Examining Algal Bloom Disturbances with High-Frequency Time Series and High-Resolution Spatial Data

Dat Tat Ha
M.S. Thesis

B.A., University of Virginia, 2019

A thesis presented to the Graduate Faculty
of the University of Virginia in Candidacy for the degree of
Master of Science

Department of Environmental Sciences

University of Virginia
November 2023

Abstract

Lakes are important ecosystems that are under threat from internal disturbances such as algal blooms and external forcings such as climate change. Algal blooms are widespread disturbances that can negatively affect ecosystems services. Understanding algal bloom dynamics can help prevent and mitigate the negative consequences including toxicity, disruption of potable water, and closure of recreational activities. As high frequency data become more accessible and cheaper, the data can be leveraged for improving the understanding of temporal and spatial patterns. This study used high frequency phytoplankton pigment and water quality data to explore disturbance and spatial dynamics. First, I applied a recently published disturbance-recovery algorithm to quantify the magnitude of algal blooms and time to recovery using high frequency pigment time series from several lakes. Results for the first study indicate that the algorithm can detect disturbance and recovery in experimental and monitored lakes. The algorithm performs best for experimental lakes where reference data is available, whereas the algorithm detects disturbances that are intense and occur at dissimilar times for monitored lakes. Second, I analyzed high resolution spatial-temporal data to identify spatial variability and hotspots in an experimental lake undergoing nutrient addition and in an adjacent, unperturbed reference lake. Results for the second study indicate that there is no long-lasting spatial structure and low spatial heterogeneity for both an experimental lake undergoing an algal bloom and a similarly sized reference lake. These studies show how advanced technology can reveal temporal and spatial dynamics of lakes, specifically those undergoing algal bloom disturbances.

Acknowledgements

There are many people who have helped me throughout this journey that I am forever grateful for. I first want to thank my advisor, Mike Pace, who gave me the opportunity to come back to UVA for grad school. I'm extremely thankful for your patience, mentorship, and support throughout this entire process.

I'd like to thank everyone in the Pace Lab and the Cascade research group: Cal Buelo, Spencer Tassone, Jon Walter, Steve Carpenter, David Ortiz, Emily Stanley, Danny Szydlowski, and Grace Wilkinson. I consider myself extremely lucky to be able to learn and collaborate with some of the brightest and funniest limnologists in the field. A special thank you to Cal, who, during my third year of undergrad, provided me the opportunity to go to northern Wisconsin and UNDERC, and opened my eyes to a new world.

Lastly, I want to thank my close friends and family for all their love and support: Tricia Swetnam, Steven Swetnam, Chris Helzer, Mary Solokas, and Stephen Stern, Trang Pham, Duong Ha, Thanh Ha, and Becca Hostetter.

Table of Contents

Abstract..... 2
Acknowledgements.....3
Table of Contents 4
List of Tables and Figures 5

Introduction..... 6

Chapter 1 - Quantifying algal blooms with high frequency data and disturbance-recovery algorithm
.....
Abstract..... 10
Introduction..... 11
Methods..... 12
Results.....15
Discussion..... 17
References..... 21

Chapter 2 - Identifying spatial variability in experimental and reference lakes
.....
Abstract.....37
Introduction..... 38
Methods..... 39
Results..... 42
Discussion.....45
References..... 49

Conclusions..... 58

List of Tables and Figures

Table 1.1. Steps of applying disturbance recovery algorithm…… 24
Table 1.2. Lakes analyzed for algal bloom disturbance and recovery…… 25
Table 1.3. Disturbances and recoveries from the experimental lakes…….. 26
Table 1.4. Disturbances and recoveries from the monitored lakes…… 27
Figure 1.1. Disturbance and recovery of Peter Lake in 2015…… 28
Figure 1.2. Disturbances and recoveries of Peter Lake in 2019…… 29
Figure 1.3. Disturbance and recovery of Tuesday Lake in 2015…….. 30
Figure 1.4. Disturbances and recoveries of Lake Lillinonah…… 31
Figure 1.5. Disturbances and recoveries of Seneca Lake…… 32
Figure 1.6. Disturbances and recoveries of Lake Erie (Toledo buoy) …… 33
Figure 1.7. Disturbances and recoveries of Muskegon Lake…… 34
Figure 1.8. Disturbances and recoveries of Lake Sammamish…… 35
Figure 1.9. Relationship between disturbance length and peak disturbance severity…… 36

Table 2.1. Temporal versus spatial coefficient of variation…… 52
Figure 2.1. Time series of chlorophyll-a for Peter and Paul Lakes …… 53
Figure 2.5. Map of the 5 studied variables for Peter and Paul Lakes in 2019……54
Figure 2.3. Time series of coefficients of variation of the 5 studied variables …… 55
Figure 2.4. Time series of the mean values and Moran's I for Peter and Paul Lakes in 2019…….56
Figure 2.6. Hotspot analysis of dissolved oxygen for Peter Lake in 2019…… 57

Introduction

Lakes are important ecosystems that provide habitat for a wide range of species, sources of drinking water and recreation for humans, and have high rates of biogeochemical cycling that contribute to greenhouse gas production (Tranvik et al. 2009, Moss 2012). Due to climate warming, lakes are undergoing rapid changes from loss of ice cover (Sharma et al. 2019), warming surface temperatures (O'Reilly et al. 2015), and alterations in mixing regimes (Woolway et al. 2019). Additionally, enhanced inputs of nutrients and sediments are well known causes of eutrophication that often result in algal blooms negatively affecting water quality. With these threats and many others to lakes, monitoring waterbodies is essential for evaluating changes in resilience and implementing management interventions to mitigate negative impacts.

High frequency sensors have become ubiquitous for water quality monitoring. They allow researchers and managers to measure dynamics at fine temporal and spatial scales. Widely available sensors can quantify variables such as temperature, dissolved oxygen, and chlorophyll-a at temporal resolution of minutes. Prior studies have used high frequency data to detect indicators of regime shifts (Carpenter et al., 2011), diel patterns and variability of temperature (Woolway et al. 2015), and diel variability of methane emissions (Sieczko et al. 2020). Additionally, Pace et al. (2017) and Wilkinson et al. (2018) used high frequency data from in-situ sensors to test for early warnings of algal blooms in experimentally fertilized lakes. High frequency data are especially useful for assessing disturbances relative to background natural variability.

In addition to high frequency temporal data, sensors can also provide high resolution spatial data. One example is the Fast Limnological Automated Measurement (FLAMe,) system which collects high resolution water quality data by drawing water past sensors in a moving boat (Crawford et al., 2015). The FLAMe system simultaneously records spatial location enabling resolution of spatial heterogeneity for lakes and rivers. Such measurements overcome limitations of single point sampling that may misrepresent how an ecosystem is behaving or responding to a disturbance if the system exhibits spatial heterogeneity. Loken et al. (2019) used the FLAMe to look at spatial and temporal variability of carbon dioxide and methane flux and concentration in Lake Mendota and found that the single point sampling at the lake center overestimated CO_2 and underestimated CH_4 emissions, respectively. Additionally, ecosystem metabolism in two small temperate lakes varied by 1–2 orders of magnitude depending on the measurement location within the lakes (Van de Bogert et al., 2012).

Algal blooms can develop rapidly and unpredictably, as well as exhibit spatial heterogeneity as evident in satellite imagery of large lakes. High resolution sampling of blooms allows better understanding of dynamics with the potential to contribute to better management of problems associated with excessive algal biomass. In this study, I analyzed existing high frequency time series and high resolution spatial-temporal data with a focus on understanding of algal blooms as disturbances in lakes. In the first chapter, I use high frequency time series data and a recently published disturbance-recovery algorithm (Walter et al. 2022) to quantify bloom magnitudes (disturbance) and return time to non-bloom conditions (recovery). I analyzed high frequency data of phytoplankton pigments from both whole lake nutrient addition studies and routine monitoring. The first goal was to test if the method adequately

detects blooms in experimental lakes where nutrient inputs were controlled, and phytoplankton dynamics measured at both high frequency and daily time scales. The second goal was to explore application of the method to monitoring data where nutrient inputs are unknown and variables other than water quality parameters are unmonitored.

For the second chapter, I analyzed high resolution spatial data to quantify spatial variability and tested for hotspots in an experimental lake undergoing a bloom and a reference lake where no bloom occurred. The goal was to determine if small lakes exhibit spatial heterogeneity and if heterogeneity increased during an algal bloom.

References

Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R., Brock, W. A., Cline, T., Coloso, J., Hodgson, J.R., Kitchell, J.F., Seekell, D.A., Smith, L., & Weidel, B. (2011). Early warnings of regime shifts: a whole-ecosystem experiment. Science, 332(6033), 1079-1082

Crawford, J. T., L. C. Loken, N. J. Casson, C. Smith, A. G. Stone, and L. A. Winslow. 2015. "High-Speed Limnology: Using Advanced Sensors to Investigate Spatial Variability in Biogeochemistry and Hydrology." Environmental Science & Technology 49(1): 442– 50.

Loken, L. C., Crawford, J. T., Schramm, P. J., Stadler, P., Desai, A. R., & Stanley, E. H. (2019). Large spatial and temporal variability of carbon dioxide and methane in a eutrophic lake. Journal of Geophysical Research: Biogeosciences, 124(7), 2248-2266.

Moss, B. (2012). Cogs in the endless machine: lakes, climate change and nutrient cycles: a review. Science of the Total Environment, 434, 130-142.

O'Reilly, C. M. et al. Rapid and highly variable warming of lake surface waters around the globe. Geophys. Res. Lett. 42, 10773–10781 (2015).

Pace, M. L., Batt, R. D., Buelo, C. D., Carpenter, S. R., Cole, J. J., Kurtzweil, J. T., & Wilkinson, G. M. (2017). Reversal of a cyanobacterial bloom in response to early warnings. Proceedings of the National Academy of Sciences, 114(2), 352–357

Sharma, S., et al. (2019). Widespread loss of lake ice around the Northern Hemisphere in a warming world. Nat. Clim. Change 9, 227–231.

Sieczko, A. K., Duc, N. T., Schenk, J., Pajala, G., Rudberg, D., Sawakuchi, H. O., & Bastviken, D. (2020). Diel variability of methane emissions from lakes. Proceedings of the National Academy of Sciences, 117(35), 21488-21494

Tranvik, L. J., et al. (2009). Lakes and Reservoirs as Regulators of Carbon Cycling and Climate. Limnology and Oceanography, 54(6), 2298 - 2297

Van de Bogert, M. C., Bade, D. L., Carpenter, S. R., Cole, J. J., Pace, M. L., Hanson, P. C., & Langman, O. C. (2012). Spatial heterogeneity strongly affects estimates of ecosystem metabolism in two north temperate lakes. Limnology and Oceanography, 57(6), 1689– 1700.

Woolway, R., Jones, I., Feuchtmayr, H., & Maberly, S. (2015). A comparison of the diel variability in epilimnetic temperature for five lakes in the English Lake District. Inland Waters, 5(2), 139–154

Woolway, R. I., & Merchant, C. J. (2019). Worldwide alteration of lake mixing regimes in response to climate change. Nature Geoscience, 12(4), 271 - 278.

I. Quantifying algal blooms with high frequency data and disturbance-recovery algorithm

Abstract

Algal blooms are disturbances to aquatic ecosystems that can impact water quality and ecosystem services. Quantifying algal blooms is difficult due to lack of data and concentration thresholds can vary based on the water body. Additionally, there is no standard approach to detecting algal bloom disturbance magnitudes and recovery times, making it challenging to compare disturbances within and among aquatic systems. Using high frequency phycocyanin and chlorophyll-a data and a newly developed disturbance-recovery algorithm, I quantified the magnitude and duration of algal bloom disturbances. I applied the algorithm to nutrient enriched experimental lakes with detailed algal data and non-experimental lakes with pigment sensor time series that extended over many years. Algal bloom onsets, magnitudes, and recoveries were detected and consistent with observed dynamics in the experimental lakes, facilitating within and cross-system comparisons. In non-experimental lakes, I identified the most severe/intense disturbances and those that occurred at unusual times. Recovery time and peak disturbance magnitude differed among non-experimental lakes. Lakes with phycocyanin and chlorophyll-a time series rarely had concurrent disturbances in both variables. Detecting algal bloom disturbances and recoveries requires appropriate reference data, and this problem is more difficult in non-experimental lakes due to the complexities of reference data selection, missing values, and pigment sensor errors. Overall, the approach shows promise in quantifying algal bloom dynamics where long-term high frequency data are available.

Introduction

Ecosystems are variably disturbed and often undergo recovery (Romme et al., 1998, Turner et al., 1998). Disturbances can affect an ecosystem's resilience causing shifts from a stable state to an alternative state (Gunderson 2000). A critical question is whether ecosystems maintain the ability to recover after repeated and large disturbances or if there are changes, especially more extended recoveries, that might portend loss of resilience and pending transformation of ecosystem state (Pickett et al., 1989, Scheffer et al., 2009). Another important general problem is determining the magnitude of a disturbance and analyzing if recovery time is related to magnitude (Weathers et al., 2016). A third general problem is comparing disturbances both magnitudes and recovery times within and among ecosystems. Quantifying the magnitude of disturbance and the time to recovery is required to address these issues.

Excessive developments of phytoplankton biomass (referred to as blooms) represent a form of ecosystem disturbance caused by conditions that promote rapid growth and accumulation (Paerl et al., 2001). Algal blooms occur due to high rates of nutrient loading accompanied by long water residence time, warm temperatures (usually), low grazing rates, low mortality due to disease (Paerl, 1988, Paerl et al., 2013). Algal blooms are significant disturbances to aquatic ecosystems and can impair water quality, result in fish die-offs, and create toxins dangerous to humans and animals. Cyanobacteria are the most prominent bloom forming species in inland waters (Sukenik et al., 2021).

Walter et al. (2022) developed a method for detecting disturbance and recovery using high frequency data. The method employs long reference time series and applies an algorithm

that quantifies differences in the empirical cumulative distribution functions of moving windows over reference and test (putative disturbance) periods. The method ultimately allows quantification of disturbance occurrence, timing, recovery time, and severity. Buelo et al. (2023) recently used the algorithm to analyze high-frequency and long-term salinity and dissolved oxygen data across 19 estuaries and found that disturbances due to tropical cyclones varied and that disturbance magnitude and duration are related to cyclone and site properties.

In this study, I used this disturbance and recovery algorithm and long-term high frequency time series of chlorophyll-a and phycocyanin from experimental and monitored lakes to ask the following questions:

1) Can the algorithm detect known phytoplankton blooms in experimental lakes with nutrient additions and in this context, what are the best reference time series to facilitate detecting these disturbances?

2) Can the disturbance-recovery algorithm detect anomalous phytoplankton blooms in monitored lakes that have long time series?

