SPATIOTEMPORAL DYNAMICS OF THE HYDROLOGIC DRIVERS OF HARMFUL ALGAL BLOOMS IN TWO PIEDMONT RESERVOIRS, GEORGIA

by

ABIGAIL S. KNAPP

(Under the Direction of Adam M. Milewski)

ABSTRACT

Freshwater resources are at risk globally from increased eutrophication and toxic algal blooms. These cyanobacterial and algal events typically occur in the summer when the reservoir is well-stratified and rich in nutrients from erosion and runoff. Prediction of bloom occurrence requires an understanding of the factors which trigger bloom formation, and on which spatiotemporal scales. The biochemical and physical contributors have been studied in depth, but there is still uncertainty in the temporal and spatial scales at which climate or hydrologic contributors are important. We identified the primary hydrologic drivers of harmful algal bloom biomass within Lake Allatoona and Lake Lanier, Georgia, using a combined in-situ and satellite remote sensing dataset. Linear and multiple regression suggest that in shallower and more energetic reaches of the reservoir, drought severity is the primary contributor to biomass, while storm duration drives growth of harmful algal blooms in deeper, more lacustrine environments.

INDEX WORDS: Cyanobacteria, blooms, reservoir, hydrology, precipitation, drought, eutrophication, Lake Allatoona, Lake Lanier
SPATIOTEMPORAL DYNAMICS OF THE HYDROLOGIC DRIVERS OF HARMFUL ALGAL BLOOMS IN TWO PIEDMONT RESERVOIRS, GEORGIA

by

ABIGAIL S. KNAPP

B.S., The University of Georgia, 2013

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2018
SPATIOTEMPORAL DYNAMICS OF THE HYDROLOGIC DRIVERS OF
HARMFUL ALGAL BLOOMS IN TWO PIEDMONT RESERVOIRS, GEORGIA

by

ABIGAIL S. KNAPP

Major Professor:  Adam M. Milewski
Committee:
    Richard H. Becker
    Elizabeth A. Kramer
    Deepak R. Mishra

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2018
DEDICATION

To Christopher Knapp.

_Inversnaid_

This darksome burn, horseback brown,
His rollrock highroad roaring down,
In coop and in comb the fleece of his foam
Flutes and low to the lake falls home.

A windpuff-bonnet of fáwn-fróth
Turns and twindles over the broth
Of a pool so pitchblack, féll-frówning,
It rounds and rounds Despair to drowning.

Degged with dew, dappled with dew
Are the groins of the braes that the brook treads through,
Wiry heathpacks, flitches of fern,
And the beadbonny ash that sits over the burn.

What would the world be, once bereft
Of wet and of wildness? Let them be left,
O let them be left, wildness and wet;
Long live the weeds and the wilderness yet.

-Gerard Manley Hopkins
ACKNOWLEDGEMENTS

There are many more people to whom I owe a great deal, but in summary; I would like to acknowledge my advisor for inspiring and guiding me, my committee for excellent lessons and advice, my lab colleagues for acceptance and encouragement, and Rachel Rotz for showing me how it’s properly done. To my fellow graduate students, thanks for the boosts, tips, and commiseration. To my department, thank you for the support and opportunities. To my professors, thank you for the many important lessons and also humorous anecdotes. To Dr. Marta Patiño, Dr. Paul Schroeder, Dr. Doug Crowe, and Dr. Mike Roden, thanks for the great working environment and the chance to learn how to teach. To my colleagues in Geography—Sam Weber, Abhishek Kumar, and Ben Page—thanks for putting up with my questions and for helping a fellow bloom enthusiast. To the Little River Marina and to my co-adventurers, thanks for some of the best fieldwork I could ask for. To Dr. Shannon Quinn, thank you for the cookies and mentorship. Thanks to my family and friends for cheering me on. Thanks to Dr. John Dowd for introducing me to boundary conditions, and to Todd Rasmussen for introducing me to G. E. Hutchinson’s books. Thanks to Christopher Knapp especially for the love and support.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>v</td>
</tr>
<tr>
<td>1  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2  LITERATURE REVIEW</td>
<td>10</td>
</tr>
<tr>
<td>3  WHEN BLOOMS LIKE IT STORMY: THE SPATIOTEMPORAL DYNAMICS OF</td>
<td>25</td>
</tr>
<tr>
<td>FRESHWATER HARMFUL ALGAL BLOOMS IN TWO PIEDMONT RESERVOIRS</td>
<td></td>
</tr>
<tr>
<td>4  REMOTE SENSING OF CYANOHABS IN SOUTHEASTERN RESERVOIRS: A</td>
<td>69</td>
</tr>
<tr>
<td>COMPARISON OF METHODS</td>
<td></td>
</tr>
<tr>
<td>5  CONCLUSIONS AND RECOMMENDATIONS</td>
<td>116</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>122</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
<tr>
<td>A  A GUIDE TO FIELD FILTRATION FOR HPLC ANALYSIS</td>
<td>141</td>
</tr>
<tr>
<td>B  STANDARD PRECIPITATION INDEX TUTORIAL</td>
<td>148</td>
</tr>
<tr>
<td>C  R FUNCTIONS</td>
<td>165</td>
</tr>
<tr>
<td>D  OLI TO SENTINEL 2B VICARIOUS CALIBRATION</td>
<td>168</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 3.1: Multiple regression models by site show a variety of significant linear relationships, ranging between 12% and 74% variance explained .................................................................44

Table 3.2: Multiple regression models of remotely sensed Chl-α and PC data by site show a variety of significant linear relationships, ranging between 18% and 91% variance explained. ..................................................................................................................45

Table 4.1: The primary productivity and turbidity at the two reservoirs shows a range consistent with mesotrophic to eutrophic conditions...........................................................................................................74

Table 4.2: Selected algorithms from the main algorithm families including their model performance from the original study.................................................................................................................79

Table 4.3: The three tested algorithms for modeling harmful algal blooms were compared with in-situ chlorophyll-α data ...........................................................................................................88

Table 4.4: FLAASH, C2RCC, the DS Histogram method, and the DS Cloud Shadow (Region of Interest) method were compared by their qualitative and quantitative performance........95

Table 4.C.1: The percent root mean square error by band for each of the compared atmospheric corrections shows that the C2RCC method has the highest performance, followed by Dark Subtract using cloud shadows, and FLAASH.........................................................115
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>The drivers of cyanobacterial Harmful Algal Blooms may be grouped into biochemical, physical, and climatohydrological families</td>
<td>2</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Unit Hydrograph from the Upper Etowah watershed shows a steep rising limb and complex falling limb</td>
<td>7</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>A conceptual model of the interactions between the drivers of cyanoHABs and other harmful algal blooms based on the literature</td>
<td>12</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>The ideal conditions for bloom onset within a lake or reservoir include a well-developed thermocline, long retention times, and plentiful orthophosphate</td>
<td>29</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Study location showing data sources and sites in the Etowah and Chattahoochee River Watersheds</td>
<td>32</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>The overall workflow for the study progresses from dataset compilation and processing to comparison of all model parameters with cyanoHABs across a variety of spatial and temporal scales</td>
<td>35</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>The major embayments of eastern Allatoona are formed by the Etowah River, Little River, and Kellogg Creek tributaries</td>
<td>37</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>The sites at Lake Lanier are at the two main tributaries, the Chestatee River and the Chattahoochee River, as well as the several embayments formed by minor inflows</td>
<td>39</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Allatoona water quality, from top left, counterclockwise</td>
<td>41</td>
</tr>
</tbody>
</table>
Figure 3.7: The Standard Precipitation Index for Lake Allatoona and Lake Lanier show several droughts within the last decade, notably from 2007-2009 .............................................................48

Figure 3.8: Standard Precipitation Index, precipitation, drought severity, and drought duration for Lake Allatoona and Lake Lanier during the study period ............................................................49

Figure 3.9: The variance-explained of each significant linear model is shown by site ............53

Figure 3.A.1: In-situ and satellite remote sensing coverage of algal bloom proxy data, from 2008 to 2017… ..........................................................................................................................66

Figure 3.A.2: Bivariate plots showing the spread and slope of several of the high-performance chlorophyll-α models ........................................................................................................67

Figure 3.A.3: HABs were observed in satellite photography at 31 locations along the shoreline at Lake Allatoona (Google Earth, October 2012) .................................................................68

Figure 4.1: The remote sensing scientist’s choice of sensor, preprocessing, and algorithms is a balance between greater accessibility (ease-of-use, public access) and greater fidelity (spectral resolution, lower RMSE) ..................................................................................................75

Figure 4.2: Historical data were obtained from five sites, three of which were ground truthed, at Lake Allatoona ..................................................................................................................77

Figure 4.3: The methodology used in this study. Methods in white were only compared and were not used in the final products ..................................................................................................80

Figure 4.4: Sample distribution of the Landsat scenes from 2008 to 2017 .................................80

Figure 4.5: Modeled chlorophyll-α and phycocyanin show a positive linear relationship with measured chlorophyll-α concentrations ..........................................................................................84

Figure 4.6: Chlorophyll-α concentrations measured at Lake Allatoona on July 17 and July 24, 2017 show good agreement with coincident field spectra .................................................................85
Figure 4.7: From top, a cloudless true color image of the Etowah River, the Little River, and the Kellogg Creek reaches of Lake Allatoona, and coincident chlorophyll-α, floating algae, and PC index products ..................................................................................................................................................86

Figure 4.8: Modeled chlorophyll-α index based on the areal maximum of Duan algorithm agree with the bloom event in late summer of 2012, and show an overall smoother trend ..........87

Figure 4.9: Modeled phycocyanin in Lake Allatoona through time (Vincent et al., 2004) ..........90

Figure 4.10: The spectral shape retrieved using various atmospheric corrections may be either dampened or exaggerated. ........................................................................................................................................93

Figure 4.A.1: The Floating Algal Index shows a reverse relationship with stormflow ..........113

Figure 4.B.1: Chlorophyll-α Index scenes through time (Duan et al., 2007) .........................114

Figure 4.C.1: Dark Subtract Histogram method applied to scenes with minimal cloud cover before manual radiometric corrections result in reduction in the Rayleigh-dominated wavelengths........................................................................................................................................115

Figure 5.1: An updated conceptual model shows our findings that rainfall duration and stormflow are statistically significant drivers of HABs in freshwater Piedmont reservoirs........117

Figure 5.2: A view of the watershed includes lowered shoreline due to drought, gravity currents carrying sediments, and upriver storm events all interacting to influence the physicochemical conditions in the reservoir........................................................................................................................................119
CHAPTER 1
INTRODUCTION

Of the world’s water resources, less than 3% is accessible as liquid freshwater. As the climate shifts and the influence of human activities grows, freshwater reserves across the planet are increasingly at risk. Growing agricultural, industrial and municipal water use requires ever more clean water, whether from surficial or groundwater sources, yet even as demand increases, water quality is decreasing. Surficial water is particularly vulnerable to eutrophication, the state in which a body of water is over-enriched in nutrients (Paerl et al., 2001).

Eutrophication can be a natural process induced by the flux of nutrients and sediments from mountains to valleys through river systems (Heisler et al., 2008). Decaying organic matter, both suspended and dissolved, load water bodies in the lowlands with plentiful carbon, nitrogen, and phosphorus which sustain plants and other primary producers. The silt and clay sediments born by the same flow are deposited within the basin, and may over geologic timescales reduce the lake depth until a meadow is formed (Hutchinson, 1957).

However, human activity may also cause eutrophication through change in land cover and both direct and indirect contamination of the streams and rivers. As vegetation is stripped from the banks of streams and rivers, and natural lake-shores are developed, large amounts of sediments and runoff flow into the waterways. The sediments limit light and retard photosynthesis, while the runoff may also contain fertilizer and animal waste. In some areas, sewage and industrial waste are directly introduced. These accelerated processes can result in a shift in the trophic state of a lake on a scale of years to decades, and the resulting increase in
Figure 1.1: The drivers of cyanobacterial Harmful Algal Blooms may be grouped into biochemical, physical, and climato-hydrological families.
productivity is often accompanied by large blooms of harmful algae and cyanobacteria (Paerl, 2008; Paerl and Huisman, 2008).

Cyanobacterial harmful algal blooms, or CyanoHABs, are phytoplanktonic colonies that often form thick, visible mats of blue and green biomass on the surface or float as turbid masses within the epilimnic zone of lakes or ponds, disrupting the ecological balance and causing human health hazards (Lampert, 1987; Hudnell, 2010). CyanoHABs can displace predecessor phytoplankton and may be toxic, leading to disruptions up and down the trophic levels of the food web (Paerl and Otten, 2013). CyanoHABs can also cause a sharp reduction in dissolved oxygen within the water body during respiration and decay, causing extensive fish kill (Paerl et al., 2001, Drever, 1982). This oxygen uptake may also negatively alter the geochemical cycling between water and sediments (Hupfer and Lewandowski, 2008).

However, the most serious concerns during a cyanoHAB event come from the cyanotoxins themselves. According to the EPA, livestock, pets, and humans who bathe in or ingest cyanoHAB-contaminated waters may experience symptoms ranging from respiratory irritation to death, as cyanobacterial species can produce dermatoxins, liver-damaging hepatotoxins, neurotoxins, and are being studied as potential carcinogens (Hudnell, 2010; Merel et al., 2013; Westrick, 2014). While many cyanobacterial genera introduce cyanotoxins into the water upon lysis, or cell death and rupture (Merel et al., 2013), there are several varieties which release the toxins during growth (Westrick, 2014). Treatment of water for cyanotoxins is expensive, but the need will only expand with our growing consumption of water resources and the increase of cyanoHAB occurrence (Hudnell, 2010; He et al., 2012).

Prevention of cyanoHAB events is in the interest of public health officials, both to protect humans and animals from direct exposure, but also to preserve the resource which may be
invaluable for the surrounding region’s household, industrial, and agricultural use. Since the 1970s Clean Water Act, regulations in the US, as in much of the post-industrial world, have increasingly been focused on prevention of eutrophication by regulating release of nutrients from industrial waste and agricultural fertilization into marine and freshwaters (Paerl, 2008; Schindler, 2008; Hudnell, 2010). These efforts require well-orchestrated collaboration between regulatory agencies, researchers, and water users, a process which will always need improvement (Watson, 2016). The issue is global, as some of the world’s largest lakes—Lake Taihu, China, Lake Winnipeg, Canada, Lake Erie, Michigan, and Lake Victoria in Africa’s Great Rift Valley are all presently threatened by severe eutrophication and recurring cyanoHABs (Paerl, 2008; McCullough et al., 2012).

Research on cyanoHABs includes remediation, prevention, and most recently, prediction through a fuller understanding of the triggers which set off new and perennial blooms. The challenge in predicting cyanoHAB events lies in the complexity of the conditions which lead up to a bloom event (Figure 1.1). Within the lake, sunlight, nutrients, temperature, and water residence time are important risk factors (Coveney et al., 2005; Heisler et al., 2008; Paerl, 2008; Hudnell, 2010). On the watershed scale, land use, geomorphology, geology, hydrology, climate and human engineering all interact to promote or inhibit harmful algal growth (Fink and Mitsch, 2007; Jones and Elliot, 2007; Mihaljevic and Stevic, 2011; McCullough et al., 2012; Reichwaldt and Ghadouani, 2012; Bowling et al., 2013; Huang et al., 2014; Istvanovics et al., 2014; Karan et al., 2014).

This study focuses on the effects of these interconnected drivers in the setting of Lake Allatoona and Lake Sidney Lanier, two flood-control U. S. Army Corp of Engineers (USACE) reservoirs in the southeastern Piedmont. While cyanoHABs commonly occur in natural lakes,
reservoirs are especially susceptible to anthropogenic eutrophication through point- and nonpoint-source nutrient pollution. Water depth control and shallow basin morphology render them even more vulnerable to cyanoHAB events during warm and dry months (Bloesch, 1995).

Both reservoirs are located primarily within the Piedmont physiographic province in North Georgia. The Piedmont is a swath of fractured metamorphic and igneous bedrock which forms the base of the Appalachian Mountains and is characterized by gentle relief and hill-and-valley topography (Hatcher, 1978). The climate within the province is temperate to subtropical, with temperatures varying from a winter average of 43° Fahrenheit to a summer average of 77° (NOAA, 2016). The average annual precipitation for the reservoirs is approximately 54 inches, with the maximal rainfall in the winter months and sporadic July thunderstorms, and a warm dry period in the late summer (NOAA, 2016; Stevens, 1993).

The Piedmont is covered by a dense network of perennial streams and river channels, which are fed by both rainfall and a surficial aquifer hosted within layers of saprolite, soil, and alluvium. The average yield from bedrock wells in the Piedmont is 15 gallons per minute, necessitating the use of reservoirs for public water supply in Atlanta and surrounding urban areas (Miller, 1990; Stevens, 1993). The reservoirs throughout the Southeast were built to mitigate damage caused by heavy floods, but are also the primary source of drinking water for much of urban North Georgia, in addition to providing wastewater dilution, hydroelectricity, and recreational services (Stevens, 1993).

Lake Allatoona lies within the Etowah Watershed of the Coosa River Basin at the head of the Alabama-Coosa-Tallapoosa (ACT) river system. The nearby highly developed cities of Canton and Cartersville share the watershed with mixed forest, suburban neighborhoods, croplands, poultry farms, and a quarry (UGA CAES, 2009). Nutrient contributions from the
major industries and farms are monitored and limited by NPDES permits to ensure that Total Maximum Daily Loads (TMDL) to the Etowah and Little River are not exceeded. The Little River receives point source loading of 18,500 pounds of phosphorus annually from a wastewater reclamation facility, two water pollution control plants, and a poultry plant, accounting for over 60% of the TMDL for the embayment (TMDL ID 9825: EPD, 2004).

Allatoona Dam was built below the conjunction of the Etowah River and tributaries Little River, Allatoona Creek, and Noonday Creek. The dam was constructed in 1949, with a drainage area of 1,110 square miles and a storage volume of 367,000 acre-feet. The reservoir ranges from a normal operating head of 190 feet at the dam to an average depth of 10 feet at the inflowing Etowah River, and the high drainage-to-surface area ratio results in a relatively flashy response within the watershed, meaning much of the energy produced by storm events anywhere within the watershed is quickly transmitted through discharge to the reservoir (EPD, 1998). A unit hydrograph constructed from a storm event in September 2010 is consistent with a runoff response in a relatively small watershed with urban land-cover and impervious surfaces (Figure 1.2).

