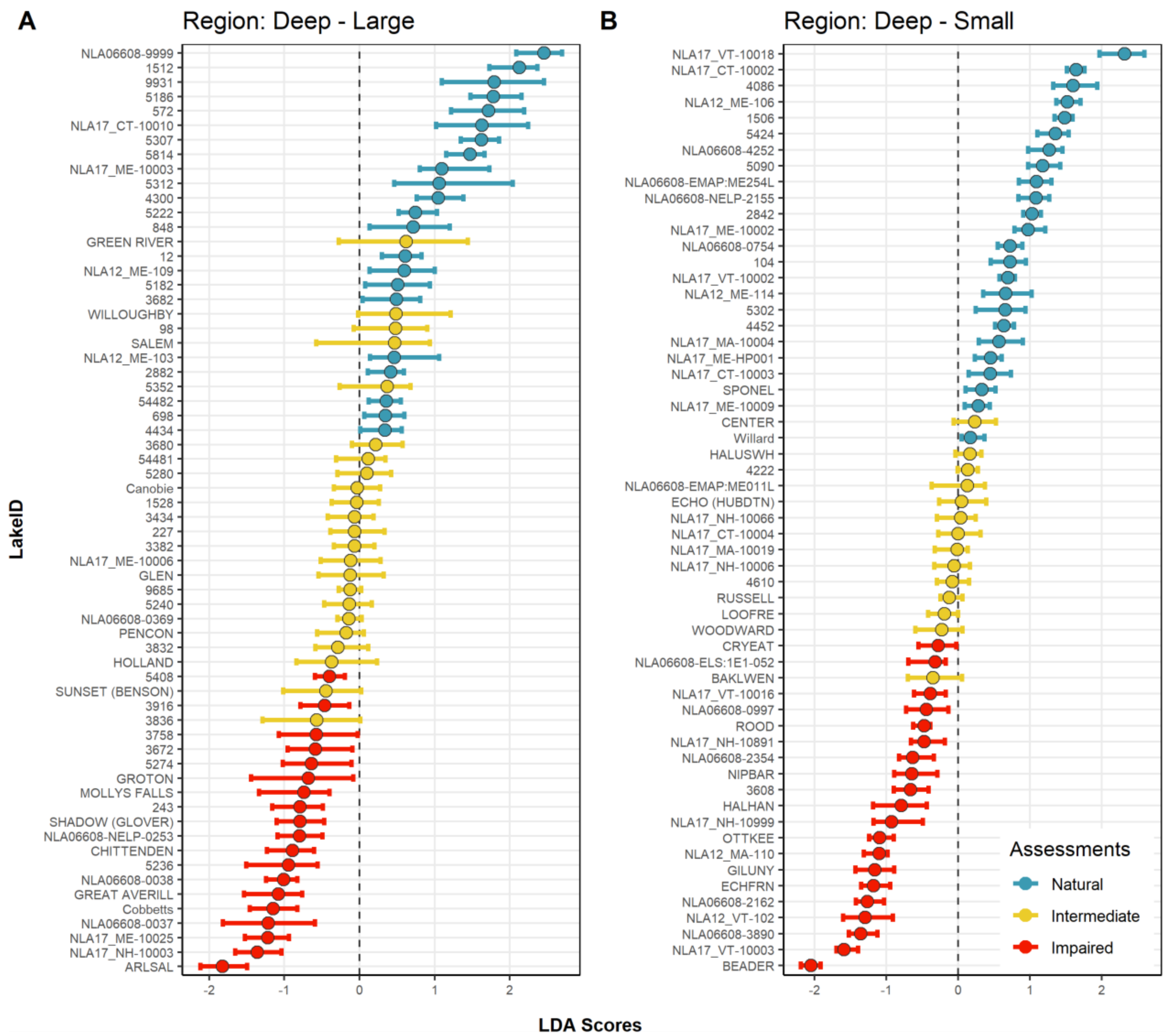


Figure 5.5. Mean LDA scores and bootstrapped 95% confidence intervals (CIs) from models for New England regional deep and shallow lakes with habitat condition assessment results. The horizontal bars show the range of the CIs. See Table 5.2 for depth and size criteria.

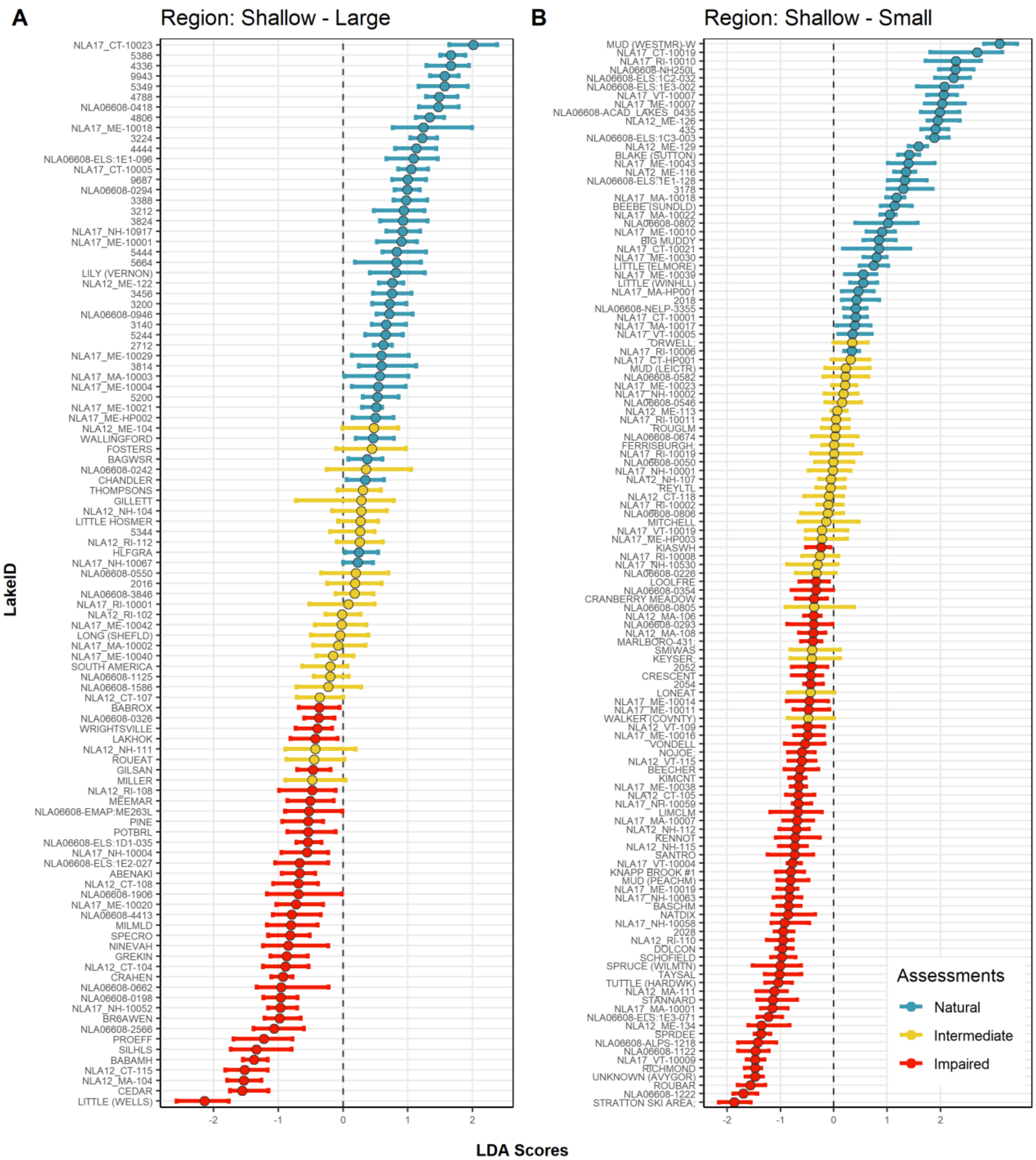


Figure 5.6. Mean LDA scores and bootstrapped 95% confidence intervals (CIs) from models for New England regional deep and shallow lakes with habitat condition assessment results. The horizontal bars show the range of the CIs. See Table 5.2 for depth and size criteria.

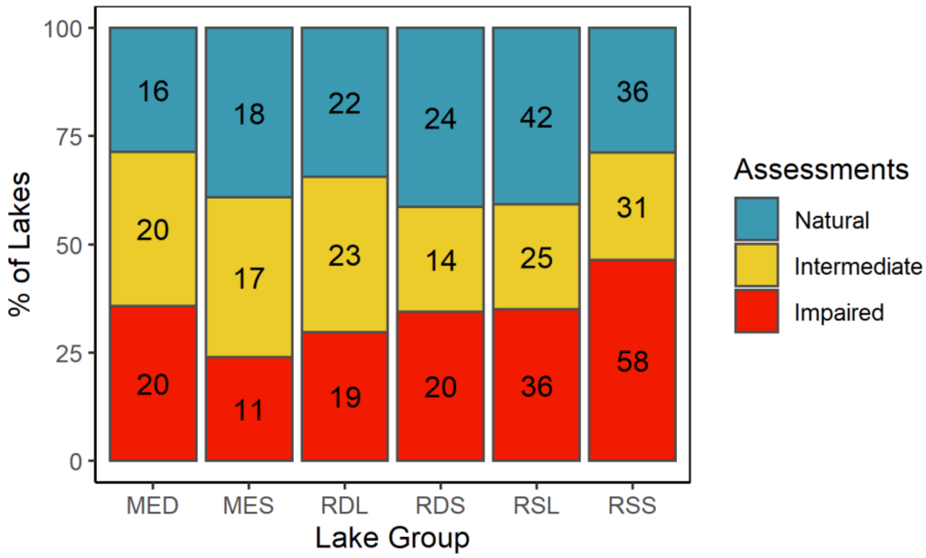


Figure 5.7. Proportion of assessment designations within each lake type. Numbers within each bar segment indicate number of lakes within each assessment category. MED = Maine deep, MES = Maine shallow; RDL = regional deep large, RDS = regional deep small, RSL = regional shallow large, RSS = regional shallow small. See Table 2 for depth and size criteria.

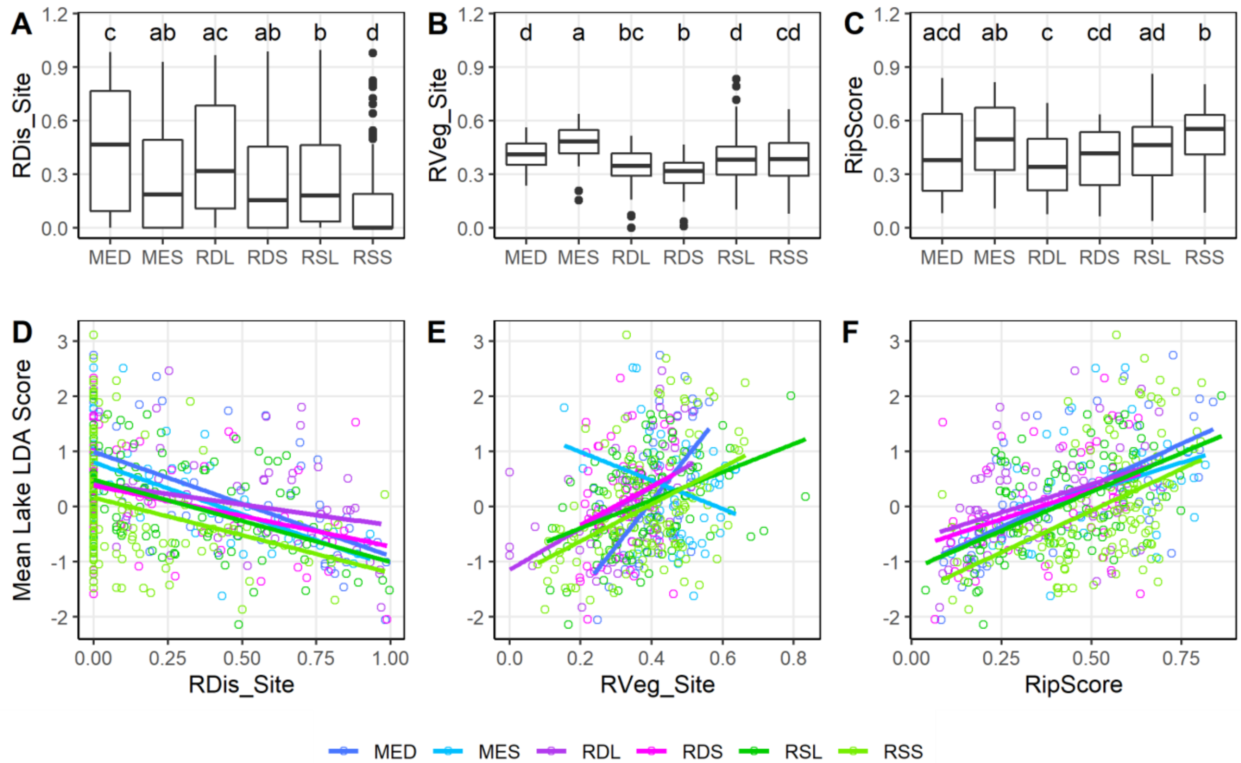


Figure 5.8. A-C: Boxplots of site-level riparian metrics used in the analysis, averaged for each lake within each lake group. *RDis_Site* measures human influence, *RVeg_Site* measures riparian vegetation condition, and *RipScore* is the combination of *RDis_Site* and *RVeg_Site* (see Table 2). Different lowercase letters above boxplots indicate significant differences from Tukey HSD *post-hoc* tests (α = 0.05). D-F: Scatterplots of mean riparian metrics vs. mean LDA model scores for each lake with linear regression lines for each lake group. MED = Maine Deep, MES = Maine Shallow; RDL = Regional Deep Large, RDS = Regional Deep Small, RSL = Regional Shallow Large, RSS = Regional Shallow Small.

Discussion

This study shows that littoral habitat observations collected according to the NLA methods may be used to perform assessments of habitat condition for individual lakes. Models using only Maine lake data had better performance than those based on New England Regional lakes, but all models correctly predicted site condition with >80% accuracy as indicated by validation success rates (Table 5.3). These models provide needed assessment tools that allow resource managers to objectively determine whether or not lake shoreland development is adversely affecting littoral habitat.

We were able to increase our confidence in assessments of Maine lakes by recording cover class observations at a slightly finer scale of assessment, especially in shallow lakes. In the Maine state-level surveys, the 10-40% cover class (used in NLA studies) was divided into 10-25% and 25-40% groups, which likely contributed to the increased LDA model performance compared to that for the Regional models (Table 5.3). Combining observations of habitat cover that occur from 10-40% may obscure important but fine-scale differences among sites. These results suggest that differences in habitat structure at within this coverage range may be most important in shallow lakes where there may be more variety habitat structure and stasis in condition, perhaps due to reduced wave action or lessened littoral slope in shallow areas of lakes (Francis and Schindler 2006, Marburg et al. 2006). Full substrate metrics, which were not used in the regional models, were important in the Maine-only models and may have contributed to increased model performance as well. For future single-lake habitat assessments, it may be most appropriate to develop models that partition cover classes into six or more groups and focus the assessments at smaller regional scales than considered here in the regional models. If new model development in is not possible due to data or resource constraints, this study has shown that a Regional index may be applied using existing NLA PHAB methodology that maintains a suitable level of performance (>80% validation success rates for all Regional models) for most lake littoral habitat assessment applications.

The separation of lakes into morphological groups helped to increase model performance. Deeds et al. (2020) showed the importance of hydrogeomorphic variables in shaping lake conditions across Regional landscapes. Maximum depth was likely an important grouping variable here because of the association of lake depth with several related variables, including lake surface area which was specifically used to partition deep and shallow lakes in the Regional assessment. Deeper lakes are generally larger, which increases the fetch of the lake, permitting greater wave action. Increased wave energy on lakeshores shapes littoral habitat by suspending and removing smaller sediment particles and organic debris, leaving coarser-grained substrates, and redistributing woody habitat to calmer areas of lakes (Marburg et al. 2006). Lake depth may also relate to the slope of the littoral area, which can redistribute woody habitat into deeper depths. Deeper waters may not always be evaluated in PHAB surveys which only extend 10 m from shore.

Lewin et al. (2014) found a more pronounced effect of shoreland development in deeper lakes, attributing the relationship to the proportionally smaller littoral zones found in deep lakes. This was reflected in our regional models, as the two Regional Deep lake models had higher rates of prediction success compared to their Regional Shallow Lake counterparts (with the exception of training percent correct in Reginal Deep-Small lakes). The Maine Deep lakes had increased percent correct in training data but similar percent correct values in the validation datasets. These results suggest that there may be a closer association between littoral habitat condition and lakeshore alteration in deeper lakes (Table 5.3).

While the results of this study cannot be used for general condition assessments of littoral habitat due to the non-randomized selection for many lakes in the dataset, the proportion of assessment categories determined in each model type may be instructive for understanding model results. The Regional analysis showed that larger proportions of Shallow-Small lakes were designated as *impaired* compared to the other three Regional lake groups (Figure 5.7). Regional Shallow-Small lakes had significantly lower lakeshore disturbance (*RDis_Site*) and significantly higher *RipScores* (combination of *RDis_Site* and riparian vegetation condition) than other Regional lakes, indicating that these lakes have comparatively more natural shorelands than other Regional lakes, overall (Figure 5.8A, C). The higher proportion of *impaired* assessments in Regional Shallow-Small lakes may be due in part to the wider variation in habitat condition among these lakes compared to other lakes in New England. Several of the surveyed shallow lakes in this dataset are small, remote ponds with minimal human alteration, while others are located near urban centers and receive a great deal of development pressure. This is reflected in the wider ranges and higher upper end of LDA values among these lakes; Regional Shallow-Small lakes have a maximum LDA of 3.108, which is 0.653-1.097 higher than the maximum scores of other three Regional lake models (Table D1). High values at the upper end of LDA scores of Regional Shallow-Small lakes created a sub-group of 12 lakes with very high LDA scores (Figure 5.6B). Most lakes in this group did not approach the most natural habitat conditions observed, as the top-end LDA scores here were so high. This was reflected in the model results, as the Regional Shallow-Small lake group is the only one with negative mean LDA value and which has the majority of lakes in the *impaired* assessment category (Table D1). A closer examination of lake condition and habitat assessments within this group may help to determine if all lakes assessed as *impaired* have truly degraded habitat or if some assessments are a function of a wider variability of lakeshore conditions within this group.

LDA model scores generally increased (more natural littoral habitat conditions) with decreasing RDis_Site (human influence, Figure 5.8D) and decreased with lower RipScores (human influence and riparian vegetation condition; Figure 5.8F). However, there were exceptions to these associations which resulted in habitat assessments that did not reflect expected habitat conditions (e.g., Upper Jo-Mary Lake in Maine: mean RipScore = 0.723, mean LDA Score = -0.847; result = *impaired*; Fig. 4A and Table D1). Lakes that had high RipScores but low LDA scores indicated that some lakes may have natural factors that influence assessment of littoral habitat condition. For example, shorelines and littoral areas with dense boulder coverage, as is the case in Upper Jo-Mary Lake, typically have homogenous substrate composition and are largely unsuitable for plant growth. This lake is also relatively large (748 ha), so it likely has a large enough fetch to redistribute woody habitat to more protected areas (Marburg et al. 2006). Naturally sandy lake shores can be common in this area of Maine, as well as other parts of New England. These conditions affect multiple metrics that are important in Maine Deep LDA models (sandy shores: SUB_SAND; substrate particle size distribution: SUB_PSIZE; substrate homogenization: SUB_VAR; lack of observed woody habitat and macrophytes: FC_LWH, LK_LITCVRQ, and LK_SMNTCVR;). These results highlight the importance of considering multiple aspects of lakeshores that affect littoral habitat before final condition assessments are made. Shoreland disturbance metrics such as RipScore and RDis_Site may be used to help determine if poor littoral habitat scores are due to shoreland alteration, natural conditions, or other factors such a water level fluctuation.

Manual lake water level fluctuations occur in some lakes in New England, usually either for macrophyte control (Cooke et al. 2005), flood control, or power generation at hydropower facilities (Mjelde et al. 2013). Drawdown can have adverse effects on littoral habitat, especially when coupled with shoreland development (Evtimova and Donohue 2016, Carmignani and Roy 2021). We did not specifically address the effect of water level drawdowns in our littoral habitat condition assessment models, but poor littoral habitat scores in some lakes with otherwise minimally developed shorelands (e.g., Chittenden, Green River, Little Averill) may be explained by seasonal water level fluctuations. Linkages among manual water level manipulations, shoreland development, and littoral habitat condition should continue to be investigated.

SUB_SAND, the percentage of sand coverage within the littoral plot substrate, was present in all models and was among the most important metrics in each lake group except for Maine-Shallow (Table 5.4). The presence of sand coincided with increased shoreland disturbance in all models (Figure 5.3). Sand may be present in abundance within the littoral area and associated with residential development for a variety of reasons. It may be natural, as sandy shore sites may be more desirable for residential home placement and recreational lake use. However, sand in the littoral area may also be due to increased erosion from shoreland development and land uses (e.g., unmitigated earthwork or improperly maintained dirt roads and driveways; Garrison et al. 2010), or the manual placement of sand for beach areas along the shore. Regardless of how the sand arrived in the littoral area, sand represents poor littoral habitat with little complexity, and our results suggest that higher coverage of sand is associated with greater shoreland development and impaired habitat condition.

LK_SMNTVR, which occurred in all six LDA models, represents a variety of habitat structures in the littoral zone (Table 5.2). This is a lake-wide metric, meaning all 10 sites receive the same score, calculated as the average of SMNTCVR across all sites for a lake. Two other lake-wide metrics (LK_LITCVRB, LK_LITCVRQ) also occurred in all lake group models. The importance of a lake-wide metric across all model types signifies that there can be effects of shoreland development that affect littoral habitat throughout entire lakes, such as extensive loss of riparian vegetation which can diminish the amount of CWH in the littoral zone (Christensen et al. 1996, Jennings et al. 2003, Francis and Schindler 2006, Marburg et al. 2006, Dustin and Vondracek 2017, Chhor et al. 2020). HCI (Habitat Complexity Index, present in four models) is site-specific and captures the effects of localized shoreland development. The mix of whole-lake and site-specific metrics (HCI and others) in our models suggests that there are interacting effects of shoreland disturbance operating at multiple scales that influence littoral habitat condition in lakes.

Metrics incorporating several habitat variables (e.g., site-specific: TLWH, FC_Nat, EVENNESS, LITCVRB, HCI, SMNCVR; and whole-lake: LK_SMNTCVR, LK_LITCVRB, LK_LITCVRQ) occurred in all models. These metrics all incorporate various aspects of littoral habitat structure: live and downed trees, various sizes of woody structure, substrate type and complexity, and macrophyte abundance. These types of metrics highlight the natural variability observed in littoral habitat structures. As shoreland development increases, the values of these metrics decrease, signifying homogenization of habitat structure and loss of habitat types from the littoral area. Lakes in need of habitat rehabilitation that are evaluated with habitat models, which depend more heavily on these types of metrics (i.e., low ΔTrain% and ΔValid% values), will likely benefit greatest from a lake-wide approach to shoreland vegetation restoration and maintenance of natural littoral macrophyte beds.