3) Is there a relationship between disturbance magnitude and recovery time?

Methods

The algorithm identifies disturbance and recovery by quantifying the empirical distribution function difference between a time series variable observed within moving windows and a reference distribution representing the normal state of the system (Walter et al., 2022). The method then quantifies how atypical these differences are by benchmarking them against the distribution of differences observed in moving windows of the same width

during a reference period (Table 1). An increase in the difference from the reference state past a specified threshold defines the onset of disturbance, and a return toward the reference state, past another threshold, defines recovery. The method requires setting several aspects of the analysis including the reference data, the width and time step (stride) of the moving window, and the thresholds for disturbance and recovery. For all analyses in this study, I used a window width of 10 days, stride of 1, disturbance threshold of z-score=2.5, and disturbance length minimum of 1 day.

Phytoplankton pigment data, specifically chlorophyll-a and phycocyanin, are important water quality variables that can describe algal bloom conditions. Chlorophyll-a is commonly used as a proxy for primary production and phytoplankton biomass. Phycocyanin is a pigment unique to cyanobacteria and is related to biomass of that taxa. High concentrations of chlorophyll-a indicate an algal bloom, and high concentrations of phycocyanin indicate a cyanobacterial bloom.

I first tested the algorithm with chlorophyll-a data collected from experimental lakes as part of nutrient addition studies (Table 2). Peter, Tuesday, and Paul lakes are lakes located at the University of Notre Dame Environmental Research Center in the Upper Peninsula of Michigan, United States. These lakes have been used for a variety of ecosystem manipulation studies with Paul Lake serving as an unmanipulated reference (Carpenter and Pace 2018). For the manipulations considered here, Peter and Tuesday Lakes received several years of nutrient additions and Paul Lake was a reference lake that never received nutrients. Peter and Tuesday lakes were fertilized daily with inorganic nitrogen and phosphorus during the summers of 2013, 2014, and 2015. Additionally, Peter Lake received daily nutrient additions in the summer of

2019. Details of addition methods, loading rates, and nutrient ratios are in Wilkinson et al. (2018).

In the experimental lake analysis, data from non-fertilization years of Peter and Tuesday lakes were used as a reference set for the respective lake when it is in a low nutrient state. For Peter Lake, there were 491 reference data points from 22 non-experimental years collected at weekly to daily time-intervals. For Tuesday Lake, there were 214 observations from 18 non-experimental years. Additionally, I tested the responses of Peter and Tuesday lakes using Paul Lake data as a reference.

For all the years of nutrient addition manipulations, chlorophyll-a was measured daily by manually collecting water samples that were then filtered, frozen, methanol extracted, and measured in the laboratory using a Turner Trilogy benchtop fluorometer (Holm-Hansen & Riemann, 1978). These daily observations provided the data for analysis of bloom disturbances.

In a second part of this study, I used data from five monitored lakes to assess whether variable and unusual disturbances were detectable from high frequency time series. In the monitored lake analysis, I analyzed each year compared to all other years as reference data. Five lakes were considered that had suitable long-term, high frequency data (Table 2). In lakes Lillinonah, Erie, and Muskegon chlorophyll-a and phycocyanin data were available while Seneca and Sammamish lakes only had data for chlorophyll-a. High frequency pigment data from all the monitored lakes were measured with different versions of a YSI EXO sonde with the pigment sensors of that instrument (Zamyadi et al., 2012). The data were selected based on the criteria that there were at least 4 years of summer pigment data. The data collection frequency across the monitored lakes ranged from 15 minutes to daily, therefore, I analyzed all data using daily

averages. The data were further inspected to remove outliers, negative values, and observations marked with low quality flags from the data provider. The algorithm tested each year using the other years in the dataset as the reference years with the goal of identifying blooms either having an unusual magnitude/recovery or occurring at anomalous times.

Results

Both the experimental lakes and reference lakes typically have low chlorophyll concentrations (< 10 $\mu g\ L^{-1}$) indicative of oligotrophic to mesotrophic conditions. Specifically, the experimental lakes had mean chlorophyll-a concentrations during non-fertilization years of 3.81 $\mu g\ L^{-1}$ (Peter Lake) and 6.10 $\mu g\ L^{-1}$ (Tuesday Lake). The mean chlorophyll-a concentration for the reference lake, Paul, was 3.16 $\mu g\ L^{-1}$, close to the value of Peter Lake.

In Peter Lake, algal blooms varied among the 4 nutrient fertilization years in terms of detection, magnitude, and recovery. There was no detected disturbance in 2013 despite summer-long additions of nutrients. Mean and maximum chlorophyll-a concentrations were 3.86 and 9.77 $\mu g\ L^{-1}$, respectively. There was one disturbance detected in 2014 on DOY 215 and the lake remained in a disturbed (or bloom) state through the remaining measurement period (ending DOY 250). The peak z-score for the 2014 disturbance was 4.97. The mean and maximum chlorophyll-a concentration in 2014 were 5.78 and 13.34 $\mu g\ L^{-1}$, respectively. In 2015, Peter Lake experienced a much earlier and larger disturbance starting on DOY 167. In this year nutrient additions were halted on DOY 180 (instead of continued) as a part of the study design (Pace et al. 2017). The lake recovered to baseline conditions after 41 days on DOY 208. (Figure 1.1). The peak z-score of 28.4 for the 2015 disturbance was much higher than for 2014

reflecting the brief but high maximum chlorophyll-a concentration of 40.2 μg L^{-1}. In 2019, Peter Lake experienced two disturbances and one recovery. The algorithm detected a disturbance on DOY 198 and recovery on DOY 229 (Figure 1.2). Nutrient additions were ongoing throughout this study in contrast with the more limited nutrient addition in the 2015 manipulation of Peter Lake (see above). The peak z-score was 38.7 and the max chlorophyll-a concentration was 49.0 μg L^{-1}. There was a second smaller disturbance detected on DOY 237 and no recovery occurred prior to the end of sampling on DOY 248.

Bloom disturbances occurred in Tuesday Lake in all three years of nutrient addition. Disturbances began on DOY 204 and 162 in 2013 and 2014, respectively. The peak z-scores for 2013 and 2014 were 13.18 and 8.13, respectively. In 2015 there were two disturbances. The first disturbance was relatively small with a peaked z-score of 4.48 but lasted 45 days. The second disturbance detected on DOY 218 was large with a peak z-score 46.0. Maximum chlorophyll-a concentrations were 74.7 μg L^{-1} and the lake did not recover prior to the end of sampling.

Using Paul Lake as a reference for Peter and Tuesday Lakes in the algorithm produced different results. When using Paul Lake as a reference for Peter Lake compared to using non-manipulated Peter Lake data, the algorithm detected more disturbances. The algorithm detected one disturbance in 2013, two additional disturbances in 2014, and one more disturbance in 2019. The main disturbance in 2015 and 2019 lasted longer by 5 and 4 days, respectively. When using Paul Lake as a reference for Tuesday Lake compared to using non-manipulated Tuesday Lake data, the algorithm detected a disturbance at the beginning of each of the time series and never detected recovery. Paul Lake, the reference lake, never received

nutrient additions. The algorithm detected no disturbance in 2013, 2014, or 2015. One disturbance in 2019 was detected on DOY 219 with a peak z-score of 5.65 when chlorophyll-a reached 5.78 $\mu g\ L^{-1}$. Maximum chlorophyll-a concentration was 8.28 $\mu g\ L^{-1}$ and the lake did not recover prior to the end of sampling.

The algorithm detected disturbances in the non-experimental, monitored lakes that differed from 'normal' conditions based on using all other years as a reference. Lake Lillinonah, Lake Erie, and Muskegon Lake had disturbances in both types of pigment data. Collectively, the three lakes had eight chlorophyll-a and six phycocyanin disturbances, but these disturbances were typically not concurrent with the exception of Lake Lillinonah in 2014 and Muskegon Lake in 2012. Overall, there were six phycocyanin and thirty-one chlorophyll-a disturbances among the five lakes. The peak z-scores ranged from 2.52-33.18 and the disturbance length ranged from 7-71 days. There was a significant, positive relationship between peak severity and disturbance length (Figure 1.9), with a correlation coefficient of $r = 0.49$ ($p < 0.004$). While there is no established period for onset of negative impacts from blooms, recovery times were often > 20 days (17 cases) including some of the more oligotrophic of the monitored lakes.

Discussion

The method successfully detected disturbance and recovery in experimental lakes and identified unusual disturbances in monitored lakes. A benefit of applying the approach is in objectively quantifying the exact date and pigment concentration where a disturbance was detected as well as a time to recovery. The z-score provides a magnitude of the disturbance which allows comparison of disturbances among years and lakes.

The experimental lakes represent an idealized case where nutrient inputs were controlled reference data were extensive. Paul Lake served as a reference system and had only one small disturbance in one year. Using Paul as the source of reference data was problematic because of differences both in the mean values of chlorophyll-a and the variability between the lakes. The early disturbances in Tuesday Lake using data from Paul Lake as reference for all three tested years was caused by the normally higher chlorophyll concentrations in Tuesday Lake compared to Paul Lake, indicating the difficulty of using a separate lake as reference data for the disturbance-recovery analysis. I did not try creating a specific offset based on differences in means for Peter and Tuesday lakes which could be one way of incorporating data from a separate reference lake. Using the data from the same lake in non-nutrient years as a reference proved more effective. Fortunately, there are many years and hundreds of observations from Peter and Tuesday lakes to establish non-nutrient enrichment conditions. The results suggest our approach would work well for quantifying algal blooms in lakes with long-term data that were transitioning from infrequent to more frequent blooms.

In the experimental lakes, blooms differed among lakes and years. The cause of these differences has been previously analyzed and include effects of dissolved organic matter and grazers (Pace et al. 2019; Carpenter et al. 2022). In some of the nutrient addition years, two blooms were observed, specifically for Tuesday Lake, 2015 and Peter Lake 2019. Both lakes were fertilized throughout the season, but blooms can collapse due to disease, temperature, or grazing (rather than nutrient exhaustion) and there can also be a shift in the dominant bloom forming algae (Mayer et al., 1997, Dokulil & Teubner, 2000, Zhang et al., 2016). The important point from this study is these dynamics were quantifiable. The method provides a way to

objectively measure important features of blooms which as evidenced by the experimental lakes can be quite variable.

Applying the algorithm to monitored lakes is more difficult compared to experimental lakes. With this application, normal algal bloom disturbances may occur regularly such that reference data include periods of high algal biomass. Without having a non-bloom reference data set similar to the experimental lakes considered above, all disturbances will not be detected. Instead, the most severe/intense disturbances are measurable or those that occur at a dissimilar time. For example, phytoplankton biomass is routinely high in Muskegon Lake, and the system would be considered eutrophic or hypereutrophic (Carlson, 1991). Muskegon Lake had intense harmful algal blooms in the past but has slowly improved with reduced blooms in recent years (Mancuso et al., 2021). The algorithm only detected chlorophyll-a disturbances in 2012, 2014, and 2015, and phycocyanin disturbances in 2011 and 2012 (Figure 1.7). These 2012, 2014, and 2015 blooms had high chlorophyll-a concentrations, but other years had blooms that were smaller in magnitude, such as those in 2011 and 2023 and therefore not detected by the analysis. Such blooms would normally be considered disturbances but there isn't a non-bloom (or low algal biomass) reference dataset to compare as was the case for the experimental lakes. This problem also makes it difficult to evaluate whether there were concurrent chlorophyll-a and phycocyanin disturbances. Despite these limitations, the algorithm still detected the most intense disturbances and quantified disturbance initiation and recovery dates.

Disturbance length and disturbance magnitude were positively correlated for the monitored lakes. The more intense the algal bloom, the longer it takes for the system to

recover. The duration of the algal bloom is important to consider and has consequences for system processes like anoxia and toxicity effects (Paerl et al., 2011). Comparing bloom recovery in water bodies over time would be of interest in determining whether stability is declining under eutrophication or climate warming or alternatively recovery times might decrease under management to reduce nutrient loading.

One issue with using the algorithm and detecting pigment disturbances are missing values and pigment sensor errors. Sensor calibration and periods of sensor failure result in interrupted time series. Various methods of interpolation can be applied to impute missing values. For the monitored lakes, erroneous values were filled in with linear interpolation. Pigment sensor measurements rely on optical fluorescence properties. The intensity of the fluorescence emitted by the pigments is directly related to the concentration of the pigments in the water and can be affected by turbidity, dominant algal group, or fouled sensors (Chang et al., 2012, Zamyadi et al., 2012). Additionally, chlorophyll-a and phycocyanin can interfere with each other's measurement (Bowling et al., 2016). These factors might cause erroneous measurements and impact the detection of bloom metrics and could explain the lack of concurrent chlorophyll-a and phycocyanin blooms in the monitored lake.

There is conflicting evidence about whether algal blooms are increasing at regional and global scales (Ho et al., 2019, Wilkinson et al., 2022, Topp et al., 2021). Ho et al. (2019) found increases in phytoplankton blooms globally whereas Wilkinson et al. (2022) found no increase in algal bloom metrics in a regional analysis. Topp et al. (2021) found improving water clarity in thousands of U.S. lakes. The differences among these studies are most likely due to differences in the lakes studied and due to measurement methods (e.g., satellite remote sensing versus in

situ water samples). As improved technology and monitoring systems enable longer and more accurate time series, application of disturbance and recovery analysis to blooms over large scales might provide a way of evaluating not only trends but features of blooms. Eventually, the question of whether lakes are becoming greener due to increased phytoplankton could be extended to asking whether lakes are having more severe and more frequent blooms. The magnitudes and duration of blooms have important consequences for human uses of inland waters and better knowledge of these dynamics will enable better management to mitigate and reduce negative impacts.

References

Bowling, L. C., Zamyadi, A., & Henderson, R. K. (2016). Assessment of in situ fluorometry to measure cyanobacterial presence in water bodies with diverse cyanobacterial populations. Water research, 105, 22-33.

Buelo, C. D., A. F. Besterman, J. A. Walter, M. L. Pace, D. T. Ha, S. J. Tassone. (2023). Quantifying disturbance and recovery in estuaries: tropical cyclones and high frequency measures of oxygen and salinity. Estuaries and Coasts.

Carlson, R. E. (1991). Expanding the trophic state concept to identify non-nutrient limited lakes and reservoirs. Enhancing the states's lake management programs, 59-71.

Carpenter, S. R., & Pace, M. L. (2018). Synthesis of a 33-yr series of whole-lake experiments: Effects of nutrients, grazers, and precipitation-driven water color on chlorophyll. Limnology and Oceanography Letters, 3(6), 419-427.

Carpenter, S. R., Pace, M. L., & Wilkinson, G. M. (2022). DOC, grazers, and resilience of phytoplankton to enrichment. Limnology and Oceanography Letters, 7(6), 466-474.