Lake Allatoona has experienced several periods of hydrologic drought in the past half-century, as defined by rainfall indices. Most recently, a prolonged drought reaching a maximum Standard Precipitation Index of -3.05 resulted in extreme storage shortages in the year 2007. The reservoir pool height at the dam reached a low of 780 feet above sea level—60 feet below summer conservation level and more than 50 feet below the decadal average (USGS 02393500; USACE 1991). As a result of periodic low-flows and continuing plentiful nutrients in the reservoir, several major algal blooms have also occurred during the past two decades (Dirnberger, 2005; Ceballos, 2007).
Adjacent and to the east of the Etowah Watershed, Lake Sidney Lanier – man-made reservoir in the state of Georgia-- resides within the Upper Chattahoochee Watershed. Lake Lanier was formed in 1957, when the Chattahoochee River was impounded 48 river miles north of Atlanta (EPD, 2008). South of the dam, the Chattahoochee River flows through Atlanta and connects to the larger Apalachicola-Chattahoochee-Flint (ACF) system, before emptying into the Gulf of Mexico. Compared to the Etowah Watershed’s 1,110 square miles, the Upper Chattahoochee Watershed drains a total of 1,040 square miles. Despite the smaller watershed, the Lanier reservoir is more than three times the size of Allatoona, with a surface area of 73 square miles and a storage volume at static levels of 1,917,000 acre-feet of water (EPD, 2008). Lanier is fed by many streams and embayments, resulting in a highly dendritic perimeter with a shoreline of 540 miles to its 44 river miles (EPD, 2008). The reservoir is 156 feet deep at the dam pool, and provided an average of 68 million gallons per day (MGPD) to Gwinnett County alone in

Figure 1.2: Unit Hydrograph from the Upper Etowah watershed shows a steep rising limb and complex falling limb.
2015 (Gwinnett County Government, 2017). Threats to the public water services at Buford Dam include droughts, most recently in 2007 (Campana et al., 2012).

**Scientific Problem and Hypothesis**

Freshwater lakes and ponds across the world and in the Southeastern Piedmont are threatened by seasonal blooms of harmful and toxic cyanobacteria, collectively known as CyanoHABs, through cultural eutrophication by biological, chemical, and hydrogeological mechanisms. The understanding of the drivers of these CyanoHABs is necessary in order to predict, prevent, and rehabilitate these occurrences. In addition to the scientific interest posed by these questions, the application of hydrologic and meteorologic conditions to HAB predictions may be used by property owners to avoid toxin ingestion, by governmental agencies to develop drought-specific nutrient policies, and by resource managers to further improve dam release decisions.

Two of the primary drivers of algal bloom growth are precipitation, which controls direct and indirect nutrient flux through variation in discharge and runoff, and drought, which may promote cyanobacterial colonies through increased thermal stratification, surface water temperature, and residence time (Paerl and Huisman, 2008). CyanoHABs are known to be affected by the annual wet and dry seasons, but the effect of precipitation and drought on a longer timescale have yet to be elucidated (Reichwaldt and Ghadouani, 2012). I hypothesize that the long-term patterns of precipitation, especially storm intensity and frequency, and hydrological drought severity, correlate with and drive cyanobacterial HAB onset and severity in the Southeastern Piedmont Lakes Allatoona and Lanier.
The goals of this study are the following:

a. Determine the relative impacts of precipitation intensity, frequency, and duration and drought severity, frequency, and duration on cyanoHAB growth in Lake Allatoona and Lake Lanier using linear and multiple regression analysis.

b. Integrate remote sensing (Landsat 5, 7, and 8 and Sentinel 2) and in situ (USGS, EPD, NOAA, and EPA) datasets to study the effect of precipitation and drought on cyanoHAB growth in varying spatiotemporal scales and with improved spatiotemporal resolution.

Thesis Organization

The first chapter of this thesis provides background information and the scientific questions and objectives sought in this research. The second chapter reviews and summarizes the key literature on the drivers of cyanobacterial harmful algal blooms. The third chapter details the methods, results, and conclusions from the bio-hydrologic models developed in this project. The fourth chapter compares methods of remote sensing cyanoHABs in reservoirs, outlining challenges and recommendations for similar projects. The fifth chapter summarizes the implications of this research for Piedmont reservoirs and provides recommendations for further study.
CHAPTER 2

LITERATURE REVIEW

The appearance and severity of HAB events is controlled by natural and anthropogenic factors on local and regional scales (Figure 1.1). Immediately surrounding the colony, plentiful light, bioavailable nutrients, and high temperatures will promote growth of the HAB. The nutrients are delivered by the hydrology of the connecting streams and rivers, and at different rates depending on the land use and land cover within the watershed. The climate within the region drives the rainfall and temperature changes through time, and in turn may shift in response to human activity. To understand their role in HAB and phytoplankton growth, these systems must be considered not only separately but as interacting controls (Figure 2.1).

**Nutrients as a Control**

The biological-chemical drivers of HAB growth are well-studied and have been extensively considered from the perspectives of microbiologists, ecologists, and biochemists. The chemical makeup of the average phytoplankton, used to approximate the nutrient needs of algae and bacteria, is described in the Redfield ratio as:

\[
106CO_2 + 16HNO_3 + H_3PO_4 + 122H_2O + 18H^+ + \text{trace elements} + \text{solar energy} \leftrightarrow (CH_2O)_{106}16NH_3H_2PO_4 + 133O_2
\]  

\textbf{Eqn. 2.1}
The necessary ingredients for this simplified photosynthesis reaction include carbon (C), nitrogen (N), phosphorus (P), and light (Redfield et al., 1963). In fact, the ratios of these nutrients, their bioavailability, and their concentrations are the most important chemical factors in HAB appearance, biomass, duration, and even toxicity (Coveney et al., 2005; Heisler et al., 2008; Paerl and Otten, 2013).

Much of the carbon needed to sustain a lake’s phytoplankton come from dissolved organic carbon (DOC) provided by plant litter and algal decay in the surrounding lotic network of streams and rivers (Kjeldsen et al., 1998; Kalff, 2002). In flowing freshwater, organic matter is carried along as it decays, providing carbon and other nutrients to the nearby phytoplankton, and eventually, to the downstream lake-dwelling organisms in cyclical uptakes and releases (Newbold et al., 1983). DOC in lakes and reservoirs is primarily affected by land use and climate parameters, with increased concentrations in forested watersheds with high soil carbon density but decreased concentrations in watersheds with higher average runoff (Sobek et al., 2007). Urban land use may also play a role, as sulfur air pollutants have shown a significant inverse relationship with DOC concentrations in many North American and European lakes (Evans et al., 2005; Monteith et al., 2007).

The importance of the two limiting macronutrients for photosynthesis, N and P, have traditionally been thought to vary with the salinity of the body of water in question. Most authors cite P as the limiting nutrient in freshwater bodies, N as the limiting factor in marine and bay environments, and a ratio of N:P to describe the nutritional needs of estuarine waters, depending on the salinity gradient (Heisler et al., 2008; Paerl and Huisman, 2013). Total and bioavailable nutrients are also thought to influence the toxicity of the cyanoHAB colonies, although whether the blooms produce more toxins under stress as other phytoplankton can (Lundgren et al., 2013)
or during episodes of hypereutrophication (Gobler et al., 2007) has not yet been elucidated. However, the exact relationship of P, N, and the ratio of N:P required for freshwater bloom onset, severity, and recurrence has been debated since the 1975 publication of the Vollenweider model. The mathematical formula which Vollenweider constructed linking phytoplanktonic biomass to P is the basis for the modern understanding of the effect of nutrient load on algal blooms and has been praised for its elegance and utility (Reynolds, 1992; Lowe and Steward, 2011). From Vollenweider’s model to a recent 37-year study of nutrient loads to an experimental lake, much prominent research has focused solely on P (Schindler et al., 2008), to the extent of

Figure 2.1: A conceptual model of the interactions between the drivers of cyanoHABs and other harmful algal blooms based on the literature. The key drivers of the system are in bold.
warning that restricting N-loading is a costly and fruitless methodology (Schindler et al., 2008). However, studies investigating the response of algal growth to varying N:P ratios successfully argue that reality is a more complex system, and that under certain conditions N could play an important role (Heisler et al., 2008; Xu et al., 2010, Paerl and Otten, 2013; Smith et al., 2015). Several studies (Paerl and Paul, 2012; Paerl and Otten, 2013) report that the N requirements for a large cyanobacterial colony cannot be satisfied by diazatrophy (N-fixing) alone, and therefore cyanohAB growth must be dependent to some degree on an external source of N. Under conditions of higher N availability, the cyanobacteria showed a higher ratio of non-diazotrophic genera. Under conditions of low P input, it is possible that observed cyanobacterial growth could be nourished by N-loading (Pearl and Otten, 2013), or, as Schindler et al. (2008) argue, P-cycling from lake sediment could be to blame.

From a hydrological perspective, the two arguments are important to consider, as the physical cycles of N and P differ significantly. Nitrogen cycling primarily involves biologic and atmospheric processes as N$_2$ is released into and removed from the air by microbial activity. Phosphate is released from the lithosphere by weathering or mining processes, and the return of phosphate to rock through sorption to sediment may be affected by sedimentation, saturation, and mixing events.

There is also some evidence towards the increased mobilization of P through the mechanism of nitrate leaching and subsequent interaction between nitrate, iron, sulfate, and phosphate (Smolders et al., 2010). Therefore, although most nutrient limiting efforts focus on P (Coveney et al., 2005; Schindler et al., 2008; Conley et al., 2009; Monteagudo and Moreno, 2016) and models use P to describe growth rates (Lowe and Steward, 2011; Jones and Elliot, 2007), both macronutrients N and P must be considered as controls on cyanohAB growth.
(Conley et al., 2009; Paerl and Paul, 2012; Paerl and Otten, 2013), and therefore, hydrologic controls for both nutrient cycles must be researched.

However, the bioavailability of P is strongly linked to the dissolved oxygen (DO) gradients in a water column. In a well-stratified lake, the hypolimnion may develop anoxic, reducing conditions from the decay of organic matter and the respiration of benthic organisms. Iron hydroxides within the sediment may then be converted from the ferric to ferrous state, releasing the phosphate ions (Drever, 1982). Until readsobered, the P may remain available for uptake by the local phytoplankton (Filippelli, 2008). In the upper layers of the lake dominated by photosynthesis, oxidation may prevail, driving the resorption of P. This cyclical sequestration is an important source and sink of P (Ansari and Gill, 2014). Therefore, iron may interact with phosphorus to release or trap the P supply, acting as a control on the algal growth (Ansari and Gill, 2014). DO also directly affects the bioavailability of some micronutrients essential to HAB growth. Chi et al. noted that the absence of dissolved iron in well-oxygenated waters can retard cyanobacteria productivity (2015).

The nutrients may still not be converted to biomass without the energy from sunlight, which has been shown in several case studies to be an especially important factor to cyanoHAB growth in highly eutrophic systems (Dzialowski et al., 2011; Fanesi et al., 2016). Not only are some cyanobacteria well-adapted to actively pursuing light within the water column, but in some nutrient rich lakes, light may be more limiting than nitrogen or phosphate (Havens et al., 2003; Biddanda et al., 2015). Additionally, light availability is inversely proportional to turbidity caused by seiches and stormflows (Bloesch, 1995).
Hydrology as a Control

Several large bodies of freshwater around the world are no longer limited in N and P. These hypereutrophic lakes, such as Lake Taihu, China, undergo annual HAB events which are primarily controlled by hydrology, morphology, and meteorology (Huang et al., 2014; Li et al., 2014; Zhang et al., 2016). The most important hydrological controls on cyanoHAB growth are widely agreed to be precipitation and the relationship between discharge and retention time (Jones and Elliot, 2007; Reichwaldt and Ghadouani, 2012; Paerl and Otten, 2013), although lake morphology, basin morphology, aquifer geometry and hydraulic conductivities, dams, and weirs have all been considered as additional variables (Bowling, et al., 2013; Huang et al., 2014; Istvanovics, et al., 2014). The majority of these research efforts were published within the last decade, and span numerical and computer modeling, field monitoring, and remote sensing techniques.

Discharge, a measure tightly correlated with precipitation, is a well-covered but not a fully described control on cyanoHAB growth. Daily discharge in the southeastern United States is naturally at its highest volume during “flushing season”, during the colder months of October-January (Langbein and Wells, 1955). The conflicting effects of discharge are decreased retention time and increased nutrient load in the receiving lake (Jones and Elliot, 2007; Reichwaldt and Ghadouani, 2012; Paerl and Huisman, 2009). Discharge directly affects the retention time of the lake’s water, increasing mixing and decreasing vertical stratification. The short-term effect of a high discharge event is to significantly reduce cyanoHAB biomass (Srifa et al., 2016; Paerl and Huisman, 2009; Bowling, et al., 2013; Mihaljevic and Stevic, 2011; Royer, et al., 2006). Data gathered from a model based on Lake George, Florida show an inverse relationship between discharge rates and cyanobacterial biomass (Srifa et al., 2016). These data are not enough to
show that discharge causes a net loss of cyanoHAB biomass, as the decrease in mass also coincided with the decrease in temperature through the winter flushing season. However, many other studies have linked increased discharge to a short-term decrease in biomass, separate from the annual decrease due to cyclical temperatures and flushing (Lundgren et al., 2013). High discharge events such as flooding causes flushing to occur through the river channel and in the near-shore areas of a lake, which disrupts the thermal stratification that cyanobacteria exploit, as well as breaking the colony into smaller bodies and even moving the bloom to other areas in the lake where conditions may not be optimal (Lundgren et al., 2013).

The cyanobacterial colonies are usually only deterred, rather than destroyed, by high discharge events. Paerl and Huisman (2009) show that cyanoHAB growth returns after flushing events and several other studies corroborate, with evidence that the nutrient loading from increased discharge even from the very disruptive process of storms, flooding and other extreme events is exploited by phytoplanktonic growth after the lake water becomes calm again (Ahn et al., 2002; Bowling et al., 2013). In fact, allochthonous loading due to the extreme events may be the most significant source of nutrient inflow for some lakes. Royer et al. (2006) reported simulated data indicating that reduction of nutrient input during periods of low discharge did not significantly alter the total annual concentrations of N and P, suggesting that the high discharge events accounted for the vast majority of nutrient loading in their experiment. In another study, modeled data of nutrient concentration in Lake Winnipeg indicated that flood events were the most significant source of nutrient inflow in their study area (McCullough et al., 2012). Some lakes have accumulated these sediments through time, building a large store of adsorbed phosphorus which may be released later through mechanical or chemical disturbance of the sediments (Bloesch, 1995; Ceballos, 2007). This autochthonous nutrient cycling is a special
concern for bloom control efforts. Whether by the oxygenation of adsorbed sediments or the recycling of organic detritus, the resuspended sediments can provide a pulse of nutrients to the phytoplankton, triggering increased growth (Mayhew et al., 2001; Hupfer and Lewandowski, 2008, Holbach et al., 2013). High-flow events may even trigger bloom events through dilution, as in the normally brackish Swan River estuary in Western Australia (Robson and Hamilton, 2003).

Engineered flushing events have been used as a mitigation technique to clear HAB colonies from freshwater, with varying success (Paerl et al., 2001). Paerl and Huisman noted that flushing flows can have impacts on vertical and chemical stratification beyond simply removing HABs from the local waters (2009). In the study of Murray River, Australia, flushing caused by natural and dam release events were suggested to be responsible for “seeding” downriver lakes with cyanobacteria that had originated in lakes upriver (Bowling et al., 2013). It is possible that in the future, cyanobacteria may develop resistance to physical separation and disruption; in any case, the full ramifications of using artificial flushing as a cyanoHAB mitigation technique are underexplored in the literature.

Inversely proportional to high discharge, long retention times can be shown to correlate with greater algal growth (Maceina and Bayne, 2003; Jones and Elliot, 2007; Elliot et al., 2009; Longphuirt et al., 2016). According to one study using PROTECH to create a model based on the UK lake Blelham Tarn, retention time could have a complex relationship with cyanoHAB growth (Jones and Elliot, 2007). PROTECH is a comprehensive tool that has been used successfully in several studies by combining a phytoplanktonic growth equation and a physical model which takes into consideration thermal stratification and inflow. Jones and Elliot (2007) used the model to investigate the effect of increased retention time on phytoplanktonic
communities, using chl-α concentration as a proxy. They ran their model holding nutrient concentration to be constant at all retention times, and then as a function of retention time, increasing with increasing discharge. When the ratio of variable nutrient load to constant load was plotted against increasing retention time, the data resulted in a negatively sloped exponential curve. These results suggest that the shorter the retention time, the more important the nutrient input becomes to chl-α growth. In addition to chl-α concentration, phytoplankton composition was considered. In any year, bimodal peaks of chl-α were observed, coinciding with the traditional springtime green algal blooms and the summertime cyanobacterial blooms. These data were verified by field observations from Blelham Tarn in 1974. Interestingly, the cyanobacterial or summer bloom had a significantly higher rate of concentration increase than did the spring bloom with increasing retention time, although the effect was one of decreasing returns. Elliot et al. (2009) later returned to the question of the relationship between nutrient loading, retention time and bloom success. They modeled Lake Bassenwaite, UK, using both PROTECH and the Vollenweider model, gathering data which was consistent with their 2007 findings that with shorter retention times, nutrient load was of greater importance. In addition, they found the Vollenweider model to be simplistic with regards to the effect of flushing on algal bloom biomass. Vollenweider deals with biomass loss from flushing as a function of nutrient loss from the event; however, the more complex PROTECH model also integrates the direct physical loss of biomass during flushing. Comparing the two modeling methods, using the somewhat generalized Vollenweider method does not appear to be sufficient when attempting to discern the effects of complex multifactor relationships in a specific lake, whereas PROTECH takes into consideration additional variables such as sunlight and shading (Elliot et al., 2009).
Long term drought may cause low-flow, high retention conditions in a lake or reservoir, increasing the thermal stratification and water column stability and affecting the phytoplankton dynamics for months or years (Dirnberger, 2011; Camacho et al., 2016). Drought has been shown to have an impact on bloom growth even in monsoon climates: in a study of the Taechung Reservoir, Korea, An and Jones reported that a cyanobacteria-dominated bloom coincided with an unusually mild monsoon season, and concluded that the resulting well-developed thermocline and high surface temperatures contributed significantly to the surface scum (An and Jones, 2000).