All models contained at least one macrophyte-related metric (EMRGT, FLOAT, FLTEMG, TOTMAC, SUBFLT, VCI, SHANNON, and LK_LITCVRQ, which incorporates FLTEMG). Regional large lakes (deep and shallow) had macrophyte metrics among the most important, but we did not observe an increased importance of macrophytes in our deep lake models. This contrasts with the findings of Beck et al. (2013) who showed that the effect of shoreland development on macrophyte communities was especially pronounced on deeper, more highly developed lakes. However, our metrics generally showed decreased macrophyte abundance at sites with greater riparian disturbance (Figure 5.3). Macrophyte communities across sites and lakes may be differentially affected by a variety of land use practices (Chhor et al. 2020) and water chemistry, especially differences in alkalinity (Mikulyuk et al. 2017). Numerous other studies also found that emergent macrophytes were especially affected by land use along lakeshores (Alahuhta et al. 2014, Dustin and Vondracek 2017, Chhor et al. 2020). In conventional biological assessments, which are commonly focused on detecting species-level changes in community composition, it would be more informative to use species-oriented metrics such as community richness or tolerance values to better elucidate the effects of shoreland development on macrophyte communities. For basic habitat assessments such as this, where macrophytes are being considered largely for their contribution to littoral habitat structure and function for multiple trophic levels, evaluation of functional groups (i.e., emergent, floating, submergent) should be adequate. Use of the shoreland disturbance metrics described here, in conjunction with species-oriented evaluations of littoral macrophytes, may help to advance understanding of the linkages between shoreland development and the biological condition of littoral macrophyte communities.

We used 95% bootstrapped confidence intervals (CIs) based on single-site evaluations to accommodate wide varieties of types and condition of habitat in lakes. The LDA models effectively compare the habitat observed in a lake to an expected natural condition, and the CIs assess if there is enough natural habitat around the lake that is accessible to lake fauna within a specified confidence interval. Lakes may have variable levels of shoreland development along different sections of the shore. High within-lake variability occurred in some of our lakes, which influenced the condition assessments. For example, some Maine Deep lakes (e.g., Woodbury, Jamie's, and Chamberlain) all have negative mean LDA scores suggesting "impaired" conditions, but the range of the confidence interval places them in the intermediate category (Figure 5.4B). As with any assessment decision, best professional judgment should always be exercised to ensure that each lake is properly evaluated through an informed decision-making process. This also applies to placing lakes in the appropriate groups for model calculation. Analyses for lakes that are borderline between multiple groups based on depth or size may benefit from evaluations with applicable models for properly informed condition assessments.

The goal of many lake assessments is to help prioritize actions for the rehabilitation of impaired lake condition. In the event of littoral habitat impairment, the rehabilitation is likely more straightforward than other impairments such as nutrient enrichment due to internal P loading or non-point source pollution from the watershed. The assessments described here focus on shoreland development as the primary stressor that is degrading littoral habitat. Therefore, stressor identification, which can be costly and time-consuming for other impairments, is likely unnecessary. Lake resource managers may address most littoral habitat impairments with shoreland vegetative buffer restoration activities. Maine has a successful program in place that addresses proper lakeshore property management (LakeSmart; Cole et al. 2018). LakeSmart provides a model for other states to create similar programs, as has been done in Vermont (LakeWise; dec.vermont.gov).

Lakes designated as *impaired* with respect to their littoral habitat may serve as a priority for agencies and resource managers to focus lakeshore improvement efforts. Lakes with *natural* littoral habitat may be targeted for conservation priorities. The *intermediate* designation may be beneficial for lake-focused associations to prioritize lakeshore protection efforts locally. The work required to improve lakeshore condition, and subsequently littoral habitat condition, may be considerably less than in lakes with *impaired* habitat status. Moreover, restoring lakeshore condition has other obvious benefits to lake condition, such as prevention of erosion and reduction of nutrient runoff, particularly phosphorus. For organizations with limited resources, striving to make the greatest gains in lake condition, shoreland rehabilitation on *intermediate* lakes may yield a valuable cost-benefit opportunity.

The habitat surveys described here may be readily adopted by community scientist groups with minimal training, following the model of volunteer lake water quality monitors that have contributed greatly to the general knowledge of lake condition for decades (Bigham Stevens et al. 2015, Poisson et al. 2020). These surveys have relatively simple observations, no required taxonomy (as is the case with macrophyte or macroinvertebrate identification), and no laboratory fees. If done properly, the surveys require less than a day of field work and need not be repeated within weeks or months, in contrast to Secchi disk transparency, dissolved oxygen, and determination of lake phosphorus concentration. There is potential for greatly increasing capacity for agency littoral habitat assessments by engaging lake-focused community groups. Furthermore, littoral habitat assessments such as this can yield almost instantaneous results, which would be a considerable benefit for community scientists who can grow restless waiting to learn the results and interpretation from their monitoring efforts.

There are many opportunities to refine the surveys and models described here to enhance the success of littoral habitat condition assessments. We improved assessment confidence by increasing the number of cover classes from five to six, but there may be finer scale separations in cover classes that could provide further benefits. Additionally, NLA methodology surveys only 10 sites per lake regardless of lake size (Kaufmann et al. 2014a, 2014b). We considered each site on a lake independently, and grouped lake information only after sites were evaluated. It is possible that these assessments may benefit from more than 10 sites, possibly based on the surface area or shoreline length of lakes. This may be especially relevant for large lakes that can hold a wide variety of littoral habitat types and shoreland condition. Future work should include power analyses of the number of sites necessary to maximize both model effectiveness and efficiency of field efforts.

Conclusions

The models described here may be used to determine the degree of departure of littoral habitat from an established reference condition, which may be the goal of many agency assessments (Stoddard et al. 2006). This work builds upon numerous studies that have established a linkage between diminished littoral habitat and the conversion of lakeshores from natural vegetation to residential land use. By incorporating findings from those studies and the established methods developed through the NLA, we established a new assessment framework to help improve the condition of lakes with diminished littoral habitat due to anthropogenic shoreland development.

CHAPTER 6. CONCLUSIONS

This dissertation examined the interacting roles of natural and anthropogenic factors that shape lake condition in Maine. The influence of anthropogenic activity on lake condition is well documented, but various aspects of lake water and habitat quality may be further explained by natural characteristics of the watershed or lake. Other aspects of lake condition are influenced by large-scale factors that are not lake or watershed specific, such as weather patterns. Littoral habitat is adversely affected by local-scale residential development on lakeshores. The overall goal of this dissertation was to advance our collective understanding of lake condition and how lakes may be variably affected by natural and anthropogenic factors, thereby allowing for better-informed lake assessments and development of more comprehensive, achievable lake management goals.

Chapter Summaries

In Chapter 1, the rationale behind lake assessments and the long-term record of lake monitoring in Maine was discussed. Among the benefits of such a long empirical history of lake assessment is the comprehensive dataset on Maine lakes that now exists. This dataset may be used to pursue new research questions, including those focused on the interacting roles of various factors that influence lake condition. In this dissertation, we developed our understanding of which aspects of lake condition are related to natural features (e.g., watershed geology or lake morphology), which aspects may be more attributable to anthropogenic influences (watershed land use), and how these factors may interact to shape the condition of Maine lakes. We focused our analyses on how the condition of lakes is affected by natural features of landscapes and watersheds (Chapter 2), the role of lake morphometry in the development of seasonal hypolimnetic anoxia (Chapter 3), the differentiation between regional and local-scale effects on temporal lake clarity trends (Chapter 4), and how anthropogenic shoreland disturbance influences littoral habitat (Chapter 5). We addressed these research questions by using data from hundreds of thousands of lake monitoring events from the past several decades as well as developing new assessment methods to describe the natural littoral habitat condition.

In Chapter 2, a lake classification was developed that tested the importance of various hydrologic, geologic, and morphometric characteristics on shaping background lake water quality characteristics, as measured by total epilimnetic phosphorus (P) and specific conductivity. We found that aggregated Level IV ecoregions (Omernik 1987, 1995 and 2004; Omernik and Griffith 2014) combined with categories based on maximum lake depth captured relevant patterns in water quality conditions observed in Maine lakes. Using land cover data, the expected ranges of water quality values in minimally disturbed, moderately disturbed, and highly disturbed watersheds were defined for reference, intermediate, and altered watershed conditions, respectively. These expected ranges of total P and specific conductivity will be helpful in future lake assessments because they offer more context for interpretation of lake monitoring data and trends. The classification also offers a framework for future lake studies and may be especially useful for those investigations requiring data from similar reference lakes that represent the minimally disturbed condition.

Needs addressed: This study enhances lake data assessment capabilities by providing more context about important patterns in natural and anthropogenic features of lake watersheds across Maine.

In Chapter 3, a logistic regression model was used to identify the morphological and stratification variables that best predict likelihood of seasonal hypolimnetic anoxia (defined as dissolved oxygen < 2 mg/L consistent to the lake bottom) developing in low nutrient lakes (TP < 15 µg/L). Maximum lake depth and the proportion of lake area beneath the epilimnion were used in the primary predictive model, which correctly predicted anoxic condition in 84% of lakes. The addition of epilimnetic TP as a third model variable increased model performance, but only slightly. These models filled a need to identify which lakes may be experiencing seasonal anoxia for reasons other than excess nutrient concentrations. Evaluations of lake condition now may include an assessment of whether hypolimnetic anoxia is occurring because of natural lake features or if it is likely being induced or exacerbated by nutrient enrichment. This allows for more comprehensive appraisals of individual lake condition, which can support better informed lake management decisions.

Needs addressed: These results will support well-informed assessments of lake condition, which will help to prioritize management and remediation efforts by determining if hypolimnetic anoxia is a result of natural lake characteristics or human activities.

In Chapter 4, natural variation in lake water clarity over time was defined using smoothed Generalized Additive Model curves of Secchi Disk Transparency (SDT) data based on reference lakes defined in Chapter 2. As these reference lakes represent lakes in the most undisturbed watersheds within each lake type, any variation in water clarity over time is likely due to regional factors, such as precipitation (patterns and amounts) and seasonal air temperature. The reference curves were then used to compare SDT data from non-reference lakes by calculating residual values between non-reference lake SDT data and the smoothed SDT curves from reference lakes of their respective lake type. This “divergent trend analysis” offers an assessment tool that allows for a comparison to a shifting baseline reference condition over time, so that lake condition may be properly evaluated against a reference condition through time. This addressed an uncertainty that often arose in assessments of lake clarity trends, as temporal clarity patterns were often hypothesized to be related to weather patterns. A Dynamic Factor Analysis (DFA) revealed that cumulative precipitation during the stratification period (April – August) accounted for the most yearly variability in the reference lake clarity data among weather variables tested. Our results suggest that baseline lake clarity will decrease in the northeast US concurrent with predicted increases in precipitation with climate change (Fernandez et al. 2020).

Needs addressed: These findings will support better-informed interpretations of clarity trend data for individual lakes. This study provides a framework for similar interpretations of other water quality parameters where data exist.

Chapter 5 described a new assessment method that was developed to evaluate the effect of anthropogenic shoreland development on littoral habitat condition in lakes. This was done by creating multi-metric indices of littoral habitat condition that focus on residential development as the stressor. Linear Discriminant Analysis (LDA) models were built to find the combination of littoral habitat metrics that best predicted the shoreland condition at each site. Bootstrapped confidence intervals based on LDA scores from each site on a lake were used to make assessment designations of *natural*, *intermediate,* or *impaired* to characterize the littoral habitat condition for individual lakes. This research fills an assessment need in the northeastern U.S., where shoreland development adversely impacted littoral habitat (USEPA 2009, 2016; Kauffman et al. 2014a), but no established methodology to evaluate this linkage on individual lakes has been widely used. This gap in lake assessment tools is of increased importance in Maine where lake habitat condition is included in the water quality statute for lakes, stating that lake "*habitat must be characterized as natural*" (MRS 38, §465-A). Despite this regulatory requirement, no numerical definition of "natural" lake habitat previously existed for Maine lakes.

Needs addressed: The ability to quantify specific effects of shoreland development on lakes with systematic evaluations of littoral habitat will allow for the prioritization of shoreland condition rehabilitation activities on lakes that need it most.

Suggestions for Future Research

This dissertation provides a framework for future research opportunities and raises numerous research questions regarding the assessment of Maine's lakes, and may be applicable to lakes in other regions where necessary data exist. The lake classification (Chapter 2) provides a framework for future studies by partitioning lakes into *a priori* groupings, further stratified by degree of human alteration, that may respond in similar ways to their environment and alterations therein. For example, the classification was used in Chapter 4 to establish reference baselines of SDT through time. The reference lakes of each lake type showed variable trends in lake clarity during the same period, reflecting the variation in background lake clarity across Maine even in the most undisturbed lakes. The classification framework and strategy may also be useful in future studies examining the biological condition of lakes, especially with respect to the structure and function of planktonic communities.

The likelihood of lakes to exhibit natural anoxia (Chapter 3) will be an important assessment component in determining the vulnerability of lakes in Maine to experiencing deteriorated conditions. Future work with this predictive model should incorporate sediment chemistry information. Lakes that have extractable Al:Fe and Al:P ratios that favor sediment P release under anoxic conditions (Kopáček et al. 2005) *and* are expected to exhibit seasonal anoxia may reveal an increased propensity to release sediment-bound P as well as other redox sensitive elements. This additional information in lake condition assessments will inform lake management decisions involving P loading allowances and restoration targets. These results may help in determining management action strategies for high nutrient lakes as well. It may now be possible to identify lakes where developing eutrophication is exacerbated by naturally-occurring anoxia, or lakes that will still experience anoxia even if excesses in nutrients and organic matter in water and sediments are reduced. These models can also support the development of long-term restoration plans where the feasibility of artificial hypolimnetic aeration or chemical treatments are being considered for the sequestration of sediment P.

The divergent trend analysis (Chapter 4) describes a method that may be used to examine any number of environmental parameters over time against a reference baseline. Although analyses may be limited by data availability, this type of trend analysis may be applied to other measures of lake condition such as total P, chlorophyll-*a*, or dissolved organic carbon. For parameters that may not have robust enough datasets for a divergent trend analysis for each lake type, data from various types may be combined into aggregations of reference lake types that contain adequate sample size for establishing reference baselines (e.g., all reference coastal lakes, all reference deep lakes, all reference lakes in Maine, etc.). We tested several weather variables in our DFA model to explain variation in the reference lake water clarity trends attributable to regional weather patterns. This type of analysis could be expanded to answer different study questions by including various other datasets, if available, such as temporal changes in human population density or watershed land use.

The littoral habitat assessments (Chapter 5) represent a new method of lake condition assessment and provide a measure for the stressor of shoreland development. This method may be used to place lakes into assessment categories (*impaired*, *natural*, etc.) based on habitat condition, as specified in the water quality statue for Maine lakes (MRS 38, §465-A). Lakes that are listed as *impaired* for habitat in Maine's Integrated Report to the US EPA (305(b) report, ME DEP 2016) will be officially prioritized for rehabilitation of shoreland condition. This assessment framework creates opportunities for the establishment of biological criteria in Maine lakes. Future research on biological communities of macrophytes or macroinvertebrates in lake littoral zones could use a similar approach as described here (e.g., using the same 10 sites along shorelines), and the metrics established here could be used to define the gradient of disturbance and habitat condition necessary for various taxonomic groups. Establishing biological criteria for lakes would connect our understanding of how shoreland development affects the *structure* of littoral habitat to how the *function* of that habitat is affected. The Maine lakes statute (MRS 38, §465-A) indicates that habitat should be characterized as natural; the intent was likely that lake biological communities should exist in a natural state, as supported by habitat in a natural condition. More comprehensive assessments of lake biological communities, especially in the littoral zone, would offer a direct measure of whether the habitat was truly supporting natural biological communities. Large-magnitude lake water level drawdowns (> 1 m), which occur on <100 Maine lakes, were not considered in these models but have an important effect on lake littoral habitat (Carmignani and Roy

2021). Future work on littoral habitat condition could focus on the creation of a littoral assessment method specifically for lakes with large water level drawdowns in Maine. There are further opportunities to refine the assessment method with different numbers of cover classes or by establishing the necessary number of evaluation sites for lakes of different sizes or shoreline lengths. Now that the method is established, general condition assessments of lake littoral habitat in Maine may be pursued; a study has already begun to investigate the littoral habitat condition within each of the lake classification types detailed in Chapter 2 by surveying lakes chosen with a randomized probability-based lake selection method.

Lake assessment methods that focus on parameters not directly addressed in this dissertation may still benefit from the application of the analyses described here. Cyanobacteria blooms in Maine lakes have become more prevalent in recent years. Even some lakes that have been historically oligo- or mesotrophic have experienced sudden cyanobacteria blooms that pose a threat to public health through the presence of cyanotoxins (Cottingham et al., 2015; Favot et al., 2019). Much research regarding cyanobacteria is currently focused on environmental indicators that may help predict the occurrences of blooms, and especially the release of cyanotoxins (Descy et la. 2016, Bukowska et al. 2017, Zhao et al. 2019). As discussed in Chapter 2, the geochemical composition of lake watersheds contributes to the overall trophic condition of lakes, and therefore may influence the likelihood of a lake to exhibit cyanobacteria blooms. This may be particularly relevant with respect to mobilization of Al and P from watershed soils, and the adsorbing and desorbing of P based on changes in pH through interactions with groundwater and surface waters (Reinhardt et al. 2004, McDonald et al. 2019). Additionally, there may be predictive potential related to the Al:Fe ratios in lake sediment and the likelihood for low-nutrient lakes to experience seasonal hypolimnetic anoxia, as discussed in Chapter 3. Further biological assessments of cyanobacteria species present in various lake types, perhaps through targeted environmental DNA (eDNA) sampling and analysis, may help to determine which lakes may be especially at risk of sudden cyanobacteria blooms.

The most collected samples for lake monitoring (SDT, TP, and Chlorophyll-*a*) are taken because they evaluate lake trophic condition. However, SDT and TP are surrogates for trophic state. Chlorophyll-*a* is commonly considered a direct measure of trophic condition, but there are often confounding factors affecting its relationship with lake TP (Yuan and Jones 2020). These measurements should continue because of the rich datasets that are based on them, and because the simplicity of measurement (especially SDT) keep community scientists engaged by contributing meaningful data to lake assessments. However, with advancing technology, many other measures may be added to lake monitoring which offer more direct measures of the biological condition of lakes. For example, eDNA samples can provide an efficient inventory of species that are likely present in a system, and this represents an opportunity for rapid characterization of planktonic communities in lakes (Ruppert et al. 2019, Yang et al. 2020, He et al. 2021). While taxonomic work is required to maintain adequate libraries of DNA information to validate eDNA sample results, the requirement of having a taxonomic specialist optically identify organisms from each sample is unnecessary. A potential drawback of planktonic community characterization, despite the wealth of information generated from the data, is that multiple samples are needed over the course of the open-water season to adequately characterize the composition of these communities. Further, diel variation in organism location in the water column also may confound results of lake plankton community assessments. Sampling regimes that account for diel and seasonal variability in plankton communities may not be possible in all situations, even with highly efficient eDNA sampling techniques. For holistic planktonic community assessments completed in one sampling event, sediment diatoms may offer the most effective and efficient sampling method. These samples offer an evaluation of the community of lake diatoms that have sunken to the bottom of the lake, with their siliceous exoskeletons still intact. Through species identification based on the exoskeletons, the community of living diatoms may be reconstructed from recent years (surficial sediment) or historic periods (sediment below surface), to evaluate trophic status condition and changes in the biological communities over time (Vermaire and Gregory-Eaves 2007, Beck et al. 2016). There is potential for eDNA methods to be applied to sedimented plankton communities as well, streamlining efficiency for this powerful assessment method (Ibrahim et al. 2021).

It will be informative to examine multiple components of lake characteristics in a singular analysis – both established methods (e.g., sediment chemistry, flushing rate, average depth, watershed soil types) and those developed here (i.e., lake type, likelihood of anoxia, divergent clarity trend results, and shoreland condition) to determine which factors may be of greatest importance in determining lake condition. Lake condition is likely dependent upon multiple related factors, and the importance of these interacting factors may vary among the lake types identified in Chapter 2. A lake "vulnerability index" has been used in Maine since the 1970s, largely based on Vollenweider (1975), which helps determine which lakes are more susceptible to increases in P concentration from increased P watershed inputs. With the advancement of lake assessment methods discussed above, there is an opportunity to refine the vulnerability index to include a wider variety of information and learn more about the relative risks involved with watershed disturbance for different types of lakes. This refinement could help further focus assessment needs, and restoration action priorities and objectives for lakes across Maine.

The assessment methods and models described here may have applicability in other regions. Many details focused on parameters important to the condition of Maine lakes, but most would be readily exchanged for factors relevant to lakes in other geographical areas of study. For example, if the lake classification framework described here (Chapter 2) was to be applied elsewhere, it is likely that the collection of influencing parameters would change. Large watershed geographical groupings may be more important than ecoregions in areas where riverine reservoirs are more common than natural lakes, and lake size may be more important than lake depth. Different traits of lakes and their landscapes may be more informative in other regions as well, such as residence time, elevation, and lake watershed size. Factors contributing to the extent of anoxia (Chapter 3) may change in areas where lakes do not experience seasonal mixing or winter ice cover. Further, the development of natural anoxia may be irrelevant in areas where low-nutrient lakes are rare. However, the logistic regression models used here may be useful for predicting other lake responses such as vulnerability to invasive species infestations, mixing frequency, or occurrence of harmful algae blooms. The comparison of temporal water quality patterns to a reference baseline through time (Chapter 4) has broad applicability for a number of parameters, although the application may be limited in some areas without suitable reference lakes. In the absence of reference lakes, the response variable under consideration (e.g., SDT) may be compared to the trends observed in the entire population of lakes, thereby offering a baseline comparison to similar waterbodies. The littoral habitat condition assessments (Chapter 5) have wide applicability outside of the region considered here; lakes are affected by human watershed alteration worldwide (Mammides 2020). The development of additional metrics should be pursued, especially in other regions, which may provide models that better address habitat condition and variability in those areas.

Despite technological advancements in lake assessment methods, support for conventional methods should not be reduced. For instance, the Secchi disk is a powerful tool not only for gathering lake assessment data, but also for providing beginning volunteer lake monitors with an entry-level task that connects them more deeply to a lake that they care about. The capacity for lake assessment in Maine is undeniably powered by the engagement of community lake monitors, a group largely comprised of volunteers. The Maine DEP may be responsible for final assessments of lake condition, but without the rich dataset of information gathered by these community scientists those assessments would be impossible. Therefore, efforts should continue that keep the Maine lakes community engaged in lake monitoring and lake-related issues. This includes prioritizing the funding of community-based lake monitoring programs which often are forced to operate with insufficient funds. Fostering the enthusiasm of lake advocates across the state will help to preserve and protect the lakes of Maine for generations to come.

REFERENCES

Abell JM, Özkundakci D, Hamilton DP, Jones JR. 2012. Latitudinal variation in nutrient stoichiometry and chlorophyll-nutrient relationships in lakes: a global study. Fundamental and Applied Limnology/Archiv für Hydrobiologie, 181:1-14.

Ahtiainen, M. 1992. The effects of forest clear-cutting and scarification on the water quality of small brooks. Hydrobiologia. 243/244: 465-473.

Alahuhta J, Kanninen A, Hellsten S, Vuori KM, Kuoppala M, Hämäläinen H. 2014. Variable response of functional macrophyte groups to lake characteristics, land use, and space: implications for bioassessment. Hydrobiologia. 737:201-214.