Chang, D. W., Hobson, P., Burch, M., & Lin, T. F. (2012). Measurement of cyanobacteria using in-vivo fluoroscopy–Effect of cyanobacterial species, pigments, and colonies. Water research, 46(16), 5037-5048.

Dokulil, M. T., & Teubner, K. (2000). Cyanobacterial dominance in lakes. Hydrobiologia, 438, 1-12.

Gunderson, L. H. (2000). Ecological resilience—in theory and application. Annual review of ecology and systematics, 31(1), 425-439.

Ho, J. C., Michalak, A. M., & Pahlevan, N. (2019). Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature, 574(7780), 667-670.

Holm-Hansen, O., and B. Riemann. 1978. “Chlorophyll a Determination: Improvements in Methodology.” Oikos 30(3): 438.

Klug, J. L., & Whitney, K. (2015, January 1). Long-term Trends in Water Quality in a New England Hydroelectric Impoundment. Northeastern Naturalist, 22(2), 273 - 272.

Mancuso, J. L., Weinke, A. D., Stone, I. P., Hamsher, S. E., Woller-Skar, M. M., Snyder, E. B., & Biddanda, B. A. (2021). Bloom and bust: Historical trends of harmful algal blooms in Muskegon Lake, Michigan, a Great Lakes estuary. Freshwater Science, 40(3), 463-477.

Mayer, J. et al. (1997). Seasonal successions and trophic relations between phytoplankton, zooplankton, ciliate and bacteria in a hypertrophic shallow lake in Vienna, Austria. Shallow Lakes’ 95: Trophic Cascades in Shallow Freshwater and Brackish Lakes, 165-174.

Pace, M. L., Batt, R. D., Buelo, C. D., Carpenter, S. R., Cole, J. J., Kurtzweil, J. T., & Wilkinson, G. M. (2017). Reversal of a cyanobacterial bloom in response to early warnings. Proceedings of the National Academy of Sciences, 114(2), 352–357

Pace, M. L., Carpenter, S. R., & Wilkinson, G. M. (2019). Long-term studies and reproducibility: Lessons from whole-lake experiments. Limnology and Oceanography, 64(S1), S22-S33.

Paerl, H. W., Fulton, R. S., Moisander, P. H., & Dyble, J. (2001). Harmful freshwater algal blooms, with an emphasis on cyanobacteria. The Scientific World Journal, 1, 76-113.

Paerl, H. W., Hall, N. S., & Calandrino, E. S. (2011). Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. Science of the total environment, 409(10), 1739-1745.

Paerl, H. W., & Otten, T. G. (2013). Harmful cyanobacterial blooms: causes, consequences, and controls. Microbial ecology, 65, 995-1010.

Pickett, S. T. A., Kolasa, J., Armesto, J. J., & Collins, S. L. (1989). The ecological concept of disturbance and its expression at various hierarchical levels. Oikos, 129-136.

Romme, W. H., Everham, E. H., Frelich, L. E., Moritz, M. A., & Sparks, R. E. (1998). Are large, infrequent disturbances qualitatively different from small, frequent disturbances? Ecosystems, 1, 524-534.

Scheffer, M., et al. (2009). Early-warning signals for critical transitions. Nature, 461(7260), 53-59.

Topp, S. N., Pavelsky, T. M., Stanley, E. H., Yang, X., Griffin, C. G., & Ross, M. R. (2021). Multi-decadal improvement in US lake water clarity. Environmental Research Letters, 16(5), 055025.

Turner, M. G., & Dale, V. H. (1998). Comparing large, infrequent disturbances: what have we learned? Ecosystems, 493-496.

Walter, J. A., Buelo, C. D., Besterman, A. F., Tassone, S. J., Atkins, J. W., & Pace, M. L. (2022). An algorithm for detecting and quantifying disturbance and recovery in high-frequency time series. Limnology & Oceanography, Methods, 20(6), 338 - 349.

Wilkinson, G. M., Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R. D., Buelo, C. D., & Kurtzweil, J. T. (2018). Early warning signals precede cyanobacterial blooms in multiple whole-lake experiments. Ecological Monographs, 88(2), 188-203.

Wilkinson, G. M., Walter, J. A., Buelo, C. D., & Pace, M. L. (2022). No evidence of widespread algal bloom intensification in hundreds of lakes. Frontiers in Ecology and the Environment, 20(1), 16-21.

Zamyadi, A., et al. (2012). Cyanobacterial detection using in vivo fluorescence probes: Managing interferences for improved decision-making. Journal-American Water Works Association, 104(8), E466-E479.

Zhang, M., Zhang, Y., Yang, Z., Wei, L., Yang, W., Chen, C., & Kong, F. (2016). Spatial and seasonal shifts in bloom-forming cyanobacteria in Lake Chaohu: Patterns and driving factors. Phycological Research, 64(1), 44-55.

Table 1.1. Steps of applying disturbance recovery algorithm adapted from Walter et al., 2022

Steps
1. Define reference period (entire period), test period (1 day), and test window width (10 days) 2. Get empirical cumulative distribution function for reference period 3. Find mean and standard deviation of the difference statistic for reference period 4. Quantify difference statistic in test period 5. Express the difference statistic during test period as a Z score benchmarked against reference period 6. Assign disturbance (z=2.5) and recovery based on thresholds of Z score

Table 1.2. Table of lakes analyzed for algal bloom disturbance and recovery

Lake	Year	Lat/ Long	Type	Data Source
Peter (Gogebic County, Michigan)	2013-2015, 2019	46.25261, -89.50384	Experimental	https://portal.edirepository.org/nis/mapbrowse?packageid=knb-lter-ntl.413.1
Tuesday (Gogebic County, Michigan)	2013-2015	46.25131, -89.50343	Experimental	
Paul (Gogebic County, Michigan)	2013-2015, 2019	46.25118, -89.49727	Reference for Experimental	
Lillinonah (Fairfield, Connecticut)	2011-2019	41.45943, -73.29963	Monitored	https://portal.edirepository.org/nis/mapbrowse?packageid=edi.559.1 https://portal.edirepository.org/nis/mapbrowse?packageid=edi.560.1
Sammamish (King County, Washington)	2000-2022	47.59931, -122.09603	Monitored	https://green2.kingcounty.gov/lake-buoy/Data.aspx
Seneca (Geneva, NY)	2006-2022	42.80018, -76.95728	Monitored	http://fli-data.hws.edu/buoy/seneca/data.php
Muskegon (Muskegon County, Michigan)	2011-2022	43.23823, -86.28053	Monitored	https://www.gvsu.edu/wri/buoy/data-index.htm
Erie (Toledo buoy near Oregon, Ohio)	2014-2022	41.70203 -83.26145	Monitored	https://uglos.mtu.edu/

Table 1.3. Table of disturbances and recoveries of chlorophyll-a from the experimental lakes. NA indicates a bloom was not observed. In some cases when a bloom was detected no recovery occurred prior to the end of sampling for the year.

Lake	Year	Disturbance #	Peak z-score	Disturbance Length (days)
Peter	2013	NA	NA	NA
Peter	2014	1	4.97	No recovery
Peter	2015	1	28.43	41
Peter	2019	1	38.66	31
Peter	2019	2	12.27	No recovery
Tuesday	2013	1	13.18	No recovery
Tuesday	2014	1	8.13	No recovery
Tuesday	2015	1	4.48	45
Tuesday	2015	2	45.98	No recovery
Paul	2013	NA	NA	NA
Paul	2014	NA	NA	NA
Paul	2015	NA	NA	NA
Paul	2019	1	5.65	No recovery

Table 1.4. Table of disturbances and recoveries of chlorophyll-a from the monitored lakes

Lake	Year	Disturbance #	Peak z-score	Disturbance Length (days)	Variable
Lillinonah	2011	1	2.59	20	Chlorophyll-a
Lillinonah	2012	1	2.79	13	Chlorophyll-a
Lillinonah	2013	1	5.32	20	Chlorophyll-a
Lillinonah	2014	1	6.09	15	Chlorophyll-a
Seneca	2015	1	2.77	14	Chlorophyll-a
Seneca	2015	2	9.81	27	Chlorophyll-a
Seneca	2015	3	5.24	17	Chlorophyll-a
Seneca	2015	4	3.17	NA	Chlorophyll-a
Seneca	2017	1	6.43	12	Chlorophyll-a
Seneca	2017	2	15.10	34	Chlorophyll-a
Seneca	2017	3	3.67	12	Chlorophyll-a
Seneca	2017	4	8.16	26	Chlorophyll-a
Erie (Toledo, OH)	2018	1	6.06	53	Chlorophyll-a
Erie (Toledo, OH)	2018	2	33.18	52	Chlorophyll-a
Muskegon	2012	1	4.85	23	Chlorophyll-a
Muskegon	2014	1	13.39	NA	Chlorophyll-a
Muskegon	2015	1	3.028	15	Chlorophyll-a
Sammamish	2001	1	7.86	63	Chlorophyll-a
Sammamish	2009	1	3.47	26	Chlorophyll-a
Sammamish	2011	1	4.67	34	Chlorophyll-a
Sammamish	2012	1	9.45	71	Chlorophyll-a
Sammamish	2013	1	4.27	16	Chlorophyll-a
Sammamish	2013	2	2.62	7	Chlorophyll-a
Sammamish	2015	1	5.01	52	Chlorophyll-a
Sammamish	2015	2	6.33	33	Chlorophyll-a
Sammamish	2015	3	2.99	9	Chlorophyll-a
Sammamish	2017	1	2.52	17	Chlorophyll-a
Sammamish	2018	1	2.94	20	Chlorophyll-a
Sammamish	2019	1	3.01	10	Chlorophyll-a
Sammamish	2022	1	6.04	14	Chlorophyll-a
Lillinonah	2014	1	10.82	71	Phycocyanin
Lillinonah	2015	1	3.23	36	Phycocyanin
Erie (Toledo, OH)	2015	1	14.97	46	Phycocyanin
Erie (Toledo, OH)	2017	1	5.22	10	Phycocyanin
Muskegon	2011	1	3.17	38	Phycocyanin
Muskegon	2012	1	4.54	17	Phycocyanin
Muskegon	2012	2	4.85	NA	Phycocyanin

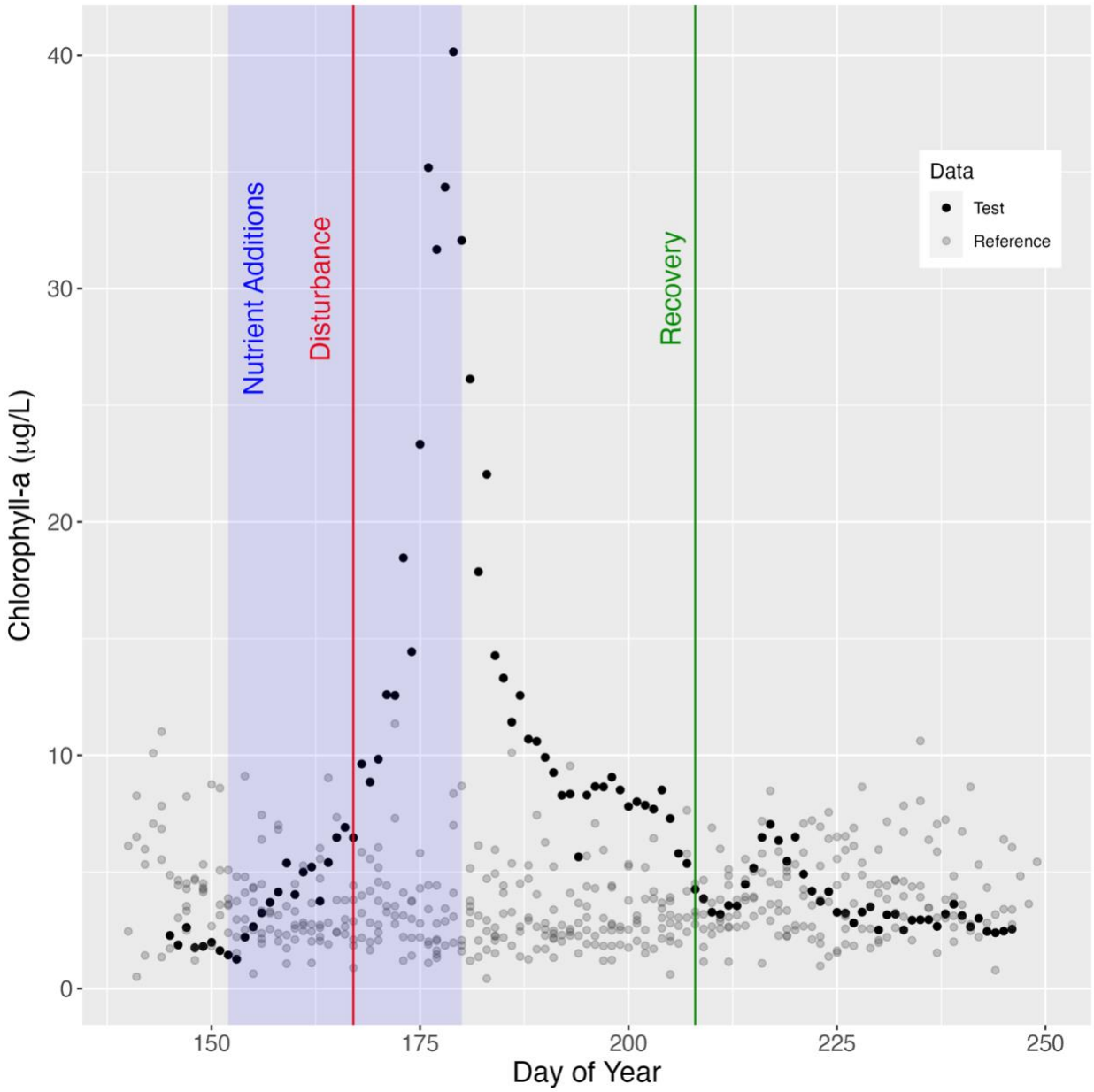

Figure 1.1. Disturbance and recovery of Peter Lake in 2015. Black points are the test data from 2015 and open gray circles are reference data from many prior years when the lake was not fertilized. Blue shaded area indicates period of nutrient addition. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