While cyanobacteria thrive most readily in the calm environment of a lake, basin slope can allow for cyanoHAB proliferation even in rivers. This natural lowland morphology was shown to have a significant impact on recurrent blooms in such locations as Murray River, Australia, where the river basin has low bed gradients and is dominated by slow, meandering channels (Bowling, et al., 2013). The Murray River, the flow through which is also altered by weirs, dams, and diversions, hosted major bloom events which coincided with the drought-caused low-flow periods in 2009 and 2010. Although dams and weirs increased retention times to an average of 8 days in some sections, this was not determined by the authors to be enough of an increase to have influenced the blooms. The low flow velocity in the river, coupled with high temperatures, were the two factors during these blooms that can be tied to the survival of the cyanobacteria which was present over a 1000 km length of the river. In their concluding remarks, the authors also tentatively hypothesized that the river turbulence would have been low due to the low bed shear stress. Istvanovics, et al. (2014) also found in their study of a model based on the Somes River in Eastern Europe that catchment topology was a control for phytoplanktonic growth. From their data, which used flow velocity for channels and runoff, retention time, water
depth, nutrient concentrations and phytoplanktonic growth, they were able to show that growth was partially a function of channel slope (Istvanovics, et al., 2014). This is intuitive, as channel slope affects flow velocity and retention time, but the ability of phytoplankton to thrive in rivers has in the past been in doubt (Bowling et al., 2013). Lake surface area, depth, volume, and basin gradient are all essentially constant and have been suggested by Pitois et al. (2014a) as the structure on which risk modeling can be built when paired with meteorological and climate-change data to simulate volume and timing changes in precipitation.

In addition to basins, the morphologies of lakes and aquifers have also been correlated with nutrient or cyanobacterial concentration (Karan et al., 2014). One mechanism for diffuse nutrient transport is through groundwater flow, as reported by Karan et al (2014). Their study used a FEFLOW model to observe a nitrate plume across a groundwater-lakebed boundary in an agricultural site in Lake Hempen, Denmark. The constraints on nitrate concentration in the lake from aquifer flow were the ratio of hydraulic conductivities between lake bed and underlying aquifer, where the lake bed had a thin layer of low conductivity material near the lowest gradient, deepest area of the lake floor. The depth and fetch of a lake is another important parameter in bloom formation, as wind may churn seiches deep enough to resuspend bottom sediments. This autochthonous nutrient cycling has been observed in Lake Taihu, China, as well as Lake Okeechobee, Florida (Canfield and Hoyer, 1988; Xu et al., 2010).

**Land Use as a Control**

If hydrology provides the transport for nutrients to the phytoplankton within lakes and rivers, land use controls the nutrient supply. Agricultural and urban land use is associated with increases in soil erosion, total N and P in watershed streams, and HAB growth, while natural
land cover such as forest canopy and wetlands are effective nutrient sinks (Graham, 2001; Verheyen et al., 2015). Agricultural land use such as row crops, pastures, and livestock operations are major sources of N and P both from fertilizers and waste products, and generally replaces historical forested lands (Cordell et al., 2009). Commercial fertilizers also may contain phosphates designed for higher bioavailability and water solubility (Rehm et al., 2002; Cordell et al., 2009). If this material is applied to fields which are then exposed to a precipitation or storm event, the runoff may contain high levels of dissolved P or P sorbed to suspended particles from the cropland soil (Paerl et al., 2001; Delpla et al., 2011). Livestock farming also produces a huge amount of N and P from the waste, which may be injected into the waterways if there are no vegetation buffers or filters between the concentrated feeding operations and nearby streams (Beaver et al., 2014). In one study, models showed that land use changes from forests to crops and urban areas showed similar effects to those from climate change, including increased mineral P in the watershed streams and increased streamflow (El-Khoury et al., 2014).

Urban land use increase the flow of several pollutants into streams and rivers, but HAB growth is primarily positively affected by nutrients supplied from commercial fertilization, poor sewage management, and point source inputs, which all may contribute P to the watershed (Tromboni and Dodds, 2017).

The types of land use or land cover in a watershed not only determine the magnitude and types of nutrients flowing into the freshwater, but also the velocity. Deforestation and urban development increase the impervious surfaces in a watershed, magnifying the effect of runoff, erosion, and nutrient transport (Arnold Jr. et al., 2007). Decreased average baseflow in some urbanized watersheds may also promote HAB growth through low-flow conditions (Anderson et al., 2006).
Climate as a Control

As our planet continues to undergo anthropogenic climate change, researchers have increasingly begun to link global climate models to a potential for increased cyanoHAB productivity (Meis and Jones, 2009; Pitois et al., 2014). The meteorological dynamics which will be affected the most by climate change model predictions are mean temperature and extreme hydrologic events (Paerl and Paul, 2012). The forecasted rise in global temperature will favor cyanobacterial growth and proliferation in new geographical regions, as Paerl and Huisman (2008) state in their paper, “Blooms Like it Hot”. Likewise, in a study of Lake George, Florida Srifa et al. (2016) collected chlorophyll-α mass, precipitation, and water temperature data from 1993 to 2010. Srifa et al. found that cyanoHABs made up 90% of the biomass during most of the May – September growth season for cyanobacteria, correlating with the season of highest temperatures. Paerl and Otten (2013) further explored the correlation between climate and cyanoHAB success, and they find that cyanobacteria can outcompete other phytoplankton in warming temperatures. One mechanism is suggested by a study of several varieties of phytoplankton, in which the authors showed that cyanobacteria use light energy more efficiently in increased temperature when compared to green algae (Fanasi et al., 2016). Algae and cyanobacteria usually exist alongside a number of other phytoplankton species in a healthy lake, as well as zooplankton which may graze on the nontoxic primary producers and in turn feed a variety of finfish and shellfish. In a eutrophic lake under increasingly warm conditions, a HAB may vastly outcompete the other phytoplankton in the summer growth period, whether by increased release of a toxin such as microcystin which damages the other organisms or simply by superior adaptation to temperatures over 25˚ C (Paerl, 2008).
Warmer temperatures also promote thermal stratification (Paerl and Huisman, 2008; Paerl and Otten, 2013). Cyanobacteria genera are also known which have the ability to regulate their vertical position in the lake using pockets of gas in their colonial structures, thus allowing them to position themselves in the ideal vertical thermal strata within the body of water, and increased stratification will likely give an advantage to buoyant organisms (Paerl, 2008). Climate models also predict more frequent events of high precipitation and flooding as well as more severe and prolonged drought; in some regions, both storms and droughts will occur (Paerl, 2008; Paerl and Huisman, 2009; Reichwaldt and Ghadouani, 2012). Cyanobacterial blooms are likely to be directly affected by both extremes (Anderson et al., 2006; Mihaljevic and Stevic, 2011), which may become an even stronger control of bloom growth in increasingly nutrient-limited conditions (Yang et al., 2016).

Anderson et al. (2006) in their paper modeling nutrient loss from land to streams used a general circulation model for climate change and incorporating the new conditions into a NAM runoff model. The site modeled was the lowland, low energy regime catchment for the River Gjern in Denmark, with loamy to sandy soil and high agricultural activity. Their data show a marked increase in runoff to streams with an overall increase in the entire catchment and significantly different results between sub-catchments depending on their baseline for the runoff during a dry year. Those areas with previously low dry-year runoff showed an increase in runoff volume of more than 40%, with overall increases mainly occurring in the winter. These data have important implications for the nutrient transport of all similar watersheds, considering that the winter flushing period and runoff is the annual-scale event which enriches the basin’s ponds and lakes. Increased incidence of drought will also lead to lower base-flow streams and rivers in
the future, which were affected more than others in the modeled study by increased amplitude of
the flushing season (Anderson et al., 2006).

**Areas for Further Research**

There are two notable areas in hydrologic research on cyanobacterial onset and success
that could be improved by further research. The influence of dams, weirs, and river rerouting
were addressed briefly in the literature, usually as a secondary note regarding increased retention
time but not coming to an understanding of the complexities that forced flushing, forced
retention, and the influence of manmade wetlands or other diverted or isolated river sections
might have on cyanobacterial proliferation. The timing of flood to drought and the influences of
a high disturbance event followed by a period of dry, still, and nutrient-rich summer were a part
of many of the discharge-focused studies, but the relative effects of droughts and storms on
HABs on different spatial and temporal scales have not yet been fully elucidated.
CHAPTER 3

WHEN BLOOMS LIKE IT STORMY: SPATIOTEMPORAL DYNAMICS OF HARMFUL ALGAL BLOOMS IN FRESHWATER RESERVOIRS

ABSTRACT

Harmful algal blooms (HABs), especially those composed of cyanobacteria, are a global threat to eutrophic and mesotrophic freshwater reservoirs. While extreme hydrologic events such as flooding and drought are thought to control the onset and success of these blooms, the spatiotemporal dynamics of these relationships are still unclear. In this study hydrologic events and HABs in Lake Allatoona and Lake Lanier, Georgia were characterized using historical and satellite datasets from 2008 to 2017 and modeled using linear and multiple regression. Results showed that the impact of stormflow and rainfall events varied systematically from riverine to lacustrine reaches of the two reservoirs on weekly and monthly scales. Storm duration and positive drought indices were the most significant and best-fitting predictors of algal bloom biomass in deeper reaches of the two reservoirs, suggesting that HABs in more lacustrine environments are better equipped for wet and stormy regimes than has been previously hypothesized.
INTRODUCTION

Reservoirs are vital to freshwater resource management in the Southeast, especially the storage of water for municipal treatment and recreational use. Due to their often shallow and nutrient-rich waters, reservoirs are at risk of eutrophication and seasonal cyanobacterial harmful algal blooms (Heiskary, 1992). These cyanoHABs often pose water quality hazards, from hypoxia and fish kill to reduced biodiversity and toxicity (Paerl and Huisman, 2008; Magrann et al., 2012). The onset and success of these blooms in freshwater bodies are generally associated with local limnologic conditions such as high temperatures, plentiful bioavailable phosphorus, and long water residence times (Figure 3.1) (Schindler, 1977; Meis, 2009; Paerl et al., 2016). Other factors in cyanoHAB appearance include predation, solar radiation, basin, and lake morphology (Waylett et al., 2013; Maheaux et al., 2016; Wang and Chen, 2016). Regional conditions which may promote cyanoHABs include extreme hydrologic events such as severe or long-lasting drought or frequent, high-intensity storms (Padisak and Toth, 1988; Bouvy et al., 1999).

The well-established paradigm for hydrologically-instigated bloom events is the occurrence of drought or storms in a lake with temperature and nutrient conditions already conducive to cyanoHAB growth (Bloesch, 1995; Robson and Hamilton, 2003). Severe drought reduces inflow to a lake, increasing retention times and therefore thermal stratification of the water column. Prolonged drought may also increase nutrient availability, as drought can reduce water depth, allowing for seiche- and current-driven bottom sediment resuspension, nutrient desorption, and biological uptake (Hutchinson, 1957; Bloesch, 1995; Li et al., 2014). Orthophosphate, the primary freshwater limiting nutrient, is supplied to freshwater bodies by sediment transport during high-intensity storms, along with dissolved nitrate (Spieles and Mitch,
1999; Schindler, 2008). While these factors are recognized, there is still considerable uncertainty in the relationship between the blooms and the interaction between drought and precipitation, on temporal or spatial scales within a freshwater system (Reichwaldt and Ghadouani, 2012). The ideal lag time trigger between a storm and bloom event, or the importance of a hydrologic process as a function of the location within the lake, are dynamics which have yet to be fully characterized but are essential for management and prediction of cyanoHAB recurrence.

The current literature addressing the hydrologic drivers of cyanoHABs focus on precipitation and climate impacts on discharge and retention time (Jones and Elliot, 2007; Bowling et al., 2013; Robson and Hamilton, 2013), with several studies also branching into watershed lithology and dam and weir mechanics (Huang et al., 2014; Istvanovics, et al., 2014; Karan et al., 2014). The compounding effects of local nutrient sources, geomorphology, or temperature may even result in seemingly contradictory results in these studies, necessitating identification of the fundamental mechanics behind each relationship rather than simplified empirical models (Dzialowski et al., 2001; Reichwaldt and Ghadouani, 2012; Camacho et al., 2016). Discharge, as an example, has been modeled as a source of nutrient rejuvenation, through autochthonous and allochthonous processes (Royer et al., 2006; Ceballos, 2007; Glibert et al., 2008; Holbach et al., 2012; McCullough et al., 2012). In these scenarios, discharge promotes cyanoHAB growth, although the positive effect comes after a temporal lag (Bowling et al., 2013; Paerl and Otten, 2013). However, frequent flushing of the system may suppress cyanoHAB growth through colony dispersion, relocation to cooler or deeper waters, or oxygenation (Paerl et al., 2011; Lundgren et al., 2013; Zhang et al., 2016). Inversely, when discharge declines and the reservoir retention time increases, thermal stratification and an increase in cyanoHAB biomass may take place (Srifa et al., 2016). The relationship between long retention times and severe
blooms has been observed by computational models and in the field and is attributed to the increased time within optimal growth conditions (Elliot et al., 2009; Michalak et al., 2013; Longphuirt et al., 2016).

**Figure 3.1:** The ideal conditions for bloom onset within a lake or reservoir include a well-developed thermocline, long retention times, and plentiful orthophosphate.

The two Southeastern Piedmont reservoirs, Lake Allatoona and Lake Sidney Lanier, were chosen for the study as representative of the ideal limnologic conditions for the recurrence of cyanoHABs in reservoirs (Figure 3.2) (Zeng and Rasmussen, 2001; Ceballos, 2007; Radcliffe et al., 2009). Both locations experience recurring algal- and cyano- dominated blooms between May and October (Ceballos, 2007; Dirnberger, 2011). These reservoirs provide drinking water to Gwinnett and Cobb County, as well as flood control, hydroelectric power, and recreational services. Lake Allatoona is a 0.45 km$^3$ (367,000 acre-feet) reservoir created by a flood control dam at the intersection of Allatoona Creek and the Etowah River roughly 48 km northwest of Atlanta, Georgia (USACE, 1952). Allatoona is 27 km long but only 2 km wide at the widest point. Much of the eastern reaches of Lake Allatoona are shallow (4.5 – 6 m), dominated by the forces of the Etowah River tributary (Figure 3.2, Figure 3.4).
Lake Lanier is a 1.3 km$^3$ (1,050,000 acre-feet) reservoir above the Buford Dam and is directly northeast of Atlanta (USACE, 1959). Lanier is a larger reservoir though the surrounding watershed is smaller (2,694 km$^2$) than Allatoona (2,903 km$^2$). Both reservoirs are primarily in the Georgia Piedmont, with a fraction of Allatoona also lying in the Blue Ridge and Ridge and Valley regions (Figure 3.2) (Griffith et al., 2001). The climate within the province is temperate to subtropical, with temperatures varying from a winter average of 6° Celsius (43° Fahrenheit) to an average of 25° C (77° F) in the summer (NOAA, 2016). The average annual precipitation for the reservoirs is approximately 137 cm (54 inches), with the maximal rainfall in the winter months and sporadic July thunderstorms, and a warm dry period in the late summer (Kunkel et al., 2013; NOAA, 2016).

The Southeastern Piedmont has experienced several periods of meteorological drought in the past half century (Campana et al., 2012). In 2007, several months with little to no rain reduced the Lake Allatoona dam pool to a low of 238 m above sea level—18 m below the summer conservation level and more than 15 m below the decadal average (USGS 02393500; USACE, 1952). As a result of periodic low-flows and continuous nutrient influx to the reservoir, several significant algal blooms have occurred in the Allatoona reservoir during the past two decades (Dirnberger, 2005; Ceballos, 2007; Dirnberger, 2011). We compared the relative drivers of blooms within Lake Allatoona to those in Lake Lanier to infer regional trends and build a conceptual model of HABs within mesotrophic to eutrophic reservoirs.

The purpose of this study was to determine the relationship between hydrologic dynamics (precipitation intensity, frequency, and duration; drought severity, frequency, and duration) and the onset and propagation of HABs in a typical reservoir characterized by geochemical heterogeneity, varied river hydraulics, and a diverse phytoplanktonic community. We tested
whether long-term patterns of precipitation—especially storm intensity and frequency—and hydrological drought severity correlate with and drive HAB onset and severity in the Southeastern Piedmont Lakes Allatoona and Lanier. In-situ historic datasets, remote sensing techniques, and field data were integrated to maximize the study’s spatial and temporal resolution. These relationships were further investigated using dimensional reduction and linear regression on both in-situ and remotely sensed datasets from 2008 to 2017.

**Methods**

The hydrologic drivers of HABs considered in this study were drought, precipitation, and streamflow. The study locations were two reservoirs, Lake Allatoona and Lake Lanier, during the years 2008 to 2017. In-situ chlorophyll-α concentrations were obtained from the Georgia Environmental Protection Division (Georgia EPD), which collected water samples monthly from April to October for the years 2008-2014. Drought conditions at those locations were characterized using precipitation records from the USGS and NOAA, the resulting Standard Precipitation Index (SPI), and the NOAA Palmer Drought Severity Index (PDSI). Drought events were categorized by severity, duration, and intensity (Figure 3.3). Independent storm events were also identified and categorized by intensity, frequency, and duration. Daily, weekly, and monthly streamflow data were gathered from USGS records for the major tributaries of the two reservoirs (e.g., Etowah River, Little River, and Chattahoochee River; Figure 3.4). Physical and chemical water quality data gathered by the Georgia EPD were also obtained for the same duration (Figure 3.3). Harmful algal blooms were characterized by in-situ chlorophyll-α data as well as modeled chlorophyll-α and phycocyanin, derived from Landsat satellite scenes from 2008 to 2017. The purpose of these scenes was to increase the length and the density of the
cyanoHAB dataset. The photosynthetic pigment chlorophyll-α is found in all phytoplanktonic blooms, while phycocyanin is almost exclusively produced by cyanoHABs (Vincent et al., 2004). Continuous chlorophyll-α data were used as a proxy for algal biomass and were analyzed using ordinary least squares regression (OLSR), with both bivariate and multivariate models fit to the data.