Allison PD. 2012. Logistic Regression Using SAS: Theory and Application, Second Edition. SAS Institute, Mar 30, 2012. 348 pp.

Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67:1-48.

Beck KK, Medeiros AS, Finkelstein SA. 2016. Drivers of Change in a 7300-Year Holocene Diatom Record from the Hemi-Boreal Region of Ontario, Canada. PloS One. 11:8.

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF. 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data. 5:180214.

Beck MW, Vondracek B, Hatch LK. 2013. Between-and within-lake responses of macrophyte richness metrics to shoreline development. Lake and Reservoir Management. 29:179-193.

Bigham Stephens DL, Carlson RE, Horsburgh CA, Hoyer MV, Bachmann RW, Canfield DE. 2015. Regional distribution of Secchi disk transparency in waters of the United States. Lake and Reservoir Management. 31:55-63.

Brauns M, Gücker B, Wagner C, Garcia XF, Walz N, Pusch MT. 2011. Human lakeshore development alters the structure and trophic basis of littoral food webs. Journal of Applied Ecology. 48:916-925.

Brezonik, PL, Bouchard RW, Finlay JC, Griffin CG, Olmanson LG, Anderson JP, Arnold WA, Hozalski R. 2019. Color, chlorophyll a, and suspended solids effects on Secchi depth in lakes: implications for trophic state assessment. Ecological Applications. 29: e01871

Brothers S, Köhler J, Attermeyer K, Grossart HP, Mehner T, Meyer N, Scharnweber K, Hilt S. 2014. A feedback loop links brownification and anoxia in a temperate, shallow lake. Limnology and Oceanography 59:1388-1398.

Bukowska A, Kaliński T, Koper M, Kostrzewska-Szlakowska I, Kwiatowski J, Mazur-Marzec H, Jasser I. 2017. Predicting blooms of toxic cyanobacteria in eutrophic lakes with diverse cyanobacterial communities. Scientific reports. 7:1-2.

Burnham KP, Anderson DR. 2002. Model Selection and Multimodel Inference: A Practical Information-theoretic Approach, 2nd ed. Springer-Verlag.

Butler RG, Demaynadier PG. 2008. The significance of littoral and shoreline habitat integrity to the conservation of lacustrine damselflies (Odonata). Journal of Insect Conservation. 12:23—36.

Canfield Jr DE, Bachmann RW, Stephens DB, Hoyer MV, Bacon LC, Williams S, Scott M. 2016. Monitoring by citizen scientists demonstrates water clarity of Maine (USA) lakes is stable, not declining, due to cultural eutrophication. Inland Waters. 6: 11-27.

Canty A, Ripley B. 2021. boot: Bootstrap Functions. R package version 1.3-28. https://cran.r-project.org/web/packages/boot/index.html

Carmignani JR, Roy AH. 2021. Annual winter water-level drawdowns influence physical habitat structure and macrophytes in Massachusetts, USA, lakes. Ecosphere. 12:e03442.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications. 8:559-568.

Carpenter SR, Pace ML. 2018. Synthesis of a 33-yr series of whole-lake experiments: Effects of nutrients, grazers, and precipitation-driven water color on chlorophyll. Limnology and Oceanography Letters. 3: 419-427.

Charlton MN. 1980. Hypolimnion oxygen consumption in lakes: discussion of productivity and morphology effects. Canadian Journal of Fisheries and Aquatic Sciences. 37:1531–1539.

Chen D, Hu M, Wang J, Guo Y, Dahlgren RA. 2016. Factors controlling phosphorus export from agricultural/forest and residential systems to rivers in eastern China, 1980–2011. Journal of Hydrology 533:53-61.

Cheruvelil, KS, Soranno PA, Bremigan MT. 2008. Grouping lakes for water quality assessment and monitoring: the roles of regionalization and spatial scale. Environmental Management. 41:425-440.

Chhor AD, Glassman DM, Smol JP, Vermaire JC, Cooke SJ. 2020. Ecological consequences of shoreline armoring on littoral fish and benthic macroinvertebrate communities in an Eastern Ontario lake. Aquatic Sciences. 82:1-13.

Christensen DL, Herwig BR, Schindler DE, Carpenter SR. 1996. Impacts of lakeshore residential development on coarse woody debris in north temperate lakes. Ecological Applications. 6:1143-1149.

Clark BJ, Dillon PJ, Molot LA, Evans HE. 2002. Application of a Hypolimnetic Oxygen Profile Model to Lakes in Ontario. Lake and Reservoir Management 18:32-43.

Coats R. 2010. Climate change in the Tahoe basin: regional trends, impacts and drivers. Climatic Change 102: 435–466.

Coldsnow KD, Relyea RA. 2018. Toxicity of various road-deicing salts to Asian clams (Corbicula fluminea). Environmental toxicology and chemistry. 37:1839-1845.

Cole FR, Junker A, Bevier CR, Shannon M, Sarkar S, Nyhus PJ. 2018. Assessing LakeSmart, a community-based lake protection program. Journal of Environmental Studies and Sciences. 8:264-280.

Cooke GD, Welch EB, Peterson SA, Nichols SA. 2005. Restoration and management of lakes and reservoirs. Third edition. Taylor & Francis, Boca Raton, Florida, USA.

Cornett RJ, Rigler FH. 1980. The areal hypolimnetic oxygen deficit: An empirical test of the model 1. Limnology and Oceanography. 25:672-679.

Cottingham KL, Ewing HA, Greer ML, Carey CC, Weathers KC. 2015. Cyanobacteria as biological drivers of lake nitrogen and phosphorus cycling. Ecosphere. 6:1.

Couture R, Wit HA, Tominaga K, Kiuru P, Markelov I. 2015. Oxygen dynamics in a boreal lake responds to long-term changes in climate, ice phenology, and DOC inputs. Journal of Geophysical Research: Biogeosciences. 120:2441-2456.

Crowder LB, Cooper WE. 1982. Habitat structural complexity and the interaction between bluegills and their prey. Ecology. 63:1802-1813.

Danielson TJ, Loftin CS, Tsomides L, DiFranco JL, Connors B, Courtemanch DL, Drummond F, Davies SP. 2012. An algal model for predicting attainment of tiered biological criteria of Maine's streams and rivers. Freshwater Science. 31:318-340.

Danielson TJ, Tsomides L, Suitor D, Difranco JL, Connors B. 2016. Effects of urbanization on aquatic life of Maine streams. Augusta (ME): Maine Department of Environmental Protection.

D'Arcy P, Carignan R. 1997. Influence of catchment topography on water chemistry in southeastern Québec shield lakes. Canadian Journal of Fisheries and Aquatic Sciences. 54:2215-2227.

Davis CJ, Ruhmann EK, Acharya K, Chandra S, Jerde CL. 2015. Successful survival, growth, and reproductive potential of quagga mussels in low calcium lake water: is there uncertainty of establishment risk? PeerJ. 3:e1276.

Davis JC. 1975. Minimal dissolved oxygen requirements of aquatic life with emphasis on Canadian species: a review. Journal of the Fisheries Board of Canada. 32:2295-2332.

Deeds J, Amirbahman A, Norton SA, Bacon LC. 2020. A hydrogeomorphic and condition classification for Maine, USA, lakes. Lake and Reservoir Management. 36:22-138.

Deeds J, Amirbahman A, Norton SA, Suitor DG, Bacon LC. 2021a. Predicting anoxia in low-nutrient temperate lakes. Ecological Applications. e02361.

Deeds J, Amirbahman A, Norton SA, Bacon LC, Hovel RA. 2021b. Shifting baselines and cross-scale drivers of lake water clarity: Applications for lake assessment. Limnology and Oceanography. 9999:1-14.

Dennis J. 1986. Phosphorus export from a low-density residential watershed and an adjacent forested watershed. Lake and Reservoir Management. 2:401-407.

Descy JP, Leprieur F, Pirlot S, Leporcq B, Van Wichelen J, Peretyatko A, Teissier S, Codd GA, Triest L, Vyverman W, Wilmotte A. 2016. Identifying the factors determining blooms of cyanobacteria in a set of shallow lakes. Ecological Informatics. 34:129-38.

DeVito KJ, Creed IF, Rothwell RL, Prepas EE. 2000. Landscape controls on phosphorus loading to boreal lakes: Implications for the potential impacts of forest harvesting. Canadian Journal of Fisheries and Aquatic Sciences. 57:1977-1984.

Diffenbaugh NS, Pal JS, Trapp RJ, Giorgi F. 2005. Fine-scale processes regulate the response of extreme events to global climate change. Proceedings of the National Academy of Sciences. 102:15774-15778.

Dugan HA, Bartlett SL, Burke SM, Doubek JP, Krivak-Tetley FE, Skaff NK, Summers JC, Farrell KJ, McCullough IM, Morales-Williams AM, Roberts DC. 2017 Salting our freshwater lakes. Proceedings of the National Academy of Sciences. 114:4453-4458.

Dupuis D, Sprague E, Docherty KM, Koretsky CM. 2019. The influence of road salt on seasonal mixing, redox stratification and methane concentrations in urban kettle lakes. Science of The Total Environment. 661:514-521.

Dustin DL, Vondracek B. 2017. Nearshore habitat and fish assemblages along a gradient of shoreline development. North American Journal of Fisheries Management. 37:432-444.

Ellis AW, Greene TR. 2019. Synoptic climate evidence of a late-twentieth century change to earlier spring ice-out on Maine Lakes, USA. Climatic Change. 153:323-339.

Environmental Systems Research Institute (ESRI). 2017. ArcGIS Release 10.6.1. Redlands, CA.

Evtimova VV, Donohue I. 2016. Water-level fluctuations regulate the structure and functioning of natural lakes. Freshwater biology. 61:251-264.

Everett RA, Ruiz GM. 1993. Coarse woody debris as a refuge from predation in aquatic communities. Oecologia. 93:475-486.

Favot EL, Rühland KM, DeSellas AM, Ingram R, Paterson AM, Smol JP. 2019. Climate variability promotes unprecedented cyanobacterial blooms in a remote, oligotrophic Ontario lake: evidence from paleolimnology. Journal of Paleolimnology. 62:31–52.

Fee EJ, Hecky RE, Kasian SEM, Cruikshank DR. 1996. Physical and chemical responses of lakes and streams. Limnology and Oceanography. 41: 912-920.

Fernandez I, Birkel S, Schmitt C, Simonson J, Lyon B, Pershing A, Stancioff E, Jacobson G, Mayewski P. 2020. Maine's Climate Future 2020 Update. Orono, ME: University of Maine.

Fernandez IJ, Schmitt CV, Birkel SD, Stancioff E, Pershing AJ, Kelley JT, Runge JA, Jacobson GL, Mayewski PA. 2015. Maine's Climate Future: 2015 Update. Orono, ME: University of Maine.

Fisher R, Wilson SK, Sin TM, Lee AC, Langlois TJ. 2018. A simple function for full-subsets multiple regression in ecology with R. Ecology and Evolution. 8:6104-6113.

Fox J, Monette G. 1992. Generalized collinearity diagnostics. Journal of the American Statistical Association. 87:178–183.

Francis TB, Schindler DE. 2006. Degradation of littoral habitats by residential development: woody debris in lakes of the Pacific Northwest and Midwest, United States. AMBIO: A Journal of the Human Environment. 35:274-280.

Fraterrigo JM, Downing JA. 2008. The influence of land use on lake nutrients varies with watershed transport capacity. Ecosystems. 11:1021-1034.

French RH, Cooper JJ, Vigg S. 1982. Secchi disk relationships. Water Resources Bulletin. 18:121-123.

Garrison PJ, LaLiberte GD, Ewart BP. 2010. The importance of water level changes and shoreland development in the eutrophication of a shallow, seepage lake. Proceedings of the Academy of Natural Sciences of Philadelphia. 160:113-126

Garrison, PJ, Wakeman RS. 2000. Use of paleolimnology to document the effect of lake shoreland development on water quality. Journal of Paleolimnology. 24:369-393.

Greco, DA, Arnott SE, Fournier IB, Schamp BS. 2021. Effects of chloride and nutrients on freshwater plankton communities. Limnology and Oceanography Letters. 1-8.

Griffith, GE, Omernik JM, Bryce SA, Royte J, Hoar WD, Homer JW, Keirstead D, Metzler KJ, Hellyer G. 2009. Ecoregions of New England (color poster with map, descriptive text, summary tables, and photographs): Reston, Virginia, U.S. Geological Survey (map scale 1:1,325,000).

Griffiths ML, Bradley RS. 2007. Variations of twentieth-century temperature and precipitation extreme indicators in the northeast United States. Journal of Climate. 20:5401-5417.

Gunn JM, Snucins ED, Yan ND, Arts MT. 2001 Use of water clarity to monitor the effects of climate change and other stressors on oligotrophic lakes. Environmental Monitoring and Assessment. 67:69-88.

Hatzenbeler GR, Kampa JM, Jennings MJ, Emmons EE. 2004. A comparison of fish and aquatic plant assemblages to assess ecological health of small Wisconsin lakes. Lake and Reservoir Management. 20:211-218.

Hayse JW, Wissing TE. 1996. Effects of stem density of artificial vegetation on abundance and growth of age-0 bluegills and predation by largemouth bass. Transactions of the American Fisheries Society. 125:422-433.

He L, Zhang L, Yang J, Zhao Z, Zhou X, Feng Q, Zhang X, Xu S. 2021. Monitoring of plankton diversity in Dianchi Lake by environmental DNA technology. InIOP Conference Series: Earth and Environmental Science 651:042023

Hein CL, Öhlund G, Englund G. 2012. Future distribution of Arctic char Salvelinus alpinus in Sweden under climate change: effects of temperature, lake size and species interactions. Ambio. 41:303-12.

Helmus MR, Sass GG. 2008. The rapid effects of a whole-lake reduction of coarse woody debris on fish and benthic macroinvertebrates. Freshwater Biology. 53:1423-1433.

Helsel DR, Hirsch RM, Ryberg KR, Archfield SA, Gilroy EJ. 2020. Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chapter A3, 458 p.

Herlihy AT, Sobota JB, McDonnell TC, Sullivan TJ, Lehmann S, Tarquinio E. 2013. An *a priori* process for selecting candidate reference lakes for a national survey. Freshwater Science 32:385-396

Hijmans RJ. 2019. raster: Geographic Data Analysis and Modeling. R package version 2.8-19. https://CRAN.R-project.org/package=raster

Hintz WD, Fay L, Relyea RA. 2022. Road salts, human safety, and the rising salinity of our fresh waters. Frontiers in ecology and the environment. 20:22-30.

Hodgkins GA, James IC, Huntington TG. 2002. Historical changes in lake ice-out dates as indicators of climate change in New England, 1850–2000. International Journal of Climatology: A Journal of the Royal Meteorological Society. 22:1819-1827.

Hollister J, Stachelek J. lakemorpho: Calculating lake morphometry metrics in R [version 1; peer review: 2 approved]. F1000Research 2017, 6:1718

Holmes E, Ward E, Scheuerell M, Wills K. 2020. *MARSS: Multivariate Autoregressive State-Space Modeling*. R package version 3.10.12, https://CRAN.R-project.org/package=MARSS.

Horne AJ. 2020 Hypolimnetic oxygenation 4. Effects on turbidity in Camanche Reservoir and its downstream fish hatchery. Lake and Reservoir Management. 36:360-375.

Hutchinson GE. 1957. A Treatise on Limnology. 1. John Wiley and Sons, Toronto.

Ibrahim A, Capo E, Wessels M, Martin I, Meyer A, Schleheck D, Epp LS. 2021. Anthropogenic impact on the historical phytoplankton community of Lake Constance reconstructed by multimarker analysis of sediment-core environmental DNA. Molecular Ecology. 30:3040-3056.

Jennings MJ, Bozek MA, Hatzenbeler GR, Emmons EE, Staggs MD. 1999. Cumulative effects of incremental shoreline habitat modification on fish assemblages in north temperate lakes. North American Journal of Fisheries Management. 19:18-27.

Jones JR, Knowlton MF, Obrecht DV, Cook EA. 2004. Importance of landscape variables and morphology on nutrients in Missouri reservoirs. Canadian Journal of Fisheries and Aquatic Sciences. 61:1503-1512.

Jones KB, Neale AC, Nash MS, Van Remortel RD, Wickham JD, Riitters KH, O'Neill RV. 2001. Predicting nutrient and sediment loadings to streams from landscape metrics: a multiple watershed study from the United States Mid-Atlantic Region. Landscape Ecology. 16:301–312.

Kaufmann PR, Hughes RM, Van Sickle J, Whittier TR, Seeliger CW, Paulsen SG. 2014b. Lakeshore and littoral physical habitat structure: A field survey method and its precision. Lake and Reservoir Management. 30:157-176.

Kaufmann PR, Peck DV, Paulsen SG, Seeliger CW, Hughes RM, Whittier TR, Kamman NC. 2014a. Lakeshore and littoral physical habitat structure in a national lakes assessment. Lake and Reservoir Management. 30:192-215.

Kaushal SS, Groffman PM, Likens GE, Belt KT, Stack WP, Kelly VR, Band LE, Fisher GT. 2005. Increased salinization of fresh water in the northeastern United States. Proceedings of the National Academy of Sciences. 102:13517-13520.

Knoll LB, Williamson CE, Pilla RM, Leach TH, Brentrup JA, Fisher TJ. 2018. Browning-related oxygen depletion in an oligotrophic lake. Inland Waters 8:255-263.

Kopáček J, Borovec J, Hejzlar J, Ulrich KU, Norton SA, Amirbahman A. 2005. Aluminum control of phosphorus sorption by lake sediments. Environmental Science & Technology 39:8784-8789.

Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. 2006. World map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift, 15:259-263.

Kraemer BM, Anneville O, Chandra S, Dix M, Kuusisto E, Livingstone DM, Rimmer A, Schladow SG, Silow E, Sitoki LM, Tamatamah R. 2015. Morphometry and average temperature affect lake stratification responses to climate change. Geophysical Research Letters. 42:4981-4988.

Lake BA, Coolidge KM, Norton SA, Amirbahman A. 2007. Factors contributing to the internal loading of phosphorus from anoxic sediments in six Maine, USA, lakes. Science of the Total Environment 373:534-541.

Lawson ZJ, Gaeta JW, Carpenter SR. 2011. Coarse woody habitat, lakeshore residential development, and largemouth bass nesting behavior. North American Journal of Fisheries Management. 31:666-670.

Lepore J. 2013. Local and cumulative influences of docks on littoral habitat structure. Master's thesis. University of Minnesota, St. Paul.

Lewin WC, Mehner T, Ritterbusch D, Brämick U. 2014. The influence of anthropogenic shoreline changes on the littoral abundance of fish species in German lowland lakes varying in depth as determined by boosted regression trees. Hydrobiologia. 724:293—306.

Livingstone DM, Imboden DM. 1996. The prediction of hypolimnetic oxygen profiles: a plea for a deductive approach. Canadian Journal of Fisheries and Aquatic Sciences. 53:924-932.

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. (2019). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.0.

Maine Department of Environmental Protection. 2004. Maine Lake Assessment Quality Assurance Program Plan. Augusta (ME).

Maine Department of Environmental Protection. 2015. Maine Lake Assessment Quality Assurance Program Plan. Augusta (ME).

Maine Department of Environmental Protection. 2018. 2016 Integrated Water Quality Monitoring and Assessment Report. Augusta (ME).

Maine Geolibrary, 2005. IMPERV. Publication date: 18 Oct 2005.

Maine Geolibrary, 2006. Maine Land Cover Dataset. Publication date: 23 May 2006.

Maine Geolibrary, 2011a. Lake Depth. Publication date: 13 October 2011.

Maine Geolibrary, 2011b. Census Blocks 2010. Publication date: 6 April 2011.

Maine Office of GIS. 2011. LAKEDEPTH. Augusta, Maine.

Mallakpour I, Villarini G. 2017. Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA. Theoretical and Applied Climatology. 130:345-363.

Mammides C. 2020. A global assessment of the human pressure on the world's lakes. Global Environmental Change. 63:102084.

Mangiafico SS. 2016. Summary and Analysis of Extension Program Evaluation in R, version 1.18.1. rcompanion.org/handbook/.

Marburg AE, Turner MG, Kratz TK. 2006. Natural and anthropogenic variation in coarse wood among and within lakes. Journal of Ecology. 94:558—568.

Martin SL, Soranno PA, Bremigan MT. 2011. Comparing hydrogeomorphic approaches to lake classification. Environmental Management. 48: 957-974.

Mazerolle MJ. 2017. AICcmodavg: model selection and multimodel inference based on (Q)AIC(c). R package version 2.2-2.

McCullough IM, Cheruvelil KS, Collins SM, Soranno PA. 2019. Geographic patterns of the climate sensitivity of lakes. Ecological Applications. 29:e01836.

McCullough IM, Loftin CS, Sader SA. 2013. Landsat imagery reveals declining clarity of Maine's lakes during 1995–2010. Freshwater Science. 32: 741-752.

McDonald GJ, Norton SA, Fernandez IJ, Hoppe KM, Dennis J, Amirbahman A. 2019. Chemical controls on dissolved phosphorus mobilization in a calcareous agricultural stream during base flow. Science of the Total Environment. 660:876-885.

McLeod AI. 2015. Kendall: Kendall rank correlation and Mann-Kendall trend test. R package version 2.2. https://CRAN.R-project.org/package=Kendall

McMahon JS. 1990. The biophysical regions of Maine: patterns in the landscape and vegetation. [master's thesis]. [Orono (ME)]: University of Maine.

Merrell K, Howe EA, Warren S. 2009. Examining shorelines, littorally. Lakeline. 13:8-13.

Merugu C, Seetharaman R. (2013), Comparative analysis of land use and lake water quality in rural and urban zones of south Chennai, India. Environment, Development and Sustainability. 15:511-528.

Michael HJ, Boyle KJ, Bouchard R. 1996. Water quality affects property prices: a case study of selected Maine lakes. Maine Agricultural and Forest Experiment Station Miscellaneous Report 398. University of Maine. Orono, ME.

Mikulyuk A, Barton M, Hauxwell J, Hein C, Kujawa E, Minahan K, Nault ME, Oele DL, Wagner KI. 2017. A macrophyte bioassessment approach linking taxon-specific tolerance and abundance in north temperate lakes. Journal of Environmental Management. 199:172-180.

Mikulyuk A, Sharma S, Van Egeren S, Erdmann E, Nault ME, Hauxwell J. 2011. The relative role of environmental, spatial, and land-use patterns in explaining aquatic macrophyte community composition. Canadian Journal of Fisheries and Aquatic Sciences. 68:1778-1789.

Miler O, Porst G, Mcgoff E, Pilotto F, Donohue L, Jurca T, Solimini A, Sandin L, Irvine K, Aroviita J, Clarke, R. 2015. An index of human alteration of lake shore morphology. Aquatic Conservation: Marine and Freshwater Ecosystems. 25:353-364.

Mjelde M, Hellsten S, Ecke F. 2013. A water level drawdown index for aquatic macrophytes in Nordic lakes. Hydrobiologia. 704:141-151.

Molot LA, Dillon PJ, Clark BJ, Neary BP. 1992. Predicting end-of-summer oxygen profiles in stratified lakes. Canadian Journal of Fisheries and Aquatic Sciences, 49:2363-2372.