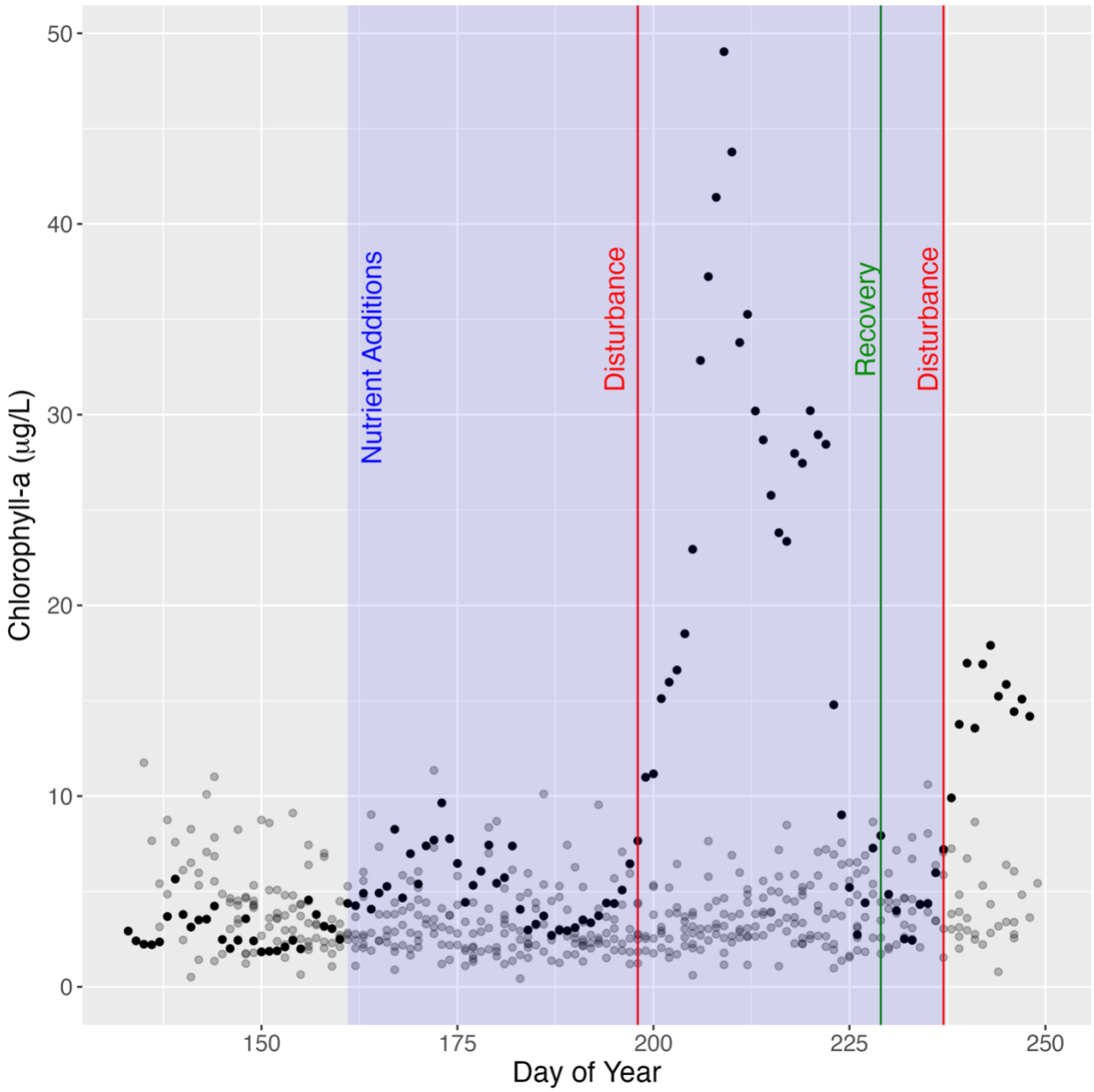

Figure 1.2. Disturbances and recoveries of Peter Lake in 2019. Blue shaded area indicates nutrient addition period. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

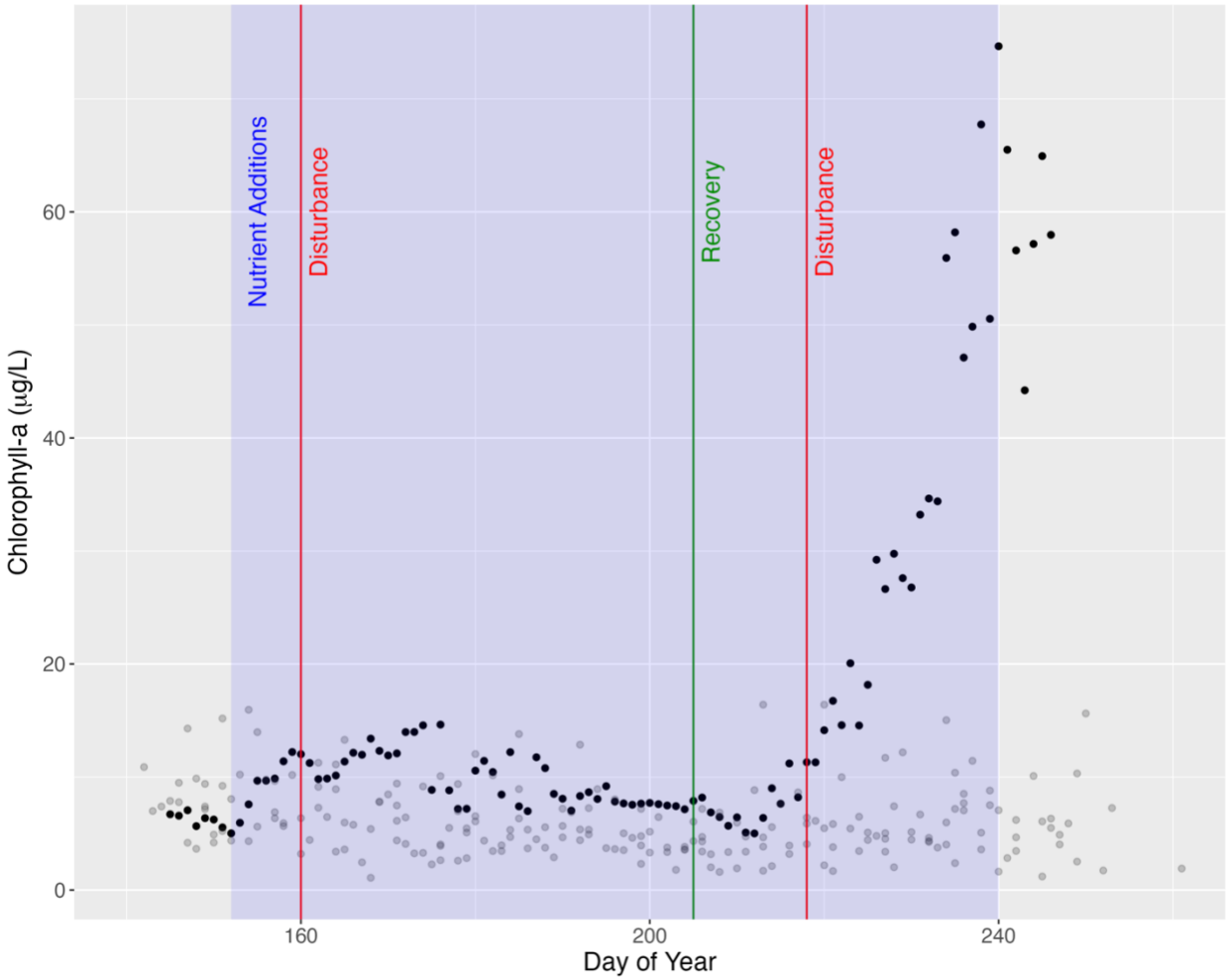

Figure 1.3. Disturbance and recovery of Tuesday Lake in 2015. Black points are the test data and open gray circles are reference data from many prior years when the lake was not fertilized. Blue shaded area indicates period of nutrient addition. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

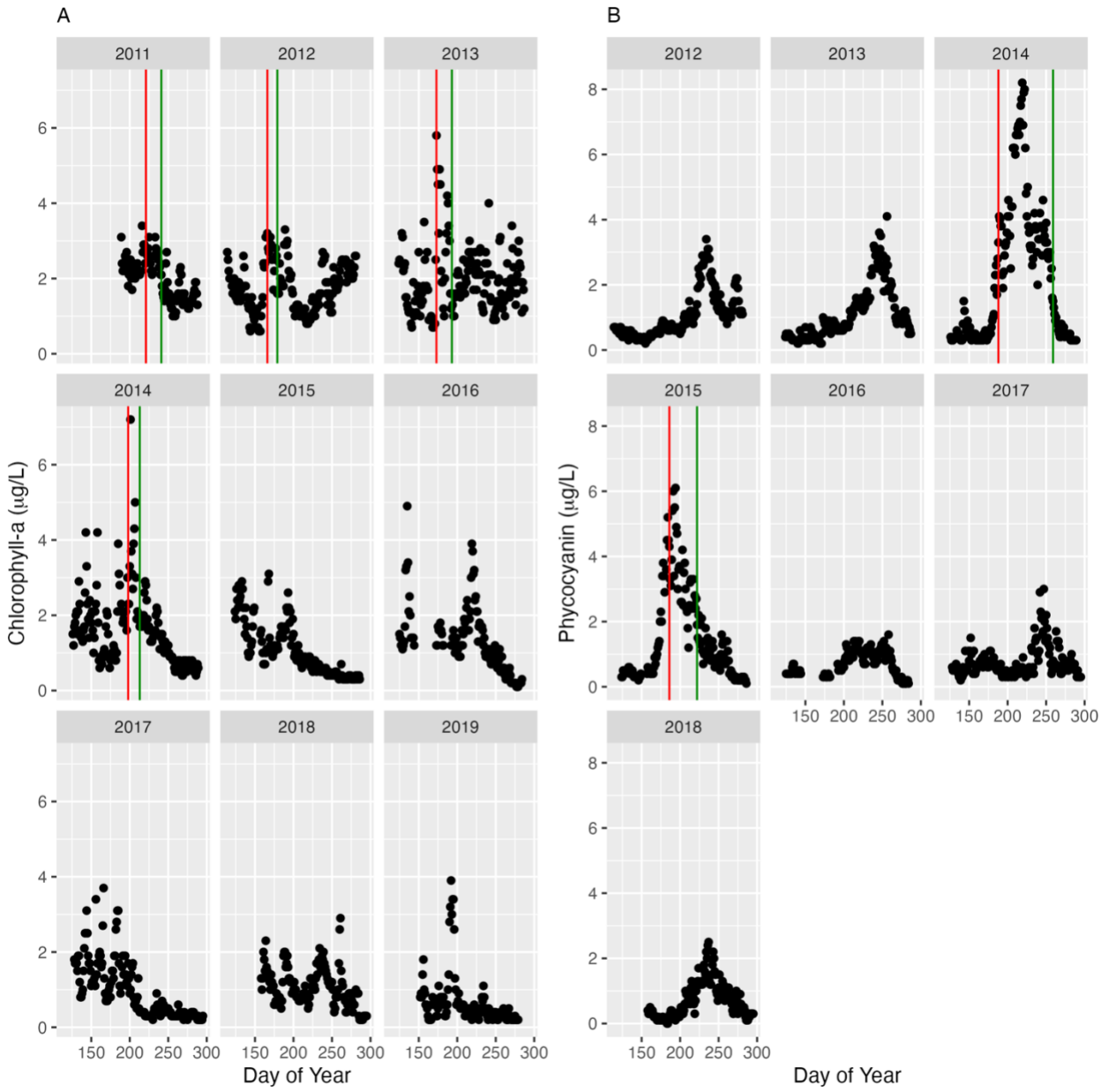

Figure 1.4. Disturbances and recoveries of Lake Lillinonah's a) chlorophyll-a and b) phycocyanin. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

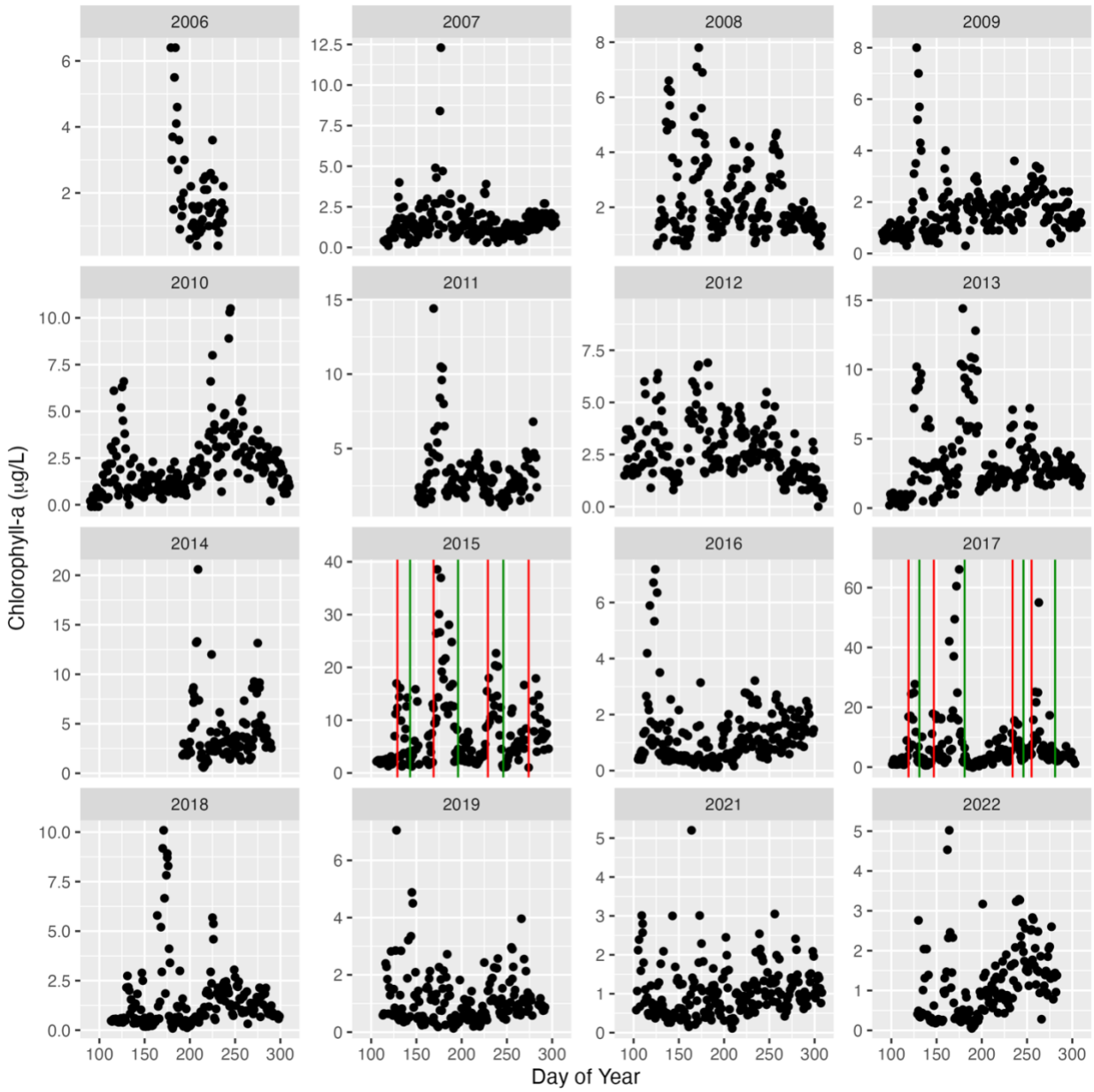

Figure 1.5. Disturbances and recoveries of chlorophyll-a in Seneca Lake. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