Figure 3.2: Study location showing data sources and sites in the Etowah and Chattahoochee River Watersheds. Five sites within the Allatoona reservoir and nine sites in the Lanier reservoir were analyzed for spatiotemporal relationships between hydrologic drivers and harmful algal blooms.
Finally, discrete bloom events were characterized and compared to explanatory variables using Random Forest classification, a machine learning method which builds tree models to rank explanatory variables by importance (Breiman, 2001).

**Drought**

Drought is defined using a variety of qualitative and quantitative metrics which often cross disciplines (Campana et al., 2012). In this study drought was characterized using hydrological and meteorological indices, including the PDSI, Palmer Hydrologic Drought Index (PHDI), Palmer Z-Index (ZNDX), and Palmer Modified Drought Index (PMDI), and the SPI (Figures 3.7 and 3.8). These are all well-tested for representing the temporal variation in wet and dry periods at a location, as well as describing the magnitude of drought events, and are widely used by the National Oceanic and Atmospheric Administration (NOAA) and the US Departments of Commerce and Agriculture (Alley, 1984). The original PDSI classifies monthly weather conditions from >4, or extremely wet, to <-4, or extreme drought, using a water budget estimated from potential evapotranspiration, recharge, combined available moisture capacity, soil moisture loss, and runoff in a simulated soil bilayer (Alley, 1984). The results are then scaled based on historical temperature and precipitation for the region. While caution is necessary for using single indices for defining drought events, the indices are a widely recognized and robust metric when used in concert (Alley, 1984, Campana et al., 2012). The four Palmer Indices used in this study were obtained from the National Centers for Environmental Information (NCEI) for Georgia Climate Division II and cover January 2007 to October 2017.

The SPI was constructed for the two reservoirs using precipitation data from the nearest rain gauge with a 30-year or longer record (Figure 3.7). The strength of using the SPI is the ability to compare values between regions, as the index does not represent the absolute values of
precipitation, but rather a data-transformed probability of each month’s precipitation volume, based on historical norms (Guttman, 1999). The SPI also does not assume a Gaussian data distribution for precipitation values and is easily calculated for the time scale of interest (Guttman, 1999). The 3-, 6-, and 12-month SPI were constructed using data from Adairsville, near Lake Allatoona, and Gainesville, near Lake Lanier, using software obtained from the National Drought Mitigation Center (NDMC, 2017). A gamma distribution is fitted to the data by the software, obtaining monthly z-scores for precipitation which are then used to interpret drought (Edwards, 1997). Following McKee et al. (1993), a drought event was considered to have begun when the SPI first fell below -1, ending when the values next became positive.

Drought severity, intensity, and duration were obtained from the 12-month SPI datasets, which are more representative of hydrologic drought than the 3-month index, a proxy for soil moisture, or the 6-month index, which has characteristics of both timescales (Svoboda et al., 2012). Drought severity was defined as the absolute value of the cumulative area under the curve of the index and drought duration as the cumulative number of months since the event started. The drought intensity was defined as the running lowest SPI per event (Spinoni et al., 2014). SPI data for the region were also obtained on 1- to 24- month scales from the NCEI, but linear regression showed that the local SPI often explained a higher percent of the chlorophyll-α variance.

Precipitation

Precipitation data were obtained from the watersheds of the Etowah River, the main tributary to Lake Allatoona, and the Chattahoochee River, which forms the valley of Lake Lanier (Figure 3.4). Unit hydrographs were constructed for the upper catchments. The lag time for the
Figure 3.3: The overall workflow for the study progresses from dataset compilation and processing to comparison of all model parameters with cyanoHABs across a variety of spatial and temporal scales.

propagation of energy from observed storm events to the reservoirs was estimated to be three hours for Lake Allatoona, and approximately six hours for Lake Lanier. Independent storm events were also obtained with the separation method proposed by Restrepo-Posada and Eagleson (1982). A record of 15-minute precipitation depths was fitted with a gamma distribution, from which a Poisson cumulative distribution function (CDF) of storm arrivals was obtained using the bootstrap method. From the CDF, a coefficient of variation was obtained,
which represents the minimum period necessary to classify the next storm as independent. This method is sensitive to missing data, so the nearest rain-gauge with a continuous record of at least ten years was used for each reservoir’s catchment. Independent storm depth and duration were obtained, from which intensity (depth divided by duration), duration, and frequency (return period) were determined.

**Temperature and Water Quality**

Air temperature data were obtained for the two reservoirs from the USGS, and water quality parameters including alkalinity, turbidity, total phosphorus, total nitrogen, and Kjeldahl nitrogen were gathered from the Georgia EPD. Average depths in the reservoirs at each data site were obtained from NOAA Raster Navigational Charts.

**Algal Blooms**

CyanoHABs may be identified by lab analysis, or by in-situ fluorometric measurement of the absorption by their key pigment, phycocyanin (Mishra et al., 2009). In longitudinal studies, the pigment chlorophyll-α is commonly used as a proxy for algal bloom biomass, both for in-situ and remote sensing applications. If a relationship is established between cyanobacteria cell density and chlorophyll-α concentration within a freshwater body, quantitative characterization of the cyanoHAB may be used (Mishra and Mishra, 2012). In-situ chlorophyll-α data for both Allatoona and Lanier were obtained from the Georgia EPD. Measurements were collected monthly from 2008 to 2014 during the regional bloom season of April through October at designated sites within major reservoir embayments.
Figure 3.4: The major embayments of eastern Allatoona are formed by the Etowah River, Little River, and Kellogg Creek tributaries. A field survey of water quality was conducted at these three sites, shown above.

Landsat 5, 7, and 8 satellite data over the same time period were processed to obtain remotely-sensed chlorophyll-α and phycocyanin products on a broader spatial and temporal scale. The Thematic Mapper (TM, Landsat 5), Enhanced Thematic Mapper (ETM, Landsat 7), and Operational Land Imager (OLI, Landsat 8) products are 250 kilometer wide scenes with
visible, near-infrared, and shortwave infrared bands with a red/green/blue spatial resolution of 30 meters and an overpass at the site every 16 days. Using these satellite products allows observation of the development and size of a bloom and as the in-situ data represented late spring through summer only, the year-round record provided by the Landsat imagery allows observation of bloom conditions during atypical seasons. Cloud-free scenes were collected from 2008 to 2017 and corrected to minimize variations caused by weather, season, and instrument.

Atmospheric corrections were performed in ENVI, using band-by-band Dark Object Subtract following the method used in Vincent et al., 2004 (Figure 3.A.4). Each scene was radiometrically corrected per band to convert the product to surface reflectance. Finally, a calibration was applied to minimize differences in gain and offset between Landsat missions, using in-situ reflectance measurements of a pseudo-invariant feature within the scene (Ortiz et al., 2017). Using the resulting reflectance products, two pigment detection algorithms were applied. CyanoHAB biomass was estimated using a unique empirical approach developed in Lake Erie for modeling phycocyanin (Vincent et al., 2004). Cyanobacteria are differentiated from other phytoplankton using this distinguishing pigment; however, the spectral resolution of Landsat only allows for indirect detection of phycocyanin. The algorithm for the detection of chlorophyll-α was chosen from several similar models based on application similarity and performance, which used the near-infrared to red band ratio (Duan, 2007). Once each scene was processed, the data were extracted at each embayment using ArcGIS software and used as chlorophyll-α and phycocyanin indices. In addition to the continuous data, the presence or absence of a bloom was determined for each scene. The remotely sensed data and in-situ measurements did not fall on the same days within the study period of 2008 to 2017, so field measurements were taken to determine the relationship between satellite and in-situ data.
Figure 3.5: The sites at Lake Lanier are at the two main tributaries, the Chestatee River and the Chattahoochee River, as well as the several embayments formed by minor inflows.

The field study took place on Lake Allatoona in July and August of 2017 (Figure 3.6). On days of Landsat 8 overpass, coincident water quality and surface reflectance measurements were taken at the Etowah River, Kellogg Creek, and Little River embayment sites used by the Georgia EPD in the 2008 to 2014 survey (Figure 3.4). Surface water temperature, pH, dissolved oxygen,
and electro-conductivity were measured using a YSI 556 sonde probe, and Secchi disk depth and weather conditions were recorded. Turbidity was measured using a Hach 2100P turbidimeter, and in-situ hyperspectral surface reflectance was obtained using a GER-1500 handheld spectroradiometer. Water samples were collected using 1000 mL brown Nalgenes, labeled, and chilled on wet ice. Four field replicates from each site were vacuum filtered onto 47 mm Whatman GF/F glass microfiber filters, wrapped in foil, cataloged, and frozen on dry ice, resulting in two samples each for phycocyanin and chlorophyll-α pigments. The samples were then shipped for laboratory analysis by high-performance liquid chromatography following the methods of Pinckney et al., (1996) and Cotti-Rausch et al., (2016). The concentrations of phycocyanin were below detection limits, and so were not calculated from the field water samples. These pigment data were then compared with satellite and field spectra and a relationship was obtained between the chlorophyll-α and other parameters at Lake Allatoona.

Modeling

Quantitative data from historical, remotely-sensed and in-situ datasets for Lake Lanier and Lake Allatoona were checked for distribution, outliers, and kurtosis. Variance in water quality, residence time, and nutrient supply is often high between arms of a dendritic reservoir, and the mean chlorophyll-α was found to be statistically different depending on the collection site (Welch’s Heteroscedastic F=14.2, DF = 4 and 91.9, p<<0.05), so each embayment was treated as a separate sub-lake for analysis. The resulting 14 embayments—five from Allatoona and nine from Lanier—each had in-situ and remotely sensed proxies for algal biomass including historical chlorophyll-α concentrations, the chlorophyll-α index, and the phycocyanin index. The independent variables representing drought, storms, and flow each had several spatiotemporal
Figure 3.6: Allatoona water quality, from top left, counterclockwise. Open water near the Etowah River site; two HABs on the Little River arm of the reservoir; Little River tributary upriver of Lake Allatoona. Lower left photograph by James Tullius, 2017, used with permission.

scales, so sensitivity analysis was performed to remove redundant inputs. In-situ and remotely-sensed chlorophyll-α were modeled as a function of each explanatory variable using ordinary least squares regression (OLSR), and those parameters which explained at least 10% of the variance in any case were preserved. Additional dimensional reduction was performed in R using Random Forest regression (Liaw and Wiener, 2002; R Core Team, 2018). The best individual parameter predictors for each site were chosen for further modeling using multiple regression.

For each site, the data subset was first checked for quadratic and higher-order relationships, complex interactions between predictor variables, and overparametrization. A
generalized additive model was first used to fit non-parametric linear regression smoothers to each predictor variable and the response variable to visually test whether a significant nonlinear relationship existed. A tree model was fitted to the data to visualize the relationship structure and to check for complex interactions between variables. The model strength for each embayment was compared to those obtained from bivariate modeling.

In addition to continuous data, remotely-sensed data were categorized by bloom presence. The bloom-present scenes were then compared to bloom-absent scenes using Welch’s Heteroscedastic F test to determine whether there was a statistically significant difference in explanatory variables during the remotely sensed bloom events.

**Results**

**Dimensional Reduction**

Of the hydrological, physical, and chemical parameters analyzed, several showed meaningful relationships with HAB biomass during dimensional reduction. Random Forest modeling was first used to explore the relative importance of several independent continuous variables on the levels of chlorophyll-α pigment concentration in the Etowah River and Little River sites on Lake Allatoona. At both sites, the parameters which grew tree models with the lowest mean error were inorganic nitrogen and 12-Month SPI, while turbidity and air temperature could be left out with minimal error increase. We considered the results of the Random Forest trees and the bivariate models in combination and set aside the most impactful parameters for choosing best-fit models for each site.
Linear Modeling of In-Situ Data

Stormflow

While total discharge into Lake Allatoona and Lake Lanier from their main tributaries did not show strong relationships with HAB biomass, flow above a baseflow value was used as an explanatory variable in several models in both reservoirs (Table 3.1). We defined stormflow in the Etowah Watershed as flow exceeding 28 m$^3$ s$^{-1}$ (1000 cfs, cubic feet per second) in the Etowah River and 4 m$^3$ s$^{-1}$ (141 cfs) in the Little River tributaries of Lake Allatoona (USGS 2391860, USGS 2392780). A unit hydrograph constructed near Canton, Georgia showed a steep rising limb and long falling limb, and a lag time of three hours between peak measured rainfall and peak discharge. Stormflow conditions in the Chattahoochee River at the mouth of Lake Lanier were defined as discharge measurements beyond 21 m$^3$ s$^{-1}$ (740 cfs, USGS 02331600).

Phytoplankton biomass showed a positive linear relationship with stormflow at several embayments, on both 7-day and 30-day scales, with spatial variation in the strength of the relationship (Table 3.1, Figure 3.9). Lake Allatoona’s eastern Little River and Etowah River embayments are relatively shallow and riverine, while the Allatoona Creek, Kellogg Creek, and Dam Pool are deeper (Figure 3.2). During stormflow events, algae did not show a strong relationship with weekly average discharge (WAQ, Table 3.1) in the tributary-dominated eastern sites, while the western Allatoona Creek showed a majority of phytoplankton variation may be explained by the Etowah’s inflow volume (Figure 3.2, Table 3.1). The statistical significance and model performance of stormflow on phytoplankton biomass as a function of stormflow increase steadily with increasing distance from the Etowah inflow ($r^2=0.06$, $r^2=0.12$, $r^2=0.27$, $r^2=0.25$, $r^2=0.56$; Table 3.1).
Table 3.1: Multiple regression models by site show a variety of significant linear relationships, ranging between 12% and 74% variance explained. Notable explanatory variables include 3 Month SPI, 12 Month SPI, PDSI, Alkalinity, Temperature, Temperature Index (TI), Weekly Average Stormflow (WAQ), Monthly Average Stormflow (MAQ), Monthly Maximum Storm Duration (MXD), and Monthly Minimum Storm Duration (MND).

<table>
<thead>
<tr>
<th>Site</th>
<th>Driver *</th>
<th>Number of Variables</th>
<th>Model</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
<th>F</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etowah River</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= -0.20 \times 12$ M SPI $+ 1.81$</td>
<td>0.13</td>
<td>-</td>
<td>5.6</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>Little River</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= -0.29 \times 12$ M SPI $+ 1.81$</td>
<td>0.18</td>
<td>0.13</td>
<td>3.6</td>
<td>2.34</td>
<td></td>
</tr>
<tr>
<td>Kellogg Creek</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= -0.20$ 12 M SPI $+ 2.26$</td>
<td>0.13</td>
<td>-</td>
<td>2.3</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>Allatoona Dam</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= -0.07$ (Temp $+$ Alk) $- 0.27$ (SPI) $- 0.71$</td>
<td>0.22</td>
<td>0.18</td>
<td>2.0</td>
<td>2.36</td>
<td></td>
</tr>
<tr>
<td>Allatoona Creek</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 0.54$ ln (PDSI) $+ 0.57$ ln (WAQ) $- 6.73$</td>
<td>0.63</td>
<td>0.56</td>
<td>7.4</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Bolling Bridge</td>
<td>S</td>
<td>3</td>
<td>ln (Ch-a) $= 0.46$ MD$^2$ $- 0.97$ MD $- 0.41$ 13 M SPI $- 0.44$</td>
<td>0.50</td>
<td>0.41</td>
<td>5.4</td>
<td>3.16</td>
<td></td>
</tr>
<tr>
<td>Lanier Bridge</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 0.37$ 3 M SPI $- 0.19$</td>
<td>0.45</td>
<td>-</td>
<td>14.5</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Balsas Creek</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 0.44$ ln (MND) $+ 1.44$</td>
<td>0.51</td>
<td>-</td>
<td>11.4</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>Mud Creek</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 2.73$ ln (MAQ) $- 7.71$</td>
<td>0.74</td>
<td>-</td>
<td>17.2</td>
<td>1.6</td>
<td></td>
</tr>
</tbody>
</table>

All Regression Models by Site: Lake Allatoona

<table>
<thead>
<tr>
<th>Site</th>
<th>Driver *</th>
<th>Number of Variables</th>
<th>Model</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
<th>F</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etowah River</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= -0.34$ ln (Little River WNQ) $+ 3.36$</td>
<td>0.15</td>
<td>-</td>
<td>6.5</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>Little River</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= 0.02$ 12 M SPI $+ 1.81$</td>
<td>0.13</td>
<td>-</td>
<td>5.6</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>1</td>
<td>ln (Ch-a) $= 0.55$ ln (WAQ) $- 1.10$</td>
<td>0.06</td>
<td>-</td>
<td>0.6</td>
<td>1.9</td>
<td>0.45</td>
</tr>
<tr>
<td>Little River</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= -0.29$ 12 M SPI $+ 1.81$</td>
<td>0.18</td>
<td>0.13</td>
<td>3.6</td>
<td>2.34</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>1</td>
<td>ln (Ch-a) $= 0.45$ ln (WAQ) $- 0.43$</td>
<td>0.12</td>
<td>-</td>
<td>12.74</td>
<td>1.9</td>
<td>0.29</td>
</tr>
<tr>
<td>Kellogg Creek</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= 0.16$ 12 M SPI $+ 2.26$</td>
<td>0.13</td>
<td>-</td>
<td>5.3</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>1</td>
<td>ln (Ch-a) $= 0.90$ ln (WAQ) $- 2.13$</td>
<td>0.27</td>
<td>-</td>
<td>3.0</td>
<td>1.9</td>
<td>0.10</td>
</tr>
<tr>
<td>Allatoona Dam</td>
<td>D</td>
<td>1</td>
<td>ln (Ch-a) $= -0.16$ 12 M SPI $+ 1.12$</td>
<td>0.12</td>
<td>-</td>
<td>4.0</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>1</td>
<td>ln (Ch-a) $= 0.98$ ln (WAQ) $- 2.24$</td>
<td>0.25</td>
<td>-</td>
<td>3.0</td>
<td>1.9</td>
<td>0.12</td>
</tr>
<tr>
<td>Allatoona Creek</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 0.54$ ln (PDSI) $+ 0.97$ ln (WAQ) $- 6.73$</td>
<td>0.65</td>
<td>0.56</td>
<td>7.4</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 1.04$ ln (WAQ) $- 2.72$</td>
<td>0.56</td>
<td>-</td>
<td>11.3</td>
<td>1.9</td>
<td></td>
</tr>
</tbody>
</table>

