Inorganic carbon cycling in Lake Superior and responses to anthropogenic carbon dioxide

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Daniel Edward Sandborn

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Elizabeth Colquitt Minor

October, 2024

Acknowledgements

None of this could have happened without the support and encouragement of many people, to whom I owe a debt of gratitude. It takes a village:

I wish to thank my advisor, Liz Minor, whose tireless support and insightful curiosity sets an example of mentorship I will cherish and aspire to. The Minor Lab, Large Lakes Observatory, WRS graduate students, Duluth water science community, and my fellow denizens of RLB 15 were the bedrock supporting my time in Duluth.

Sandy Brovold, Julie Agnich, Erik Hendrickson, Sarah Grosshuesch, Kim Lundy, Michael Jacob, and the Captain and crew of the *RV Blue Heron* deserve special thanks for sharing their knowledge and occasionally stretching their patience. Thanks as well to Luke Busta, Melissa Maurer-Jones, Courtney Larson, Craig Hill, and Jay Austin for being so generous with advice and assistance. Quinn, Peter, Audrey, and John, you taught me so much about supporting others. Nolan, David, Levi, Emily, and Sarah, you helped make Duluth home.

Emily Bockmon and the Bockmon lab inspired me to dream ocean-sized dreams during my time at Cal Poly, where many others also helped nurture my love of chemistry, including Kristen Meisenheimer, Phil Bailey, Greg Scott, Michael Heying, Matthew Zoerb, and Corinne Lehr. Aditi Sengupta helped me see the big picture of what a career in science could be. I am immensely appreciative of the generous support of Bill and Linda Frost, the Frost Fund, and the Frost Scholars community, who played a foundational role in enabling my science then and now.

The support and assistance of the University of Minnesota Office of the Vice President for Research, National Oceanic and Atmospheric Administration, Cooperative Institute for Great Lakes Research, National Park Service, Minnesota SeaGrant, Minnesota Pollution Control Agency, the Association for the Sciences of Limnology and Oceanography, the International Association for Great Lakes Research, and the Ocean Carbon and Biogeochemistry Program is gratefully acknowledged.

Joe Albrecht, Sierra Bedwell, Lacey Chappel, and Max Morgan: our friendship is one of the most important things in the world to me, and I would not have accomplished this without you. To my parents, Michelle and Steve, my sisters Theresa and Melissa, and Nesbert, Marcus, and Ethan: this is dedicated to you with love, and hope.

Dedication

To Mom, Dad, Theresa, Melissa, Nesbert, Ethan, and Marcus.

We are all in this greenhouse together.

Carl Sagan, *Congressional testimony*, 1985

Abstract

The rapid increase of Earth's atmospheric concentration of carbon dioxide drives myriad changes to marine chemistry and ecology, yet its influence in inland waters is relatively unknown. This research advances the understanding of inorganic carbon cycling in Lake Superior, the largest fresh water body on Earth by surface area, and describes how its carbon biogeochemistry responds to perturbations on seasonal to interannual scales. Improved instrumentation for measuring inorganic carbon parameters in freshwater environments was developed and deployed. These tools enabled analysis of the first multi-year underway time series describing drivers of Lake Superior's inorganic carbon cycle, demonstrating a sustained rise in surface water carbon dioxide partial pressure (pCO_2) at approximately the same rate as the atmospheric increase. Machine learning modeling produced an observation-based model of Lake Superior pCO_2 and air-sea CO_2 flux spanning the lake surface over 2019-2023, over which period Superior demonstrated periods of sustained net CO_2 influx and efflux which nearly balanced on annual scales. This result brought Superior's carbon budget closer to closure. Finally, aragonite saturation state was proposed as a useful variable to synthesize inorganic carbon cycling trends with calcifying species presence in inland waters in the face of the potential for lake acidification. The findings of these research projects advance the understanding of Lake Superior's shifting carbon cycle in light of atmospheric CO_2 invasion and acidification, with ecologic outcomes dependent on humanity's choices in the coming decades.

Contents

List of Tables

excepting those in publications

List of Figures

excepting those in publications

Chapter 1

Introduction

The study of chemistry in large bodies of water entails two very different perspectives: the *micro* view of molecules as the fabric of physical reality, and the *macro* view of large lakes and oceans characterized by emergent phenomena arising from and feeding back to the *micro* level. The field of biogeochemistry is defined by the interdependence of chemical systems with both *macro* and *micro* biological and geological contexts (Bianchi, 2021). Together with the disciplines of limnology (the study of inland waters) and oceanography (the study of oceans), this work draws on the expertise and knowledge of conventionally disparate yet closely-related fields to answer questions about carbon cycling in large lakes. In focusing on the biogeochemistry of inorganic carbon — carbon dioxide (CO_2), carbonates, and related species — this work describes freshwater resources in the context of a continuing anthropogenic perturbation of Earth's carbon cycle.

The present climate crisis arises from a fundamental reshaping of biogeochemical cycling at a scale and rate unprecedented in human history (Zeebe, 2012). No aspect of human society nor the ecosystems in which it is embedded will remain unaffected by this perturbation; consequences such as sea level rise, ocean acidification, vanishing ice sheets, and increased frequency of extreme

weather events are already being felt (Gissi et al., 2021). From an Earth science perspective, dramatic shifts of a biogeochemical cycle regulating Earth's climate represent an opportunity to understand the mechanisms connecting element cycles to global change, which further inform climate change mitigation and adaptation efforts (Cavallaro et al., 2018). The global carbon cycle is many-faceted, encompassing organic and inorganic carbon transfers across reservoirs including the atmosphere, hydrosphere, biosphere, geosphere, and cryosphere. The scope of the present work is limited mainly to inorganic carbon in the hydrophere, specifically surface waters in contact with the rising atmospheric CO_2 concentration.

Shifting water chemistry resulting from CO_2 invasion of surface waters remains most advanced in the oceanographic literature, particularly in open, offshore regions relatively free from riverine alkalinity delivery (Gomez et al., 2021), sediment buffering (Archer et al., 1998), biological feedbacks, and noisier coastal acidification signals (Duarte et al., 2013). Despite (and indeed because of) these obstacles, the chemical oceanography community's ability to detect the results of "The Other CO_2 Problem" (Doney et al. (2009); denoted in reference to the greenhouse effect as the "Primary" problem) has advanced considerably with the development of instruments, platforms, and observation strategies measuring these signals with increasing precision, accuracy, and spatiotemporal resolution (Newton et al., 2015; Sutton et al., 2022). Detection of perturbations to inorganic carbon cycling is a task of matching the character of a signal (rate of change, signal-to-noise ratio, seasonality) to instrument deployments with the power to detect it in a given period (Carter et al., 2019). In the context of these marine advances in carbon biogeochemical analysis, this work turns its attention to lakes, particularly large lakes, defined as inland water bodies with a surface area greater than 500 km^2 (Wetzel, 2001), which are in contact with the same anthropogenic atmospheric CO_2 increase as the surface ocean.

Most lakes are net CO_2 sources to the atmosphere as a result of allochthonous

organic and inorganic carbon inputs (Cole et al., 1994) which mediate the balance of autotrophy and heterotrophy along a latitudinal gradient over which increasing latitude favors heterotrophy and CO_2 efflux (Alin & Johnson, 2007). Low-latitude lakes would be expected to favor autotrophy and CO_2 influx, yet studies of tropical large lakes such as Lake Malawi indicated variability in CO_2 flux direction over space and time (Ngochera & Bootsma, 2020).

CO_2 air-water flux may vary in magnitude and direction over space and time, but the inorganic carbon concentration of a lake still remains subject to atmospheric CO_2 changes because an increased partial pressure of atmospheric CO_2 depresses CO_2 efflux, thus resulting in (at least transiently) higher lake CO_2 concentrations effected by buildup of internal CO_2 (Phillips et al., 2015). In smaller or more terrestrially-influenced systems, such a buildup (and associated inorganic carbon equilibrium changes; Section 1.2) may be difficult to detect against the noise of other inorganic carbon cycle drivers such as diel and seasonal temperature and biological variability, allochthonous carbon inputs during periods of high runoff, and other biological and geochemical feedbacks. These sources of noise mirror those in the coastal ocean (Challener et al., 2016; Hofmann et al., 2011; Reimer et al., 2023), so the task of discerning a signal of atmospheric CO_2 influence in lacustrine systems may be amenable to similar strategies as in marine systems. Large lakes present a suitable testing ground for this idea (Sterner, 2021), which motivates the methods of this dissertation. The ideal large lake for study of atmospheric CO_2-driven changes must resemble oceans to facilitate transfer of the principles and technology of chemical oceanography, yet be accessible and sufficiently well-studied to provide the foundation for a campaign of observation and modeling of inorganic carbon cycling.

One candidate is Lake Superior, by merit of its size, cold deep waters subject to twice-annual destratification and homogenization, long history of limnology research (Elster, 1974), and accessibility to a community of researchers. It is the largest lake on Earth by surface area, enabling it to experience tides

(Mortimer et al., 1976) and coriolis-driven circulation phenomena (Bennington et al., 2010). It is deep enough to exhibit euphotic and aphotic zones with a deep chlorophyll maximum (Kovalenko et al., 2019), and internal waves propagating along pycnoclines (Austin, 2013). In many other ways, Lake Superior is unlike an ocean, resembling instead a marginal sea. Strong riverine influences on nearshore nutrient and sediment delivery create regions of altered primary ecosystem function (Delvaux, 2017), primary productivity (Sterner, 2010) and organic matter remineralization (McManus et al., 2003). It has extremely soft water chemistry due to largely igneous mineralogy within its catchment (Chapra et al., 2012; Weiler, 1978). Without a halocline to contribute to stratification, Lake Superior experiences full column mixing twice annually when high winds and seasonal temperature swings deepen its seasonally-variable mixed layer from 10-50 m to its full depth of 406 m (Austin & Allen, 2011; Bennett, 1978). This mixing maintains high oxygen concentrations throughout the water column and penetrating up to 12 cm into sediments (Li et al., 2012). These characteristics together demonstrate the utility of biogeochemical study of Lake Superior as a system amenable to the techniques of marine *and* lacustrine science. Indeed, oceanographic observation may be a necessity in vast lakes presenting such spatial and temporal heterogeneity. In return, studying carbon cycling in a large water body with a lower alkalinity and buffer capacity relative to the world ocean may allow new insights into ocean carbon biogeochemistry by comparison with large lakes. The motivation for this dissertation derives from significant gaps in knowledge of Lake Superior's carbon cycle in the context of a perturbed global carbon cycle.

The carbon cycles of the five Laurentian Great Lakes (hereafter Great Lakes) are understudied, leading to unbalanced carbon budgets and barring their inclusion in regional and global carbon budgets despite a fruitful history of biogeochemical research (Atilla et al., 2011; Bennington et al., 2012; Brothers & Sibley, 2018; Cavallaro et al., 2018; Cotner et al., 2004; McKinley et al., 2011; Urban, 2005). The reasons for these gaps in knowledge are discussed below,

but the consequences of underconstrained carbon biogeochemical budgets for Earth's largest freshwater ecosystems motivate scientific work to address this gap at three levels: local, regional, and global.

At the local level, carbon biogeochemistry interacts with all aspects of element cycling, notably with nutrient nitrogen, phosphorous, and trace metal cycling. These in turn play controlling roles in the balance of production and respiration characterizing a lake's trophic state. Examples of carbon cycling altering nutrient and trophic dynamics may be found in systems subject to organic carbon loading, which has led in the western Lake Erie basin to periodic bottom water anoxia and internal nutrient loading (Karim et al., 2011). Similarly, changes to inorganic carbon concentrations can lead to carbon limitation in systems subject to rapid CO_2 drawdown, for example after an algal bloom (James & Lee, 1974). Excess dissolved inorganic carbon modifies the acidity of a water body. The negative ecological effects of acidification are explored more fully in the ocean acidification (Doney et al., 2009), acid rain (Likens & Bormann, 1974) and mine drainage (McKnight & Feder, 1984) literature, as well as in the discussion below in Section 1.4.

At the regional scale, underconstrained budgets of carbon cycling impair science and management ability to predict the effects of shifting climate and element cycling on populations and water quality (Cavallaro et al., 2018). The invasion of the Great Lakes by Dreissenid mussels provides a dramatic example of an ecosystem engineer rapidly changing whole-lake organic carbon metabolism and inorganic carbon and nutrient cycling. These changes may have impacts on species dynamics, trophic states, and patterns of alkalinity regulation with significant importance to water resources and fisheries management (Cuhel & Aguilar, 2013; Hecky et al., 2004; Li et al., 2021; Lin & Guo, 2016).

The global scale of Earth climate and carbon cycling also interacts with large lake biogeochemistry in a two-way feedback reflecting perturbations to

each system (Williamson et al., 2009). The Great Lakes in particular are significant players in the regional carbon cycle, but as previously discussed, the extent to which they act as sources or sinks to the atmosphere is uncertain. Feedback of the lakes to the global increase in atmospheric CO_2, whether positive (e.g. amplified CO_2 efflux) or negative (e.g. storage of carbon) is similarly unclear, and hinders their integration into terrestrial carbon budgets and earth system models which help define, predict, and mitigate the continuing climate crisis. This framework of local, regional, and global-scale relationships of carbon cycling in the Great Lakes shapes the projects presented in this work. Each tackles one or more of the obstacles to better understanding of carbon cycling in these spheres of influence.

The goals of this work are shaped and informed by three strategies which enable study of Lake Superior's inorganic carbon system:

- Develop and deploy analytical platforms to measure Lake Superior's inorganic carbon cycle with the precision, accuracy, and resolution necessary to describe seasonal and interannual trends.

- Combine established observational products with new measurements to produce descriptions of Superior's inorganic carbon system variability in space and time.

- Integrate inorganic carbon system trends with ecological consequences and feedbacks to enable a fuller understanding of limnology in a world of changing carbon biogeochemistry.

To these ends, observation and modeling projects were carried out over the course of this dissertation research. They are described more fully in the following chapters:

- **Chapter 2** describes the development of an instrument for measuring total alkalinity in freshwater and seawater which enhances the accessibility

of this inorganic carbon parameter. This chapter was published as Sandborn et al. (2023) in the journal *Limnology and Oceanography: Methods*.

- **Chapter 3** presents an analysis of Lake Superior surface pCO_2 describing seasonal and interannual trends informed by the first multiannual time series of underway pCO_2 obtained on Lake Superior. This chapter was published as Sandborn and Minor (2024) in the *Journal of Geophysical Research: Biogeochemistry*.

- **Chapter 4** expands on the underway pCO_2 analysis with a gap-filling model of surface pCO_2 and CO_2 air-sea flux which creates a continuous observation-based product predicting Lake Superior surface pCO_2 spanning 5 years.

- **Chapter 5** synthesizes observations of inorganic carbon parameters and *Dreissenia polymorphia* presence to discuss aragonite saturation state as a predictor of the relationship between chemical drivers and calcifying bivalve habitat.

- **Chapter 6** presents a final discussion of the foregoing work, and puts forward future research directions and related projects.

The remainder of this chapter reviews the present scientific knowledge of carbon cycling in the Great Lakes with emphasis on the inorganic carbon cycle of Lake Superior, describes the theoretical and practical aspects of inorganic carbon chemistry in natural waters, integrates potential shifts in carbon cycling with ecological context and feedbacks, and lays out gaps in knowledge and opportunities with which the rest of this work interfaces.

1.1 Great Lakes carbon cycling

The Great Lakes make up the largest contiguous mass of surface fresh water on Earth. Lakes Superior, Michigan, Huron, Erie, and Ontario (Figure 1.1)

Table 1.1: Laurentian Great Lakes listed with volumes (Fuller & Shear, 1995), mean DIC concentrations (McKinley et al., 2011), and dissolved inorganic carbon reservoir masses. Of note are the Lakes Michigan and Superior with nearly-equal masses each representing approximately a third of the inorganic carbon reservoir of the Great Lakes.

Lake	Volume (km^3)	DIC Concentration ($\mu mol\,C\,kg^{-1}$)	DIC Reservoir (Tg C)
Superior	12,100	860	125
Huron	3540	1569	66.7
Michigan	4920	2165	128
Erie	484	1782	10.4
Ontario	1640	1817	35.8
Total	22,684		366

together contain approximately 23,000 km^3 of water, or about 20% of Earth's available supply. Nearly 40 million humans reside within the lakes' 521,830 km^2 watershed, comprising two countries, and historic and contemporary lands of indigenous and first nations (Fergen et al., 2022). They range in maximum depth from 406 m (Superior) to 64 m (Erie), and in surface area from 82,100 km^2 (Superior) to 18,960 km^2 (Ontario) (Fuller & Shear, 1995). Their waters act as reservoirs for approximately 366 Tg C inorganic carbon, dominated by Lakes Michigan and Superior (Table 1.1).

In addition to their geographic superlatives, the Great Lakes are the setting for a diversity of ecologic communities interacting along gradients of climate (subarctic to humid continental, Beck et al. (2018)), geology (Precambrian shield to Quaternary glacial deposits, Kesler (2019)), trophic status and nutrient availability (Sterner, 2021). These factors together demonstrate an interconnected web of interactions which contributes to abiding scientific interest in the Great Lakes region. The overarching narrative produced by this interest has been one of change: of climate change and glaciation before and during human settlement (Saarnisto, 1974), of response to the pressures of human presence stretching back at least 15,000 yr (Becerra-Valdivia & Higham, 2020),

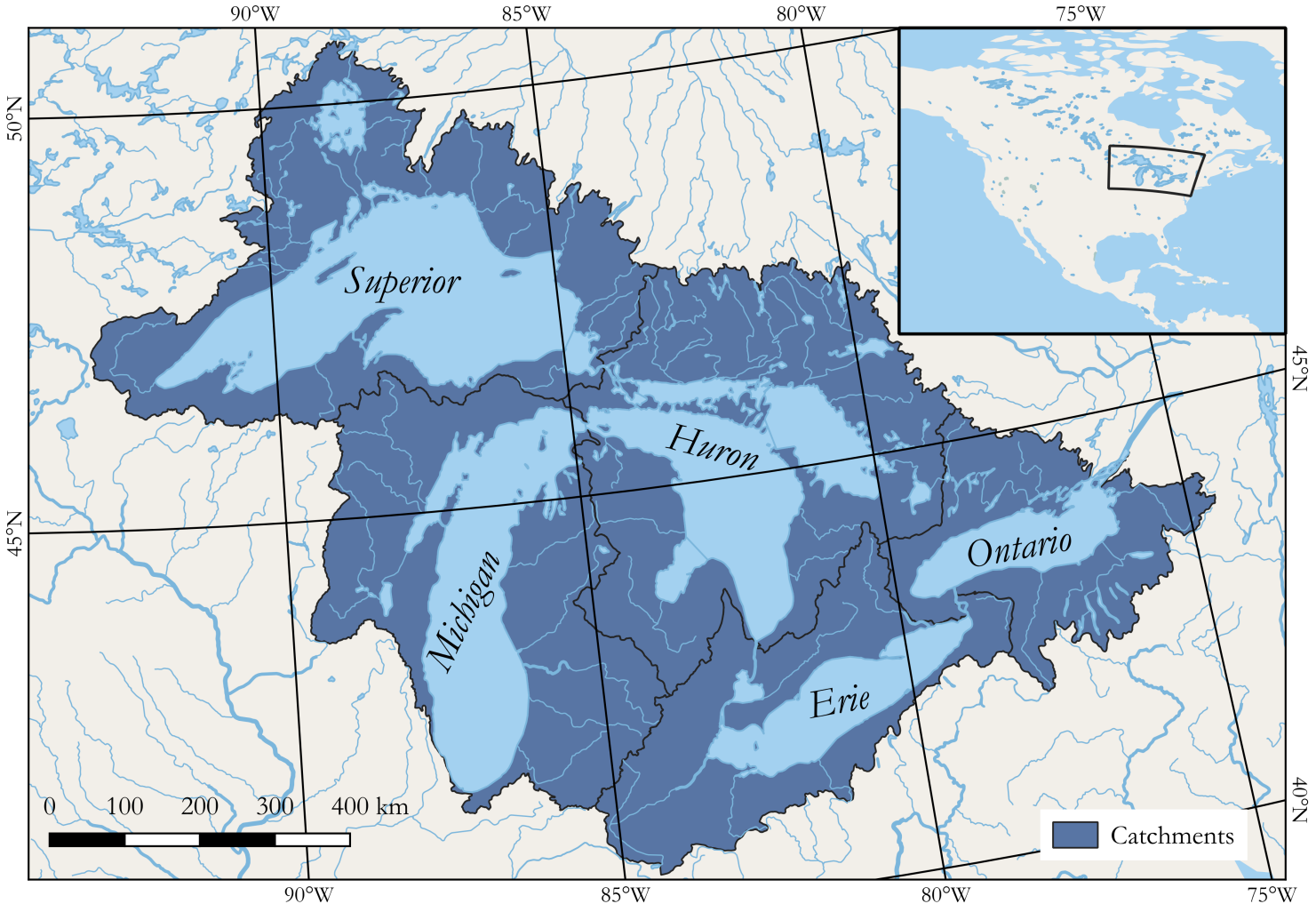

Figure 1.1: The Laurentian Great Lakes, with their catchments shaded and major rivers drawn. Produced using data from Natural Earth and the United States Geological Survey.

of water quality declines and improvements (Hartig et al., 2020), of challenges to fisheries and other natural resources (Regier et al., 1999), and of broad scale ecosystem shifts associated with invasive species (Escobar et al., 2018). In a reductive sense, all of these changes connect with shifts in the global carbon cycle: the industrial revolution which enabled exponentiation of human impact was itself enabled by utilization of energy from fossil fuels which increased atmospheric CO_2 concentrations.

Such reasoning, however, ignores the intermediate connections between changes in the Great Lakes' ecology, physical characteristics, and biogeochemical element cycling underlying, responding to, and feeding back to external forcing. It is the goal of this dissertation to illuminate these connections and describe a more mechanistic connection between the external forcing of atmospheric carbon dioxide, lacustrine carbon cycling, and the character of large lake ecology and physicochemical characteristics. To do so necessitates beginning at the *micro* level of the inorganic carbon complex, which will set the stage for an interface with the *macro* scale of global change and the Great Lakes.

1.2 The inorganic carbon complex

Inorgnic carbon plays a pivotal role in determining the chemical composition of natural waters by responding to and regulating biological and geochemical processes. It exists as aqueous CO_2/carbonic acid (hereafter $CO_2{}^*$ as the two species are functionally indistinguishable), bicarbonate ion ($HCO_3{}^{2-}$), carbonate ion ($CO_3{}^{2-}$), carbonate minerals in contact with natural waters (e.g. calcite, aragonite, ikaite, dolomite, vaterite), and various ion pairs in solution, collectively termed the "inorganic carbon complex" (Wetzel, 2001). These species participate in a series of chemical equilibria described by thermodynamic equilibrium coefficients K, the values of which depend on environmental conditions including temperature, pressure, and ionic strength. For most practical purposes, K which is a function of species' activities (denoted with braces)

at equilibrium is replaced by K', the apparent equilibrium constant which is a function of species' concentrations (denoted with brackets) at equilibrium. Given an aqueous system at equilibrium with atmospheric pCO_2 and sediment $CaCO_3$, and neglecting all other species present in natural waters, the inorganic carbon complex can be summarized at a surface level by the reactions:

$$CO_2(g) \overset{K_0'}{\rightleftharpoons} CO_2^* \quad K_0' = \frac{[CO_2^*]}{[CO_2(g)]} \tag{1.1}$$

$$CO_2^* \overset{K_1'}{\rightleftharpoons} HCO_3^- + H^+ \quad K_1' = \frac{[HCO_3^-][H^+]}{[CO_2^*]} \tag{1.2}$$

$$HCO_3^- \overset{K_2'}{\rightleftharpoons} CO_3^{2-} + H^+ \quad K_2' = \frac{[CO_3^{2-}][H^+]}{[HCO_3^-]} \tag{1.3}$$

$$CaCO_3 \overset{K_{sp}'}{\rightleftharpoons} CO_3^{2-} + Ca^{2+} \quad K_{sp}' = [CO_3^{2-}][Ca^{2+}] \tag{1.4}$$

$$H_2O \overset{K_w'}{\rightleftharpoons} H^+ + OH^- \quad K_w' = [H^+][OH-] \tag{1.5}$$

Here, $[H^+]$ encompasses the complex of acidic species (hydronium ions, bisulfate, hydrofluoric acid) described in Section 1.3.3. In natural waters containing many species interacting with the inorganic carbon complex (e.g. borate, fluoride, sulfate, sulfide, organic acids, etc.), this scheme represents a simplified view of the inorganic carbonate system and its primary chemical species and reactions. The accurate determination of inorganic carbon speciation and associated biogeochemical cycling necessitates accounting for these additional species' concentrations and equilibrium reactions; see Section 1.3.5. The inclusion of equilibrium with $CaCO_3$ here is significant; Lake Superior lacks carbonate minerals in its surface sediments in contrast to the remaining Great Lakes (Thomas & Dell, 1978), which carries implications for its reaction to acidification (Section 1.4) as well as for its suitability for invasive calcifying species

(Chapter 5). It should be emphasized that the carbon cycle comprises both the inorganic carbon complex as well as organic carbon cycling. A brief overview of their interactions is necessary to establish the context for the inorganic carbon complex in natural waters, and motivate further discussion of analytical approaches (Section 1.3) and applications 1.4 of inorganic carbon cycling science.

1.2.1 Interaction of inorganic and organic carbon

Organic carbon interacts with inorganic carbon cycling at many points in a water body's carbon cycle. Although the inorganic and organic pools are separate, they become closely associated by direct exchanges between their pools as well as by co-association with other biogeochemical processes. These two categories neatly (if somewhat arbitrarily) divide the present treatment of interactions between organic and inorganic carbon cycles, which are discussed in greater depth in the cited references.

Organic carbon, in both dissolved (DOC) and particulate (POC) forms, may represent a relatively major or minor portion of a lake's carbon pool, depending on the dynamic balance of organic carbon production, loading, and remineralization. Production of organic carbon occurs via autotrophy, in which inorganic carbon is reduced and organic carbon is synthesized. Organic carbon is remineralized biotically (e.g. respiration) and abiotically (e.g. photochemical degradation) to produce dissolved inorganic carbon. The changes to inorganic carbon concentration and speciation resulting from organic matter production and degradation may alter pH, affect carbon availability for primary producers, alter CO_2 gas flux, and shift the balance of particulate inorganic carbon and dissolved inorganic carbon in sediments and the water column. The give-and-take of coupled organic and inorganic carbon cycles results in carbon recycling or turnover, which is measured as an indicator of biological activity varying across lake systems in correlation with environmental factors

such as climate and water retention time. Lake Superior's organic carbon pool of approximately 16.2 – 21.5 Tg C (Cotner et al., 2004) is dwarfed by its inorganic carbon pool of approximately 125 Tg C (Table 1.1) yet the short span of its organic carbon turnover time (8 – 60 yr) relative to its hydrologic residence time c. 172 yr implies rapid and repeated internal carbon recycling (Minor & Oyler, 2021). These exchanges between inorganic and organic pools are combined with other carbon inputs and outputs in Lake Superior's carbon budget, in which respiration and production constitute the largest terms and propagate the most uncertainty to implied air-sea CO_2 flux (Urban, 2005). Chapter 4 provides an updated look at Lake Superior's carbon budget with a discussion of the imprint of organic carbon processing on implied and observed CO_2 flux.

Indirect interactions between lake organic and inorganic carbon cycles occur as a result of other biogeochemical element cycles. Representative examples may be found in nutrient cycling, particularly nitrogen and trace metal cycles. Conversion among the various species of nitrogen in natural waters (nitrate, nitrite, ammonia, organic nitrogen, dinitrogen) facilitates growth and energy transfer as organisms use these species as substrates for respiration, terminal electron acceptors, and essential building blocks of organic matter. The chemical reactions changing the oxidation state of nitrogen (e.g. nitrification, anammox) produce or consume dissolved inorganic carbon and alkalinity in characteristic ratios proportional to the amount of organic matter produced or consumed, shifting the equilibrium of the inorganic carbon complex (Wolf-Gladrow et al., 2007; L. Xue & Cai, 2020). Similarly, changes to a water body's pH alter trace metal speciation, which has been observed to lead to altered iron-mediated photochemical oxidation of DOC (Sulzberger, 2015) as well as changes in organic carbon solubility as a function of trace metal complexation (Winterdahl et al., 2014).

These examples of interplay between organic and inorganic carbon cycles suffice to illustrate their coupling in both direct and indirect means. The projects in the following chapters continue to engage with this subject, for example in

the discussion of organic alkalinity as a player in the inorganic carbon complex in Chaapter 2, and the inference of productivity and respiration as direct drivers of pCO_2 variability in Chapters 3 and 4.

1.3 Analytical approaches to inorganic carbon cycling

Several chemical parameters are defined and directly measured in water to determine the state of the inorganic carbon complex: dissolved inorganic carbon (DIC), total alkalinity (A_T), the partial pressure of CO_2 (pCO_2), and pH. While others (e.g. CO_3^{2-}; Byrne and Yao (2008)) are occasionally directly measured, their theoretical and analytical details are not described here.

1.3.1 Dissolved inorganic carbon

Dissolved inorganic carbon is the total amount of carbon dioxide dissolved in sample of water, given as the sum of its dissolution products' concentrations (Equation 1.6). The units of DIC throughout this work are µmol C kg^{-1}.

$$DIC = [CO_2^*] + [HCO_3^-] + [CO_3^{2-}] \tag{1.6}$$

DIC is perhaps the most intuitive inorganic carbon parameter, as it represents a tangible quantity of matter dissolved in water. DIC has long been one of the most well-defined chemical quantities used by inorganic carbon chemists. It is most commonly measured by sparging CO_2 from a solution of known mass by addition of a strong acid, followed by quantification of the evolved gas by coulometry or non-dispersive infrared light absorption. The imprint of the increasing atmospheric CO_2 burden in the surface ocean is captured by increasing DIC, but seldom measured or represented as such due to difficulty in discriminating the anthropogenic CO_2 signal from background preindustrial DIC, productivity, and remineralization (Gruber et al., 1996; Sabine et al., 2004). Subsequent chapters include the results of DIC analyses of Lake

Superior water samples collected, preserved, and analyzed via oceanographic best practices (Dickson et al., 2007) and used mainly for the purposes of constraining equilibrium calculation of other inorganic carbon parameters.

1.3.2 Total alkalinity

Total alkalinity is defined as "a measure of the proton deficit of the solution relative to an arbitrarily defined zero level of protons" (Dickson, 1981) which can then be calculated at the chemical species level as the sum of the concentrations of the proton acceptors minus the sum of the concentrations of the proton donors. The oceanographic zero proton level (established with reference to existing titration methods as well as convenient gaps among major species' pK_a values) is pH = 4.5, which for the most common inorganic species in natural waters implies (Wolf-Gladrow et al., 2007):

$$\begin{aligned} A_T = {} & [HCO_3^-] + 2[CO_3^{2-}] + [B(OH)^{4-}] + [OH^-] + [HPO_4^{2-}] \\ & + 2[PO_4^{3-}] + [SiO(OH)_3^-] + [HS^-] + 2[S^{2-}] + [NH_3] \\ & - [H^+] - [HSO_4^-] - [HF] - [H_3PO_4] \end{aligned} \quad (1.7)$$

In practice, A_T is determined via acidimetric titration in which the addition of a known amount of hydrochloric acid (sulfuric acid introduces error by slightly modifying A_T; see Appendix B) decreases the pH of a sample monitored potentiometrically (Dyrssen & Sillén, 1967) or spectrophotometrically (Yao & Byrne, 1998). The calculation of alkalinity from these titration data is typically accomplished by single-endpoint (though this is not recommended; Rounds and Wilde (2012)) or multi-point linear interpolation of a Gran function (Gran et al., 1950) in inland waters, but non-linear interpolation of a modified Gran function better accounts for multiple chemical equilibria and electrode response function variability and uncertainty close to the zero proton

level, and so has become standard in marine work (Dickson, 1992). A closely-related concept in freshwater work is "acid neutralizing capacity", which denotes the alkalinity of an unfiltered water sample (Rounds & Wilde, 2012).

In most natural waters, HCO_3^- is the main source of A_T, with CO_3^{2-} increasing in importance with pH (Eq. 1.3). Each body of water has a distinct set of non-carbonate proton acceptors and donors; in estuarine and marine waters, such species as borate, fluoride, (bi)sulfate, and others play increasingly large roles in determining A_T (and thus the state of the entire inorganic carbon complex) with increasing concentrations frequently estimated from salinity. A recent development is the recognition of organic acid-base systems which exert considerable influence on the inorganic carbon complex in waters across the salinity spectrum (S. Liu et al., 2020). These have been termed organic (or excess) alkalinity, but details on their occurrence, structure, and chemical properties are difficult to define in the same way as the other A_T species, because the dissolved organic matter (DOM) in a given parcel of water is a unique and complex ensemble of organic molecules with variable dissociation and ion pairing properties. Taking these issues into account requires an extended discussion of alkalinity measurement in freshwater systems, found in Chapter 2, which develops a custom instrument advancing high-quality and accessible A_T measurement across the salinity spectrum.

Alkalinity is perhaps the least-intuitive of the inorganic carbon parameters presented here by merit of its functional definition, multiple interpretations, and computational complications. Despite these barriers, alkalinity remains one of the most commonly-measured water chemistry parameters in inland and marine waters alike. The multiple methods by which it is measured (and thus defined) present serious difficulties to its incorporation into inorganic carbon system equilibrium calculations in inland waters (S. Liu et al., 2020), especially compounded with uncertainties and errors associated with pH analysis.

1.3.3 pH

pH is a measurement of the chemical activity of hydrogen ions in solution. It is frequently depicted as a measurement of acidity (actually a separate water chemistry parameter defined as negative alkalinity; Stumm and Morgan (1996)). This parameter is the focus of sustained attention in the context of increasing atmospheric pCO_2, as atmospheric CO_2 invasion of surface waters shifts carbonic acid dissociation equilibria (Equations 1.1-1.3), decreasing pH. The definition of a thermodynamically-valid and computationally-useful pH in terms of chemical concentrations carries a host of problems that have led to the creation of multiple pH scales. The differences among these scales are not always recognized across disciplines, creating barriers to accurate computation with reported pH values.

The concept of pH as a logarithmic transform of hydrogen ion concentration is taught in introductory science courses:

$$\mathrm{pH} = -\log_{10}[\mathrm{H}^+] \tag{1.8}$$

This definition suffers from the fact that hydrogen ions are bound to water molecules to form hydronium ions, and interacting with all other ions in solution to modify the true activity of H^+. The NBS (National Bureau of Standards) scale was created in 1957 in an attempt to approach a pH defined by the activity of the hydronium ion (R. Bates & Wichers, 1957) using standard buffer solutions with assigned pH values. The NBS scale only approaches the true activity scale in dilute media, in which unpredictable glass electrode liquid junction potentials introduce uncertainty on the order of 0.02 units (French et al., 2002; S. Liu et al., 2020). In media with higher ionic strengths (such as seawater) and lower ionic strengths (such as Lake Superior) even larger drift, bias, and uncertainty can arise (Stauffer, 1990; Young et al., 2022). To solve the scale dilemma in marine work, several pH scales were proposed which defined pH in standard buffer solutions with ionic strengths around that of seawater

(c. 0.7). The Free scale (Eq. 1.9), Seawater scale (Eq. 1.10), and Total scale (Eq. 1.11) differ in their inclusion or exclusion of bisulfate or hydrofluoric acid species.

$$\mathrm{pH_{free}} = -\log_{10}[\mathrm{H^+}] \tag{1.9}$$

$$\mathrm{pH_{seawater}} = -\log_{10}([\mathrm{H^+}] + [\mathrm{HSO_4^-}] + [\mathrm{HF}]) \tag{1.10}$$

$$\mathrm{pH_{total}} = -\log_{10}([\mathrm{H^+}] + [\mathrm{HSO_4^-}]) \tag{1.11}$$

For media with known (or estimated) concentrations of total sulfate and fluoride, these values are readily inter-converted, but this depends on pH values being denoted with their scales, which isn't always the case. At low ionic strength (i.e. fresh water), the differences among free, seawater, and total scales disappear, but the offset associated with the NBS scale does not because the reference states of the NBS and oceanic pH scales assume that the activity coefficient of H^+ approaches unity when $[H^+]$ approaches zero in pure water or seawater, respectively (Millero, 1986; Zeebe & Wolf-Gladrow, 2001). Two methods are commonly employed for measurement of pH in natural waters: potentiometry and spectrophotometry. The latter has come to be preferred by investigators seeking discrete pH measurements with sufficient accuracy for equilibrium calculations and detection of ocean acidification and other modes of variability for which measurement uncertainty associated with glass electrode potentiometry (in all but the best-controlled experiments; Easley and Byrne (2012)) bars the latter's use. Recent development of instruments based on ion-sensitive field effect transistors (ISFET; Bagshaw et al. (2021) and Martz et al. (2010)) have expanded the possibilities for deployed and autonomous pH measurement, but their drift and need for calibration contribute to spectrophotometric methods' continuing use in point measurements.

pH is measured spectrophotometrically in succeeding chapters, aligning

with current best practices including correction for dye perturbation of samples, use of purified indicator dye, and flow-through cuvettes minimizing gas exchange (Carter et al., 2013; X. Liu et al., 2011). Discrete measurements of pH, A_T and DIC produced in this work complement the continuous and autonomous underway measurements of pCO_2 employed in Chapters 2, 3, and 4.

1.3.4 pCO_2 and CO_2 air-sea flux

The partial pressure of CO_2 of a given atmosphere in equilibrium with a parcel of water is denoted as pCO_2, given in this work in µatm. This parameter is measured with high accuracy in the headspace of a discrete water sample or in an autonomous flow-through system connected to a vessel or mooring. Chapters 2, 3, and 4 include pCO_2 measured using a SuperCO2 instrument equipped with dual showerhead equilibrator connected to the underway system of the *RV Blue Heron*, and Chapter 4 includes pCO_2 measured using a CO_2-Pro instument deployed on a mooring and equipped with a membrane-based equilibration system; these two types of equilibrators represent the most common methods in use. pCO_2 is frequently discussed with fugacity (fCO_2), which is the *activity* of CO_2 and nearly equal (within a few parts per thousand) to pCO_2 under standard conditions, at which the CO_2 fugacity coefficient a approaches unity and

$$fCO_2 = a \times pCO_2 \tag{1.12}$$

Along with its advantages as a time series variable, pCO_2 boasts a direct mechanistic connection to CO_2 flux across the air-water interface, the rate of which has been experimentally shown to be proportional to the difference in surface water CO_2 concentration relative to the overlying atmosphere, scaled by a gas flux velocity k (an empirical function of temperature, salinity, and/or wind speed approximating complex boundary layer processes) and converted

from partial pressures to concentrations by Henry's constant K_o, resulting in a bulk gas flux parameterization:

$$\mathrm{CO_2\,Flux} = k \times \Delta[\mathrm{CO_2}]_{water-air} = kK_o \times (p\mathrm{CO_{2water}} - p\mathrm{CO_{2air}}) \tag{1.13}$$

The choice of k is by no means trivial (Ho et al., 2011; Wanninkhof, 1992). A condensed discussion of various air-sea gas flux parameterizations is presented in Chapter 3, which attempts to convey the uncertainty attached to application of a gas flux model to large lakes. Applications of $p\mathrm{CO_2}$ underway time series measurements are developed further in Chapter 3 and incorporated into the training and validation of a machine learning predictive model in Chapter 4.

1.3.5 Equilibrium calculations with inorganic carbon parameters

Implicit in the definitions and choice of measurable inorganic carbon parameters is the assumption that one may calculate the value of any measurable inorganic carbon parameter (and many more unmeasurable) from any two others, given the additional context of temperature, depth, salinity, and reasonable assumptions of the relevant equilibrium coefficients and non-carbonate species' concentrations. A suite of software packages have been developed for this task, most of which are essentially equivalent in ability and result (Orr et al., 2015); this work uses PyCO2SYS (Humphreys et al., 2020). These tools, however, were designed for marine and estuarine settings. Adaptation to inland waters necessitates consideration of other pertinent differences, such as the definition of parameters ($\mathrm{A_T}$, pH), choice of carbonic acid equilibrium coefficients, and matrix differences. While software designed for inland water equilibria (i.e. PHREEQC; Pötter et al. (2021)) has a long history of use, it was not found suitable for this work given its lack of uncertainty propagation and other shortcomings.

Differences in inorganic carbon parameter definitions have been discussed

in previous sections, and are best exemplified by the competing definitions of pH and alkalinity. When a reported value is known or justifiably-assumed to conform to the definitions of one of the four pH scales (NBS, free, etc.) or the rigourous definition of A_T, the appropriate choices can be made in the software; otherwise, additional error is propagated to output values. This and other sources of measurement uncertainty have contributed to longstanding systematic and unsystematic errors associated with freshwater inorganic carbon equilibrium calculations (Golub et al., 2017).

The choice of carbonic acid apparent equilibrium coefficients best-suited for inland waters falls largely to comparison of the salinity range across which they were constrained against that of a study system. Of the coefficient sets in common use, only three are considered valid for salinities less than 1: those of Millero (1979), Cai and Wang (1998), and Waters et al. (2014). Millero's work was a linear extrapolation of marine equilibrium coefficients to infinite dilution, while Cai and Wang constrained their coefficients by observations along estuarine salinity gradients, and Waters et al. modeled equilibria using a Pitzer ion-association model constrained by buffers along a wide salinity gradient. Previous work indicated that the coefficients of Cai and Wang led to less error in pCO_2 measured and calculated from pH and A_T (Minor & Brinkley, 2022). This work found greatest consistency in the coefficients of Waters et al. (which weren't available to Minor & Brinkley) when well-defined pH and A_T measurements were combined with adjustment for the chemical matrix of Lake Superior by specifying major ionic concentrations (Chapter 2). These assumptions inevitably contribute uncertainty to the calculated inorganic carbon parameters, which must be combined with that generated from measurement uncertainty in a repeatable and rigorous fashion.

Propagation of Uncertainty

Efforts to observe inorganic C cycling in aquatic systems can be helped or hampered by the choice of observable parameters, specifically by how their measurement uncertainties propagate to a desired calculated parameter (Orr et al., 2018). Uncertainty in an analytical sense is defined as a function of both precision and accuracy, representing an interval within which a true value may be found. For calculated parameters, the uncertainties of variables within its function must be combined and propagated. The uncertainty u_y of an output variable y which is a function of input variables $x_1, x_2, ...x_n$ can be found via Taylor series expansion (neglecting covariance; cf. Dickson and Riley, 1978; Orr et al., 2018) as

$$u_y = \sqrt{\sum_{i=0}^{n} \left(\left(\frac{\partial y}{\partial x_i} \right)^2 u_{x_i}^2 \right)} \tag{1.14}$$

Application of this formula to calculation of uncertainty propagated from measured inorganic carbon parameters to calculated parameters often necessitates many terms and complicated partial derivatives. As a simple example, DIC can be explicitly calculated from A_T and pH (Appendix B in Zeebe and Wolf-Gladrow (2001), neglecting borate and other minor equilibria):

$$\mathrm{DIC} = \frac{\left(A_T - \frac{K_W^*}{[H^+]} \right) \left(1 + \frac{K_1^*}{[H^+]} + \frac{K_1^* K_2^*}{[H^+]^2} \right)}{\frac{K_1^*}{[H^+]} + 2\frac{K_1^* K_2^*}{[H^+]}} \tag{1.15}$$

Applying Equation 1.14 to Equation 1.15 results in a lengthy series of partial differential equations containing the uncertainties of A_T, $[H^+]$, and the various equilibrium constants. This analytical process is unwieldy in manual application, so numerical methods are employed to produce propagated uncertainties more efficiently. This work makes use of propagation routines built into inorganic carbon equilibrium calculation software PyCO2SYS to illustrate the value of improved measurement accuracy and precision (Chapter 2, Fig. 1)

and careful choice of measured parameters propagating with minimal relative error (Chapter 4, Section 1.3).

The analytical tools described in the preceding sections are key elements in research observing and modeling the inorganic carbon complex of natural waters. The following sections present the motivation and strategies for their use in this work.

1.4 Shifts in inorganic carbon cycling

With the mechanisms and theoretical aspects of the inorganic carbon complex summarized, their roles in hypothesized carbon cycle shifts in the Great Lakes are considered, along with strategies for detecting shifts which underlie the succeeding chapters. The role of atmospheric CO_2 in shifting surface water carbon cycling has been described, along with attendant acidification, atmospheric CO_2 invasion, modification by feedbacks such as temperature and alkalinity changes, and ecological consequences of these and other interacting stressors of aquatic communities. Observing and predicting these effects in inland waters, specifically in the Great Lakes, motivates consideration of their connections to this dissertation.

1.4.1 CO_2 invasion

CO_2 invasion of surface water is a rise in DIC driven by a parallel rise in atmospheric concentrations over a defined period which may be sub-seasonal to multiannual. This phenomenon leads to the surface ocean acting as a sink for approximately 30% of anthropogenic CO_2 emissions (Friedlingstein et al., 2023). As previously argued, any surface water body must negotiate this same equilibrium and potentially undergo atmospheric CO_2 invasion. Detecting CO_2 invasion in Lake Superior, specifically over interannual periods, involves several assumptions and observational requirements.