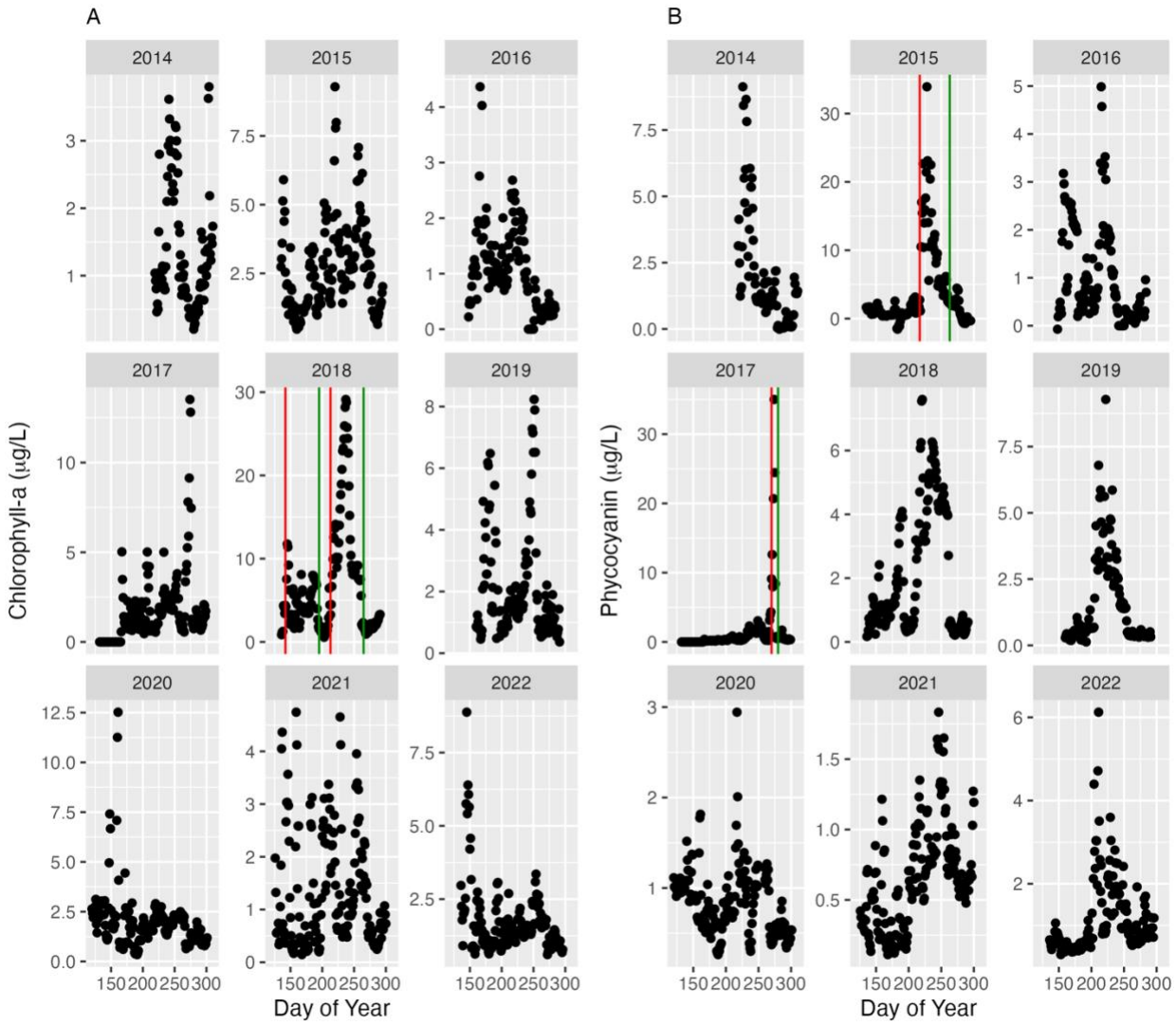

Figure 1.6. Disturbances and recoveries of Lake Erie's (Toledo buoy) a) chlorophyll-a and b) phycocyanin. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

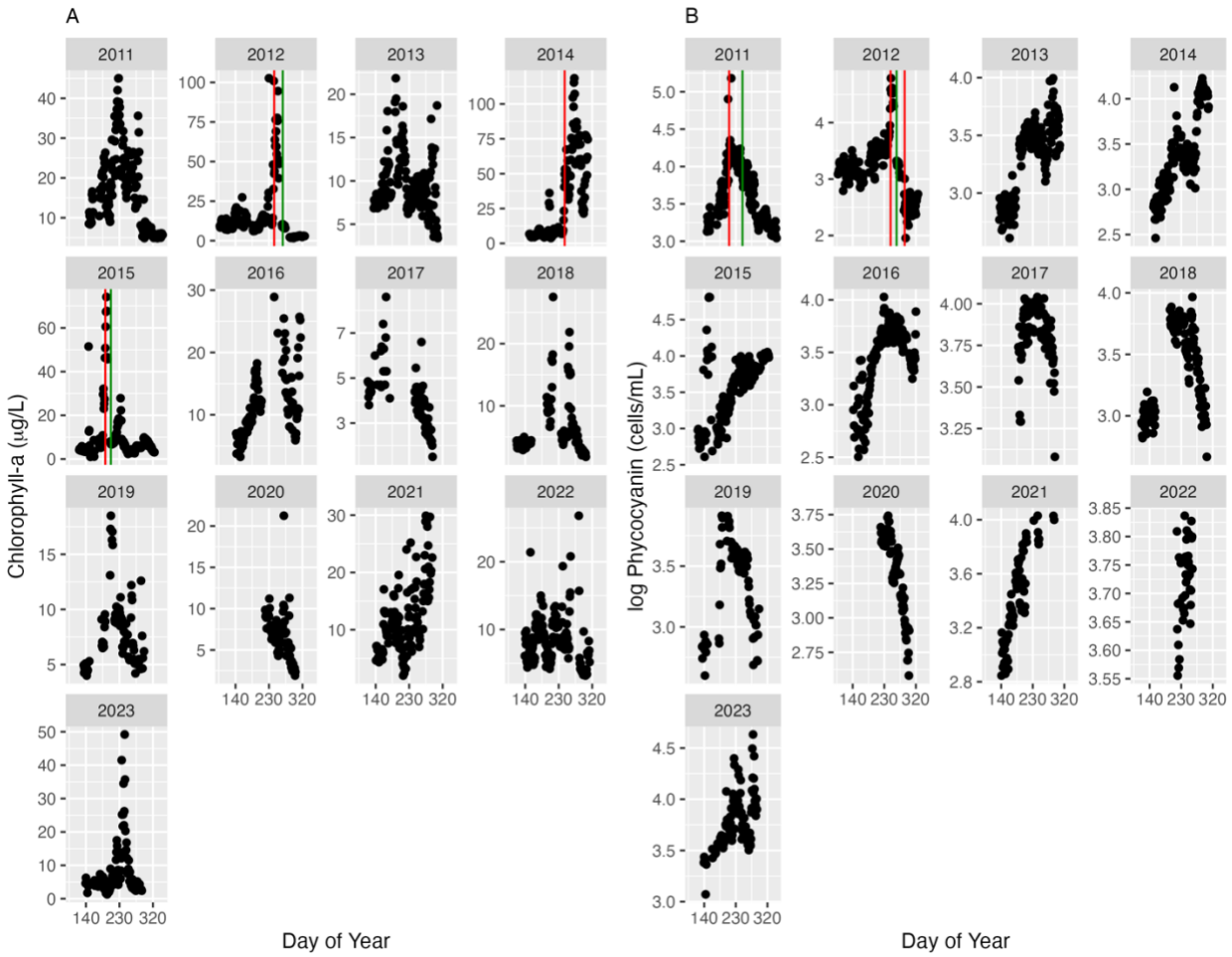

Figure 1.7. Disturbances and recoveries of Muskegon Lake's a) chlorophyll-a and b) phycocyanin. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

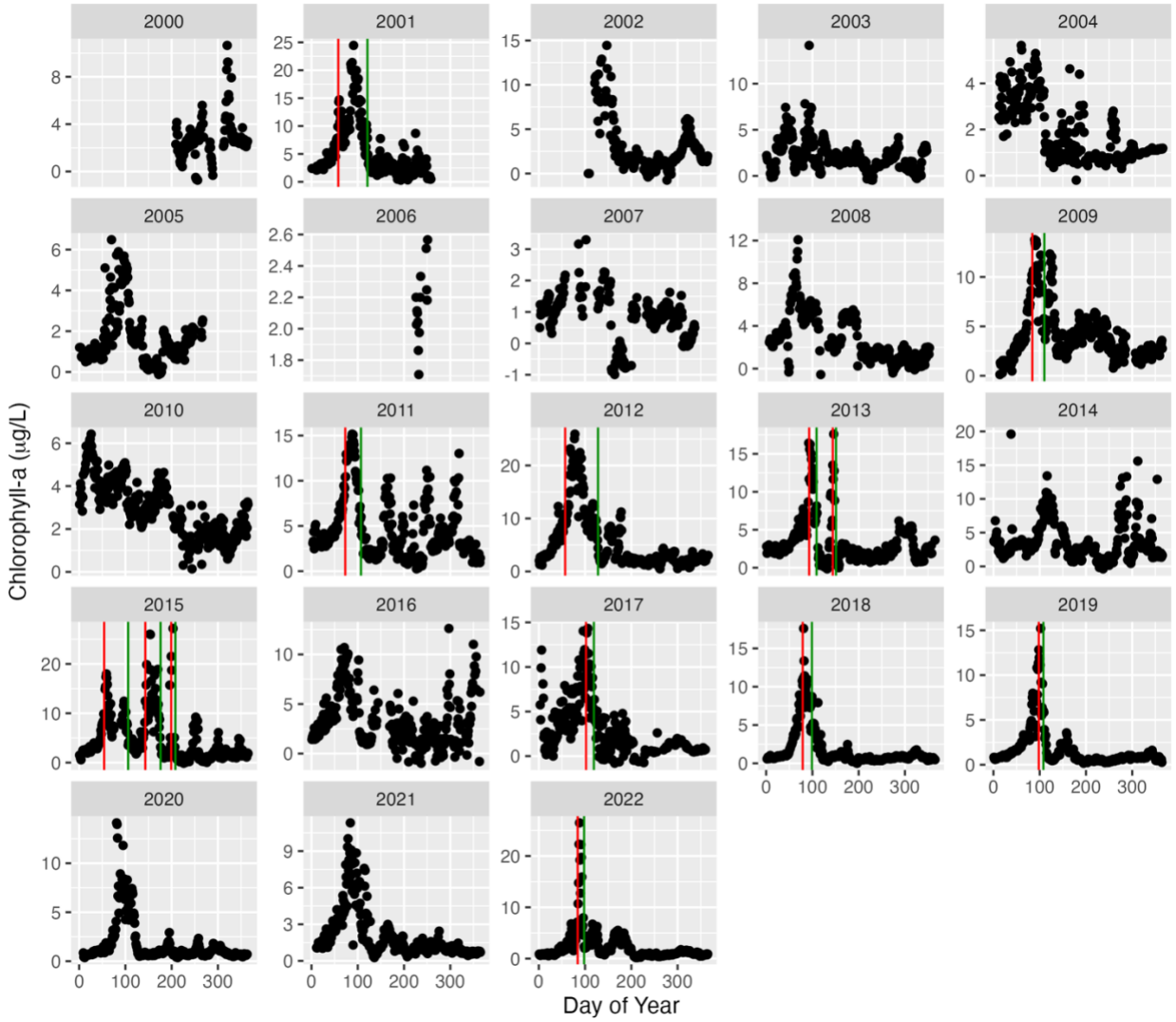

Figure 1.8. Disturbances and recoveries of Lake Sammamish's chlorophyll-a. Red vertical lines indicate disturbances and green vertical lines indicate recovery.

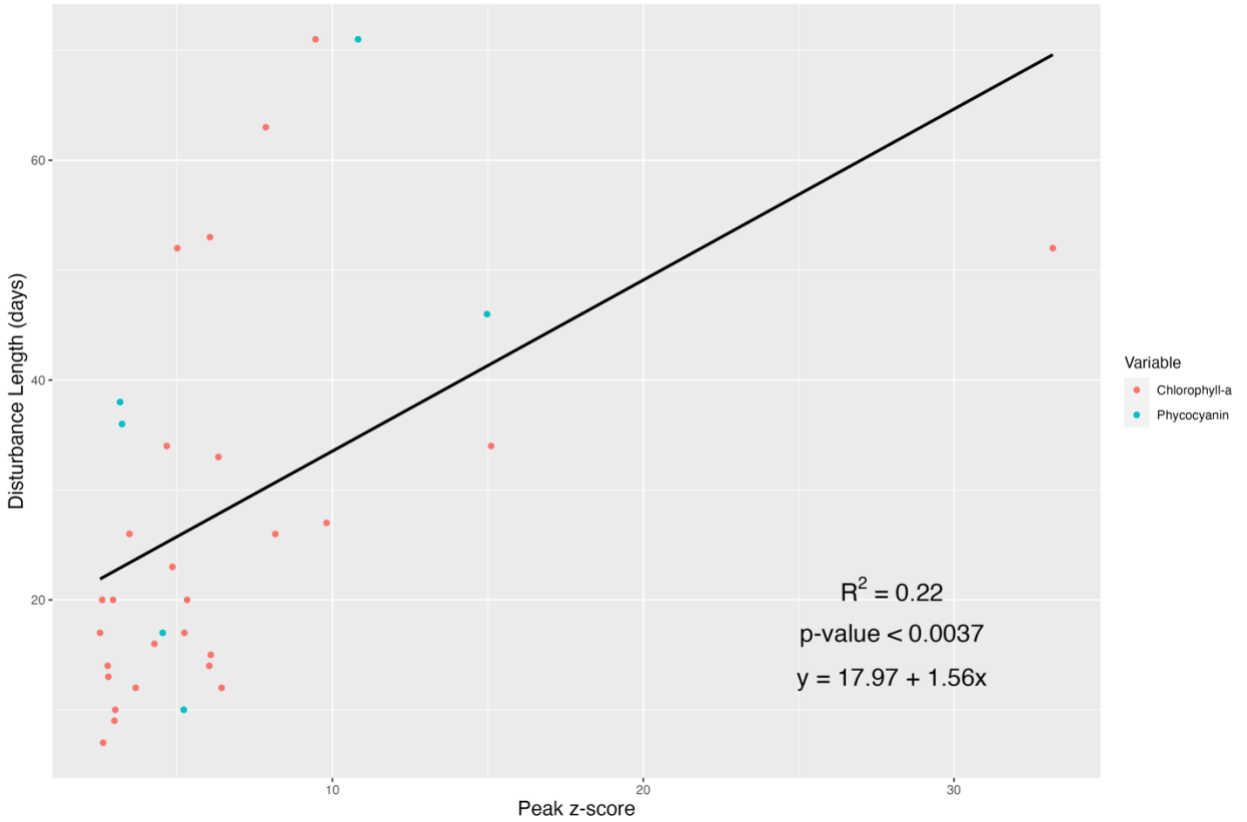

Figure 1.9. Relationship between disturbance length and peak disturbance severity for chlorophyll-a and phycocyanin from the monitored lakes.