All Regression Models by Site: Lake Lanier

<table>
<thead>
<tr>
<th>Site</th>
<th>Driver *</th>
<th>Number of Variables</th>
<th>Model</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
<th>F</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bolling Bridge</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 0.46$ MD$^2$ $- 0.97$ MD $- 0.43$ 13 M SPI $+ 0.44$</td>
<td>0.50</td>
<td>0.41</td>
<td>5.4</td>
<td>3.16</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.37$ 3 M SPI $- 0.19$</td>
<td>0.31</td>
<td>-</td>
<td>8.0</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.26$ NOAA 1 M SPI $+ 0.19$</td>
<td>0.27</td>
<td>-</td>
<td>6.5</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.11$ ZNDX $+ 1.51$</td>
<td>0.27</td>
<td>-</td>
<td>8.6</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.25$ NOAA 3 M SPI $+ 0.22$</td>
<td>0.20</td>
<td>-</td>
<td>4.6</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.11$ PDSI $+ 1.66$</td>
<td>0.19</td>
<td>-</td>
<td>4.2</td>
<td>1.18</td>
<td>0.06</td>
</tr>
<tr>
<td>Lanier Bridge</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 0.04$ TI $+ 4.30$</td>
<td>0.46</td>
<td>-</td>
<td>15.3</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.37$ 3 M SPI $- 0.19$</td>
<td>0.45</td>
<td>-</td>
<td>14.5</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.24$ NOAA 3 M SPI $+ 0.55$</td>
<td>0.26</td>
<td>-</td>
<td>6.4</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.18$ NOAA 2 M SPI $+ 0.86$</td>
<td>0.23</td>
<td>-</td>
<td>5.3</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Balsas Creek</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 0.44$ ln (MND) $- 1.44$</td>
<td>0.51</td>
<td>-</td>
<td>11.4</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.23$ NOAA 1 M SPI $+ 0.12$</td>
<td>0.27</td>
<td>-</td>
<td>5.5</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>2</td>
<td>ln (Ch-a) $= 0.29$ MXQ $- 0.74$</td>
<td>0.21</td>
<td>-</td>
<td>4.0</td>
<td>1.15</td>
<td>0.06</td>
</tr>
<tr>
<td>Mud Creek</td>
<td>S</td>
<td>2</td>
<td>ln (Ch-a) $= 2.73$ ln (MAQ) $- 7.71$</td>
<td>0.74</td>
<td>-</td>
<td>17.2</td>
<td>1.6</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Best-Fit Regression Models by Site, In-Situ Data

Statistical significance: * $p < 0.05$  ** $p < 0.01$  *** $p < 0.001$  

a D. Drought  S. Storms

b F-statistic  c Degrees of freedom
Table 3.2: Multiple regression models of remotely sensed Chl-α and PC data by site show a variety of significant linear relationships, ranging between 18% and 91% variance explained. Best-fitting explanatory variables include Daily Average Stormflow (DAQ), Drought Intensity, Monthly Average Storm Intensity (MAI), Little River Daily Average Stormflow (Little River DAQ), Drought Severity, and Monthly Total Storm Intensity (MTI). Chl-α and PC values were modeled using empirical models (Vincent et al., 2004; Duan et al., 2007).

<table>
<thead>
<tr>
<th>Site</th>
<th>Driver</th>
<th>Model</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
<th>$F$</th>
<th>DF$^a$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etowah River</td>
<td>D</td>
<td>In (PC) = -0.0017 DAQ + 4.52</td>
<td>0.43</td>
<td>- 9.2</td>
<td>1.12</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Little River</td>
<td>S</td>
<td>In (PC) = 0.13 MAI + 2.92</td>
<td>0.41</td>
<td>- 7.6</td>
<td>1.11</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Kellogg Creek</td>
<td>D</td>
<td>In (PC) = -0.012 Little River DAQ + 4.92</td>
<td>0.62</td>
<td>- 13.2</td>
<td>1.8</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Allatoona Dam</td>
<td>D</td>
<td>In (Chl-α) = 0.59 Drought I + 0.23</td>
<td>0.27</td>
<td>- 8.9</td>
<td>1.24</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Allatoona Creek</td>
<td>D</td>
<td>In (PC) = 0.038 Drought S + 0.56</td>
<td>0.91</td>
<td>- 29.0</td>
<td>1.3</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Boiling Bridge</td>
<td>S</td>
<td>In (Chl-α) = 0.001 MTI + 0.85</td>
<td>0.40</td>
<td>- 6.0</td>
<td>1.9</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Mud Creek</td>
<td>S</td>
<td>In (PC) = -35.00 Drought I + 31.47</td>
<td>0.24</td>
<td>- 6.5</td>
<td>1.21</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

All Regression Models by Site: Lake Allatoona

<table>
<thead>
<tr>
<th>Site</th>
<th>Driver</th>
<th>Model</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
<th>$F$</th>
<th>DF$^a$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etowah River</td>
<td>D</td>
<td>In (PC) = -0.0017 DAQ + 4.52</td>
<td>0.43</td>
<td>- 9.2</td>
<td>1.12</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Little River</td>
<td>S</td>
<td>In (PC) = 0.13 MAI + 2.92</td>
<td>0.41</td>
<td>- 7.6</td>
<td>1.11</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Kellogg Creek</td>
<td>D</td>
<td>In (PC) = -0.012 Little River DAQ + 4.92</td>
<td>0.62</td>
<td>- 13.2</td>
<td>1.8</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Allatoona Dam</td>
<td>D</td>
<td>In (Chl-α) = 0.59 Drought I + 0.23</td>
<td>0.27</td>
<td>- 8.9</td>
<td>1.24</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Allatoona Creek</td>
<td>D</td>
<td>In (PC) = 0.038 Drought S + 0.56</td>
<td>0.91</td>
<td>- 29.0</td>
<td>1.3</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Boiling Bridge</td>
<td>S</td>
<td>In (Chl-α) = 0.001 MTI + 0.85</td>
<td>0.40</td>
<td>- 6.0</td>
<td>1.9</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Mud Creek</td>
<td>S</td>
<td>In (PC) = -35.00 Drought I + 31.47</td>
<td>0.24</td>
<td>- 6.5</td>
<td>1.21</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

All Regression Models by Site: Lake Lantier

<table>
<thead>
<tr>
<th>Site</th>
<th>Driver</th>
<th>Model</th>
<th>$R^2$</th>
<th>Adj. $R^2$</th>
<th>$F$</th>
<th>DF$^a$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boiling Bridge</td>
<td>S</td>
<td>In (Chl-α) = 0.081 MTI + 0.85</td>
<td>0.40</td>
<td>- 6.0</td>
<td>1.9</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Mud Creek</td>
<td>S</td>
<td>In (PC) = -35.00 Drought I + 31.47</td>
<td>0.24</td>
<td>- 6.5</td>
<td>1.21</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Statistical significance: * < 0.05    ** < 0.01    *** < 0.001

$^a$ F-statistic  
$^b$ Degrees of freedom
Lake Lanier’s phytoplankton communities showed positive linear relationships with the average daily, weekly, and monthly Chattahoochee River discharge in several embayments from the north-most and riverine Chattahoochee site to the stream-fed and urban embayments near the city of Gainesville (Figure 3.9). Several sites within Lake Lanier showed relationships between higher chlorophyll-α and monthly average stormflow (MAQ), although only the Mud Creek relationship is statistically significant ($r^2 = 0.74$; Table 3.1). Many sites were best modeled using linear OLSR, except the Allatoona Dam, Allatoona Creek, and Bolling Bridge sites (Table 3.1).

**Storms**

The kinetic energy of the peak discharge events originates with recent storm events within the watershed and propagates towards the mouth of the reservoir, bringing nutrients and disrupting the water column stability (Bloesch, 1995). We characterized these storms using intensity (inches/day), duration (decimal days), and frequency (return period, years) on 7- and 30-day scales. While the return period or intensity of storm events did not show statistically significant relationships with in-situ chlorophyll-α data in this study period, storm duration was an important driver of phytoplanktonic growth at both Bolling Bridge and Balus Creek within Lake Lanier (Table 3.1).

Bolling Bridge, at the mouth of the Chestatee River, showed a quadratic relationship with maximum storm duration on a 30-day scale, with curvature in the duration term confirmed using a generalized additive model. At this site, the link between algal biomass, storms, and stormflow was fit with the model:

$$\text{Ln} (\text{Chl} - \alpha) = 0.46 (\text{MXD}^2) - 0.97 (\text{MXD}) + 0.41 (3 \text{M SPI}) - 0.44 \quad \text{Eq. 3.1}$$
The independent variables are Monthly Maximum Storm Duration (MXD) and 3-Month SPI (3M SPI). This multivariate model was significant (p=0.009) and explained 50% of the variation within our data, however, we also fit a simpler, but still adequate model to the data (r²=0.31, Table 3.1).

Drought

Drought was defined by a suite of indices such as the regional PDSI and several related or modified indices (MPDSI, PDHI, ZNDX), as well as local SPI built from nearby precipitation gauges for each reservoir. Drought severity (cumulative integral area of the index), duration (months), and drought intensity (cumulative maximum severity) were defined and compared to in-situ chlorophyll-α data within each reservoir but did not show significant interaction during our study period. The indices themselves—both SPI and PDSI—were more closely related to chlorophyll-α at both Allatoona and Lanier. In Lake Allatoona, the Etowah River and the Allatoona Dam Pool showed a slight negative relationship with meteorological drought, with increased chlorophyll-α production during long-term precipitation shortages as modeled by 12-Month SPI (r²=0.13, adj. r²=0.18; Table 3.1). A positive relationship with short-term meteorological drought was observed in Lake Lanier at both major tributary sites (r²=0.45, r²=0.41).

Hydrologic drought, modeled by the local SPI, was a major driver of phytoplankton growth in four reaches of Lake Allatoona (Table 3.1). In addition, we obtained a good fit when modeling phytoplankton as a function of both higher PDSI and weekly average stormflow at the Allatoona Creek embayment (r²=0.65, adj. r²=0.56; Table 3.1). While seven of the nine best-fit models were driven in part by drought indices, only four of these showed a positive response to increasing dryness (more negative index).
Figure 3.7: The Standard Precipitation Index for Lake Allatoona and Lake Lanier show several droughts within the last decade, notably from 2007-2009.

 Temperature and Chemistry

Air temperature was poorly linked to phytoplanktonic mass in Lake Lanier, with only two out of nine sites exhibiting subtle negative relationships at the joining of the Chattahoochee and Chestatee Rivers with the reservoir. Chlorophyll-α is statistically related to the ecoregional Temperature Index at Lanier Bridge; however, the effect size of the index is small ($r^2=0.46$, Table 3.1). Chlorophyll-α did not show a linear relationship with the air temperature at Lake Allatoona.
Figure 3.8: Standard Precipitation Index, precipitation, drought severity, and drought duration for Lake Allatoona and Lake Lanier during the study period. The two reservoirs share a climate region, but Lake Allatoona has experienced more prolonged and extreme drought in the last decade.
Water quality parameters such as total nitrogen, total phosphorus, and turbidity did not correlate with the concurrent chlorophyll-α concentrations, although alkalinity did show a possible role in phytoplankton biomass at the Allatoona Dam Pool as a factor in a complex interaction (adj. $r^2 = 0.18$, Table 3.1). Note that precedent water quality conditions were not obtained in this study.

**Linear Modeling of Remotely Sensed Data**

Remotely sensed chlorophyll-α and phycocyanin were obtained from modeled indices by reach or embayment within the two reservoirs. Lake Allatoona was subdivided into the Etowah River, Little River, Kellogg Creek, Allatoona Dam Pool, and Allatoona Creek reaches (Figures 3.2 and 3.4). Lake Lanier was divided into the Bolling Bridge, Lanier Bridge, Browns Bridge, Flowery Branch, and Buford Dam Pool reaches. The values for the two pigment indices were extracted only from pixels at least 5 m from the shoreline and were compared to stormflow and storm and drought dynamics. Bivariate models fitted to the data showed significant linear relations in six of the ten reaches, and in most cases agreed with the models using the in-situ chlorophyll-α dataset (Table 3.1, Table 3.2). Bloom presence did not show a statistically significant relationship with hydrologic drivers, and so only continuous data were used for modeling.

Both storms and drought were significant predictors of remotely sensed chlorophyll-α and phycocyanin within Allatoona and Lanier (Figure 3.9). In Lake Allatoona, phycocyanin may be predicted using average storm intensity on a 30-day scale, resulting in a positive relationship with an $r^2$ of 0.41 (Table 3.2). Chlorophyll-α in Lake Lanier also showed an increase in predicted concentration with an increase in total storm intensity with an $r^2$ of 0.40. Drought was an even better predictor of both remotely sensed chlorophyll-α and phycocyanin, with positive
relationships between drought and HABs at all sites in Lake Allatoona except for the Etowah River reach (Figure 3.9). At Allatoona Creek, predicted phycocyanin index increased the most per increased drought intensity ($r^2 = 0.85$, Table 3.2), as did chlorophyll-α, though with a lesser fit ($r^2 = 0.21$). A departure from these results, predicted phycocyanin at Lake Lanier’s Mud Creek embayment decreased significantly with increased drought intensity ($r^2=0.24$).

**Discussion**

**Lake Allatoona Models**

The conditions which promote HAB growth, especially one which is primarily cyanobacterial, are nutrients readily available within the water column, long residence times leading to a strong thermocline, and plentiful solar radiation to provide the energy for a burst of photosynthesis and pigment production (Figure 3.1; Schindler, 1977; Heiskary, 1992; Paerl and Huisman, 2008). The volume of nutrients, especially phosphate, within a freshwater system is driven by high-kinetic events and so our hypothesized system should derive nutrients primarily from short, intense rainfall (Delpla et al., 2011). We also hypothesized that greater biomass would require the water column stability that comes from long, severe drought. However, our results reflect a system driven by peaks in tributary stormflow, long-lasting rainfall, and by wetter regimes rather than long-term drought in many of the reservoir sites (Table 3.1). An example of these is seen in the Allatoona Creek branch of Lake Allatoona, where an increase in the PDSI and related increase in rainfall was associated with higher chlorophyll-α concentrations (Table 3.1). This positive relationship between stormflow and phytoplankton growth contrasts with the decreased biomass with decreased retention time described in several other reservoirs (An and Jones, 2000; Jones and Elliott, 2007; Zhang et al., 2016). In the case of one study with
high temporal resolution, blooms were reported to decrease after storm events, and increase after approximately five days (Ahn et al., 2002). While our resolution is not daily, our results may also be explained by a storm event which triggers a bloom on a lag of five days, but we also see a correlation within months with higher stormflow and greater storm duration, often spread between several events, and higher phytoplankton biomass (Table 3.1). This suggests that there may be systematic mechanisms which counteract disruption of the water column.

The response of blooms to stormflow in Lake Allatoona is spatially explicit, with increasing strength in the relationship with increasing distance from the tributary mouth (Figure 3.9). The tributaries in the east where the Etowah and the Little River join the reservoir show little response to high peaks of discharge, and while the effect size and significant increase westward, only the far Allatoona Creek, south of the dam, shows a statistically significant relationship to the Etowah inflow volumes (for relative distances, Figure 3.2). The seemingly uncorrelated inflow-water quality relationship at the eastern reaches may be a result of coupled and opposing processes. Increased flow leads to renewed allochthonous and autochthonous loading but the same turbulent forces increase sediment suspension, blocking inbound sunlight (Ceballos, 2007). The disruption of chlorophyll-α production at the riverine reach of the Allatoona is consistent with the rapid peak rainfall to peak discharge time in the Etowah Watershed. More distant from the source, the Allatoona Creek branch may be benefiting from the same rainfall events and increased nutrient runoff without the non-algal turbidity or thermal mixing caused by the Etowah River.

Among the in-situ dataset, no well-fitted or significant relationships were found between storm dynamics and Lake Allatoona. This may not be a true result but rather a consequence of
Figure 3.9: The variance-explained of each significant linear model is shown by site. The spatial variance of the relationship strengths appears to be a function of distance from the main inflow site at each reservoir. Note that $r^2$ is shown here without the direction of the slope.

...the event discretization for storm identification which resulted in a small sample size for the period 2008 – 2014. In the Lake Lanier dataset, the larger number of sampling sites (Figure 3.2) resulted in a greater sample size for comparison with the limited storm dataset and was resolved into statistically significant models (Table 3.1). As the relationships to storm dynamics in Lake...
Lanier follow those with stormflow, it is reasonable to assume that the Allatoona storm relationships would also show increased driver importance with increased distance from the tributary (Figure 3.9).

**Lake Lanier Models**

The chlorophyll-α concentrations in several Lanier embayments are driven by storm duration on a monthly scale (Table 3.1, Table 3.2). However, at Lanier, the phytoplankton’s affinity for storm events is not confined to cove-shaped regions. Chestatee River, the smaller of two tributaries at the head of Lanier, also presents higher chlorophyll-α concentrations during lengthy storms, provided that their duration exceeds a threshold at the Bolling Bridge site (Figure 3.9). This suggests the flushing hydraulic force attributed to storm duration is not enough to disrupt the levels of free-floating algae and bacteria in even this seemingly riverine stretch of the reservoir, potentially due to low hydraulic velocity in comparison to the mouth of the Etowah. The depth at Bolling Bridge is also approximately 27 m, contrasted with the average 8 m of water at the Etowah site. If density differences between the river and reservoir are high enough, whether by suspended load or temperature differential, a current may be formed (Middleton, 1993). The inflow at Lanier might then be too deep a current to disrupt the epilimnion, while still causing a spike in nutrients.