Müller B, Bryant LD, Matzinger A, Wüest A. 2012. Hypolimnetic oxygen depletion in eutrophic lakes. Environmental Science and Technology. 46:9964-9971.

Muscutt, DA, Withers P. 1996. The phosphorus content of rivers in England and Wales. Water Research. 30:1258-1268.

Nandana P, Gharabaghi B, Howard K. 2013. Groundwater chloride response in the Highland Creek watershed due to road salt application: A re-assessment after 20 years. Journal of Hydrology. 479:159-168.

Neale MW and Rippey B. 2008. A comparison of environmental and biological site classifications for the prediction of macroinvertebrate communities of lakes in Northern Ireland. Aquatic Conservation: Marine and Freshwater Ecosystems. 18:729-741.

Newman DJ, Perault DR, Shahady, TD. 2006. Watershed development and sediment accumulation in a small urban lake. Lake and Reservoir Management. 22:303-307.

Nielsen A, Trolle D, Søndergaard M, Lauridsen TL, Bjerring R, Olesen JE, Jeppesen E. 2012. Watershed land use effects on lake water quality in Denmark. Ecological Applications. 22:1187-1200.

Nieratko DP. 1992. Factors controlling phosphorus loading to lakes in Maine: a statistical analysis. [Master of Science thesis]. [Orono (ME)]: University of Maine.

Norton SA, Amirbahman A, Bacon L, Ewing HA, Novak M, Nurse A, Retelle M, Stager JC, Yates M. 2019. Chemical, biological, and trophic status of temperate lakes can be strongly influenced by the presence of late-glacial sediments. Lake and Reservoir Management. DOI: 10.1080/10402381.2019.1669747

Norton, SA, Brakke DF, Kahl JS, and Haines TA. 1989. Major influences on lake water chemistry in Maine. In: Tucker R, Marvinney R, editors. Studies in Maine Geology. Augusta (ME): Maine Geological Survey. 109-124.

Nürnberg GK, Shaw M. 1999. Productivity of clear and humic lakes: nutrients, phytoplankton, bacteria. Hydrobiologia. 382:97-112.

Nürnberg GK. 1995. Quantifying anoxia in lakes. Limnology and Oceanography. 40:1100-1111.

Nürnberg GK. 1997. Coping with water quality problems due to hypolimnetic anoxia in Central Ontario Lakes. Water Quality Research Journal of Canada. 32:391-405.

Nürnberg GK. 2004. Quantified hypoxia and anoxia in lakes and reservoirs. The Scientific World Journal 4:42-54.

Nürnberg GK. 2019. Quantification of anoxia and hypoxia in water bodies (2nd edition). *In* Encyclopedia of Water: Science, Technology, and Society, Edited by Patricia A. Maurice. John Wiley & Sons, Inc.

Olivero-Sheldon A, Anderson MG. 2016. Northeast Lake and Pond Classification. The Nature Conservancy, Eastern Conservation Science, Eastern Regional Office. Boston, MA.

Omernik JM, Griffith GE. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management 54:1249-1266.

Omernik JM. 1977. Nonpoint source stream nutrient level relationships: A nationwide study. EPA/600/3-77-105. U.S. EPA. Environmental Research Laboratory, Corvallis, OR. 151

Omernik JM. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Annals of the American Association of Geographers. 77:118-125.

Omernik JM. 1995. Ecoregions: A spatial framework for environmental management. In: Davis, WS and Simon TP, editors. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Boca Raton (FL): Lewis Publishers, p. 49-62.

Omernik JM. 2004. Perspectives on the nature and definition of ecological regions. Environmental Management. 34:S27-S38.

O'Reilly CM, Sharma S, Gray DK, Hampton SE, Read JS, Rowley RJ, Schneider P, Lenters JD, McIntyre PB, Kraemer BM, Weyhenmeyer GA. 2015. Rapid and highly variable warming of lake surface waters around the globe. Geophysical Research Letters. 42:10-773.

Ostendorp W, Hofmann H, Teufel L, Miler O. 2020. Effects of a retaining wall and an artificial embankment on nearshore littoral habitats and biota in a large Alpine lake. Hydrobiologia. 847:365-389.

Peng CYJ, Manz BD, Keck J. 2001. Modeling categorical variables by logistic regression. American Journal of Health Behavior 25:278-284.

Peng CYJ, Lee KL, Ingersoll GM. 2002. An introduction to logistic regression analysis and reporting. The Journal of Educational Research. 96:3-14.

Petrone KC. 2010. Catchment export of carbon, nitrogen, and phosphorus across an agro-urban land use gradient, Swan-Canning River system, southwestern Australia. Journal of Geophysical Research. 115:1-16.

Poisson AC, McCullough IM, Cheruvelil KS, Elliott KC, Latimore JA, Soranno PA. 2020. Quantifying the contribution of citizen science to broad-scale ecological databases. Frontiers in Ecology and the Environment. 18:19-26.

Porcal P, Amirbahman A, Kopáček J, Norton SA. 2010. Experimental photochemical release of organically bound aluminum and iron in three streams in Maine, USA. Environmental monitoring and assessment. 171:71-81.

Prairie YT, Kalff J. 1986. Effect of catchment size on phosphorus export. JAWRA Journal of the American Water Resources Association. 22: 465–470.

Psenner R, Pucsko R. 1988. Phosphorus fractionation: advantages and limits of the method for the study of sediment P origins and interactions. Arch Hydrobiol Beih Ergebn Limnol. 30:43-59.

Puth MT, Neuhäuser M, Ruxton GD. 2015. On the variety of methods for calculating confidence intervals by bootstrapping. Journal of Animal Ecology. 84:892—897.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for statistical Computing, Vienna, Austria.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Radomski P, Goeman TJ. 2001. Consequences of human lakeshore development on emergent and floating-leaf vegetation abundance. North American Journal of Fisheries Management. 21:46-61.

Read EK, Patil VP, Oliver SK, Hetherington AL, Brentrup JA, Zwart JA, Winters KM, Corman JR, Nodine ER, Woolway RI, Dugan HA, Jaimes A, Santoso AB, Hong GS, Winslow LA, Hanson PC, Weathers KC. 2015. The importance of lake-specific characteristics for water quality across the continental United States. Ecological Applications. 25:943-955.

Read JS, Hamilton DP, Jones ID, Muraoka K, Winslow LA, Kroiss R, Wu CH, Gaiser E. 2011. Derivation of lake mixing and stratification indices from high-resolution lake buoy data. Environmental Modelling & Software. 26:1325-1336.

Read JS, Rose KC. 2013. Physical responses of small temperate lakes to variation in dissolved organic carbon concentrations. Limnology and Oceanography. 58:921-931.

Reckhow H. 1977. Empirical lake models for phosphorus: Development, applications, limitations and uncertainty. *In* D. Scavia and A. Robertson (ed.), Lake ecosystem modeling. Ann Arbor. 193-221.

Reckhow KH, Chapra SC. 1983. Engineering approaches for lake management; Volume 1: Data analysis and empirical modeling. Butterworth Publishers Boston.

Reinhardt RL, Norton SA, Handley M, Amirbahman A. Dynamics of P, Al, and Fe during high discharge episodic acidification at the Bear Brook Watershed in Maine, USA. *In* Biogeochemical Investigations of Terrestrial, Freshwater, and Wetland Ecosystems across the Globe 2004 (pp. 311-323). Springer, Dordrecht.

Rippey B, McSorley C. 2009. Oxygen depletion in lake hypolimnia. Limnology and oceanography 54:905-916.

Robertson DM, Ragotzkie RA. 1990. Changes in the thermal structure of moderate to large sized lakes in response to changes in air temperature. Aquatic Sciences. 52:360-380.

Rose KC, Greb SR, Diebel M, Turner MG. 2017. Annual precipitation regulates spatial and temporal drivers of lake water clarity. Ecological Applications. 27:632-643.

Rose KC, Winslow LA, Read JS, Hansen GJ. 2016. Climate-induced warming of lakes can be either amplified or suppressed by trends in water clarity. Limnology and Oceanography Letters. 1:44–53.

Rosfjord CH, Webster KE, Kahl JS, Norton SA, Fernandez IJ, Herlihy AT. 2007. Anthropogenically driven changes in chloride complicate interpretation of base cation trends in lakes recovering from acidic deposition. Environmental Science & Technology. 41:7688-7693.

Roth BM, Kaplan IC, Sass GG, Johnson PT, Marburg AE, Yannarell AC, Havlicek TD, Willis TV, Turner MG, Carpenter SR. 2007. Linking terrestrial and aquatic ecosystems: the role of woody habitat in lake food webs. Ecological Modelling. 203:439-452.

Ruppert KM, Kline RJ, Rahman MS. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. 2019. Global Ecology and Conservation. 17:e00547.

Sass GG, Kitchell JF, Carpenter SR, Hrabik TR, Marburg AE, Turner MG. 2006. Fish community and food web responses to a whole-lake removal of coarse woody habitat. Fisheries. 31:321-330.

Savino JF, Stein RA. 1982. Predator-prey interaction between largemouth bass and bluegills as influenced by simulated, submersed vegetation. Transactions of the American Fisheries Society. 111:255-266.

Scheffer M, Hosper SH, Meijer ML, Moss B, Jeppesen E. 1993. Alternative equilibria in shallow lakes. Trends in Ecology & Evolution. 8:275-279.

Schindler DE, Geib SI, Williams MR. 2000. Patterns of fish growth along a residential development gradient in north temperate lakes. Ecosystems. 3:229-237.

Schindler DW, Bayley SE, Parker BR, Beaty KG, Cruikshank DR, Fee EJ, Schindler EU, Stainton MP. 1996. The effects of climatic warming on the properties of boreal lakes and streams at the Experimental Lakes Area, northwestern Ontario. Limnology and Oceanography. 41:1004-1017.

Schindler DW. 2001. The cumulative effects of climate warming and other human stresses on Canadian freshwaters in the new millennium. Canadian Journal of Fisheries and Aquatic Sciences. 58:18-29.

Schmidt W. 1928. Über Temperatur und Stabilitätsverhältnisse von Seen, Geograf. Ann. 10, 145–177.

Schwefel R, Steinsberger T, Bouffard D, Bryant LD, Müller B, Wüest A. 2018. Using small-scale measurements to estimate hypolimnetic oxygen depletion in a deep lake. Limnology and Oceanography. 63:S54-S67.

Shadish WR, Zuur AF, Sullivan KJ. 2014. Using generalized additive (mixed) models to analyze single case designs. Journal of school psychology. 52:149-178.

Snucins E, Gunn J. 2000. Interannual variation in the thermal structure of clear and colored lakes. Limnology and Oceanography. 45:1639-1646.

Soranno PA, Cheruvelil KS, Wagner T, Webster KE, Bremigan MT. 2015. Effects of land use on lake nutrients: the importance of scale, hydrologic connectivity, and region. PLoS ONE. 10:1-22.

Soranno PA, Hubler SL, Carpenter SR, Lathrop RC. 1996. Phosphorus loads to surface waters: a simple model to account for spatial pattern of land use. Ecological Applications. 6:865-878.

Speiss AN. 2018. qpcR: Modelling and Analysis of Real-Time PCR Data. R package version 1.4-1. https://cran.r-project.org/package=qpcR.

Steedman RJ. 2000. Effects of experimental clearcut logging on water quality in three small boreal forest lake trout (*Salvelinus namaycush*) lakes. Canadian Journal of Fisheries and Aquatic Sciences. 57:92-96.

Stockwell JD, Doubek JP, Adrian R, Anneville O, Carey CC, Carvalho L, De Senerpont Domis LN, Dur G, Frassl MA, Grossart HP, Ibelings BW. 2020. Storm impacts on phytoplankton community dynamics in lakes. Global Change Biology. 26:2756-2784.

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications. 16:1267–1276.

Stoddard JL, Van Sickle J, Herlihy AT, Brahney J, Paulsen S, Peck DV, Mitchell R, Pollard AI. 2016. Continental-scale increase in lake and stream phosphorus: Are oligotrophic systems disappearing in the United States? Environmental Science & Technology. 50:3409-3415.

Strom KM. 1931. Feforvatn. A physiographic and biological study of a mountain lake. Archiv für Hydrobiologie 22:491-536.

Sutherland JW, Norton SA, Short JW, Navitsky C. 2018. Modeling salinization and recovery of road salt-impacted lakes in temperate regions based on long-term monitoring of Lake George, New York (USA) and its drainage basin. Science of the Total Environment. 637-638:282-294.

Taranu ZE, Gregory-Eaves I. 2008. Quantifying relationships among phosphorus, agriculture, and lake depth at an inter-regional scale. Ecosystems. 11:715-725.

Thompson WB, Borns HW. 1985. Surficial Geologic Map of Maine. 1:250,000. Department of Conservation. Maine Geological Survey. Augusta, ME.

Thornton MM, Shrestha R, Wei Y, Thornton PE, Kao S, Wilson BE. 2020. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1840

Twardochleb LA, Olden JD. 2016. Human development modifies the functional composition of lake littoral invertebrate communities. Hydrobiologia. 775:167-184.

United States Department of Agriculture – Natural Resource Conservation Service, 2012a. USDA/NRCS - National Geospatial Center of Excellence. 1981-2010 Annual Average Precipitation by State. Washington (DC).

United States Department of Agriculture – Natural Resource Conservation Service, 2012b. USDA/NRCS - National Geospatial Center of Excellence. 1981-2010 Annual Average Maximum Temperature by State. Washington (DC).

United States Department of Agriculture – Natural Resource Conservation Service, 2012c. USDA/NRCS - National Geospatial Center of Excellence. 1981-2010 Annual Average Minimum Temperature by State. Washington (DC).

United States Environmental Protection Agency. 2000. Ambient water quality criteria recommendations: Information supporting the development of state and tribal Nutrient criteria For Lakes and reservoirs in nutrient ecoregion VIII. EPA-822-B-00-010. U.S. EPA Office of Water, Washington, D.C., USA.

United States Geological Survey (USGS) and U.S. Department of Agriculture, Natural Resources Conservation Service. 2013. Federal Standards and Procedures for the National Watershed Boundary Dataset (WBD) (4th ed.): Techniques and Methods 11–A3, 63 p.

USEPA. 2009. National Lakes Assessment: A Collaborative Survey of the Nation's Lakes. EPA 841-R-09-001. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, D.C. https://www.epa.gov/sites/default/files/2013-11/documents/nla_newlowres_fullrpt.pdf

USEPA. 2016. National Lakes Assessment 2012: A Collaborative Survey of Lakes in the United States. EPA 841-R-16-113. U.S. Environmental Protection Agency, Washington, DC. https://nationallakesassessment.epa.gov/

USEPA. 2017. National Lakes Assessment 2017. Field Operations Manual. EPA 841-B-16-002. U.S. Environmental Protection Agency, Washington, DC. https://www.epa.gov/sites/default/files/2021-01/documents/nla_2017_fom_version_1.1_2017_04_06.pdf

Uttormark PD, Wall JP. 1975. Lake classification, a trophic characterization of Wisconsin lakes. Corvallis (OR). National Environmental Research Center.

Vaithiyanathan P, Correll DL. 1992. The Rhode River watershed: Phosphorus distribution and export in forest and agricultural soils. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America. 21:280-288.

Venables WN, Ripley BD. 2002. Modern Applied Statistics with S, Fourth edition. Springer, New York. ISBN 0-387-95457-0, https://www.stats.ox.ac.uk/pub/MASS4/.

Vermaire JC, Gregory-Eaves I. 2007. Reconstructing changes in macrophyte cover in lakes across the northeastern united states based on sedimentary diatom assemblages. Journal of Paleolimnology. 39:477-490.

Verpoorter C, Kutser T, Seekell DA, Tranvik LJ. 2014. A global inventory of lakes based on high-resolution satellite imagery. Geophysical Research Letters. 41:6396-6402.

Vollenweider RA, Janus LL. 1982. Statistical models for predicting hypolimnetic oxygen depletion rates. Mem. Ist. Ital. Idrobiol. 40:1-24.

Vollenweider RA. 1975. Input-output models. Schweizerische Zeitschrift für Hydrologie. 37:53-84.

Weaver MJ, Magnuson JJ, Clayton MK. 1997. Distribution of littoral fishes in structurally complex macrophytes. Canadian Journal of Fisheries and Aquatic Sciences. 54:2277-2289.

Welch EB, Perkins MA. 1979. Oxygen deficit-phosphorus loading relation in lakes. Journal (Water Pollution Control Federation). 1:2823-2828.

Wetzel RG. 2001. Limnology: Lake and River Ecosystems. Gulf Professional Publishing.

Wilkerson E, Hagan JM, Whitman AA. 2010. The effectiveness of different buffer widths for protecting water quality and macroinvertebrate and periphyton assemblages of headwater streams in Maine, USA. Canadian Journal of Fisheries and Aquatic Sciences. 67:177-190.

Wilson TA, Norton SA, Lake B, Amirbahman A. 2008. Sediment geochemistry of Al, Fe, and P for two oligotrophic Maine lakes during a period of acidification and recovery. Science of the Total Environment. 404:269-275.

Winslow L, Read J, Woolway R, Brentrup J, Leach T, Zwart J, Albers S, Collinge D. 2019. rLakeAnalyzer: Lake Physics Tools. R package version 1.11.4.

Wood S. 2021. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation. R package version 1.8-34. https://cran.r-project.org/web/packages/mgcv/index.html

Wright RF, Norton SA, Brakke DF, Frogner T. 1988. Acidification of stream water by whole-catchment experimental addition of dilute seawater. Nature. 334:422-424.

Yang J, Zhang X. 2020. eDNA metabarcoding in zooplankton improves the ecological status assessment of aquatic ecosystems. Environment international. 134:105230.

Yuan LL, Jones JR. 2019. Modeling hypolimnetic dissolved oxygen depletion using monitoring data. Canadian Journal of Fisheries and Aquatic Sciences. 77:814-823.

Yuan LL, Jones JR. 2020. Rethinking phosphorus–chlorophyll relationships in lakes. Limnology and Oceanography. 65:1847-1857.

Zeileis A, Hothorn T. 2002. Diagnostic Checking in Regression Relationships. R News. 2:7

Zhao CS, Shao NF, Yang ST, Ren H, Ge YR, Feng P, Dong BE, Zhao Y. 2017. Predicting cyanobacteria bloom occurrence in lakes and reservoirs before blooms occur. Science of the total environment. 670:837-48.

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. 2009. Mixed effect models and extentions in ecology with R. Springer Science and Buisness Media, New York.

Zuur AF, Tuck ID, Bailey N. 2003. Dynamic factor analysis to estimate common trends in fisheries time series. Canadian Journal of Fisheries and Aquatic Sciences. 60:542-552.

APPENDIX A. CLASSIFICATION

Table A. 1. Two-way ANOVA results for lake type (Eco4 + Depth) and three versions of condition classes based on various WQI thresholds specific to each lake type: v1 = 10th/90th percentiles, v2 = 25th/75th percentiles, v3 = 25th/90th percentiles to delineate altered, intermediate, and reference lakes.

Total Phosphorus

	df	Sum Sq.	Mean Sq.	F	p
Class	4	4.92	1.23	51.22	<0.001
WQI.v1	2	1.49	0.74	30.95	<0.001
Class:WQI.v1	8	0.74	0.09	3.86	<0.001
Residuals	434	10.43	0.02		
Class	4	4.92	1.23	51.85	<0.001
WQI.v2	2	1.65	0.82	34.67	<0.001
Class:WQI.v2	8	0.71	0.09	3.74	<0.001
Residuals	434	10.31	0.02		
Class	4	4.92	1.23	52.50	<0.001
WQI.v3	2	1.74	0.87	37.14	<0.001
Class:WQI.v3	8	0.74	0.09	3.95	<0.001
Residuals	434	10.18	0.02		

Specific Conductivity

	df	Sum Sq.	Mean Sq.	F	p
Class	4	4.12	1.03	37.46	<0.001
WQI.v1	2	1.36	0.68	24.75	<0.001
Class:WQI.v1	8	0.85	0.11	3.85	<0.001
Residuals	404	11.12	0.03		
Class	4	4.12	1.03	40.81	<0.001
WQI.v2	2	2.36	1.18	46.77	<0.001
Class:WQI.v2	8	0.76	0.10	3.77	<0.001
Residuals	404	10.20	0.03		
Class	4	4.12	1.03	40.84	<0.001
WQI.v3	2	2.22	1.11	44.03	<0.001
Class:WQI.v3	8	0.91	0.11	4.50	<0.001
Residuals	404	10.19	0.03		

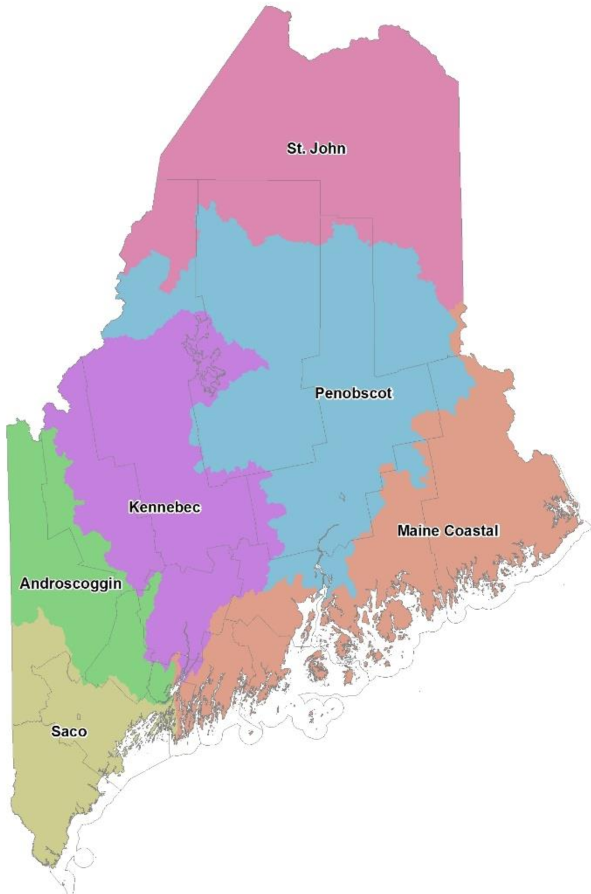
Table A. 2. Total Phosphorus (TP, in µg/L) summary data for each lake type and condition class within the Eco4-depth classification scheme. These data are illustrated in Figures 2.2 and 2.3. DL = Deep Lakes (≥10 m), SL = Shallow Lakes (< 10 m). DL = Deep Lakes; SL = Shallow Lakes; Ref = Reference; Int = Intermediate; Alt = Altered.