First, it is assumed that Lake Superior approaches chemical equilibrium with respect to atmospheric CO_2 repeatedly. This is implied by biogeochemical modeling by Bennington et al. (2012) which illustrated repeated achievement of surface water atmospheric equilibrium. Second, it is assumed that no other internal or external source of CO_2 exists with an interannual trend large enough to overwhelm the atmospheric signal. Despite research indicating that metabolic components of Lake Superior's carbon budget exhibit significant interannual variability (Brothers & Sibley, 2018), no sustained source or sink is evident. Third, it is assumed that capturing seasonal variability in observations of inorganic carbon parameters will allow deconvolution of an interannual trend from smaller-period variability, which has been a persistent obstacle to previous research pioneering this inquiry (Minor & Brinkley, 2022; Minor et al., 2019). Lastly, it is assumed that decreasing measurement uncertainty through the use of sampling and analytical techniques with improved accuracy and precision will increase signal-to-noise ratios.

These assumptions inform the work in Chapter 2, which describes improved instrumentation for A_T analysis paired with DIC and pH analyses of Lake Superior water. They also come into play in Chapter 3, which isolates an interannual trend in observed Lake Superior surface water pCO_2 relative to previous modeling work which parallels the atmospheric trend. The inorganic carbon parameter measurement techniques and trends are once again used in Chapter 4 to produce and validate a spatially-comprehensive multi-annual description of surface water pCO_2 seasonal cycling and CO_2 flux, which together support continuing, repeated, and rapid atmospheric equilibrium of Lake Superior surface waters with rising atmospheric CO_2. This work interfaces further with hypothesized lake acidification as a symptom of atmospheric CO_2 acidification.

1.4.2 Acidification

It has been hypothesized that the Great Lakes will experience acidification mirroring Ocean Acidification (Phillips et al., 2015), but the equilibrium of atmospheric and lake CO_2 and interactions with other variables (warming waters, decreasing ice cover, river alkalinity delivery, and variable lake levels) remains relatively poorly-understood. Discussion of the Great Lakes acidification hypothesis would benefit from reinterpretation of the forecast of Phillips et al. using updated atmospheric CO_2 forcing and integration of a novel feedback in the form of $CaCO_3$ compensation.

The years since publication of Phillips et al. have seen updates to the Intergovernmental Panel on Climate Change (IPCC) CO_2 forecasts (IPCC, 2022) and better-constrained lake warming trends (P. Xue et al., 2022). For the sake of brevity and focus on the best-understood systems, acidification forecasts are presented for lakes Superior and Michigan using the core assumptions of Phillips et al: no seasonality and equilibrium with atmospheric CO_2 in addition to a further simplifying factor of constant 4 °C water temperature. These assumptions have a physical basis during and after spring destratification, when surface waters warm to the temperature of maximum density and cause whole-column mixing. The most important difference between lakes Superior and Michigan in terms of acidification regulation is the total absence of particulate inorganic carbon in the water and sediments of Lake Superior and its prevalence in Michigan (evidenced by whiting events, the frequency of which has decreased in recent years in lakes Michigan and Ontario in part due to Dreissenid sequestration of $CaCO_3$; Barbiero et al., 2006). Particulate inorganic carbon as $CaCO_3$ represents a potent buffer to decreasing pH, well-studied in the oceans as the primary millennial-scale sink of anthropogenic CO_2 (Archer et al., 1998). A buffering process that will take millennia in the ocean (characterized by millennial-scale overturning) may take only months or years in a dimictic lake characterized by twice-annual full-column water

mixing which exposes carbonate sediment to recently-ventillated waters. An extended description or model of the kinetics and thermodynamics of this sediment buffer is beyond the present work, which presents a simplified acidification model for the purpose of motivating and contextualizing the following chapters.

Mirroring the work of Phillips et al., constant A_T was assumed for both lakes in the scenario without $CaCO_3$ compensation, which was used together with an assumption of equilibrium with overlying atmospheric CO_2 concentrations from IPCC scenarios to calculate pH at a constant 4 °C using PyCO2SYS software, using the coefficients and constants discussed in Section 1.3.5. For the case of $CaCO_3$ compensation in Lake Michigan, constant $CaCO_3$ (calcite) saturation state was assumed, implying neutralization of the atmospheric CO_2 signal to maintain present saturation state conditions (mirroring the simplified buffering model presented by Sarmiento and Gruber (2006), their Chapter 10). The resulting pH forecast, extending to 2120, is presented in Figure 1.2.

The most striking aspects of this simplified acidification forecast are the wide spread of acidification results over the next century as well as the substantial negative feedback of $CaCO_3$ compensation at work in Lake Michigan. The spread in acidification is driven by the diverging scenarios formulated in the most recent IPCC report, implying that incremental climate change mitigation efforts may bring about significant gains in preventing lake acidification. The rates and magnitudes of projected acidification essentially match those of surface ocean acidification over the same time span (Jiang et al., 2023). $CaCO_3$ compensation approximately halves acidification in Lake Michigan, but does not eliminate it completely. This forecast lacking seasonality likely ignores important effects, such as $CaCO_3$ dissolution in undersaturated hypolimnia during stratified periods (which was responsible for varved $CaCO_3$-containing sediments in Lake Superior during the last glaciation; Thomas and Dell, 1978).

This acidification forecast remains a decidedly tentative hypothesis, subject to confirmation of atmospheric CO_2 invasion, pH trends, and feedbacks such

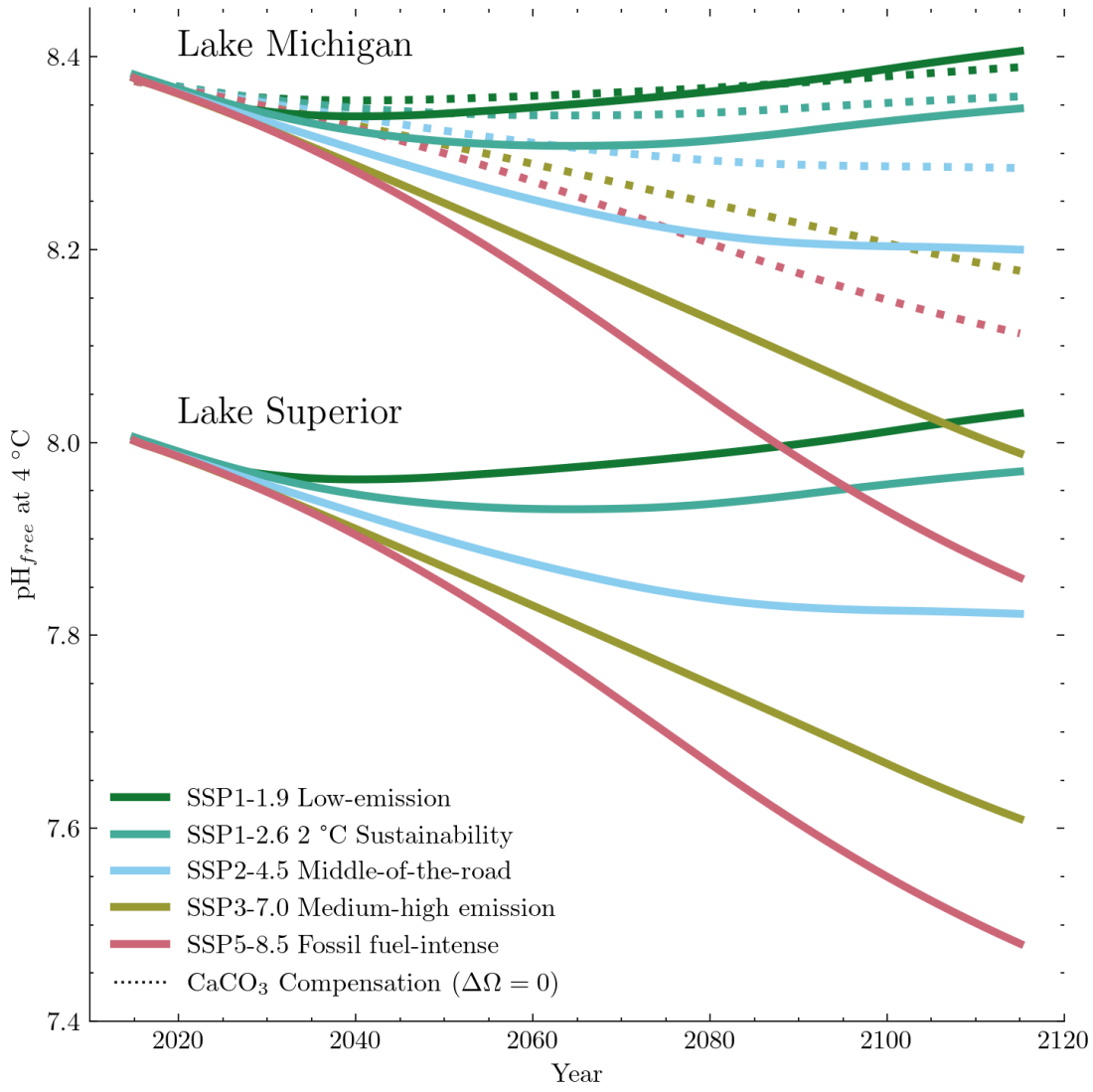

Figure 1.2: Simplified acidification forecast for Great Lakes Superior and Michigan, after Phillips et al. (2015). IPCC predictions of atmospheric CO_2 concentrations for 5 pathways were used to force calculations of inorganic carbon equilibria at a constant 4 °C which assumed constant total alkalinity (solid lines) or constant $CaCO_3$ saturation state (dotted lines, Michigan only). PyCO2SYS software was used to perform equilibrium calculations.

as $CaCO_3$ compensation. Elements of this dissertation include development of tools and strategies for observing any alkalinity feedback to acidification in Chapter 2, illustrating large seasonal variability of calculated pH in Chapter 3 which may obscure present and future trends, discussing and quantifying spatial heterogeneity of inorganic carbon parameters in Chapter 4 which complicate presentation of mean lake conditions, and engaging with ecological outcomes of acidification in Chapter 5. This last subject deserves further context in a brief discussion of Great Lakes ecosystems and a shifting inorganic carbon cycle.

1.4.3 Ecological interactions

With global drivers and chemical mechanisms of the changing CO_2 cycle established, a brief overview of ecological outcomes is necessary to reconnect the *micro* scale of the succeeding studies to the *macro* context of lacustrine ecosystems. Just as gaps exist in the scientific knowledge of inland water CO_2 cycling in the face of the anthropogenic perturbation, so too does ecological knowledge of effects of this phenomenon suffer from insufficient study. Once again reference is taken from the marine world, in which higher pCO_2, lower pH, changes to oxygen availability, warming waters, and related co-stressors have spurred intensive study of their effects at all levels of ecological organization. In brief, oceans which are "warming up, turning sour, losing breath" (Gruber, 2011) stand to harm ecosystem structure and function by degrading the survival and reproduction of organisms, shifting or diminishing their populations, and fracturing stressed communities, which contributes to the picture of ecosystems threatened by intersecting stressors (Breitberg et al., 2015).

These findings have been developed by observational and experimental studies informed by well-constrained models of marine chemistry and climate change. In general, the analogous changes in inland water carbon cycling are more

poorly-constrained for reasons that have already been discussed, which contributes to a dearth of limnological studies of aquatic ecosystems undergoing similar stresses. Recent studies have made headway towards alleviating this gap in knowledge, postulating or observing ecological responses to rising pCO_2 (L. C. Weiss et al., 2018) and acidification (Johannesson et al., 2023). Higher water pCO_2 may cause metabolic challenges (Jeffrey et al., 2018) and zooplankton behavioral changes (L. C. Weiss et al., 2018). Decreasing pH may affect organisms via acidosis, changes to mineral and trace metal solubility, and homeostasis challenges (Hasler et al., 2018).

The postulated and observed ecological effects of shifting inland water inorganic carbon chemistry represent challenges as well as opportunities for limnologists and biogeochemists seeking to understand how aquatic systems respond to perturbation. The diversity of inland water ecosystem structures and their varying sensitivities to atmospheric CO_2 invasion would seem to imply myriad possible outcomes of this threat, in seeming contrast to the monolithic treatment of the ocean acidification trend superimposed upon the world's oceans. Closer reading of the marine literature indicates a diversity of sensitivities and feedbacks to anthropogenic CO_2 invasion at chemical and ecological levels (Doney et al., 2020). Further study of the manifold inland water environments in the context of the CO_2 signal may contribute meaningfully to the understanding of *all* Earth surface waters in a period of shifting inorganic carbon cycling.

1.5 Summary

In this introductory chapter, context for the following work was introduced by outlining the increasing atmospheric CO_2 signal as a potential driver of large lake carbon cycling shifts. Several hypotheses were enumerated and technical approaches designed to address them, introducing and motivating a series of observation and modeling projects described in the next four chapters, which

seek to advance knowledge of inorganic carbon biogeochemistry in important surface water resources.

One important point remains to be made: the social context and ethical import of work to address the continuing climate crisis. This work is not motivated solely by academic interest in chemical limnology and oceanography. Its genesis is the pressing need for understanding given a clear and present danger of ignorance and inaction. The climate crisis is a symptom of a planetary-scale experiment with element cycling which has rapidly (on geological timescales) effected a departure from the stable Holocene climate which allowed humanity to develop and thrive (Broecker, 1985). Much has been written and much remains to be discovered about the drivers, mechanisms, and outcomes of this perturbation to Earth's carbon cycle, but the reason it must be uncovered and shared is one of ethical imperatives. The widely-acknowledged ethical imperative to mitigate and adapt to climate change is predicated on altruism informed by the best empirical evidence of Earth's future in a changing climate (Hoegh-Guldberg et al., 2019). This implies a corollary imperative to perform work informing and underlying these efforts. It is hoped that the succeeding chapters incrementally yet meaningfully advance this goal.

Chapter 2

Total Alkalinity Measurement Using an Open-Source Platform

This chapter was published in May 2023 in the journal Limnology and Oceanography: Methods under a CC BY 4.0 license, and is reproduced without alteration of the text on the following pages.

Sandborn, Daniel E.[1,2*]; Minor, Elizabeth C.[2,3]; Hill, Craig[2,4]

[1]Water Resources Science Program, University of Minnesota, Saint Paul, MN USA 55108

[2]Large Lakes Observatory, University of Minnesota Duluth, Duluth MN, USA 55812

[3]Department of Chemistry and Biochemistry, University of Minnesota Duluth, Duluth MN, USA 55812

[4]Mechanical and Industrial Engineering Department, University of Minnesota Duluth, Duluth MN, USA 55812

Abstract

Total alkalinity is a pivotal water quality parameter dictating the response of natural waters to acid-base system perturbations such as ocean acidification

and acid mine drainage. Its value as a biogeochemical and ecological variable is enhanced not just by high measurement quality, but also by measurement accessibility. This research demonstrates an instrument that advances the accessibility of high-precision, high-accuracy total alkalinity measurement using open-source and low-cost instrumentation. Repeated testing of water samples from Lake Superior demonstrated a measurement precision (standard deviation or s.d.) of 3.0 µmol kg^{-1}. Analysis of standards and reference materials demonstrated an uncertainty of 5.3 µmol kg^{-1} (s.d.) as well as robustness to fresh- and saltwater matrices. This instrument adds to the wealth of inorganic carbon measurement technologies in marine and lacustrine settings and stands to enhance the ability of both communities to generate accurate and accessible measurements of total alkalinity.

Introduction

Alkalinity in natural waters plays a central role in determining pH and CO_2 flux, as well as biogeochemical responses to perturbations including acidification (Lerman & Stumm, 1989), whiting events (Morse et al., 2003; Müller et al., 2016), and phytoplankton blooms (Verspagen et al., 2014). As a concentration of buffering species, alkalinity moderates acidification, such that aquatic systems with relatively low alkalinity have been found to be more vulnerable to anthropogenic acidification (Shadwick et al., 2013). Total alkalinity (A_T) or its closely related analogues and proxies (acid neutralizing capacity, phenolphthalein alkalinity, carbonate alkalinity, and others) are frequently measured in scientific studies of lacustrine and marine systems and as water quality parameters by regulatory bodies. This study considers only total alkalinity, which was defined at a zero level of protons of pH = 4.5 by Dickson (1981) as:

$$A_T = [HCO_3^-] + 2[CO_3^{2-}] + [B(OH)^{4-}] + [OH^-] + [HPO_4^{2-}] + 2[PO_4^{3-}] + [SiO(OH)_3^-] + [HS^-] + 2[S^{2-}] + [NH_3] - [H^+] - [HSO_4^-] - [HF] - [H_3PO_4] \quad (2.1)$$

where concentration is in units of moles kg^{-1} to ensure independence from temperature and pressure. The zero proton level separates proton acceptors (with pK_a values above 4.5) from proton donors (with pK_a values below 4.5), and is an arbitrary but careful choice designed to enable use of A_T as a parameter of equilibrium calculations. In most fresh, oxic waters, this value is close to, *but not identical* to that of carbonate alkalinity (A_C), which excludes all terms except the inorganic carbon, proton, and hydroxide concentrations:

$$A_C = [HCO_3^-] + 2[CO_3^{2-}] + [OH^-] - [H^+] \quad (2.2)$$

In many systems including oceans, saline lakes, euxinic waters, highly eutrophic bodies, and polluted waters, the other terms in equation 1 play a larger relative role; in typical marine waters, A_C/A_T = ~93-95% (Zeebe & Wolf-Gladrow, 2001), while in Mono Lake, California, a similar value is obtained for its saline, alkali waters (Oxburgh et al., 1991). In these and similar systems, A_C is not a useful approximation of A_T for the purposes of equilibrium calculations and studies of carbon cycling. Methods of A_T determination involve titration of a water sample with a strong acid to a pH endpoint, which may be measured potentiometrically (Thompson & Anderson, 1940), spectrophotometrically (Yao & Byrne, 1998), or conductometrically (Park et al., 1963). An improvement upon single-endpoint titration methods is represented by Gran titration, which involves acidimetric titration of a water sample past a defined zero proton level to determine the endpoint titrant volume with a high degree of precision (Dyrssen & Sillén, 1967). This method has been further developed into open-

and closed-cell alkalinity titrations with standardized procedures adopted by the oceanographic community (Dickson et al., 2007). Further applications of alkalinity measurement may be found in management (Davis & Simon, 1995) and aquaculture (Somridhivej & Boyd, 2016) contexts, in which accurate measurements of alkalinity enhance its use as an environmental variable.

This study follows metrological convention in defining precision and uncertainty (Eurachem Working Group on Uncertainty in Chemical Measurement, 2012). Precision denotes the variability of measurements about a mean value caused by random error. Accuracy is a describes how a measurement differs from the true or accepted value. Measurement uncertainty is a function of both precision and accuracy, and is denoted by an interval within which the true value of a measurand may be found. In this work, precision is estimated using the standard deviation (s.d.) of replicate measurements, accuracy is determined empirically from a measurement of standards, and standard measurement uncertainty is expressed as the standard deviation of the difference between a measured value and the known value for a suite of standards. This study estimates measurement uncertainty via empirical means (Type A determination) and propagation of uncertainties (Type B determination) contributing to the uncertainty in A_T. Standard errors (s.e.) are also indicated where applicable, for example, describing uncertainty in linear regression coefficients.

Chemical oceanographers have advanced the definition and measurement of A_T rapidly over the past century in pursuit of minimizing measurement uncertainty, which makes A_T more valuable as a parameter of oceanic interaction with anthropogenic CO_2 and climate change (Dickson, 1992). Faced with continuing ocean acidification and a changing carbon cycle, the Global

Ocean Acidification Observing Network (GOA-ON) set A_T measurement uncertainty goals of 10 μmol kg^{-1} to observe short term ("weather"-level) variability and 2 μmol kg^{-1} to observe long-term ("climate"-level) variability (Newton et al., 2015). Such goals are made feasible by the oceanographic community's widespread utilization of standardized analytical procedures and reference materials (Bockmon & Dickson, 2015). No such goals, methodologies, or reference materials have been elaborated for inland waters, despite calls for improved measurement and standardization of inorganic carbon chemistry measurements in the face of climate change and other anthropogenic impacts on biogeochemical cycling (Phillips et al., 2015). It is not yet clear how suitable Certified Reference Materials of ocean water are as analytical standards for A_T determination in water with different major ion concentrations and ratios than the ocean, such as the freshwater Laurentian Great Lakes, or saline lakes such as Mono Lake.

A high level of analytical finesse is necessary not just for detecting trends and variability in A_T, but also for accuracy in equilibrium calculations (such as calcium carbonate saturation state) that use other measured inorganic carbon parameters including pH, dissolved inorganic carbon (DIC), and partial pressure of CO_2 (pCO_2). This latter parameter drives CO_2 flux, and pCO_2 can be estimated from equilibrium calculations or measured directly without propagating error from these calculations. That said, investigators of CO_2 dynamics in water must also measure at least two of the above-listed inorganic carbon parameters to constrain the inorganic carbon system and quantify the competing biogeochemical drivers of pCO_2. The potential pairs of parameters (pH-pCO_2, pH-A_T, pH-DIC, pCO_2- A_T, pCO_2-DIC, and DIC- A_T) vary in their suitability for equilibrium calculations as some pairs (e.g. pH-pCO_2) exhibit significant covariance when calculating other parameters (e.g. A_T) (Orr et al., 2018).

Estimations of pCO_2 as a function of measured A_T and pH have been used to constrain CO_2 flux models of lakes and rivers thanks to the great quantity

of measurements of these two parameters produced over the preceding century (Cole et al., 1994); however, insufficient attention has been given to the quality of these measurements until recently. Unpredictable biases in glass-electrode pH measurements caused by liquid junction potentials have been demonstrated to bias carbon cycling observations especially in low-ionic strength waters (Golub et al., 2017), leading to the nascent yet growing use of spectrophotometric pH determination in inland waters (Young et al., 2022). A similar advancement in technology is needed for lake and river alkalinity measurements, which are produced via a variety of methods with little published accounting for comparability, uncertainty, and bias. Biases in glass-electrode pH measurement are likely to introduce uncertainty into single-endpoint potentiometric A_T titrations; less so the titration calculation method described in Dickson (2007) and utilized in this work: the liquid junction potential is subsumed into the calculated Nernst equation E^* in a non-linear stepwise regression carried out in the software associated with the instrument described in this work.

Uncertainty (u) in pCO_2 propagated from pH and A_T can be visualized in an error space diagram (Orr et al. 2018), which demonstrates the dual dependence of pCO_2 uncertainty on both pH and A_T measurement quality (Figure 1). Uncertainty in inorganic carbonate parameters can be directly propagated from u(pH) and u(A_T) values using PyCO2SYS software in conditions like late summer surface waters of Lake Superior. GOA-ON guidelines for "climate" level uncertainty in pH and A_T (approximately ±0.003 and ±2 µmol kg^{-1}, respectively) produce a propagated uncertainty in pCO_2 of only ±2 µatm, while "weather" level uncertainty (approximately ±0.02 and ±10 µmol kg^{-1}) is associated with a pCO_2 uncertainty of ±17 µatm. This latter value is greater than the diel variability in pCO_2 observed in early summer 2001 in Lake Superior by Atilla et al. (2011), greater than the annual variability of Earth's atmospheric pCO_2, and greater than the mean daily difference between atmospheric and surface water pCO_2 in Lake Superior for large portions of the 5-year period

modeled by (Bennington et al., 2012). The uncertainty associated with measurements of pH and A_T in the Great Lakes National Program Office (GLNPO) biannual surveys of Laurentian Great Lakes is estimated as 0.2 pH units and 50 µmol kg^{-1} (Minor & Brinkley, 2022), which propagate to a pCO_2 uncertainty of ±163 µatm. Uncertainty in measured inorganic carbon parameters including pH and A_T has the potential to hinder carbon cycling studies and obscure biogeochemical observations of inland waters. This highlights the importance of development and application of improved measurement capabilities for both measurements.

A broad-scale challenge facing scientists and policymakers around the world is the disproportionate under-sampling of inorganic carbon parameters in inland waters and developing coastal regions (McDonald et al., 2013; Raymond et al., 2013). In reference to this gap in knowledge, Kim et al. (2022) emphasize the importance of adopting a combination of open-source and low-cost solutions to make high-quality carbon and greenhouse gas data generation more accessible. Development of new instruments, analyses, and technologies makes the prospect of a comprehensive global carbon observation network more feasible (Harmon, 2020). This study attempts to remove some of the barriers to increased quantity and quality of A_T measurements in heretofore under-sampled waters by demonstrating a new A_T instrument based upon proven open-cell titration techniques and optimized to utilize open-source and low-cost components for the measurement of A_T in inland or marine waters. This instrument expands the accessibility of vital water chemistry measurements and bridges a gap between oceanographic and limnological understanding of inorganic carbon chemistry.

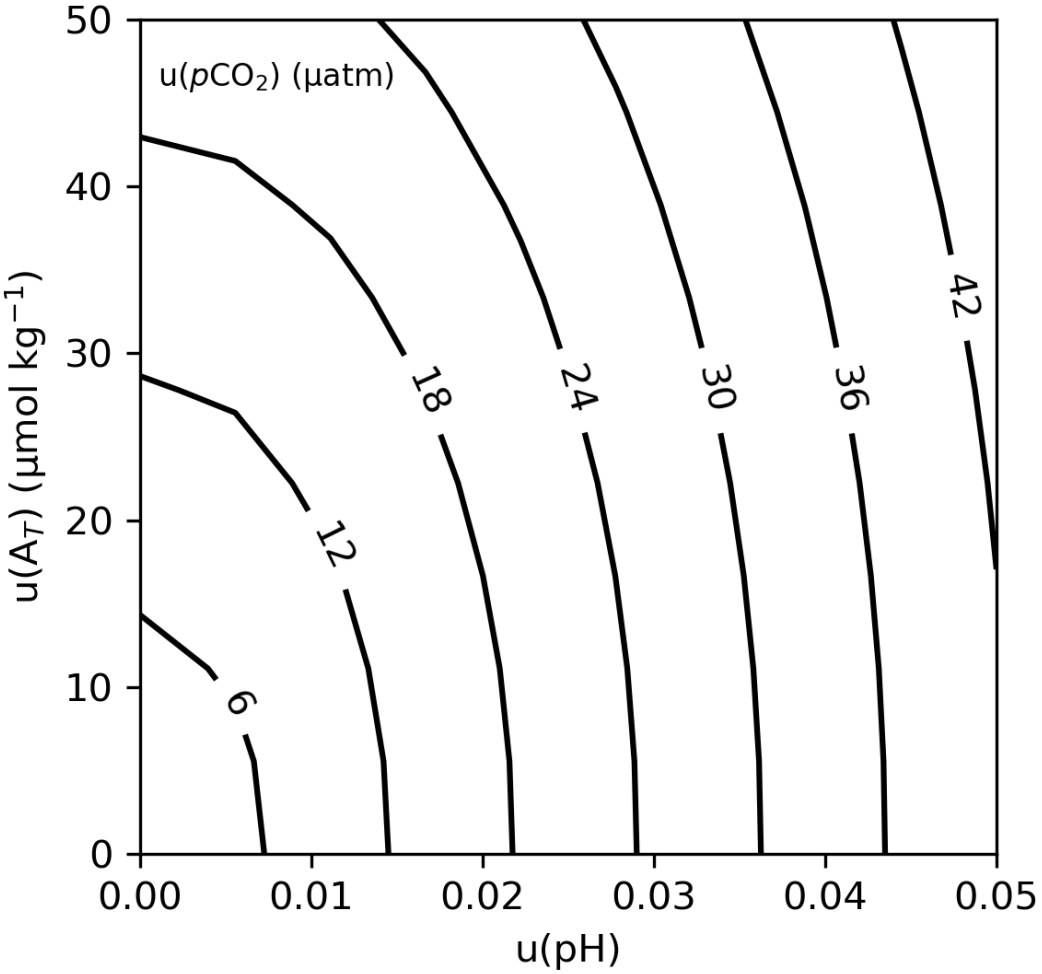

Figure 2.1: Error space diagram demonstrating the dependence of u(pCO_2), u(pH), and u(A_T) in conditions of late summer surface waters of lake superior. A_T = 850 µmol kg^{-1}, pH_T = 7.95, practical salinity = 0.05, and in-situ water temperature = 10 °C. u(pCO_2) is indicated with contours as µatm. The carbonate equilibrium constants of Cai et al. (1998) were used based upon improved system consistency in over-constrained measurements (Minor and Brinkley 2022). Code for this diagram is available in Data S1.

Materials

An instrument for A_T analysis was constructed from commercially available and custom-made components. Instrument functions are controlled by a Raspberry Pi microcomputer, an open-source platform for both user interface and data acquisition (Raspberry Pi 3 Model B+, Raspberry Pi Ltd.; any Raspberry Pi model currently available should work as long as a 40-pin GPIO header is available with standard Raspberry Pi pin layout). The Raspberry Pi is supplemented with a "DAQ HAT" (Data AcQuisition, Hardware Attached on Top) module (MCC128, Measurement Computing Corporation) that collects 16-bit analog voltage signals across a user-defined range (± 1 VDC for this case) from a combination glass pH electrode (Ecotrode Plus, Metrohm) via a custom amplifier circuit designed by Texas Instruments (Instruments, 2013) (Supporting Figure S1). While not a requirement, the custom circuit was assembled on a Raspberry Pi breadboard HAT, mounting it directly above the MCC-128 DAQ HAT and condensing the setup into a compact solution. The titration flask temperature is monitored with a thermistor (DS18B20, Analog Devices) with a resolution of 0.0625 °C and calibrated to an accuracy better than 0.1 °C against a NIST-calibrated thermistor (RBRsolo T, RBR). Titration is carried out in a 250 mL jacketed borosilicate beaker (Wilmad-LabGlass) atop a stir plate with a continuous flow of water controlled by a water pump (a gear drive pump was used here, but an aquarium pump suffices) recirculating water from a 20 L reservoir acting as a heat sink. This system employs titration flask temperature stabilization rather than temperature control in order to eliminate costly bath temperature controllers and increase affordability and accessibility of the method. Acid titrant is dispensed with a manual titrator (Hach 1690001) calibrated to have a volumetric uncertainty of 1.25 µL over a working range of 0-3 mL. The titrator calibration followed the procedure specified by Dickson et al. (2007). In brief, a calibration curve was constructed by dispensing volumes of

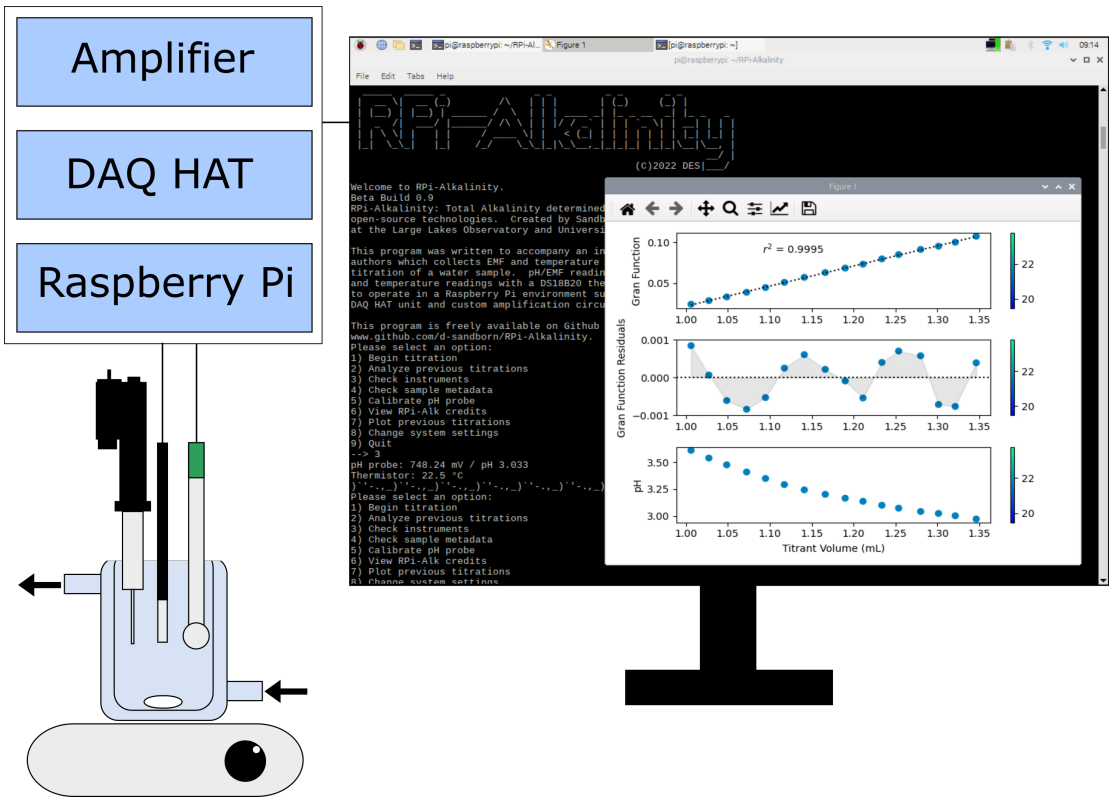

Figure 2.2: Instrument components, including stir plate, jacketed beaker, stir bar, digital titrator, thermistor, pH probe, Raspberry Pi microcomputer, DAQ HAT, amplifier circuit, and computer monitor. Water recirculation reservoir, aquarium pump, and electrical wiring not shown.

ultrapure water, then weighing the dispensate, correcting for buoyancy, converting to volume via a density factor, and plotting expected vs. calculated volume to obtain a type-I linear regression with an uncertainty of volume delivered indicated above as the standard error of the dependent variable. Instrument components are diagrammed in Figure 2 and pictured in Supporting Figure 2.

All instrument functions are coordinated within an open-source Python program, "RPi-Alkalinity" which provides a text-based interface for users to conduct titrations, plot results, and manipulate instrument settings. Details

of the RPi-Alkalinity software functions are provided in the Supporting Information section and on its Github page (*https://github.com/d-sandborn/RPi-Alkalinity*). While users need no knowledge of Python coding to operate the instrument (which enhances technical accessibility), this interface promotes the flexibility of this instrument to adaptations in hardware and procedures. Instrument components are listed in the Supporting Information to aid replication of this work. All numerical and statistical analyses in this study were completed using Python 3.7, utilizing various packages: PyCO2SYS (Humphreys et al. 2020), Calkulate (**humphreysmatthewp.CalkulateTotalAlkalinity2022**), Pandas (Team, 2023), Numpy (C. R. Harris et al., 2020), Scipy (Virtanen et al. 2020), Scikit-learn (Pedregosa et al. 2011), and Matplotlib (Hunter 2007).

Procedures

A_T sample analyses were conducted as similarly as possible to the specifications of the oceanographic community-standard Best Practices SOP 3b open-cell titration (Dickson et al. 2007). The pH electrode was calibrated in NIST-traceable pH buffers (Fisherbrand) before use, requiring 15 minutes at the beginning of every analysis session. A 100 g aliquot of water ±0.001 g was dispensed into a jacketed beaker, which was lightly covered with paraffin film to limit evaporation, connected to the recirculation system, and allowed to come to a stable temperature. The pH electrode and thermistor were suspended in the sample, and titration began after pH and temperature readings were stable within 0.002 pH units and 0.05 °C, which usually took 4 minutes for samples already at lab temperature. If probe readings were unstable before the next step, over- or under-acidification could result from measured pH biases. Hydrochloric acid (HCl) titrant solution was added to the sample until it reached a pH of 3.8, after which it was degassed by stirring vigorously with vortex-induced bubbles for 6 minutes. Following the degassing process, acid was added with a digital titrator (after ensuring the burette was free of bubbles)

in increments of at least 12.5 μL and the probe potential measured after each addition. The titration was concluded once the sample reached a pH of 3.0, typically after 15-20 additions. Overall titration time per sample was generally 15 minutes, and a 7.5 hour work period yielded 20-24 sample analyses with time allotted for instrument warmup, probe calibration, and standard analyses. The Python package Calkulate was integrated into the Raspberry Pi interface to calculate A_T from paired measurements of dispensed titrant volume, electrode potential, and sample temperature over the pH range 3.5-3.0. Analyses in which the titrand temperature changed by more than 0.05 °C were rejected.

Titrant acid was prepared by diluting concentrated HCl (Macron, ACS grade) with ultrapure (18.2 MΩ cm) water to approximately 0.25 mol kg^{-1}. Titrant acid was stored in acid-rinsed and combusted 500 mL borosilicate media bottles with minimal headspace, then sealed tightly using the bottles' plastic caps and electrical tape around the cap-bottle interface until use. This titrant was standardized against standard solutions as part of the Assessment section, below. The standard titration procedure was the same as the sample titration procedure, except that a 'provisional' A_T was calculated with an assumed [HCl] = 0.25 mol kg^{-1}. This provisional A_T informed the calculation of titrant concentration via the stepwise acid standardization procedure described in Best Practices SOP 3a (Dickson et al. 2007). A standardized acid titrant concentration of 0.25237 ± 0.00029 (s.d.) mol kg^{-1} was calculated for the batch of HCl used in all titrations in this study. The instrument software facilitates recalculation with revised sample masses, salinities, acid concentrations, and other parameters.

Three standard solutions were prepared or obtained: Na_2CO_3 solutions prepared from a solid, Certified Reference Material (CRM) batch 196 obtained from the laboratory of Dr. A. Dickson at Scripps Institute of Oceanography, and dilutions of CRM 196 in ultrapure water, referred to as "sub-reference materials" (SRM) hereafter. Na_2CO_3 standards were prepared with solid Na_2CO_3

(Nacalai Tesque special analytical grade) dried in an oven at 270 °C for 4 hours then cooled and stored in a desiccator. Additions of Na_2CO_3 with masses known to within 0.001 mg (measured on a Sartorius M2P microbalance) were dissolved into ultrapure water to create standard solutions with a mass ±0.001 g. SRM solutions were created with salinity and alkalinity gradients spanning 2-17 and 135-1119 µmol kg^{-1} respectively. Aliquots of CRM solution with masses known to within 0.01 g were dispensed into acid-rinsed and combusted 500 mL borosilicate flasks with Apiezon-greased ground glass closures. These solutions were diluted to approximately 300 g total volume with ultrapure water with a mass known to within 0.01 g. Alkalinity and salinity were assumed to act conservatively with respect to dilution with ultrapure water. Buoyancy corrections were made to all gravimetric measurements (G. L. Harris, 2019), which impacted standards' salinities negligibly and A_T concentrations by less than 2 µmol kg^{-1}. Certified Reference Materials were analyzed as provided.

All standards were analyzed identically, but two sets of thermodynamic constants were employed in the calculation of A_T from their titration curves via the Calkulate Python package. Freshwater samples and Na_2CO_3 standards were analyzed using the carbonic acid dissociation constants reported in Millero (1979), along with estimated concentrations of total sulfate (Chapra et al., 2012) and silicate (Johnson & Eisenreich, 1979) for Lake Superior samples. The carbonic acid dissociation constants of Lueker et al. (2000), the total borate to salinity ratio of K. Lee et al. (2010), the total fluoride to salinity ratio of Perez and Fraga (1987), and the bisulfate dissociation constants of Dickson (1990) were used for CRM and SRM analyses. It is crucial for the accuracy of A_T analysis that suitable sets of constants be used. To choose these, non-carbonate alkalinity must be parameterized by salinity for marine and estuarine samples (which are assumed to have relatively constant proportions of major ions to each other) or otherwise determined (for example, by ion chromatography) for inland waters that have varying major ion compositions.

Assessment

Measurement precision and accuracy were investigated with analyses of paired samples of water from Lake Superior. Measurement uncertainty was explored with standard solutions of Na_2CO_3, and ocean water CRM and SRM solutions. A subset of standard solutions was first used to standardize the acid titrant, after which measurement uncertainty was estimated from a separate subset as the standard deviation of errors from the differences between measured and known A_T. An additional estimate of measurement uncertainty was formulated from propagation of estimated uncertainties associated with the calculation of A_T.

An estimation of measurement precision was obtained via analysis of two sets of lake water samples. Simultaneously with underway pCO_2 measurement with a SuperCO2 system (Sunburst Sensors), water samples were collected from the surface of Lake Superior through the underway water system of the R/V Blue Heron at two open-water sites in the Western Arm in October 2021. Site 1 was at N 46° 53.25' W 91° 53.44' near the outlet of the French River into Lake Superior, while Site 2 was at N 46° 57.44' W 91° 53.06' in the Western Arm of Lake Superior. Water was dispensed from the underway system through silicone tubing into 500 mL borosilicate flasks with a full-bottle overflow, given a 5 mL headspace, and poisoned with 140 µL saturated $HgCl_2$ solution (Ricca). Practical salinity was calculated for each sample from underway thermosalinograph-measured conductivity and temperature using the equations of Hill et al. (1986). DIC was then analyzed via coulometry (CM150, UIC, Inc.). Particulate inorganic carbon is non-existent in Lake Superior, so we assume total inorganic carbon equals dissolved inorganic carbon. For alkalinity measurements, samples were filtered through a GF/F to remove particulate matter. From these two filtered water batches, 100-150 g aliquots (n=5) were

analyzed for A_T. Analysis results (Table 1) demonstrate a repeatable measurement precision of 3.0 µmol kg^{-1} (s.d.) associated with A_T analysis by this instrument. The measured values of A_T and DIC are indistinguishable from the values given in Chapra et al. (2012) and Zigah et al. (2011). The discrepancy between measured and observed pCO_2 is likely due in part to the influence of proton-binding dissolved organic matter on carbonate equilibria (Kuliński et al., 2014). If pCO_2 at Site 1 is recalculated in PyCO2SYS using an assumed DOM concentration of 100 µmol C L^{-1} (Zigah et al. 2011) along with assumptions of 14% of DOM providing acidic functional groups with a bulk K_a of 2.94 x 10^{-8} (both measured in the Baltic Sea by Kuliński et al. 2014) then the recalculated pCO_2 of 309 µatm closely approximates the measured value. We do not assert that DOM characteristics from the Baltic Sea accurately represent Lake Superior, but the increase in pCO_2 resulting from reasonable assumptions of DOM properties is of the right magnitude to explain the discrepancy between our measured and calculated values.

Standard A_T solutions were analyzed in approximately 100 g aliquots with masses ±0.001 g. 31 solutions provided 78 analyses (Supporting Table 3). These data were randomly split into a training dataset (75% or n = 58) and a testing dataset (25% or n = 20) with the "train_test_fit" function of the scikit-learn Python package in order to avoid overfitting and provide an unbiased assessment of uncertainty. The acid titrant concentration was calculated via a stepwise linear regression from the training dataset, as previously noted. This titrant standardization allowed calculation of measured A_T in the testing dataset, after which the measurement uncertainty was then evaluated as the standard deviation of the distribution of errors (measured A_T - known A_T) in the testing dataset. This train/test process was repeated 100 times with different random dataset splits to yield an acid titrant concentration of 0.25237 ± 0.00029 (s.d.) mol kg^{-1} and a mean measurement uncertainty of 6.3 µmol kg^{-1}. The calibrated A_T measurements of all 78 trials illustrate a standard curve (Figure 3a) with a type-I linear regression slope (1.000 ± 0.001 s.e.) insignificantly

Table 2.1: Measured or calculated inorganic carbon parameter values for samples of Lake Superior surface water. Measurement precision is denoted as ± SD for measured values. Uncertainty in calculated pCO_2 was propagated from uncertainty in DIC observed from measurements of reference materials and observed uncertainty in AT from this study. The carbonate equilibrium constants of Cai et al. (1998) were used based upon improved system consistency in over-constrained measurements (Minor and Brinkley 2022).

	SST (°C)	Practical salinity	A_T (µmol kg^{-1})	DIC (µmol kg^{-1})	pCO_2 (µatm) measured	pCO_2 (µatm) calculated
Site 1	14.3	0.045	838.9 ± 3.0 (n=5)	834.2 ± 2.2	312	149 ± 48
Site 2	15.3	0.046	839.9 ± 3.0 (n=5)	828.4 ± 2.8	275	108 ± 31

different from unity and an intercept (0.87 ± 1.2 s.e. μmol kg^{-1}) indistinguishable from the origin. This study employed standards spanning a wide range of A_T and salinity values encountered in inland, estuarine, and marine waters. Carbonate standards displayed a mean absolute error of 4.0 μmol kg^{-1}, while SRM solutions displayed a larger mean absolute error of 7.3 μmol kg^{-1}. SRM and CRM standards performed similarly to carbonate standards in terms of within-bottle measurement precision, but a higher mean absolute error was observed among SRM solutions (Figure 3b), which is likely due to random errors during preparation of the SRM solutions by dilution of the CRM standard. It is recommended that all A_T analyses with this instrument be referenced to and corrected with multiple reliable reference materials with a range of A_T exceeding that of samples. Titrant standardization and correction similar to the method used above is already built into the Calkulate package, which can calibrate titrant concentration and apply a correction based upon results of standard analyses.