The data are correlated both with increased storm duration but also increased drought indices (wetter regimes), though only related to drought duration, intensity, or severity within remotely sensed datasets (Table 3.1, Table 3.2). The index is likely the more robust parameter as it represents a condition which does not change as quickly as duration or severity, both cumulative values, or change as sporadically as intensity, a measurement of the highest severity a drought event had achieved to the collection date. The indices are good predictors of variation in
chlorophyll-α in both reservoirs, but as with stormflow, there are notable differences from tributary mouth to the far reach (Figure 3.9). In Allatoona, drier conditions are related to higher biomass from the Etowah to the dam pool sites. In contrast, wetter conditions with lower evapotranspiration, as modeled by PDSI, are related to biomass within the remaining sites. This relationship is likely driven by the reduced disruption of the epilimnion under drought conditions. The tributary site on Lanier’s Chattahoochee River, likely because of its depth, is consistent with the more lacustrine Allatoona sites, and higher SPI values were accompanied with high algal biomass. High SPI was also a significant predictor in the multivariate model fitted to the smaller tributary, Chestatee River, at the Bolling Bridge site (Figure 3.2). These results suggest that a wetter regime dominated by multi-day precipitation events is a driver of HAB success at Lake Lanier. This is especially the case in sites protected from turbidity and mixing disruptions at the source, whether by the vertical depth to the inflow current or by the lateral distance along the reservoir length (Figure 3.9).

Remote Sensing Models

We employed remote sensing to expand the data past the Georgia EPD’s in-situ campaign. While the chlorophyll-α concentrations measured at the lake were taken during the typical bloom season of April through October, satellite data were obtained for the entire span of the year, from 2008 through 2017. The dataset size was limited by cloud and wind contamination; however, several interesting and statistically significant results were still observed. Both chlorophyll-α, a proxy for all phytoplankton, and phycocyanin, a pigment nearly exclusive to cyanobacteria, were observed to have a positive relationship with drought at the Allatoona Creek site, counter to the in-situ data.
One consideration in the modeled phycocyanin data is the use of the empirical Vincent equation, which was originally developed at Lake Erie for use in eutrophic waters. We considered the use of several algorithms, including empirical ones, to be valuable as an additional line of experimentation alongside in-situ methods, but where a model disagrees between in-situ chlorophyll-α and modeled water quality data from remote sensing, the in-situ results have a higher degree of certainty.

A seasonality factor may be responsible for the difference in the relationship. The remotely sensed dataset expanded the observation of the phytoplankton to include the winter season, during which the majority of the Georgia Piedmont experiences the highest precipitation per month and heavy sediment transport from the watershed uplands to the lake basins (NCDC, 2016). While sporadic rains might encourage summer blooms, winter conditions drop lake temperatures and inhibit HAB growth (Paerl and Huismann, 2008).

These data were carefully used but there are several limitations and sources of error to consider. The analysis covered seven years of in-situ and nine years of Landsat satellite data. While the satellite data were normalized from scene to scene to reduce residuals from instrument aging, and the Georgia EPD data were collected using systematic procedures, additional heterogeneity might have been introduced through a change in land use and smog levels in the region. Data loss through cloud or sun glint contamination also restricted the sample size. However, long-term studies are still valuable despite the complexities, as precipitation regimes through decadal time scales are increasingly connected to water quality in our rivers, lakes, and reservoirs.
Other Model Parameters

While several other physical, chemical, and biological parameters (e.g. alkalinity, turbidity, and fecal coliform) within the reservoir were compared with the in-situ phytoplankton, no well-fitted or significant models were produced outside of the hydrologic factors. Future models would be more likely to resolve linear relationships between chlorophyll-α and pH, water temperature, and dissolved oxygen as a function of depth.

Model Implications

We modeled HAB growth as a function of stormflow, drought severity, intensity, and duration, and storm intensity and duration on multiple spatiotemporal scales. These drivers control HAB onset and propagation through the two main mechanisms of nutrient inflow, which promotes growth, and water column disruption, which inhibits growth. In Lake Allatoona, both positive and negative relationships with increased precipitation are seen, depending on the site location. We anticipated that increased storm intensity would promote HAB growth given sufficient distance from the main tributary at the Etowah River. When comparing Lake Allatoona to Lake Lanier, we expected to see similar results, with drought-driven HAB growth near the tributary mouth and storm-driven HAB growth further downriver. However, every Lake Lanier site showed a positive relationship with increased storm events and wetter regimes. The difference between the riverine reaches near Allatoona’s tributary and that of Lake Lanier is attributed to the depth of the water column. We therefore hypothesize that the controlling factor between storm-driven HABs and drought-driven HABs is the depth of the site, rather than the distance from the tributary. This hypothesis could be tested in the future using a combination of residence time analysis and vertical in-situ measurement of physical and chemical characteristics at several of the sites. We hypothesize that future study will show that at shallower, more
riverine sites, vertical water stability is more likely to be disrupted by storm events, while at
dereper, lacustrine sites, the stormwater currents may slip under the epilimnion and leave the
photic zone relatively undisturbed. While the apparent exceptions to this hypothesis are the
Kellogg Creek site and the deep but drought-driven Allatoona dam pool site, the Allatoona Dam
releases occur at the upper layer of the pool, resulting in more disruption during times of high
rainfall. These results reinforce a paradigm in which depth controls the relative importance of
hydrologic drivers of HAB growth in temperate reservoirs.

**Conclusion**

- The hypothesized precipitation and drought dynamics important to the
  phytoplankton biomass within reservoirs were high storm intensity, low storm
duration, and high drought severity characterized by a negative PDSI and SPI.
- The results show that while hydrologic processes are significant drivers of high
  chlorophyll-α, wetter climatic conditions and longer-lasting rainfall events are the
  better predictor in many of the sites, rather than dry climatic periods and short,
  intense storms.
- In Lake Allatoona’s shallower reaches, drought is a better predictor of HAB
  biomass, while the HABs in the deeper sites of Lake Lanier show more growth
  with increased storm duration.
- The control of depth over the hydrologic drivers of HABs may be a result of the
  vertical separation of inflow from the photic zone, which will be less disrupted at
  a site with a deeper water column.
- Our results have implications for the effectiveness of flushing for HAB mitigation within mesotrophic and eutrophic reservoirs.

**Acknowledgments**

The authors wish to acknowledge Dr. Deepak Mishra, Dr. Richard Becker, Dr. Elizabeth Kramer, Abhishek Kumar, Sam Weber, and Dr. Susan Wilde, Dr. Dayton Wilde; Dr. Koray Yilmaz, Stephanie Fulton, Doog Becker, the Little River Marina, and the Water Resources & Remote Sensing lab for field support; Dr. John Dowd and Dr. Todd Rasmussen; and Dr. Tammi Richardson and Eric Lachenmyer for HPLC laboratory analysis.

This research was funded by the 2017 Geological Society of America Student Research Grant and 2017 Watt’s Wheeler Research Grant.

**References**


42. NDMC, 2017, National Drought Mitigation Center, University of Nebraska, drought.unl.edu.


Figure 3.A.1: In-situ and satellite remote sensing coverage of algal bloom proxy data, from 2008 to 2017.
Figure 3.A.2: Bivariate plots showing the spread and slope of several of the high-performance chlorophyll-α models.
Figure 3.A.3: HABs were observed in satellite photography at 31 locations along the shoreline at Lake Allatoona (Google Earth, October 2012). These images were captured seven days before a Landsat scene showing high levels of chlorophyll-α throughout the eastern reaches of the reservoir (top right).
CHAPTER 4

REMOTE SENSING OF CYANOHABS IN SOUTHEASTERN RESERVOIRS: A COMPARISON OF METHODS

Abstract

Cyanobacterial harmful algal blooms (cyanoHABs) are an increasing concern to resource managers and water resource scientists alike. CyanoHABs are frequently studied using remote sensing techniques, through which data may be processed rapidly, but with challenges in spectral, spatial, and temporal resolutions. These difficulties may be mitigated by using several algorithms in conjunction and comparing results and confidences. In this study, we resolve bloom events over a period of nine years through a combination of several satellite-based techniques and historical in-situ data. Three bloom events were identified within the Lake Allatoona scenes and one in the Lake Lanier scenes from 2008 – 2017. Chlorophyll-α and phycocyanin indices tested showed agreement with field chlorophyll-α measurements ($r^2 = 0.85$, $r^2 = 0.78$). Compared approaches include three atmospheric correction methods and three algorithm families. This combined approach may reduce uncertainty in remote based studies of cyanoHABs in reservoirs and small water bodies, especially useful in the case of sparse historical ground data.
**Introduction**

The remote sensing of cyanobacterial harmful algal blooms (cyanoHABs) has been a goal for water resource scientists for decades, as a tool which would allow for high spatiotemporal resolution in research as well as rapid detection and mitigation (Dekker, 1993; Gomez et al., 2011; Agha et al., 2012; Frolov et al, 2013; Li et al, 2014). Satellite monitoring of freshwater bodies has been practiced and evaluated since the launch of the first Landsat (Ritchie et al., 1976; Frazier, 2000). Dozens of satellite missions have produced data which have been used to detect freshwater cyanoHABs, including the Landsat series, MERIS and MODIS, Sentinel, Hyperion, and Aster (Matthews, 2011; Ogashawara et al., 2013).

Direct detection of cyanoHABs relies on the photosynthetic pigment phycocyanin (PC), produced primarily in cyanobacteria (Dekker, 1993). Lacustrine environments host a broad array of phytoplankton and therefore detection of photosynthetic bodies is not sufficient unless a monoculture can be verified (Aguirre-Gomez et al., 2001). PC has a signature absorption feature at 620 nm; however, in the case of low PC concentrations, the presence of colored dissolved matter (CDM) or suspended particulate matter (SPM) may complicate the estimation of PC from the spectral data (Simis et al., 2007; Hunter et al., 2008). Nonetheless, the challenge has been taken up by many researchers. MERIS and other instruments with the spectral resolution to directly detect the absorption and reflectance of chlorophyll-α and PC are therefore highly valuable to cyanoHAB sensing, and MERIS-like sensor data is widely used alongside field data for calibration and validation of empirical or semi-empirical models such as the PC3 and nested band-ratio models (Simis et al., 2007; Mishra et al., 2009).

However, many satellites which have high spatial resolution (Sentinel 2, Landsat 8) do not have equipment which can directly sense PC (Vincent et al., 2004). Instead, algorithms are
often developed which indirectly estimate or proxy the presence and concentration of the blooms. The accuracy of these methods depends heavily on sensor technology, correction methods, sample size, and individual lake biogeochemical properties, such as PC to chlorophyll-α (chl-α) ratio and CDM concentrations (Hunter et al., 2010, Li et al., 2015).

Several of these techniques are also location-sensitive; while these algorithms may be developed for broad geospatial application, each algorithm and each inland body may have different linear relationships to actual cyanobacterial cell concentration (Stumpf et al., 2016). It is valuable, then, to compare methodologies using a single dataset and region, especially in the case that the satellite data can be compared to in-situ PC and chlorophyll-α data.

After sensor choice, the algorithm used to model PC and chlorophyll-α will have the greatest impact on algal bloom identification and differentiation. There are few algorithms designed for truly universal use among freshwater cyanoHAB monitoring, and the number of algorithms to choose from is currently rapidly growing (Ogashawara et al., 2013). Even with distinctive, mathematically predictable absorption and reflection features, chlorophyll-α and PC are not typically directly modeled using a single band (Duan et al., 2007). As a result of the spectral complexity of turbid water, single spectral features are rarely robust enough to explain the variation in measured pigments, whether from the interference of CDM and SPM (also known as gelbstoff and tripton, respectively) or overlapping chlorophyll-α and PC absorption (Dekker and Peters, 1993; Kubiak et al., 2016). A variety of approaches have been developed pigment algorithms, from site-based empirical fits to spectral shape models, described here as broad families. The most well-known and widely used methods can be categorized as empirical, semi-empirical, semi-analytical, and analytical algorithms (Dekker, 1993; Ogashawara et al., 2013).
Preprocessing is key to error reduction when using historical, reservoir-scale satellite data. The spatial resolution, with dendritic and often small surface areas, is susceptible to cloud and wind (sun glint) contamination and thus each scene is a valuable sample of data. The satellite data provided from the Landsat missions are available preprocessed with geometric corrections, which may be verified in ArcGIS or similar software. Radiometric and atmospheric corrections vary depending on the index or algorithm chosen but should be used consistently across all scenes to reduce heterogeneity (Wynne and Stumpf, 2015). Absolute corrections seek the values as measured at the surface, but there are also relative methods, which may increase within-scene error, but reduce among-scene variability (Coppin and Bauer, 1994).

The study sites were the reservoirs Allatoona and Lanier, which are mesotrophic to eutrophic and host several varieties of cyanobacteria alongside diatoms, dinoflagellates and green algae during the cyanoHAB season of May through October (EPA, 2009). These reservoirs exhibit typical dendritic shorelines, with highest concentrations of blooms within these enclaves. The primary aim of this study is to evaluate several methodologies for the detection and characterization of cyanoHABs in these two Piedmont reservoirs. The relative performance of these algorithms in this environment would be of significant use to water resource managers, who rely on these bodies of freshwater which spatially and temporally range from high-clarity to extremely turbid (Dirnberger, 2011; Zeng and Rasmussen, 2005).

In-situ data collected by the Georgia Environmental Protection Division are available from 2008-2014 and include chl-α concentration and water quality parameters. Chl-α and PC sampling and ground-sensed remote sensing reflectance were performed on pass-over dates for Landsat 8 and Sentinel 2 sensors in 2017 in order to calibrate or evaluate both the set of applicable algorithms. With these 2008 – 2014 data, we first performed spatiotemporal
difference analysis on processed scenes in ArcGIS. The 2008 - 2014 scenes were secondly compared to in-situ chl-α data, in order to confirm the presence of phytoplankton. Thirdly, the scenes from 2017 were processed through each algorithm and then compared to reflectance data and pigment analyses from field sampling.

Table 4.1: The primary productivity and turbidity at the two reservoirs shows a range consistent with mesotrophic to eutrophic conditions.

Table 4.1: Site Characteristics

<table>
<thead>
<tr>
<th>Reservoir</th>
<th>Site Name</th>
<th>Historical Chl-α , µg/L</th>
<th>Measured Chl-α , µg/L</th>
<th>Turbidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allatoona</td>
<td>Etowah River</td>
<td>0.6 – 25.0</td>
<td>2.0 – 4.5</td>
<td>3.75 – 5</td>
</tr>
<tr>
<td></td>
<td>Little River*</td>
<td>2.2 – 25.0</td>
<td>2.8 – 7.1</td>
<td>3 – 3.25</td>
</tr>
<tr>
<td></td>
<td>Kellogg Creek</td>
<td>1.1 – 14.0</td>
<td>1.2 – 4.0</td>
<td>5.5 - 7</td>
</tr>
<tr>
<td></td>
<td>Allatoona Dam</td>
<td>1.1 – 13.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Allatoona Creek</td>
<td>1.3 – 16.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lanier</td>
<td>Little River**</td>
<td>2.6 – 24.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Boling Bridge</td>
<td>2.1 – 12.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Lanier Bridge</td>
<td>0.5 – 15.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Brown’s Bridge</td>
<td>1.4 – 11.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Flat Creek</td>
<td>2.1 – 9.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Balus Creek</td>
<td>1.6 – 7.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Mud Creek</td>
<td>0.4 – 6.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Flowery Branch</td>
<td>0.5 – 8.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Six Mile Creek</td>
<td>0.6 – 8.9</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* Little River at Allatoona  **West Fork Little River at Lanier
Figure 4.1: The remote sensing scientist’s choice of sensor, preprocessing, and algorithms is a balance between greater accessibility (ease-of-use, public access) and greater fidelity (spectral resolution, lower RMSE).

**Methods**

**Study Site**

The study sites were two southeastern Piedmont reservoirs, Lake Allatoona and Lake Sidney Lanier. The morphology of the reservoir basins is derived from former river valleys within metasedimentary and metaigneous bedrock terranes, and as such, both reservoirs have highly developed shorelines (Kesler, 1950). Both reservoirs range from mesotrophic to eutrophic states (Table 4.1; EPD, 2004; Zeng and Rasmussen, 2005; Ceballos, 2007). The phytoplankton of Lake Allatoona have been identified as belonging to several divisions, including the
Chrysophyta (golden algae), Chlorophyta (green algae), Euglenophyta (flagellates), Pyrrhophyta (diatoms), and the Cyanophyta (cyanobacteria), including the *Oscillatoria* and *Anabaena* genera (EPA, 2009).

**Verification and Characterization**

First-order verification of bloom presence is currently possible through both photogrammetry and soft data sources (Kumar and Bhandarkar, 2017). When a likely bloom is identified, satellite data may be reviewed and ground validation undertaken. The value of water quality report from social media, citizen science platforms, or high resolution photogrammetry (Google Earth Pro, Google Earth Engine) is the ease and cost-effectiveness, from which the most promising case studies may be selected for rigorous analysis, monitoring, and field testing (Leeuw et al., 2013; Seltenrich, 2014; Boddula and Mishra, 2015). Both smartphone application and web-based platforms for cyanoHAB reports such as CyanoTracker, CyanoScope, and BloomWatch are valuable resources for determining where active blooms may be observed (Lee et al., 2016; Scott et al., 2016). In this study, literature describing algal and cyanoHAB occurrence at Lake Allatoona and Lake Lanier were combined with reports of water quality trends (Dirnberger and Weinberger, 2005; Ceballos, 2007; Radcliffe, 2009; Dirnberger, 2011).

Additional characterization of bloom occurrence was obtained from Google Earth Pro historical imagery from 1993-2017. Each available orthoimage was scanned for surface scum features and non-algal turbidity events (Figure 4.8).