Lake Type	# Lakes	Min	1st.Q	Median	Mean	3rd.Q	Max	Lakes w/o Data	Interquartile Range
Coastal-DL-Ref	22	4.7	6.8	8.0	8.4	8.6	18.2	1	6.8 - 8.6
Coastal-DL-Int	147	4.0	7.0	8.7	8.9	10.3	20.0	38	7.0 - 10.3
Coastal-DL-Alt	57	5.6	9.7	12.2	14.4	17.6	33.5	9	9.7 - 17.6
All Coastal Lakes	*226*	*4.0*	*7.4*	*9.1*	*10.3*	*11.4*	*33.5*	*48*	*7.4 - 11.4*
Coastal-SL-Ref	16	7.6	8.3	9.3	11.4	13.1	23.1	1	8.3 - 13.1
Coastal-SL-Int	100	7.3	9.8	12.1	13.9	15.4	45.7	23	9.8 - 15.4
Coastal-SL-Alt	39	7.0	12.0	15.3	17.8	20.5	56.5	4	12.0 - 20.5
All Coastal Ponds	*155*	*7.0*	*9.7*	*12.7*	*14.7*	*17.6*	*56.5*	*28*	*9.7 - 17.6*
Inland-SL-Ref	3	7.0	7.1	7.3	7.3	7.4	7.5	1	7.1 - 7.4
Inland-SL-Int	18	6.0	9.2	11.1	11.5	13.0	19.3	7	9.2 - 13.0
Inland-SL-Alt	7	7.5	8.5	10.5	10.3	11.8	13.1	1	8.5 - 11.8
All Inland Ponds	*28*	*6.0*	*7.8*	*10.6*	*10.7*	*12.2*	*19.3*	*9*	*7.8 - 12.2*
Inland-DL-Ref	15	3.7	5.5	6.3	6.4	7.4	8.2	4	5.5 - 7.4
Inland-DL-Int	94	4.0	6.0	6.9	7.2	8.0	13.7	29	6.0 - 8.0
Inland-DL-Alt	37	4.4	5.6	8.0	8.3	9.2	16.3	12	5.6 - 9.2
All Inland Lakes	*146*	*4.0*	*7.4*	*9.1*	*10.3*	*11.4*	*33.5*	*146*	*7.4 - 11.4*
Northern-Ref	3	5.7	6.5	7.3	7.3	8.2	9.0	1	6.5 - 8.2
Northern-Int	26	6.3	9.2	13.0	13.4	16.8	24.6	8	9.2 - 16.8
Northern-Alt	4	16.7	17.2	19.5	48.4	50.7	138.0	0	17.2 - 50.7
All Northern Lakes	*33*	*5.7*	*9.0*	*13.2*	*18.7*	*17.1*	*138.0*	*9*	*9.0 - 17.1*

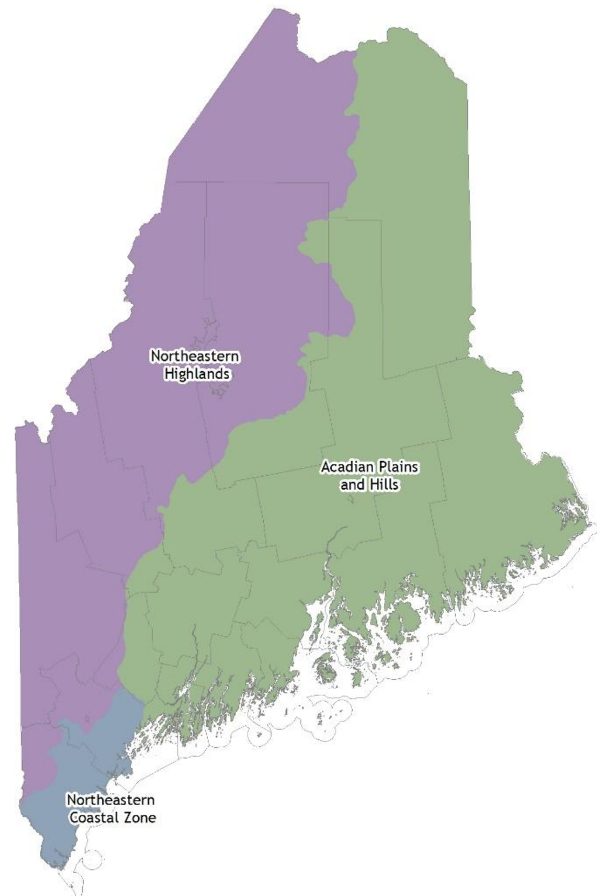
Table A. 3. Specific Conductivity (SpCon, in µS/cm) summary data for each lake type and condition class within the Eco4-depth classification scheme. These data are illustrated in Figures 2.2 and 2.3. DL = Deep Lakes (≥10 m), SL = Shallow Lakes (< 10 m). DL = Deep Lakes (>=10 m), SL = Shallow Lakes (< 10 m). DL = Deep Lakes; SL = Shallow Lakes; Ref = Reference; Int = Intermediate; Alt = Altered.

Lake Type	# Lakes	Min	1st.Qu	Median	Mean	3rd.Qu.	Max	Lakes w/o Data	Interquartile Range
Coastal-DL-Ref	22	19.0	26.2	32.5	34.2	39.1	76.3	5	26.2 - 39.1
Coastal-DL-Int	147	25.7	39.3	47.5	48.8	57.3	98.0	34	39.3 - 57.3
Coastal-DL-Alt	57	28.0	47.3	64.5	66.3	79.5	155.4	17	47.3 - 79.5
All Coastal Lakes	*226*	*19.0*	*38.2*	*47.6*	*51.5*	*61.2*	*155.4*	*56*	*38.2 - 61.2*
Coastal-SL-Ref	16	17.3	22.2	27.3	30.1	31.3	71.1	3	22.2 - 31.3
Coastal-SL-Int	100	18.5	40.3	52.3	57.8	71.1	219.6	25	40.3 - 71.1
Coastal-SL-Alt	39	27.3	48.4	70.2	80.0	85.6	347.6	13	48.4 - 85.6
All Coastal Ponds	*155*	*17.3*	*32.9*	*52.0*	*59.7*	*71.9*	*347.6*	*41*	*32.9 - 71.9*
Inland-SL-Ref	3	19.8	19.8	19.8	19.8	19.8	19.8	2	19.8 - 19.8
Inland-SL-Int	18	22.0	29.1	30.7	32.6	34.4	56.8	8	29.1 - 34.4
Inland-SL-Alt	7	17.4	34.1	37.0	40.4	41.4	74.9	1	34.1 - 41.4
All Inland Ponds	*28*	*17.4*	*28.9*	*31.7*	*34.6*	*37.7*	*74.9*	*11*	*28.9 - 37.7*
Inland-DL-Ref	15	22.6	23.5	29.4	30.4	37.1	42.0	5	23.5 - 37.1
Inland-DL-Int	94	18.8	24.0	27.6	29.2	33.1	46.5	35	24.0 - 33.1
Inland-DL-Alt	37	19.3	29.2	32.7	36.7	38.6	61.5	11	29.2 - 38.6
All Inland Lakes	*146*	*19.0*	*38.2*	*47.6*	*51.5*	*61.2*	*155.4*	*51*	*38.2 - 61.2*
Northern-Ref	3	49.7	61.1	72.6	72.6	84.0	95.5	1	61.1 - 84.0
Northern-Int	26	23.5	31.7	51.6	57.5	76.3	131.3	9	31.7 - 76.3
Northern-Alt	4	79.1	129.7	159.8	146.6	176.7	187.6	0	129.7 - 176.7
All Northern Lakes	*33*	*23.5*	*33.0*	*59.9*	*74.3*	*90.9*	*187.6*	*10*	*33.0 - 90.9*

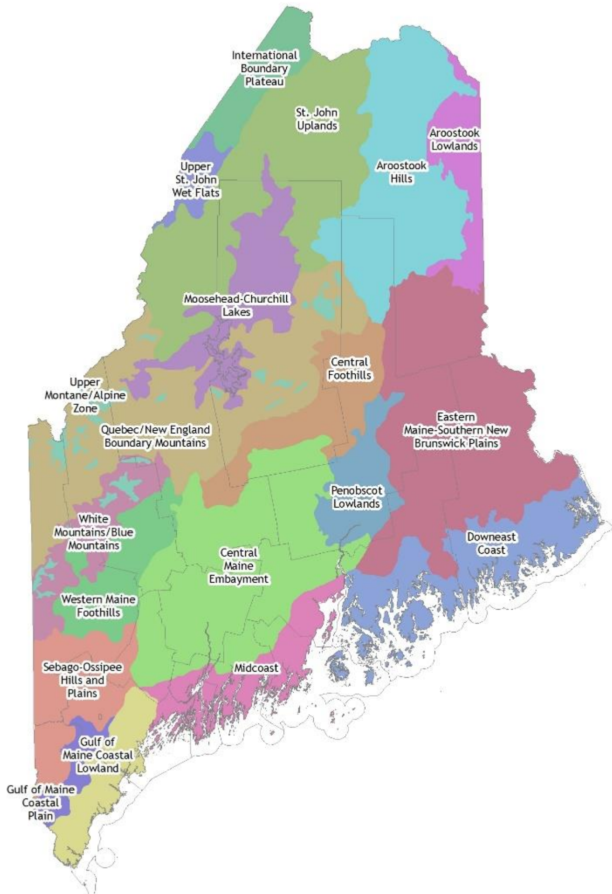
Figure A.1. Maps showing geographical schemes for lake classification schemes tested with linear mixed effect modeling.



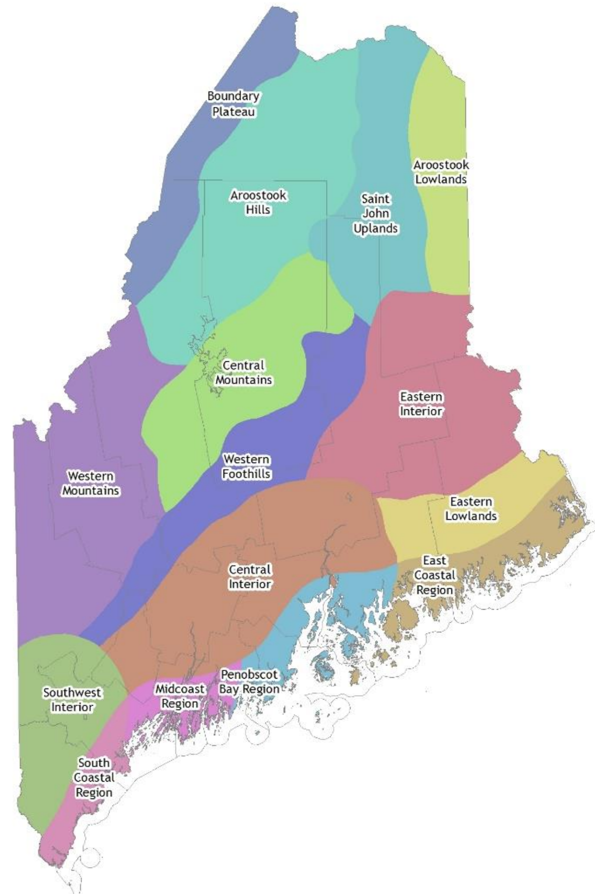
A.1a) HUC4 Areas



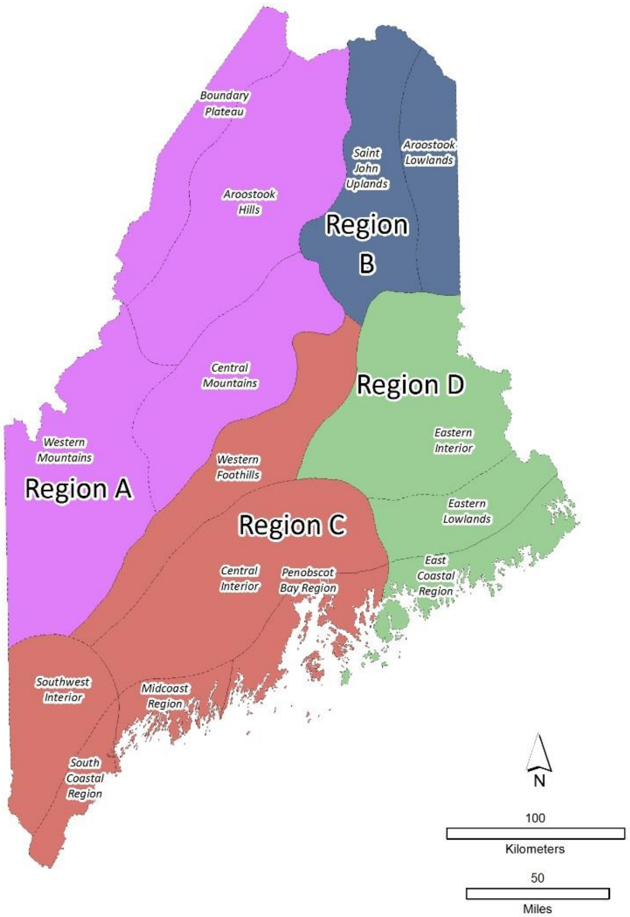
A.1b) Level III Ecoregions



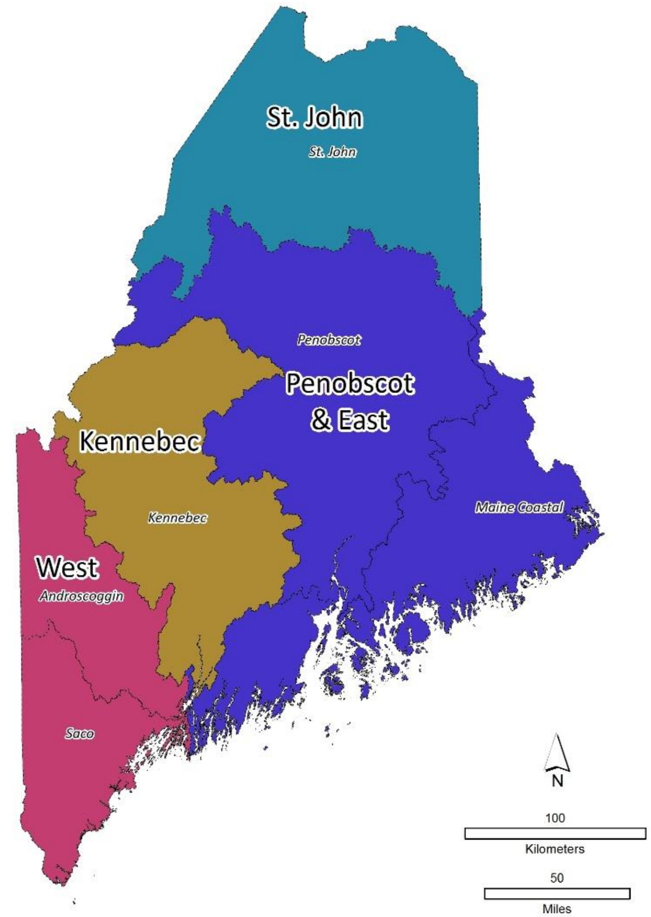
A.1c) Level IV Ecoregions



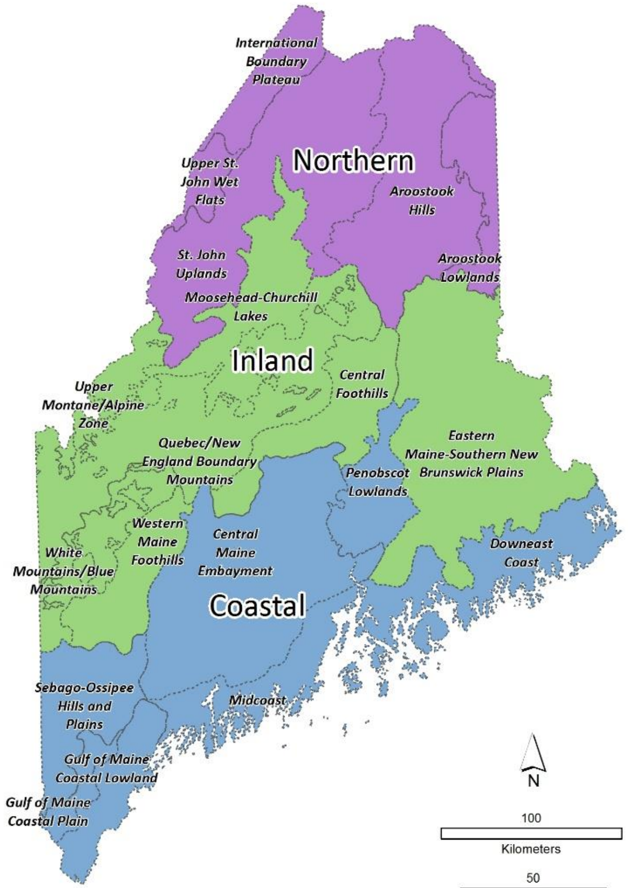
A.1d) Maine Biophysical Regions



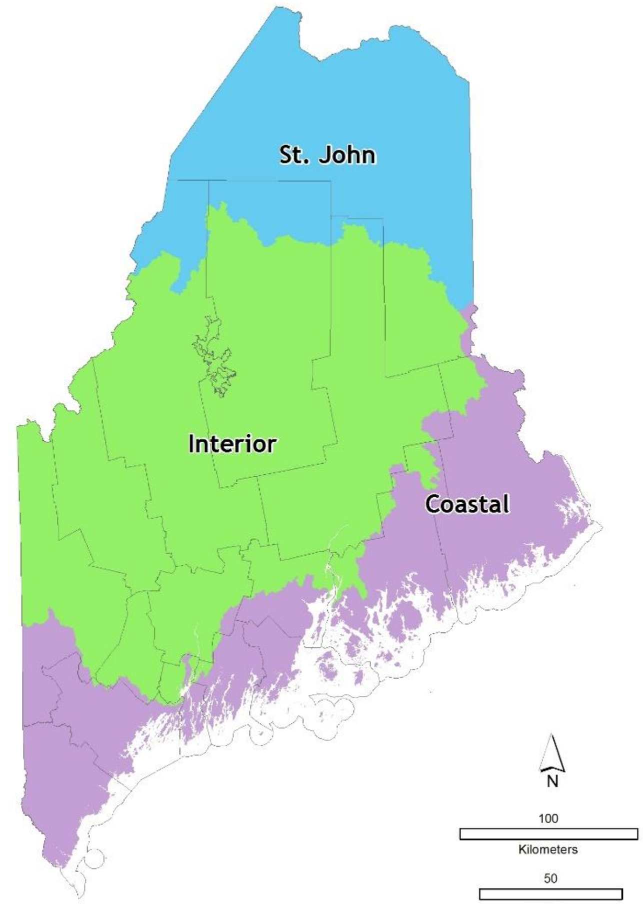
A.1e) Maine Biophysical Region aggregated areas
(“Biop”, from cluster analysis)



A.1f) HUC4 aggregated areas, from cluster analysis



A.1g) Level IV Ecoregion aggregated areas ("Eco4", from cluster analysis)



A.1h) Ecological Drainage Units (EDU) based on HUC4 areas

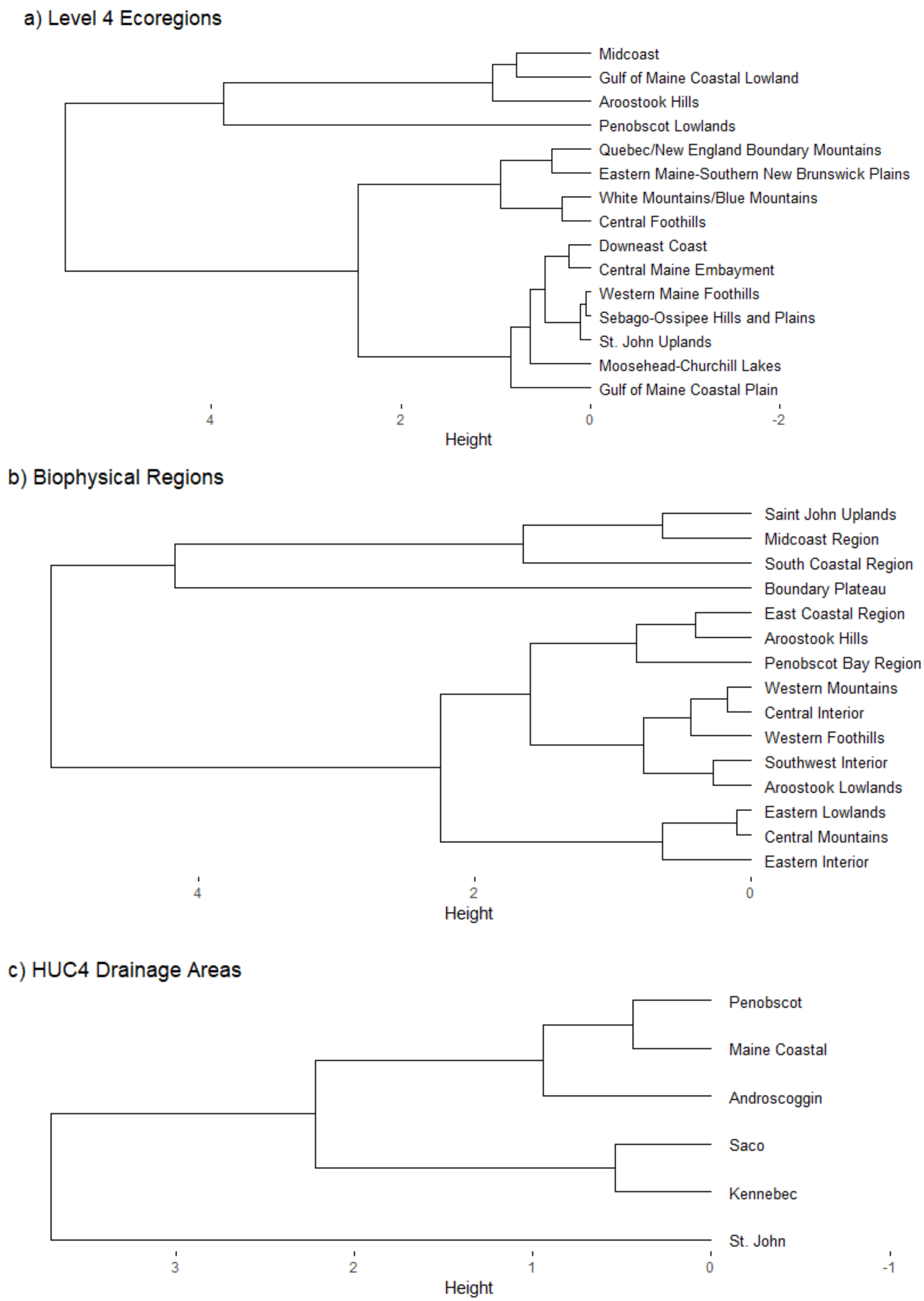


Figure A. 2. Cluster dendrograms used to guide formation of regional groupings for (a) Level IV Ecoregions, (b) biophysical regions, and (c) HUC4 drainage areas. Areas that clustered together and were geographically adjacent were grouped together in classification trials.

APPENDIX B. ANOXIA

B1. Logistic Regression

The basic equation for logistic regression is:

$$Logit(Y) = \alpha + (\beta_n \times X_n)$$

where the logit is the natural logarithm (ln) of the odds of *Y* happening, α is the y-intercept, X_n is the predictor, and β_n is the coefficient for variable *n*. In Logistic regression, *X* for each variable *n* can be categorical or continuous, but *Y* is always categorical. The null hypothesis of logistic regression models is that all β values equal zero, meaning that there is no linear relationship in the population. A rejection of the null hypothesis indicates that at least one $\beta > 0$ and that a linear relationship exists between *X* and the logit(*Y*) (Peng et al. 2002). The logit is converted to a probability of response, π:

$$\pi = \frac{e^{Logit(Y)}}{1 + e^{Logit(Y)}}$$

where π varies from 0 to 1. In application of the model, values of π are comparted to a probability threshold to evaluate the predicted response for each case.