Analysis of CRM batch 196 resulted in a mean measured A_T that was 4.5 μmol kg^{-1} (n = 2) higher than the certified value of 2215.32 μmol kg^{-1} and falls within the stated uncertainty of this instrument. While a difference in ionic strength existed between the reference material (I ~ 0.7 M) and titrant acid (I ~ 0.25 M), the decrease in ionic strength of the sample matrix does not exceed 1% for a typical titration of a 100 g sample. Nonlinearity in the Gran function due to changing proton activity and liquid junction potential might be expected given a change in titrand ionic strength during titration, but none was observed; R^2 values of linear regressions of the Gran function over titrant volume exceeded 0.9995 for each CRM and SRM analysis and were indistinguishable from the freshwater background standards in shape or behavior.

An additional estimation of measurement uncertainty associated with A_T analysis by this instrument can be developed via propagation of uncertainties through the alkalinity calculation software employed by this study. A_T is calculated as a function of titrant mass (m_{HCl}), titrant concentration (C_{HCl}), and

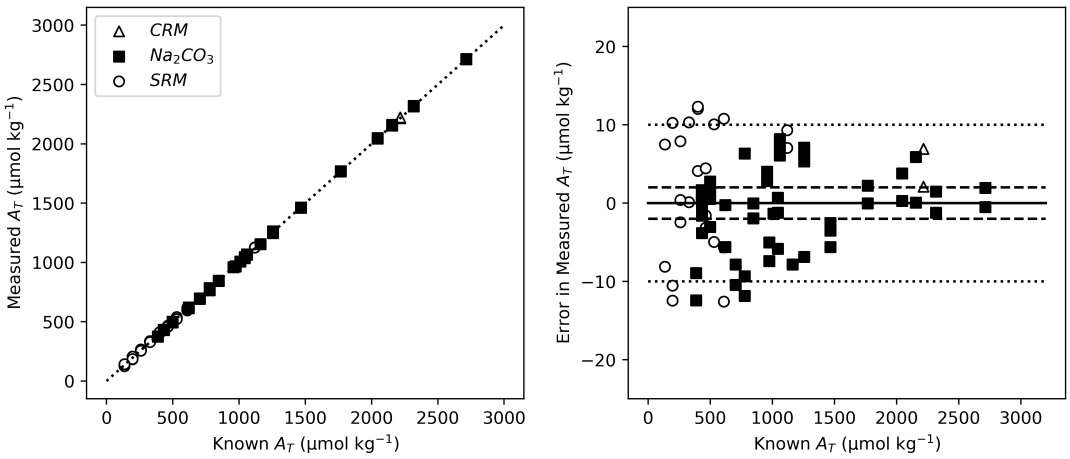

Figure 2.3: (a) Standard curve for total alkalinity analysis of standard solutions. A 1:1 line is plotted for reference, which is indistinguishable from the type-I linear regression of these data. (b) Control chart indicating error in measured AT for all standards. Dashed lines indicate the "climate" goal of ±2 µmol kg^{-1} for A_T measurement uncertainty indicated by Newton et al. (2015). Dotted lines indicate the "Weather" goal of ±10 µmol kg$_{-1}$.

sample mass (m_{Sample}) via the equation:

$$A_T = \frac{C_{HCl} \cdot m_{HCl}}{m_{Sample}} \tag{2.3}$$

Titrant concentration was determined to be 0.25237 ± 0.00029 (s.d.) mol kg^{-1} and sample masses are typically ~100 and measured to ±0.001 g. The uncertainty in acid titrant mass (u(m_{HCl})) can be determined from the uncertainty in the x-intercept of the Gran function. u(m_{HCl}) is dependent upon, e.g., temperature drift, probe calibration drift, electronic noise, and evaporation effects. Temperature drift is eliminated for this instrument by its recirculating system, while evaporative concentration is insignificant over the course of a 15-minute titration in laboratory conditions. Probe calibration drift is of little concern because the A_T calculation algorithm does not rely upon the pH probe's calibration, but instead calculates its standard potential for the Nernst equation for each titration (Dickson et al. 2007).

The noise-related uncertainty in the Gran function intercept can be estimated by a jackknife resampling analysis of titration curves, in which A_T analyses of standards and reference materials were each resampled by removing one point of their titration curves, after which A_T was calculated with the remainder of the curve. This was done repeatedly for 10 samples of Lake Superior water, once for each titration point, yielding a mean A_T and standard deviation for each, from which a corresponding mean titrant mass of titrant and its uncertainty was calculated. Jackknife resampling analysis code is provided in the Supplementary Information. The mean standard error of titrant mass was 32 µg.

Given the uncertainty estimates for each term in Equation 3, the resulting uncertainty in A_T (u(A_T)) in equations 5 and 6) can be propagated for a sample with a mass of 100.000 g and A_T = 1000. µmol kg^{-1}, requiring 0.400 g titrant (as described in equation 4):

$$0.001\frac{mol}{kg} = \frac{0.25051\frac{mol}{kg} \cdot 0.400\ g}{100.000\ g} \tag{2.4}$$

$$u\left(A_T\right) = A_T\sqrt{\left(\frac{u\left(C_{HCl}\right)}{C_{HCl}}\right)^2 + \left(\frac{u\left(m_{HCl}\right)}{m_{HCl}}\right)^2 + \left(\frac{u\left(m_{Sample}\right)}{m_{Sample}}\right)^2} \tag{2.5}$$

$$= 1000\frac{\mu mol}{kg}\sqrt{\left(\frac{0.00029\ mol\ kg^{-1}}{0.25237\ mol\ kg^{-1}}\right)^2 + \left(\frac{0.000032\ g}{0.39919\ g}\right)^2 + \left(\frac{0.001\ g}{100.000\ g}\right)^2} \tag{2.6}$$

The uncertainty in A_T is dominated by the titrant calibration, which displays a relative uncertainty of 1.1 ppt. The resulting propagated standard uncertainty of 1.2 µmol kg^{-1} approximates and contributes to the experimentally-determined uncertainty of 5.3 µmol kg^{-1}. We note that the Python package Calkulate also provides a function estimating uncertainty in measured A_T as a function of titration curve variability, but it does not propagate error from other sources.

The effect of decreasing sample volume can be estimated by recalculating Equation 6 for the volume of interest; halving the sample volume (as well as the required titrant) nearly doubles the propagated uncertainty to 1.9 µmol kg^{-1}. The pH probe used in this study has a geometry which limits sample volume to greater than approximately 60 mL. Smaller sample volumes may decrease the time needed for CO_2 sparging and amount of acid titrant consumed, leading to less time and expense per sample but larger measurement uncertainty.

Regardless of whether experimentally-determined or propagated measurement uncertainties are considered, the A_T instrument described in this study achieved a measurement uncertainty well within the "weather" goal of the GOA-ON framework, justifying its use to produce high-quality low-cost open source measurements of a crucial carbon cycling parameter.

Discussion

This study describes a development in the science of A_T measurement and demonstrates the capabilities of an instrument designed upon open-source, low-cost principles. Replicate A_T analysis of lake water samples indicated a measurement precision of 3.0 μmol kg^- (s.d.), while analyses of standard solutions demonstrated an empirical measurement uncertainty of 5.3 μmol kg^{-1} (s.d.). This result was reinforced by a propagation of uncertainties associated with titration, which resulted in a propagated measurement uncertainty of 1.2 μmol kg^{-1}. We hypothesize that the discrepancy between the two estimations of uncertainty is associated with uncertainty in the preparation of standard solutions. This instrument produces results falling within the GOA-ON "weather" goal (Newton et al. 2015), fulfilling the stated goal of this study to develop an instrument suitable for accurate and cost-effective measurement of A_T.

One significant decision made in this study when developing a method of alkalinity determination was the choice of titrant acid. Hydrochloric acid (HCl) was chosen for several reasons: it is a relatively cheap and accessible reagent; it is a strong acid; it is the standard titrant in oceanographic A_T titrations; and it carries no risk of forming protonated species at the pH range of Gran titration (pH 3.5-3.0). Sulfuric acid (H_2SO_4) has been used by many investigators as a titrant, but it forms bisulfate ions at pH levels approaching its second acid dissociation $pK_a \sim 2$. This carries some ramifications for single-point titrations (which cease at or around pH 4.5) and even more for Gran titrations, which continue to around pH 3.0. While knowledge of sulfate equilibrium constants allows accounting for bisulfate ion formation in A_T titration analysis software such as Calkulate (Humphreys, personal communication), the problem can be avoided altogether with HCl. Biases associated with sulfuric acid titrants are considered in greater detail in the Supporting Information. The acid titrant standardization determined in this study appeared to be stable over a period of months in tightly-sealed bottles, but this must be determined with periodic

standard analyses (preferably with A_T standards bracketing expected sample concentrations and spanning several orders of magnitude) and the use of control charts.

Apart from measurement uncertainties created by methodology, environmental factors such as high concentrations of particulate and dissolved organic matter also pose a challenge to high-quality measurements of inorganic carbon parameters. Proton-accepting behavior of organic matter has been termed "organic alkalinity" or "excess alkalinity". In sufficiently high concentrations, as in bog lakes, tannin-rich rivers, the coastal ocean, and eutrophic ponds, the interaction of organic matter with the inorganic carbon complex can shift concentrations and behavior of carbonate species, potentially biasing observations of carbon cycling (S. Liu et al., 2020; Sharp & Byrne, 2020). The instrument described by this study is expected to be similarly affected by organic alkalinity as the instrumentation of other investigators. While filtration methods have been suggested to deal with particulate matter (Bockmon & Dickson, 2014), dissolved organic alkalinity can be quantified by one of several titration methods (e.g. Cai et al., 1998), which this instrument could be readily modified to accomplish. The significance of organic alkalinity to the inorganic carbon complex varies according to its concentration and acid dissociation constant; organic matter with pKa values between 5 and 7 exert the greatest influence on measurements of A_T and may bias the fit of the Gran function (Sharp and Byrne 2020), potentially requiring novel titration curve fitting methods (Michałowski & Asuero, 2012) and/or inorganic carbon equilibria calculations explicitly accounting for organic alkalinity (B. Yang et al., 2015). Improved measurement of carbon cycling across the marine-lacustrine spectrum would be aided by further research should investigating the link between organic matter and the inorganic carbon complex.

The lake water samples collected in this study were poisoned with a saturated $HgCl_2$ solution because DIC and A_T were measured using the same sample. Poisoning samples for A_T analysis is unnecessary if samples are to be analyzed immediately (Dickson et al. 2007), or filtered to 0.45 μm and stored for up to 6 months (Moore et al., 2023; Mos et al., 2021). Hg (II) solutions present a significant safety and environmental hazard; the necessary precautions for handling Hg (II) may hinder the accessibility goals of this research, so it is recommended that investigators carefully consider their use of Hg-poisoned water samples. Investigators of freshwater and coastal processes should also note the potential for Hg-organic alkalinity complexes (Mos et al. 2021) to bias A_T measurements made using this instrument just as for other A_T instruments.

The open-source low-cost design philosophy guiding the development of this study ties into the broader picture of open science. This initiative seeks to increase access to both the means and products of scientific inquiry. While the drivers and outcomes of this initiative are discussed at length elsewhere (National Academies of Sciences, 2018), this research delivers a product in line with its goals. Open-source low-cost instrumentation such as that described here has the potential to increase the accessibility, reproducibility, and inclusivity of the means of science. These goals would be compromised in an instrument requiring frequent maintenance; however, over the course of more than 500 titrations in connection with this work and other projects not detailed here, the only regular maintenance involved refilling the acid burette and pH probe filling solution as per the manufacturer's instructions. The performance of the digital titrator was observed to decline over time. This titrator had been used in other studies for 4 to 5 years before being incorporated into this project and its plastic and metal parts developed considerable corrosion over the course of this work. We encourage analysts to investigate alternative titration systems or to assess the digital titrator regularly and plan on replacing it every few years.

Comments and Recommendations

The instrument described in this study represents a balancing act of open-source low-cost design philosophy and analytical finesse. The requirements of the inorganic carbon chemistry community along with the requirements of value-conscious investigators together inform the solution presented here. We assert that this instrument represents a substantial improvement upon alkalinity measurement systems available at a similar price point (see Supporting Information for comparison) in terms of accuracy and precision of results. It combines a proven procedure with innovative components to increase the accessibility of high-quality A_T measurement which will inform scientific knowledge of Earth's changing carbon cycle. The flexibility and open-source nature of this instrument's design opens the door to further modifications and improvements in instrumental analysis. The acknowledged labor-intensive operation of this instrument could be addressed with a syringe pump automated by the instrument's Raspberry Pi microcomputer in an exchange of labor for additional instrumentation cost and power needs. Portability and field deployment might also be accommodated by the design of this instrument. Given a 12 VDC power source (such as a conventional lead-acid battery) with the requisite connections to the Raspberry Pi microcomputer, stir plate, and aquarium pump, this instrument could conceivably measure A_T in the field and in locations without a reliable electricity supply. We note that the signal amplifier circuit appears to be sensitive to mechanical vibration and electromagnetic field sources, which may constrain its use to laboratory or field camp settings rather than shipboard deployments. Inclement weather could also pose a barrier to deployment if waterproofing measures are not taken as suggested in the Supporting Information. Future work should explore novel deployments as well as further cost-saving improvements to expand the horizons of inorganic carbonate parameter measurement.

This study used carbonate solutions, Certified Reference Materials, and dilutions of the latter to demonstrate the abilities of a novel A_T measurement system. These solutions were prepared with ionic strengths and A_T comparable to a wide range of inland, estuarine, and marine waters. All standard solutions were analyzed with identical methodologies, using a titrant with a deionized water background, and produced indistinguishable results when A_T (rather than merely carbonate alkalinity) was calculated with a non-linear stepwise regression method. We suggest that Certified Reference Materials may be used by investigators measuring A_T across the spectrum of fresh to saline waters *if and only if* they utilize a method capable of accounting for all A_T species; however, matching the solution matrix of standards and samples remains analytical best practice. Further research should explore the use of these and other alkalinity reference materials in a broader suite of natural waters.

This instrument recommends itself to practitioners of water analysis across the spectrum of research, regulatory, and academic sectors. It is not intended to displace the existing suite of A_T measurement systems, but to supplement and expand accurate analyses to locations and labs that may not have access to these systems. Additionally, we suggest that the analytical practices of the marine chemistry community be applied to limnological practice after suitable modification. The adoption of standardized procedures, reference materials, and high-accuracy analyses have great potential to deliver high-quality data on total alkalinity and across inorganic carbon parameters, necessary to disentangle the effects of various drivers of change for inland waters, which include land-use changes, invasive species effects, changing hydrology and changing temperature ranges and ice phenology.

Acknowledgements

The authors wish to thank Lucas Busta, Emily Bockmon, and Payton K. Kittaka for their invaluable suggestions benefitting this work. Thanks are also

due to the Associate Editor and three anonymous reviewers who provided constructive comments and suggestions that improved this publication. This research was supported by a Grant-in-Aid from the University of Minnesota Office of the Vice President for Research to E.C.M. D.E.S. was supported by a Graduate Research Fellowship from the Cooperative Institute for Great Lakes Research. Funding for this project was provided by the University of Michigan Cooperative Institute for Great Lakes Research (CIGLR), through the National Oceanic and Atmospheric Administration (NOAA) Cooperative Agreement NA17OAR4320152. This is CIGLR contribution 1207.

Author Contributions

D.E. Sandborn conceived of the study. D.E. Sandborn and C. Hill designed and assembled the instrument. D.E. Sandborn tested the instrument with input from E.C. Minor. E.C. Minor provided scientific and editorial recommendations. D.E. Sandborn wrote the manuscript with input from all authors.

Data Availability Statement

The data that support the findings of this study are available in the Supporting Information, and the code that supports replication of this instrument is openly available on Zenodo at DOI: 10.5281/zenodo.7864447, and on Github at *https://github.com/d-sandborn/RPi-Alkalinity*.

Chapter 3

Underway pCO_2 surveys unravel CO_2 invasion of Lake Superior from seasonal variability

D. E. Sandborn[1,2]; E. C. Minor[1,3]

This chapter was published in 2024 in the Journal of Geophysical Research: Biogeosciences under a CC BY 4.0 license, and is reproduced without alteration to the text on the following pages.

[1]Water Resources Science Program, University of Minnesota, Saint Paul, MN USA 55108

[2]Large Lakes Observatory, University of Minnesota Duluth, Duluth MN, USA 55812

[3]Department of Chemistry and Biochemistry, University of Minnesota Duluth, Duluth MN, USA 55812

Key Points

- Underway pCO_2 was measured in Lake Superior from 2019 to 2022 to form the first multi-year pCO_2 time series in the Laurentian Great Lakes.

- The seasonal pCO_2 cycle showed competition of thermal and non-thermal drivers and spatial heterogeneity associated with riverine influence.
- Superior's CO_2 reservoir equilibrated with the atmosphere repeatedly, indicating increasing sea surface pCO_2 with rising atmospheric CO_2.

Abstract

This study observed seasonal trends and inferred drivers of CO_2 biogeochemistry at the air-water interface of Lake Superior. Underway carbon dioxide partial pressure (pCO_2) was measured in surface water during 69 transects spanning ice free seasons of 2019-2022. These data comprise the first multiannual pCO_2 time series in the Laurentian Great Lakes. Surface water pCO_2 was closely tied to increasing atmospheric pCO_2 over a 100-day CO_2 equilibration timescale, while seasonal variability was controlled equally by thermal and non-thermal drivers during the ice-free season. Comparison to previous modeling efforts indicated that Lake Superior surface pCO_2 increased with two decades of rising atmospheric CO_2. Spatial heterogeneity in CO_2 dynamics was highlighted by a conductivity-based delineation of "riverine" and "pelagic" regimes, each of which was associated with net CO_2 influx over Julian days 100-300 on the order of 25 Gmol C. These findings refine previous estimates of Lake Superior C fluxes, support predictions of anthropogenic CO_2 invasion during the ice-free season, point to new observation strategies for large lakes, and highlight an urgent need for studies of changes to lacustrine C cycling.

Plain Language Summary

Carbon dioxide gas concentrations were measured in surface waters of Lake Superior for four years, forming the first multi-year dataset of direct observations

of carbon dioxide gas concentration in the Laurentian Great Lakes. Lake Superior's surface carbon dioxide concentration was closely tied during the ice-free season to that of the atmosphere on time scales longer than one year. Seasonal variations in carbon dioxide concentration were driven by water temperature, biological activity, river influence, and gas exchange with the atmosphere. Lake Superior released and absorbed carbon dioxide cyclically at different times during the year, absorbing more than it released from April to November. The carbon dioxide gas concentration in mixing surface waters was observed to have increased over a multi-decade period with that of the atmosphere. This work improves the scientific understanding of carbon cycling in Lake Superior and advances techniques for carbon cycle observation and modeling of other lakes in a time of large-scale changes to surface water CO_2 cycling.

Introduction

Measurements of carbon cycling in the Earth's hydrosphere are central to understanding global biogeochemical cycling and responses to perturbation (Le Quéré et al., 2013). Continuing anthropogenic emissions of carbon dioxide (CO_2) are increasing atmospheric concentrations at an unprecedented rate, which may force changes in carbonate equilibria in the oceans (R. Feely et al., 2001), in soils (Oh & Richter, 2004), in rivers (Raymond & Hamilton, 2018), and in lakes (Alin & Johnson, 2007).

Many studies of CO_2 dynamics in inland waters collect and analyze discrete water samples for inorganic C parameters including pH, dissolved inorganic carbon (DIC), total alkalinity (A_T), and partial pressure of carbon dioxide (pCO_2) (Cole et al., 1994). Direct measurements of CO_2 flux across the air-water interface are also made via floating chamber or eddy covariance methods (Podgrajsek et al., 2014). Calculation of one inorganic C parameter from two others remains fraught with uncertainty due to ongoing challenges associated with measurement and equilibrium calculations in low-ionic strength

waters (S. Liu et al., 2020; Minor & Brinkley, 2022; Young et al., 2022). Furthermore, constructing time series of discrete water chemistry measurements is time- and labor-intensive and may not resolve inorganic C cycling in water bodies with high spatial or temporal variability (Schilder et al., 2013). To bridge these gaps in observational capabilities, instruments measuring inorganic C parameters continuously or autonomously have been developed and deployed in aquatic systems spanning the lacustrine-marine spectrum (Bushinsky et al., 2019; Lynch et al., 2010). Recent years have seen applications of pH and pCO_2 underway sensors that perform with uncertainties similar to those of discrete sample measurements (Ma et al., 2019; Takeshita et al., 2018).

Inorganic C chemistry remains less-studied in inland waters than in marine systems (Phillips et al., 2015), due in part to high heterogeneity within and among water bodies, each of which may require divergent approaches to inorganic C cycling analysis depending on hydrology, climate, or terrestrial influence. The approaches used in marine systems (where variations in inorganic C parameters are often smaller and require more precise measurements) might be applied to large lakes and provide new insights into aquatic systems. In this way, large lakes may serve as stepping-stones for application and further development of oceanographic chemical techniques in inland waters. Their great volume and relatively small terrestrial influences lend them a more constant chemistry and physics than their smaller peers. The largest of lakes share with oceans similar biogeochemical features and relative importance to local and global biogeochemical cycling (Sterner et al., 2017). On the other hand, large lakes respond more rapidly than the global ocean to perturbation. Their hydrologic residence times (c. 190 years for Lake Superior) are shorter than that of the global ocean (millennia). Holomictic lakes experience full water column mixing at least annually, which represents a homogenizing driver not observed in oceans. For these reasons, large lakes can act as test systems for investigations of environmental variables, with responses occurring on more accessible spatial and temporal scales for research than the global ocean Sterner, 2021.

The Laurentian Great Lakes lie on the border of the United States of America and Canada and within the historical and contemporary lands of Native American and First Nations. They constitute the largest contiguous aquatic ecosystem on Earth (Wetzel, 2001), yet C cycling in the Great Lakes is not well-understood (Minor & Oyler, 2021). It remains unclear to what extent the Great Lakes are net sources or sinks of CO_2 to the atmosphere (McDonald et al., 2013; Urban & Desai, 2009). Alin and Johnson (2007) concluded that they are annual net CO_2 sources, but Cotner et al. (2004) and Urban (2005) highlighted apparently unbalanced C budgets in Lake Superior, the best-studied of the five lakes in terms of carbon cycling. Bennington et al. (2012) and Atilla et al. (2011) noted that studies of CO_2 cycling in Lake Superior have been biased by sparse observations restricted to the ice-free period (generally April-November ± 2 months), and could not close the cycle by modeling all C inputs and outputs. These pioneering studies were confounded by observations of inorganic C cycling that were sparse, irregular or unrepresentative of the lakes as a whole. This situation is similar to that of the Southern Ocean or South Pacific Ocean, in which limited observation hindered attempts to constrain biogeochemical budgets (Takahashi et al., 2009). Large lakes functioning as "sentinels, integrators, and regulators of climate change" (Williamson et al., 2009) exert significant influence on regional and global C budgets (Cole, 2013a) and demand more detailed study.

This research focuses on surface water pCO_2 measurements to illustrate the C cycle of Lake Superior in unprecedented spatial and temporal detail. pCO_2 in water responds to physical (temperature, pressure, salinity), chemical (pH, DIC, A_T, $CaCO_3$ dissolution/precipitation), and biological (production, respiration) drivers (Zeebe & Wolf-Gladrow, 2001), such that a comprehensive understanding of pCO_2 variability sheds light on a suite of biogeochemical functions.

As a direct driver of CO_2 flux across the air/water interface, pCO_2 in surface waters acts as an important factor of atmospheric CO_2 accumulation. The

surface ocean has maintained a lower mean pCO_2 than the atmosphere in recent decades, resulting in about a quarter of annual anthropogenic CO_2 emissions being absorbed by the ocean each year (Friedlingstein et al., 2022). In contrast, many inland waters are supersaturated with CO_2 recycled from the terrestrial C sink and emitted to the atmosphere or exported to the coastal ocean (Cole et al., 1994; Raymond et al., 2013). These marine and lacustrine C cycles are linked by C imports and exports along the terrestrial-aquatic continuum which form important elements of comprehensive global C budgets, especially in light of their perturbation by anthropogenic global change (Tian et al., 2023). Accurate predictions of climate change and mitigation efforts require an improved understanding of surface waters as sources and sinks of CO_2 and other greenhouse gases (Cavallaro et al., 2018).

Lake Superior is characterized by a small catchment-to-surface area ratio of 1.55 (Urban, 2005) and is underlain by a weathering-resistant igneous mineralogy leading to exceptionally dilute, soft, and carbonate-poor water chemistry (Weiler, 1978). Its primary production is dominated by a large proportion of picoplankton and nanoplankton, with identifiable phytoplankton species including an abundance of diatoms and small flagellated plankton (Fahnenstiel et al., 1986; Ivanikova et al., 2007; Kovalenko et al., 2019), and has been classified between oligo- and mesotrophic, with relatively high transparency and low phosphorus concentration (Sterner, 2010). Multiple aspects of the Lake Superior's limnology are changing: its waters are warming faster than the overlying atmosphere (Austin & Colman, 2008), the concentration of most of its major ions is increasing (Chapra et al., 2012), and its nutrient stoichiometry is shifting as the result of long-term nitrate accumulation (Sterner, 2011). Interannual trends in A_T, pH, and pCO_2 have proven difficult to constrain due to covariation with lake level, influence from Dreissenid calcification in tributaries, large measurement uncertainty, and spatial heterogeneity (Minor & Brinkley, 2022). These poorly-understood trends, and especially the potential for acidification (Phillips et al., 2015), contribute to the need for a sustained campaign of

spatially- and temporally-comprehensive measurements of the inorganic carbon system in Lake Superior.

In this work, underway pCO_2 measurements gathered by instrumentation aboard *RV Blue Heron* from cruises of opportunity during four consecutive field seasons (April-November 2019-2022) provided a survey of unprecedented spatial and temporal scope describing inorganic C cycling drivers and variability in Earth's largest lake by surface area. This information was used to illustrate ice-free season surface pCO_2 variability, infer trends in pCO_2 and CO_2 flux over space and time and establish the interplay of thermal and non-thermal drivers of pCO_2. The results demonstrate a pathway towards comprehensive CO_2 budgets for the Laurentian Great Lakes via novel observation strategies and improved modeling efforts.

3.1 Methods

Underway instrument datasets from 69 transects of *RV Blue Heron* were compiled. These efforts included single-day endeavors near the vessel's home port of Duluth Minnesota, as well as multi-week transects across the Laurentian Great Lakes (Figure 3.1). Water was directed from the ship's water intake line at 2 m depth through a suite of sensors measuring parameters including dry molar fraction of carbon dioxide (xCO_2), sea surface temperature (SST), and sea surface conductivity. These were combined with wind velocity, barometric pressure, and air temperature collected from an onboard meteorological station. The multi-year span considered in this study permitted evaluation of interannual variability in inorganic C biogeochemistry despite limited cruises in 2020 and 2021 due to challenges associated with the Coronavirus pandemic.

xCO_2 was measured in water from the underway system at 2 second intervals by a Sunburst Sensors SuperCO_2 instrument equipped with a dual showerhead equilibrator which recirculates air at a rate of 1 L min^{-1} through a two-stage system as described by Pierrot et al. (2009), which has been found to have

a response time around 2 minutes (Evans et al., 2019). The sample water flow rate c. 4 L min^{-1} through two c. 2 L showerhead equilibrators ensured complete equilibration with air-sea xCO_2 gradients never exceeding the manufacturer's guideline of 1000 ppm. This instrument has proven robust to wide ranges of temperature and sample pH in intercomparison experiments in which it was used as the reference instrument (Shangguan et al., 2022) and its equilibrator design aligns with established best practices (Pierrot et al., 2009).

Measurements from four standard gases with CO_2 concentrations between 0 and 1018 ppm were performed every 2 hours (Figure S1) and the 60 seconds before and after calibration removed from the time series to prevent memory effects. The slope and intercept values from a type-I linear regression of measured vs. standard xCO_2 were used to correct surface water xCO_2 before conversion to pCO_2 (Equation 3.1) A nearly-identical instrument demonstrated pCO_2 measurement uncertainty of ± 5 µatm (DeGrandpre et al., 2020). SST and conductivity were obtained from a SBE21 thermosalinograph every 2 seconds. Conductivity was converted to practical salinity using the equations of Hill et al. (1986). Previous work found such an extrapolation from seawater salinity to Lake Superior's ionic strength fit overconstrained carbonate system equilibrium calculations better than assuming freshwater carbonate dissociation constants (Minor & Brinkley, 2022). Salinity as used in this work cannot be quantitatively compared to ocean salinity due to differences in major ion ratios which invalidate the seawater equation of state. Salinity is used here only for equilibrium calculations.

Wind velocity was measured with a Young 05106 wind monitor on a mast 10 meters from the sea surface. Air temperature was obtained from a Young 41372VC thermometer. It was assumed that mast-measured windspeed (corrected for vessel velocity) approximated neutral wind speed at 10 meters (U_{10n}) sufficiently well for the determination of instantaneous CO_2 flux. Measured pCO_2 and calculated CO_2 flux were averaged for each day of each transect in

0.01° x 0.01° boxes (approximately 1.1 x 0.8 km at the latitude of Lake Superior) to normalize distributions of pCO_2 and CO_2 flux on an areal basis, prevent biases due to vessel idling, and provide observations on a scale greater than the equilibration time of the instrument.

pCO_2 was calculated as a product of ambient atmospheric pressure (p_{atm}, in atm) and xCO_2 both measured by the SuperCO_2 instrument and corrected for water vapor partial pressure (p_{H_2O}, in atm) calculated as a function of temperature assuming saturation (Dickson et al., 2007):

$$pCO_2 = xCO_2 \cdot \left(p_{atm} - p_{H_2O}\right) \tag{3.1}$$

pCO_2 was also corrected for the temperature difference between SST and the equilibrator as in Pierrot et al. (2009). CO_2 flux (in mol m^{-2} d^{-1}) was calculated as the difference between aqueous and atmospheric pCO_2, multiplied by the gas transfer velocity (k, in m d^{-1}), a function of the unitless Schmidt number Sc (Ho et al., 2006), mean squared neutral wind speed at 10 meters above the sea surface ($< U_{10\,n}^2 >$, in $m^2 s^{-2}$), and K_o, the solubility of CO_2 in water (in mol m^{-3} atm^{-1}) (R. Weiss, 1974). Positive values of CO_2 flux indicate sea-to-air efflux.

$$CO_2 \text{ Flux} = kK_o\left(pCO_{2\,aq} - pCO_{2\,atm}\right) \tag{3.2}$$

$$k = 0.266 < U_{10}^2 > \left(\frac{Sc}{600}\right)^{-0.5} \tag{3.3}$$

We compared two sources of atmospheric CO_2 concentrations for calculation of CO_2 flux: underway-measured atmospheric pCO_2 measured every 2 hours by the SuperCO2 instrument and atmospheric pCO_2 as measured at the WLEF/Park Falls Wisconsin tower (Desai, 2022). The WLEF/Park Falls time series was chosen for CO_2 air-sea flux calculations, as detailed in the Results.

There is considerable disagreement among parameterizations and measurements of k in inland waters. Varying empirical formulations of the wind speed dependence of k have been derived and applied in lakes (e.g.)Atilla et al.,

2011; Cole and Caraco, 1998; Oxburgh et al., 1991; Wanninkhof, 1992), while comparison with eddy covariance-derived values of k indicates that heat and water fluxes also play a role in determining gas flux rates in lakes (Heiskanen et al., 2014). The k parameterization in this study (Ho et al., 2006) was chosen on the grounds that Lake Superior can be understood similarly to marine environments, with a high range of wind speeds and large fetch which merit the quadratic wind dependence discussed by Wanninkhof (1992) (D. Ho, personal communication). The extent to which this parameterization may bias calculated CO_2 flux in Lake Superior is unclear, and may benefit from direct measurement of k in future studies.

Calculations were completed with Python 3.8, using Pandas Reback et al., 2022 for data structure manipulation, SciPy Virtanen et al., 2020 and Statsmodels Seabold and Perktold, 2010 for regression and statistical analysis, Numpy C. R. Harris et al., 2020 for array computation, PyCO2SYS Humphreys et al., 2020 for CO_2 system calculations, GSW-Python Firing et al., 2021 for salinity conversions, and Matplotlib Hunter, 2007 and Seaborn Waskom, 2021 for visualization.

3.2 Results

More than 6 x 10^6 measurements of xCO_2 in Lake Superior surface waters were assembled into a pCO_2 and CO_2 flux timeseries. These data spanned the lake's most significant hydrological regions, including shallow coastal zones, deep (maximum 406 m) waters, riverine outlets, and regions bordering significant human development (Figure 3.1). The most heavily-observed regions included the far western arm of Lake Superior and a cross-lake transect from Duluth to Sault Ste. Marie. Binning of pCO_2 and CO_2 flux data by grouping observations by date and 0.01° boxes yielded 1.3 x 10^4 observations.

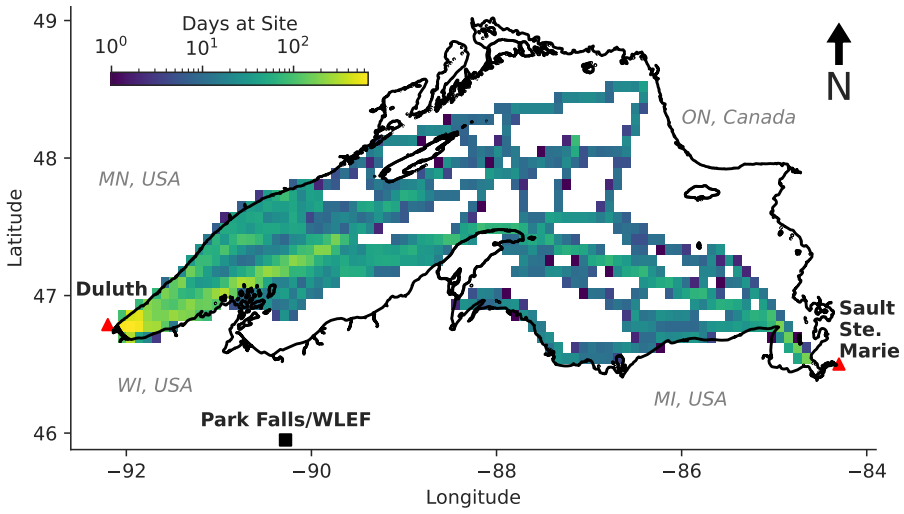

Figure 3.1: Underway measurement density transects 2019-2022, visualized as the number of occupations of approximately 5 km squares. The number of days of observation ranged from 0 to nearly 600. The cities of Duluth and Sault Ste. Marie, between which multi-lake transects traverse, are indicated by red triangles. The Park Falls/WLEF tower is denoted by a black square.

3.2.1 Underway Timeseries Overview

Mean observed SST was 11.4 °C with a median of 12.7 °C. SST varied widely among and within cruises, ranging from a maximum of 23.5 °C in July 2019 near the center of the Far Western Arm to a minimum of 0.45 °C in April 2022 in the plume of the St. Louis River Estuary. Specific conductance (SC) ranged from a near-constant 99.44 µS cm^{-1} in unstratified offshore waters to values exceeding 170 µS cm^{-1} in the plume of the St. Louis River Estuary, displaying a mean of 99.6 µS cm^{-1}, a median of 99.3 µS cm^{-1}, and a standard deviation of 2.2 µS cm^{-1}. The timing of thermal stratification in Lake Superior varied widely among locations and years (Austin et al., 2022), so observations within 0.5 °C of the temperature of maximum density of freshwater (3.98 °C) were designated as unstratified. Stratification onset occurred between late June and August, depending on year and location (Figure 3.2a). Interannual meteorological variability, including the previous winter's thermal state, exerts considerable influence on Lake Superior's summer stratification (Austin & Allen, 2011), as indicated by the historically late stratification of Lake Superior in August 2022 (J. Austin, personal communication), as well as maximum winter ice coverage which varied from 94.9% in early 2019 to 22.6%, 50.6%, and 79.5% the following years (T.-Y. Yang et al., 2020).

Surface-water DIC and pH (free scale) were calculated from measured pCO_2, SST, and an assumed A_T of 840 µmol kg^{-1} (Figure 3.2d-e) with PyCO$_2$SYS, using the carbonate constants of Waters et al. (2014). A_T is given as µmoles of hydrogen ion equivalent kg^{-1} in this work, and is largely invariant in Lake Superior (Minor and Brinkley 2022, Sandborn et al. 2023) except in regions with significant terrestrial influence; no A_T-conductivity relationship for Lake Superior has been published, so A_T was not estimated from underway data.

Calculated pH_{free} exhibited a mean of 8.075 and standard deviation of 0.093, while calculated DIC exhibited a mean of 855.0 µmol kg^{-1} and standard deviation of 8.8 µmol kg^{-1}. Seasonal variation in DIC was evident as a summertime

decrease on the order of 20 µmol kg^{-1}, followed by an autumn increase of c. 10 µmol kg^{-1}. The pH_{free} distribution fell within the range of values given in Minor and Brinkley (2022), while the mean calculated DIC was 10-40 µmol kg^{-1} higher than observations given in Zigah et al. (2011) and Sandborn et al. (2023). The discrepancy may be due to interannual DIC increases, sampling bias in the latter two studies favoring regions or periods of lower DIC, contribution of organic alkalinity, uncertainty associated with equilibrium calculation, or error propagated from uncertainty in A_T. Minor et al. (2019) Reported mean annual AT values for the 1996-2014 EPA GLNPO timeseries ranging from approximately 810 to 870 µmol kg-1. Propagation of this range in AT through our pCO_2 observations yielded a calculated DIC range of 825-885 µmol kg^{-1} and a calculated pH range of 8.059-8.090. This sensitivity analysis produced DIC and pH ranges overlapping with previous measurements, such that some of the discrepancy between calculated and previously-reported DIC could be due to error in assumed AT.

3.2.2 Atmospheric CO_2

Daily mean shipboard atmospheric xCO_2 varied seasonally in concert with the CO_2 timeseries observed at the Park Falls/WLEF tower (Desai, 2022) (Figure S5), approximately 80 km south of Chequamegon Bay, Wisconsin. Both series displayed a larger annual variability and a phase shift from the Mauna Loa CO_2 time series (Keeling & Keeling, 2017). No systematic biases in atmospheric CO_2 concentration were observed between the underway and Park Falls/WLEF time series within years, yet the underway atmospheric signal displayed a much larger variability. Several anomalies emerged in the underway atmospheric data. Atmospheric xCO_2 measurements in several cruises were consistently higher than expected despite nominal measurements of standard gases and sea surface xCO_2. These cruises included extended periods of idling, and presumably detection of exhaust CO_2 by the underway system. In another

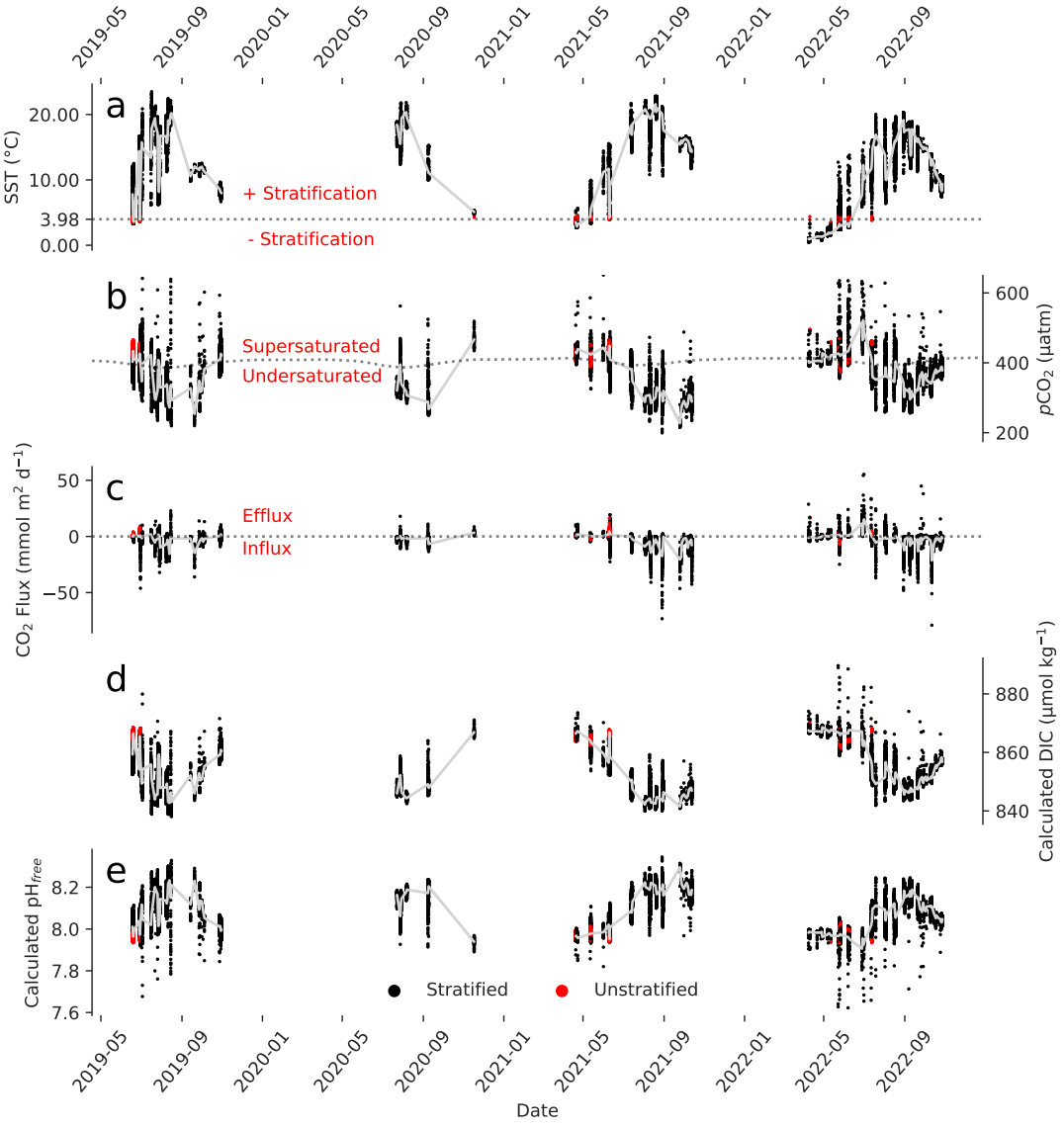

Figure 3.2: Sea surface temperature, pCO_2, calculated CO_2 flux, calculated DIC, and calculated pH_{free} observed in 0.01° boxes on transects of Lake Superior, 2019-2022. Median values for each day of observation are connected by a grey line. **a.** The 3.98 °C temperature of maximum density is indicated by a dotted line, along which lie unstratified conditions, highlighted in red. Depressed springtime surface temperatures of 2022 are visible as a delayed warming trend. **b.** The Park Falls/WLEF time series is displayed as a dotted line separating observations of CO_2 supersaturation and undersaturation. **c.** The division of CO_2 efflux vs. influx is indicated by a dotted line. **d.** DIC as calculated from pCO_2 and assumed A_T = 840 µmol kg^{-1}. **e.** pH (free scale) as calculated from pCO_2 and assumed A_T = 840 µmol kg^{-1}.

two cruises in September 2022, atmospheric (but not sea surface) xCO_2 was depressed over a period of weeks for reasons related to a filter on the air inlet. Due to these discrepancies, we chose to use daily means of nearby Park Falls/WLEF tower hourly measurements of atmospheric xCO_2 with the expectation of a well-mixed atmosphere over these scales. The occurrence of most atmospheric underway xCO_2 measurements within a close approximation of the Park Falls/WLEF timeseries validated this expectation. If atmospheric CO_2 were not well-mixed on these scales, an uncertain amount of error in calculated CO_2 flux could result, which further work should explore.

3.2.3 Wind Speed

Wind speed observed on Lake Superior (corrected for direction of travel) exhibited a skewed unimodal distribution with a peak at 4.5 m s^{-1} (Figure S2). Some bias may have been incurred by intentional planning of transects around inclement weather and targeting the ice-free season, so it was unclear how well these transects represented the true distribution of wind velocity above Lake Superior. The underway-observed wind speed distribution in 2020 stood out from other years with a lower and irregular distribution; these transects were limited in time and space (Figure S1) and are less likely to represent the true distribution of wind speed over Lake Superior. Comparison of the underway wind speed distributions with those measured offshore at the Stannard Rock Lighthouse over the same periods (Figure S3) indicates that the underway-observed wind speed distribution closely approximated that of the whole season.

The wind speed distribution peaks observed from either source were lower than the global U_{10n} distribution peak of approximately 7 m s^{-1} in Yang et al. (2022), which may imply an underestimation of CO_2 flux as parameterized by dual-tracer models as in this research. The present scarcity of research on gas

flux parameterization validity in large lake systems for which size, morphometry, and variable winds greatly influence gas flux magnitude and timing (Perolo et al., 2021; Schilder et al., 2013) does not yet allow exploration of similar biases in this research.