Chapter II. Identifying spatial variability in experimental and reference lakes

Abstract

Spatial heterogeneity in lakes is often caused by physical, biological, and chemical drivers. Larger lakes exhibit spatial heterogeneity and are often studied for their localized algal blooms, greenhouse gas emissions, or biological communities. However, spatial heterogeneity in small waterbodies is seldom studied. I analyzed high resolution spatial data in two small lakes (< 3 ha), including a nutrient enriched experimental lake and an adjacent unmanipulated lake, to better understand spatial dynamics of small water bodies and a lake undergoing an algal bloom. Repeated spatial surveys were conducted where temperature, nitrate, dissolved oxygen, pH, and phycocyanin were measured over thousands of locations in each lake. The periods of spatial surveys also included phytoplankton blooms in the experimental lake. Spatial coefficients of variation were low (<5.7%) for all lake-year-variable combinations. Moran's I, a measure of spatial autocorrelation, ranged from 0-0.7 and was similar between the bloom and non-bloom periods. Spatial autocorrelation values were typically significant but overall clustering was low due to low spatial variability. A hotspot analysis based on applying the Getis Ord GI* method indicated that hot and cold spots were random across both lakes and there was no agreement between hot/cold spots with dominant wind direction. There was no difference between spatial variability for the lake undergoing an algal bloom when comparing bloom and non-bloom periods. Additionally, there was no differences in spatial patterns between the experimental and reference lakes. Overall, the results indicate that small lakes exhibit low spatial heterogeneity and single point sampling at the center of the lake is sufficient.

Introduction

Lakes of varying sizes and shapes are spatially heterogeneous for a variety of variables including, for example, greenhouse gases and common water quality parameters (Schilder et al., 2013, Loken et al. 2019, Kislik et al., 2022, Ortiz & Wilkinson, 2021, Hou et al. 2022). Inlet nutrient contributions, physical processes (e.g., wind direction), morphology, and macrophytes contribute to creating spatial heterogeneity in lakes. Despite this known heterogeneity, most lake studies are based on measurements taken at the center of the lake or fixed points and extrapolated across the entire area of the waterbody (Stanley et al., 2019). Central and limited spatial measurements taken from lakes with large spatial heterogeneity can produce inaccurate representation of the system. Spatially distributed sampling can overcome this problem and provide more accurate estimates for standing stocks and fluxes as well as help identify hotspots or areas of concern.

Algal blooms can be patchy and spatially heterogeneous across a lake (Buelo et al., 2018, Serizawa et al., 2008). The availability of nutrients, light, and other growth-promoting conditions may vary for different parts of the water body. Additionally, the movement of water, via currents or turbulence, can influence the spatial distribution of algae.

Spatial heterogeneity related to phytoplankton blooms is most obvious and most studied in larger lakes like Lake Erie where satellite remote sensing clearly indicates variability in phytoplankton pigments (Wynne et al., 2015). Spatial heterogeneity is less commonly studied in smaller lakes. Nevertheless, previous research documented spatial heterogeneity in chlorophyll-a, water temperature, and dissolved oxygen in a moderate-size (~113 ha) temperate lake (Mackay et al., 2011) and ecosystem metabolism in small (< 5 ha) lakes (Van de

Bogert et al., 2012, Van de Bogert et al., 2007). This poses the question if there is significant and persistent spatial heterogeneity and variability in smaller lakes amongst other variables.

Previous research studying the same small lake considered by Van de Bogert et al. (2012) tested the use of spatial early warning statistics as possible indicators of pending algal blooms (Butitta et al. 2017, Buelo et al., 2022). Using the high resolution spatial-temporal data for Peter and Paul lakes from the Butitta et al. (2017) and Buelo et al. (2022) studies, I analyzed spatial variability to address additional questions (beyond those related to early warning) including:

1) Does the spatial coefficient of variation (CV) increase over a bloom cycle due to greater patchiness in phytoplankton?
2) Is there persistent spatial structure in an experimental lake undergoing an algal bloom and a reference lake experiencing no disturbance?
3) How does spatial variability compare to temporal variability?

These questions were addressed using data from years both with and without whole lake nutrient additions for an experimental and reference lake. In each year an extensive spatial sampling campaign was conducted.

Methods

The lake studies were done at the University of Notre Dame Environmental Research Center, located in the Upper Peninsula of Michigan, United States. Peter and Paul lakes are relatively small (2.6 ha for Peter Lake and 1.7 ha for Paul Lake) with bowl-shaped basins (mean depth 5.7 and 3.7 m, maximum depth approximately 18 and 12 m, respectively). Paul Lake is

upstream of Peter Lake and has been consistently used as a non-manipulated reference system, while Peter Lake have been subject to manipulations (Carpenter and Pace 2018). Peter and Paul Lakes are fed by groundwater flows and lack significant inputs and outputs, resulting in having long water residence times (Cole et al., 1998).

Spatial data were collected using the FLAMe system (Fast Limnological Automated Measurements; Crawford et al., 2015). The FLAMe has a flow-through design to collect spatial measures of variables using sensors. While the boat is in motion, surface lake water is drawn through an intake and passed through a sensor array, while also concurrently capturing GPS location data. Spatial variables included phycocyanin ($\mu g\ L^{-1}$), dissolved oxygen (D.O.) saturation (%), pH, temperature, and nitrate. Phycocyanin, DO, pH, and temperature were measured with a YSI, Inc. EXO 2 sonde and nitrate was measured by a SUNA nitrate sensor. Spatial resolution for the YSI-measured variables was < 3 m while nitrate was measured at a courser resolution of ~10 m. Data were collected in the summers of 2015, 2018, and 2019. Spatial sampling was conducted from late May to early September with weekly observations in 2015 and 2018, and three samplings per week (Monday, Wednesday, Friday) in 2019.

In 2015 and 2019, Peter Lake was fertilized with inorganic nitrogen and phosphorus to induce an algal bloom. Solutions of phosphoric acid and ammonium nitrate were prepared and distributed uniformly across the surface of the lake by pumping them into the prop-wash of a boat propelled by an electric motor. Nutrient additions started on day of year (DOY) 152 in 2015 and DOY 161 in 2019 (Pace et al., 2017, Buelo et al., 2022). The nutrient additions ended on DOY 180 in 2015 and DOY 237 in 2019. To minimize disturbance from trolling motors on the

lake, FLAMe sampling was conducted in the morning, and nutrient additions were carried out in the afternoon.

To examine spatial variability and hotspots of the lakes, I calculated the spatial coefficient of variation, spatial autocorrelation (Moran's I), and the Getis-Ord Gi* statistic for temperature, nitrate, DO, pH, and phycocyanin. Spatial coefficient of variation was calculated by formula: (Standard Deviation / Mean) * 100 for all the observations collected in a sampling event. Temporal variability was evaluated using data collected with sondes in the center of the lakes that were averaged to daily values (to remove diel variability). Additionally, wind direction and speed were collected with a local HOBO weather station located at the center of Peter Lake.

Moran's I, a measure of spatial autocorrelation, was calculated in R using the *moranfast* package based on equation (1) below. The Moran's I statistic (I) is the global spatial association between locations and how the values at one location relate to the values at other locations, normalized to constrain values between -1 and 1, where N is the total number of observations, W is the spatial weights sum (inverse distance weighting in this case), x_i and x_j are the values of the variable at location *i* and *j*, respectively, $\bar{x}$ is the mean of the variable, and w_{ij} represents the spatial weight between *i* and *j*.

$$(1) \qquad I = \frac{N}{W} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

The Getis-Ord Gi* statistic is a hot spot and cold spot analysis and indicates high and low clusters of values, based on equation (2) below, where X_j is the value for feature *j*, $w_{i,j}$ is the

spatial weight between *i* and *j*, n is the total number of features, $\bar{x}$ is the mean of the variable, and S represents the total spatial aggregation. This statistic was calculated in ArcGIS Pro using an inverse distance weighting and the mean distance threshold calculated from the Incremental Spatial Autocorrelation tool from all the sampling events for each variable. The Incremental Spatial Autocorrelation tool plots Moran's I against distance and the distance where spatial autocorrelation was the highest was selected from each sampling event-variable combination. The distance thresholds for temperature, nitrate, dissolved oxygen, pH, and phycocyanin were: 50m, 55m, 60m, 65m, and 40m, respectively.

$$(2) \qquad G_i^* = \frac{\sum_{j=1}^{n} w_{ij}x_j - \bar{x}\sum_{j=1}^{n} w_{ij}}{S\sqrt{\frac{n\sum_{j=1}^{n} w_{ij}^2 - (\sum_{j=1}^{n} w_{ij})^2}{n-1}}}$$

Results

Responses to Nutrient Addition

Peter Lake had blooms as measured by chlorophyll-a in 2015 and 2019 due to nutrient additions (Figure 2.1). In 2015, there was a brief period of nutrient addition which caused one bloom followed by a decline back to baseline after nutrient additions halted (Pace et al. 2017). In 2018 when no nutrients were added, Peter Lake had low concentrations of chlorophyll-a (Figure 2.1b). In 2019, Peter Lake underwent continuous nutrient addition which caused a bloom that collapsed and then a secondary bloom formed (Figure 2.1c). Paul Lake had low levels of chlorophyll-a for all three years (Figure 2.1). Additionally based on daily measures at a central station, Peter Lake had greater concentrations of DO, higher pH, and greater

phycocyanin during bloom periods compared to Paul Lake in 2015 and 2019 (Wilkinson et al. 2018; Buelo et al. 2022).

Spatial versus Temporal Variability

During the bloom periods, concentrations of DO, pH, and phycocyanin determined from spatial surveys were greater compared to non-bloom periods (Figure 2.2). Additionally, concentrations were greater in Peter than Paul Lake. Peter and Paul Lakes had similar values and showed similar patterns in all variables during 2018, and dynamics of temperature were the same in both lakes in all three years. Spatial variability was low in the lakes and not substantially changed by nutrient additions. Peter Lake had similar coefficients of variation (hereafter CVs) as Paul Lake for the variables considered with the exception of phycocyanin (Figure 2.3). Phycocyanin CVs were problematic due to low and negative raw values during non-bloom periods. The data indicated close to no cyanobacteria in the lakes and little spatial variability (Figure 2.2). Reliable CVs for phycocyanin could only be calculated during blooms.

Peter Lake had similar CVs during the non-bloom periods compared to bloom periods of 2015 and 2019, as well as years with blooms and the year without nutrient addition. The spatial CVs were low and ranged from 0-5.7% across all variables except for phycocyanin. A specific example is the comparison of mean CVs in Peter and Paul lakes for the year 2015 where values for temperature, DO, and pH were, respectively: 0.56% compared to 0.72%, 1.41% compared to 0.82%, and 1.74% compared to 0.63%.

Temporal and spatial CVs reflect different aspects of variability and are based on different sets of measurements. Temporal CVs reflect daily variability for each year as

influenced by seasonality, weather, and blooms when nutrients were added. The spatial CVs represent the means of a series of average CVs from the spatial surveys for each year. The two measures represent different scales, but comparison is useful in indicating the importance of time versus space in influencing variation. The temporal CVs were greater than spatial CVs for temperature, DO, and pH (Table 1). Peter and Paul Lakes had similar CVs during 2018, the non-nutrient addition year, except for temperature. pH had the greatest difference among temporal CVs between bloom and non-bloom years reflecting the effect of elevated primary production on pH in these softwater lakes.

Spatial Autocorrelation in Bloom versus Non-Bloom Periods

Here, I focused on data from 2019 because that year had the most spatial sampling events and two bloom periods (Figure 2.1). Temperature followed a seasonal dynamic in both Peter and Paul lakes in 2019 (Figure 2.4). Nitrate was low in both lakes but rose in Peter Lake as nutrient addition continued indicating incomplete consumption of this nutrient by phytoplankton (Figure 2.4). Oxygen and pH were elevated in Peter Lake during the first bloom, fell below values in Paul Lake when the bloom collapsed, and increased above Paul when the second bloom developed at the end of the sampling season (Figure 2.4).

In contrast with the differences observed due to blooms, Peter and Paul Lakes had similar values of spatial autocorrelation during the non-bloom and bloom periods (Figure 2.4). All mean Moran's I values for each lake-variable combination were statistically significant. Peter Lake had a mean Moran's I value of 0.28, 0.12, 0.25, 0.24, 0.15 for temperature, nitrate, DO, pH, and phycocyanin, respectively. Paul Lake had a mean Moran's I value of 0.30, 0.11, 0.29,

0.29, 0.15 for temperature, nitrate, DO, pH, and phycocyanin, respectively. Among all the variables, nitrate had the lowest Moran's I values with the other variables having similar ranges (Figure 2.4).

Hotspot Analysis

There was no long-lasting spatial structure in Peter and Paul Lakes based on the Getis Ord GI* analysis. Peter Lake had no obvious difference in spatial structure before, during, or after bloom in 2019 (Figure 2.5). The Getis-Ord Gi* hotspot analysis did indicate statistically significant hotspots and cold spots of dissolved oxygen, but the hotspots and cold spots changed each sampling date. Before the bloom, the DO hotspots were on the east side of the lake (Figure 2.5a). The hotspots shifted to the southeast during the bloom (Figure 2.5b). Lastly when the bloom collapsed, there were minor hotspots located in the middle and northeast side of the lake (Figure 2.5c). Similarly, Paul Lake had random hot spots and cold spots that shifted throughout the summer without evidence of persistence (Figure 2.5).

Discussion

The variables dissolved oxygen, pH, and phycocyanin had a greater range in the spatial FLAMe surveys during the 2015 and 2019 bloom periods relative to other times. There were 2 spatial sampling events during the 2019 bloom where dissolved oxygen saturation was both above and below 100% saturation, indicating that there were points in the lake where oxygen was super- and under-saturated. Despite these larger and expected ranges in variables during bloom periods, spatial CVs were low (< 10%) across all lake-year-variable combinations, even

during the bloom periods. Overall, there is low spatial variability in Peter and Paul Lakes. By comparison, the temporal CVs coefficient of variation for temperature, DO, and pH were greater than the spatial CVs for all years considered.