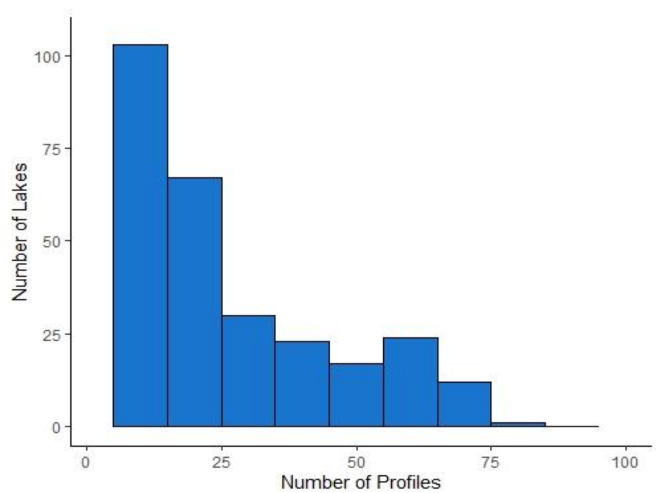


Figure B. 1. Histogram of all lake profiles from lakes with n ≥ 3 late summer profiles (1 August – 7 September) from the ME DEP lakes database (2015).

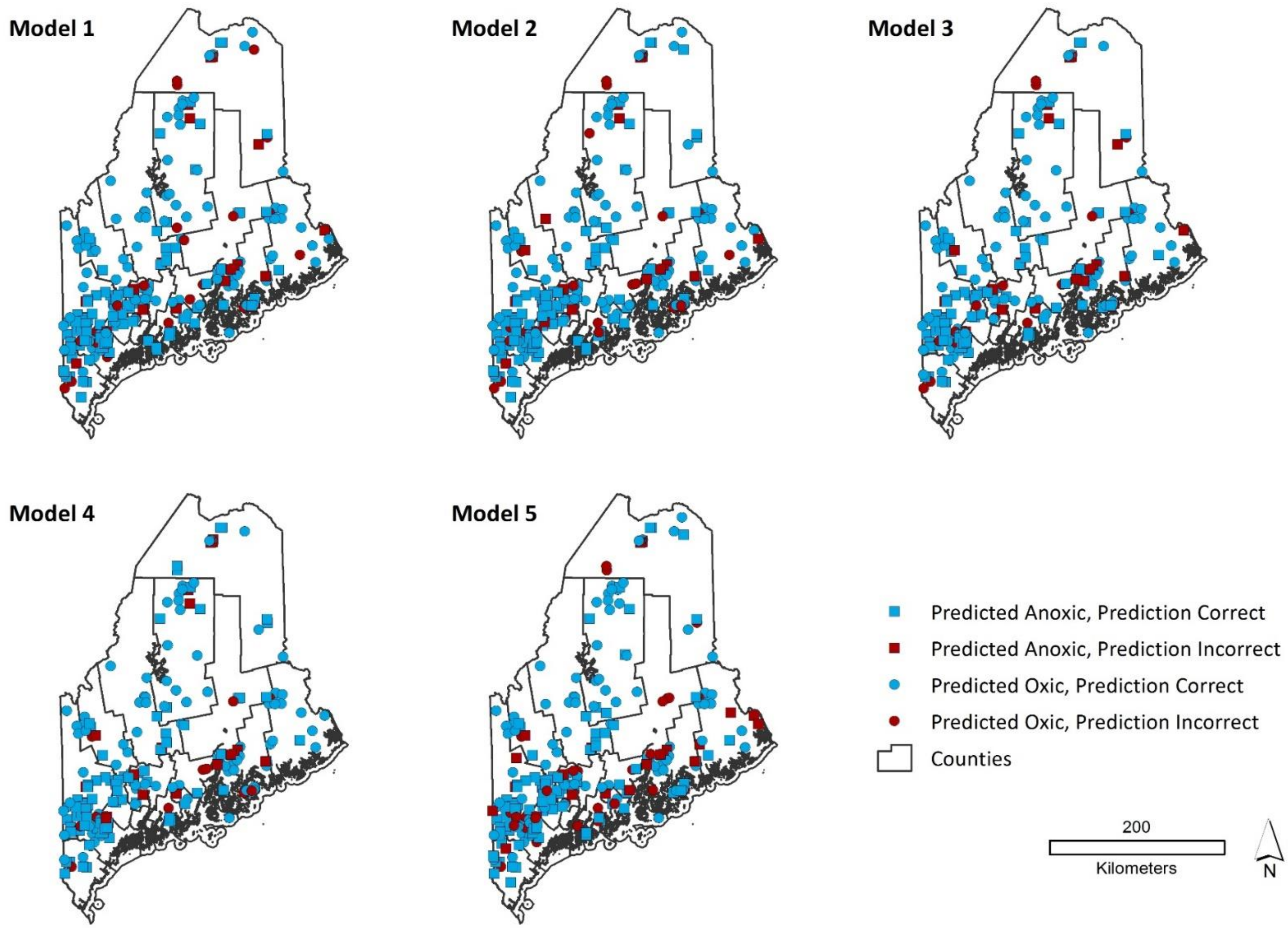


Figure B. 2. Spatial Distribution of correctly predicted lakes (blue) and incorrectly predicted lakes (red) with prediction success for each logistic regression model.

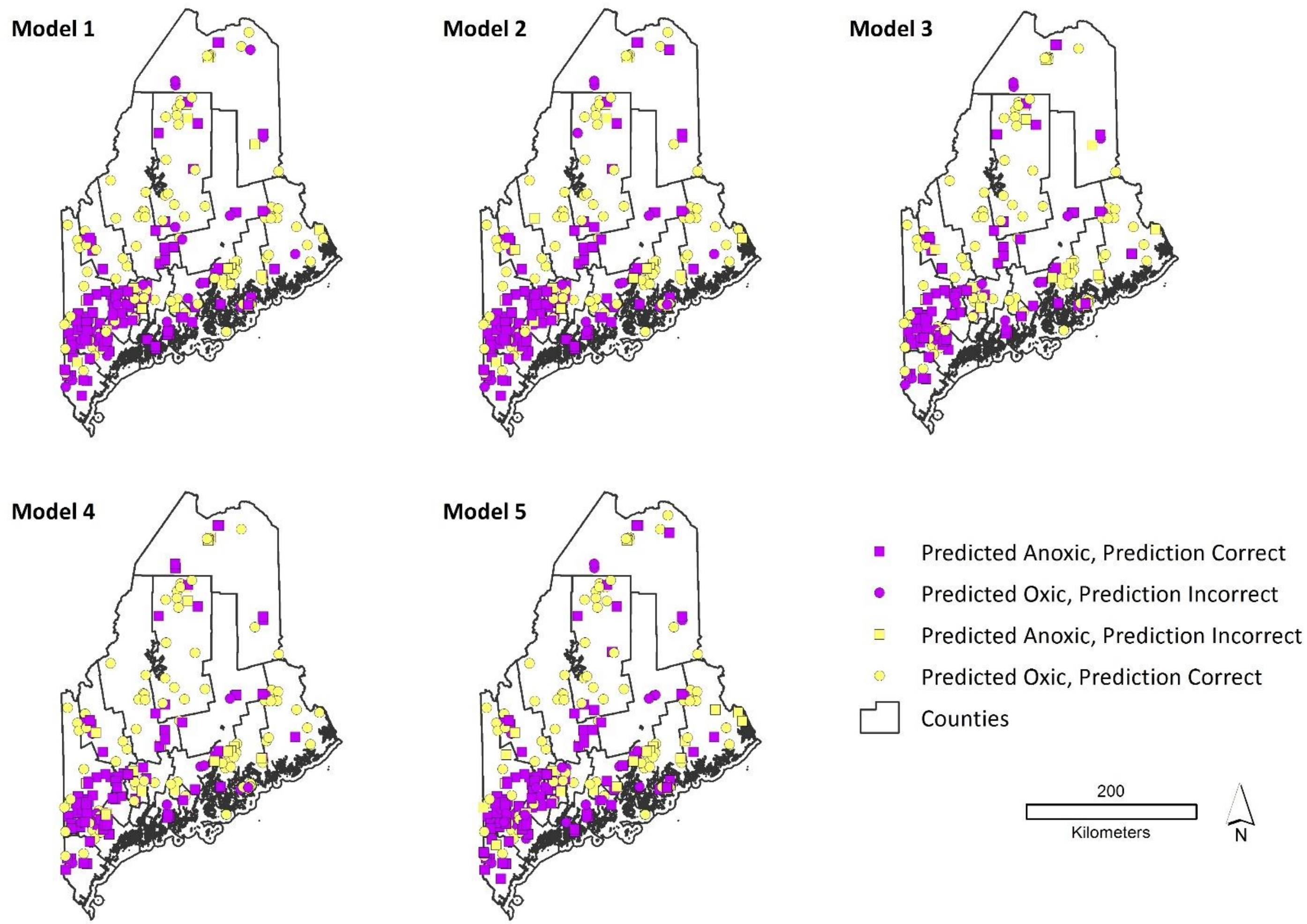


Figure B. 3. Spatial Distribution of anoxic (purple) and oxic (yellow) lakes with prediction success for each logistic regression model.

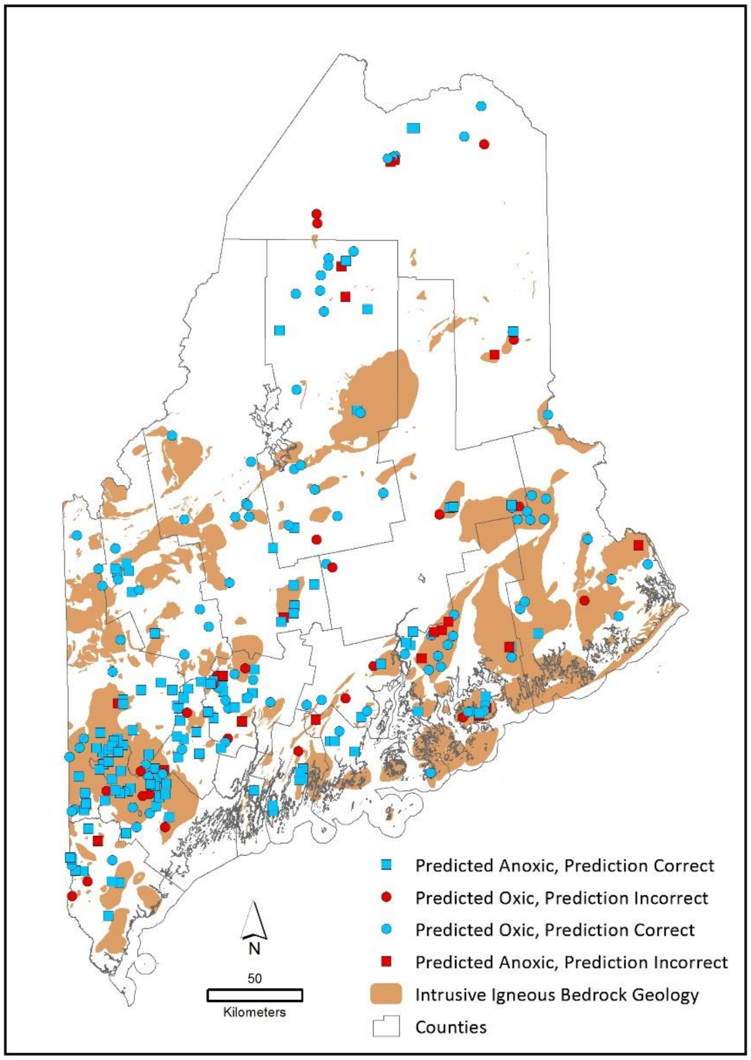


Figure B. 4. Model 1 results shown with intrusive igneous bedrock geology. A larger proportion (66%) of study lakes ($z_{max} \geq 10$m, TP ≤ 15 µg/L) are found on this type of bedrock geology, but there were no significant associations found between lakes with this geology and anoxic condition or predictive model results.

APPENDIX C. TRENDS

Table C. 1. Number of lakes used in the creation of reference lake trends, and the mean number of years of data for individual lakes in each lake type during 1999-2018. Inland Shallow and Northern Lake types include non-reference quality lakes to meet our data inclusion restrictions.

Reference Lake Type	Number of Reference Lakes	Mean years of data per Lake
Coastal Deep	10	17.3
Coastal Shallow	10	15.1
Inland Deep	6	14.7
Inland Shallow	8	13.2
Northern	6	10.3

Table C. 2. Selected traits of reference lakes for the five lake types in the study. SD = standard deviation, IQR = interquartile range (75th - 25th percentile values).

Metric	Coastal Deep			Coastal Shallow			Inland Deep			Inland Shallow			Northern		
	Mean	SD	IQR	mean	SD	IQR	mean	SD	IQR	mean	SD	IQR	mean	SD	IQR
Maximum Depth (m)	20.0	12.1	6.4	5.2	2.8	3.5	19.2	4.8	4.7	6.2	2.8	2.5	28.7	16.5	23.2
Mean Depth (m)	7.4	3.4	5.6	2.2	1.3	1.7	7.2	2.1	1.4	3.3	1.6	1.9	9.1	4.3	4.1
Catchment Area (km2)	14.5	25.9	8.3	4.1	7.8	1.1	101.1	101.3	153.8	15.3	14.3	19.5	109.6	88.0	154.3
Surface Area (ha)	173.7	241.6	266.9	28.8	25.4	9.0	1332.9	818.4	652.3	138.9	97.0	142.3	1435.6	1506.3	2621.9
Volume (m3)	18.2	33.1	15.9	1.0	0.9	0.7	100.0	76.6	109.3	4.7	3.8	6.4	164.9	194.2	333.2
Flushes yr-1	1.7	1.8	1.7	3.6	3.2	3.7	2.0	1.8	1.4	2.5	2.2	2.0	3.1	4.4	2.9
Catchment:Lake Area	10.0	6.7	11.5	11.1	9.2	7.4	7.0	4.7	5.9	9.3	6.1	5.9	21.4	29.7	7.2
Dissolved Organic Carbon (mg/L)	3.5	1.0	1.1	4.5	2.2	2.1	5.2	2.0	2.4	5.1	1.3	1.6	5.7	1.5	1.7
Total Phosphorus (ug/L)	6.8	1.8	1.4	9.3	3.1	2.7	5.7	0.8	0.4	9.3	3.3	3.2	9.4	4.4	3.4
Chlorophyll-a (ug/L)	3.4	1.4	1.8	4.1	2.1	2.8	2.9	0.9	0.6	4.3	2.8	1.2	5.2	3.4	3.4
Secchi Disk Transparency (m)	6.8	2.3	2.2	5.1	1.2	1.2	5.9	1.5	2.3	4.3	1.3	0.8	5.3	2.4	1.0
Estimated Catchment Population	104.0	177.8	90.1	24.1	38.6	13.3	6.6	10.2	9.9	75.1	79.2	93.8	147.3	104.9	149.2
Total Developed Wshed Area (%)	0.1	0.3	0.1	0.1	0.2	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.4	0.3	0.3
Total Forest + Wetland Wshed Area (%)	81.3	12.1	16.4	75.6	29.9	27.6	80.6	11.2	15.5	66.5	8.9	8.1	73.1	4.0	3.8
Total Agricultural Wshed Area (%)	1.4	1.1	1.1	0.2	0.3	0.3	0.0	0.0	0.0	1.2	1.8	2.2	8.4	8.1	7.0
Total Undeveloped Wshed Area (%)	1.4	1.1	1.1	0.2	0.3	0.3	0.0	0.0	0.0	1.2	1.8	2.2	8.4	8.1	7.0
Total Forested Wshed Area (%)	97.4	1.0	1.4	99.2	0.6	1.2	99.9	0.1	0.1	96.7	2.1	2.8	89.5	8.3	7.6
Total Wetland Wshed Area (%)	78.3	14.3	19.6	72.7	28.8	26.5	74.8	10.7	5.4	64.2	9.0	8.5	64.0	9.0	5.9

Table C. 3. Model results for different types of variance and different numbers of common trends (*M*) for Dynamic Factor Analysis models incorporating Secchi depth time series for five separate lake types. The model selected as best for this analysis, Model 4, had two common trends and diagonal and equal variance which indicates shared variance among lake types with no year-to-year correlations among lake types. Other variance types: diagonal and unequal = unique variance for each lake type and no year-to-year correlations among lake types; equal var/covariance = same variance for each lake type and same covariance among sites (most constrained); unconstrained = variances all different (least constrained).

Model #	Variance Type	*M*	logLik	AICc	ΔAICc	ΔAICc Evidence Ratio
1	equal var/covariance	1	-112.68	240.74	0.00	--
2	diagonal and unequal	1	-109.19	241.21	0.47	1.26
3	diagonal and equal	1	-114.91	242.84	2.10	2.86
4	diagonal and equal	2	-110.68	244.18	3.44	5.59
5	equal var/covariance	2	-110.42	246.27	5.53	15.84
6	diagonal and unequal	2	-106.36	246.39	5.65	16.83
7	diagonal and equal	3	-110.82	252.49	11.75	355.96
8	equal var/covariance	3	-110.41	254.51	13.76	974.98
9	diagonal and unequal	3	-106.52	255.66	14.92	1.73E+03
10	diagonal and equal	4	-110.68	257.94	17.20	5.42E+03
11	unconstrained	1	-103.33	259.00	18.26	9.25E+03
12	equal var/covariance	4	-110.41	260.38	19.64	1.84E+04
13	diagonal and unequal	4	-106.49	261.99	21.25	4.12E+04
14	unconstrained	2	-102.24	271.24	30.50	4.19E+06
15	unconstrained	3	-102.51	283.80	43.06	2.24E+09
16	unconstrained	4	-102.24	291.98	51.24	1.34E+11

Table C. 4. Model comparison table for DFA model 4 (from Table C3) and climate covariables. Model 4 (with no covariables) is Ranked 17th in model strength compared to models with covariates. Data for the 10 strongest covariate models and the no-covariate model are presented here. Legend: PRCP = mean monthly precipitation; PDAY = percent of days with any precipitation; PINT = percent of days with precipitation > 10mm; TAVG = mean of daily average air temperatures; DD = total cumulative degree days (days > 18.3°C or 65°F). Lowercase suffixes indicate time period: st = stratification season (April – August); su = summer (June - August); w = water year (previous October – September, inclusive). Spring (April-June) and winter (December-February) were also calculated but did not improve model performance. PRCPst, in bold, was chosen as the best covariate model for this analysis. Combinations of paired precipitation and air temperature covariates were also tested (e.g., PRCPst + DDAYst) but did not exceed model performance of PRCPst as the sole covariate and are not presented here.

Covariate Model Ranking	Covariate	LogLik	AIC	AICc	ΔAICc	Evidence Ratio
1	**PRCPst**	**-97.67**	**225.33**	**231.91**	**0**	**--**
2	PDAYst	-99.02	228.04	234.62	2.71	3.88
3	DDAYwa	-99.28	228.57	235.14	3.23	5.03
4	PDAYspsu	-99.92	229.84	236.41	4.50	9.49
5	PDAYsu	-100.47	230.94	237.51	5.60	16.44
6	PRCPspsu	-100.47	230.94	237.52	5.61	16.53
7	PINTst	-101.75	233.49	240.07	8.16	59.15
8	PRCPsu	-102.47	234.94	241.51	9.60	121.51
9	DDAYsp	-102.51	235.02	241.59	9.68	126.47
10	TAVGwa	-103.04	236.07	242.65	10.74	214.86
17	*No Covariates* (Model 4)	-110.68	241.36	244.18	12.27	461.74

APPENDIX D. HABITAT

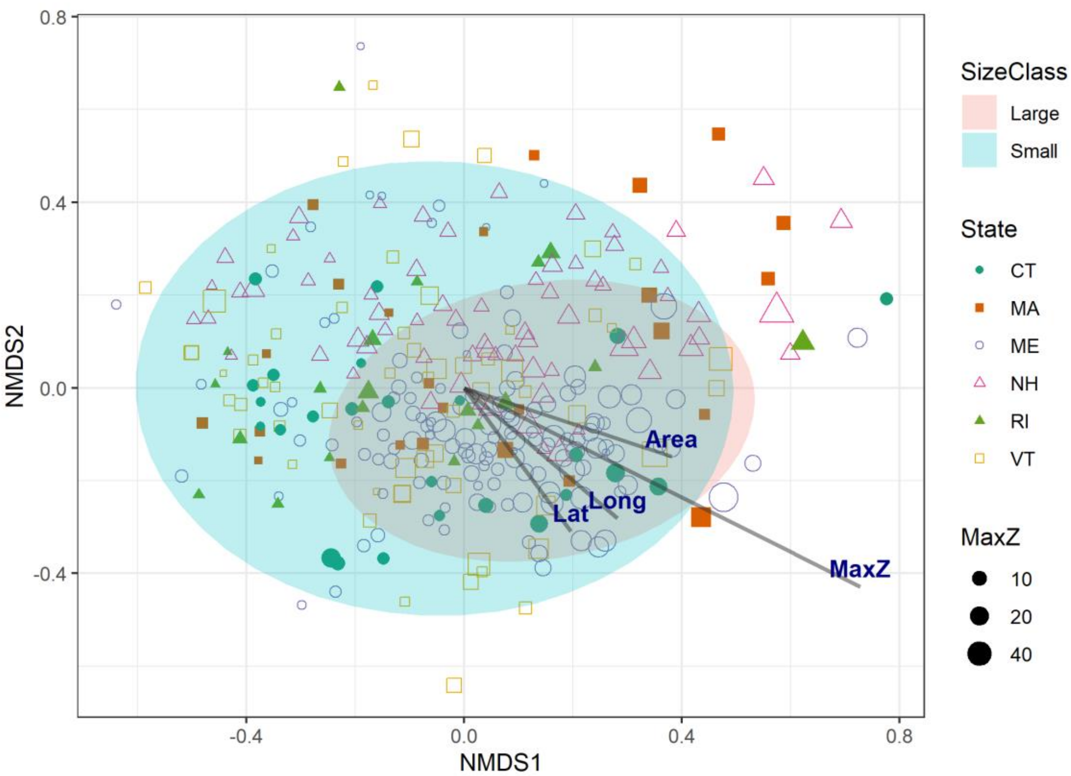


Figure D. 1. Non-metric multi-dimensional scaling (NMDS) plot of lakes based on habitat metric data. Colored areas represent size classes of lakes (Large ≥ 81 ha), colors and shapes of dots represent different states, and dot size indicates maximum lake depth (MaxZ, in m). Vectors show the relative importance (line length) and direction of association of lakes with the continuous variables Latitude (Lat), Longitude (Long), Surface area (Area) and maximum depth (MaxZ). This plot illustrates general patterns in the data that led to the creation of lake types based on depth and size criteria; there was no apparent separation in data based on state.

Table D. 1. Results of habitat model assessments for individual lakes. **LAKE ID**: Lake Identification names or codes, **MEAN LDA SCORE**: Individual lake Linear Discriminant Analysis model scores, **MEAN RIPSCORE**: mean score indicating degree of riparian disturbance (see main text and Table 3), **BCI (LOW)** and **BCI (HIGH)**: Low and High values for 95% Bootstrapped Confidence Intervals based on LDA scores from all sites on each lake, and **BCI ASSESSMENT**: the resulting lake littoral habitat assessment based on BCI scores. Regional results include Maine lakes, as calculated with Regional indices. See text for details.