Gas transfer velocities (k) calculated from the underway wind distribution displayed a mean of 1.6 m d^{-1}, about half the mean ocean value of 3.3 m d^{-1} given by Broecker and Peng (1982) and supported by revised gas transfer velocity parameterizations (e.g. Ho et al., 2006; Wanninkhof, 2014). Given k, along with the 147 m mean depth of Lake Superior (Fuller & Shear, 1995), DIC and aqueous CO_2 concentration [CO_2^*] (both in molal units), and Revelle Factor (RF, the unitless ratio of relative change in [CO_2^*] per relative change in DIC), all produced via PyCO2SYS equilibrium calculation, the characteristic timescale, or e-folding time, of CO_2 equilibration in Lake Superior (τ_{CO_2}) can be estimated (Zeebe & Wolf-Gladrow, 2001):

$$\tau_{CO_2} = \frac{\text{mixing depth}}{k} \cdot \frac{\text{DIC}}{[CO_2^*]} \cdot \frac{1}{RF} \tag{3.4}$$

τ_{CO_2} functions as an local gas exchange relaxation timescale describing a system seeking equilibrium; the extent to which the period in which a system undergoes gas exchange approaches τ_{CO_2} dictates the efficiency of the exchange (Jones et al., 2014). In the case that surface waters remain in contact with the atmosphere for periods meeting and exceeding τ_{CO_2}, equilibration approaches completion barring other processes (primary production, respiration, etc.) maintaining air-sea disequilibrium. For systems like Lake Superior in which these other processes are minimal during mixing (see Section 3.5), τ_{CO_2} gains further relevancy as a descriptor of the actual process of air-sea CO_2 equilibration as observed in this work, and may be used to compare the physicochemical drivers of mixing among systems.

During unstratified periods, mean RF was 26.9 ± 0.6, mean DIC was 867.0 ± 0.9 µmol kg^{-1}, and mean [CO_2*] was 29.6 ± 0.8 µmol kg^{-1} (all ± s.d.). The

resulting τ_{CO_2} during the unstratified period was 100. ± 4 days. This period is much smaller than that of most of the surface ocean mixed layer, indicating relatively fast CO_2 equilibrium despite Superior's deeper mixed layer. This period is similar in magnitude to the duration of the twice-annual unstratified periods in December-January and May-July (though stratification phenology varies among years; (Austin & Colman, 2008; Woolway et al., 2021)), so it is reasonable to expect that on multiannual timescales, Lake Superior maintains near-atmospheric CO_2 equilibrium. This inference depends on lake stratification and wind velocity, both of which may shift with the changing climate (P. Xue et al., 2022). Climate change effects on lake thermal state and atmospheric circulation are likely to have complex effects on lake biogeochemistry which extend to CO_2 flux behavior changes (Desai et al., 2009). This brief calculation also implies that stratified waters and high wind events may exhibit a much smaller τ_{CO_2}; assuming a mixing depth of 10 m and k of 6.4 m d^{-1}, which might be observed in a strong gale during the winter stratified season, τ_{CO_2} would decrease by a factor of c. 40. This thought experiment informed by our data emphasizes the importance of thermal structure and ephemeral wind events for determining large-scale C cycling in Lake Superior.

3.2.4 pCO_2 Variability

A continuous multiannual cycle of observed pCO_2 could not be constructed due to large gaps in the time series, so the four years of observations were collated by Julian day of year in order to illustrate seasonal CO_2 cycling variability, facilitate comparison to previous observations and models, and contrast Lake Superior with other systems. LOESS regression of observations grouped by 0.01° boxes and date of observation highlighted seasonal cycling of sea surface pCO_2 which resembled the sinusoidal behavior modeled by Bennington et al. (2012) (Figure 3.3a). This regression used a data window spanning the 5% of observations on either side of regression points and 10 iterative fits. A local

maximum was observed in the regression at 427 µatm in early June, while a local minimum was observed at 310 µatm in early September. Individual observations ranged from 199 to 1056 µatm, with a mean value of 379 ± 60 µatm (± s.d.).

Heterogeneity was visible in the range of pCO_2 values observed on a given date, with super- and under-saturated conditions observed throughout the year. This high degree of spatial and/or temporal heterogeneity obscured the seasonal pCO_2 cycle. Additionally, the high concentration of transects in the riverine-influenced Western Arm of Lake Superior may not have represented open-water conditions prevailing in the remainder of the lake. Diel variability was examined as a potential source of bias, but no significant difference between daytime and nighttime pCO_2 was found (Supporting Text S1) in contrast with observations of large diel CO_2 cycling variability in many smaller lakes (Golub et al., 2023).

Confounded spatial and seasonal variability was partly separated by SC into "riverine" and "pelagic" regimes in order to isolate open-water seasonal variability. A cutoff SC value was defined by statistically significant departure from the surface SC distribution observed in unstratified periods. In every year of observation, springtime unstratified surface SC observations formed a narrow distribution with a mean of 99.44 µS cm^{-1} and a standard deviation of 0.64 µS cm^{-1}. This value was taken to represent the mean SC of the well-mixed freshwater lake. Observations with SC 3 standard deviations greater than the unstratified period mean were considered river-influenced, as in this system, river inputs have higher ionic strength or SC due to interactions with the rocks and soil in their watersheds and due to antropogenic inputs such as road salts (Weiler, 1978). This scheme decreased the noise around the seasonal trend of surface water pCO_2 in pelagic observations (Figure 3.3b) and highlighted heterogeneity associated with riverine-influenced observations (Figure 3.3c). Potential interferences with this classification included evaporation and precipitation, which would be expected to increase and decrease surface SC,

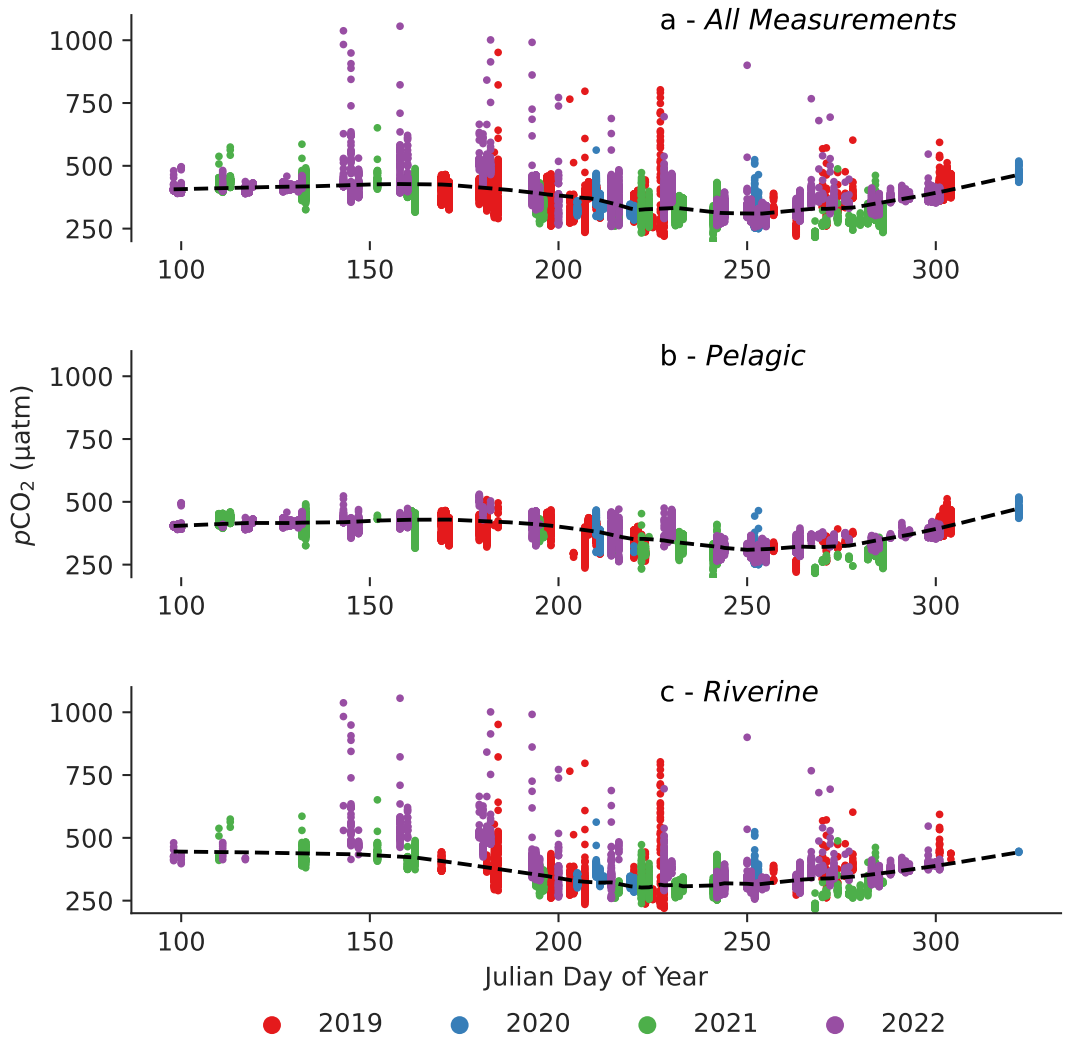

Figure 3.3: Seasonal variation collated from pCO_2 observations grouped by 0.01° squares and date during transects of Lake Superior over 2019-2022. Black dashed lines represent LOESS regressions of each time series.

respectively. For this reason, we elected not to construct a quantitative mixing relationship based on underway-measured surface SC and merely used it as a rough proxy for riverine influence.

In pelagic waters of Lake Superior during April-November the mean observed pCO_2 was 380 ± 53 µatm, while in river-influenced waters, the mean observed pCO_2 was 343 ± 38 µatm (± s.d.). The depression of riverine regime mean pCO_2 pCO_2 may have been due to promotion of primary production and CO_2 drawdown in nutrient-rich riverine-influenced waters (Minor et al., 2014; Sterner et al., 2020), yet contemporaneous observations of highly elevated pCO_2 point to a gradient of heterotrophy and autotrophy associated with riverine plumes (Delvaux, 2017). The pelagic pCO_2 cycle regression displayed a greater seasonal variability than the modeled time series of Bennington et al. (2012) (Figure 3.4). Annual pCO_2 summer minima and spring maxima were approximately 330 and 400 µatm in Model 1 of that work, compared to 313 and 428 µatm in this study's regression of collated pelagic observations from 2019-2022. Bennington et al. modeled surface water equilibrium with an atmospheric pCO_2 of 360 µatm at the end of a mixing period spanning late April-late June 1997-2001. At the end of destratification in this (2019-2022) study, a mean surface water pCO_2 of 430 ± 30 µatm (± s.d.) was observed, which was indistinguishable from contemporaneous atmospheric pCO_2. The two models presented by Bennington et al. differed in their treatment of primary production limitation, which resulted in the greatest differences after spring mixing, when this study's observations also displayed high variability. These differences in the shape of the post-mixing pCO_2 cycle may indicate a model shortcoming or a change in inorganic carbon system drivers over the preceeding two decades.

The observed increase in spring mixing period pCO_2 relative to that modeled by Bennington et al. (2012) was consistent with the magnitude of atmospheric CO_2 concentration increase (c. 2 ppm yr^{-1}, Keeling and Keeling (2017)) over the approximately 22 years separating the modeled period of Bennington

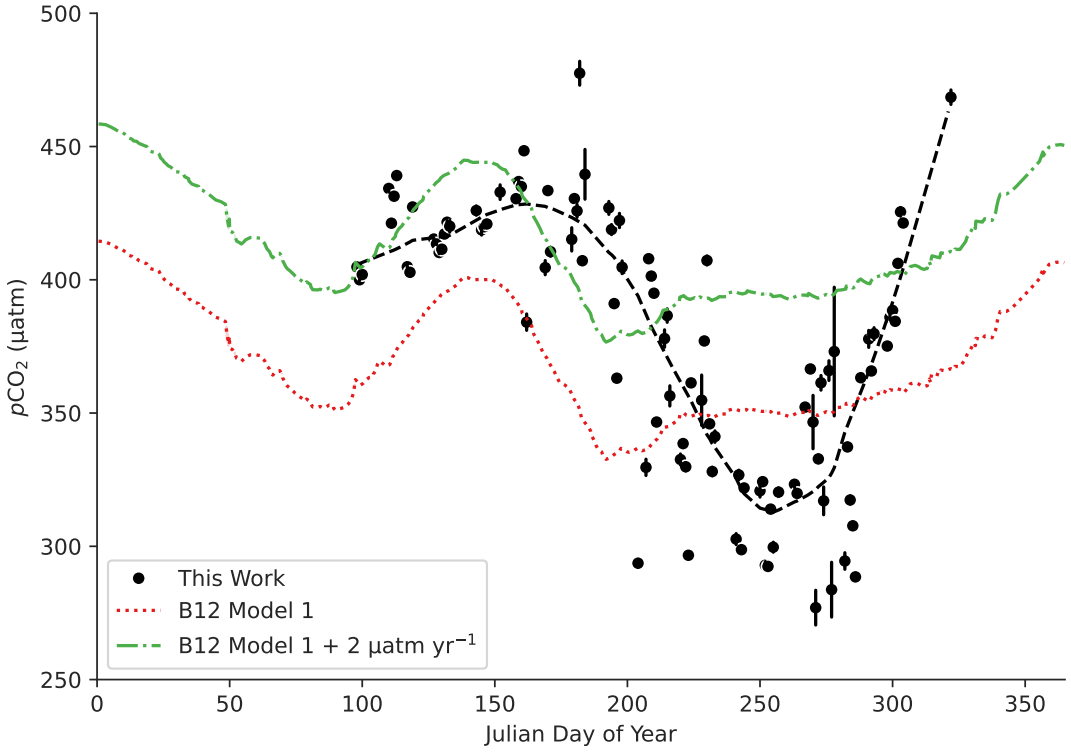

Figure 3.4: Median daily observations of pelagic surface water pCO_2 observed during 2019-2022 compared with Model 1 from Bennington et al. (2012), which described mean lake surface pCO_2 1997-2001. A 44 µatm adjustment of Model 1 to account for 22 years' atmospheric CO_2 increase (assuming 2 µatm yr^{-1}) aligned spring and mixing season modeled results with contemporary observations.

et al. and these observations, as well as the direction of increase in Lake Superior surface water pCO_2 calculated from pH and A_T over the period 1992-2019 by Minor and Brinkley (2022). The precise rate of increase of Lake Superior surface water pCO_2 over decadal timescales remains difficult to constrain, but its repeated achievement of near-atmospheric equilibrium, along with radiocarbon measurements indicating rapid (<3 years) recycling of the DIC pool Zigah et al. (2011), and a relatively brief characteristic time of CO_2 equilibration (Section 3.3), indicate that it mirrors atmospheric pCO_2 over interannual periods and may continue to do so. This atmospheric CO_2 invasion of Lake Superior occurs regardless of its net CO_2 air-sea flux due to its dynamic chemical equilibration with atmospheric CO_2.

The magnitude of seasonal variability in Lake Superior surface pCO_2 was comparable to that of subtropical ocean regions (N. R. Bates, 2001), but shifted in the year. In terms of pCO_2 phenology, Lake Superior resembled the Arctic ocean most closely, despite exhibiting a larger amplitude (Orr et al., 2022). Neighboring Lake Michigan displayed a minimum surface pCO_2 2 to 3 months earlier than Superior in a process model study (Pilcher et al., 2015). Scarcity of data from November-April prevented extrapolation to those periods, but previous, model-based studies suggest Lake Superior pCO_2 likely remains supersaturated or near-atmospheric equilibrium throughout that period (Bennington et al., 2012). Interannually-variable wintertime ice cover (White et al., 2012) may modify the expected CO_2 efflux.

3.2.5 Competing Drivers of pCO_2

Deconvoluting the pelagic pCO_2 cycle (Figure 3.3b) into inferred drivers sheds light on biogeochemical cycling in Lake Superior by isolating direct temperature control of pelagic surface pCO_2 from control by other drivers. Changes in water temperature directly control the speciation of the inorganic carbon system via changes to chemical equilibria and solubility. By removing the direct

influence of water temperature variability, the remaining variation can be ascribed to biological and physical processes including production, respiration, sea-air CO_2 gas flux, and allochthonous C inputs. The indirect influence of temperature on these processes was not normalized by this analysis to preserve the seasonality of the biological and physical drivers.

The method of Takahashi et al. (1993) was used to separate measured pCO_2 into thermal ($pCO_{2\,\mathrm{T}}$) and non-thermal ($pCO_{2\,\text{non-T}}$) signals using observed sea surface temperture (T) and the temperature-driven variation of pCO_2:

$$pCO_{2\,\mathrm{T}} = \overline{pCO_2} \cdot e^{\left(\frac{\partial ln(pCO_2)}{\partial T}[T-\overline{T}]\right)} \tag{3.5}$$

$$pCO_{2\,\text{non-T}} = pCO_2 \cdot e^{\left(\frac{\partial ln(pCO_2)}{\partial T}[\overline{T}-T]\right)} \tag{3.6}$$

These signals have been termed "temperature" and "biological"/"biophysical" effects in other studies (e.g. Fassbender et al., 2018; Minor et al., 2019). This work uses the T/non-T terminology of Atilla et al. (2011). $CaCO_3$ dissolution and precipitation were neglected in this analysis because Lake Superior is vastly undersaturated with respect to $CaCO_3$ so precipitation and dissolution processes are not relevant. Overbars indicated arithmetic mean values in the literature source, but this study analyzed an incomplete annual time series of pCO_2, so mean temperature ($\overline{T}$) and mean pCO_2 ($\overline{pCO_2}$) were adjusted to 1 °C and 400 µatm to ensure convergence of the driver signals at the beginning of the observed period. The temperature partial derivative of $ln(pCO_2)$ was calculated via PyCO2SYS, yielding an average value of 0.03606 $°C^{-1}$ for Lake Superior over the temperature range 0-20 °C (Supporting Text S2). This temperature dependence is in good agreement with values used in previous studies (0.038 $°C^{-1}$ Atilla et al., 2011; 0.0384 $°C^{-1}$ Lynch et al., 2010).

Plotting the measured, thermal, and non-thermal pCO_2 signals illustrated the interplay of these competing drivers of pCO_2 in Lake Superior (Figure 3.5).

Seasonal temperature effects were visible as the springtime increase and autumn decrease in $pCO_{2\,T}$, opposed by the summertime dip in $pCO_{2\,\text{non-T}}$. Measured pCO_2 lay suspended between the curves. The degree to which thermal vs. non-thermal drivers control pCO_2 can be conceptualized as the vertical distance between the measured curve and its two drivers; in spring, measured pCO_2 was closely tied to $pCO_{2\,T}$, indicating that most of the spring trend in pCO_2 was driven by seasonal warming. pCO_2 moved equidistant between drivers before dipping with the non-thermal curve through the summer, resulting in CO_2 influx presumably maintained by net biological C drawdown, which persisted until the end of the ice-free season.

Quantitatively, the ratio of thermal to non-thermal control of pCO_2 can be calculated (Fassbender et al., 2018; Takahashi et al., 2002) as

$$R_{\text{T non-T}^{-1}} = \frac{\max(pCO_{2\,T}) - \min(pCO_{2\,T})}{\max(pCO_{2\,\text{non-T}}) - \min(pCO_{2\,\text{non-T}})} \tag{3.7}$$

which yielded a value of 1.1 using the regressions in Figure 3.5, indicating roughly equal thermal and non-thermal driver magnitudes over the ice-free period. Interestingly, this value aligns with that of the Atlantic Ocean at the approximate latitude of Lake Superior (Fassbender et al., 2018), which raises questions about latitudinal gradients in $R_{\text{T non-T}^{-1}}$ in inland waters compared to marine systems. Minor et al. (2019) found majority non-thermal control of calculated pCO_2 from discrete samples of Lake Superior surface water in 2014-2016 in agreement with the findings of Atilla et al. (2011). This deconvolution analysis highlights the importance of non-thermal drivers of C drawdown in the latter stages of the ice-free season acting to depress pelagic surface pCO_2 despite opposition from temperature effects.

3.2.6 CO_2 Flux Variability

CO_2 air-sea flux rates calculated from observed pCO_2 displayed seasonal variation similar to that of pCO_2, but with a greater degree of variability within

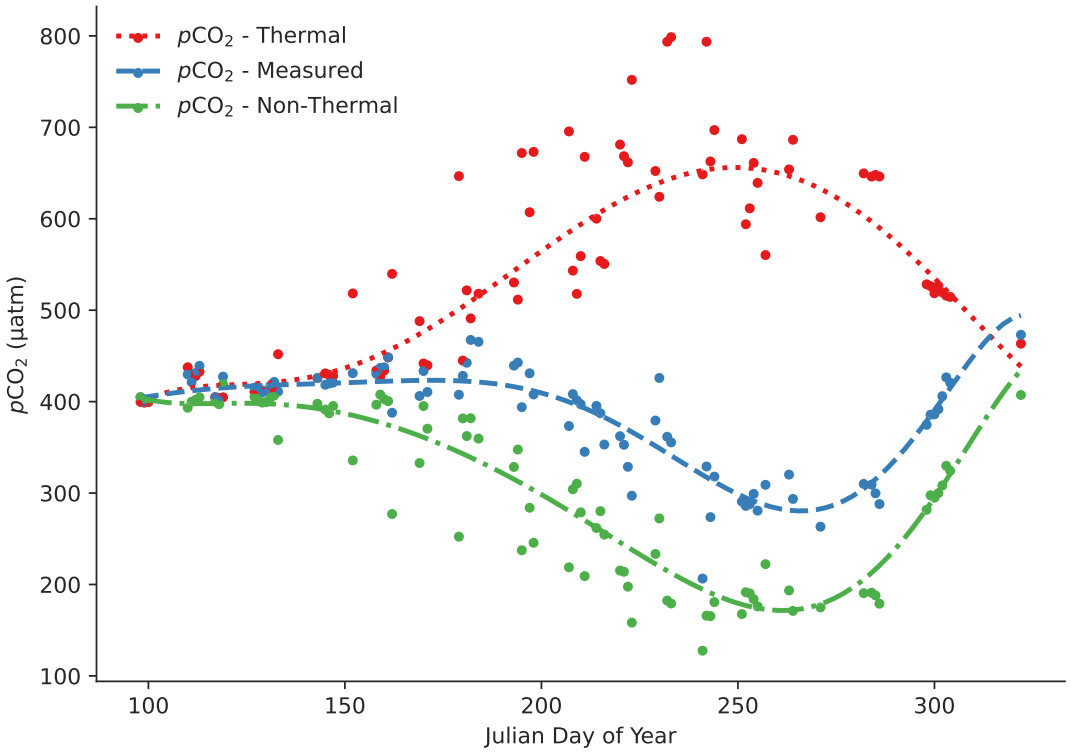

Figure 3.5: Deconvolution of median daily measured sea surface pCO_2 (blue dashed line) into non-thermal (green dash-dot line) and thermal (red dotted line) drivers. 7^{th}-order power function regressions are shown as visual aids, and their equations are given in the Supporting Information.

individual cruises. The mean calculated CO_2 flux rate was -2.7 ± 6.4 mmol C m^{-2} d^{-1}. In pelagic waters, the mean calclated flux rate was -1.9 ± 5.7 mmol C m^{-2} d^{-1}, while in riverine-influenced waters a lower value of -4.5 ± 7.5 mmol C m^{-2} d^{-1} was calculated (all ± s.d.). These values fall within the range of instantaneous CO_2 fluxes observed or modeled in previous work (Bennington et al., 2012; Delvaux, 2017). The most extreme air-sea fluxes were calculated in midsummer, when high wind speeds coupled with CO_2-undersaturated surface waters to create high rates of CO_2 drawdown exceeding 70 mmol C m^{-2} d^{-1}. Differences between pelagic and riverine-associated observations were clearest in spring, when pelagic observations tended to favor CO_2 air-sea influx while riverive observations favored efflux. This correlated with observations of springtime allochthonous C loading associated with nearshore CO_2 evasion in Lake Superior (Minor et al., 2019).

Net CO_2 air-sea flux over the observed seasons was estimated via trapezoidal integration of median daily observations of instantaneous CO_2 flux collated by Julian day across the observed time domain: Julian day 100 (April 9 or 10) through 300 (November 26 or 27). The resulting values (Table 3.1) were both multiplied by the total area of Lake Superior (8.21 x 10^{10} m^2) to yield total fluxes as it was unclear what fraction of the lake was considered pelagic vs. riverine. Uncertainty in integrated CO_2 fluxes was determined by bootstrap random resampling of observations with replacement for 100 repetitions of the integration and given as the standard deviation of the ensemble net fluxes. These values serve as rough bounds for the mean net CO_2 flux of Lake Superior throughout the ice-free seasons of 2019-2022. The contrast between spring riverine regime instantaneous CO_2 *efflux* and integrated ice-free season CO_2 *influx* speaks to temporal variability in riverine effects on inorganic C cycling in Lake Superior.

The resulting ice-free period CO_2 influx on the order of 25 Gmol C (300 Gg C, 0.30 mol C m^{-2}) was similar in magnitude but opposite in sign to the

Table 3.1: Time-integrated fluxes of CO_2 over the air-water interface of Lake Superior ascribed to Pelagic and Riverine chemical regimes for Julian Days 100-300. Total Flux values assume the whole lake surface area is either similar to pelagic or to riverine values and thus give the range within which total CO_2 lake flux is likely to fall. Uncertainties are given as standard deviations propagated via bootstrap resampling with replacement for 100 repetitions. Negative signs indicate influx.

Region	CO_2 Areal Flux (mol C m^{-2})	CO_2 Total Flux (Gmol C)
Pelagic	-0.296 ± 0.036	-24.3 ± 2.9
Riverine	-0.341 ± 0.017	-28.0 ± 1.4

only fully-annual estimate of Lake Superior air-sea CO_2 flux: a mean net annual *efflux* of 16 Gmol C yr^{-1} (190 Gg C yr^{-1}, 0.19 mol C m^{-2} yr^{-1}) over the period 1997-2001 (Bennington et al. 2012). The discrepancy is accounted for by winter supersaturation of surface pCO_2. Assuming the veracity and comparability of the above values (perhaps an invalid assumption due to the shift in model curves suggested by Figure 3.4), an efflux of 41 Gmol C (492 Gg C, 0.50 mol C m^{-2}) during Julian days 301-99 would be implied. The rough approximations of carbon budgets allowed by available observational CO_2 flux data continues to prohibit integration of Lake Superior into regional and global C budgets. There remains the possibility that the modeled annual CO_2 flux and this study's calculated sub-annual flux are not comparable due to two intervening decades of ecological and climate change, an under-constrained modeled pCO_2 cycle, and ongoing uncertainty about comparisons of measured versus calculated pCO_2 in Lake Superior. An updated observation-based and/or process model constrained by spatially- and temporally- comprehensive direct observations of pCO_2 and CO_2 flux is required for substantive comparisons of observed and modeled C cycling.

3.3 Discussion

Four years of surface pCO_2 measurements gathered on transects across Lake Superior were used to describe inorganic carbon system variability across temporal and spatial scales. Ice-free season (April-November) observations yielded a detailed account of the seasonal pCO_2 cycle, driven by thermal and non-thermal drivers acting in opposition to perturb surface pCO_2 from its interannually-repeated atmospheric equilibrium, resulting in sustained periods of CO_2 influx and efflux. Increases in surface pCO_2 over the last two decades illustrate that Lake Superior is undergoing CO_2 invasion in agreement with Phillips et al. (2015). Separating the lake into pelagic and riverine regimes on the basis of SC highlighted riverine influences on the inorganic C system which tended to increase spring pCO_2 observations and instantaneous air-sea CO_2 efflux, likely through a combination of DIC loading and promotion of respiration.

Comparison of this study's calculated areal CO_2 flux rates demonstrates the importance of Lake Superior relative to regional C budget terms. Previous eddy covariance work indicated Lake Superior's ice-free season CO_2 sink is likely less than that of the surrounding terrestrial ecosystem (Vasys et al., 2011). The mean CO_2 instantaneous air-sea influx of 2.7 mmol C m^{-2} d^{-1} found in this study supports this finding when compared to the terrestrial net ecosystem production (NEP) of 19.6 mmol C m^{-2} d^{-1} (86.1 g C m^{-2} yr^{-1}) for the Great Lakes hydrologic unit (Butman et al., 2016). This terrestiral NEP is still nearly three times larger than the mean annual Lake Superior CO_2 efflux of 6.3 mmol C m^{-2} d^{-1} modeled by Bennington et al. (2012). The absolute values of these three C flux rates remain less than an order of magnitude apart, demonstrating the importance of Lake Superior as an major player in the regional C cycle. This point is further emphasized by the sheer size of Lake Superior, which has a surface area 24% the size of the Great Lakes hydrologic unit, and a net annual Lake Superior CO_2 efflux Bennington et al., 2012 17% that of the Great Lakes hydrologic unit net lake CO_2 efflux of 9.2 Gmol C Butman et al., 2016.

Integration of measured CO_2 air-sea fluxes rates over the observed period indicated net CO_2 influxes of 24.3 ± 2.9 Gmol C (pelagic) and 28.0 ± 1.4 Gmol C (riverine), which are considered bounding values for the mean whole-lake ice-free period mean CO_2 flux during 2019-2022. The results of this study are in agreement with previous work indicating sustained periods of CO_2 influx during the summer stratified period contrasting with efflux during winter (Bennington et al., 2012), yet the contribution of the Great Lakes to regional and global C budgets remains uncertain, as recent regional and global estimates of CO_2 evasion from lakes and reservoirs have presented a broad range of values with large margins of error. Cavallaro et al. (2018) produced a North American lake and reservoir CO_2 efflux of 122 Tg C (10.1 Tmol C) yr^{-1} ± 100% which neglected the Laurentian Great Lakes, while Lauerwald et al. (2023) reported a North American natural lake CO_2 evasion of 337 Tg C (28.1 Tmol C) yr^{-1} with a range of 202-459 Tg C yr^{-1}. These recent studies represent significant advancements in the science of inland water carbon cycling, but their uncertainties indicate the potential for improvements in CO_2 cycling observation as demonstrated in this work to contribute to improved regional and global C budgets.

Future work should seek to address some of the limitations of this work. In particular, a paucity of early Spring and late Fall data hindered analysis of periods at the extremes of the observation period, which could shed light on the effects of ice-off as a driver of ice-free season CO_2 flux (cf. Ahmed et al., 2019. As previously noted, there may be some bias in wind-parameterized gas transfer velocities associated with dual-tracer experiments (M. Yang et al., 2022), such that the gas transfer velocities calculated here may be underestimates by as much as 20%. Small-scale variations in turbulence and eddy scales are also important and underexplored mechanisms affecting air-water gas transfer uncertainty in low-wind environments (Wang et al., 2015). Wind speed gas flux parameterization applications in large lakes may be an important area for future study.

3.3.1 Consequences of Increasing pCO_2

Among the most impactful findings of this research is the observation that Lake Superior surface pCO_2 maintains near-equilibrium with the overlying atmosphere over multi-year periods. Temperature variability and biogeochemical processes drive seasonal departures of pCO_2 from atmospheric equilibrium (creating a likely net annual CO_2 efflux), yet surface water pCO_2 returns to a baseline state near atmospheric equilibrium on timescales shorter than a year. This fact has several significant consequences in a world of increasing atmospheric CO_2 concentration:

First, the solubility pump of Lake Superior acts as a partial CO_2 sink which can be approximated by an equilibrium calculation: Assuming A_T = 840 µmol kg^{-1}, SST = 3.98 °C (temperature of maximum density during destratification), an initial pCO_2 = 400 µatm, and atmospheric ΔpCO_2 Δt^{-1} = 2.50 µatm yr^{-1}, then equilibrium calculation indicates ΔDIC Δt^{-1} = 0.184 µmol kg^{-1} yr^{-1}, which is multiplied by the approximate mass of Lake Superior (1.21 x 10^{17} kg) to give a CO_2 storage of 22.3 Gmol C yr^{-1} (267 Gg C yr^{-1}) due to increasing atmospheric CO_2 alone. This storage is characteristic of any body of water maintaining CO_2 equilibrium with a non-steady-state atmosphere. It acts alongside C sources (e.g. DIC loading) and sinks (e.g. C burial) to compose the net annual C budget of Lake Superior. Development of an annual net CO_2 flux using expanded observational and modeling capabilities may yield insights on all of these contributors. If atmospheric pCO_2 were stable, then Superior's annual net CO_2 efflux could be larger than it is today, mirroring the case of the pre-industrial global ocean, which likely acted as a CO_2 source instead of a sink (Cartapanis et al., 2018).

Second, Lake Superior's water chemistry will undergo changes as a result of consistently-higher pCO_2. Its weak CO_2 buffer (RF 25-30 in this work, compared to marine values 8-16; Sarmiento and Gruber, 2006) and absence of sediment carbonate buffer (unlike neighboring Lake Michigan) result in relatively

high sensitivity to atmospheric CO_2 acidification. The outcomes of hypothesized lake acidification mirror those in the ocean: decreasing pH and $CaCO_3$ saturation states, impacts on primary producer communities and carbonate-shelled organisms such as *Dreisseina* mussels, changes to metal ion activities, and other phenomena with potentially detrimental ecosystem effects (Doney et al., 2009). Trends in A_T and temperature may modify the *speciation* (e.g. $[CO_3^{2-}]$, pH) of the inorganic carbon system as well as the seasonal and spatial expression of the surface water pCO_2 cycle, but *not* the surface pCO_2 of a system at equilibrium with the atmosphere.

Third, efforts to observe Lake Superior's inorganic C system must capture a greater fraction of the annual cycle and spatial variability to constrain these changes. The twice-annual time series of chemical parameters (including glass electrode pH and Gran titration alkalinity) collected by US EPA Great Lakes National Program Office includes samples over a broad spatial scale, during periods of mean CO_2 efflux (April-May) and influx (August-September) but fails to observe intervening periods which provide necessary context for inter-annual variability of the annual pCO_2 cycle. Undersampling a complex signal like inorganic C chemistry delays detection of climate change effects (Carter et al., 2019). A more complete picture of biogeochemical parameters is sorely needed during the present period of climate change and ecological disruption. This gap in observational capabilities can be addressed by a sustained campaign of higher-quality, higher-frequency measurements of inorganic C parameters in the Laurentian Great Lakes.

3.3.2 Observational Challenges and Opportunities

Environmental and instrumental challenges limit deployment of underway pCO_2 systems as tools for biogeochemical observation on large lakes like Superior. These instruments describe only a small fraction of a water body at any given time, which complicates efforts to generalize results to the system as a whole. A

network of similar sensors equipped on moorings, vessels of opportunity, and other vehicles (e.g. drifters, saildrones, wavegliders) may be suited for more comprehensive observation. Seasonal ice cover limits winter deployment of autonomous sensors, and has long acted as a blinder focusing scientific attention on more accessible seasons (Block et al., 2018). Novel observation platforms designed to observe under-ice pCO_2 (DeGrandpre et al., 2019; C. Lee et al., 2022) demonstrate the potential to expand the horizons of inorganic C observation in seasonally ice-covered lakes. Direct measurements of gas flux may also be obtained by eddy covariance towers in the vicinity of the Great Lakes (Shao et al., 2015).

This research grappled with problems of bias in transect data due to overrepresentation of certain regions in space (the far western lake) and time (summer). Although these problems were partially addressed by regression analysis and separation of pelagic and riverine regimes, future work should consider other drivers of spatial and temporal heterogeneity, for example: dissolved organic matter and chlorophyll measured by in-situ instruments or remote sensing (e.g. Lohrenz et al., 2018; Sims et al., 2023). Expanded monitoring of pCO_2 and related chemical properties in the Laurentian Great Lakes provides a fruitful avenue for observation and modeling of CO_2 budgets in the world's largest surface freshwater resource.

3.4 Conclusions

This study provided the most comprehensive observations to date of surface pCO_2 variability in Earth's largest freshwater lake by area and demonstrated techniques for inferring C cycling drivers in an understudied system. As the present perturbation of Earth's C cycle continues, the need for such knowledge to inform water and climate policy will grow apace, requiring continuing innovation of observational and modeling capabilities. This is as true for the Laurentian Great Lakes as for the African Rift Lakes and other understudied

surface waters of the world.

A spatially-comprehensive, fully annual CO_2 flux budget is not achievable with the data presented here because of spatial and temporal gaps in the time series presented. Future work must perform more observation of neglected regions in space and time, extrapolation to unobserved domains, and generalization of observed fluxes and drivers by modeling efforts. To this end, we recommend further development of observational strategies such as underway data collection, moored and autonomous instrumentation, remote sensing, and winter limnology techniques to better constrain CO_2 flux in Superior and other large lake systems. Efforts to resolve the modeled C budgets of the Great Lakes will benefit from a greater number of CO_2 measurements to constrain and correct models (Gloege et al., 2022). Insights into the balance of productivity and respiration may result from pairing a large pCO_2 survey with measurements of other biogeochemical tracers such as dissolved oxygen (Evans et al., 2022) or primary productivity (Sterner, 2010). As ice cover of temperate lakes declines with climate change, the period amenable to transects of seasonally ice-covered lakes will grow. This disappearance of the ice cover regime is among driving forces of the sub-discipline of winter limnology, which studies a vanishing environment (Ozersky et al., 2021). It is unclear how changes in ice cover will affect annual pCO_2 fluxes in these changing lakes systems. Spatially- and temporally- comprehensive observations of element cycling in these large lakes hint at the depth and complexity of biogeochemical functions responding and feeding back to a changing planet.

Chapter 4

A neural network-based estimate of the seasonal to inter-annual variability of the Lake Superior carbon cycle

Abstract

Lake Superior emits and absorbs CO_2 with significant seasonal and interannual variability, which complicates efforts to constrain its carbon cycle. While it regains atmospheric CO_2 equilibrium repeatedly on a sub-annual scale, resulting in a sustained rise in observed pCO_2 over the last several decades, significant gaps in observation have prevented examination of variability in its C cycle on smaller temporal or spatial scales. We develop a gap-filling model of Lake Superior surface water pCO_2 and CO_2 flux in order to infer observation-based C cycling patterns and create a fully-annual CO_2 flux product. A feed-forward neural network was trained and tested on underway pCO_2 measurements spanning 2019-2023, yielding a spatially-comprehensive product describing inorganic carbon dynamics over a five-year period. Integration of calculated CO_2

air-sea flux time series over this five-year period indicated Lake Superior alternated between net annual CO_2 influx and efflux, with values ranging from $0.30\,\mathrm{Tg\,C\,yr^{-1}}$ influx to $0.06\,\mathrm{Tg\,C\,yr^{-1}}$ efflux over the span of this model. This refinement of Lake Superior's C budget juxtaposed the lake's large seasonal and interannual variability against a mean net annual balance of C sources and sinks, and opened the door to further applications of machine learning predictive models of lacustrine carbon cycling.

4.1 Introduction

The concentration of carbon dioxide (CO_2) in Earth's atmosphere has risen by more than 50% since the pre-industrial era (Friedlingstein et al., 2023). While the primary symptom of this perturbation of Earth's carbon cycle is the increasing radiative forcing of the greenhouse effect, a parallel outcome is CO_2 invasion of the surface ocean and cascading chemical and ecological outcomes (Doney et al., 2009) which combine to create a multi-stressor environment for marine ecosystems (Breitberg et al., 2015; Wallace et al., 2014). Observations of anthropogenic CO_2 invasion of inland waters have so far been limited, due in part to the difficulty of detecting climate signals amid other modes of variability (Phillips et al., 2015). This problem parallels the study of ocean acidification in coastal waters subject to both marine and terrestrial drivers (Frieder et al., 2012; Hofmann et al., 2011; Lowe et al., 2019).

Progress towards observing changes in inland water CO_2 cycling has recently been made in the Laurentian Great Lakes (hereafter Great Lakes), Earth's largest contiguous body of fresh water. These lakes on the border of the United States and Canada and within the historical and contemporary lands of First Nations and Native Americans provide invaluable water resources and ecological services to millions of people. Observations of pH and total alkalinity (A_T) spanning three decades indicated increasing calculated partial pressure of carbon dioxide (pCO_2) for some lakes for some periods, and an uncertain

trend in pH (Minor & Brinkley, 2022) connected with A_T variability or uncertainty. Rising pCO_2 in the Laurentian Great Lakes was predicted by Phillips et al. (2015), along with a hypothesized decrease in mean annual pH (ignoring seasonal variability or feedbacks such as invasive mussel calcification and sediment buffering). An interannual rise in Lake Superior surface water pCO_2 was confirmed with underway observations of pCO_2 in Lake Superior demonstrating that the lake kept pace with the concurrent rise in atmospheric CO_2 (Sandborn & Minor, 2024). This behavior was mediated by the lake's dimixis, rapid equilibration timescale, and lack of $CaCO_3$ buffer. That study was unable to generalize the time series of Lake Superior sea surface pCO_2 to unobserved regions and periods to produce a spatially- and temporally-comprehensive CO_2 air-sea flux budget due to observed spatial heterogeneity and gaps in the time series.

Despite these advances, the Great Lakes' carbon budgets are relatively poorly-constrained. It has been postulated that the Great Lakes ecosystem is a net source of CO_2 to the atmosphere on the scale of several $\mathrm{Gg\,C\,yr^{-1}}$, while the individual lakes are thought to be CO_2 sources (Michigan, Huron) or sinks (Erie, Ontario) of highly uncertain magnitudes (Alin & Johnson, 2007; McKinley et al., 2011; Urban & Desai, 2009). Even the sign of Lake Superior's net air-sea CO_2 flux is uncertain, as the two most recent estimates give a net efflux (Bennington et al., 2012) based on biogeochemical process modeling or a net influx (Brothers & Sibley, 2018) based on observation-based inference of lake metabolism. This uncertainty has barred the Great Lakes' incorporation into regional or global carbon budgets (Cavallaro et al., 2018) despite their demonstrated significance to the regional carbon budget (Butman et al., 2016; Cole, 2013a; Urban & Desai, 2009). This gap in knowledge stands as a testament to the difficulty in studying such large, energetic, seasonally ice-covered lakes.

In light of a dearth of accurate and spatially- and temporally-resolved carbon cycling data, the underway pCO_2 time series described by Sandborn and Minor (2024) is a valuable resource, but gaps in space and time among cruises

and during the ice-covered season prohibited its use in isolation to constrain a comprehensive budget of CO_2 air-sea flux. This problem parallels the shortfall of observations in under-observed regions of the ocean which led to the use of statistical tools to generate observation-based models of marine inorganic carbon cycling (Rödenbeck et al., 2015). While this barrier can to some extent be circumvented by process modeling (e.g. coupled biogeochemical-hydrodynamic products; Bennington et al., 2012) and addressed by novel observing platforms (C. Lee et al., 2022), machine learning techniques such as extreme gradient boosting, regression trees, and feedforward neural network (FFNN) techniques have utility in gap-filling global ocean biogeochemical datasets and generalizing driver variables into unobserved domains in time and space (Bennington et al., 2022; Gloege et al., 2022; Keppler et al., 2023; Landschützer et al., 2013). This work produces a pCO_2 model based on a FFNN trained by underway pCO_2 and deployed to predict sea surface pCO_2 for Lake Superior across a period of 5 years.

Lake Superior resembles the global ocean for the purposes of this model in several important ways. The majority of Surface Ocean CO_2 Atlas (SOCAT) observations (Bakker et al., 2016) used for similar predictive models are clustered around shipping lanes, particularly around North America, Europe, and East Asia. Underway data in Lake Superior similarly display repeated transects and clustering in the southwestern quarter of the lake (Figure 1, Sandborn and Minor, 2024). Underway pCO_2 measurements in Lake Superior are also collected using essentially identical instrumentation to that generating much of the SOCAT underway data. These qualities enhance the spatial coverage of data while illustrating temporal variability at select locations. The physical, chemical, and biological characteristics of this oligotrophic large lake also resemble those of the coastal ocean more closely than perhaps any other fresh water body (Sterner, 2021). These qualities aid the transfer of marine gap-filling model techniques to a large lake environment. Using the output of the FFNN model described in this work, the first multiannual, spatially-comprehensive

observation-based estimate of CO_2 air-sea flux is produced and compared to previous estimates, resulting in a refined carbon budget for Lake Superior.

4.2 Data and Methods

Surface water pCO_2 was measured during 90 transects of the *RV Blue Heron* across Lake Superior during the ice-free seasons (roughly April to November) of 2019-2023, and archived and accessed via Rolling Deck to Repository (University of Minnesota Duluth, 2023). Collection and processing of this data was described by Sandborn and Minor (2024). Briefly, measurements of pCO_2 were collected at 2 second intervals by a SuperCO2 instrument (Sunburst Sensors) equipped with a dual showerhead equilibrator. After correction based on standard gas analysis, and aggregation as daily mean values in a $0.02° \times 0.02°$ resolution grid, 9.3×10^4 measurements of surface water pCO_2 with an estimated measurement uncertainty of ± 5 µatm were obtained.

Input parameters for the FFNN predictive model were identified by comparing available gridded parameters against a set of requirements: input parameters must have some proxy relationship with the output parameter pCO_2, they must span Lake Superior over the domain 2019-2023, and they must contain no gaps in space or time. Similar marine studies have utilized remotely-sensed, observed, or modeled gridded products describing sea surface temperature (SST), salinity, mixed layer depth, solar irradiance, chlorophyll-a concentration, and atmospheric CO_2 concentration, among other parameters, together with their respective climatologies and anomalies. Lake Superior presented obstacles to the use of many of these input parameters. Winter ice and cloud cover prevented great confidence in remotely-sensed predictors (Mouw et al., 2013). Salinity was less useful as a biogeochemical tracer in this freshwater system, though elevated conductivity has been used as an indication of riverine influence (Weiler, 1978). Mixed layer depth was not well-defined in any gridded products available for the entire period 2019-2023 (one potential source,

the Lake Superior Observational Forecast System hydrodyamic FVCOM model, wasn't available before October 2022).