**Sensors, Algorithms, and Corrections**

Once the study site was chosen, the remote sensing methodology was designed. For the highly developed shores of Lake Allatoona and Lake Lanier, spatial resolutions of less than 50 m
per pixel are unlikely to capture unmixed pixels within the shallow embayments where blooms initially form. The sensors aboard the Landsat and Sentinel missions obtain data between 10 and 60 m per pixel within the visible, near infrared, and shortwave infrared wavelengths. Additionally, Landsat missions 5, 7, and 8 provide temporal coverage of the sites every 16 days from 1984 to the present. The modern Landsat bands are often used to develop empirical models of chlorophyll-α and the cyanobacterial pigment PC, which are reparametrized for each study site through calibration and validation measurements (Matthews, 2011).
Landsat data from 2008 to 2017 were obtained for the two reservoirs, spanning the Landsat Thematic Mapper (TM), Enhanced Thematic Mapper (ETM), and Operational Land Imager (OLI) datasets. The Landsat 7 ETM sensor’s Scan Line Corrector (SLC) suffered a mechanical failure in 2003, resulting in partial loss of data in each scene along regular EW trending bands. These missing data affect central and western Lake Allatoona as well as the entirety of Lake Lanier. TM and OLI data were therefore prioritized to minimize sampling bias introduced by the SLC data gaps. Scenes were further culled when contaminated by clouds, haze, or sun glint. The final scene sample consisted of 30 dates for Lake Allatoona and 21 dates for Lake Lanier, evenly distributed throughout the study period (Figure 4.4).

Atmospheric corrections were applied to each scene to reduce among-scene variability and improve quantitative estimations of cyanoHAB pigments. Several atmospheric processing techniques were compared across scenes and with field spectra, but the final dataset was corrected using the DS histogram method (Chavez, 1988). Histograms of the digital numbers for each band within each scene were obtained in ENVI, and the dark object was taken to be the band-specific values below the first slope change.

Initially, surface reflectance products were obtained through the ENVI radiometric calibration function, but after the automated tool returned negative values, radiometric corrections were reprocessed by manual calculation (Equation 4.1). While a manual workflow may introduce human errors due to repetitive tasks, the assumptions and settings within automated processes may be poorly documented or difficult to debug.

\[ \rho_\lambda = \frac{(\pi \cdot L_\lambda \cdot d^2)}{(ESUN_\lambda \cdot \sin \alpha)} \]

*Eqn. 4.1*
Table 4.2: Selected algorithms from the main algorithm families including their model performance from the original study.

<table>
<thead>
<tr>
<th>Algorithm Family</th>
<th>Algorithm Structure</th>
<th>Name (Citation)</th>
<th>Sensor</th>
<th>Training Range (μg/L)</th>
<th>Model Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical</td>
<td>PC = −9.21(B3/B1) + 29.7(B4/B3) – 118(B4/B3) – 6.81(B5/B3) – 41.9(B7/B3) – 14.7(B7/B4) + 47.7</td>
<td>Vincent, 2004**</td>
<td>Landsat ETM</td>
<td>7 – 20</td>
<td>0.78 ~26</td>
</tr>
<tr>
<td></td>
<td>ln(Chl-α) = -1.7(B1/B3) + 9.65</td>
<td>Brezonik et al., 2005*</td>
<td>Landsat TM</td>
<td>2 – 279</td>
<td>0.88 -</td>
</tr>
<tr>
<td></td>
<td>Chl-α = 925(B5/B4) – 77.157</td>
<td>Watanabe et al., 2015*</td>
<td>Landsat OLI</td>
<td>1 – 500</td>
<td>0.75 -</td>
</tr>
<tr>
<td>Semi-Empirical</td>
<td>Chl-α ∝ B4/B3</td>
<td>Duan et al., 2007**</td>
<td>Landsat TM</td>
<td>6 – 58</td>
<td>0.69 -</td>
</tr>
<tr>
<td></td>
<td>PC ∝ Rs(709)/Rs(620)</td>
<td>Simis et al., 2005, 2007; adapted by Ogashawara et al., 2013</td>
<td>MERIS</td>
<td>8 – 36; 68 – 857</td>
<td>0.77 -</td>
</tr>
<tr>
<td>Semi-Analytical</td>
<td>FAI = Rrs(859) – [Rrs(645) + (Rrs(645) + (Rrs(1640) – Rrs(645))(859 – 645)(1640 – 645))]</td>
<td>FAI: Hu, 2009 **</td>
<td>MODIS</td>
<td>-</td>
<td>- -</td>
</tr>
<tr>
<td></td>
<td>Chl-α ∝ [Rrs(708) – Rrs(665)] / [Rrs(708) + Rrs(665)]</td>
<td>NDCI1: Mishra and Mishra, 2012</td>
<td>MODIS, Sentinel 2</td>
<td>14 – 28</td>
<td>0.93 4.83</td>
</tr>
<tr>
<td></td>
<td>Chl-α = 86.115(NDCI1) + 194.325(NDCI1)^2 + 14.039</td>
<td>NDCI4: Page et al., 2018*</td>
<td>Landsat OLI, Sentinel 2</td>
<td>15 – 55</td>
<td>- -</td>
</tr>
<tr>
<td></td>
<td>PC ∝ [Rrs−1(615)−Rrs−1(600)] × Rrs(725)</td>
<td>Hunter et al., 2010</td>
<td>CASI-2, AISA</td>
<td>5 – 10; 27 – 36</td>
<td>0.832 29.8</td>
</tr>
<tr>
<td>Inversion</td>
<td>PC ∝ [Rrs−1(629) – Rrs−1(659)] × Rrs(724)</td>
<td>PC3: Mishra et al., 2009</td>
<td>MERIS</td>
<td>68 – 3032</td>
<td>0.99 -</td>
</tr>
<tr>
<td></td>
<td>Li et al., 2013</td>
<td>AISA</td>
<td>0 – 300</td>
<td>0.93 -</td>
<td></td>
</tr>
</tbody>
</table>

**Algorithms assessed at Lake Allatoona, Georgia
*Other algorithms appropriate for small or dendritic reservoirs

Rrs – remote sensing reflectance
Rrc – Rayleigh corrected reflectance
Figure 4.3: The methodology used in this study. Methods in white were only compared and were not used in the final products.

Figure 4.4: Sample distribution of the Landsat scenes from 2008 to 2017. No scenes from 2012 were cloud-free over Lake Lanier.
After radiometric and atmospheric correction, cross-generational corrections were applied to standardize all scenes from 2008 to 2017. The goal was not to process each scene for individual error reduction as much as time series normalization so that equivalent comparisons could be drawn in the analysis. Each scene was not processed for individual error reduction as much as for normalization across the time series, by selecting and using a pseudo invariant feature (PIF) within the scene (Schott et al., 1988; Ortiz et al., 2017). Blooms were identified using the Duan and Vincent algorithms and closest historical water quality measurement. Only one of the overpass dates corresponded to water sampling dates during the 2008-2014 Georgia EPD field campaign, but six dates fell within a three-day margin. Chlorophyll-α index data were obtained from the two products using single pixels; 10-pixel averages; and maximum, average, range, and standard deviations within the spatial subset. A cross-sensor calibration for a coincident overpass between Landsat 8 and Sentinel 2B was attempted to obtain NDCI data for Landsat data following Page et al (2018), but the calibration was not completed due to contamination by sun glint.

**Bloom Identification**

Blooms within reservoirs may be easily detectable surface scums as in the hypereutrophic Lake Taihu, or may consist of cells dispersed throughout the photic zone. In the latter case, which is the most commonly found in Lake Allatoona and Lake Lanier, bloom geometry and spatial development may be difficult to obtain from remote sensing, and the distinction between bloom and clear water must be made a priori using state guidelines or percent normal statistical definitions. Chlorophyll-α concentrations of 20 μg/L or more for calibrated data, or an index more than 2 standard deviations from the mean in uncalibrated data were set as thresholds for bloom presence.
The resulting remote-sensing chlorophyll-α and phycocyanin data were brought into R and compared to water quality and hydrologic parameters. Using both the remote and in-situ datasets, linear and multiple linear regression models were developed and relative effect sizes of the various drivers of the bloom events were ranked. Following the choice of sensors, algorithms, and correction methods, ground-truthing or calibration-validation procedures are necessary to relate satellite based observations with quantitative cyanoHAB estimation.

**Calibration and Validation**

In this study, field data were collected from Lake Allatoona at three EPD sites from the 2008-2014 water quality survey (Figure 4.2). The sites were positioned on reservoir reaches controlled by the Etowah River and Little River tributaries, as well as the lesser Kellogg Creek. Water quality data were measured and site characteristics were recorded (Table 4.1; Green et al., 2015). At each site, three samples and 3 replicates were collected from the top 20 centimeters of water and stored in brown Nalgene flasks for pigment analysis by high performance liquid chromatography (HPLC), following the EPD Standard Operating Procedure for collection of chlorophyll-α (Cutcliff, 2008). The samples were immediately labeled and stored on wet ice. Samples were then gently filtered by vacuum pump through GF/F filters and stored in labeled foil on dry ice. On return from the field, samples were transferred to -80 degree freezer until shipped to the external laboratory. Pigment analysis was performed by HPLC according to Pickney et al., with modifications to the preparation according to Cotti-Rausch (Pickney et al., 1996; Cotti-Rausch et al., 2016). The resulting chlorophyll-α concentration data was then compared to field spectra and remote sensing datasets and calibrated to a linear best-fit model.

Field spectra were processed by the NDCI and compared alongside the FAI, the Duan et al. Chlorophyll Index, and the Vincent et al. Phycocyanin Index, which were then modeled as a
function of chlorophyll-α in four calibration ordinary least-squares regression models (Figure 4.5, Vincent et al., 2004; Duan et al., 2007; Hu et al., 2009). The indices were also compared to each other using linear regression (Figure 4.6).

**Error Analysis**

Regression model error was evaluated by determination coefficient ($r^2$) and root mean square error (RMSE). The resulting model performance, seen in Table 4.3, was used to compare the FAI, Duan et al., and Vincent algorithms for their efficacy in mesotrophic Piedmont reservoirs (Vincent et al., 2004; Duan et al., 2007; Hu et al., 2009).

**Results**

**Atmospheric Correction**

There are many atmospheric correction methods available, the most common of which are dark subtraction and radiative transfer models such as FLAASH. We compared field spectra to coincident scenes corrected using several of these methods and compared their performance by spectral shape, $r^2$, RMSE, and level of correction (Figure 4.10, Table 4.4). The atmospheric correction method chosen for this study was the histogram Dark Subtract method, which was used in the original Vincent et al. study (2004). This method retrieved reasonable reflectances for scenes with no cloud cover, but overcorrected scenes with partial cloud cover such as the scene of ground-truthed data (Figure 4.10, Table 4.4).

**Calibration**

The field spectra at the three tributary sites show a typical, low-PC cyanoHAB and mixed algal signature, as do the in-situ chlorophyll-α measurements. The PC samples were below
Figure 4.5: Modeled chlorophyll-α and phycocyanin show a positive linear relationship with measured chlorophyll-α concentrations. The FAI shows a negative linear relationship with the same chlorophyll-α, consistent with reservoir-scale trends showing that FAI does not act as a proxy for chlorophyll-α within these at-depth algal blooms.
Figure 4.6: Chlorophyll-α concentrations measured at Lake Allatoona on July 17 and July 24, 2017 show agreement with coincident field spectra. Field spectra taken on August 18, 2017 show greater variability when modeled as a function of chlorophyll-α; this increased error is likely a result of the small sample size.

detection limits for HPLC analysis, but chlorophyll-α concentrations were within the range often seen in historical data for the two reservoirs. Chlorophyll-α data were compared with remote sensing data processed through three algorithms, and show positive relationships with the Vincent and Duan indices (Figure 4.4), but not with the FAI data. This negative relationship seen between the FAI and the measured chlorophyll-α is consistent with trends seen between the FAI and historical water quality data (Figure 4.A.1). Field spectra were also taken at the three tributary sites, and resampled to simulated Landsat bands. These data were then processed using
the NDCI and compared to the in-situ chlorophyll-α, resulting in linear best-fits with a positive slope ($r^2 = 0.89$, $r^2 = 0.91$, $r^2 = 0.14$; Figure 4.5).

Figure 4.7: From top, a cloudless true color image of the Etowah River, the Little River, and the Kellogg Creek reaches of Lake Allatoona, and coincident chlorophyll-α, floating algae, and PC index products.
Bloom Phenology

The greatest strength of remote sensing cyanobacterial Harmful Algal Blooms (cyanoHABs) is the ability to model spatiotemporal development of a bloom and quantify the area, concentration, or cell counts. In this study, we were able to define the approximate edges of a bloom based on modeled chlorophyll-α concentration from calibrated, ground-truthed data using the Duan and Vincent algorithms applied to scenes at both reservoirs. The blooms did not produce defined surface scum during our field measurements or during the study period, and so while we used the Floating Algal Index, confidence in the resulting bloom locations and extents is not high. Blooms consistently occurred within the warm season of May through October, and were observed in the deep waters rather than the embayments (Figure 4.7). Data compiled using areal statistics showed similar trends, but with smoother values (Figure 4.8).

Figure 4.8: Modeled chlorophyll-α index based on the areal maximum of Duan algorithm agree with the bloom event in late summer of 2012, and show an overall smoother trend.
Comparison with Historical Ground Data

The secondary set of data was obtained from a survey of water quality from 2008 – 2017 (Table 4.1). These data included blooms identified based on a minimum concentration of chlorophyll-α, and align with the remotely sensed blooms in several cases. The most noticeable agreement between datasets, which did not have many coincident dates, is October 31, 2012. On this date, chlorophyll-α and phycocyanin are modeled throughout the eastern reaches of Lake Allatoona using the two best performing algorithms. Supplementary evidence from a Google Earth orthoimage shows surface-scum forming blooms on October 23, 2012, just one week before. The blooms may be observed in 31 separate coves within the reservoir and match the color, shape, and locations of cyanoHABs observed in the field on August 18, 2017 and are consistent with in-situ chlorophyll-α measurements from October 22, 2012 at the Etowah River and Little River sites (Table 4.1).

Table 4.3: The three tested algorithms for modeling harmful algal blooms were compared with in-situ chlorophyll-α data.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Reference</th>
<th>Average $R^2$</th>
<th>RMSE</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floating Algal Index</td>
<td>Hu et al., 2009</td>
<td>0.40</td>
<td>4.68</td>
<td>-</td>
</tr>
<tr>
<td>Phycocyanin Index</td>
<td>Vincent et al., 2004</td>
<td>0.86</td>
<td>7.58</td>
<td>+</td>
</tr>
<tr>
<td>Chlorophyll-α Index</td>
<td>Duan et al., 2007</td>
<td>0.87</td>
<td>3.81</td>
<td>+</td>
</tr>
</tbody>
</table>
Discussion

Comparison of Sensor Choice

The first challenge in a satellite-based remote sensing study is the prioritization between spatial, spectral, temporal, and radiometric resolution. Spatial resolution is one of the most important considerations when working with reservoirs, as the dendritic shape leads to higher “mixed pixels” per scene than a glacial lake. High spatial resolution is necessary in studies of agricultural ponds, rivers, and reservoirs, where narrow targets and dendritic shorelines may cause difficulty in differentiating between water and land dominated pixels (Zhang et al., 2014). So-called medium resolution satellite sensors such as the Landsat TM, ETM, and OLI (28.5 m per pixel) or the Sentinel 2 series (30 m per pixel) are able to resolve major rivers, ponds, and most reservoirs, while the widely used low-resolution MODIS or Sentinel 3 data (250 m, 300 m per pixel) result in only one or two unmixed pixels per site in the cases of Lake Allatoona and Lake Lanier.

The importance of spatial resolution goes beyond static sample size. During seasonal changes in depth, reservoirs often experience drastic shoreline shifts. For the most robust selection of these water boundaries during image analysis, an unsupervised classification may be used in the case of medium resolution data. Precise boundary selection reduces mixed pixels, bottom-sediment reflectance, and the effects of shadows from nearby buildings or trees.

Spectral resolution is the second greatest limiting factor in multispectral remote sensing. While sensor technology is improving in each sphere, the choice is often still between satellite pixel size and band coverage, as in the Sentinel mission series. Sentinel 2 offers ten to 20 m resolution in the visible and near infrared wavelengths, but cannot directly detect the 620 nm PC absorption feature used in cyanoHAB monitoring. Sentinel 3 bears the Ocean and Land Color
Figure 4.9: Modeled phyocyanin in Lake Allatoona through time (Vincent et al., 2004).
Instrument (OLCI), with full resolution of 300 m and 21 bands, including a 10 nm band centered on 620 nm. Landsat has been used since first launch for water quality monitoring, and has been employed in many applications for monitoring cyanoHABs in particular (Vincent et al., 2004; Matthews, 2011; Isenstein et al., 2014; Watanabe et al., 2015). While the bands do not directly measure cyanoHAB pigment absorption or reflection features, algorithms have been developed and calibrated to leverage the relatively high spatial resolution, trading some taxonomic differentiation for better imaging of the development and spatial extents of the bloom features (Vincent et al., 2004). These indirect spectral models do have an effect on the signal to noise ratio for detection of PC, chl-α, CDM, and SPM. A study of hyperspectral and multispectral discrimination between phytoplankton communities used principal component analysis (PCA) showed that multispectral sensors distinguish the groups poorly in the presence of CDM and SPM, commonly in high concentration in eutrophic to hypertrophic reservoirs (Hunter et al., 2008). Downscaling of spectral resolution has primarily been achieved using extensive field surveys with hyperspectral imaging, such as with the Airborne Imaging Spectrometer for Applications (AISA) (Li et al., 2010).

A third dynamic to consider in satellite data is the temporal resolution, both of the return period over a single site and the total length of the mission. Longitudinal studies in cyanoHAB development, a key to developing robust conceptual models of the system, may require years of data, while studies in the development and degradation of a single bloom event benefit from high pass-over density, increasing the ability to observe subtle changes through time in a fast-forming bloom (Matthews et al., 2010; Lunetta et al., 2015; Matthews and Bernard, 2015; Urquhart et al., 2017). MERIS and MODIS have extremely high temporal resolution, while the Landsat missions pass over a site once over 16 days, or only about 23 times per year. The Sentinel 2 mission, with
sensors 2A and 2B in staggered orbits, are designed to provide a medium temporal resolution, with an interval as short as 5 days. Temporal range or length of previous data acquisition is an important factor for historical studies. Landsat is the only satellite which combines the long-running data with moderate spatial resolution, with data still in use from the TM mission beginning in 1984. The freely available and well-documented Landsat instruments are widely tested and calibrated across generations by the use of updated gains and offsets. MODIS is similarly long-lived, with data available from 1999 to present. Temporal resolution also includes the days between overpass at a single site. Of the freely available satellite datasets, MODIS has the densest temporal coverage, with passes every 2 days for most locations. The Sentinel 2 missions also provide superior overpass coverage, returning to a site every ten days in the United States, and approximately every five days in Europe. Sentinel 2A and Sentinel 2B are the first in a series of satellites with identical payloads, whose tandem flight paths are responsible for their imaging efficiency. While the widely used medium-resolution Landsat missions 7 and 8 have staggered 16-day orbits, the ETM sensor’s 2003 Scan Line Corrector failure means that Landsat 7 data is viable only for select sites in the center of the study site swath.