All Moran's I correlations were positive and statistically significant indicating clustering of values in space but because of the small magnitude of variation, spatial structure was relatively weak. I expected the least amount of spatial structure for temperature because temperature is a physical variable. I expected greater spatial variability for biologically influenced variables such as pH, DO (% saturation), and phycocyanin. There were a few sampling events where there were peaks in Moran's I, such as DOY 170 for Peter Lake in 2015. In this case, Moran's I was 0.70 for phycocyanin as the bloom was developing (i.e., rapid increase in cyanobacteria), but the value dropped to 0.18 and 0.24 for the immediately following sampling dates that occurred during the bloom (Figure 2.1a, Figure 2.4). The noisy and fluctuating Moran's I values suggests that the spatial autocorrelation patterns were random and biologically influenced variables did not tend to have greater spatial structure than a physical variable like temperature.

The hot and cold spots moved around the lake for each variable and sampling event, with no obvious pattern. Dominant wind direction did not match the hot or cold spots. Additionally, the Getis Ord GI* hotspot analysis is extremely sensitive to the distance threshold selected and can significantly skew the results. The distance thresholds used for this analysis ranged from 40-65 meters.

The lack of spatial heterogeneity in Peter and Paul Lakes could be attributed to fast mixing. These lakes are small, bowl shaped, lack extensive macrophytes. These factors

contribute to the low heterogeneity. Overall, the results indicate mechanisms that create structure were slow relative to the mechanisms that homogenize structure in these lakes. In other lakes, spatial structure could arise from recruitment of bloom forming algae in the littoral, persistent directional winds, morphometric features that promote differential growth, or heterogeneity in nutrient concentrations. In Peter and Paul lakes, rapid mixing and apparent rapid spatially uniform algal growth (based on oxygen and pH dynamics) create homogeneity. Additionally, the effort to distribute nutrients uniformly as a methodology of the whole lake manipulations probably contributed to the lack of spatial heterogeneity as observed for nitrate. In natural systems, nutrients from the watershed often enter the system via tributaries (Rast et al., 1983, Robertson et al., 2011, Mooney et al., 2020) and act like point-sources contributing to spatial pattern (Loken et al., 2019).

Van de Bogert et al. (2012) measured large spatial heterogeneity in ecosystem metabolism in Peter Lake. Calculating ecosystem metabolism involves high frequency diel measurements and there is likely considerable measurement error related to vertical and horizontal oxygen dynamics. Nevertheless, the dissolved oxygen concentrations were remarkably uniform spatially in the current study. Thus, the large variation in spatial metabolism rates is not easily explained from the spatial data of this study. Surveys done at one time during the morning might be insufficient to capture dynamics affecting oxygen concentrations over a 24-hour cycle.

Some studies of small lakes have observed low spatial variation while in others spatial heterogeneity was significant. Ray et al. (2023) found low spatial heterogeneity of dissolved carbon dioxide and methane in ponds and shallow lakes (depths <5m). A single sample for

these variables is representative of the entire waterbody (Ray et al., 2023). Conversely, Ortiz & Wilkinson (2021) found large spatial heterogeneity in chlorophyll-a, phycocyanin, dissolved oxygen, and pH, and the greatest variability during an algal bloom in a shallow lake, specifically where there were macrophytes. Given differences among waterbodies, it is important to consider the factors that might create spatial heterogeneity in any given case. Based on the results from this study, small, spatially simple lakes exhibit low spatial heterogeneity in temperature, nitrate, dissolved oxygen, pH, and phycocyanin and for many purposes spatial variation could be ignored.

The increasing capability of satellite and drone remote sensing for studying lakes should aid future research on spatial heterogeneity. Remote images can identify variation in certain features like phytoplankton pigments and complement field measurements to indicate the degree of spatial variability of lakes (Kislik et al., 2022, Fernandez-Figueroa et al., 2022, Zhou et al., 2023). Remote sensing and spatial surveys can complement one another to provide a better understanding of spatial heterogeneity (Powers et al., 2023). The degree to which a lake expresses spatial heterogeneity will determine the potential value of spatial information from in situ sampling using measurement approaches like FLAMe and satellites remote sensing. Most lakes are small (Downing et al., 1984) similar to those considered in this study. If the patterns observed in this study generalize to other small lakes, spatial variation is low and likely can be ignored for most purposes.

References

Buelo, C. D., Carpenter, S. R., & Pace, M. L. (2018). A modeling analysis of spatial statistical indicators of thresholds for algal blooms. Limnology and Oceanography Letters, 3(5), 384– 392.

Buelo, C. D., Pace, M. L., Carpenter, S. R., Stanley, E. H., Ortiz, D. A., & Ha, D. T. (2022). Evaluating the performance of temporal and spatial early warning statistics of algal blooms. Ecological Applications, 32(5), 1 - 15.

Butitta, V. L., S. R. Carpenter, L. C. Loken, M. L. Pace, and E. H. Stanley. 2017. "Spatial Early Warning Signals in a Lake Manipulation." Ecosphere 8(10): e01941.

Carpenter, S. R., & Pace, M. L. (2018). Synthesis of a 33-yr series of whole-lake experiments: Effects of nutrients, grazers, and precipitation-driven water color on chlorophyll. Limnology and Oceanography Letters, 3(6), 419-427.

Cole, J. J., & Pace, M. L. (1998). Hydrologic variability of small, northern Michigan lakes measured by the addition of tracers. Ecosystems, 1, 310-320.

Crawford, J. T., L. C. Loken, N. J. Casson, C. Smith, A. G. Stone, and L. A. Winslow. 2015. "High-Speed Limnology: Using Advanced Sensors to Investigate Spatial Variability in Biogeochemistry and Hydrology." Environmental Science & Technology 49(1): 442– 50.

Downing, J. A., Prairie, Y. T., Cole, J. J., Duarte, C. M., Tranvik, L. J., Striegl, R. G., ... & Middelburg, J. J. (2006). The global abundance and size distribution of lakes, ponds, and impoundments. Limnology and oceanography, 51(5), 2388-2397.

Fernandez-Figueroa, E. G., Wilson, A. E., & Rogers, S. R. (2022). Commercially available unoccupied aerial systems for monitoring harmful algal blooms: A comparative study. Limnology and Oceanography: Methods, 20(3), 146-158.

Hou, P., Chang, F., Duan, L., Zhang, Y., & Zhang, H. (2022). Seasonal Variation and Spatial Heterogeneity of Water Quality Parameters in Lake Chenghai in Southwestern China. Water (20734441), 14(10), 1640 - 1655.

Kislik, C., Dronova, I., Grantham, T. E., & Kelly, M. (2022). Mapping algal bloom dynamics in small reservoirs using Sentinel-2 imagery in Google Earth Engine. Ecological Indicators, 140.

Loken, L. C., Crawford, J. T., Schramm, P. J., Stadler, P., Desai, A. R., & Stanley, E. H. (2019). Large spatial and temporal variability of carbon dioxide and methane in a eutrophic lake. Journal of Geophysical Research-Biogeosciences, 124(7), 2248– 2266.

Mackay, E. B., Jones, I. D., Thackeray, S. J., & Folkard, A. M. (2011). Spatial heterogeneity in a small, temperate lake during archetypal weak forcing conditions. Fundamental and applied limnology, 179(1), 27-40.

Mooney, R. J., Stanley, E. H., Rosenthal, W. C., Esselman, P. C., Kendall, A. D., & McIntyre, P. B. (2020). Outsized nutrient contributions from small tributaries to a Great Lake. Proceedings of the National Academy of Sciences, 117(45), 28175-28182.

Ortiz, D. A., & Wilkinson, G. M. (2021). Capturing the spatial variability of algal bloom development in a shallow temperate lake. Freshwater Biology, 66(11), 2064 - 2075.

Pace, M. L., Batt, R. D., Buelo, C. D., Carpenter, S. R., Cole, J. J., Kurtzweil, J. T., & Wilkinson, G. M. (2017). Reversal of a cyanobacterial bloom in response to early warnings. Proceedings of the National Academy of Sciences, 114(2), 352-357.

Powers, S. M., Barnard, M. A., Macleod, M. S., Miller, L. A., & Wagner, N. D. (2023). Spatially Intensive Patterns of Water Clarity in Reservoirs Determined Rapidly With Sensor-Equipped Boats and Satellites. Journal of Geophysical Research: Biogeosciences, 128(10), e2023JG007650.

Ray, N. E., et al. (2023). Spatial and temporal variability in summertime dissolved carbon dioxide and methane in temperate ponds and shallow lakes. Limnology and Oceanography.

Rast, W., & Lee, G. F. (1983). Nutrient loading estimates for lakes. Journal of Environmental Engineering, 109(2), 502-517.

Robertson, D. M., & Saad, D. A. (2011). Nutrient inputs to the Laurentian Great Lakes by source and watershed estimated using SPARROW watershed models 1. JAWRA Journal of the American Water Resources Association, 47(5), 1011-1033.

Schilder, J., Bastviken, D., Van Hardenbroek, M., Kankaala, P., Rinta, P., Stoetter, T., & Heiri, O. (2013). Spatial heterogeneity and lake morphology affect diffusive greenhouse gas emission estimates of lakes. Geophysical Research Letters, 40(21), 5752 - 5756.

Serizawa, H., Amemiya, T., & Itoh, K. (2008). Patchiness in a minimal nutrient—Phytoplankton model. Journal of Biosciences, 33(3), 391– 403.

Stanley, E. H., Collins, S. M., Lottig, N. R., Oliver, S. K., Webster, K. E., Cheruvelil, K. S., & Soranno, P. A. (2019). Biases in lake water quality sampling and implications for macroscale research. Limnology and Oceanography, 64(4), 1572– 1585.

Sukenik, A., & Kaplan, A. (2021). Cyanobacterial harmful algal blooms in aquatic ecosystems: A comprehensive outlook on current and emerging mitigation and control approaches. Microorganisms, 9(7), 1472.

Van de Bogert, M. C., Carpenter, S. R., Cole, J. J., & Pace, M. L. (2007). Assessing pelagic and benthic metabolism using free water measurements. Limnology and Oceanography: Methods, 5(5), 145-155.

Van de Bogert, M. C., Bade, D. L., Carpenter, S. R., Cole, J. J., Pace, M. L., Hanson, P. C., & Langman, O. C. (2012). Spatial heterogeneity strongly affects estimates of ecosystem metabolism in two north temperate lakes. Limnology and Oceanography, 57(6), 1689– 1700.

Weathers, K. C., et al. (2016). Frontiers in ecosystem ecology from a community perspective: The future is boundless and bright. Ecosystems, 19, 753-770.

Wilkinson, G. M., Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R. D., Buelo, C. D., & Kurtzweil, J. T. (2018). Early warning signals precede cyanobacterial blooms in multiple whole-lake experiments. Ecological Monographs, 88(2), 188-203.

Wynne, T. T., & Stumpf, R. P. (2015). Spatial and temporal patterns in the seasonal distribution of toxic cyanobacteria in western Lake Erie from 2002–2014. Toxins, 7(5), 1649-1663.

Zhou, T., Li, Y., Jiang, B., Alatalo, J. M., Li, C., & Ni, C. (2023). Tracking spatio-temporal dynamics of harmful algal blooms using long-term MODIS observations of Chaohu Lake in China from 2000 to 2021. Ecological Indicators, 146.

Table 2.1 Temporal versus spatial coefficients of variation. Coefficient of variation is reported in percentage.

Lake	Year	Temperature		DO		pH	
		Temporal	Spatial	Temporal	Spatial	Temporal	Spatial
Peter	2015	15.2	0.56	8.62	1.41	13.4	1.74
Paul	2015	15.2	0.72	4.85	0.82	4.01	0.63
Peter	2018	10.6	0.70	3.85	1.08	0.86	0.50
Paul	2018	20.5	1.02	3.52	0.87	0.89	0.24
Peter	2019	20.0	0.54	8.95	1.15	13.4	0.79
Paul	2019	20.5	0.41	5.8	0.76	1.71	0.58

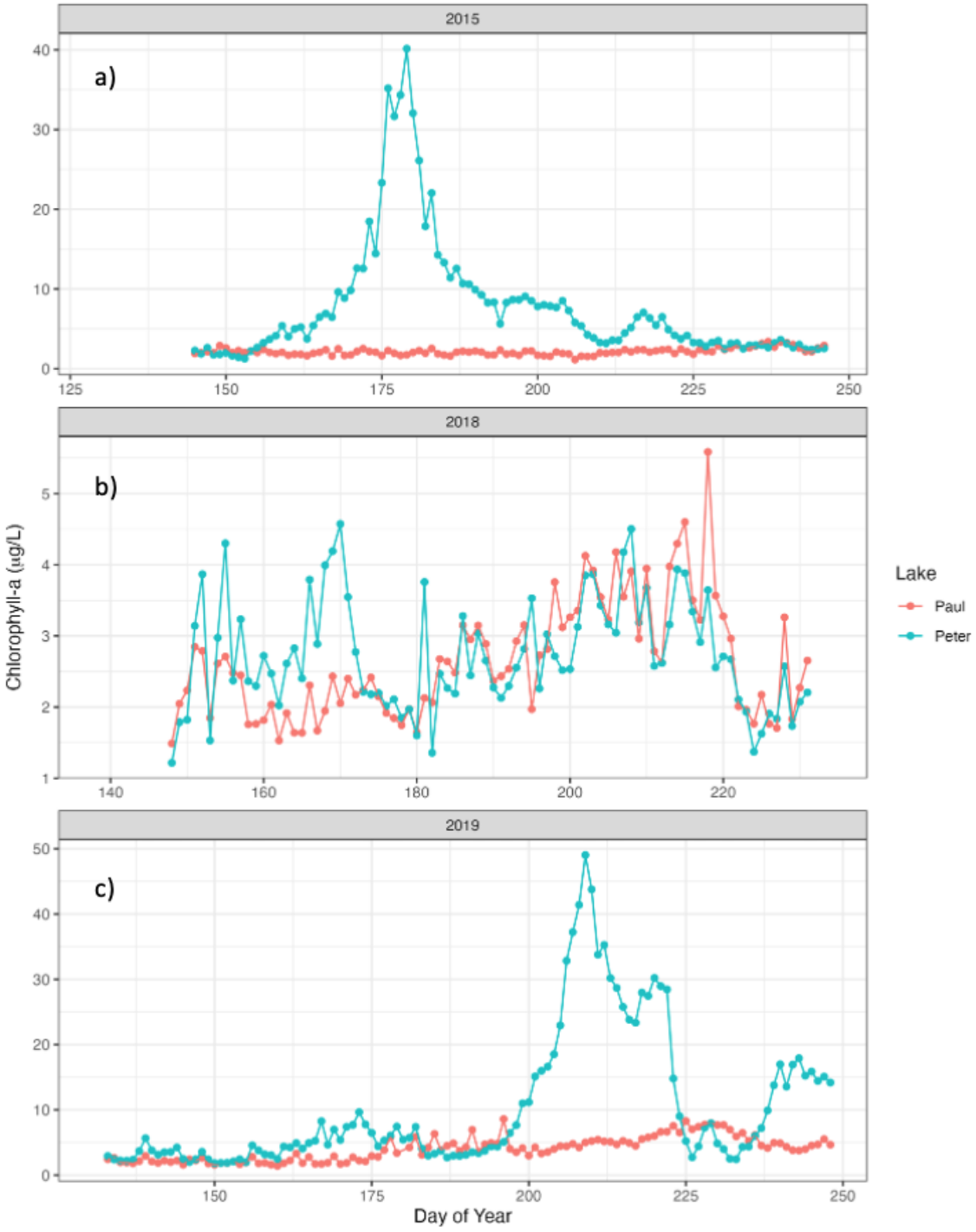

Figure 2.1. Time series of chlorophyll-a for Peter and Paul Lakes in a) 2015, b) 2018, and c) 2019. Peter Lake received nutrient additions in 2015 and 2019.