To overcome these challenges, other directly-observed or modeled variables were used as prediction parameters, each of which has been linked to carbon cycling in Lake Superior. After an iterative process seeking to increase the predictive capability of this model while minimizing error, the variables SST, downward photosynthetically active radiation at the air-water interface (PAR), parameterized primary production, and wind speed at 10 meters (U_{10}) were chosen (Table 4.1). The ability of these variables to act as proxies for Lake Superior pCO_2 has been established by previous studies describing water temperature and primary productivity as direct drivers of sea surface pCO_2 (Atilla et al., 2011; Minor et al., 2019; Sandborn & Minor, 2024), wind speed controlling air-sea CO_2 flux, and incident sunlight interacting with lake metabolism (Brothers & Sibley, 2018) and organic mater photodegradation in large lakes (Minor & Oyler, 2021).

The Great Lakes Surface Environmental Analysis (Woude & Liu, 2024) provided SST and ice coverage fraction (the latter of which was used only to estimate ice cover inhibition of gas flux) on an approximately $0.02° \times 0.02°$ grid with hourly resolution, which was aggregated as daily means. The North American Regional Reanalysis (Mesinger et al., 2006) provided daily mean downward PAR and U_{10} on an approximately $0.3° \times 0.3°$ grid, which were aggregated as daily means and bilinearly interpolated to match the SST resolution. Primary production was calculated from SST and PAR using the equation of Sterner (2010), which has been demonstrated to predict primary production better than chlorophyll a parameterizations (R.W. Sterner, personal communication). Explicit inclusion of primary production as a function of two other input parameters aligned with knowledge-guided machine learning theory by removing the need for implicit model inference of patterns of productivity from SST and PAR. Atmospheric dry CO_2 mixing fraction (xCO_2) was obtained from daily means of hourly observations of atmospheric xCO_2 observed at the WLEF

tall tower (Desai, 2022) and converted to pCO_2. These parameters all spanned the period January 2019-December 2023.

Calculations were completed with Python 3.11, using the packages Pandas (Reback et al., 2022), SciPy (Virtanen et al., 2020), Statsmodels (Seabold & Perktold, 2010), Numpy (C. R. Harris et al., 2020), PyCO2SYS (Humphreys et al., 2020), SeaFlux (Gregor & Humphreys, 2021), GSW-Python (Firing et al., 2021), Keras (Chollet, 2015), Scikit-learn (Pedregosa et al., 2011), xESMF (Zhuang et al., 2023), Cartopy (Elson et al., 2023), Matplotlib (Hunter, 2007), SHAP (Lundberg & Lee, 2017), and Seaborn (Waskom, 2021).

4.2.1 Model construction and training

The 4 input parameters were subjected to distribution standardization by removing the mean of the training data and scaling to unit variance. The FFNN was created with 4 input neurons (corresponding to the input parameters), two hidden layers of 128 and 32 neurons with rectified linear unit activation, and a single output neuron with linear activation. A dropout fraction of 0.2 between hidden layers (denoting that 20% of the training data was not used by the second hidden layer) and early stopping (in which model training ceased when model loss failed to decrease for 5 consecutive epochs) were used to prevent overtraining. The model was trained on a random selection of 80% of the underway observations, and tested with the remaining 20%. Model training sought to minimize mean absolute error of predictions. A train vs. test comparison of model predictions (Figure 4.1) displayed a mean absolute error of 27.0 µatm. The mean of the test pCO_2 observations was 371 µatm with a standard deviation of 61 µatm, while the mean of predicted test pCO_2 was 371 µatm with a standard deviation of 44 µatm, demonstrating comparable distribution means but a restricted prediction variance. The bimodal distribution of predictions closely resembled that of the test data, indicating that a major mode of seasonal variability (corresponding to spring conditions near atmospheric

pCO_2 equilibrium followed by summer CO_2 drawdown) was well-represented. The model captured more than two-thirds of test data variability, implying that while the majority of variability was accounted for, further improvements could be possible.

4.2.2 Model validation

Futher model validation was accomplished by comparison of model output against pCO_2 observed at a mooring or calculated from bottle sampling performed independently from model training and testing observations. The surface pCO_2 time series generated by this model overlapped with a deployment of a mooring in the western arm of Lake Superior at N 46.81° W 91.84° from October 2022 to May 2023. This mooring was equipped with a CO_2-Pro instrument (Pro-Oceanus) measuring pCO_2 with a manufacturer-specified accuracy of ±0.5 % at 10 m depth with hourly resolution.

The first source of bottle measurements used to validate the model presented in this work came from paired measurements of pH and A_T collected by the United States Environmental Protection Agency Great Lakes National Program Office (GLNPO) in 2019, 2021, 2022, and 2023. These data were obtained from the Great Lakes Environmental Database System, and were estimated to have uncertainties of ±50 µmol kg^{-1} for A_T and ±0.2 for pH (Minor & Brinkley, 2022).

Bottle samples were also collected from surface waters in western Lake Superior in 2020-2022, using standard oceanographic procedures (Dickson et al., 2007). Water was collected from Niskin bottles closed at approximately 2 m depth through pre-rinsed tubing into 500 mL flasks (for DIC and A_T) or 250 mL flasks (for pH) with a full-bottle overflow, given a 5 mL headspace, poisoned with 140 µL saturated $HgCl_2$ (Ricca), and sealed using Apiezon grease and rubber bands. pH was analyzed spectrophotometrically using purified *meta*-cresol purple dye within 30 minutes of sampling (Lai et al., 2016). DIC

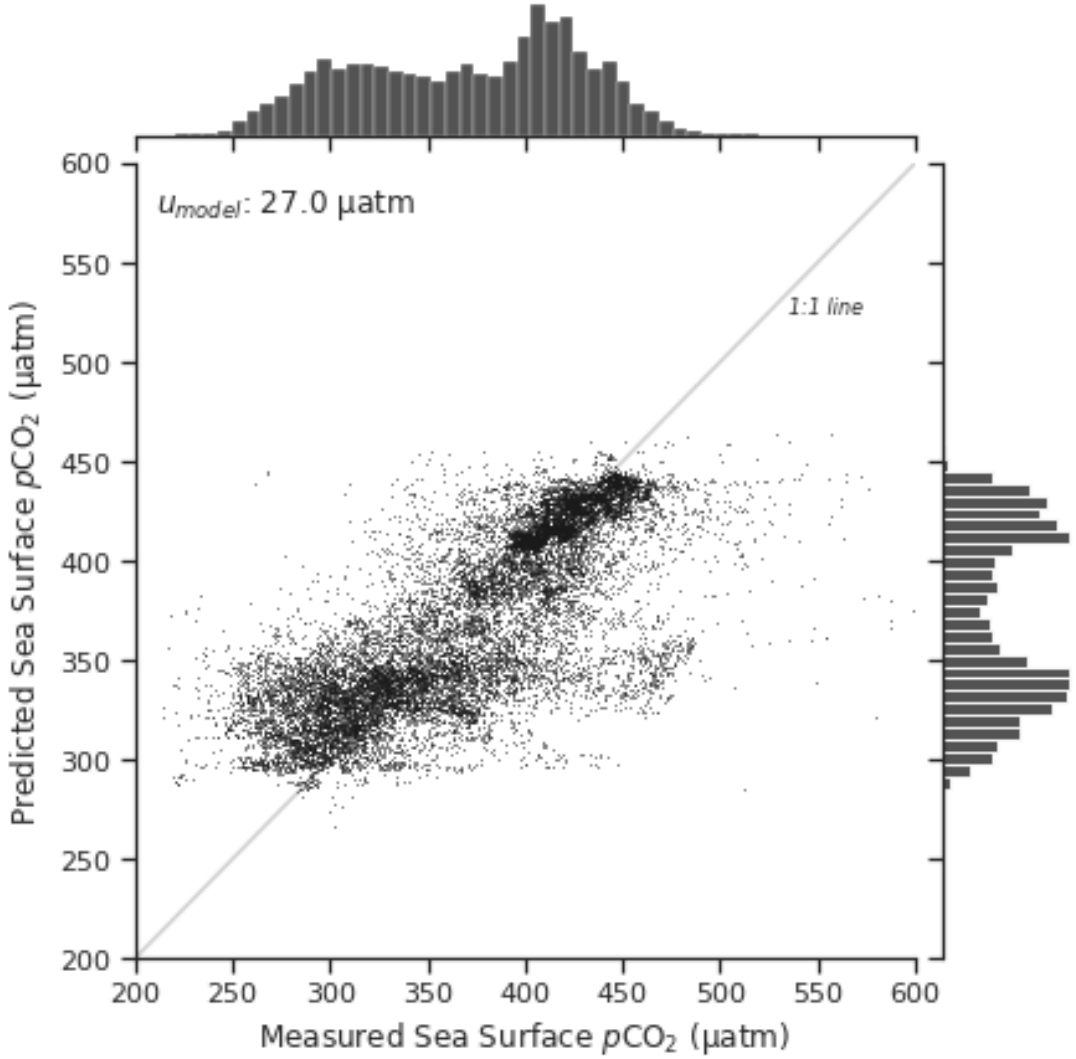

Figure 4.1: Blue Heron underway pCO_2 plotted against pCO_2 predicted via FFNN. The mean absolute error of this regression was 27.0 µatm.

was analyzed coulometrically (UIC CM150). A_T was analyzed as in Sandborn et al. (2023). DIC and A_T analyses of Certified Reference Materials demonstrated an uncertainty better than $\pm 5\,\mu mol\,kg^{-1}$ as judged by mean absolute error, while pH was assumed to have measurement uncertainty better than ±0.01.

For GLNPO bottle samples, pCO_2 was calculated from A_T and pH (DIC was not measured). The combination of A_T and DIC propagated with approximately 10 times the uncertainty of DIC and pH or A_T and pH (*cf.* Orr et al., 2018), so pCO_2 for the authors' bottle samples was calculated from the latter two pairs. The carbonate equilibrium constants of Waters et al. (2014) were used, along with typical sulfate (Chapra et al., 2012) and silica (Johnson & Eisenreich, 1979) concentrations. Total borate and fluoride concentrations were assumed to be zero. Salinity was calculated from conductivity (Hill et al., 1986) solely for equilibrium constant estimation, as major ion concentrations were explicitly specified.

4.2.3 Air-sea CO_2 flux estimation

Model output pCO_2 was used to estimate daily mean air-sea CO_2 flux, parameterized as the difference between modeled sea surface pCO_2 and observed atmospheric $pCO_{2\,atm}$ at the WLEF tall tower, multiplied by the gas transfer velocity k, a function of Schmidt number Sc (Ho et al., 2006), mean squared neutral wind speed at 10 meters above the sea surface ($< U^2_{10\,n} >$), the solubility of CO_2 in water K_o (R. Weiss, 1974), and the fraction of ice cover f_{ice}. Positive values of CO_2 flux denote efflux.

$$CO_2\ \text{Flux} = kK_o\,(pCO_2 - pCO_{2\,\text{atm}})\,(1 - f_{ice}) \tag{4.1}$$

$$k = 0.266 < U^2_{10} > \left(\frac{Sc}{600}\right)^{-0.5} \tag{4.2}$$

Table 4.1: Neural network model parameter metadata. Resolution indicates the temporal and spatial resolution of the provided gridded product before daily averaging and regridding to match the SST product. Regridding was accomplished via bilinear interpolation. For explanation of uncertainty associated with underway pCO_2 and model-predicted pCO_2 see *Model Uncertainty Analysis*. Note negative PAR values indicate downward radiation.

Input Parameter	**Uncertainty**	**Range**	**Source**
Sea surface pCO_2	5 µatm	198 — 1055 µatm	underway instrumentation
SST		0.45 — 23.54 °C	Great Lakes Surface Environmental Analysis
U_{10}		0 — 19.6 m s^{-1}	North American Regional Reanalysis
PAR		−45.8 — −388.0 W m^{-2} d^{-1}	North American Regional Reanalysis
Primary productivity		0.46 — 4.75 mg C m^{-2} d^{-1}	*derived*; Sterner (2010)
Output Parameter	**Uncertainty**	**Range**	
Sea surface pCO_2	27.5 µatm	283 — 515 µatm	
CO_2 flux		−96 — 31 mmol C m^{-2} d^{-1}	

Total net air-sea CO_2 flux over Lake Superior was calculated via integration of daily mean air-sea CO_2 flux using SeaFlux (Gregor & Humphreys, 2021) which integrated over the area of the model grid (77,110 km^2). The model grid area was 93.9% of the true surface area of Lake Superior, (82,100 km^2; Fuller and Shear, 1995), so a further factor of 0.939^{-1} was applied to total net air-sea CO_2 flux to scale it to the true surface area. The smaller size of the model grid stemmed from neglect of grid squares containing land, implying an offshore bias in the output of this model. The effect of this bias may be to increase or decrease net CO_2 flux depending on time and location, as nearshore riverine influence may enhance primary production and respiration of terrestrial organic material, shifting pCO_2 in opposing directions.

4.2.4 Model uncertainty analysis

Overall pCO_2 product uncertainty (u_{pCO_2}) was estimated after Landschützer et al. (2013) as the root sum of squared uncertainties of contributors to prediction uncertainty: uncertainty of the input observations (u_{obs}), uncertainty in aggregating observations into a grid (u_{agg}), and uncertainty in model prediction (u_{model}).

$$u_{pCO_2} = \sqrt{u_{obs}^2 + u_{agg}^2 + u_{model}^2} \tag{4.3}$$

u_{obs} was assigned based on the estimated measurement uncertainty (±5 µatm). u_{agg} was assigned based on the mean standard deviation among aggregated underway pCO_2 observations in the training data (±1.9 µatm). u_{model} was assigned as the mean absolute error (±27.0 µatm) of the train vs. test comparison (Figure 4.1). The resulting overall pCO_2 product uncertainty was ±27.5 µatm, dominated by u_{model}. This uncertainty was used to estimate whole-lake annual air-sea CO_2 flux uncertainty in a Monte Carlo reanalysis, described in the presentation of total air-sea CO_2 flux below.

4.3 Results

Daily mean sea surface pCO_2 and CO_2 flux were predicted for Lake Superior over the period 2019-2023 with a spatial resolution of 0.02°. Predicted pCO_2 ranged from 283 µatm to 515 µatm with a mean of 402 µatm and standard deviation of 46 µatm. Predicted CO_2 flux ranged from $-96\,\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$ to $31\,\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$ with a mean of $0.4\,\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$ and standard deviation of $4.0\,\mathrm{mmol\,C\,m^{-2}\,d^{-1}}$.

Mean daily values of pCO_2 and CO_2 flux and predictor variables were averaged over the entire model grid to produce mean time series (Figure 4.2). The mean annual pCO_2 time series agreed qualitatively with the double sinusoidal behavior modeled by Bennington et al. (2012), featuring two local minima falling in mid-February and late August, and two local maxima falling in early June and early December (Figure 4.3). The late August pCO_2 minimum aligned with the primary production maximum calculated here as well as by Sterner (2010), and the December pCO_2 maximum similarly aligned with the calculated primary production minimum. The mid-winter pCO_2 minimum aligned closely with the SST minimum, as expected from the increasing solubility of CO_2 with falling SST. Separating the temperature-driven effects of CO_2 solubility from the model-predicted time series (cf. Sandborn and Minor, 2024) illustrated how the climatology of predicted pCO_2 was driven by both thermal and non-thermal (primary productivity, respiration, mixing, CO_2 flux, etc.) drivers working in opposition to perturb pCO_2 from a mean value near atmospheric equilibrium, creating these four pCO_2 time series local minima and maxima (Figure 4.3). This analysis aided in explaining the timing of the spring pCO_2 maximum, which appeared with rising SST before falling sharply after biophysical drivers including primary productivity created the summer CO_2 drawdown. Similarly, the late-fall pCO_2 maximum appeared in spite of decreasing SST due to the influence of biophysical drivers likely including ventilation of hypolimnetic high-pCO_2 waters with weakening stratification.

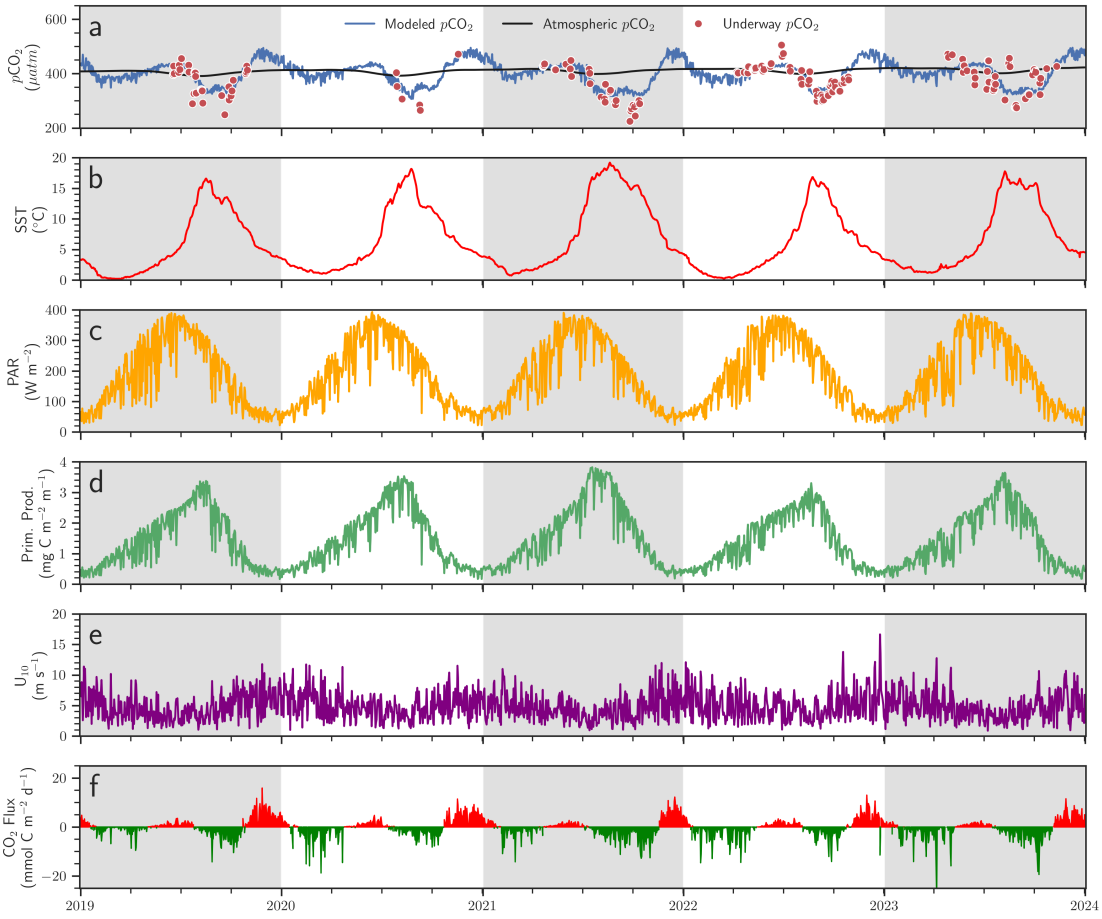

Figure 4.2: Time series of daily mean values of model input parameters (3a-e) and modeled lake surface pCO_2 (3a) and CO_2 flux (3f). Efflux is indicated as positive values.

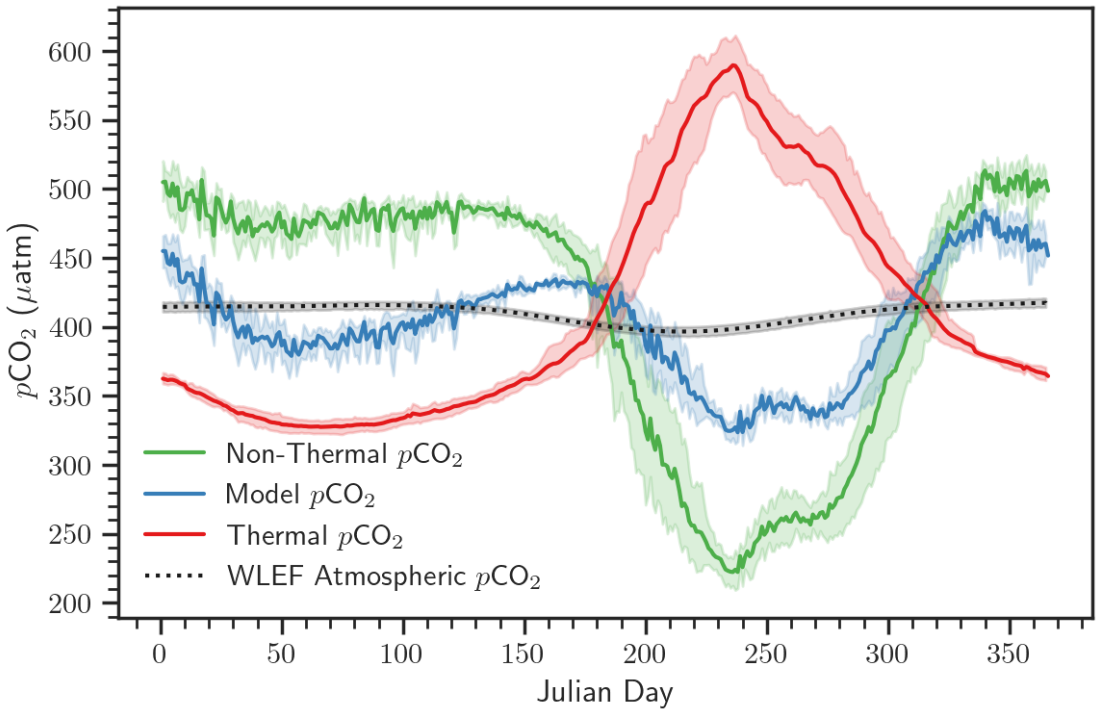

Figure 4.3: Mean annual sea surface pCO_2 cycle modeled for Lake Superior, with Thermal and Non-Thermal decomposition components and WLEF atmospheric pCO_2 cycle over the period 2019-2023. Solid lines indicate mean, and bounding strips indicate standard deviation.

A cursory analysis of model prediction parameter importance and effect was performed using Shapley additive explanation (SHAP) which provided approximate explanations of parameter effects on individual predictions. Mean absolute parameter effects (SHAP values) over the training dataset indicated the overriding importance of PAR, with a mean absolute SHAP value of 24 µatm, followed by SST and primary productvity with mean absolute SHAP values 18 µatm and 16 µatm, trailed by U_{10} with a mean absolute SHAP value of 2 µatm. The relative magnitudes of these four SHAP values indicated the relative importance of the four input parameters in determining predicted pCO_2. Comparison of input parameter values to corresponding SHAP values (Figure 4.4) indicated elevated primary productivity, SST, and negative PAR (corresponding with greater downward irradiation) were associated with decreased pCO_2. This relationship was expected for primary productivity and PAR as a proxies of CO_2 drawdown and seasonality. The case for SST was more complicated, for (as indicated previously) CO_2 solubility was expected to decrease with higher SST, which contrasted with the SHAP value analysis indicating that the SST input parameter acted as a proxy for other processes affecting pCO_2 such as spatial heterogeneity, seasonality, and mixing dynamics.

No significant trend was observed in a linear regression of deseasonalized mean sea surface pCO_2 or air-sea CO_2 flux. Interannual variability, clearly visible in wide error strips around mean annual pCO_2 (especially in spring and fall; Figure 4.3) likely overwhelmed the relatively minute atmospheric increase over 5 years, necessitating a longer time series for detection of such a trend (Carter et al., 2019).

4.3.1 Total air-sea CO_2 flux

Integration of CO_2 flux rates over space and time for the modeled period yielded a mean net annual CO_2 influx of $0.1398\,\mathrm{Tg\,C\,yr^{-1}}$ for the period 2019-2023. Year-by-year, this varied from an maximum influx of $0.296\,\mathrm{Tg\,C\,yr^{-1}}$ in 2023 to

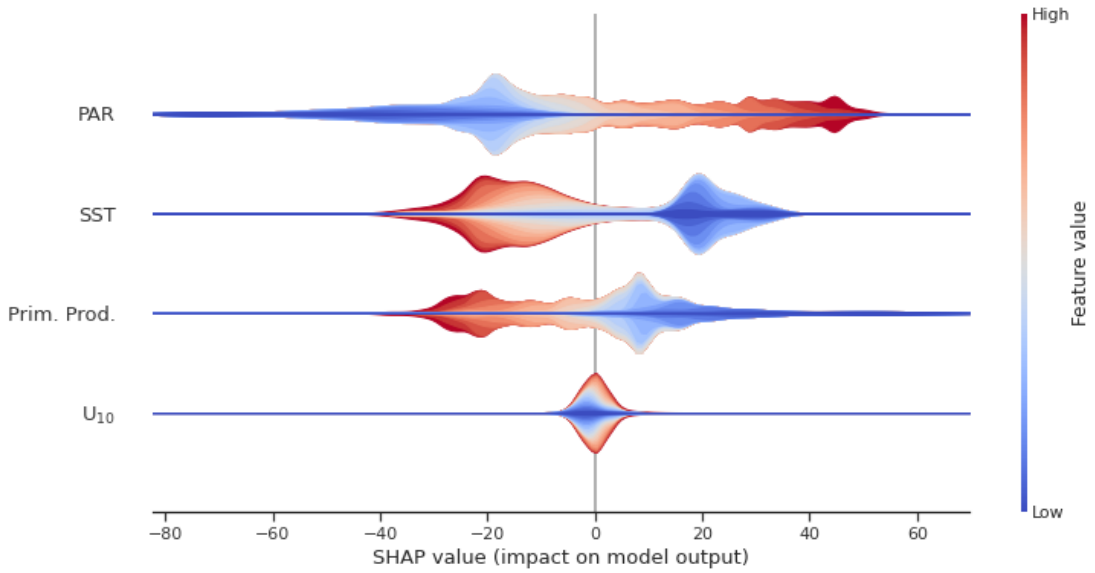

Figure 4.4: Comparison of SHAP values associated with prediction of pCO_2 by the four model input parameters. SHAP values on the x-axis indicate impact on model output, such that a positive SHAP value for a given parameter for a given prediction indicates that it positively biases the output value.

a maximum efflux of 0.058 Tg C yr^{-1} in 2019 (the only year in which a mean net annual efflux was predicted). Monte Carlo reanalysis simulated uncertainty of modeled CO_2 flux from u_{pCO_2} estimated above. Values pulled from a random gaussian distribution with a mean of 0 and a standard deviation of u_{pCO_2} were added to grid cells' modeled pCO_2 values, and this perturbed model output integrated to produce a value of total annual air-sea CO_2 flux. This process was repeated for 50 iterations, resulting in a distribution of total annual air-sea CO_2 flux with a standard deviation of 0.0001 Tg C yr^{-1} which was taken as the uncertainty of net annual CO_2 flux associated with this model. The variability of the five years' annual net CO_2 flux exceeded this value, implying that the modeled range represented significant interannual variability.

Net annual CO_2 flux predicted for the five years of this model fell between the mean annual 0.19 Tg C yr^{-1} efflux modeled for the period 1997-2001 by Bennington et al. (2012) and the mean annual 0.4 Tg C yr^{-1} influx inferred by Brothers and Sibley (2018) for the period 1968-2016. The annual mean ice-free season (Julian days 100-300) CO_2 net influx of 0.22 Tg C yr^{-1} was a close match to the 0.3 Tg C yr^{-1} estimated for the same period during the years 2019-2022 by Sandborn and Minor (2024). The influence of modes of interannual climate variability (El Niño-Southern Oscillation and Pacific Decadal Oscillation) on the metabolism of Lake Superior remained difficult to infer from model output. Brothers and Sibley (2018) associated a strong El Niño in 1997-1998 with a period of net heterotrophy in 1998-2001. In contrast, this model predicted a change from net CO_2 efflux in 2019 (associated with a weak El Niño) to net CO_2 influx in the following years (marked by a moderate La Niña 2020-2023 followed by a transition to a strong El Niño). Nearly the entire modeled period was associated with a negative-phase Pacific Decadal Oscillation, like the period 1998-2001.

The mean pCO_2 supersaturation predicted during winter is likely poorly-constrained by a lack of wintertime training data, yet comparison to the modeled time series of Bennington et al. (2012) supported this model's prediction

in both the location of the local minima and maxima (which agree within 2-3 weeks), the duration of super- and under-saturated periods, and the magnitude of springtime pCO_2 under- and super-saturation with respect to the atmosphere, which were approximately 25 µatm in both models at both times. The conclusion in our previous work (Sandborn & Minor, 2024) that Lake Superior pCO_2 has increased with atmospheric pCO_2 is reinforced by this parallel result.

4.3.2 Validation

Validation of modeled pCO_2 by mooring-observations and bottle samples demonstrated close agreement of the model predictions with in-situ conditions during both the ice-free season as well as in winter.

Comparison of mooring-measured and modeled pCO_2 time series at a mooring location (Figure 4.5) over the period of sensor deployment indicated substantive agreement in the fall and winter. The mean absolute error between modeled and observed pCO_2 was 32.5 µatm over October 2022-May 2023, yet during October 2022-December 2022 it was only 12.7 µatm. After January 2023 model output exceeded the measured value by an average of 40 µatm. The extent to which this was due to model error, sensor drift, or stratification was unclear. The CO_2-Pro sensor employed a zero-point correction, but previous deployments have indicated drift on the order of 5 ppm over months (Arruda et al., 2020). Wintertime negative stratification would be expected to isolate sensor measurement (at 10 m depth) from surface waters after early December when surface waters cooled to 4 °C, which approximately aligned with the divergence of mooring and model pCO_2, perhaps indicating mooring observation of more CO_2-rich hypolimnion-influenced water. Very weak winter stratification and variable winter mixed layer depths would be expected at the mooring location (J. Austin, personal communication), so more detailed analysis may be needed to understand physical mixing drivers of wintertime pCO_2variability

in Lake Superior.

This comparison of winter observations with model prediction highlights the need for improved winter observations to help validate models as in this work, overcoming one of the most significant barriers to formulating annual C budgets in seasonally ice-covered lakes.

Bottle samples provided surface water pCO_2 calculated from paired measurements of DIC and pH or A_T and pH obtained simultaneously with underway pCO_2 measurements. Comparison of calculated-modeled pCO_2 error (ΔpCO_2) against modeled values (Figure 4.6) indicated a mean absolute error of 12.7 µatm. Residuals calculated as modeled-observed pCO_2 did not appear to illustrate a significant trend vs. underway-measured pCO_2. GLNPO measurements were paired with co-located modeled pCO_2 predictions, which demonstrated a mean absolute error of 180 µatm, the magnitude of which was dominated by a minority of observations which displayed values far in excess of modeled values. Outlier points represented both spring and summer observations from multiple years (though dominated by observations in 2019 and 2021), and displayed no geographic pattern such as proximity to shore or major river outlets. The median absolute error for pCO_2 calculated from GLNPO bottle samples was 27 µatm, which differed insignificantly from the total model uncertainty (27.5 µatm) as well as uncertainty propagated from measurement uncertainty of GLNPO A_T and pH (c. 225 µatm) thus demonstrating close correspondence of observed and modeled pCO_2.

4.4 Discussion

This work detailed the development of the first neural network model of inorganic carbon cycling in a large lake. Despite the pioneering efforts of parallel oceanographic work which proved and advanced the utility of this method (e.g. Landschützer et al., 2013), much work remains in order to fully develop this mode of observation-based modeling as a useful biogeochemical tool for the

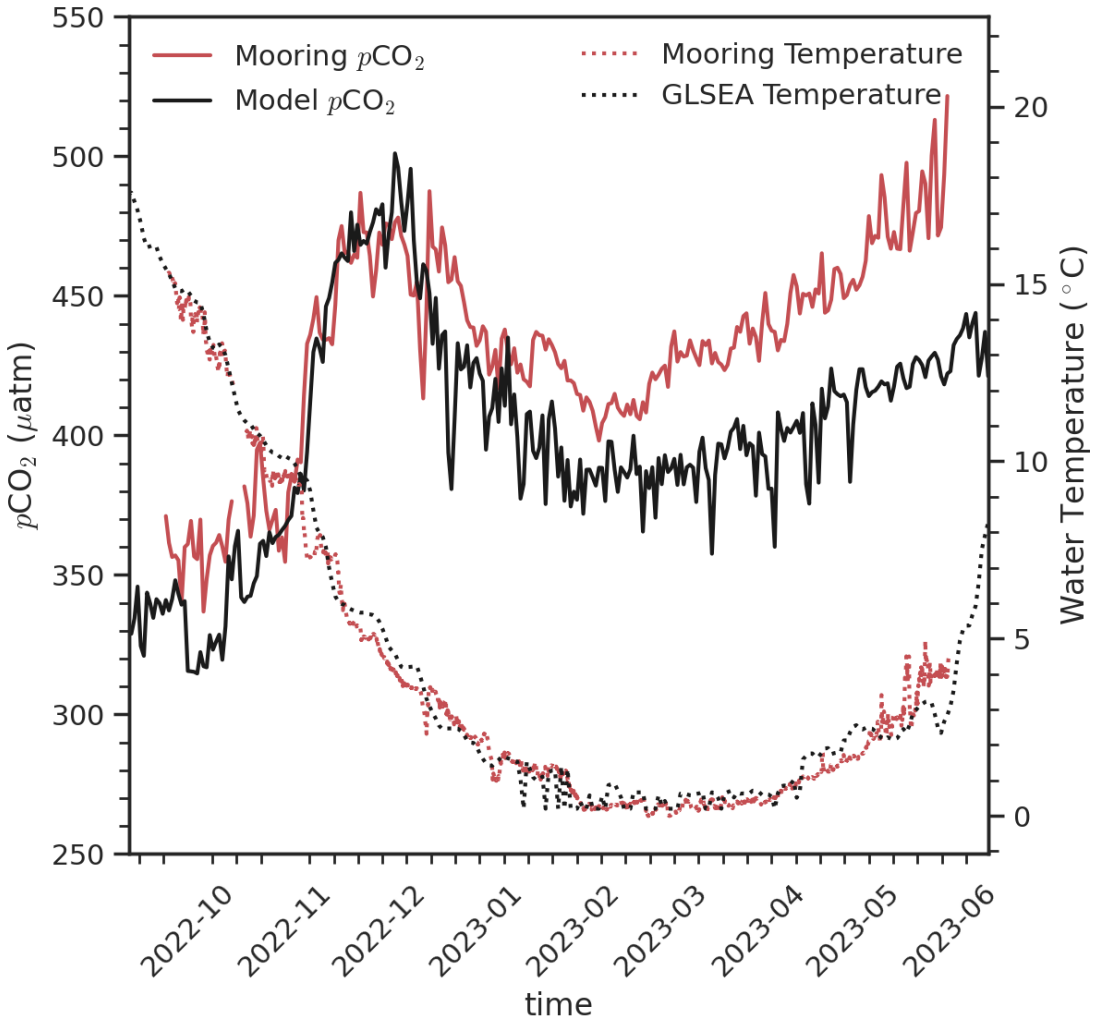

Figure 4.5: Time series of pCO_2 observed at an overwinter mooring and modeled at the same location and time period. The red line indicates hourly observations of pCO_2 at 10 m depth, while the black line indicates daily mean pCO_2 modeled values. Water temperature is indicated as dotted red (mooring-observed at 10 m) and black (Great Lakes Surface Environmental Analysis) lines.

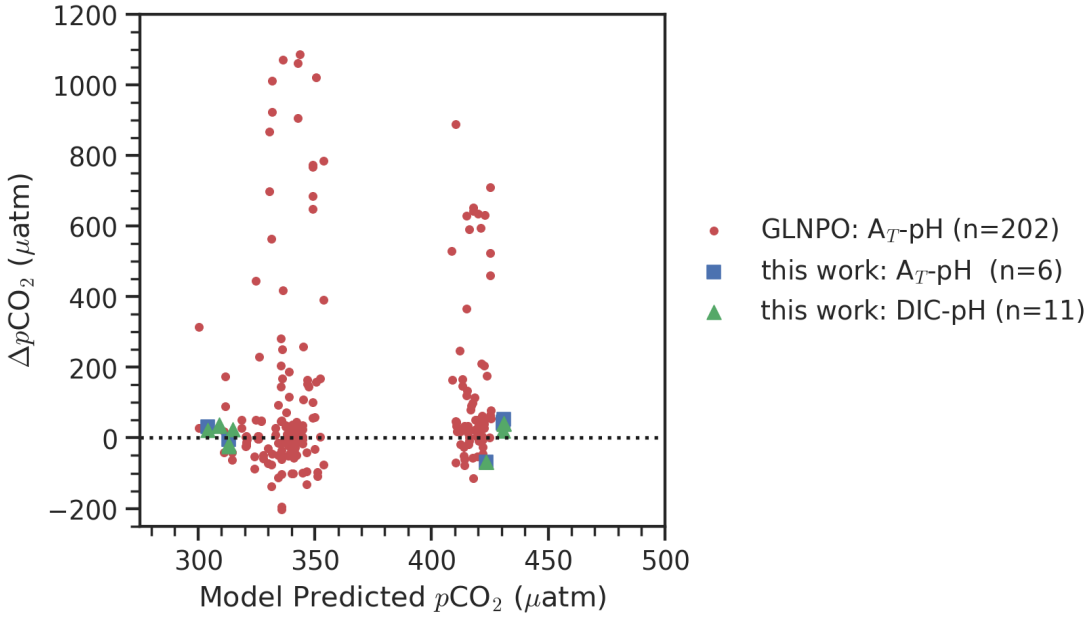

Figure 4.6: Plot of modeled-calculated pCO_2 error (ΔpCO_2) against modeled pCO_2. pCO_2 values calculated from GLNPO samples are in red circles. pCO_2 calculated from authors' measurements of A_T and pH or DIC and pH are in blue squares and green triangles, respectively.

study of inland waters.

Lake Superior provided an ideal test case for this work, as it remains the best-studied of the Great Lakes in terms of carbon biogeochemistry, and its large size and relatively-minimal terrestrial influence minimized noise that might overwhelm model training. While the few available gridded products suitable as model prediction parameters led to a model with sufficient accuracy and precision for the promulgation of a refined carbon budget term, wintertime uncertainties dominated the results. These problems and more are likely to apply to other large lakes to which a machine learning predictive model might be applied.

4.4.1 Toward a refined carbon budget

Comparison of the range net annual air-sea CO_2 flux values produced by this model for the period 2019-2023 to other known terms (Table 4.2) produces a refined picture of Lake Superior's carbon budget which approaches closure. The mass balance of carbon sinks and sources from the literature implies a range of net annual carbon flux between $0.95\,\mathrm{Tg\,C\,yr^{-1}}$ influx and $2.35\,\mathrm{Tg\,C\,yr^{-1}}$ efflux, which we assume is accounted for in a steady-state system by air-sea CO_2 gas flux. The modeled values in this study fall within this range and form a smaller range, decreasing the uncertainty and error of Lake Superior's carbon budget, despite the fact that the modeled net annual air-sea CO_2 flux in this work suffers from a lack of wintertime training data and unquantified associated error. Groundwater inputs and weathering are neglected in this budget.

Of the remaining terms, the largest uncertainties are associated with community respiration. Bennington et al. (2012) modeled an average annual community respiration rate of $5.45\,\mathrm{Tg\,C\,yr^{-1}}$, which when doubled (as suggested in that work to account for the inclusion of all tributary inputs, of which their work included 56% by flow volume) approaches the $10.78\,\mathrm{Tg\,C\,yr^{-1}}$ inferred

by Brothers and Sibley (2018). The separation of components of Lake Superior's carbon budget in time is another source of uncertainty. Warming waters (Austin & Colman, 2008), increasing pCO_2 (Minor & Brinkley, 2022; Sandborn & Minor, 2024), decreasing ice cover (P. Xue et al., 2022), changes in light penetration (Brothers & Sibley, 2018), and other modes of interannual variability may decrease comparability of terms in Lake Superior's carbon budget. Despite this shortcoming, calculation of Lake Superior's carbon residence time as sum of the lake's carbon pool of 125 Tg DIC and its organic carbon pool of 18.85 Tg C (see Chapter 1) over the average carbon efflux of 11.4 Tg C yr^{-1} yields a period of 12.6 yr, between the the DIC turnover time of 3-5 years and the DOC turnover time upper bound of <60 years found by Zigah et al. (2011).

The relationship between Lake Superior net annual CO_2 air-sea flux and El Niño-Southern Oscillation remains unclear, as the transition from El Niño to La Niña in 2019-2020 was associated with transition from net annual CO_2 air-sea influx to efflux in contrast with previously-observed lake metabolism relationships (Brothers & Sibley, 2018). This study's support for negligible annual net air-sea CO_2 flux over Lake Superior contrasts with latitudinal trends in large lake gas flux which indicate increasing efflux with latitude (Alin & Johnson, 2007). Sustained observation and modeling of biogeochemical cycling in Lake Superior and other large lakes may shed light on the interface of climate variability, climate change, and shifting element cycles.

4.4.2 Prospects of predictive modeling

The FFNN model in this work was trained on pCO_2 observations spanning the ice-free season, roughly April-November each year, yet its predictions extended into winter months and unobserved locations with the aid of modeled and observed parameters. Extrapolation carries the risk of introducing additional uncertainty into predicted values. Our choice of a machine learning

Table 4.2: Mass balance of estimated carbon fluxes (as total organic and inorganic carbon) into and out of Lake Superior. All values as Tg C yr^{-1}. The net flux of carbon from literature sources implies an range of air-sea CO_2 fluxes which overlaps with the values produced by this work.

all Tg C yr^{-1}	C_{in}	C_{out}	Notes & References
Primary Production	9.73		Sterner (2010)
In/outflow	0.4-0.9	0.1	Urban (2005)
Atmospheric Deposition	0.1-0.4		Cotner et al. (2004) and Urban (2005)
Sediment Burial		0.48	McManus et al. (2003)
Community Respiration		9.5-12.0	Brothers and Sibley (2018)
Totals	10.23-11.03	10.08-12.58	
Implied air-sea CO_2 flux	0.95 in – 2.35 out		$C_{out} - C_{in}$
Modeled air-sea CO_2 flux	0.30 in – 0.06 out		*this work*

model thought to be more robust to generalization beyond the domain of training data and inclusion of model parameters acting as proxies of in-situ drivers of pCO_2 improved model robustness, leading to satisfactory validation against independent measurements.

The model developed here represents an essentially statistical approach to predicting inorganic carbon cycling. Like other statistical-predictive models, it is not driven by explicit biogeochemical processes and as an empirical product its output resists inference of drivers. The resulting blindness of the model to *a priori* drivers can be a significant advantage as well as a central weakness.

Neural network and similar machine learning models have demonstrated superior ability to generalize and apply complex relationships compared to more transparent statistical alternatives, e.g. multiple linear regression (Vance et al., 2022) and random forest regression (Sharp et al., 2022). FFNN models are relatively easy to deploy with minimal assumptions of predictor or output behavior. These advantages must be weighed against the "black box" nature of this class of models; namely, that the complex nonlinear relationships among predictors and output are difficult to deconvolute. SHAP analysis as deployed in this work can aid in inferring input parameter importance and estimated output effect. Other strategies devised to aid inference of model behavior and underlying relationships include backward optimization and layerwise relevance propagation (Toms et al., 2020). Model output can also be analyzed as if it were actual observation, as in the discussion of seasonal trends, but once again nonlinearity may intercede to confuse quantitative interpretation. Inference of high-level phenomena like patterns of large-ecosystem element cycling then must rely upon antireductive (*cf.* Popper, 1966) interpretation of output at a high level at the cost of losing reductive tools which might illustrate drivers of these phenomena.

Overcoming limitations of machine learning models remains an active field of study in the Earth sciences (Karpatne et al., 2019). Knowledge-guided machine learning products (Daw et al., 2021) and hybrid models incorporating

feedback between earth system models and predictive machine learning products (Bennington et al., 2022; Gloege et al., 2022) have achieved improved model accuracy and better representation of biogeochemical processes. Both avenues of research are candidates for further development in the pursuit of improved knowledge of carbon cycling in large lakes.

4.5 Conclusions

The predictive gap-filling model of surface pCO_2 developed in this work takes a step towards closing the carbon budget of Lake Superior by producing the first observation-based, multi-year, spatially-comprehensive estimate of net annual air-sea CO_2 flux, which varied over the period 2019-2023 on both sides of zero net CO_2 flux. Leveraging the predictive capacity of available gridded input variables (SST, PAR, U_{10}, and calculated primary productivity) allowed this model to generalize patterns of pCO_2 variability to unobserved regions of Lake Superior, with consistency verified by moored sensor and bottle observations. Insofar as the resulting picture of pCO_2 and CO_2 flux expands the capabilities of present observation capacity, it also emphasizes the value of improved and expanded observations to inform statistical-predictive models such as the one presented here, as well as process-based biogeochemical models.