Maine Deep

LAKE NAME	MIDAS (LAKE ID)	MEAN LDA SCORE	MEAN RIPSCORE	BCI (LOW)	BCI (HIGH)	BCI ASSESSMENT
DEBOULLIE	1512	2.428	0.727	1.943	2.857	Natural
PARKER	5186	2.385	0.591	2.092	2.728	Natural
THREECORNERED	5424	2.177	0.437	1.771	2.532	Natural
ELLIS	4086	2.053	0.800	1.798	2.293	Natural
TORSEY	5307	1.861	0.349	1.328	2.293	Natural
LITTLE JIM	5090	1.850	0.659	1.577	2.108	Natural
B POND	3276	1.799	0.838	1.512	2.129	Natural
WADLEIGH	572	1.676	0.784	1.140	2.402	Natural
LITTLE BIG WOOD	2630	1.594	0.516	1.220	2.202	Natural
BIG REED	2842	1.485	0.627	1.245	1.748	Natural
TROUT	3212	1.295	0.702	1.089	1.515	Natural
BLACK	1506	1.245	0.767	0.878	1.496	Natural
GRASS	104	1.169	0.801	0.747	1.657	Natural
ECHO	5814	1.145	0.255	0.802	1.410	Natural
SEBEC	848	0.797	0.235	0.365	1.102	Natural
TWIN #1	2026	0.760	0.688	-0.081	1.299	Intermediate
TOGUS	9931	0.714	0.266	0.273	1.159	Natural
NEQUASSET	5222	0.580	0.502	-0.407	1.298	Intermediate
NAHMAKANTA	698	0.317	0.797	-0.028	0.707	Intermediate
PEAKED MOUNTAIN	1254	0.281	0.435	-0.019	0.505	Intermediate
GARDNER	1528	0.226	0.657	-0.055	0.546	Intermediate
UPPER NARROWS	98	0.153	0.289	-0.383	0.912	Intermediate
LOWER SO BR	4222	0.074	0.633	-0.172	0.383	Intermediate
PORTER	12	0.072	0.284	-0.324	0.549	Intermediate
WILSON - KN	3832	0.070	0.377	-0.202	0.394	Intermediate

MARANACOOK	5312	0.054	0.253	-0.410	0.664	Intermediate
CHAMBERLAIN	2882	-0.071	0.679	-0.467	0.220	Intermediate
JAMIES	5302	-0.109	0.534	-1.091	0.321	Intermediate
BUBBLE	4452	-0.110	0.589	-0.519	0.278	Intermediate
VARNUM	3680	-0.221	0.468	-0.551	0.081	Intermediate
FLYING	5182	-0.232	0.203	-0.767	0.628	Intermediate
TUNK	4434	-0.344	0.660	-0.758	-0.036	Impaired
WOODBURY	5240	-0.372	0.213	-0.742	0.427	Intermediate
THREE CORNER	5384	-0.375	0.434	-0.668	0.023	Intermediate
WEBBER	5408	-0.402	0.324	-0.914	0.010	Intermediate
LOWER HADLOCK	4610	-0.419	0.380	-0.864	0.042	Intermediate
WASSOOKEAG	227	-0.427	0.194	-0.710	0.011	Intermediate
TRICKEY	3382	-0.514	0.185	-0.755	-0.040	Impaired
SALMON	5352	-0.593	0.372	-0.868	-0.244	Impaired
TRIPP	3758	-0.645	0.273	-1.016	-0.039	Impaired
CHINA W	54482	-0.695	0.624	-0.788	-0.507	Impaired
WEBB	3672	-0.782	0.246	-0.980	-0.591	Impaired
WILSON - FR	3682	-0.800	0.198	-1.170	-0.178	Impaired
PHILLIPS	4300	-0.812	0.110	-1.125	-0.471	Impaired
BRETTUNS	3608	-0.885	0.141	-1.150	-0.619	Impaired
PLEASANT	3446	-0.898	0.181	-1.050	-0.556	Impaired
COBBOSSECONTEE	5236	-0.993	0.166	-1.388	-0.541	Impaired
PENNESSEEWASSEE	3434	-1.000	0.168	-1.382	-0.472	Impaired
ANDROSCOGGIN LAK	3836	-1.057	0.392	-1.420	-0.713	Impaired
CHINA E	54481	-1.060	0.178	-1.348	-0.831	Impaired
BRANDY	9685	-1.103	0.089	-1.319	-0.689	Impaired
GREAT	5274	-1.158	0.209	-1.585	-0.706	Impaired
MESSALONSKEE	5280	-1.191	0.190	-1.597	-0.667	Impaired
UPPER MARY JO	243	-1.261	0.723	-1.440	-1.075	Impaired
LONG	1682	-1.820	0.140	-2.044	-1.606	Impaired

Maine Shallow

LAKE NAME	MIDAS	MEAN LDA SCORE	MEAN RIPSCORE	BCI (LOW)	BCI (HIGH)	BCI ASSESSMENT

UNNAMED	9486	2.514	0.594	1.726	3.006	Natural
GOULD	5474	2.507	0.511	1.486	3.308	Natural
CRANBERRY	3066	1.794	0.444	0.909	2.108	Natural
FOURTH ST JOHN	2416	1.758	0.613	0.799	2.174	Natural
MCLEAN	1550	1.365	0.423	1.023	1.816	Natural
SKITACOOK	1730	1.310	0.375	0.374	1.843	Natural
8TH DEBSCONEAG	608	1.129	0.711	0.627	1.850	Natural
SECOND BUTTERMILK	836	0.994	0.770	0.508	1.680	Natural
MYRICK	4416	0.975	0.640	0.697	1.325	Natural
LITTLE WATCHIC	3398	0.965	0.624	0.360	1.668	Natural
HUDSON	2724	0.847	0.689	0.402	1.249	Natural
LOWER MIDDLE BR	4494	0.754	0.647	0.061	1.353	Natural
ROCKY	2018	0.752	0.816	0.532	1.261	Natural
PERLEY	3140	0.732	0.475	0.260	1.383	Natural
GILMAN	4	0.690	0.293	-0.044	1.783	Intermediate
JAYBIRD	3178	0.678	0.638	0.338	1.235	Natural
FARRINGTON	3200	0.673	0.534	-0.011	1.174	Intermediate
NELSON	3610	0.567	0.554	0.160	1.328	Natural
KATAHDIN	2016	0.518	0.785	-0.266	1.254	Intermediate
MARTINS #2	2054	0.473	0.750	0.310	0.591	Natural
LITTLE	7871	0.442	0.387	-0.296	0.944	Intermediate
COFFEELOS	2712	0.420	0.684	0.132	0.895	Natural
PICKERAL	9687	0.397	0.724	-0.226	0.808	Intermediate
HAVENER	5718	0.339	0.578	-0.226	0.901	Intermediate
MARTINS #1	2052	0.281	0.692	-0.655	0.725	Intermediate
CUSHMAN	3224	0.118	0.321	-0.367	0.447	Intermediate
BLACK	351	0.001	0.300	-0.228	0.226	Intermediate
ATWOOD	4250	-0.074	0.653	-0.540	0.476	Intermediate
JIMMY	5244	-0.121	0.327	-0.563	0.496	Intermediate
TWIN #2	2028	-0.197	0.786	-0.497	0.131	Intermediate
CROWELL	5200	-0.197	0.471	-0.533	0.120	Intermediate
LONG	5444	-0.212	0.303	-0.568	0.057	Intermediate
PARKER	3388	-0.466	0.188	-0.850	0.005	Intermediate
WOOD	435	-0.482	0.478	-0.952	-0.055	Impaired
WOOD	3456	-0.500	0.146	-0.649	-0.299	Impaired

NORTH	5344	-0.520	0.680	-1.575	-0.082	Impaired
LOVEJOY	5664	-0.585	0.248	-1.721	0.369	Intermediate
POCASSET	3824	-0.623	0.249	-0.980	-0.297	Impaired
SEWALL	9943	-0.665	0.386	-0.931	-0.455	Impaired
ABRAMS	4444	-0.706	0.225	-1.274	0.793	Intermediate
HOBBS	4806	-0.858	0.325	-1.198	-0.481	Impaired
COCHNEWAGON	3814	-0.860	0.110	-1.249	-0.452	Impaired
EAST	5349	-0.865	0.333	-1.388	-0.315	Impaired
HORSESHOE	4788	-0.926	0.681	-1.376	-0.145	Impaired
ALAMOOSOOK	4336	-1.040	0.166	-1.263	-0.835	Impaired
DYER LONG	5386	-1.620	0.385	-2.126	-1.249	Impaired

Region – Deep Large

LAKE NAME	LAKE ID	STATE	MEAN LDA SCORE	MEAN RIPSCORE	BCI (LOW)	BCI (HIGH)	BCI ASSESSMENT
MAIDSTONE LAKE	NLA06608-9999	VT	2.455	0.453	2.090	2.698	Natural
DEBOULLIE	1512	ME	2.131	0.638	1.729	2.368	Natural
TOGUS	9931	ME	1.795	0.245	1.098	2.454	Natural
PARKER	5186	ME	1.784	0.522	1.480	2.156	Natural
WADLEIGH	572	ME	1.718	0.687	1.222	2.192	Natural
NONE	NLA17_CT-10010	CT	1.630	0.575	1.023	2.246	Natural
TORSEY	5307	ME	1.625	0.316	1.346	1.856	Natural
ECHO	5814	ME	1.470	0.236	1.153	1.666	Natural
LITTLE BIG WOOD POND	NLA17_ME-10003	ME	1.096	0.648	0.802	1.730	Natural
MARANACOOK	5312	ME	1.061	0.579	0.463	2.043	Natural
PHILLIPS	4300	ME	1.052	0.412	0.765	1.382	Natural
NEQUASSET	5222	ME	0.744	0.447	0.522	1.035	Natural
SEBEC	848	ME	0.716	0.219	0.134	1.203	Natural
GREEN RIVER	GREEN RIVER	VT	0.621	0.345	-0.278	1.441	Intermediate
PORTER	12	ME	0.611	0.261	0.298	0.828	Natural

SEBASTICOOK LAKE	NLA12_ME-109	ME	0.599	0.223	0.137	1.001	Natural
FLYING	5182	ME	0.509	0.192	0.078	0.941	Natural
WILSON - FR	3682	ME	0.495	0.453	0.043	0.807	Natural
WILLOUGHBY	WILLOUGHBY	VT	0.487	0.176	-0.016	1.213	Intermediate
UPPER NARROWS	98	ME	0.481	0.265	-0.077	0.903	Intermediate
SALEM	SALEM	VT	0.467	0.244	-0.575	0.936	Intermediate
CHAIN OF PONDS	NLA12_ME-103	ME	0.462	0.478	0.142	1.059	Natural
CHAMBERLAIN	2882	ME	0.418	0.598	0.113	0.592	Natural
SALMON	5352	ME	0.370	0.336	-0.263	0.681	Intermediate
CHINA W	54482	ME	0.358	0.671	0.122	0.552	Natural
NAHMAKANTA	698	ME	0.344	0.698	0.067	0.600	Natural
TUNK	4434	ME	0.340	0.582	0.012	0.563	Natural
VARNUM	3680	ME	0.216	0.417	-0.097	0.576	Intermediate
CHINA E	54481	ME	0.119	0.380	-0.313	0.345	Intermediate
MESSALONSKEE	5280	ME	0.098	0.485	-0.295	0.420	Intermediate
CANOBIE	Canobie	NH	-0.027	0.134	-0.339	0.275	Intermediate
GARDNER	1528	ME	-0.036	0.579	-0.371	0.258	Intermediate
WASSOOKEAG	227	ME	-0.064	0.184	-0.389	0.335	Intermediate
PENNESSEEWASSEE	3434	ME	-0.064	0.162	-0.424	0.189	Intermediate
TRICKEY	3382	ME	-0.065	0.177	-0.340	0.198	Intermediate
PLEASANT LAKE	NLA17_ME-10006	ME	-0.115	0.571	-0.514	0.280	Intermediate
GLEN	GLEN	VT	-0.121	0.503	-0.548	0.321	Intermediate
BRANDY	9685	ME	-0.126	0.255	-0.274	0.024	Intermediate
WOODBURY	5240	ME	-0.134	0.200	-0.469	0.165	Intermediate
CASPIAN LAKE	NLA06608-0369	VT	-0.142	0.186	-0.293	0.029	Intermediate
PENCON	PENCON	NH	-0.177	0.504	-0.562	0.061	Intermediate
WILSON - KN	3832	ME	-0.287	0.340	-0.586	0.120	Intermediate
HOLLAND	HOLLAND	VT	-0.368	0.326	-0.837	0.236	Intermediate
WEBBER	5408	ME	-0.396	0.295	-0.593	-0.192	Impaired
SUNSET (BENSON)	SUNSET (BENSON)	VT	-0.448	0.197	-1.014	0.022	Intermediate
SQUARE	3916	ME	-0.462	0.248	-0.784	-0.135	Impaired
ANDROSCOGGIN LAKE	3836	ME	-0.566	0.353	-1.291	0.010	Intermediate

TRIPP	3758	ME	-0.574	0.252	-1.075	-0.023	Impaired
WEBB	3672	ME	-0.584	0.228	-0.956	-0.094	Impaired
GREAT	5274	ME	-0.639	0.197	-1.020	-0.108	Impaired
GROTON	GROTON	VT	-0.681	0.168	-1.440	-0.080	Impaired
MOLLYS FALLS	MOLLYS FALLS	VT	-0.739	0.297	-1.334	-0.396	Impaired
UPPER MARY JO	243	ME	-0.792	0.635	-1.163	-0.486	Impaired
SHADOW (GLOVER)	SHADOW (GLOVER)	VT	-0.794	0.097	-1.102	-0.466	Impaired
DUCK LAKE	NLA06608-NELP-0253	ME	-0.796	0.553	-1.091	-0.493	Impaired
CHITTENDEN	CHITTENDE N	VT	-0.888	0.211	-1.229	-0.602	Impaired
COBBOSSECONTEE	5236	ME	-0.944	0.161	-1.505	-0.558	Impaired
ISLAND POND	NLA06608-0038	VT	-1.008	0.118	-1.241	-0.827	Impaired
GREAT AVERILL	GREAT AVERILL	VT	-1.078	0.154	-1.537	-0.764	Impaired
COBBETTS	Cobbetts	NH	-1.148	0.103	-1.459	-0.828	Impaired
LAKE WARAMAUG	NLA06608-0037	CT	-1.211	0.095	-1.818	-0.594	Impaired
LONG LAKE	NLA17_ME-10025	ME	-1.217	0.556	-1.526	-0.940	Impaired
SWAMP POND	NLA17_NH-10003	NH	-1.359	0.514	-1.651	-1.037	Impaired
ARLSAL	ARLSAL	NH	-1.826	0.077	-2.114	-1.495	Impaired

Region – Deep Small

LAKE NAME	LAKE ID	STATE	MEAN LDA SCORE	MEAN RIPSCORE	BCI (LOW)	BCI (HIGH)	BCI ASSESSMENT
LONG POND	NLA17_VT-10018	VT	2.324	0.536	1.975	2.596	Natural
RIGA LAKE	NLA17_CT-10002	CT	1.648	0.558	1.516	1.766	Natural
ELLIS	4086	ME	1.603	0.629	1.328	1.944	Natural
HOWARD POND	NLA12_ME-106	ME	1.523	0.086	1.370	1.706	Natural
BLACK	1506	ME	1.488	0.604	1.347	1.595	Natural
THREECORNERED	5424	ME	1.356	0.339	1.109	1.541	Natural

SPRING LAKE - VT	NLA06608-4252	VT	1.272	0.249	0.975	1.457	Natural
LITTLE JIM	5090	ME	1.181	0.515	0.975	1.425	Natural
BIG REED POND	NLA06608-EMAP:ME254L	ME	1.095	0.493	0.847	1.302	Natural
PUSHINEER POND	NLA06608-NELP-2155	ME	1.088	0.420	0.841	1.273	Natural
BIG REED	2842	ME	1.030	0.496	0.909	1.154	Natural
PEAKED MOUNTAIN POND	NLA17_ME-10002	ME	0.975	0.466	0.785	1.219	Natural
WALLUM LAKE	NLA06608-0754	MA	0.726	0.318	0.552	0.897	Natural
GRASS	104	ME	0.725	0.629	0.455	0.948	Natural
SILVER LAKE	NLA17_VT-10002	VT	0.695	0.560	0.577	0.791	Natural
FLYING POND	NLA12_ME-114	ME	0.664	0.208	0.353	1.026	Natural
JAMIES	5302	ME	0.660	0.419	0.248	0.939	Natural
BUBBLE	4452	ME	0.642	0.474	0.520	0.783	Natural
TULLY LAKE	NLA17_MA-10004	MA	0.576	0.577	0.288	0.902	Natural
B POND	NLA17_ME-HP001	ME	0.459	0.530	0.237	0.607	Natural
WEST HILL POND	NLA17_CT-10003	CT	0.451	0.585	0.151	0.738	Natural
SPONEL	SPONEL	NH	0.332	0.485	0.105	0.521	Natural
THREE CORNER POND	NLA17_ME-10009	ME	0.287	0.567	0.094	0.445	Natural
CENTER	CENTER	VT	0.235	0.309	-0.056	0.532	Intermediate
WILLARD	Willard	NH	0.177	0.453	0.050	0.370	Natural
HALUSWH	HALUSWH	NH	0.169	0.561	-0.034	0.325	Intermediate
LOWER SO BR	4222	ME	0.138	0.571	-0.005	0.277	Intermediate
FEMALE POND	NLA06608-EMAP:ME011L	ME	0.130	0.370	-0.367	0.378	Intermediate
ECHO (HUBDTN)	ECHO (HUBDTN)	VT	0.050	0.353	-0.259	0.397	Intermediate
HILLS POND	NLA17_NH-10066	NH	0.039	0.558	-0.293	0.247	Intermediate
MORRIS RESERVOIR	NLA17_CT-10004	CT	0.005	0.538	-0.272	0.314	Intermediate
LOVELLS POND	NLA17_MA-10019	MA	-0.011	0.592	-0.321	0.140	Intermediate

PEQUAWKET POND	NLA17_NH-10006	NH	-0.052	0.494	-0.326	0.166	Intermediate
LOWER HADLOCK	4610	ME	-0.077	0.297	-0.294	0.158	Intermediate
RUSSELL	RUSSELL	NH	-0.122	0.449	-0.240	0.062	Intermediate
LOOFRE	LOOFRE	NH	-0.189	0.206	-0.413	0.003	Intermediate
WOODWARD	WOODWARD	VT	-0.228	0.167	-0.593	0.063	Intermediate
CRYEAT	CRYEAT	NH	-0.274	0.115	-0.550	-0.022	Impaired
SPRING LAKE - ME	NLA06608-ELS:1E1-052	ME	-0.318	0.432	-0.692	-0.168	Impaired
BAKLWEN	BAKLWEN	NH	-0.345	0.235	-0.700	0.060	Intermediate
LAKE RESCUE	NLA17_VT-10016	VT	-0.383	0.519	-0.614	-0.167	Impaired
BEEBE LAKE	NLA06608-0997	VT	-0.442	0.126	-0.721	-0.133	Impaired
WHITE LAKE	NLA17_NH-10891	NH	-0.470	0.534	-0.653	-0.181	Impaired
ROOD	ROOD	VT	-0.470	0.371	-0.622	-0.386	Impaired
GORTON POND	NLA06608-2354	RI	-0.627	0.192	-0.817	-0.335	Impaired
NIPBAR	NIPBAR	NH	-0.639	0.157	-0.888	-0.283	Impaired
BRETTUNS	3608	ME	-0.662	0.108	-0.896	-0.409	Impaired
HALHAN	HALHAN	NH	-0.787	0.269	-1.183	-0.433	Impaired
THIRD CONNECTICUT LAKE	NLA17_NH-10999	NH	-0.924	0.534	-1.177	-0.489	Impaired
OTTKEE	OTTKEE	NH	-1.087	0.371	-1.240	-0.893	Impaired
NORWICH POND	NLA12_MA-110	MA	-1.098	0.147	-1.310	-0.978	Impaired
GILUNY	GILUNY	NH	-1.160	0.554	-1.428	-0.886	Impaired
ECHFRN	ECHFRN	NH	-1.176	0.119	-1.348	-0.940	Impaired
YAWGOO POND	NLA06608-2162	RI	-1.264	0.379	-1.421	-1.030	Impaired
LAKE PARKER	NLA12_VT-102	VT	-1.294	0.130	-1.599	-0.908	Impaired
BEACH POND	NLA06608-3890	RI	-1.355	0.179	-1.518	-1.123	Impaired
LITTLE AVERILL POND	NLA17_VT-10003	VT	-1.585	0.636	-1.689	-1.389	Impaired
BEADER	BEADER	NH	-2.048	0.064	-2.189	-1.915	Impaired
LONG POND	NLA17_VT-10018	VT	2.324	0.536	1.975	2.596	Natural
RIGA LAKE	NLA17_CT-10002	CT	1.648	0.558	1.516	1.766	Natural

ELLIS	4086	ME	1.603	0.629	1.328	1.944	Natural
HOWARD POND	NLA12_ME-106	ME	1.523	0.086	1.370	1.706	Natural
BLACK	1506	ME	1.488	0.604	1.347	1.595	Natural
THREECORNERED	5424	ME	1.356	0.339	1.109	1.541	Natural
SPRING LAKE - VT	NLA06608-4252	VT	1.272	0.249	0.975	1.457	Natural
LITTLE JIM	5090	ME	1.181	0.515	0.975	1.425	Natural
BIG REED POND	NLA06608-EMAP:ME254L	ME	1.095	0.493	0.847	1.302	Natural
PUSHINEER POND	NLA06608-NELP-2155	ME	1.088	0.420	0.841	1.273	Natural

Region – Shallow Large

LAKE NAME	LAKE ID	STATE	MEAN LDA SCORE	MEAN RIPSCORE	BCI (LOW)	BCI (HIGH)	BCI ASSESSMENT
BILLINGS LAKE	NLA17_CT-10023	CT	2.011	0.861	1.641	2.384	Natural
DYER LONG	5386	ME	1.666	0.592	1.501	1.893	Natural
ALAMOOSOOK	4336	ME	1.664	0.487	1.288	1.945	Natural
SEWALL	9943	ME	1.568	0.578	1.344	1.790	Natural
EAST	5349	ME	1.567	0.548	1.168	1.927	Natural
HORSESHOE	4788	ME	1.482	0.588	1.282	1.768	Natural
GRASSY POND	NLA06608-0418	NH	1.471	0.634	1.173	1.785	Natural
HOBBS	4806	ME	1.339	0.499	1.127	1.565	Natural
HAVENER POND	NLA17_ME-10018	ME	1.238	0.639	0.767	1.996	Natural
CUSHMAN	3224	ME	1.220	0.306	1.041	1.452	Natural
ABRAMS	4444	ME	1.129	0.536	0.807	1.443	Natural
MIDDLE CHAIN POND	NLA06608-ELS:1E1-096	ME	1.090	0.673	0.669	1.464	Natural
HALLS POND	NLA17_CT-10005	CT	1.052	0.747	0.850	1.315	Natural
PICKERAL	9687	ME	0.995	0.617	0.761	1.281	Natural
DERBY LAKE	NLA06608-0294	VT	0.995	0.232	0.800	1.187	Natural