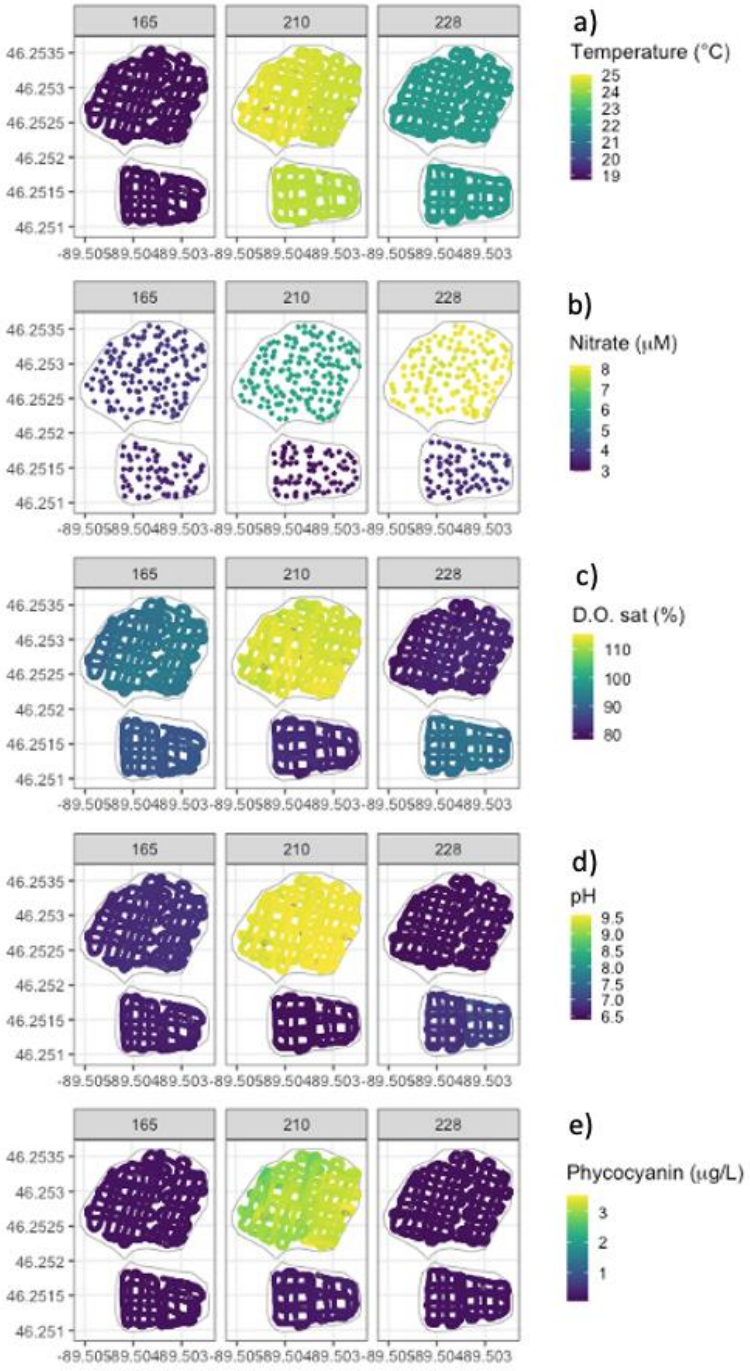

Figure 2.2. Values of a) temperature, b) nitrate, c) dissolved oxygen saturation, d) pH, and e) phycocyanin for Peter and Paul Lakes in 2019 before, during, and after the bloom. Peter Lake is north of Paul Lake.

Figure 2.3. Time series of spatial coefficient of variation of temperature, nitrate, dissolved oxygen, and pH for Peter and Paul Lakes for the years studied.

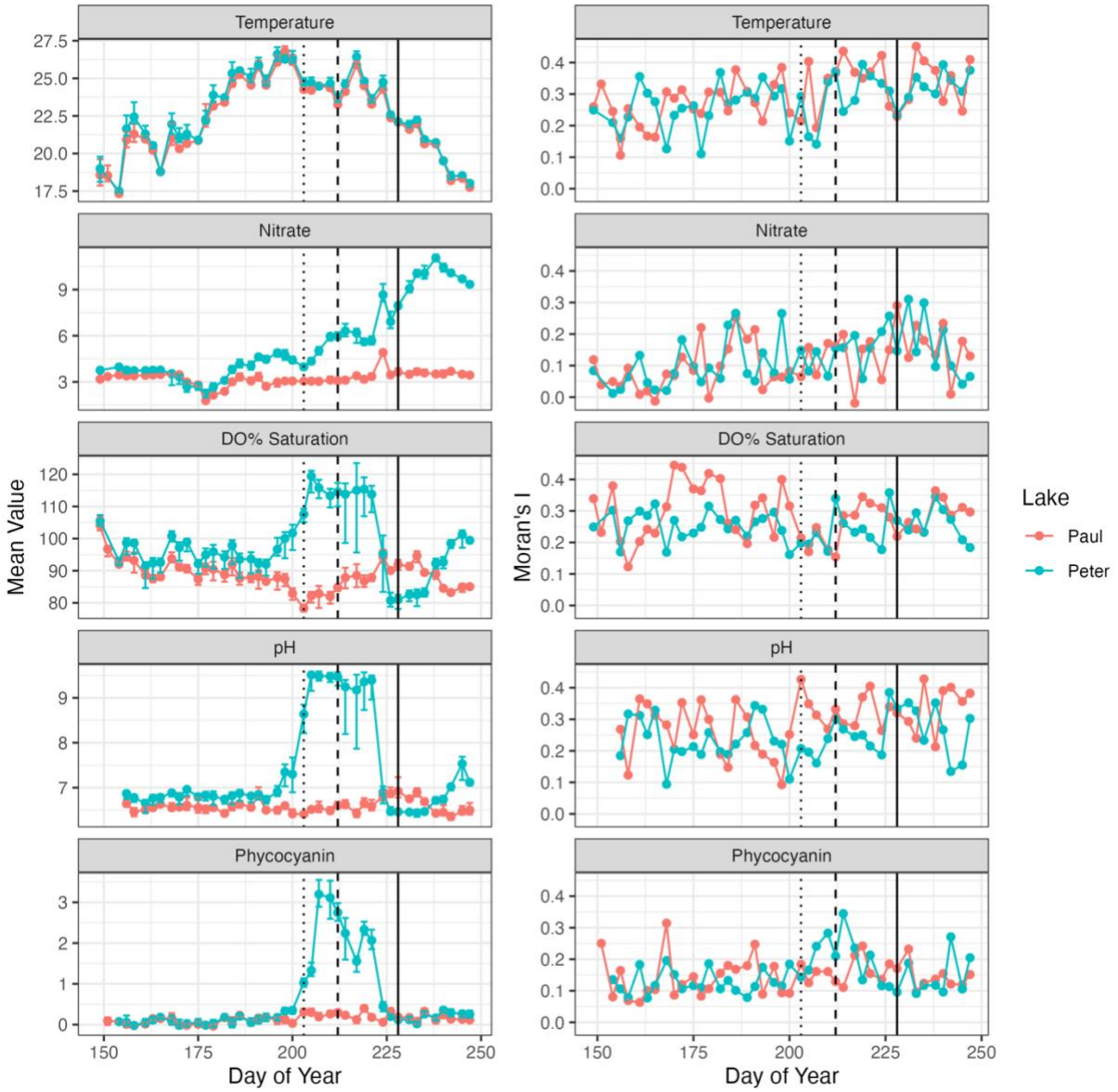

Figure 2.4. Time series of the mean value and Moran's I of temperature, nitrate, DO% Saturation, pH, and phycocyanin for Peter and Paul Lakes in 2019. The left column are time series of the mean values of the five studied variables. Error bars represent the range of the observations for each sampling event. The right column are time series of the Moran's I value of the five variables. The dotted, dashed, and solid lines represent bloom onset, peak, and crash, respectively.

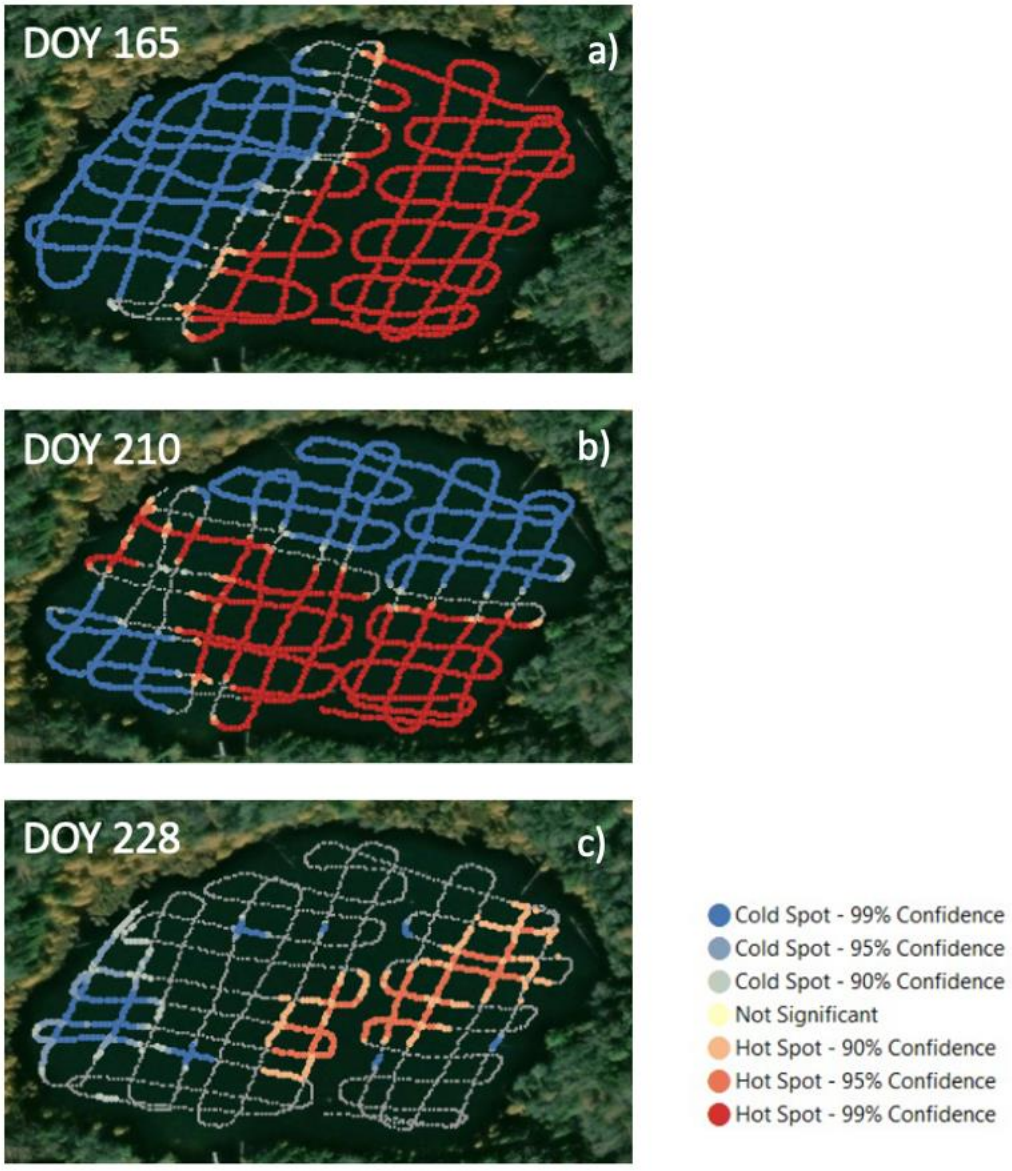

Figure 2.5. Hotspot analysis of dissolved oxygen for Peter Lake in 2019 on three dates indicated by day of year (DOY) a) before the bloom: DOY 165, b) during the bloom: DOY 210 , c) after the bloom collapsed: DOY 228. Hot and cold spots represent areas where high and low values cluster together, respectively.

Conclusions

Advances in water quality measurement technology increasingly enables scientists to examine fine temporal and spatial dynamics of waterbodies. In my study, I used high frequency phytoplankton pigment data and a disturbance-recovery algorithm to quantify disturbances and recovery in experimental and monitored lakes. Bloom onset, disturbance magnitude, and recovery time were measured and compared for two experimental lakes across several different nutrient addition manipulations. The blooms differed among lakes and years. Further, I demonstrated the method can be applied to monitored lakes and identify unusual disturbances. A next step would be to consider applying the method in real time to detect disturbances as they occur and to quantify their dynamics. Even more exciting would be the possibility of studying management interventions upon bloom detection to evaluate whether actions are effective.

I used high resolution spatial data to quantify and compare spatial heterogeneity in an experimental and reference lake. Surprisingly there was little spatial variability or persistent spatial structure in Peter and Paul Lakes even during bloom periods. An oft criticized approach in limnology of only sampling at the center of the lake appears effective for lakes like Peter and Paul. Spatial variation is so low that it can be ignored at least for the variables I considered. A next step in studying spatial heterogeneity in lakes would be to make similar spatial analyses across a large gradient of size, nutrient, and morphometry, using either the FLAMe or remote sensing. Such a study would allow one to characterize for what variables and conditions lakes transition from having low to significant spatial heterogeneity.