The central result of this work, an observation-based description of Lake Superior vacillating between net annual CO_2 influx and efflux, underscores the scientific understanding of Superior's inorganic carbon cycle as a surprisingly sensitive system to interannual climate variability in juxtaposition to its vastness, long hydrological residence time and decade-scale organic carbon turnover time (Minor & Oyler, 2021). The resulting illustration of a large lake sensitively integrating and feeding back to unknown modes of interannual climate variability resembles the work of Brothers and Sibley (2018) which inferred a relationship with ENSO, yet this work found no such correlation over

a 5-year period. This result may indicate more complex factors mediating interannual variability of inorganic carbon cycling in Lake Superior (and perhaps other large lakes) which deserve further study at the scale of individual lakes as well as the collection of large lakes across gradients of geography and climate.

While the five-year model product of pCO_2 and CO_2 flux spanning Lake Superior presented here helps unravel seasonal, interannual, and spatial variability in carbon cycling, it is fundamentally limited in predictive power outside the observed domain. This work supports previously-modeled and -inferred winter pCO_2 supersaturation, but can provide only a rough estimate of that season's net CO_2 efflux, and consequently the mean net CO_2 flux of the multiannual time series. Despite these drawbacks, we contend that this model output is consistent with the best knowledge of Lake Superior's C budget, and refines and extends it beyond previous work.

Acknowledgments

Thanks to Payton Kittaka, Ham Lam and the Minnesota Supercomputing Institute, Daniel Dominguez, the Chief Scientists and Primary Investigators of the transects providing underway data, and the Captain and crew of the *RV Blue Heron* for their assistance improving this work. This research was aided by a grant from the Great Lakes Observation System which supported mooring installation and maintenance.

Chapter 5

A survey of aragonite saturation state in lakes and association with calcifying mussel habitat

Abstract

A survey of aragonite saturation state in lakes of the world is presented with a discussion of its potential use as a predictor of invasive calcifying mussel habitat in inland waters. Calcium carbonate saturation state describes the thermodynamic stability of a calcium carbonate mineral (calcite, aragonite) in water as a function of pH, alkalinity, temperature, pressure, and ionic strength or salinity. Each of these variables have been discussed separately as predictors of calcifying organism survival and reproduction, yielding sometimes conflicting results. The history of calcite and aragonite saturation states as useful predictors of marine calcifying species ecology indicates that a similar index may have use in a lacustrine setting. Data from the Surface Water Chemistry database and the Great Lakes National Program Office water quality data record is presented to illustrate saturation state variability among and within

lakes. Cross-referencing lakes' median aragonite saturation against observations of *Dreissena polymorphia* in the Nonindigenous Aquatic Species database indicated a relationship between aragonite saturation and Zebra mussel mussel habitat characterized by lower distribution limits of aragonite saturation, pH, alkalinity, and calcium ion. Future observational and experimental work should investigate saturation state as a useful chemical-ecological variable in inland waters.

5.1 Introduction

We propose the use of a mineral solubility index to better summarize the geochemistry of inland waters' calcium carbonate cycling and connect chemical drivers with the ecology of calcifying bivalve species. Calcium carbonate saturation state (Ω) has been widely applied in the marine literature to describe precipitation and dissolution behavior of calcite (Ω_{Ca}) and aragonite (Ω_{Ar}) together with the biology of calcifying plankton, corals, and bivalves (R. A. Feely et al., 2004). The use of Ω in inland waters has been limited largely to estuarine research, which has demonstrated its use in illustrating inorganic carbon chemistry across the salinity spectrum (Reimer et al., 2023). The goal of this research is to briefly highlight the utility of an underutilized parameter bridging chemical context and ecological application.

Ω is calculated as the product of aqueous Ca^{2+} and $CO_3{}^{2-}$ activities (denoted with braces) divided by the solubility product K_{sp} of the mineral phase (calcite or aragonite):

$$\Omega = \frac{\left\{CO_3{}^{2-}\right\}\left\{Ca^{2+}\right\}}{K_{sp}} \tag{5.1}$$

but more commonly as the product of aqueous Ca^{2+} and $CO_3{}^{2-}$ concentrations (denoted with brackets) divided by the apparent solubility product $K_{sp}{}'$:

$$\Omega = \frac{\left[CO_3{}^{2-}\right]\left[Ca^{2+}\right]}{K_{sp}{}'} \tag{5.2}$$

Ω is a mechanistic driver of $CaCO_3$ solubility grounded in thermodynamic theory. Values greater than 1 indicate saturated conditions favoring precipitation, while undersaturated values less than 1 favor dissolution, with rates of dissolution and precipitation determined by the degree of super/undersaturation and environmental conditions (Adkins et al., 2021). The tolerance of calcifying organisms to undersaturated conditions (with respect to the mineral phase of their shells) depends on species, life stage, and metabolism, as the energetic cost of calcification increases with degree of undersaturation (Waldbusser et al., 2015). This thermodynamic description of $CaCO_3$ thus defines a kinetic constraint on biotic calcification which costs organisms energy to overcome (Waldbusser et al., 2016). While Ca^{2+} and $CO_3{}^{2-}$ concentrations are individually predictive of marine bivalve health, Ω can be a better predictor of marine mussel habitat and survivorship, especially for sensitive veliger life stages (Gazeau et al., 2010; Gazeau et al., 2011; Ragg et al., 2019). These marine studies worked at the species level, at which marine calcifying species' habitats are similarly negatively affected by increasingly acidic ocean waters (R. A. Feely et al., 2012). Community and ecosystem sensitivities to changes in saturation state have proven difficult to generalize in light of acclimation, adaptation, and interactions with other climate disruptions (Doney et al., 2020). These studies together provide the foundation for the central hypothesis of this work: that $CaCO_3$ saturation state has useful predictive capacity for the habitat of inland water calcifying species, particularly invasive mussels.

Many calcifying species in inland waters exhibit sensitivity to chemical parameters tied to Ω, including pH, Ca^{2+} concentration, alkalinity, and conductivity, which has led to these measurable parameters' use (together or separately) as habitat predictors of freshwater shell-forming invasive species including *Limnoperna fortunei* (Oliveira et al., 2010), *Dreissena polymorpha* and *bugensis* (Karatayev et al., 2015; Mcmahon, 1996; Mellina & Rasmussen, 1994; Ramcharan et al., 1992), *Orconectes rusticus*, *Viviparus georgianus*, and *Cipangopaludina chinensis* (Latzka et al., 2015). This work focuses on Zebra mussels (*Dreissena polymorpha*) as exemplar aquatic invasive species with potential Ω sensitivity, but the results inform work with other species. Quagga mussels (*Dreissena bugensis*) are generally more sparsely-observed, yet it is thought that their greater tolerance of colder, deeper waters, softer substrates, and perhaps $CaCO_3$ chemistry may indicate distinct chemical-ecological habitats (Karatayev & Burlakova, 2022).

Published chemical habitat requirements for *Dreissena* mussels are listed exhaustively by Mackie and Claudi (2009). Significant disagreement is evident the literature values, complicating their use as ecological variables. Calcium-based limits have been observed or proposed repeatedly, with lower limits varying from $28.3\,\mathrm{mg\,L^{-1}}$ (Ramcharan et al., 1992) to as low as $8.5\,\mathrm{mg\,L^{-1}}$ (Hincks & Mackie, 1997). Conflicting results have also been obtained among North American and European Dreissenid populations (Mackie & Claudi, 2009). These results have not led to well-defined Ca^{2+} habitat limits (Cohen & Weinstein, 2001). Reported pH limits are similarly confounded, with Dreissenid invasion potential bounded by lower limit pH values ranging from 7.0 (Sepulveda et al., 2023) to 7.8 (Mackie & Claudi, 2009), corresponding to a relative change in $[H^+]$ of 84%. Temperature, alkalinity, and conductivity limits have also been proposed, but are more poorly-defined (Nalepa & Schloesser, 2013).

Further inspiration for this work derives from observations of co-variance of pH and Ca^{2+} limits (Hincks & Mackie, 1997) and previous suggestions of mineral solubility indices (Latzka et al., 2015; Prescott et al., 2014). Decreasing pH

(assuming constant alkalinity) decreases $[CO_3^{2-}]$, increasing $[Ca^{2+}]$ necessary to maintain a given $CaCO_3$ saturation; the inverse is also true (Equation 5.2). The problems associated with multiple conflicting habitat predictors demonstrate the need for a more comprehensive evaluation of chemical drivers of Dreissenid habitat, which Ω may fulfill.

Calculation of Ω requires measurement or parameterization of $[Ca^{2+}]$, temperature, and salinity (if significant), and constraining an equilibrium calculation of $[CO_3^{2-}]$ by measuring at least two of its parameters: pH, dissolved inorganic carbon (DIC), total alkalinity (A_T), and partial pressure of CO_2 in water (pCO_2) (Zeebe & Wolf-Gladrow, 2001). The resulting saturation state carries the advantage of synthesizing disparate chemical variables into a single value, yet suffers from a not-insignificant sampling and analysis requirement which this study demonstrates.

Both under- and super-saturated values of Ω_{Ar} are observed in the Laurentian Great Lakes (hereafter Great Lakes). Whiting events in lakes Ontario, Erie, and Michigan (Peng & Effler, 2011) are the product of supersaturation. Lake Superior alone is thought to remain undersaturated with respect to aragonite, but large variability in previous estimates of Great Lakes Ω_{Ar} values and ranges has obscured discussion of drivers and consequences of Great Lakes $CaCO_3$ cycling (Sterner, 2021). It is likely that invasive species' spread westward has been hampered by the cold, oligotrophic, and corrosive waters of Lake Superior (Grigorovich et al., 2003), but the role of $CaCO_3$ undersaturation in inhibiting calcifying species has seldom been explicitly discussed in relation to Lake Superior. Just as in oceans, spatial and seasonal variability can play a large role in determining lacustrine saturation states, which may vary depending on biological activity, riverine input, and physical conditions.

The uncertainty surrounding Ω in inland waters and its relationship with calcifying species biology and ecology represents a knowledge gap that hinders the understanding of invasive species management and ecological forecasting in a changing world. This study poses four central questions which explore the

utility of $CaCO_3$ saturation state as a chemical, ecological, and management tool:

1. How is $CaCO_3$ saturation state calculated?
2. What is the distribution of $CaCO_3$ saturation state in inland waters?
3. How does $CaCO_3$ saturation state compare to other chemical predictors of calcifying species habitat?
4. What science is needed to enable the use of $CaCO_3$ saturation state as a research and management tool?

These research questions are advanced in this communication with brief examples from water chemistry and invasive species databases, new measurements, and reanalysis and synthesis of literature. Although not comprehensive in geographic scope or consideration of calcifying species, this work aims to lay the foundation for such advances.

5.2 Methods

Water chemistry data for 236 inland lakes and ponds were obtained from the SWatCh Version 2 database (Rotteveel et al., 2022), which comprises measurements from more than 33,000 inland water sites around the world. Only freshwater (salinity ≤ 1) lakes and ponds which had contemporaneous observations of pH and water temperature, $[Ca^{2+}]$, and at least one of alkalinity (total, carbonate, or acid neutralizing capacity) or inorganic carbon (total or dissolved) were accepted, yielding 1461 observations. Ω_{Ar} was calculated from each set of contemporaneous observations using PyCO2SYS software (Humphreys et al., 2020). Only Ω_{Ar}, rather than Ω_{Ca}, was calculated here and elsewhere in this work in the pursuit of brevity, and in light of the fact that Dreissenid mussel shells are composed of aragonite. Carbonate equilibrium constants from

Waters et al. (2014) were used, salinity was estimated for the purposes of equilibrium calculation from conductivity after Hill et al. (1986), and total borate and fluoride concentrations were assumed to be zero. The resulting values of $[CO_3{}^{2-}]$ and Ω_{Ar} were summarized by taking the median value for each body. The resulting data describe lakes and ponds in Austria, Italy, Lithuania, and Canada. These data were supplemented with the LAGOS-US LIMNO (hereafter LAGOS) database (Shuvo et al., 2023) which provided paired in-situ measurements of pH, water temperature, alkalinity, and $[Ca^{2+}]$ for 3666 freshwater lakes in the contiguous United States. $[CO_3{}^{2-}]$ and Ω_{Ar} were calculated as above.

The Great Lakes were treated separately using Great Lakes National Program Office survey data from 1986 to 2023 (Great Lakes National Program Office, 2023). Contemporaneous observations of water temperature, pH, and A_T were paired with measured or estimated $[Ca^{2+}]$ and $[CO_3{}^{2-}]$ and Ω_{Ar} calculated as above. Estimations of $[Ca^{2+}]$ were calculated as the mean $[Ca^{2+}]$ for each lake given for the year 2009 by Chapra et al. (2012) multiplied by the ratio of the measured specific conductivity to the mean lake specific conductivity reported by Chapra et al. Comparison of calculated to estimated $[Ca^{2+}]$ indicated a median relative error of less than 3.8% using this method. This error was considered negligible in light of the domination of within-lake Ω_{Ar} variance by $[CO_3{}^{2-}]$ (Section 5.3.1).

Comparison of Ω_{Ar} with Zebra mussel habitat was enabled by the Nonindigenous Aquatic Species database (Benson et al., 2023) which catalogs observations of Zebra mussels (*Dreissena polymorphia*) in water bodies of the United States. Considering 2609 observations in lakes or ponds and cross-referencing them against the same lakes in the LAGOS dataset yielded 134 lakes or ponds with confirmed observations of Zebra mussel presence and sufficient chemical observation to calculate median $[CO_3{}^{2-}]$ and Ω_{Ar} values.

5.3 Results

Three databases providing standardized water chemistry descriptions of surface lakes and ponds around the world informed this survey of calcium carbonate saturation state distribution. Two classes of lakes were considered separately: the Great Lakes provided a demonstration of spatiotemporal variability in Ω_{Ar} *within* five large lakes spanning a gradient of geochemical and hydrological context, while the remaining lakes considered in this work provided a demonstration of variability in Ω_{Ar} *among* lakes of the world. A subset of this latter category were further classified by observations of Zebra mussels. The chemical characteristics of the lakes in which these mussels invaded enabled discussion of chemical drivers of invasive calcifying mussel habitat.

5.3.1 Distribution of Ω_{Ar} in the Great Lakes

The large geochemical and hydrological context spanned by the Great Lakes was reflected in the distribution of Ω_{Ar}, $[Ca^{2+}]$, and $[{CO_3}^{2-}]$ observed in each lake (Figure 5.1). Of 26,936 observations, 4786 were of Lake Superior, 6234 were of Lake Michigan, 5901 were of Lake Huron, 6401 were of Lake Erie, and 3614 were of Lake Ontario. The observed or calculated $[Ca^{2+}]$ for each lake displayed a relatively narrow distribution, except in the case of Lake Erie, for which a subset of observations located in the highly-temporally-variable western basin displayed depressed $[Ca^{2+}]$. The mean values of Ω_{Ar} generally bounded those calculated by Sterner (2021), which did not consider spring and summer values separately. Variability in Ω_{Ar} was due largely to the relatively broad range of $[{CO_3}^{2-}]$ observed in each system.

Seasonal differences in Ω_{Ar}, $[Ca^{2+}]$, $[{CO_3}^{2-}]$, and pH distributions (logarithmically transformed to highlight relative changes among and within lakes) indicated the large role of seasonal variability in depressing Spring $[{CO_3}^{2-}]$ and thus Ω_{Ar} to a lower and more limited range in each lake (Figure 5.2). Superior

was the only system which remained undersaturated with respect to aragonite (i.e. $\Omega_{Ar} < 1$) at nearly all observed stations throughout the year, though median spring Ω_{Ar} was undersaturated with respect to aragonite for Lakes Huron, Erie, and Ontario as well. A statistical summary of Ω_{Ar}, $[CO_3^{2-}]$, and $[Ca^{2+}]$ was compiled for each lake and season (Table 5.1). Similar summaries may be found for pH, A_T, and Ca^{2+} elsewhere (Chapra et al., 2012; Minor & Brinkley, 2022; Minor et al., 2019).

One potential problem with this overview of Ω_{Ar} in the Great Lakes is the uncertainty in $[CO_3^{2-}]$ propagated from uncertainty in pH and A_T. Assuming a pH uncertainty of ±0.2 and an A_T uncertainty of $\pm 50\,\mu mol\,kg^{-1}$ (Minor & Brinkley, 2022), the mean propagated uncertainty in $[CO_3^{2-}]$ for all five lakes was $7.6\,\mu mol\,kg^{-1}$, and the mean propagated uncertainty in Ω_{Ar} for all five lakes was 0.6, both representing a relative uncertainty of 49%. This magnitude of obscured seasonal variability of $CaCO_3$ under- or super-saturation and inhibited time series or spatial variability analysis in this work. Strategies for circumventing measurement uncertainty challenges are discussed below.

The survey of Ω_{Ar} in the Great Lakes provided here expanded on previous descriptions of $CaCO_3$ chemistry in the world's largest contiguous freshwater body with particular focus on Ω_{Ar} variability *within* lakes. The 30-year GLNPO observational dataset illustrated significant variability in Ω_{Ar} on seasonal scales through the action of variable $[CO_3^{2-}]$ rather than the more constant $[Ca^{2+}]$. Examining Ω_{Ar} variability *among* lakes required chemical characterization of many more lakes, fulfilled in this work by the SWatCh and LAGOS datasets.

5.3.2 Distribution of Ω_{Ar} in world lakes

A distribution of world lake Ω_{Ar} was produced for lakes described by the SWatCh and LAGOS datasets, which indicated patterns and variability in $CaCO_3$ solubility chemistry among lakes, a third dataset category representing the intersection of LAGOS and NAS, is discussed in Section 5.3.3. In contrast to

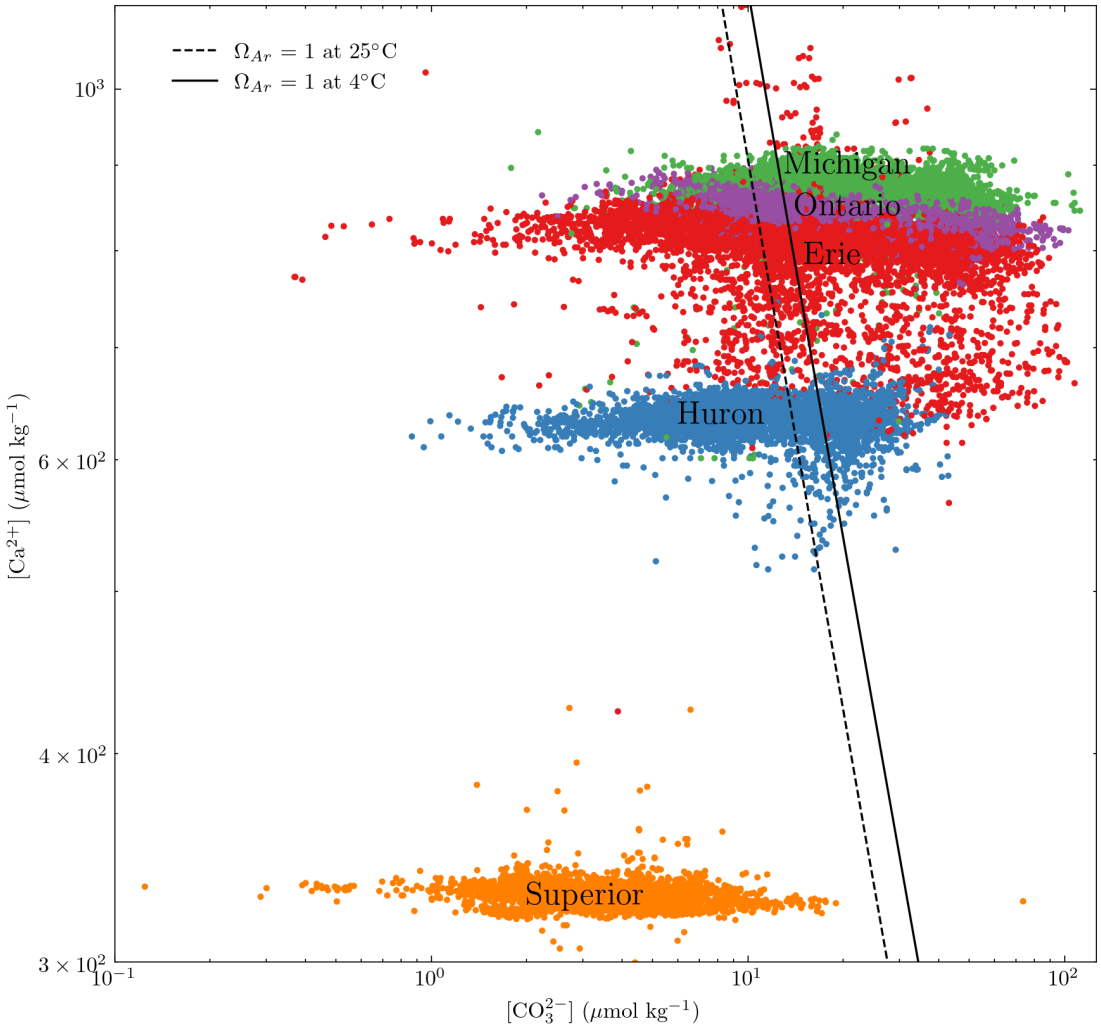

Figure 5.1: Plot of observations of Laurentian Great Lake concentrations of Ca^{2+} and ${CO_3}^{2-}$ calculated from GLNPO observations obtained in early Spring and late Summer campaigns in each lake during most years 1986-2023. Contour lines indicate $\Omega_{Ar} = 1$ (saturation) at 4 °C (dashed) and 25 °C (solid). From highest to lowest mean $[Ca^{2+}]$, the lakes Michigan, Ontario, Erie, Huron, and Superior display relatively little variation in $[Ca^{2+}]$, generally no more than a tenth of an order of magnitude (except for Erie) but much greater variability in $[{CO_3}^{2-}]$ (exceeding an order of magnitude), which is largely the effect of temporal variability. Note logarithmic transformation of the axes, which enables illustration of a linear aragonite saturation and better differentiates the five lakes across 3 orders of magnitude ($[{CO_3}^{2-}]$) and less than 1 order of magnitude ($[Ca^{2+}]$).

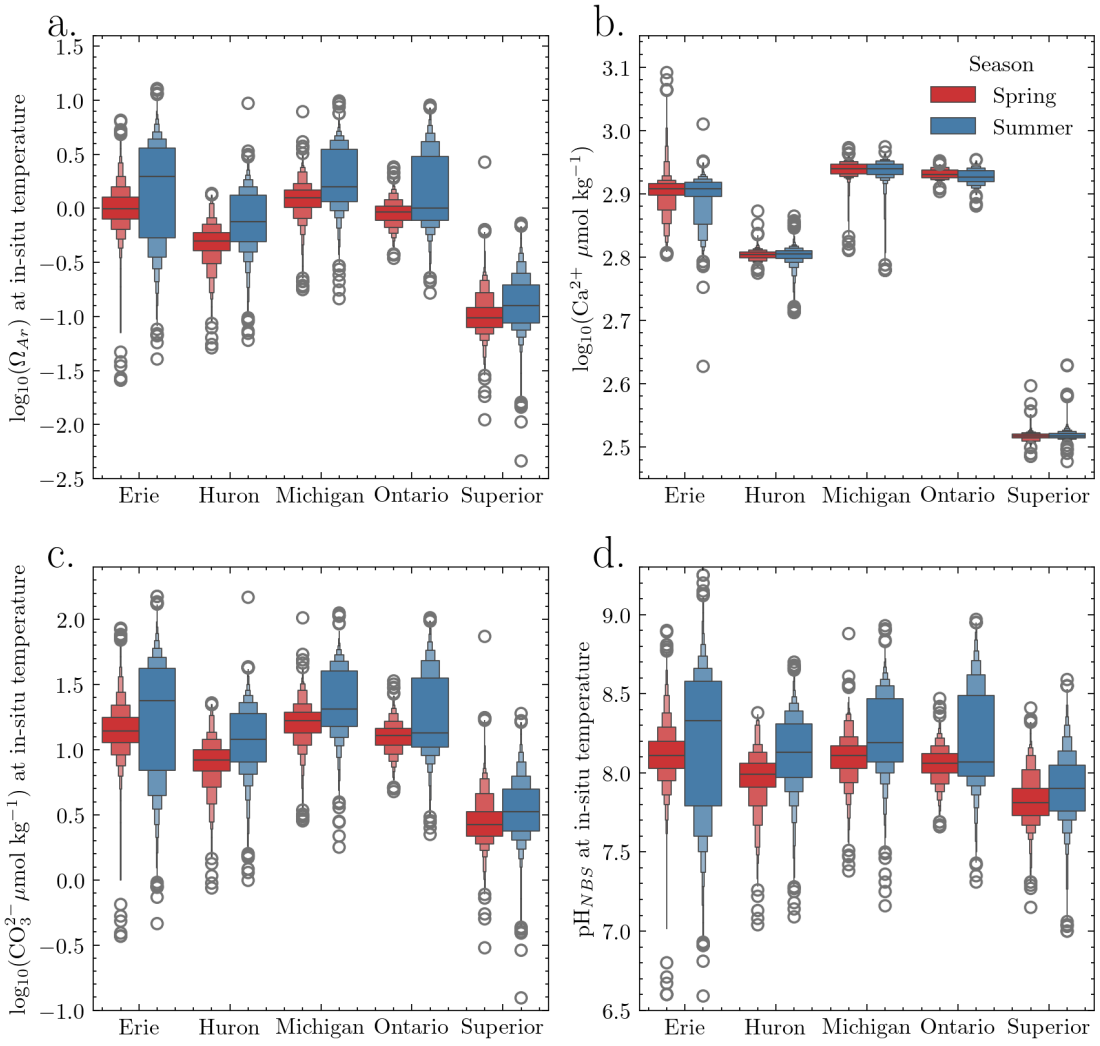

Figure 5.2: Decile plot of seasonal distributions of Decile plot of distributions of **a.** Ω_{Ar} **b.** Ca^{2+}, **c.** ${CO_3}^{2-}$, and **d.** pH in the Laurentian Great Lakes from GLNPO observations. Median values are indicated by the line in the center of the largest box, with decile values indicated by lines above and below. Logarithmic transforms were performed as indicated on axis labels in order to aid comparison of distribution differences and ranges across each lake and season.

Table 5.1: Statistical summary of Ω_{Ar} and $[CO_3{}^{2-}]$ observed or calculated in the Laurentian Great Lakes from the GLNPO dataset, 1986-2023.

		Superior		**Michigan**		**Huron**		**Erie**		**Ontario**	
		Spring	*Summer*	*Spring*	*Summer*	*Spring*	*Summer*	*Spring*	*Summer*	*Spring*	*Summer*
	count	2054	2732	2706	3528	2560	3341	2998	3403	1524	2090
Ω_{Ar}	1^{st} quartile	0.079	0.087	1.02	1.2	0.41	0.49	0.80	0.5	0.78	0.8
	mean	0.112	0.154	1.29	2.6	0.52	0.93	1.12	2.2	0.94	1.9
	median	0.097	0.126	1.26	1.6	0.50	0.75	0.99	2.0	0.93	1.0
	3^{rd} quartile	0.121	0.197	1.47	3.5	0.60	1.32	1.27	3.6	1.05	3.0
	std. dev.	0.079	0.094	0.47	1.4	0.21	0.56	0.59	1.8	0.25	1.7
$[CO_3{}^{2-}]$ ($\mu mol\,kg^{-1}$)	1^{st} quartile	2.2	2.4	13.5	15.	6.9	8.1	11.3	6.	10.8	10.
	mean	3.1	4.0	17.1	27.	8.7	13.9	15.6	26.	13.0	23.
	median	2.7	3.4	16.7	20.	8.3	12.0	13.9	24.	12.8	13.
	3^{rd} quartile	3.3	5.0	19.4	40.	9.9	19.0	17.6	42.	14.5	36.
	std. dev.	2.2	2.3	6.1	16.	3.5	7.7	8.2	21.	3.5	19.

previously-considered distributions of Ω_{Ar} in the Great Lakes (Figure 5.1) which varied largely due to variations in $[CO_3^{2-}]$ within lakes, median Ω_{Ar} in 3493 world lakes varied approximately linearly against log-transformed $[CO_3^{2-}]$ and log-transformed $[Ca^{2+}]$ among lakes (Figure 5.4b, c). A linear regression describing the relationship between calcium (y) and carbonate (x) ion concentrations gave a slope of 0.355 ± 0.005 with an intercept 2.301 ± 0.005 for the LAGOS data and a slope of 0.376 ± 0.005 with an intercept 2.562 ± 0.009 for the SWatCh data ($\pm$ standard error of type-I linear regression). These lines roughly paralleled the charge balance approximated by $2[Ca^{2+}] = [HCO3^{-}]$ and previously observed in various major rivers of the world (Holland, 1978), but with the HCO_3^- axis transformed to CO_3^{2-} assuming a range of water pCO_2 levels (indicated in Figure 5.4) to illustrate the linear relationship between log-$[Ca^{2+}]$ and log-$[CO_3^{2-}]$. Departures from this charge balance relationship may result from runoff richer in other ions, e.g. K^+, $Fe^{2+/3+}$, Al^{3+}, the ratios of which vary over continental settings (Mackenzie & Garrels, 1971).

Both datasets included lakes supersaturated and undersaturated with respect to aragonite. Mean Ω_{Ar} values of 3.1 for LAGOS lakes and 2.5 for SWatCh lakes indicated distributions skewed towards supersaturation, yet the standard deviation of Ω_{Ar} approached or exceed the mean values, demonstrating the wide range of aragonite solubility behavior found in world lakes (Table 5.2). The shape of the Ω_{Ar} distribution in both datasets was essentially the same, along with that of the $[CO_3^{2-}]$, $[Ca_2^+]$, and pH distributions (Figure 5.3b, c), though each of the four parameters exhibited a higher median value in SWatCh lakes relative to LAGOS lakes. An analysis of error or uncertainty as demonstrated for the Great Lakes in the previous section was not carried out, as the means for inferring uncertainty (standards, replication, etc.) were not reported.

Estimation of the world lake distribution of Ω_{Ar} enabled by consideration of the LAGOS and SWatCh datasets illustrated covariance of Ω_{Ar} with both $[CO_3^{2-}]$ and $[Ca_2^+]$ in a reproducible fashion predicted by charge balance.

Table 5.2: Statistical summary of Ω_{Ar} in world lakes, calculated from observations in the SWatCh or LAGOS dataset, together with the subset of the LAGOS dataset overlapping with NAS database observations (denoted NAS+LAGOS) of Zebra mussels (*D. polymorphia*) in lakes and ponds. 5^{th} and 10^{th} percentiles are provided for the NAS+LAGOS observations to indicate a range of potential chemical habitat limits, as discussed in Section 5.3.3

		NAS+LAGOS	**SWatCh**	**LAGOS**
	count	131	233	3260
Ω_{Ar} ($\mu mol\,kg^{-1}$)	5^{th} percentile	0.5		
	10^{th} percentile	1.0		
	1^{st} quartile	2.6	0.1	<0.1
	mean	5.0	3.1	2.5
	median	4.6	2.8	0.4
	3^{rd} quartile	6.6	5.0	3.9
	std. dev.	3.3	2.9	3.9

Combination of these chemical characterizations of $CaCO_3$ solubility in lakes with information about the habitat of a calcifying mussel next led to inference of ecologic behavior associated with chemical context.

5.3.3 *Dreissena polymorphia* and Ω_{Ar}

Calculated Ω_{Ar} values in the Great Lakes and selection of world lakes spanned a large range of undersaturated to supersaturated conditions with respect to aragonite. Together with a set of related water quality parameters previously proposed as predictors of calcifying species habitat including pH and $[Ca^{2+}]$, these variables were compared in order to evaluate their relative ability to constrain Dreissenid habitat. The distributions of pH, $[Ca^{2+}]$, $[{CO_3}^{2-}]$, and Ω_{Ar} in the intersection of NAS and LAGOS data (representing invaded lakes) indicated a restricted distribution of each variable relative to world lakes represented in SWatCh and LAGOS data sets (Figure 5.3).

This lack of observation of Zebra mussel populations in lakes with relatively lower pH, $[Ca^{2+}]$, and $[{CO_3}^{2-}]$ implied an association between chemical environment and biological habitat for this species in agreement with previous

studies. A novel finding in this work was the parallel association of Ω_{Ar} and Zebra mussel observations. The 10^{th} percentile in-situ Ω_{Ar} value in invaded lakes was 1.0, precisely at aragonite saturation (Table 5.2), while the 5^{th} percentile Ω_{Ar} value was 0.5, reflecting the presence of Zebra mussels in a handful of lakes with median conditions reflecting aragonite undersaturation. These two values illustrated a band of slightly-aragonite-undersaturated conditions associated with an apparent boundary of Zebra mussel habitat. In line with the understanding of biological effects of $CaCO_3$ undersaturation on calcifying organisms, the region $0.5 \leq \Omega_{Ar} \leq 1.0$ may present a region of increasing kinetic constraint upon biotic calcification by Zebra mussels requiring increasing energy expenditure.

The commonly-cited $[Ca^{2+}]$ constraint of Zebra mussel habitat of $12\,mg\,L^{-1}$ or $300\,\mu mol\,kg^{-1}$ (Mcmahon, 1996) overlaps with the cluster of invaded lakes in Figure 5.4, yet excludes several invaded lakes which were found to have median Ω_{Ar}near or above saturation. This result underlines potential pitfalls of single-parameter delineations of invasive species habitat, which has been echoed by other research (Karatayev et al., 2015). A comprehensive test of predictive ability was not possible in this observational survey, which considered lakes classified by whether Zebra mussels had been observed, rather than whether a lake had been invaded, or whether Zebra mussels survived and reproduced. Among lakes without observed populations of Zebra mussels (i.e. not part of the NAS data set), there existed the twin possibilities of an unobserved invasion in a chemically-suitable lake (a false negative), or a failed invasion of an chemically-suitable lake (a false positive). For this reason, the data presented in this work was used only to suggest likely chemical drivers of Zebra mussel habitat limits indicated by conditions in invaded lakes, rather than a causal connection. This result suggests several opportunities for further experimental study (Section 5.4.1).

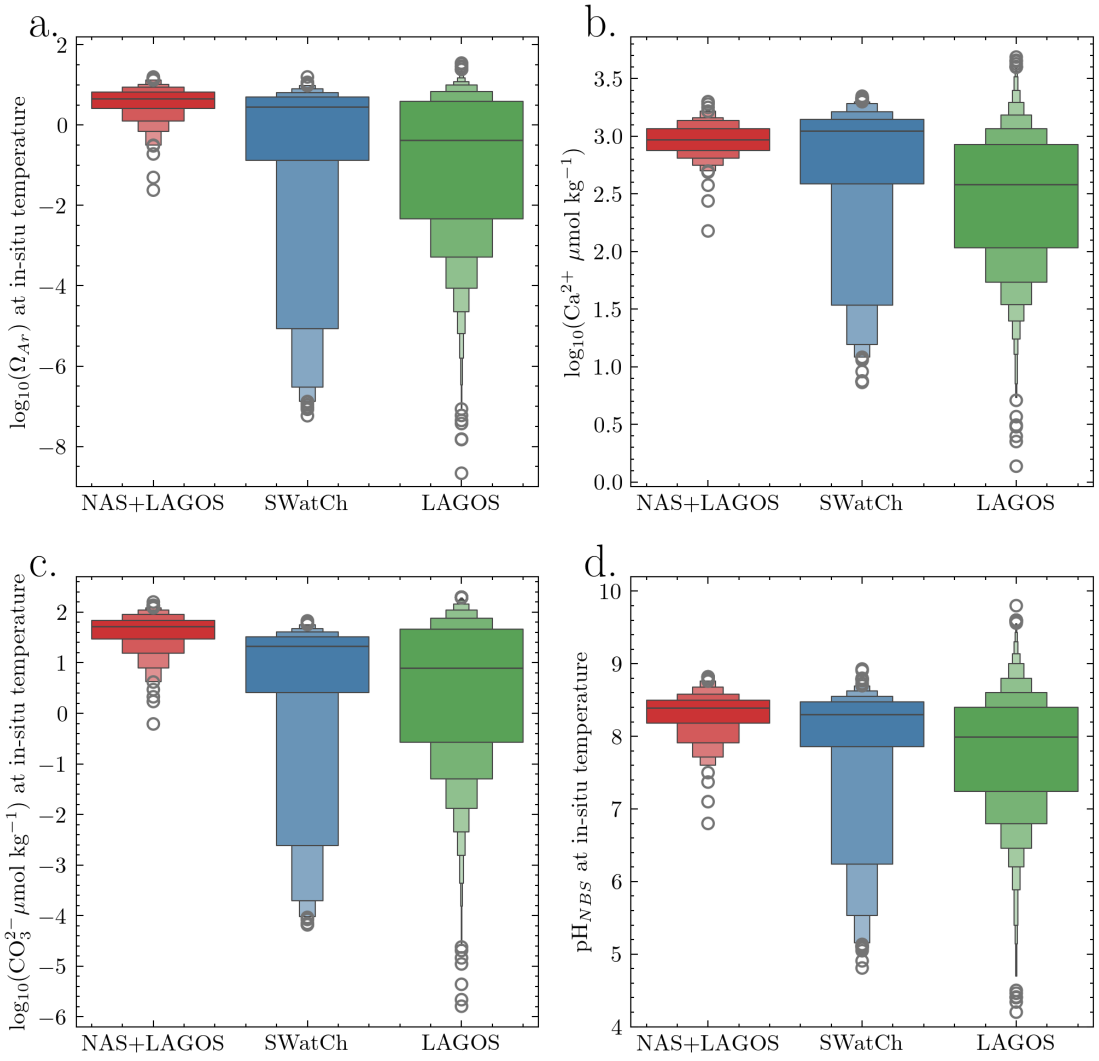

Figure 5.3: Decile plot of distributions of **a.** Ω_{Ar} **b.** Ca^{2+}, **c.** $CO_3{}^{2-}$, and **d.** pH in world lakes, calculated from the SWatCh or LAGOS datasets, as well as the subset of the LAGOS dataset overlapping with NAS database obervations of Zebra mussels (denoted NAS+LAGOS) in lakes and ponds. Median values are indicated by the line in the center of the largest box, with decile values indicated by lines above and below. Logarithmic transforms were performed as indicated on axis labels in order to aid comparison of distribution differences and ranges across each lake and season.

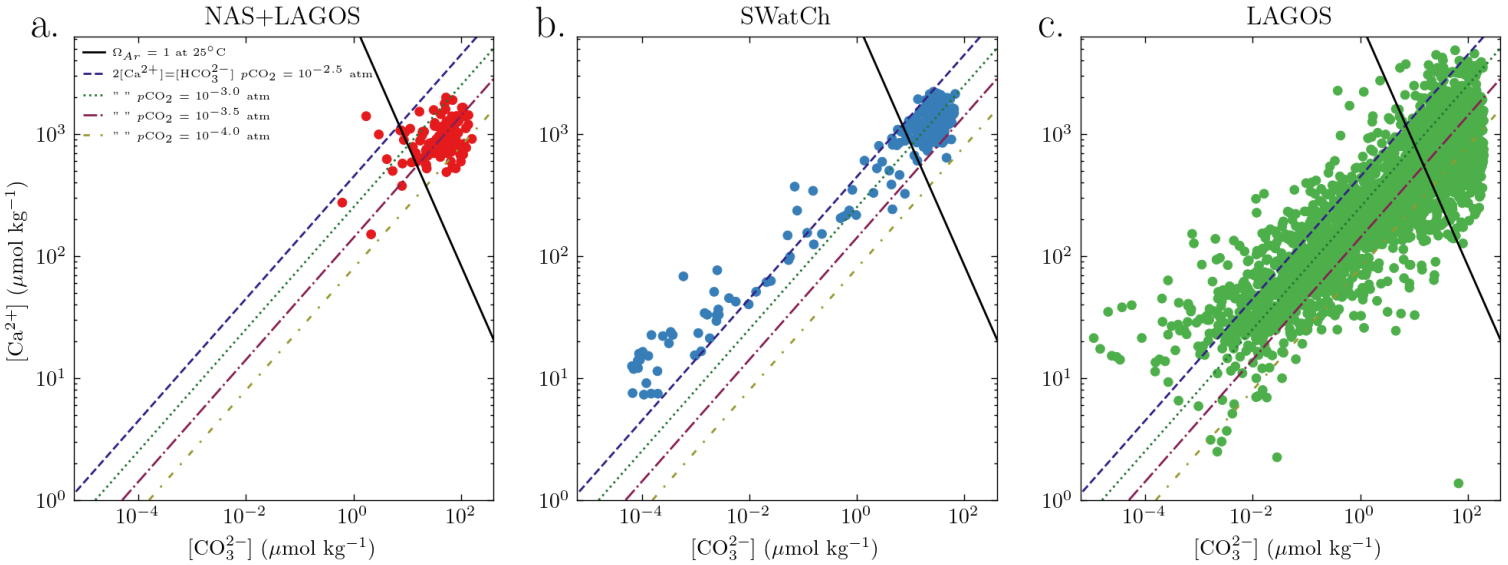

Figure 5.4: Plot of median observations of world lake concentrations of Ca^{2+} and $CO_3{}^{2-}$ calculated from the SWatCh or LAGOS datasets, as well as the subset of the LAGOS dataset overlapping with NAS database observations of Zebra mussels (denoted NAS+LAGOS) in lakes and ponds. Contour lines indicate $\Omega_{Ar} = 1$ (saturation) at 4 °C (dashed) and 25 °C (solid) assuming pH = 8.0. Note logarithmic transformation of the axes, which enables illustration of a linear aragonite saturation and better differentiates the lakes across 7 orders of magnitude ($[CO_3{}^{2-}]$) and 4 orders of magnitude ($[Ca^{2+}]$). An apparently linear relationship is apparent between log-transformed Ca^{2+} and log-transformed $CO_3{}^{2-}$ in the SWatCh and LAGOS data.

5.4 Conclusions

This letter briefly outlined the theoretical and practical basis for the use of a mineral saturation state in lakes to describe interesting and useful aspects of $CaCO_3$ chemistry among and within inland water bodies. It addition to the utility of Ω (and its aragonite- and calcite-specific forms Ω_{Ar} and Ω_{Ca}) as chemical variables synthesizing multiple aspects of inorganic carbon and calcium cycling in water, its use as an ecological variable was proposed and supported by observations indicating a association of Ω_{Ar} with Dreissenid mussel habitat. Further study is needed to demonstrate causal mechanisms as well as to establish values of Ω in natural waters with management applications.

Several further opportunities and barriers pertaining to expanded use of Ω in inland waters bear mentioning.

Accumulation of CO_2 in Earth's atmosphere has been driving acidification of the surface ocean. It is projected that the Great Lakes may also experience higher aqueous CO_2 concentrations driving a pH decrease between 0.29 and 0.49 units by the year 2100 assuming no alkalinization or sediment buffer. This could contribute to a decrease in Ω_{Ar} as observed in the global ocean (Phillips et al., 2015). Three processes at work in inland waters may soften or reverse a Ω_{Ar} trend: first, river alkalinization may locally increase ${CO_3}^{2-}$ concentration faster than acidification decreases it (Gomez et al., 2021); second, increased Ca^{2+} inputs from weathering, land use changes, and road salting contribute to higher Ω_{Ar} values (Kaushal et al., 2018); third, the internal $CaCO_3$ buffer in lakes with carbonate sediments represents another bulwark to acidification (Turk & Adams, 1983); last, warming waters increase Ω_{Ar} by shifting inorganic carbon speciation. The impact of changing $CaCO_3$ cycling on aquatic ecosystems (and calcifying species in particular) is an open question, as the connections between water chemistry, species survival, and ecosystem health are not well-understood. Recent work supporting hypothesized atmospheric CO_2 invasion of Lake Superior (Sandborn & Minor, 2024) underlines a pressing need

for study.

This survey of Ω_{Ar} in lakes and ponds and potential connection to Dreissenid presence relied heavily upon the valuable work of teams coordinating the SWatCh, LAGOS, GLNPO, and NAS databases. These databases, in turn, comprised the work of many observers working with varying methods and definitions of chemical analysis differing by agency, nation, instrument manufacturer, or other factors. While the coordinated datasets in use for this project were well-standardized, with erroneous features removed or flagged, there still remained the question of varying measurement uncertainty and its propagation through the equilibrium calculations that underlaid this work.

Several sources of uncertainty factored into the equilibrium calculations of Ω, which pose a barrier to application of saturation state as an ecological variable. First, the set of inorganic carbon equilibrium coefficients used may introduce some bias. Second, the uncertainty of measurements of carbonate parameters propagates to Ω; this uncertainty can be considerable in measurements obtained with potentiomeric pH instruments or various alkalinity titration methods (Golub et al., 2017; Sandborn et al., 2023; Young et al., 2022). Third, Ω can change rapidly in an aquatic system as biogeochemical processes produce or consume CO_3^{2-}. This last challenge mirrors that associated with pH monitoring, and similar solutions (continuous monitoring and careful time series analysis) apply (Pfister et al., 2014). Organic alkalinity species may also contribute to measured A_T, but are not routinely accounted for in equilibrium calculations. The presence of organic alkalinity can introduce biases in Ω with a direction and magnitude dependent upon the choice of inorganic carbon input parameters and acid-base properties of the dissolved organic matter; bias in Ω is minimized by calculations based on A_T and pH (Sharp & Byrne, 2020). The influence of organic alkalinity on inorganic carbon equilibria is understudied in inland waters, and deserves closer study. Effects of Mg^{2+} and other "hard" ions on mineral solubility were not considered in this work, but marine studies indicate that higher solubility of high-Mg biogenic carbonate minerals may put

some calcifiers at risk in an acidifying environment (Andersson et al., 2008), while incorporation of Mg into biogenic carbonates varies with environmental Mg^{2+}/Ca^{2+} ratios and species (Ries, 2010). Development of mineral saturation state as a freshwater tool should consider how these marine findings may apply to inland waters with varying Mg^{2+}/Ca^{2+} ratios and calcifying species.