PARKER	3388	ME	0.972	0.449	0.780	1.304	Natural
TROUT	3212	ME	0.936	0.615	0.476	1.257	Natural
POCASSET	3824	ME	0.923	0.248	0.575	1.305	Natural
LONG POND	NLA17_NH-10917	NH	0.920	0.534	0.671	1.191	Natural
LOWER MIDDLE BRANCH POND	NLA17_ME-10001	ME	0.903	0.559	0.523	1.143	Natural
LONG	5444	ME	0.826	0.289	0.612	1.283	Natural
LOVEJOY	5664	ME	0.824	0.248	0.182	1.207	Natural
LILY (VERNON)	LILY (VERNON)	VT	0.813	0.260	0.413	1.268	Natural
LITTLE LAKE	NLA12_ME-122	ME	0.764	0.497	0.556	0.932	Natural
WOOD	3456	ME	0.756	0.404	0.464	1.056	Natural
FARRINGTON	3200	ME	0.718	0.468	0.454	0.980	Natural
KNOWLTON POND	NLA06608-0946	CT	0.716	0.559	0.521	1.076	Natural
PERLEY	3140	ME	0.662	0.422	0.447	0.975	Natural
JIMMY	5244	ME	0.655	0.315	0.340	0.926	Natural
COFFEELOS	2712	ME	0.619	0.585	0.476	0.753	Natural
COCHNEWAGON	3814	ME	0.594	0.398	0.248	1.131	Natural
SKITACOOK LAKE	NLA17_ME-10029	ME	0.594	0.580	0.134	1.014	Natural
BUCKLEY DUNTON LAKE	NLA17_MA-10003	MA	0.568	0.745	0.027	1.011	Natural
HUDSON POND	NLA17_ME-10004	ME	0.541	0.620	0.137	0.968	Natural
CROWELL	5200	ME	0.532	0.429	0.309	0.860	Natural
BLACK LAKE	NLA17_ME-10021	ME	0.511	0.619	0.294	0.604	Natural
CROSBY POND	NLA17_ME-HP002	ME	0.499	0.599	0.151	0.778	Natural
SOPER POND	NLA12_ME-104	ME	0.479	0.552	-0.007	0.850	Intermediate
WALLINGFORD	WALLINGFO RD	VT	0.462	0.575	0.198	0.784	Natural
FOSTERS	FOSTERS	VT	0.446	0.390	-0.105	0.976	Intermediate
BAGWSR	BAGWSR	NH	0.374	0.545	0.083	0.608	Natural
PACHAUG POND	NLA06608-0242	CT	0.355	0.301	-0.254	1.051	Intermediate
CHANDLER	CHANDLER	VT	0.340	0.378	0.058	0.634	Natural
THOMPSONS	THOMPSON S	VT	0.304	0.440	-0.085	0.584	Intermediate

GILLETT	GILLETT	VT	0.285	0.552	-0.728	0.784	Intermediate
HIGHLAND LAKE	NLA12_NH-104	NH	0.282	0.253	-0.170	0.676	Intermediate
LITTLE HOSMER	LITTLE HOSMER	VT	0.272	0.265	-0.078	0.543	Intermediate
NORTH	5344	ME	0.264	0.588	-0.199	0.489	Intermediate
SCITUATE RESERVOIR	NLA12_RI-112	RI	0.259	0.520	-0.093	0.615	Intermediate
HLFGRA	HLFGRA	NH	0.243	0.263	0.034	0.542	Natural
UPPER KIMBALL POND	NLA17_NH-10067	NH	0.230	0.603	0.001	0.463	Natural
ARMINGTON LAKE	NLA06608-0550	NH	0.197	0.279	-0.339	0.699	Intermediate
KATAHDIN	2016	ME	0.184	0.663	-0.248	0.595	Intermediate
SCHOOL HOUSE POND	NLA06608-3846	RI	0.180	0.708	-0.110	0.471	Intermediate
KEECH POND	NLA17_RI-10001	RI	0.085	0.588	-0.517	0.487	Intermediate
BELLEVILLE POND	NLA12_RI-102	RI	-0.013	0.388	-0.269	0.271	Intermediate
SMITHS MILLPOND	NLA17_ME-10042	ME	-0.019	0.433	-0.435	0.361	Intermediate
LONG (SHEFLD)	LONG (SHEFLD)	VT	-0.044	0.484	-0.496	0.390	Intermediate
NORTON RESERVOIR	NLA17_MA-10002	MA	-0.074	0.749	-0.461	0.346	Intermediate
FOURTH SAINT JOHN POND	NLA17_ME-10040	ME	-0.154	0.547	-0.410	0.158	Intermediate
SOUTH AMERICA	SOUTH AMERICA	VT	-0.197	0.534	-0.631	0.072	Intermediate
WONONPAKOOK LAKE	NLA06608-1125	CT	-0.199	0.372	-0.454	0.085	Intermediate
STUMP POND	NLA06608-1586	RI	-0.230	0.351	-0.717	0.282	Intermediate
MESSERSCHMIDT POND	NLA12_CT-107	CT	-0.358	0.568	-0.711	0.003	Intermediate
BABROX	BABROX	NH	-0.366	0.581	-0.683	-0.054	Impaired
GROTON RESERVOIR	NLA06608-0326	CT	-0.374	0.188	-0.597	-0.133	Impaired
WRIGHTSVILLE	WRIGHTSVILLE	VT	-0.397	0.290	-0.731	-0.170	Impaired
LAKHOK	LAKHOK	NH	-0.426	0.533	-0.811	-0.080	Impaired
SPAULDING POND	NLA12_NH-111	NH	-0.427	0.248	-0.889	0.196	Intermediate
ROUEAT	ROUEAT	NH	-0.447	0.416	-0.870	0.019	Intermediate

GILSAN	GILSAN	NH	-0.465	0.513	-0.710	-0.190	Impaired
MILLER	MILLER	VT	-0.477	0.282	-0.891	0.037	Intermediate
FLAT RIVER RESERVOIR	NLA12_RI-108	RI	-0.477	0.296	-0.984	-0.110	Impaired
MEEMAR	MEEMAR	NH	-0.504	0.446	-0.860	-0.145	Impaired
FOURTH MACHIAS LAKE	NLA06608-EMAP:ME263L	ME	-0.526	0.460	-0.893	-0.010	Impaired
PINE	PINE	VT	-0.539	0.382	-0.936	-0.305	Impaired
POTBRL	POTBRL	NH	-0.539	0.387	-0.857	-0.113	Impaired
HALFWAY POND	NLA06608-ELS:1D1-035	MA	-0.544	0.490	-0.721	-0.329	Impaired
SIP POND	NLA17_NH-10004	NH	-0.554	0.531	-0.949	-0.229	Impaired
TOMHEGAN POND	NLA06608-ELS:1E2-027	ME	-0.669	0.654	-1.044	-0.228	Impaired
ABENAKI	ABENAKI	VT	-0.672	0.395	-0.946	-0.429	Impaired
LONG MEADOW POND	NLA12_CT-108	CT	-0.688	0.182	-1.075	-0.384	Impaired
SLACK RESERVOIR	NLA06608-1906	RI	-0.690	0.199	-1.173	-0.014	Impaired
GILMAN POND	NLA17_ME-10020	ME	-0.725	0.441	-1.026	-0.306	Impaired
BOWDISH RESERVOIR	NLA06608-4413	RI	-0.794	0.190	-1.083	-0.340	Impaired
MILMLD	MILMLD	NH	-0.805	0.501	-1.172	-0.390	Impaired
SPECRO	SPECRO	NH	-0.819	0.307	-1.157	-0.513	Impaired
NINEVAH	NINEVAH	VT	-0.848	0.254	-1.234	-0.230	Impaired
GREKIN	GREKIN	NH	-0.870	0.434	-1.117	-0.543	Impaired
ROSELAND LAKE	NLA12_CT-104	CT	-0.889	0.309	-1.238	-0.533	Impaired
CRAHEN	CRAHEN	NH	-0.928	0.356	-1.109	-0.779	Impaired
BACK LAKE	NLA06608-0662	NH	-0.954	0.155	-1.334	-0.224	Impaired
HINKLEY'S POND	NLA06608-0198	MA	-0.961	0.173	-1.231	-0.703	Impaired
BRINDLE POND	NLA17_NH-10052	NH	-0.968	0.562	-1.154	-0.705	Impaired
BR6AWEN	BR6AWEN	NH	-0.981	0.244	-1.208	-0.650	Impaired
CHAPMAN POND	NLA06608-2566	RI	-1.066	0.258	-1.380	-0.604	Impaired
PROEFF	PROEFF	NH	-1.220	0.175	-1.690	-0.781	Impaired

SILHLS	SILHLS	NH	-1.337	0.094	-1.731	-0.791	Impaired
BABAMH	BABAMH	NH	-1.376	0.157	-1.544	-1.164	Impaired
COLUMBIA LAKE	NLA12_CT-115	CT	-1.517	0.038	-1.815	-1.163	Impaired
LONG POND	NLA12_MA-104	MA	-1.530	0.187	-1.787	-1.262	Impaired
CEDAR	CEDAR	VT	-1.557	0.179	-1.743	-1.151	Impaired
LITTLE (WELLS)	LITTLE (WELLS)	VT	-2.140	0.198	-2.569	-1.762	Impaired
BILLINGS LAKE	NLA17_CT-10023	CT	2.011	0.861	1.641	2.384	Natural

Region – Shallow Small

LAKE NAME	LAKE ID	STATE	MEAN LDA SCORE	MEAN RIPSCORE	BCI (LOW)	BCI (HIGH)	BCI ASSESSMENT
MUD (WESTMR)-W	MUD (WESTMR)-W	VT	3.108	0.569	2.818	3.437	Natural
NONE	NLA17_CT-10019	CT	2.689	0.648	1.809	3.160	Natural
PEACE DALE RESERVOIR	NLA17_RI-10010	RI	2.292	0.690	1.718	2.759	Natural
TROUT POND	NLA06608-NH250L	NH	2.281	0.586	1.972	2.618	Natural
CROOKED POND	NLA06608-ELS:1C2-032	NH	2.245	0.806	1.898	2.557	Natural
OTTER POND	NLA06608-ELS:1E3-002	ME	2.078	0.735	1.561	2.407	Natural
NONE	NLA17_VT-10007	VT	2.063	0.633	1.752	2.314	Natural
EIGHTH DEBSCONEAG POND	NLA17_ME-10007	ME	2.033	0.636	1.707	2.462	Natural
LAKE WOOD	NLA06608-ACAD_LAKES_0435	ME	1.986	0.666	1.631	2.352	Natural
HORSESHOE LAKE	NLA12_ME-126	ME	1.951	0.684	1.762	2.361	Natural
WOOD	435	ME	1.907	0.434	1.642	2.146	Natural

RODERIQUE POND	NLA06608-ELS:1C3-003	ME	1.881	0.733	1.754	2.154	Natural
WEBSTER POND	NLA12_ME-129	ME	1.590	0.544	1.402	1.753	Natural
BLAKE (SUTTON)	BLAKE (SUTTON)	VT	1.409	0.469	1.208	1.607	Natural
GOULD POND	NLA17_ME-10043	ME	1.398	0.550	1.014	1.890	Natural
TEN THOUSAND ACRE POND	NLA12_ME-116	ME	1.354	0.793	1.134	1.528	Natural
HALFMILE POND	NLA06608-ELS:1E1-128	ME	1.329	0.694	1.015	1.746	Natural
JAYBIRD	3178	ME	1.303	0.564	1.008	1.851	Natural
NONE	NLA17_MA-10018	MA	1.175	0.803	0.984	1.323	Natural
BEEBE (SUNDLD)	BEEBE (SUNDLD)	VT	1.142	0.455	0.881	1.466	Natural
BASSETT POND	NLA17_MA-10022	MA	1.049	0.784	0.873	1.163	Natural
WALLACE POND	NLA06608-0802	MA	1.017	0.496	0.400	1.571	Natural
SECOND BUTTERMILK POND	NLA17_ME-10010	ME	0.904	0.699	0.618	1.149	Natural
BIG MUDDY	BIG MUDDY	VT	0.848	0.574	0.559	1.164	Natural
HALLS POND	NLA17_CT-10021	CT	0.846	0.625	0.172	1.435	Natural
NELSON POND	NLA17_ME-10030	ME	0.801	0.657	0.557	0.994	Natural
LITTLE (ELMORE)	LITTLE (ELMORE)	VT	0.747	0.594	0.486	1.023	Natural
MCLEAN LAKE	NLA17_ME-10039	ME	0.552	0.576	0.211	0.797	Natural
LITTLE (WINHLL)	LITTLE (WINHLL)	VT	0.549	0.585	0.305	0.814	Natural
NONE	NLA17_MA-HP001	MA	0.460	0.585	0.155	0.747	Natural
ROCKY	2018	ME	0.422	0.709	0.148	0.855	Natural
UPPER POND	NLA06608-NELP-3355	ME	0.411	0.617	0.187	0.617	Natural
BISSONNETTE POND	NLA17_CT-10001	CT	0.406	0.716	0.205	0.624	Natural
NONE	NLA17_MA-10017	MA	0.395	0.626	0.052	0.689	Natural
MILE POND	NLA17_VT-10005	VT	0.350	0.553	0.087	0.712	Natural
ORWELL;	ORWELL;	VT	0.347	0.601	-0.006	0.638	Intermediate

BARBER POND	NLA17_RI-10006	RI	0.336	0.729	0.198	0.473	Natural
PATTACONK RESERVOIR	NLA17_CT-HP001	CT	0.314	0.621	-0.041	0.666	Intermediate
MUD (LEICTR)	MUD (LEICTR)	VT	0.223	0.486	-0.157	0.672	Intermediate
UNION POND	NLA06608-0582	CT	0.216	0.085	-0.193	0.642	Intermediate
ATWOOD POND	NLA17_ME-10023	ME	0.208	0.592	-0.030	0.425	Intermediate
LITTLE GREENOUGH POND	NLA17_NH-10002	NH	0.179	0.697	-0.170	0.443	Intermediate
GRIST MILLPOND	NLA06608-0546	MA	0.151	0.471	-0.157	0.511	Intermediate
MUD POND	NLA12_ME-113	ME	0.066	0.590	-0.042	0.239	Intermediate
BLACKAMORE POND	NLA17_RI-10011	RI	0.043	0.630	-0.196	0.281	Intermediate
ROUGLM	ROUGLM	NH	0.031	0.447	-0.224	0.275	Intermediate
PLAINFIELD POND	NLA06608-0674	MA	0.024	0.410	-0.410	0.444	Intermediate
FERRISBURGH;	FERRISBURGH;	VT	0.007	0.389	-0.218	0.351	Intermediate
POSNEGANSET POND	NLA17_RI-10019	RI	0.003	0.608	-0.427	0.511	Intermediate
ADDER POND	NLA06608-0050	NH	-0.010	0.721	-0.352	0.364	Intermediate
NONE	NLA17_NH-10001	NH	-0.018	0.623	-0.475	0.312	Intermediate
DEEP HOLE POND	NLA12_NH-107	NH	-0.052	0.630	-0.276	0.207	Intermediate
REYLTL	REYLTL	NH	-0.064	0.572	-0.327	0.206	Intermediate
WAUREGAN RESERVOIR	NLA12_CT-118	CT	-0.088	0.563	-0.557	0.177	Intermediate
SLATERSVILLE RESERVOIRS	NLA17_RI-10002	RI	-0.104	0.679	-0.302	0.167	Intermediate
TURTLEHEAD POND	NLA06608-0806	VT	-0.110	0.490	-0.615	0.170	Intermediate
MITCHELL	MITCHELL	VT	-0.145	0.479	-0.660	0.468	Intermediate
NONE	NLA17_VT-10019	VT	-0.221	0.526	-0.524	0.251	Intermediate
PERLEY POND	NLA17_ME-HP003	ME	-0.223	0.586	-0.523	0.245	Intermediate
KIASWH	KIASWH	NH	-0.237	0.711	-0.521	-0.064	Impaired
RANDALL POND	NLA17_RI-10008	RI	-0.258	0.609	-0.594	0.088	Intermediate

KILBURN POND	NLA17_NH-10530	NH	-0.307	0.622	-0.871	0.074	Intermediate
VOYDATCH POND	NLA06608-0226	NH	-0.330	0.241	-0.711	0.031	Intermediate
LOOLFRE	LOOLFRE	NH	-0.334	0.190	-0.642	-0.092	Impaired
HAYNES RESERVOIR	NLA06608-0354	MA	-0.337	0.677	-0.791	-0.014	Impaired
CRANBERRY MEADOW	CRANBERRY MEADOW	VT	-0.369	0.298	-0.710	-0.132	Impaired
BEARDSLEY POND	NLA06608-0805	CT	-0.371	0.266	-0.904	0.382	Intermediate
LAKE KENOSIA	NLA06608-0293	CT	-0.381	0.113	-0.870	-0.026	Impaired
SHAW POND	NLA12_MA-106	MA	-0.381	0.498	-0.554	-0.247	Impaired
UNNAMED	NLA12_MA-108	MA	-0.384	0.298	-0.650	-0.159	Impaired
MARLBORO-431;	MARLBORO -431;	VT	-0.395	0.351	-0.613	-0.231	Impaired
SMIWAS	SMIWAS	NH	-0.407	0.449	-0.812	0.111	Intermediate
KEYSER;	KEYSER;	VT	-0.413	0.197	-0.814	0.113	Intermediate
MARTINS #1	2052	ME	-0.419	0.614	-0.781	-0.122	Impaired
MARTINS #2	2054	ME	-0.437	0.659	-0.556	-0.202	Impaired
CRESCENT	CRESCENT	VT	-0.437	0.311	-0.783	-0.220	Impaired
LONEAT	LONEAT	NH	-0.440	0.311	-0.857	0.015	Intermediate
LITTLE POND	NLA17_ME-10014	ME	-0.460	0.556	-0.884	-0.104	Impaired
LITTLE WATCHIC POND	NLA17_ME-10011	ME	-0.478	0.650	-0.757	-0.087	Impaired
WALKER (COVNTY)	WALKER (COVNTY)	VT	-0.485	0.227	-0.863	0.006	Intermediate
SHIPPEE POND	NLA12_VT-109	VT	-0.491	0.485	-0.755	-0.178	Impaired
NONE	NLA17_ME-10016	ME	-0.491	0.545	-0.738	-0.184	Impaired
VONDELL	VONDELL	VT	-0.542	0.380	-0.920	-0.173	Impaired
UNNAMED	NLA12_VT-115	VT	-0.599	0.382	-0.853	-0.343	Impaired
NOJOE;	NOJOE;	VT	-0.599	0.382	-0.855	-0.358	Impaired
BEECHER	BEECHER	VT	-0.635	0.414	-0.923	-0.291	Impaired
KIMCNT	KIMCNT	NH	-0.656	0.419	-0.839	-0.525	Impaired
MYRICK LAKE	NLA17_ME-10038	ME	-0.658	0.605	-0.808	-0.521	Impaired

BEARDSLEY POND	NLA12_CT-105	CT	-0.667	0.384	-0.897	-0.362	Impaired
NONE	NLA17_NH-10059	NH	-0.670	0.658	-0.773	-0.426	Impaired
LIMCLM	LIMCLM	NH	-0.671	0.399	-1.192	-0.224	Impaired
DRY POND	NLA17_MA-10007	MA	-0.689	0.667	-0.947	-0.383	Impaired
MILL POND	NLA12_NH-112	NH	-0.704	0.224	-1.022	-0.467	Impaired
KENNOT	KENNOT	NH	-0.722	0.641	-1.085	-0.261	Impaired
DUCK POND	NLA12_NH-115	NH	-0.731	0.558	-1.028	-0.501	Impaired
SANTRO	SANTRO	NH	-0.732	0.362	-1.245	-0.385	Impaired
RICHMOND POND	NLA17_VT-10004	VT	-0.778	0.593	-0.869	-0.615	Impaired
KNAPP BROOK #1	KNAPP BROOK #1	VT	-0.807	0.296	-1.089	-0.558	Impaired
MUD (PEACHM)	MUD (PEACHM)	VT	-0.816	0.538	-1.056	-0.474	Impaired
CRANBERRY POND	NLA17_ME-10019	ME	-0.829	0.435	-1.060	-0.683	Impaired
ELLSWORTH POND	NLA17_NH-10063	NH	-0.835	0.618	-1.128	-0.609	Impaired
BASCHM	BASCHM	NH	-0.854	0.576	-1.060	-0.622	Impaired
NATDIX	NATDIX	NH	-0.856	0.455	-1.156	-0.353	Impaired
OTTARNIC POND	NLA17_NH-10058	NH	-0.922	0.530	-1.169	-0.464	Impaired
TWIN #2	2028	ME	-0.953	0.687	-1.111	-0.757	Impaired
EPPELEY POND	NLA12_RI-110	RI	-0.953	0.666	-1.256	-0.771	Impaired
DOLCON	DOLCON	NH	-0.966	0.412	-1.082	-0.772	Impaired
SCHOFIELD	SCHOFIELD	VT	-0.974	0.437	-1.173	-0.720	Impaired
SPRUCE (WILMTN)	SPRUCE (WILMTN)	VT	-1.005	0.127	-1.524	-0.612	Impaired
TAYSAL	TAYSAL	NH	-1.022	0.216	-1.293	-0.605	Impaired
TUTTLE (HARDWK)	TUTTLE (HARDWK)	VT	-1.043	0.381	-1.285	-0.784	Impaired
BESSE BOG RESERVOIR	NLA12_MA-111	MA	-1.118	0.425	-1.459	-0.882	Impaired
STANNARD	STANNARD	VT	-1.144	0.417	-1.436	-0.687	Impaired
DICKS POND	NLA17_MA-10001	MA	-1.150	0.643	-1.365	-0.868	Impaired
CARRY POND	NLA06608-ELS:1E3-071	ME	-1.228	0.568	-1.431	-0.969	Impaired

LEVEL HILL ROAD FARM POND	NLA12_ME-134	ME	-1.360	0.146	-1.601	-0.828	Impaired
SPRDEE	SPRDEE	NH	-1.370	0.673	-1.482	-1.187	Impaired
GOOSE POND	NLA06608-ALPS-1218	ME	-1.423	0.505	-1.798	-1.080	Impaired
KETTLE BROOK RESERVOIR #1	NLA06608-1122	MA	-1.469	0.404	-1.792	-1.223	Impaired
NONE	NLA17_VT-10009	VT	-1.475	0.447	-1.634	-1.299	Impaired
RICHMOND	RICHMOND	VT	-1.476	0.407	-1.670	-1.357	Impaired
UNKNOWN (AVYGOR)	UNKNOWN (AVYGOR)	VT	-1.477	0.425	-1.657	-1.333	Impaired
ROUBAR	ROUBAR	NH	-1.566	0.408	-1.804	-1.286	Impaired
LONG POND	NLA06608-1222	MA	-1.698	0.113	-1.881	-1.435	Impaired
STRATTON SKI AREA;	STRATTON SKI AREA;	VT	-1.866	0.159	-2.157	-1.555	Impaired

BIOGRAPHY OF THE AUTHOR

Jeremy Deeds was born in Waterville, Maine on April 20, 1979. He was raised in Windsor, Maine and graduated from Erskine Academy in 1997. He attended the University of Maine at Farmington and graduated in 2001 with a Bachelor's degree in Environmental Science. He attended Kent State University and graduated with a Master's degree in Aquatic Ecology in 2003. Jeremy is currently an aquatic ecologist with the Lakes Assessment Section in the Maine Department of Environmental Protection. Jeremy is a candidate for the Doctor of Philosophy degree in Ecology and Environmental Science from the University of Maine in May 2022.