This work provided a practical demonstration of the analytical burden needed to observe Ω_{Ar} in inland waters; namely, that the inorganic carbon system be constrained by measurement of at least two of DIC, A_T, pH, or pCO_2, along with Ca^{2+}, water temperature, and preferably conductivity/salinity. This must be repeated in order to observe spatiotemporal variability of Ω_{Ar} in large or seasonally-variable systems. For some monitoring efforts, these measurements may already be at hand. A central challenge in this work was the dearth of high-quality simultaneous measurements constraining equilibrium calculation of $CO_3{}^{2-}$ from, e.g. pH and A_T. For both observational and experimental efforts to use Ω_{Ar} (or Ω_{Ca}) as an ecological variable, high quality inorganic carbon and Ca^{2+} measurements will be central to observing accurate Ω. Improved and more routine chemical characterization of $CaCO_3$ saturation state is a necessary step in any effort to broaden this variable's use in limnology, ecology, and lake management.

5.4.1 Next steps

We demonstrated how $CaCO_3$ saturation state may be calculated, outlined its distribution in a selection of lakes, and illustrated its potential as a chemical predictor of calcifying species habitat. Our last goal is to suggest research needed to enable the use of $CaCO_3$ saturation state as a research and management tool. More comprehensive knowledge of $CaCO_3$ cycling among and within inland waters could be gained by expanded observation of inorganic carbon parameters (pH, A_T, DIC, pCO_2) in lakes and rivers, taking advantage of

recent improvements in analytical techniques achieved in both marine and lacustrine communities. Chemical responses to changing climate, hydrology, and biogeochemical cycles become increasingly important to observe and model in the context their relationship to natural resources and ecosystems. Calcifying species (both invasive and native; see Introduction) have been demonstrated to be sensitive to the concentrations and changes in chemical parameters driving $CaCO_3$ saturation state in both observational and experimental study. This survey points to a similar association between Ω_{Ar} and *Dreissena polymorphia* habitat. Experimental and observational study should be employed to examine any mechanistic or causal relationship of $CaCO_3$ saturation and calcifying species habitat, survival, and reproduction.

Acknowledgments

Thanks to Ted Ozerzky, Audrey Huff, and Lyubov Burlakova for their invaluable suggestions benefiting this work. D.E.S. was supported by a Graduate Research Fellowship from the Cooperative Institute for Great Lakes Research. Funding for this project was provided by the University of Michigan Cooperative Institute for Great Lakes Research (CIGLR), through the National Oceanic and Atmospheric Administration (NOAA) Cooperative Agreement, NA17OAR4320152.

Chapter 6

Conclusion and Discussion

In this dissertation, observations and models of inorganic carbon cycling in Lake Superior provided a clearer picture of the lake's carbon biogeochemistry in the context of a shifting global carbon cycle. The rising CO_2 concentration in Earth's atmosphere and its interactions with the surface ocean drove the hypothesis that Lake Superior's inorganic carbon cycle is interacting and responding to the atmospheric CO_2 trend in a parallel fashion. Individual projects engaged with this hypothesis via a variety of means, but each was rooted in oceanographic practice grounded in lacustrine context. This overarching motif of marine-freshwater interdisciplinarity aiming at "bridging the salty divide" (Cole, 2013b) benefited each of the four projects.

6.1 Summary and synthesis of work

Improved observation of inorganic carbon cycling was promoted by the development of an instrument for the analysis of A_T adhering to established marine best practices, while providing a more accurate, precise, thermodynamically-consistent, and accessible measurement on an open-source platform. Chapter 2 demonstrated analytical techniques which could contribute to constraining

seasonal and interannual trends in alkalinity, indicating hydrological changes, and illustrating acid-buffering behavior in light of to atmospheric CO_2 invasion and potential acidification as reviewed in Chapter 1 and supported in Chapter 3. Several aspects of alkalinity in inland waters were raised but not fully explored, including the presence of organic or excess alkalinity in the Great Lakes and implications for inorganic carbon equilibria, the relationship between alkalinity and conductivity or ionic strength in these systems, and development of reference materials (or use of existing certified reference materials) to aid freshwater A_T and DIC analysis, all of which merit further study.

Using improved inorganic carbon parameter analytical methods (including the instrument demonstrated in Chapter 2), the next project presented the first multi-year pCO_2 time series produced on Lake Superior by underway instrumentation. An analysis of seasonal and interannual variability in Lake Superior surface pCO_2 revealed prolonged periods of CO_2 influx and efflux which varied in space in association with riverine influence. These observations served as the first direct observation of multiannual-scale pCO_2 increase, supporting the hypothesis of atmospheric CO_2 invasion of Lake Superior. Significant gaps in space and time between transects, notably in winter, inhibited improvement of Superior's as-yet unbalanced carbon budget. In light of this dearth of observation, which has similarly stymied previous attempts to understand Lake Superior's carbon cycle, an observation-based model was developed in Chapter 4.

An model of surface water pCO_2 spanning Lake Superior over the period 2019-2023 was assembled as an observation-based gap-filling mechanism, with the goal of describing sub-seasonal to multi-annual variability in unobserved domains. A feed-forward neural network model resembling similar marine efforts (e.g. Landschützer et al., 2013) was trained on underway pCO_2 observations to predict daily mean pCO_2 over the entire lake surface from 2019-2023 with a spatial resolution of $0.02° \times 0.02°$. Comparison of model output with bottle and moored sensor observation validated model predictions both in and

out of the training data spatiotemporal domain. This allowed inference of annual net whole-lake CO_2 air-sea fluxes which were consistent with a balanced carbon budget. This model further supported the picture of a large lake at multi-annual equilibrium with an increasing atmospheric CO_2 concentration. Ecological ramifications of this phenomenon motivated the discussion of the relationship between shifting inorganic carbon chemistry, $CaCO_3$ solubility, and calcifying bivalves in Chapter 5.

The final project in this work presented a discussion of aragonite saturation state as a chemical predictor of calcifying species habitat in lakes, with a focus on invasive Zebra mussels (*Dreissena polymorphia*). This transfer of a mineral saturation state from marine science to a freshwater setting was motivated by its utility to synthesize several variables which have been individually demonstrated to predict the habitat of calcifying bivalves. In light of shifting inorganic carbon chemistry described in previous chapters, saturation state as a chemical-ecological parameter gained further importance. Literature values describing calcium carbonate saturation state in the Great Lakes and several thousand world lakes illustrated variability of aragonite saturation state within lakes driven by carbonate ion concentration, while variability among lakes demonstrated an essentially linear distribution of saturated and undersaturated conditions dictated by the charge balance of calcium and bicarbonate ions. Comparison of lakes' aragonite saturation states against Zebra mussel observations illustrated association of habitat bounds with aragonite saturation state similar to other predictor variables. More work remains to establish any utility of calcium carbonate saturation states as useful chemical-ecological variables, such as further observation of saturation state and mussel habitat, as well as experimental frameworks studying mussel survival and reproduction under conditions of varying saturation state. Comparison of the work of marine biologists studying ocean acidification (e.g. Riebesell et al., 2011) may be of use for future work.

6.2 Next steps

The foregoing chapters prompt further scientific questions regarding carbon cycling in Lake Superior, the Laurentian Great Lakes, and other large lake systems in general. In addition to further research suggested in the foregoing section, several other questions arise which should be addressed by the Earth science community.

It should be determined to what extent the finding of atmospheric CO_2 invasion in Lake Superior may apply to other systems. The invasion of Lake Superior described in this work occurs because of repeated equilibration with a rising atmospheric signal, mediated by a relatively-short characteristic equilibration timescale. Apparent dominance of Lake Superior surface water pCO_2 by the atmospheric CO_2 reservoir on multiannual scales illustrated a fundamental behavior of the lake's carbon cycle. An observation-based model developed in Chapter 4 took advantage of a unique source of underway pCO_2 data as well as moored pCO_2 and bottle samples to create a gap-filling model of pCO_2 and CO_2 flux spanning Lake Superior throughout 5 years, demonstrating an application of expanded observation which supported a better-constrained carbon budget for Earth's largest lake. Other lake systems which maintain super- or under-saturated CO_2 concentrations relative to the atmosphere do not necessarily undergo atmospheric CO_2 invasion. The CO_2 concentration of a supersaturated lake (as most are, most of the time) may vary interannually due to allocthonous carbon loading and internal metabolism, or it may achieve some predictable relationship with atmospheric CO_2 which results in an interannually-rising CO_2 concentration for all or part of the year. Observation and time series analysis resolving an interannual trend (if any) may discriminate lakes undergoing CO_2 invasion from those not doing so.

This dissertation was motivated in part by the specter of lake acidification, yet only the action of its driver, CO_2 invasion, was substantively supported by the research presented above. Trends of lake pH have proven more difficult to

discern than pCO_2 for several reasons: first, surface water pH is mediated by the atmospheric equilibrium of the increasing surface water pCO_2 signal and interannually variable alkalinity, which likely responds to hydrologic drivers including lake level. Second, available time series (i.e. Great Lakes National Program Office surveys) do not resolve seasonal variability with twice-annual observations falling on variable dates superimposed upon shifts in pCO_2 cycle phenology potentially linked to climate and lake level variability. This inhibits deseasonalizing time series data as an essential step towards interannual trend detection. Third, the measurement uncertainty associated with many historic inorganic carbon parameter measurements reduces the statistical power of trend-finding efforts; continuing adoption of more precise and accurate tools including spectrophotometric pH, deployed pCO_2 sensors, and improved total alkalinity and dissolved inorganic carbon methods will strengthen present and future observational efforts. Fourth, the dearth of wintertime observations prevents understanding of year-round inorganic carbon cycling dynamics. Under-ice sensor deployments and winter limnology campaigns have begun to alleviate this gap in knowledge while climate change is shifting the experience of winter.

The sum of these obstacles leaves large lake limnologists with the daunting task of measuring shifts in CO_2 cycling and acidification among lakes of broadly variable chemical contexts. This task is matched in importance by the need to understanding biogeochemical and ecological outcomes of a yet uncertain acidification trend. The way forward, it is hoped, is cleared somewhat by the examples detailed in this work: more and higher-quality measurements of inorganic carbon parameters, improved spatial and temporal resolution and coverage, integration with observation-based and biogeochemical process models, and dialog with ecological effects and feedbacks.

References

Adkins, J. F., Naviaux, J. D., Subhas, A. V., Dong, S., & Berelson, W. M. (2021). The Dissolution Rate of $CaCO_3$ in the Ocean. *Annu. Rev. Mar. Sci.*, *13*(1), 57–80. https://doi.org/10.1146/annurev-marine-041720-092514

Ahmed, M., Else, B. G. T., Burgers, T. M., & Papakyriakou, T. (2019). Variability of surface water pCO_2 in the Canadian Arctic Archipelago from 2010 to 2016. *J. Geophys. Res. Oceans*, *124*(3), 1876–1896. https://doi.org/10.1029/2018JC014639

Alin, S. R., & Johnson, T. C. (2007). Carbon cycling in large lakes of the world: A synthesis of production, burial, and lake-atmosphere exchange estimates. *Global Biogeochem. Cycles*, *21*. https://doi.org/10.1029/2006GB002881

Andersson, A., Mackenzie, F., & Bates, N. (2008). Life on the margin: Implications of ocean acidification on Mg-calcite, high latitude and cold-water marine calcifiers. *Mar. Ecol. Prog. Ser.*, *373*, 265–273. https://doi.org/10.3354/meps07639

Archer, D., Kheshgi, H., & Maier-Reimer, E. (1998). Dynamics of fossil fuel CO_2 neutralization by marine $CaCO_3$. *Global Biogeochem. Cycles*, *12*(2), 259–276. https://doi.org/10.1029/98GB00744

Arruda, R., Atamanchuk, D., Cronin, M., Steinhoff, T., & Wallace, D. W. R. (2020). At-sea intercomparison of three underway pCO_2 systems. *Limnology & Ocean Methods*, *18*(2), 63–76. https://doi.org/10.1002/lom3.10346

Atilla, N., McKinley, G. A., Bennington, V., Baehr, M., Urban, N. R., DeGrandpre, M. D., Desai, A. R., & Wu, C. (2011). Observed variability of Lake Superior pCO_2. *Limnol. Oceanogr.*, *56*(3), 775–786. https://doi.org/10.4319/lo.2011.56.3.0775

Austin, J. A. (2013). Observations of near-inertial energy in Lake Superior. *Limnol. Oceanogr.*, *58*(2), 715–728. https://doi.org/10.4319/lo.2013.58.2.0715

Austin, J. A., & Allen, J. (2011). Sensitivity of summer Lake Superior thermal structure to meteorological forcing. *Limnology & Oceanography*, *56*(3), 1141–1154. https://doi.org/10.4319/lo.2011.56.3.1141

Austin, J. A., & Colman, S. (2008). A century of temperature variability in Lake Superior. *Limnol. Oceanogr.*, *53*(6), 2724–2730. https://doi.org/10.4319/lo.2008.53.6.2724

Austin, J. A., Hill, C., Fredrickson, J., Weber, G., & Weiss, K. (2022). Characterizing temporal and spatial scales of radiatively driven convection in a deep, ice-free lake. *Limnol. Oceanogr.*, lno.12203. https://doi.org/10.1002/lno.12203

Bagshaw, E. A., Wadham, J. L., Tranter, M., Beaton, A. D., Hawkings, J. R., Lamarche-Gagnon, G., & Mowlem, M. C. (2021). Measuring pH in low ionic strength glacial meltwaters using ion selective field effect transistor (ISFET) technology. *Limnol Oceanogr Methods*, lom3.10416. https://doi.org/10.1002/lom3.10416

Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S.-i., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., … Xu, S. (2016). A multi-decade record of high-quality pCO_2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT). *Earth Syst. Sci. Data*, *8*(2), 383–413. https://doi.org/10.5194/essd-8-383-2016

Barbiero, R. P., Tuchman, M. L., & Millard, E. S. (2006). Post-dreissenid increases in transparency during summer stratification in the offshore waters of Lake Ontario: Is a reduction in whiting events the cause? *J. Great Lakes Res.*, *32*(1), 131–141. https://doi.org/10.3394/0380-1330(2006)32[131:PIITDS]2.0.CO;2

Bates, N. R. (2001). Interannual variability of oceanic CO_2 and biogeochemical properties in the Western North Atlantic subtropical gyre. *Deep Sea Res. Part II Top. Stud. Oceanogr.*, *48*(8-9), 1507–1528. https://doi.org/10.1016/S0967-0645(00)00151-X

Bates, R., & Wichers, E. (1957). Precise intercomparison of acids by differential potentiometric titration with hydrogen electrodes. *J. Res. Natl. Bur. Stan.*, *59*(1), 9. https://doi.org/10.6028/jres.059.002

Becerra-Valdivia, L., & Higham, T. (2020). The timing and effect of the earliest human arrivals in North America. *Nature*, *584*(7819), 93–97. https://doi.org/10.1038/s41586-020-2491-6

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci Data*, *5*(1), 180214. https://doi.org/10.1038/sdata.2018.214

Bennett, E. B. (1978). Characteristics of the Thermal Regime of Lake Superior. *J. Great Lakes Res.*, *4*(3-4), 310–319. https://doi.org/10.1016/S0380-1330(78)72200-8

Bennington, V., Galjanic, T., & McKinley, G. A. (2022). Explicit Physical Knowledge in Machine Learning for Ocean Carbon Flux Reconstruction: The pCO_2 -Residual Method. *J Adv Model Earth Syst*, *14*(10). https://doi.org/10.1029/2021MS002960

Bennington, V., McKinley, G. A., Kimura, N., & Wu, C. H. (2010). General circulation of Lake Superior: Mean, variability, and trends from 1979 to 2006. *J. Geophys. Res.*, *115*(C12), C12015. https://doi.org/10.1029/2010JC006261

Bennington, V., McKinley, G. A., Urban, N. R., & McDonald, C. P. (2012). Can spatial heterogeneity explain the perceived imbalance in Lake Superior's carbon budget? *J. Geophys. Res.*, *117*(G03020). https://doi.org/10.1029/2011JG001895

Benson, A. J., Raikow, D., Larson, J., Fusaro, A., Bogdanoff, A. K., & Elgin, A. (2023). *Dreissena polymorpha (Pallas, 1771)*. U.S. Geological Survey. Gainesville, FL. https://nas.er.usgs.gov/queries/FactSheet.aspx?speciesID=5

Bianchi, T. S. (2021). The evolution of biogeochemistry: Revisited. *Biogeochemistry*, *154*(2), 141–181. https://doi.org/10.1007/s10533-020-00708-0

Block, B. D., Denfeld, B. A., Stockwell, J. D., Flaim, G., Grossart, H.-P. F., Knoll, L. B., Maier, D. B., North, R. L., Rautio, M., Rusak, J. A., Sadro, S., Weyhenmeyer, G. A., Bramburger, A. J., Branstrator, D. K., Salonen, K., & Hampton, S. E. (2018). The unique methodological challenges of winter limnology. *Limnol Oceanogr Methods*, lom3.10295. https://doi.org/10.1002/lom3.10295

Bockmon, E. E., & Dickson, A. G. (2014). A seawater filtration method suitable for total dissolved inorganic carbon and pH analyses. *Limnol. Oceanogr. Methods*, *12*(4), 191–195. https://doi.org/10.4319/lom.2014.12.191

Bockmon, E. E., & Dickson, A. G. (2015). An inter-laboratory comparison assessing the quality of seawater carbon dioxide measurements. *Mar. Chem.*, *171*, 36–43. https://doi.org/10.1016/j.marchem.2015.02.002

Breitberg, D., Salisbury, J., Bernhard, J., Cai, W.-J., Dupont, S., Doney, S., Kroeker, K., Levin, L., Long, W. C., Milke, L., Miller, S., Phelan, B., Passow, U., Seibel, B., Todgham, A., & Tarrant, A. (2015). And on Top of All That… Coping with Ocean Acidification in the Midst of Many Stressors. *Oceanog.*, *25*(2), 48–61. https://doi.org/10.5670/oceanog.2015.31

Broecker, W. S. (1985). *How to build a habitable planet*. Eldigio Press.

Broecker, W. S., & Peng, T.-H. (1982). *Tracers in the Sea*. Lamont-Doherty Earth Observatory, Columbia University.

Brothers, S., & Sibley, P. (2018). Light may have triggered a period of net heterotrophy in Lake Superior. *Limnol. Oceanogr.*, *63*(4), 1785–1798. https://doi.org/10.1002/lno.10808

Bushinsky, S. M., Takeshita, Y., & Williams, N. L. (2019). Observing changes in ocean carbonate chemistry: Our autonomous future. *Curr. Clim. Change Rep.*, *5*(3), 207–220. https://doi.org/10.1007/s40641-019-00129-8

Butman, D., Stackpoole, S., Stets, E., McDonald, C. P., Clow, D. W., & Striegl, R. G. (2016). Aquatic carbon cycling in the conterminous United States and implications for terrestrial carbon accounting. *Proc. Natl. Acad. Sci. U.S.A.*, *113*(1), 58–63. https://doi.org/10.1073/pnas.1512651112

Byrne, R. H., & Yao, W. (2008). Procedures for measurement of carbonate ion concentrations in seawater by direct spectrophotometric observations of Pb(II) complexation. *Mar. Chem.*, *112*(1), 128–135. https://doi.org/10.1016/j.marchem.2008.07.009

Cai, W.-J., & Wang, Y. (1998). The chemistry, fluxes, and sources of carbon dioxide in the estuarine waters of the Satilla and Altamaha Rivers, Georgia. *Limnol. Oceanogr.*, *43*(4), 657–668. https://doi.org/10.4319/lo.1998.43.4.0657

Cai, W.-J., Wang, Y., & Hodson, R. E. (1998). Acid-Base Properties of Dissolved Organic Matter in the Estuarine Waters of Georgia, USA. *Geochimica et Cosmochimica Acta*, *62*(3), 473–483. https://doi.org/10.1016/S0016-7037(97)00363-3

Cartapanis, O., Galbraith, E. D., Bianchi, D., & Jaccard, S. L. (2018). Carbon burial in deep-sea sediment and implications for oceanic inventories of carbon and alkalinity over the last glacial cycle. *Clim. Past*, *14*(11), 1819–1850. https://doi.org/10.5194/cp-14-1819-2018

Carter, B. R., Radich, J. A., Doyle, H. L., & Dickson, A. G. (2013). An automated system for spectrophotometric seawater pH measurements: Automated spectrophotometric pH measurement. *Limnol. Oceanogr. Methods*, *11*(1), 16–27. https://doi.org/10.4319/lom.2013.11.16

Carter, B. R., Williams, N. L., Evans, W., Fassbender, A. J., Barbero, L., Hauri, C., Feely, R. A., & Sutton, A. J. (2019). Time of detection as a metric for prioritizing between climate observation quality, frequency, and duration. *Geophys. Res. Lett.*, *46*(7), 3853–3861. https://doi.org/10.1029/2018GL080773

Cavallaro, N., Shrestha, G., Birdsey, R., Mayes, M. A., Najjar, R. G., Reed, S. C., Romero-Lankao, P., & Zhu, Z. (2018). *Second state of the carbon cycle report*. U.S. Global Change Research Program. https://doi.org/10.7930/Soccr2.2018

Challener, R. C., Robbins, L. L., & McClintock, J. B. (2016). Variability of the carbonate chemistry in a shallow, seagrass-dominated ecosystem: Implications for ocean acidification experiments. *Mar. Freshwater Res.*, *67*(2), 163. https://doi.org/10.1071/MF14219

Chapra, S. C., Dove, A., & Warren, G. J. (2012). Long-term trends of Great Lakes major ion chemistry. *J. Great Lakes Res.*, *38*(3), 550–560. https://doi.org/10.1016/j.jglr.2012.06.010

Chollet, F. (2015). *Keras*. https://keras.io

Cohen, A. N., & Weinstein, A. (2001). *Zebra Mussel's Calcium Threshold and Implications for its Potential Distribution in North America*. San Francisco Estuary Institute. Richmond, California.

Cole, J. J. (2013a). *Freshwater ecosystems and the carbon cycle*. International Ecology Institute.

Cole, J. J. (2013b). Bridging the Salty Divide? *Limnol. Oceanog. Bull.*, *22*(3), 84–84. https://doi.org/10.1002/lob.201322384

Cole, J. J., & Caraco, N. F. (1998). Atmospheric exchange of carbon dioxide in a low-wind oligotrophic lake measured by the addition of SF_6. *Limnol. Oceanogr.*, *43*(4), 647–656. https://doi.org/10.4319/lo.1998.43.4.0647

Cole, J. J., Caraco, N. F., Kling, G. W., & Kratz, T. K. (1994). Carbon Dioxide Supersaturation in the Surface Waters of Lakes. *Science*, *265*(5178), 1568–1570. https://doi.org/10.1126/science.265.5178.1568

Cotner, J. B., Biddanda, B. A., Makino, W., & Stets, E. (2004). Organic carbon biogeochemistry of Lake Superior. *Aquatic Ecosystem Health & Management*, *7*(4), 451–464. https://doi.org/10.1080/14634980490513292

Cuhel, R. L., & Aguilar, C. (2013). Ecosystem Transformations of the Laurentian Great Lake Michigan by Nonindigenous Biological Invaders. *Annu. Rev. Mar. Sci.*, *5*(1), 289–320. https://doi.org/10.1146/annurev-marine-120710-100952

Davis, W. S., & Simon, T. P. (Eds.). (1995). *Biological assessment and criteria: Tools for water resource planning and decision making*. Lewis Publishers.

Daw, A., Karpatne, A., Watkins, W., Read, J., & Kumar, V. (2021, September 28). *Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling*. arXiv: 1710.11431 [physics, stat]. Retrieved June 13, 2024, from http://arxiv.org/abs/1710.11431

DeGrandpre, M. D., Evans, W., Timmermans, M.-L., Krishfield, R., Williams, B., & Steele, M. (2020). Changes in the Arctic Ocean Carbon Cycle With Diminishing Ice Cover. *Geophys. Res. Lett.*, *47*(12). https://doi.org/10.1029/2020GL088051

DeGrandpre, M. D., Lai, C.-Z., Timmermans, M.-L., Krishfield, R. A., Proshutinsky, A., & Torres, D. (2019). Inorganic carbon and pCO_2 variability during ice formation in the Beaufort Gyre of the Canada Basin. *J. Geophys. Res. Oceans*, *124*(6), 4017–4028. https://doi.org/10.1029/2019JC015109

Delvaux, J. M. (2017). *River Influence on the Nearshore Ecosystem of Western Lake Superior* [M.S.]. The University of Wisconsin - Milwaukee. www.proquest.com/dissertations-theses/river-influence-on-nearshore-ecosystem-western/docview/1987937549/se-2
10621639.

Desai, A. R. (2022). *AmeriFlux BASE US-PFa Park Falls/WLEF*. AmeriFlux AMP. https://doi.org/10.17190/AMF/1246090

Desai, A. R., Austin, J. A., Bennington, V., & McKinley, G. A. (2009). Stronger winds over a large lake in response to weakening air-to-lake temperature gradient. *Nature Geosci*, *2*(12), 855–858. https://doi.org/10.1038/ngeo693

Dickson, A. G. (1981). An exact definition of total alkalinity and a procedure for the estimation of alkalinity and total inorganic carbon from titration data. *Deep Sea Research Part A. Oceanographic Research Papers*, *28*(6), 609–623. https://doi.org/10.1016/0198-0149(81)90121-7

Dickson, A. G. (1990). Standard potential of the reaction: $AgCl(s) + 0.5H_2(g) = Ag(s) + HCl(aq)$, and the standard acidity constant of the ion HSO_2^- in synthetic sea water from 273.15 to 318.15 K. *The Journal of Chemical Thermodynamics*, *22*(2), 113–127. https://doi.org/10.1016/0021-9614(90)90074-Z

Dickson, A. G. (1992). The development of the alkalinity concept in marine chemistry. *Mar. Chem.*, *40*(1-2), 49–63. https://doi.org/10.1016/0304-4203(92)90047-E

Dickson, A. G., & Riley, J. (1978). The effect of analytical error on the evaluation of the components of the aquatic carbon-dioxide system. *Mar. Chem.*, *6*(1), 77–85. https://doi.org/10.1016/0304-4203(78)90008-7

Dickson, A. G., Sabine, C. L., Christian, J. R., Bargeron, C. P., & Organization, N. P. M. S. (Eds.). (2007). *Guide to best practices for ocean CO_2 measurements*. North Pacific Marine Science Organization.

Doney, S. C., Busch, D. S., Cooley, S. R., & Kroeker, K. J. (2020). The Impacts of Ocean Acidification on Marine Ecosystems and Reliant Human Communities. *Annu. Rev. Environ. Resour.*, *45*(1), 83–112. https://doi.org/10.1146/annurev-environ-012320-083019

Doney, S. C., Fabry, V. J., Feely, R. A., & Kleypas, J. A. (2009). Ocean acidification: The other CO_2 problem. *Annu. Rev. Mar. Sci.*, *1*(1), 169–192. https://doi.org/10.1146/annurev.marine.010908.163834

Duarte, C. M., Hendriks, I. E., Moore, T. S., Olsen, Y. S., Steckbauer, A., Ramajo, L., Carstensen, J., Trotter, J. A., & McCulloch, M. (2013). Is Ocean Acidification an Open-Ocean Syndrome? Understanding Anthropogenic Impacts on Seawater pH. *Estuaries and Coasts*, *36*(2), 221–236. https://doi.org/10.1007/s12237-013-9594-3

Dyrssen, D., & Sillén, L. G. (1967). Alkalinity and total carbonate in sea water. A plea for *p*-*T*-independent data. *Tellus*, *19*(1), 113–121. https://doi.org/10.1111/j.2153-3490.1967.tb01464.x

Easley, R. A., & Byrne, R. H. (2012). Spectrophotometric Calibration of pH Electrodes in Seawater Using Purified m-Cresol Purple. *Environ. Sci. Technol.*, *46*(9), 5018–5024. https://doi.org/10.1021/es300491s

Elson, P., De Andrade, E. S., Lucas, G., May, R., Hattersley, R., Campbell, E., Dawson, A., Little, B., Stephane Raynaud, Scmc72, Snow, A. D., Comer, R., Donkers, K., Blay, B., Killick, P., Wilson, N., Peglar, P., Lgolston, Lbdreyer, … Havlin, C. (2023, August 5). *SciTools/cartopy: V0.22.0* (Version v0.22.0). https://doi.org/10.5281/ZENODO.1182735

Elster, H.-J. (1974). History of limnology. *SIL Communications, 1953-1996*, *20*(1), 7–30. https://doi.org/10.1080/05384680.1974.11923880

Escobar, L. E., Mallez, S., McCartney, M., Lee, C., Zielinski, D. P., Ghosal, R., Bajer, P. G., Wagner, C., Nash, B., Tomamichel, M., Venturelli, P., Mathai, P. P., Kokotovich, A., Escobar-Dodero, J., & Phelps, N. B. D. (2018). Aquatic Invasive Species in the Great Lakes Region: An Overview. *Reviews in Fisheries Science & Aquaculture*, *26*(1), 121–138. https://doi.org/10.1080/23308249.2017.1363715

Eurachem Working Group on Uncertainty in Chemical Measurement. (2012). *Quantifying uncertainty in analytical measurement* (S. Ellison & A. Williams, Eds.; 3rd). Eurachem. http://www.eurchem.org/

Evans, W., Lebon, G. T., Harrington, C. D., Takeshita, Y., & Bidlack, A. (2022). Marine CO_2 system variability along the northeast Pacific Inside Passage

determined from an Alaskan ferry. *Biogeosciences*, *19*(4), 1277–1301. https://doi.org/10.5194/bg-19-1277-2022

Evans, W., Pocock, K., Hare, A., Weekes, C., Hales, B., Jackson, J., Gurney-Smith, H., Mathis, J. T., Alin, S. R., & Feely, R. A. (2019). Marine CO_2 Patterns in the Northern Salish Sea. *Front. Mar. Sci.*, *5*, 536. https://doi.org/10.3389/fmars.2018.00536

Fahnenstiel, G. L., Sicko-Goad, L., Scavia, D., & Stoermer, E. F. (1986). Importance of Picoplankton in Lake Superior. *Can. J. Fish. Aquat. Sci.*, *43*(1), 235–240. https://doi.org/10.1139/f86-028

Fassbender, A. J., Rodgers, K. B., Palevsky, H. I., & Sabine, C. L. (2018). Seasonal asymmetry in the evolution of surface ocean pCO_2 and pH thermodynamic drivers and the influence on Sea-Air CO_2 flux. *Global Biogeochem. Cycles*, *32*(10), 1476–1497. https://doi.org/10.1029/2017GB005855

Feely, R., Sabine, C., Takahashi, T., & Wanninkhof, R. (2001). Uptake and storage of carbon dioxide in the ocean: The global CO_2 survey. *Oceanog*, *14*(4), 18–32. https://doi.org/10.5670/oceanog.2001.03

Feely, R. A., Sabine, C. L., Byrne, R. H., Millero, F. J., Dickson, A. G., Wanninkhof, R., Murata, A., Miller, L. A., & Greeley, D. (2012). Decadal changes in the aragonite and calcite saturation state of the Pacific Ocean. *Global Biogeochem. Cycles*, *26*(3), 2011GB004157. https://doi.org/10.1029/2011GB004157

Feely, R. A., Sabine, C. L., Lee, K., Berelson, W., Kleypas, J., Fabry, V. J., & Millero, F. J. (2004). Impact of Anthropogenic CO_2 on the $CaCO_3$ System in the Oceans. *Science*, *305*(5682), 362–366. https://doi.org/10.1126/science.1097329

Fergen, J. T., Bergstrom, R. D., Twiss, M. R., Johnson, L., Steinman, A. D., & Gagnon, V. (2022). Updated census in the Laurentian Great Lakes Watershed: A framework for determining the relationship between the population and this aquatic resource. *J. Great Lakes Res.*, *48*(6), 1337–1344. https://doi.org/10.1016/j.jglr.2022.03.004

Firing, E., Filipe, Barna, A., & Abernathey, R. (2021, March 23). *TEOS-10/GSW-Python: V3.4.1* (Version v3.4.1). https://doi.org/10.5281/zenodo.4631364

French, C. R., Carr, J. J., Dougherty, E. M., Eidson, L. A., Reynolds, J. C., & DeGrandpre, M. D. (2002). Spectrophotometric pH measurements of freshwater. *Analytica Chimica Acta*, *453*(1), 13–20. https://doi.org/10.1016/S0003-2670(01)01509-4

Frieder, C. A., Nam, S. H., Martz, T. R., & Levin, L. A. (2012). High temporal and spatial variability of dissolved oxygen and pH in a nearshore California kelp forest. *Biogeosciences*, *9*(10), 3917–3930. https://doi.org/10.5194/bg-9-3917-2012

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., … Zheng, B. (2023). Global Carbon Budget 2023. *Earth Syst. Sci. Data*, *15*(12), 5301–5369. https://doi.org/10.5194/essd-15-5301-2023

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., … Zheng, B. (2022). Global Carbon Budget 2022. *Earth Syst. Sci. Data*, *14*(11), 4811–4900. https://doi.org/10.5194/essd-14-4811-2022

Fuller, K., & Shear, H. (Eds.). (1995). *The Great Lakes: An environmental atlas and resource book* (3rd). Great Lakes National Program Office, U.S. Environmental Protection Agency ; Government of Canada.

Gazeau, F., Gattuso, J.-P., Dawber, C., Pronker, A. E., Peene, F., Peene, J., Heip, C. H. R., & Middelburg, J. J. (2010). Effect of ocean acidification on the early life stages of the blue mussel *Mytilus edulis*. *Biogeosciences*, *7*(7), 2051–2060. https://doi.org/10.5194/bg-7-2051-2010

Gazeau, F., Gattuso, J.-P., Greaves, M., Elderfield, H., Peene, J., Heip, C. H. R., & Middelburg, J. J. (2011). Effect of Carbonate Chemistry Alteration on the Early Embryonic Development of the Pacific Oyster (Crassostrea gigas). *PLoS ONE*, *6*(8), e23010. https://doi.org/10.1371/journal.pone.0023010

Gissi, E., Manea, E., Mazaris, A. D., Fraschetti, S., Almpanidou, V., Bevilacqua, S., Coll, M., Guarnieri, G., Lloret-Lloret, E., Pascual, M., Petza, D., Rilov, G., Schonwald, M., Stelzenmüller, V., & Katsanevakis, S. (2021). A review of the combined effects of climate change and other local human stressors on the marine environment. *Sci. Total Environ.*, *755*, 142564. https://doi.org/10.1016/j.scitotenv.2020.142564

Gloege, L., Yan, M., Zheng, T., & McKinley, G. A. (2022). Improved quantification of ocean carbon uptake by using machine learning to merge global models and pCO_2 data. *J. Adv. Model Earth Syst.*, *14*(2). https://doi.org/10.1029/2021MS002620

Golub, M., Desai, A. R., McKinley, G. A., Remucal, C. K., & Stanley, E. H. (2017). Large Uncertainty in Estimating pCO_2 From Carbonate Equilibria in Lakes. *J. Geophys. Res. Biogeosci.*, *122*(11), 2909–2924. https://doi.org/10.1002/2017JG003794

Golub, M., Koupaei-Abyazani, N., Vesala, T., Mammarella, I., Ojala, A., Bohrer, G., Weyhenmeyer, G. A., Blanken, P. D., Eugster, W., Koebsch, F., Chen, J., Czajkowski, K., Deshmukh, C., Guérin, F., Heiskanen, J., Humphreys, E., Jonsson, A., Karlsson, J., Kling, G., … Desai, A. R. (2023). Diel, seasonal, and inter-annual variation in carbon dioxide effluxes from lakes and reservoirs. *Environ. Res. Lett.*, *18*(3), 034046. https://doi.org/10.1088/1748-9326/acb834

Gomez, F. A., Wanninkhof, R., Barbero, L., & Lee, S.-K. (2021). Increasing River Alkalinity Slows Ocean Acidification in the Northern Gulf of Mexico. *Geophysical Research Letters*, *48*(24), e2021GL096521. https://doi.org/10.1029/2021GL096521

Gran, G., Dahlenborg, H., Laurell, S., & Rottenberg, M. (1950). Determination of the Equivalent Point in Potentiometric Titrations. *Acta Chem. Scand.*, *4*, 559–577. https://doi.org/10.3891/acta.chem.scand.04-0559

Great Lakes National Program Office. (2023). *EPA GLNPO Water Quality Time-series*. United States Environmental Protection Agency. https://www.epa.gov/greatlakes

Gregor, L., & Humphreys, M. P. (2021, April 1). *SeaFlux: Updated continuous integration and docs* (Version 2.0.0). https://doi.org/10.5281/ZENODO.4659162

Grigorovich, I. A., Korniushin, A. V., Gray, D. K., Duggan, I. C., & MacIsaac, H. J. (2003). Lake Superior: An invasion coldspot?, 20.

Gruber, N. (2011). Warming up, turning sour, losing breath: Ocean biogeochemistry under global change. *Phil. Trans. R. Soc. A.*, *369*(1943), 1980–1996. https://doi.org/10.1098/rsta.2011.0003

Gruber, N., Sarmiento, J. L., & Stocker, T. F. (1996). An improved method for detecting anthropogenic CO_2 in the oceans. *Global Biogeochemical Cycles*, *10*(4), 809–837. https://doi.org/10.1029/96GB01608

Harmon, T. C. (2020). Carbon gas flux to and from inland waters: Support for a global observation network. *Limnology*, *21*(3), 429–442. https://doi.org/10.1007/s10201-020-00623-1

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Harris, G. L. (2019, May). *Selected laboratory and measurement practices and procedures to support basic mass calibrations* (NIST IR 6969-2019). National Institute of Standards and Technology. Gaithersburg, MD. https://doi.org/10.6028/NIST.IR.6969-2019

Hartig, J. H., Krantzberg, G., & Alsip, P. (2020). Thirty-five years of restoring Great Lakes Areas of Concern: Gradual progress, hopeful future. *Journal of Great Lakes Research*, *46*(3), 429–442. https://doi.org/10.1016/j.jglr.2020.04.004

Hasler, C. T., Jeffrey, J. D., Schneider, E. V. C., Hannan, K. D., Tix, J. A., & Suski, C. D. (2018). Biological consequences of weak acidification caused by elevated carbon dioxide in freshwater ecosystems. *Hydrobiologia*, *806*(1), 1–12. https://doi.org/10.1007/s10750-017-3332-y

Hecky, R. E., Smith, R. E., Barton, D. R., Guildford, S. J., Taylor, W. D., Charlton, M. N., & Howell, T. (2004). The nearshore phosphorus shunt: A consequence of ecosystem engineering by dreissenids in the Laurentian Great Lakes. *Can. J. Fish. Aquat. Sci.*, *61*(7), 1285–1293. https://doi.org/10.1139/f04-065

Heiskanen, J. J., Mammarella, I., Haapanala, S., Pumpanen, J., Vesala, T., Macintyre, S., & Ojala, A. (2014). Effects of cooling and internal wave motions on gas transfer coefficients in a boreal lake. *Tellus B: Chemical and Physical Meteorology*, *66*(1), 22827. https://doi.org/10.3402/tellusb.v66.22827

Hill, K., Dauphinee, T., & Woods, D. (1986). The extension of the Practical Salinity Scale 1978 to low salinities. *IEEE J. Ocean. Eng.*, *11*(1), 109–112. https://doi.org/10.1109/JOE.1986.1145154

Hincks, S. S., & Mackie, G. L. (1997). Effects of pH, calcium, alkalinity, hardness, and chlorophyll on the survival, growth, and reproductive success of zebra mussel (*Dreissena polymorpha*) in Ontario lakes. *Can. J. Fish. Aquat. Sci.*, *54*(9), 2049–2057. https://doi.org/10.1139/f97-114

Ho, D. T., Law, C. S., Smith, M. J., Schlosser, P., Harvey, M., & Hill, P. (2006). Measurements of air-sea gas exchange at high wind speeds in the Southern Ocean: Implications for global parameterizations. *Geophys. Res. Lett.*, *33*(16), L16611. https://doi.org/10.1029/2006GL026817

Ho, D. T., Wanninkhof, R., Schlosser, P., Ullman, D. S., Hebert, D., & Sullivan, K. F. (2011). Toward a universal relationship between wind speed and gas exchange: Gas transfer velocities measured with $^3He/SF_6$ during the Southern Ocean Gas Exchange Experiment. *J. Geophys. Res.*, *116*, C00F04. https://doi.org/10.1029/2010JC006854

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Guillén Bolaños, T., Bindi, M., Brown, S., Camilloni, I. A., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Hope, C. W., Payne, A. J., Pörtner, H.-O., Seneviratne, S. I., Thomas, A., … Zhou, G. (2019). The human imperative of stabilizing global climate change at 1.5°C. *Science*, *365*(6459), eaaw6974. https://doi.org/10.1126/science.aaw6974

Hofmann, G. E., Smith, J. E., Johnson, K. S., Send, U., Levin, L. A., Micheli, F., Paytan, A., Price, N. N., Peterson, B., Takeshita, Y., Matson, P. G., Crook, E. D., Kroeker, K. J., Gambi, M. C., Rivest, E. B., Frieder, C. A., Yu, P. C., & Martz, T. R. (2011). High-Frequency Dynamics of Ocean pH: A Multi-Ecosystem Comparison. *PLoS ONE*, *6*(12), e28983. https://doi.org/10.1371/journal.pone.0028983

Holland, H. (1978). *The chemistry of the atmosphere and oceans*. Wiley.

Humphreys, M. P., Sandborn, D. E., Gregor, L., Pierrot, D., van Heuven, S., S.M.A.C., Lewis, E., & Wallace, D. (2020). PyCO2SYS: Marine carbonate system calculations in Python. *Zenodo*. https://doi.org/10.5281/zenodo.3744275

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Instruments, T. (2013). *AN-1852 Designing With pH Electrodes*. Texas Instruments Incorporated. Dassas, Texas.

IPCC. (2022). *Climate Change 2022: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC. Geneva, Switzerland.

Ivanikova, N. V., Popels, L. C., McKay, R. M. L., & Bullerjahn, G. S. (2007). Lake Superior Supports Novel Clusters of Cyanobacterial Picoplankton. *Appl. Environ. Microbiol.*, *73*(12), 4055–4065. https://doi.org/10.1128/AEM.00214-07

James, D. H., & Lee, F. G. (1974). A model of inorganic carbon limitation in natural waters. *Water Air Soil Pollut*, *3*(3), 315–320. https://doi.org/10.1007/BF00226461

Jeffrey, J. D., Hannan, K. D., Hasler, C. T., & Suski, C. D. (2018). Hot and bothered: Effects of elevated pCO_2 and temperature on juvenile freshwater mussels. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *315*(1), R115–R127. https://doi.org/10.1152/ajpregu.00238.2017

Jiang, L.-Q., Dunne, J., Carter, B. R., Tjiputra, J. F., Terhaar, J., Sharp, J. D., Olsen, A., Alin, S., Bakker, D. C. E., Feely, R. A., Gattuso, J.-P., Hogan, P., Ilyina, T., Lange, N., Lauvset, S. K., Lewis, E. R., Lovato, T., Palmieri, J., Santana-Falcón, Y., … Ziehn, T. (2023). Global Surface Ocean Acidification Indicators From 1750 to 2100. *J Adv Model Earth Syst*, *15*(3), e2022MS003563. https://doi.org/10.1029/2022MS003563

Johannesson, K. H., Horne, J. D., Misra, A., Aliperta, C., Meletis, O. V., Santore, R. C., White, C. D., Mavrommati, G., & Burdige, D. J. (2023). Acidification Of Northeastern USA Lakes From Rising Anthropogenic-Sourced Atmospheric Carbon Dioxide and Its Effects on Aluminum Speciation. *Geophysical Research Letters*, *50*(22), e2023GL104957. https://doi.org/10.1029/2023GL104957

Johnson, T. C., & Eisenreich, S. J. (1979). Silica in Lake Superior: Mass balance considerations and a model for dynamic response to eutrophication. *Geochimica et Cosmochimica Acta*, *43*(1), 77–92. https://doi.org/10.1016/0016-7037(79)90048-6

Jones, D. C., Ito, T., Takano, Y., & Hsu, W.-C. (2014). Spatial and seasonal variability of the air-sea equilibration timescale of carbon dioxide. *Global*

Biogeochemical Cycles, *28*(11), 1163–1178. https://doi.org/10.1002/2014GB004813

Karatayev, A. Y., & Burlakova, L. E. (2022). What we know and don't know about the invasive zebra (Dreissena polymorpha) and quagga (Dreissena rostriformis bugensis) mussels. *Hydrobiologia*. https://doi.org/10.1007/s10750-022-04950-5

Karatayev, A. Y., Burlakova, L. E., Mastitsky, S. E., & Padilla, D. K. (2015). Predicting the spread of aquatic invaders: Insight from 200 years of invasion by zebra mussels. *Ecological Applications*, *25*(2), 430–440. https://doi.org/10.1890/13-1339.1

Karim, A., Dubois, K., & Veizer, J. (2011). Carbon and oxygen dynamics in the Laurentian Great Lakes: Implications for the CO_2 flux from terrestrial aquatic systems to the atmosphere. *Chemical Geology*, *281*(1-2), 133–141. https://doi.org/10.1016/j.chemgeo.2010.12.006

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2019). Machine Learning for the Geosciences: Challenges and Opportunities. *IEEE Trans. Knowl. Data Eng.*, *31*(8), 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006

Kaushal, S. S., Likens, G. E., Pace, M. L., Utz, R. M., Haq, S., Gorman, J., & Grese, M. (2018). Freshwater salinization syndrome on a continental scale. *Proc Natl Acad Sci USA*, *115*(4), E574–E583. https://doi.org/10.1073/pnas.1711234115

Keeling, R. F., & Keeling, C. D. (2017). *Atmospheric monthly in situ CO_2 data - Mauna Loa Observatory, Hawaii*. UC San Diego Library Digital Collections. https://doi.org/10.6075/J08W3BHW

Keppler, L., Landschützer, P., Lauvset, S. K., & Gruber, N. (2023). Recent Trends and Variability in the Oceanic Storage of Dissolved Inorganic Carbon. *Global Biogeochemical Cycles*, *37*(5), e2022GB007677. https://doi.org/10.1029/2022GB007677

Kesler, S. (2019). *Great lakes rocks: 4 billion years of geologic history in the great lakes region*. University of Michigan Press.

Kim, D.-G., Bond-Lamberty, B., Ryu, Y., Seo, B., & Papale, D. (2022). Ideas and perspectives: Enhancing research and monitoring of carbon pools and land-to-atmosphere greenhouse gases exchange in developing countries. *Biogeosciences*, *19*(5), 1435–1450. https://doi.org/10.5194/bg-19-1435-2022

Kovalenko, K. E., Reavie, E. D., Bramburger, A. J., Cotter, A., & Sierszen, M. E. (2019). Nearshore-offshore trends in Lake Superior phytoplankton. *Journal of Great Lakes Research*, *45*(6), 1197–1204. https://doi.org/10.1016/j.jglr.2019.09.016

Kuliński, K., Schneider, B., Hammer, K., Machulik, U., & Schulz-Bull, D. (2014). The influence of dissolved organic matter on the acid–base system of the Baltic Sea. *Journal of Marine Systems*, *132*, 106–115. https://doi.org/10.1016/j.jmarsys.2014.01.011

Lai, C.-Z., DeGrandpre, M. D., Wasser, B. D., Brandon, T. A., Clucas, D. S., Jaqueth, E. J., Benson, Z. D., Beatty, C. M., & Spaulding, R. S. (2016). Spectrophotometric measurement of freshwater pH with purified *meta*-cresol purple and phenol red. *Limnol. Oceanogr. Methods*, *14*(12), 864–873. https://doi.org/10.1002/lom3.10137

Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., Sasse, T. P., & Zeng, J. (2013). A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences*, *10*(11), 7793–7815. https://doi.org/10.5194/bg-10-7793-2013

Latzka, A. W., Crawford, J. T., Koblings, A. S., Caldeira, Y., Hilts, E., & Vander Zanden, M. J. (2015). Representing calcification in distribution models for aquatic invasive species: Surrogates perform as well as $CaCO_3$ saturation state. *Hydrobiologia*, *746*(1), 197–208. https://doi.org/10.1007/s10750-014-2001-7

Lauerwald, R., Allen, G. H., Deemer, B. R., Liu, S., Maavara, T., Raymond, P., Alcott, L., Bastviken, D., Hastie, A., Holgerson, M. A., Johnson, M. S., Lehner, B., Lin, P., Marzadri, A., Ran, L., Tian, H., Yang, X., Yao, Y., & Regnier, P. (2023). Inland Water Greenhouse Gas Budgets for RECCAP2: 2. Regionalization and Homogenization of Estimates. *Global Biogeochemical Cycles*, *37*(5), e2022GB007658. https://doi.org/10.1029/2022GB007658

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., … Zeng, N. (2013). The global carbon budget 1959–2011. *Earth Syst. Sci. Data*, *5*(1), 165–185. https://doi.org/10.5194/essd-5-165-2013

Lee, C., DeGrandpre, M. D., Guthrie, J., Hill, V., Kwok, R., Morison, J., Cox, C., Singh, H., Stanton, T., & Wilkinson, J. (2022). Emerging technologies and approaches for in situ, autonomous observing in the Arctic. *Oceanog.*, *35*(3-4), 210–221. https://doi.org/10.5670/oceanog.2022.127

Lee, K., Kim, T.-W., Byrne, R. H., Millero, F. J., Feely, R. A., & Liu, Y.-M. (2010). The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. *Geochimica et Cosmochimica Acta*, *74*(6), 1801–1811. https://doi.org/10.1016/j.gca.2009.12.027

Lerman, A., & Stumm, W. (1989). CO_2 storage and alkalinity trends in lakes. *Water Research*, *23*(2), 139–146. https://doi.org/10.1016/0043-1354(89)90037-7

Li, J., Crowe, S. A., Miklesh, D., Kistner, M., Canfield, D. E., & Katsev, S. (2012). Carbon mineralization and oxygen dynamics in sediments with deep oxygen penetration, Lake Superior. *Limnol. Oceanogr.*, *57*(6), 1634–1650. https://doi.org/10.4319/lo.2012.57.6.1634

Li, J., Ianaiev, V., Huff, A., Zalusky, J., Ozersky, T., & Katsev, S. (2021). Benthic invaders control the phosphorus cycle in the world's largest freshwater

ecosystem. *Proc. Natl. Acad. Sci. U.S.A.*, *118*(6), e2008223118. https://doi.org/10.1073/pnas.2008223118

Likens, G. E., & Bormann, F. H. (1974). Acid Rain: A Serious Regional Environmental Problem. *Science*, *184*(4142), 1176–1179. https://doi.org/10.1126/science.184.4142.1176

Lin, P., & Guo, L. (2016). Do invasive quagga mussels alter CO_2 dynamics in the Laurentian Great Lakes? *Sci Rep*, *6*(1), 39078. https://doi.org/10.1038/srep39078

Liu, S., Butman, D. E., & Raymond, P. A. (2020). Evaluating CO_2 calculation error from organic alkalinity and pH measurement error in low ionic strength freshwaters. *Limnol. Oceanogr. Methods*, *18*(10), 606–622. https://doi.org/10.1002/lom3.10388

Liu, X., Patsavas, M. C., & Byrne, R. H. (2011). Purification and Characterization of *meta*-Cresol Purple for Spectrophotometric Seawater pH Measurements. *Environ. Sci. Technol.*, *45*(11), 4862–4868. https://doi.org/10.1021/es200665d

Lohrenz, S., Cai, W.-J., Chakraborty, S., Huang, W.-J., Guo, X., He, R., Xue, Z., Fennel, K., Howden, S., & Tian, H. (2018). Satellite estimation of coastal pCO_2 and air-sea flux of carbon dioxide in the northern Gulf of Mexico. *Remote. Sens. Environ.*, *207*, 71–83. https://doi.org/10.1016/j.rse.2017.12.039

Lowe, A. T., Bos, J., & Ruesink, J. (2019). Ecosystem metabolism drives pH variability and modulates long-term ocean acidification in the Northeast Pacific coastal ocean. *Sci Rep*, *9*(1), 963. https://doi.org/10.1038/s41598-018-37764-4

Lueker, T. J., Dickson, A. G., & Keeling, C. D. (2000). Ocean pCO_2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: Validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium. *Mar. Chem.*, *70*(1-3), 105–119. https://doi.org/10.1016/S0304-4203(00)00022-0

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30* (pp. 4765–4774). Curran Associates, Inc. http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf

Lynch, J. K., Beatty, C. M., Seidel, M. P., Jungst, L. J., & DeGrandpre, M. D. (2010). Controls of riverine CO_2 over an annual cycle determined using direct, high temporal resolution pCO_2 measurements. *J. Geophys. Res.*, *115*(G3), G03016. https://doi.org/10.1029/2009JG001132

Ma, J., Shu, H., Yang, B., Byrne, R. H., & Yuan, D. (2019). Spectrophotometric determination of pH and carbonate ion concentrations in seawater: Choices, constraints and consequences. *Analytica Chimica Acta*, *1081*, 18–31. https://doi.org/10.1016/j.aca.2019.06.024

Mackenzie, F. T., & Garrels, R. M. (1971). *Evolution of sedimentary rocks*. Norton.

Mackie, G. L., & Claudi, R. (2009, December 23). *Monitoring and Control of Macrofouling Mollusks in Fresh Water Systems*. CRC Press. https://doi.org/10.1201/9781439804414

Martz, T. R., Connery, J. G., & Johnson, K. S. (2010). Testing the Honeywell Durafet for seawater pH applications. *Limnol. Oceanogr. Methods*, *8*(5), 172–184. https://doi.org/10.4319/lom.2010.8.172

McDonald, C. P., Stets, E. G., Striegl, R. G., & Butman, D. (2013). Inorganic carbon loading as a primary driver of dissolved carbon dioxide concentrations in the lakes and reservoirs of the contiguous United States. *Global Biogeochem. Cycles*, *27*(2), 285–295. https://doi.org/10.1002/gbc.20032

McKinley, G., Urban, N. R., Bennington, V., Pilcher, D., & McDonald, C. (2011). Preliminary Carbon Budgets for the Laurentian Great Lakes. *OCB News*, *4*(2).

McKnight, D. M., & Feder, G. L. (1984). The ecological effect of acid conditions and precipitation of hydrous metal oxides in a Rocky Mountain stream. *Hydrobiologia*, *119*(2), 129–138. https://doi.org/10.1007/BF00011952

Mcmahon, R. F. (1996). The Physiological Ecology of the Zebra Mussel, *Dreissena polymorpha* , in North America and Europe. *Am Zool*, *36*(3), 339–363. https://doi.org/10.1093/icb/36.3.339

McManus, J., Heinen, E. A., & Baehr, M. M. (2003). Hypolimnetic oxidation rates in Lake Superior: Role of dissolved organic material on the lake's carbon budget. *Limnol. Oceanogr.*, *48*(4), 1624–1632. https://doi.org/10.4319/lo.2003.48.4.1624

Mellina, E., & Rasmussen, J. B. (1994). Patterns in the Distribution and Abundance of Zebra Mussel (*Dreissena polymorpha*) in Rivers and Lakes in Relation to Substrate and Other Physicochemical Factors. *Can. J. Fish. Aquat. Sci.*, *51*(5), 1024–1036. https://doi.org/10.1139/f94-102

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., & Shi, W. (2006). North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, *87*(3), 343–360. https://doi.org/10.1175/bams-87-3-343

Michałowski, T., & Asuero, A. G. (2012). New Approaches in Modeling Carbonate Alkalinity and Total Alkalinity. *Critical Reviews in Analytical Chemistry*, *42*(3), 220–244. https://doi.org/10.1080/10408347.2012.660067

Millero, F. J. (1979). The thermodynamics of the carbonate system in seawater. *Geochimica et Cosmochimica Acta*, *43*(10), 1651–1661. https://doi.org/10.1016/0016-7037(79)90184-4

Millero, F. J. (1986). The pH of estuarine waters. *Limnol. Oceanogr.*, *31*(4), 839–847. https://doi.org/10.4319/lo.1986.31.4.0839

Minor, E. C., & Brinkley, G. (2022). Alkalinity, pH, and pCO_2 in the Laurentian Great Lakes: An initial view of seasonal and inter-annual trends. *J. Great*

Lakes Res., S0380133022000211. https://doi.org/10.1016/j.jglr.2022.01.005

Minor, E. C., Forsman, B., & Guildford, S. J. (2014). The effect of a flood pulse on the water column of western Lake Superior, USA. *J. Great Lakes Res.*, *40*(2), 455–462. https://doi.org/10.1016/j.jglr.2014.03.015

Minor, E. C., & Oyler, A. R. (2021). Dissolved organic matter in large lakes: A key but understudied component of the carbon cycle. *Biogeochemistry*. https://doi.org/10.1007/s10533-020-00733-z

Minor, E. C., Tennant, C. J., & Brown, E. T. (2019). A Seasonal to Interannual View of Inorganic and Organic Carbon and pH in Western Lake Superior. *J. Geophys. Res. Biogeosci.*, *124*(2), 405–419. https://doi.org/10.1029/2018JG004664

Moore, C. S., Byrne, R. H., & Yates, K. K. (2023). Technical note: An Assessment of Hg^{II} to Preserve Carbonate System Parameters in Organic-Rich Estuarine Waters. *EGUsphere Preprint*. https://doi.org/10.5194/egusphere-2022-1493

Morse, J. W., Gledhill, D. K., & Millero, F. J. (2003). $CaCO_3$ precipitation kinetics in waters from the Great Bahama Bank. *Geochimica et Cosmochimica Acta*, *67*(15), 2819–2826. https://doi.org/10.1016/S0016-7037(03)00103-0

Mortimer, C. H., Fee, E. J., Rao, D. B., & Schwab, D. J. (1976). Free surface oscillations and tides of Lakes Michigan and Superior. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *281*(1299), 1–61. https://doi.org/10.1098/rsta.1976.0020

Mos, B., Holloway, C., Kelaher, B. P., Santos, I. R., & Dworjanyn, S. A. (2021). Alkalinity of diverse water samples can be altered by mercury preservation and borosilicate vial storage. *Sci Rep*, *11*(1), 9961. https://doi.org/10.1038/s41598-021-89110-w

Mouw, C. B., Chen, H., McKinley, G. A., Effler, S., O'Donnell, D., Perkins, M. G., & Strait, C. (2013). Evaluation and optimization of bio-optical inversion

algorithms for remote sensing of Lake Superior's optical properties. *JGR Oceans*, *118*(4), 1696–1714. https://doi.org/10.1002/jgrc.20139

Müller, B., Meyer, J. S., & Gächter, R. (2016). Alkalinity regulation in calcium carbonate-buffered lakes: Alkalinity regulation in calcium carbonate-buffered lakes. *Limnol. Oceanogr.*, *61*(1), 341–352. https://doi.org/10.1002/lno.10213

Nalepa, T., & Schloesser, D. (2013). *Quagga and zebra mussels: Biology, impacts, and control* (2nd ed.). CRC Press.

National Academies of Sciences, a. M., Engineering. (2018). *Open science by design: Realizing a vision for 21st century research*. The National Academies Press. Retrieved August 2, 2022, from doi.org/10.17226/25116

Newton, J. A., Feely, R. A., Jewett, E. B., Williamson, P., & Mathis, J. (2015, October). *Global ocean acidification observing network: Requirements and governance plan*. Global Ocean Acidification Observing Network.

Ngochera, M. J., & Bootsma, H. A. (2020). Spatial and temporal dynamics of pCO_2 and CO_2 flux in tropical Lake Malawi. *Limnol. Oceanogr.*, *65*(7), 1594–1607. https://doi.org/10.1002/lno.11408

Oh, N.-H., & Richter, D. D. (2004). Soil acidification induced by elevated atmospheric CO_2. *Global Change Biology*, *10*(11), 1936–1946. https://doi.org/10.1111/j.1365-2486.2004.00864.x

Oliveira, M., Hamilton, S., Calheiros, D., Jacobi, C., & Latini, R. (2010). Modeling the potential distribution of the invasive golden mussel Limnoperna fortunei in the Upper Paraguay River system using limnological variables. *Braz. J. Biol.*, *70*, 831–840. https://doi.org/10.1590/S1519-69842010000400014

Orr, J. C., Epitalon, J.-M., & Gattuso, J.-P. (2015). Comparison of ten packages that compute ocean carbonate chemistry. *Biogeosciences*, *12*(5), 1483–1510. https://doi.org/10.5194/bg-12-1483-2015

Orr, J. C., Epitalon, J.-M., Dickson, A. G., & Gattuso, J.-P. (2018). Routine uncertainty propagation for the marine carbon dioxide system. *Mar. Chem.*, *207*, 84–107. https://doi.org/10.1016/j.marchem.2018.10.006

Orr, J. C., Kwiatkowski, L., & Pörtner, H.-O. (2022). Arctic Ocean annual high in pCO_2 could shift from winter to summer. *Nature*, *610*(7930), 94–100. https://doi.org/10.1038/s41586-022-05205-y

Oxburgh, R., Broecker, W. S., & Wanninkhof, R. H. (1991). The carbon budget of Mono Lake. *Global Biogeochem. Cycles*, *5*(4), 359–372. https://doi.org/10.1029/91GB02475

Ozersky, T., Bramburger, A. J., Elgin, A. K., Vanderploeg, H. A., Wang, J., Austin, J. A., Carrick, H. J., Chavarie, L., Depew, D. C., Fisk, A. T., Hampton, S. E., Hinchey, E. K., North, R. L., Wells, M. G., Xenopoulos, M. A., Coleman, M. L., Duhaime, M. B., Fujisaki-Manome, A., McKay, R. M., … Zastepa, A. (2021). The changing face of winter: Lessons and questions from the Laurentian Great Lakes. *J. Geophys. Res. Biogeosci.*, *126*(6). https://doi.org/10.1029/2021JG006247

Park, Kilho., Oliphant, Malcolm., & Freund, Harry. (1963). Conductometric Determination of Alkalinity of Sea Water. *Anal. Chem.*, *35*(10), 1549–1550. https://doi.org/10.1021/ac60203a063

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peng, F., & Effler, S. W. (2011). Characterizations of the light-scattering attributes of mineral particles in Lake Ontario and the effects of whiting. *J. Great Lakes Res.*, *37*(4), 672–682. https://doi.org/10.1016/j.jglr.2011.07.002

Perez, F. F., & Fraga, F. (1987). Association constant of fluoride and hydrogen ions in seawater. *Mar. Chem.*, *21*(2), 161–168. https://doi.org/10.1016/0304-4203(87)90036-3

Perolo, P., Fernández Castro, B., Escoffier, N., Lambert, T., Bouffard, D., & Perga, M.-E. (2021). Accounting for surface waves improves gas flux estimation at high wind speed in a large lake. *Earth Syst. Dynam.*, *12*(4), 1169–1189. https://doi.org/10.5194/esd-12-1169-2021

Pfister, C. A., Esbaugh, A. J., Frieder, C. A., Baumann, H., Bockmon, E. E., White, M. M., Carter, B. R., Benway, H. M., Blanchette, C. A., Carrington, E., McClintock, J. B., McCorkle, D. C., McGillis, W. R., Mooney, T. A., & Ziveri, P. (2014). Detecting the Unexpected: A Research Framework for Ocean Acidification. *Environ. Sci. Technol.*, *48*(17), 9982–9994. https://doi.org/10.1021/es501936p

Phillips, J., McKinley, G., Bennington, V., Bootsma, H., Pilcher, D., Sterner, R. W., & Urban, N. R. (2015). The potential for CO_2-induced acidification in freshwater: A Great Lakes case study. *Oceanog.*, *25*(2), 136–145. https://doi.org/10.5670/oceanog.2015.37

Pierrot, D., Neill, C., Sullivan, K., Castle, R., Wanninkhof, R., Lüger, H., Johannessen, T., Olsen, A., Feely, R. A., & Cosca, C. E. (2009). Recommendations for autonomous underway pCO2 measuring systems and data-reduction routines. *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(8-10), 512–522. https://doi.org/10.1016/j.dsr2.2008.12.005

Pilcher, D. J., McKinley, G. A., Bootsma, H. A., & Bennington, V. (2015). Physical and biogeochemical mechanisms of internal carbon cycling in Lake Michigan. *J. Geophys. Res. Oceans*, *120*(3), 2112–2128. https://doi.org/10.1002/2014JC010594

Podgrajsek, E., Sahlée, E., Bastviken, D., Holst, J., Lindroth, A., Tranvik, L., & Rutgersson, A. (2014). Comparison of floating chamber and eddy covariance measurements of lake greenhouse gas fluxes. *Biogeosciences*, *11*(15), 4225–4233. https://doi.org/10.5194/bg-11-4225-2014

Popper, K. (1966). *Of clouds and clocks: An approach to the problem of rationality and the freedom of man*. Washington University.

Pötter, L., Tollrian, R., Wisotzky, F., & Weiss, L. C. (2021). Determining freshwater pCO_2 based on geochemical calculation and modelling using PHREEQC. *MethodsX*, *8*, 101430. https://doi.org/10.1016/j.mex.2021.101430

Prescott, K., Claudi, R., Jeff, J., & Tanya, V. (2014). Use of the calcite saturation index as an indicator of environmental suitability for dreissenid mussels. *MBI*, *5*(3), 217–224. https://doi.org/10.3391/mbi.2014.5.3.04

Ragg, N. L. C., Gale, S. L., Le, D. V., Hawes, N. A., Burritt, D. J., Young, T., Ericson, J. A., Hilton, Z., Watts, E., Berry, J., & King, N. (2019). The Effects of Aragonite Saturation State on Hatchery-Reared Larvae of the Greenshell Mussel Perna canaliculus. *Journal of Shellfish Research*, *38*(3), 779. https://doi.org/10.2983/035.038.0328

Ramcharan, C. W., Padilla, D. K., & Dodson, S. I. (1992). Models to Predict Potential Occurrence and Density of the Zebra Mussel, *Dreissena polymorpha*. *Can. J. Fish. Aquat. Sci.*, *49*(12), 2611–2620. https://doi.org/10.1139/f92-289

Raymond, P. A., & Hamilton, S. K. (2018). Anthropogenic influences on riverine fluxes of dissolved inorganic carbon to the oceans: Riverine fluxes of inorganic carbon to the oceans. *Limnol. Oceanogr.*, *3*(3), 143–155. https://doi.org/10.1002/lol2.10069

Raymond, P. A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., & Guth, P. (2013). Global carbon dioxide emissions from inland waters. *Nature*, *503*(7476), 355–359. https://doi.org/10.1038/nature12760

Reback, J., Jbrockmendel, McKinney, W., Van Den Bossche, J., Augspurger, T., Cloud, P., Hawkins, S., Roeschke, M., Gfyoung, Sinhrks, Klein, A., Hoefler, P., Petersen, T., Tratner, J., She, C., Ayd, W., Naveh, S., Darbyshire, J.,

Garcia, M., … Seabold, S. (2022, February 12). *Pandas-dev/pandas: Pandas 1.4.1* (Version v1.4.1). https://doi.org/10.5281/ZENODO.3509134

Regier, H., Whillans, T., Christie, W., & Bocking, S. (1999). Over-fishing in the Great Lakes: The context and history of the controversy. *Aquatic Ecosystem Health & Management*, *2*(3), 239–248. https://doi.org/10.1080/14634989908656959

Reimer, J. J., Medeiros, P. M., Hussain, N., Gonski, S. F., Xu, Y.-Y., Huang, T.-H., & Cai, W.-J. (2023). Carbonate Chemistry and the Potential for Acidification in Georgia Coastal Marshes and the South Atlantic Bight, USA. *Estuaries and Coasts*. https://doi.org/10.1007/s12237-023-01261-3

Riebesell, U., Fabry, V. J., Hansson, L., & Gattuso, J.-P. (2011). *Guide to best practices for ocean acidification research and data reporting.* Publications Office. LU. Retrieved April 9, 2022, from https://data.europa.eu/doi/10.2777/66906

Ries, J. B. (2010). Review: Geological and experimental evidence for secular variation in seawater Mg/Ca (calcite-aragonite seas) and its effects on marine biological calcification. *Biogeosciences*, *7*(9), 2795–2849. https://doi.org/10.5194/bg-7-2795-2010

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., Landschützer, P., Metzl, N., Nakaoka, S., Olsen, A., Park, G.-H., Peylin, P., Rodgers, K. B., Sasse, T. P., Schuster, U., Shutler, J. D., Valsala, V., Wanninkhof, R., & Zeng, J. (2015). Data-based estimates of the ocean carbon sink variability – first results of the Surface Ocean pCO_2 Mapping intercomparison (SOCOM). *Biogeosciences*, *12*(23), 7251–7278. https://doi.org/10.5194/bg-12-7251-2015

Rotteveel, L., Heubach, F., & Sterling, S. M. (2022). The Surface Water Chemistry (SWatCh) database: A standardized global database of water chemistry to facilitate large-sample hydrological research. *Earth Syst. Sci. Data*, *14*(10), 4667–4680. https://doi.org/10.5194/essd-14-4667-2022

Rounds, S. A., & Wilde, F. D. (2012). *Chapter A6. Section 6.6. Alkalinity and acid neutralizing capacity* (09-A6.6). United States Geological Survey. https://doi.org/10.3133/twri09A6.6

Saarnisto, M. (1974). The Deglaciation History of the Lake Superior Region and its Climatic Implications. *Quat. Res.*, *4*(3), 316–339. https://doi.org/10.1016/0033-5894(74)90019-2

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., & Rios, A. F. (2004). The Oceanic Sink for Anthropogenic CO_2. *Science*, *305*(5682), 367–371. https://doi.org/10.1126/science.1097403

Sandborn, D. E., & Minor, E. C. (2024). Underway pCO_2 Surveys Unravel CO_2 Invasion of Lake Superior From Seasonal Variability. *J. Geophys. Res. Biogeosci.*, *129*(5), e2023JG007877. https://doi.org/10.1029/2023JG007877

Sandborn, D. E., Minor, E. C., & Hill, C. (2023). Total alkalinity measurement using an open-source platform. *Limnol. Oceanogr. Methods*, *21*(6), 334–344. https://doi.org/10.1002/lom3.10549

Sarmiento, J. L., & Gruber, N. (2006). *Ocean biogeochemical dynamics*. Princeton University Press.

Schilder, J., Bastviken, D., van Hardenbroek, M., Kankaala, P., Rinta, P., Stötter, T., & Heiri, O. (2013). Spatial heterogeneity and lake morphology affect diffusive greenhouse gas emission estimates of lakes. *Geophys. Res. Lett.*, *40*(21), 5752–5756. https://doi.org/10.1002/2013GL057669

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with Python. *9th Python in Science Conference*.

Sepulveda, A. J., Gage, J. A., Counihan, T. D., & Prisciandaro, A. F. (2023). Can big data inform invasive dreissenid mussel risk assessments of habitat suitability? *Hydrobiologia*. https://doi.org/10.1007/s10750-023-05156-z

Shadwick, E. H., Trull, T. W., Thomas, H., & Gibson, J. A. E. (2013). Vulnerability of Polar Oceans to Anthropogenic Acidification: Comparison of Arctic and Antarctic Seasonal Cycles. *Sci Rep*, *3*(1), 2339. https://doi.org/10.1038/srep02339

Shangguan, Q., Prody, A., Wirth, T. S., Briggs, E. M., Martz, T. R., & DeGrandpre, M. D. (2022). An inter-comparison of autonomous in situ instruments for ocean CO2 measurements under laboratory-controlled conditions. *Marine Chemistry*, 104085. https://doi.org/10.1016/j.marchem.2022.104085

Shao, C., Chen, J., Stepien, C. A., Chu, H., Ouyang, Z., Bridgeman, T. B., Czajkowski, K. P., Becker, R. H., & John, R. (2015). Diurnal to annual changes in latent, sensible heat, and CO_2 fluxes over a Laurentian Great Lake: A case study in Western Lake Erie. *J. Geophys. Res. Biogeosci.*, *120*(8), 1587–1604. https://doi.org/10.1002/2015JG003025

Sharp, J. D., & Byrne, R. H. (2020). Interpreting measurements of total alkalinity in marine and estuarine waters in the presence of proton-binding organic matter. *Deep Sea Research Part I: Oceanographic Research Papers*, *165*, 103338. https://doi.org/10.1016/j.dsr.2020.103338

Sharp, J. D., Fassbender, A. J., Carter, B. R., Lavin, P. D., & Sutton, A. J. (2022). A monthly surface pCO_2 product for the California Current Large Marine Ecosystem. *Earth Syst. Sci. Data*, *14*(4), 2081–2108. https://doi.org/10.5194/essd-14-2081-2022

Shuvo, A. K., Lottig, N. R., Webster, K. E., Delany, A., Reinl, K., Gries, C., Smith, N. J., Poisson, A. C., McCullough, I. M., Collins, S. M., King, K. B., Phillips, E., Cheruvelil, K. S., & Soranno, P. A. (2023). *LAGOS-US LIMNO: Data module of surface water chemistry from 1975-2021 for lakes in the conterminous U.S.* Environmental Data Initiative. https://doi.org/10.6073/PASTA/2C58F5A50AB813919F99CC1F265F271C

Sims, R. P., Ahmed, M. M. M., Butterworth, B. J., Duke, P. J., Gonski, S. F., Jones, S. F., Brown, K. A., Mundy, C. J., Williams, W. J., & Else, B. G. T. (2023).

High interannual surface pCO_2 variability in the southern Canadian Arctic Archipelago's Kitikmeot Sea. *Ocean Sci.*, *19*(3), 837–856. https://doi.org/10.5194/os-19-837-2023

Somridhivej, B., & Boyd, C. E. (2016). An assessment of factors affecting the reliability of total alkalinity measurements. *Aquaculture*, *459*, 99–109. https://doi.org/10.1016/j.aquaculture.2016.03.032

Stauffer, R. E. (1990). Electrode pH error, seasonal epilimnetic pCO_2, and the recent acidification of the Maine lakes. *Water Air Soil Pollut.*, *50*(1-2). https://doi.org/10.1007/BF00284788

Sterner, R. W. (2010). In situ-measured primary production in Lake Superior. *J. Great Lakes Res.*, *36*(1), 139–149. https://doi.org/10.1016/j.jglr.2009.12.007

Sterner, R. W. (2011). C:N:P stoichiometry in Lake Superior: Freshwater sea as end member. *Inland Waters*, *1*(1), 29–46. https://doi.org/10.5268/IW-1.1.365

Sterner, R. W. (2021). The Laurentian Great Lakes: A biogeochemical test bed. *Annu. Rev. Earth Planet. Sci.*, *49*(1). https://doi.org/10.1146/annurev-earth-071420-051746

Sterner, R. W., Ostrom, P., Ostrom, N. E., Klump, J. V., Steinman, A. D., Dreelin, E. A., Vander Zanden, M. J., & Fisk, A. T. (2017). Grand challenges for research in the Laurentian Great Lakes. *Limnol. Oceanogr.*, *62*(6), 2510–2523. https://doi.org/10.1002/lno.10585

Sterner, R. W., Reinl, K. L., LaFrancois, B. M., Brovold, S., & Miller, T. R. (2020). A first assessment of cyanobacterial blooms in oligotrophic Lake Superior. *Limnol Oceanogr*, *65*(12), 2984–2998. https://doi.org/10.1002/lno.11569

Stumm, W., & Morgan, J. J. (1996). *Aquatic chemistry: Chemical equilibria and rates in natural waters* (3rd ed). Wiley.

Sulzberger, B. (2015). Light-Induced Redox Cycling of Iron: Roles for CO_2 Uptake and Release by Aquatic Ecosystems. *Aquat. Geochem.*, *21*(2-4), 65–80. https://doi.org/10.1007/s10498-015-9260-4

Sutton, A. J., Battisti, R., Carter, B., Evans, W., Newton, J., Alin, S., Bates, N. R., Cai, W.-J., Currie, K., Feely, R. A., Sabine, C., Tanhua, T., Tilbrook, B., & Wanninkhof, R. (2022). Advancing best practices for assessing trends of ocean acidification time series. *Front. Mar. Sci.*, *9*, 1045667. https://doi.org/10.3389/fmars.2022.1045667

Takahashi, T., Olafsson, J., Goddard, J. G., Chipman, D. W., & Sutherland, S. C. (1993). Seasonal variation of CO_2 and nutrients in the high-latitude surface oceans: A comparative study. *Global Biogeochem. Cycles*, *7*(4), 843–878. https://doi.org/10.1029/93GB02263

Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., Bates, N., Wanninkhof, R., Feely, R. A., Sabine, C., Olafsson, J., & Nojiri, Y. (2002). Global sea–air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects. *Deep Sea Res. Part II Top. Stud. Oceanogr.*, *49*(9-10), 1601–1622. https://doi.org/10.1016/S0967-0645(02)00003-6

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., … de Baar, H. J. (2009). Climatological mean and decadal change in surface ocean pCO_2, and net sea–air CO_2 flux over the global oceans. *Deep Sea Res. Part II Top. Stud. Oceanogr.*, *56*(8-10), 554–577. https://doi.org/10.1016/j.dsr2.2008.12.009

Takeshita, Y., Johnson, K. S., Martz, T. R., Plant, J. N., & Sarmiento, J. L. (2018). Assessment of autonomous pH measurements for determining surface seawater partial pressure of CO_2. *J. Geophys. Res. Oceans*, *123*(6), 4003–4013. https://doi.org/10.1029/2017JC013387

Team, P. D. (2023, April 24). *Pandas-dev/pandas: Pandas* (Version v2.0.1). https://doi.org/10.5281/ZENODO.7857418

Thomas, R., & Dell, C. (1978). Sediments of Lake Superior. *Journal of Great Lakes Research*, *4*(3-4), 264–275. https://doi.org/10.1016/S0380-1330(78)72197-0

Thompson, T., & Anderson, D. (1940). The determination of the alkalinity of seawater. *J. Mar. Res.*, *3*, 224–247.

Tian, H., Yao, Y., Li, Y., Shi, H., Pan, S., Najjar, R. G., Pan, N., Bian, Z., Ciais, P., Cai, W.-J., Dai, M., Friedrichs, M. A. M., Li, H.-Y., Lohrenz, S., & Leung, L. R. (2023). Increased Terrestrial Carbon Export and CO $_2$ Evasion From Global Inland Waters Since the Preindustrial Era. *Global Biogeochemical Cycles*, *37*(10), e2023GB007776. https://doi.org/10.1029/2023GB007776

Toms, B. A., Barnes, E. A., & Ebert-Uphoff, I. (2020). Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability. *J Adv Model Earth Syst*, *12*(9), e2019MS002002. https://doi.org/10.1029/2019MS002002

Turk, J. T., & Adams, D. B. (1983). Sensitivity to acidification of lakes in the flat tops wilderness area, Colorado. *Water Resour. Res.*, *19*(2), 346–350. https://doi.org/10.1029/WR019i002p00346

University of Minnesota Duluth. (2023). RV Blue Heron underway data repository. https://www.rvdata.us/search/vessel/Blue%20Heron

Urban, N. R. (2005). Carbon cycling in Lake Superior. *J. Geophys. Res.*, *110*(C6), C06S90. https://doi.org/10.1029/2003JC002230

Urban, N. R., & Desai, A. R. (2009). Are the Great Lakes a significant net source or sink of CO_2? *SIL Proceedings*, *1922-2010*, *30*(8), 1283–1288. https://doi.org/10.1080/03680770.2009.11923931

Vance, J. M., Currie, K., Zeldis, J., Dillingham, P. W., & Law, C. S. (2022). An empirical MLR for estimating surface layer DIC and a comparative assessment to other gap-filling techniques for ocean carbon time series.

Biogeosciences, *19*(1), 241–269. https://doi.org/10.5194/bg-19-241-2022

Vasys, V. N., Desai, A. R., McKinley, G. A., Bennington, V., Michalak, A. M., & Andrews, A. E. (2011). The influence of carbon exchange of a large lake on regional tracer-transport inversions: Results from Lake Superior. *Environ. Res. Lett.*, *6*(3), 034016. https://doi.org/10.1088/1748-9326/6/3/034016

Verspagen, J. M. H., Van de Waal, D. B., Finke, J. F., Visser, P. M., Van Donk, E., & Huisman, J. (2014). Rising CO_2 Levels Will Intensify Phytoplankton Blooms in Eutrophic and Hypertrophic Lakes (H. G. Dam, Ed.). *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0104325

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Waldbusser, G. G., Hales, B., & Haley, B. A. (2016). Calcium carbonate saturation state: On myths and this or that stories. *ICES Journal of Marine Science*, *73*(3), 563–568. https://doi.org/10.1093/icesjms/fsv174

Waldbusser, G. G., Hales, B., Langdon, C. J., Haley, B. A., Schrader, P., Brunner, E. L., Gray, M. W., Miller, C. A., & Gimenez, I. (2015). Saturation-state sensitivity of marine bivalve larvae to ocean acidification. *Nature Clim Change*, *5*(3), 273–280. https://doi.org/10.1038/nclimate2479

Wallace, R. B., Baumann, H., Grear, J. S., Aller, R. C., & Gobler, C. J. (2014). Coastal ocean acidification: The other eutrophication problem. *Estuarine, Coastal and Shelf Science*, *148*, 1–13. https://doi.org/10.1016/j.ecss.2014.05.027

Wang, B., Liao, Q., Fillingham, J. H., & Bootsma, H. A. (2015). On the coefficients of small eddy and surface divergence models for the air-water gas transfer velocity. *JGR Oceans*, *120*(3), 2129–2146. https://doi.org/10.1002/2014JC010253

Wanninkhof, R. (1992). Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.*, *97*(C5), 7373. https://doi.org/10.1029/92JC00188

Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited: Gas exchange and wind speed over the ocean. *Limnol. Oceanogr. Methods*, *12*(6), 351–362. https://doi.org/10.4319/lom.2014.12.351

Waskom, M. (2021). Seaborn: Statistical data visualization. *J. Open Source Softw.*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

Waters, J., Millero, F., & Woosley, R. (2014). Corrigendum to "The free proton concentration scale for seawater pH", [MARCHE: 149 (2013) 8–22]. *Mar. Chem.*, *165*, 66–67. https://doi.org/10.1016/j.marchem.2014.07.004

Weiler, R. (1978). Chemistry of Lake Superior. *J. Great Lakes Res.*, *4*(3-4), 370–385. https://doi.org/10.1016/S0380-1330(78)72207-0

Weiss, L. C., Pötter, L., Steiger, A., Kruppert, S., Frost, U., & Tollrian, R. (2018). Rising pCO_2 in Freshwater Ecosystems Has the Potential to Negatively Affect Predator-Induced Defenses in Daphnia. *Current Biology*, *28*(2), 327–332.e3. https://doi.org/10.1016/j.cub.2017.12.022

Weiss, R. (1974). Carbon dioxide in water and seawater: The solubility of a non-ideal gas. *Mar. Chem.*, *2*(3), 203–215. https://doi.org/10.1016/0304-4203(74)90015-2

Wetzel, R. G. (2001). *Limnology: Lake and river ecosystems* (3rd). Elsevier. Retrieved March 17, 2022, from http://site.ebrary.com/id/10606261

White, B., Austin, J. A., & Matsumoto, K. (2012). A three-dimensional model of Lake Superior with ice and biogeochemistry. *J. Great Lakes Res.*, *38*(1), 61–71. https://doi.org/10.1016/j.jglr.2011.12.006

Williamson, C. E., Saros, J. E., Vincent, W. F., & Smol, J. P. (2009). Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnol. Oceanogr.*, *54*, 2273–2282. https://doi.org/10.4319/lo.2009.54.6_part_2.2273

Winterdahl, M., Bishop, K., & Erlandsson, M. (2014). Acidification, Dissolved Organic Carbon (DOC) and Climate Change. In B. Freedman (Ed.), *Global Environmental Change* (pp. 281–287). Springer Netherlands. https://doi.org/10.1007/978-94-007-5784-4_107

Wolf-Gladrow, D. A., Zeebe, R. E., Klaas, C., Körtzinger, A., & Dickson, A. G. (2007). Total alkalinity: The explicit conservative expression and its application to biogeochemical processes. *Mar. Chem.*, *106*(1-2), 287–300. https://doi.org/10.1016/j.marchem.2007.01.006

Woolway, R. I., Sharma, S., Weyhenmeyer, G. A., Debolskiy, A., Golub, M., Mercado-Bettín, D., Perroud, M., Stepanenko, V., Tan, Z., Grant, L., Ladwig, R., Mesman, J., Moore, T. N., Shatwell, T., Vanderkelen, I., Austin, J. A., DeGasperi, C. L., Dokulil, M., La Fuente, S., … Jennings, E. (2021). Phenological shifts in lake stratification under climate change. *Nat. Commun.*, *12*(1), 2318. https://doi.org/10.1038/s41467-021-22657-4

Woude, A. V., & Liu, S. (2024). Great Lakes surface environmental analysis. https://coastwatch.glerl.noaa.gov/satellite-data-products/great-lakes-surface-environmental-analysis-glsea/

Xue, L., & Cai, W.-J. (2020). Total alkalinity minus dissolved inorganic carbon as a proxy for deciphering ocean acidification mechanisms. *Marine Chemistry*, *222*, 103791. https://doi.org/10.1016/j.marchem.2020.103791

Xue, P., Ye, X., Pal, J. S., Chu, P. Y., Kayastha, M. B., & Huang, C. (2022). Climate projections over the Great Lakes Region: Using two-way coupling of a regional climate model with a 3-D lake model. *Geosci. Model Dev.*, *15*(11), 4425–4446. https://doi.org/10.5194/gmd-15-4425-2022

Yang, B., Byrne, R. H., & Lindemuth, M. (2015). Contributions of organic alkalinity to total alkalinity in coastal waters: A spectrophotometric approach. *Mar. Chem.*, *176*, 199–207. https://doi.org/10.1016/j.marchem.2015.09.008

Yang, M., Bell, T. G., Bidlot, J.-R., Blomquist, B. W., Butterworth, B. J., Dong, Y., Fairall, C. W., Landwehr, S., Marandino, C. A., Miller, S. D., Saltzman, E. S., & Zavarsky, A. (2022). Global Synthesis of Air-Sea CO_2 Transfer Velocity Estimates From Ship-Based Eddy Covariance Measurements. *Front. Mar. Sci.*, *9*, 826421. https://doi.org/10.3389/fmars.2022.826421

Yang, T.-Y., Kessler, J., Mason, L., Chu, P. Y., & Wang, J. (2020). A consistent Great Lakes ice cover digital data set for winters 1973–2019. *Sci Data*, *7*(1), 259. https://doi.org/10.1038/s41597-020-00603-1

Yao, W., & Byrne, R. H. (1998). Simplified seawater alkalinity analysis: Use of linear array spectrometers. *Deep Sea Research Part I: Oceanographic Research Papers*, *45*(8), 1383–1392. https://doi.org/10.1016/S0967-0637(98)00018-1

Young, F. L., Shangguan, Q., Beatty, C. M., Gilsdorf, M. D., & DeGrandpre, M. D. (2022). Comparison of spectrophotometric and electrochemical pH measurements for calculating freshwater pCO_2. *Limnol. Oceanogr. Methods*, *20*, 514–529. https://doi.org/10.1002/lom3.10501

Zeebe, R. E. (2012). History of Seawater Carbonate Chemistry, Atmospheric CO_2 and Ocean Acidification. *Annu. Rev. Earth Planet. Sci.*, *40*(1), 141–165. https://doi.org/10.1146/annurev-earth-042711-105521

Zeebe, R. E., & Wolf-Gladrow, D. A. (2001). *CO_2 in seawater: Equilibrium, kinetics, isotopes*. Elsevier.

Zhuang, J., Dussin, R., Huard, D., Bourgault, P., Banihirwe, A., Stephane Raynaud, Brewster Malevich, Schupfner, M., Filipe, Levang, S., Gauthier, C., Jüling, A., Almansi, M., RichardScottOZ, RondeauG, Rasp, S., Smith, T. J., Stachelek, J., Plough, M., … Xianxiang Li. (2023, September 18).

Pangeo-data/xESMF: V0.8.2 (Version v0.8.2). https://doi.org/10.5281/ZENODO.4294774

Zigah, P. K., Minor, E. C., Werne, J. P., & McCallister, S. L. (2011). Radiocarbon and stable carbon isotopic insights into provenance and cycling of carbon in Lake Superior. *Limnol. Oceanogr.*, *56*(3), 867–886. https://doi.org/10.4319/lo.2011.56.3.0867

Appendix A

Glossary and Acronyms

Care has been taken in this thesis to minimize the use of acronyms. This appendix contains a table of necessary acronyms and their meaning.

A.1 Acronyms

Table A.1: Acronyms

Acronym	Meaning
A_T	Total Alkalinity
DIC	Dissolved Inorganic Carbon
ISFET	Ion-Sensitive Field Effect Transistors
fCO_2	Fugacity of Carbon Dioxide
FFNN	Feed-Forward Neural Network
FVCOM	Finite Volume Community Ocean Model
GLNPO	Great Lakes National Program Office
GLSEA	Great Lakes Surface Environmental Atlas

Continued on next page

Table A.1 – Continued from previous page

Acronym	Meaning
NBS	National Bureau of Standards
PAR	Photosynthetically-Active Radiation
pCO_2	Partial Pressure of Carbon Dioxide
PIC	Particulate Inorganic Carbon
POC	Particulate Organic Carbon
SHAP	Shapley Additive Explanation
SOCAT	Surface Ocean CO_2Atlas
SST	Sea Surface Temperature
U_{10}	Neutral Wind Speed at 10 m
Ω_{Ar}	Saturation State of Aragonite
Ω_{Ca}	Saturation State of Calcite