Radiometric resolution is the sensor’s color depth, or sensitivity to fine variations in reflectance caused by physicochemical differences in water quality (Richards, 1999). In studies of mesotrophic to eutrophic reservoirs, the increase in dynamic range from Landsat 5 and 7’s 8 bit data to Landsat 8 and Sentinel 2’s higher resolution (16 and 12 bit, respectively) may allow for bloom detection at the beginning of an event, during a low-concentration event, or in the case of increased depth of the bloom in the water column (Townshend et al., 1991).
Figure 4.10: The spectral shape retrieved using various atmospheric corrections may be either dampened or exaggerated. From top, methods are FLAASH, Dark Subtract Region of Interest, C2RCC, and Dark Subtract Histogram.

Comparison of Correction Methods

While preprocessing is the most difficult step in a remote sensing study, appropriate choice of corrections methods and programs is fundamental to obtaining a low-error final product (Jensen, 2009). Corrections generally begin with radiometric and atmospheric preprocessing, after which temporal corrections and artefact removal may be necessary.

Atmospheric corrections which are commonly used range in complexity from Dark
Subtraction (DS) to black-box algorithms such as the neural-network driven Case 2 Regional / Coast Color (C2RCC) (Vincent, 1972; Brockmann et al., 2016). Absolute atmospheric correction methods include image-based correction, radiative transfer modeling, and the empirical line method (Chavez, 1996; Adler-Golden, 1998; Farrand et al., 1994). The DS method assumes homogenous atmospheric effect over the scene, and that a black body exists within the scene (Chavez, 1988). This black body or “dark object” may be identified by the operator by knowledge of the area, such as where the water is deep and calm, or selected from the center of a cloud’s shadow. Any remotely sensed reflectance from this dark object is then assumed to be a result of the uniform Rayleigh scatter and is removed from each pixel. While the method yields variable accuracy, the accessibility of this technique gives it broad utility (Campbell, 1993). The assumptions made in radiative transfer modeling such as Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) or Second Simulation of a Satellite Signal in the Solar Spectrum (6S) include homogenous, spherical particles with a uniform distribution and a flat, Lambertian planetary surface (Adler-Golden et al., 1998). The FLAASH program offers a range of parametrization, including optional inputs for topography and adjacent-pixel scattering (Kruse, 2004). The Empirical Line Method (ELM) is a third absolute correction derived from scene-specific field spectra. ELM assumes a linear relationship between top-of-atmosphere (TOA) reflectance and surface reflectance across the scene, which can be modeled using field spectra across a range of reflectance values near the target site (Smith and Milton, 1999; Ortiz et al., 2017).

The final correction method, using the hypothetical Pseudo-Invariant Feature (PIF), is relative rather than absolute and is most applicable for historical or long-running studies (Schott et al., 1988). While absolute methods for atmospheric correction are useful for quantitative
analysis of specific events, they may complicate long-term change detection (Schroeder et al., 2006). In a comparative study, it was found that absolute correction increased spectral heterogeneity relative to the raw data, likely due to model assumptions of linearity (Schroeder et al., 2006).

Table 4.4: FLAASH, C2RCC, the DS Histogram method, and the DS Cloud Shadow (Region of Interest) method were compared by their qualitative and quantitative performance.

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference</th>
<th>$R^2$</th>
<th>% MRMSE</th>
<th>Shape</th>
<th>Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>-</td>
<td>0.17</td>
<td>0.078</td>
<td>Poor</td>
<td>Under</td>
</tr>
<tr>
<td>FLAASH</td>
<td>Adler-Golden, 1998</td>
<td>0.53</td>
<td>0.035</td>
<td>Good</td>
<td>Under</td>
</tr>
<tr>
<td>C2RCC</td>
<td>Brockmann et al., 2016</td>
<td>0.95</td>
<td>0.005</td>
<td>Good</td>
<td>Under</td>
</tr>
<tr>
<td>DS Histogram</td>
<td>Chavez, 1996; Vincent, 2004</td>
<td>0.38</td>
<td>0.086</td>
<td>OK</td>
<td>Over</td>
</tr>
<tr>
<td>DS Cloud Shadow</td>
<td>Chavez, 1996</td>
<td>0.53</td>
<td>0.021</td>
<td>OK</td>
<td>Variable</td>
</tr>
</tbody>
</table>

**Bold results indicate best performance in test**

The conceptual model behind a PIF correction is that if the scene contains a near-Lambertian surface, measurements of the surface may be taken to provide a baseline for comparing and correcting satellite data. The surface reflectance at the PIF is assumed constant through time. Although few truly invariant surfaces exist, parking lots and building rooftops are often used as PIF target locations (Hadjimitsis et al., 2009). The surface should be chosen to cover more than a single pixel within the scene, and the area covered by that pixel should be measured using a regular grid by handheld spectroradiometry near the sensor overpass time to remain close to the sky and solar angle conditions in the satellite imagery (Ortiz et al., 2017). Sensor-specific variance, such as the radiometric resolution of Landsat TM/ETM+ and Landsat
OLI or the shifts in gains and bias over time, are reduced when applying the PIF or other relative correction methods (Schott et al., 1988). The sensors, algorithms, and correction techniques may be combined in several ways to leverage modern computational power against the challenges of historical or reservoir-based cyanoHAB studies. After designing a remote sensing approach which best serves the chosen study site, data collection, calibration, and validation may be completed with confidence.

Several methods for radiometric and atmospheric correction and cyanoHAB modeling were tested and compared. Radiometric corrections were summarily compared, and after negative reflectance values were obtained using the automated ENVI radiometric correction tool, the final product was manually retrieved from DNs using the Band Math tool. The atmospheric corrections compared included Region of Interest (cloud shadow) DS, histogram method DS, FLAASH, and C2RCC. The results of comparison between the atmospherically corrected Landsat products and the field spectra taken during the summer of 2017 can be seen below in Table 4.4. There are numerous considerations when choosing a particular correction method, including ease of use, error accumulation, and synergy with other methods. In this study, the Vincent algorithm was followed including the preprocessing by histogram method DS for reflectance retrieval (Chavez, 1996; Vincent et al., 2004). As Vincent et al. state in their 2004 paper, the strength of the DS method is the simplicity. The assumption that there is a black body within the scene is usually reasonable in the Piedmont, as one may often choose from a shadow or deep reservoir. The histogram method does not allow for the operator to choose the black body sampling location, so the manual pixel-selection DS correction was also used and the results compared (Table 4.4). The two DS methods result in similar reflectance values on a scene wide basis, and match the field spectra reasonably well.
The two automated atmospheric correction methods tested were the FLAASH program and the C2RCC tool within SNAP (Adler-Golden, 1998; Brockmann et al., 2016). FLAASH is a well-known radiative transfer model with parameters to model the scene atmospheric thickness and conditions as well as the incoming and scattered radiance (Adler-Golden, 1998). The benefits of using a radiative transfer model are the incorporation of site-specific data including overpass date and time or climate into the calculation of atmospheric effect, building a simple but reasonably robust model of the atmospheric effects on the surface reflectance.

C2RCC, a tool developed for use by the European Space Agency, is a neural network for the correction of Sentinel 2 MSI data, also applicable to Landsat 8 OLI data. The C2RCC correction is specific to case-II waters, is a free add-on to a General Public License program, and is reasonably straightforward to use. The algorithm has also performed well in field calibration at Lake Allatoona (Table 4.4). However, the neural network is not yet grown for shortwave infrared bands, reducing the utility to a product for only select algorithms. While a future iteration may be a superior choice, the C2RCC atmospheric correction is not currently ideal for freshwater cyanobacteria sensing.

Histogram method DS, while simple, hyper-local, and easy to understand, may be too limiting in the assumption of constant atmospheric effect per scene. DOS is also subject to the operator’s choice of black body pixel or histogram slope change. Radiative transfer models are ideal for an operator with sufficient information on the scene, sensing geometry, meteorological and climactic conditions, and other parameters, but still makes major assumptions about the scattering effects of the atmosphere based on idealized water and gas particle sizes and homogeneity which may not be reflective of conditions in the scene. Balancing the assumptions and benefits with the desired outcomes, the simple histogram DS was chosen as the atmospheric
correction for the temporal study. When relative correction will be applied, the most consistent preprocessing is the most desirable. For studies of a single event, a program such as FLAASH may show superior performance in absolute correction. For use in sensing freshwater cyanoHABs, the C2RCC algorithm shows the most promise, if bands with longer wavelengths are incorporated into the program.

**Comparison of Algorithms**

The broad families within which most algorithms lie are empirical, semi-empirical, semi-analytical, and analytical. Empirical algorithms are developed from field data, where a several bands or band ratios are modeled, and the best fits determined by a technique such as PCA or multiple regression analysis (Eklundh and Singh, 1993; Vincent, 2004; Brezoni et al., 2005; Shi et al., 2015). Similar to traditional statistical models, neural networks may also be used to develop predictive models for cyanoHAB outbreaks (Cao et al., 2016). Empirical algorithms are usually regionally sensitive and may require re-parametrization before adequate results can be achieved at a new location (Sathyendranath et al., 2001, Ligi et al., 2017). Empirical algorithms have been used extensively in field-scale and regional-scale studies, often sharing similar trends among best fitting band-ratios (Matthew, 2011).

Semi-empirical methods model cyanoHAB pigments and water quality constituents using theoretical spectral relationships and field-derived coefficients. These algorithms often preform adequately at new sites but may fit better with site specific coefficients (Sathyendranath et al., 2001, Lee et al., 2002). Semi-empirical models used for detection of cyanoHABs in freshwater systems commonly use the ratio of near infrared (NIR) to red reflectance (Gitelson and Merzlyak, 1994a; Duan et al., 2007; Moses et al., 2009) One such example is the well-tested and
versatile Floating Algal Index (FAI), which may be applied to any system with red, NIR, and SWIR bands (Hu et al., 2009).

Semi-analytical models integrate field and laboratory data on non-algal optical constituents such as CDM within their algorithm (Dekker et al., 1993). Semi-analytical or quasi-analytical algorithms are similar to semi-empirical models but are designed to model the spectral contributions of individual optical constituents (Lee et al., 2002; Wynne et al., 2008). These algorithms often require MERIS-like band inputs and offer high predictive power (Becker et al., 2009; Mishra et al., 2009; Mishra and Mishra, 2012).

Analytical algorithm modeling starts from individual components within a freshwater system, and mathematically models their interactions to develop spectral best fits (Becker, 2009; Linhai et al., 2011; Ficek et al., 2012). Pure analytical algorithms require no data training or coefficients but may require hyperspectral resolution for the inputs (Randolph et al., 2008). Analytical models, while technically complex, are the most precise with the least risk of erroneous correlation (Dekker, 1993). Subsets of analytical methods, inversion algorithms use the absorption spectra of water color constituents to retrieve quantitative pigment data (Li et al., 2013; Twomey, 2013).

CyanoHAB remote sensing in small, dendritic reservoirs often requires finer spatial resolution than is offered by MERIS and MODIS sensors. However, hyperspectral sensors such as AISA, CASI-2, or CHRIS are prohibitively expensive for routine monitoring. Algorithms which use either Landsat-based sensing or a multi-algorithmic approach are therefore ideal for these studies (Table 4.2). The remote sensing of PC using Landsat data was first applied by Vincent in his 2004 study, producing a spectral ratio algorithm for use with Landsat 7 with resulting $r^2$ (adjusted) of 77.6%. Dependent on the dominance of cyanobacteria in the site’s
ecosystem, other methods may be used to indirectly target cyanobacteria concentration include IOP algorithms, indices such as the NDCI or Floating Algal Index, or band-ratio algorithms (Lee et al., 2002; Duan et al., 2007; Hu, 2009; Mishra and Mishra, 2012). Recent advancements have been made in vicarious calibration which allow for use of several methods in conjunction. Landsat mission data have often been combined for longitudinal studies (Schroeder et al., 2006), and recently, Landsat and Sentinel sensors were cross-calibrated to produce a longer running and spectrally denser dataset (Page et al., 2018). MODIS data have also been downscaled by Landsat for the development of subpixel estimation of algal coverage (Zhang et al., 2014). Caution should be used when combining sensor datasets that the data is preprocessed for relative comparison, but the resulting data may allow increased study periods, or backwards compatibility for new spectral and spatial resolution.

Three algorithms for the remote sensing of cyanoHABs in freshwater were tested in this study. The empirical phycocyanin index, the semi-empirical Duan chlorophyll-α index, and the semi-analytical FAI were fitted to satellite remote sensing from a date coincident with field spectroradiometry and water sample collection. The resulting best fits belong to the Duan et al. semi-empirical model for the detection of chlorophyll-α and the empirical algorithm developed by Vincent et al. for the indirect modeling of phycocyanin (Vincent et al., 2004; Duan et al., 2007). The Duan algorithm was used as a band-ratio model with no coefficients, while the Vincent algorithm was used with all coefficients preserved.

The Duan algorithm uses the theoretical relationship between the strongly reflecting near-infrared and strongly absorbing red wavelengths in chlorophyll-α producing organisms. The relationship is the same used in Fluorescent Line Height analytical algorithms, but was
developed as a simplified model of the near infrared-red relationship for rapid utilization in Landsat data (Duan et al., 2007).

**Approach Effectiveness**

The temporal and spatial scales at which extreme hydrological events control algal processes are thought to range from the daily effects of short-lived rainfall to multiannual drought, from the whole basin to meter-long fluctuations in bloom migration (Reichwaldt et al., 2012; Bowling et al., 2013). Hydrological studies of bloom dynamics therefore require a combination of long-term observation and large scene coverage with high resolution data. In optimizing the spectral, spatial, radiometric, and temporal resolutions of the available data, the balance is often made between spectral and spatial resolutions. While historical chlorophyll-α concentrations at Lake Allatoona have been in the 20-30 μg/L range, the observed water quality during the summer of 2017 was oligotrophic to mesotrophic (Table 4.1). These low observed concentrations resulted in a narrow range of values, and as such, no validation data were separated and tested. The effectiveness of the approach within mesotrophic lakes will often be especially affected by sample size, as the statistical comparisons improve with greater biochemical range, which is difficult to obtain in a mesotrophic to eutrophic reservoir setting. However, the combination of several remote sensing methods, verified against ground data and historical records, provided much more certainty in the occurrence and magnitude of the blooms.

While we tested the use of the FAI, we found it to be inapplicable to a mesotrophic setting (Hu, 2009). In these Piedmont reservoirs, algal density is dispersed vertically throughout the epilimnion, and is therefore affected by the sensed water column depth (Kutser et al., 2008). The FAI may have been a proxy in our study for inflowing sediment based on form and locations
of the maximum values; however, this reinforces that the algorithm should only be used in the sensing of surface scum forming blooms (Hu, 2009). Additionally, the Vincent et al. algorithm is still best used within the Lake Erie region, or perhaps reservoirs with the range of cyanobacterial densities observed in hypereutrophic settings. Within the Allatoona and Lanier reservoirs, the relationships between modeled phycocyanin and probable phycocyanin is still unquantified, and would be best studied using reparametrization and PCA methods.

Conclusion

The purpose of this review is to determine the ideal methods for the remote sensing of HABs or cyanobacteria within mesotrophic to eutrophic reservoirs such as are found across the Southeastern United States. Landsat OLI and Sentinel 2 are the best choices for sensor type due to high spatial and temporal resolution, but require algorithms which model pigments indirectly. The ideal atmospheric correction is dependent on the goal and scale of the study, but for HAB events which are directly ground truthed, using a radiative transfer model may best recreate the spectral shape seen in field measurements. For longitudinal studies, the Dark Subtract method is advisable for relative correction and temporal standardization. Algorithms which detect blooms in low-concentration events include simple semi-empirical models such as NIR/Red, which ought to be calibrated and validated for full use (Duan et al., 2007; Matthews, 2011). In the case that phycocyanin is below detection limits in water samples, the blooms may simply be classified as HABs or phytoplankton.

With an increase in tools available to iterate custom algorithms through large raster datasets, using several algorithms in remote sensing analysis is not much more time intensive than using a single method, and yields more confidence in the results. Multiple lines of evidence
for any satellite remote sensing should be built up from the appropriate dataset, preprocessing methods, and algorithms, and then strengthened by the most robust calibration and validation available for the funding and study length. This combined approach is useful in challenging remote sensing tasks, such as the characterization of HAB events within mesotrophic reservoirs.

**Acknowledgments**

The authors wish to acknowledge Abhishek Kumar, Sam Weber, Dr. Susan Wilde and Dr. Dayton Wilde; Dr. Koray Yilmaz, Stephanie Fulton, and the Water Resources & Remote Sensing lab for field support; and Dr. Tammi Richardson and Eric Lachenmyer for HPLC laboratory analysis.

This research was funded by the Geological Society of America Research Grant (2017) and the Watt’s Wheeler Research Grant (2017).

**References**


27. G.A. EPD (Georgia Department of Natural Resources Environmental Protection Division), 2004, Total maximum daily load evaluation for the little river embayment in the coosa river basin for chlorophyll-α.


42. Lee, T., Haddock, R., and Duke, D., 2016, Using citizen science to advance environmental research and monitoring in Alberta: Miistakis Institute, Mount Royal University.


Figure 4.A.1: The Floating Algal Index shows a reverse relationship with stormflow.
Figure 4.B.1: Chlorophyll-α Index scenes through time (Duan et al., 2007)
Appendix C

Table 4.C.1: The percent root mean square error by band for each of the compared atmospheric corrections shows that the C2RCC method has the highest performance, followed by Dark Subtract using cloud shadows, and FLAASH.

Table 4.C.1: Root mean square error of atmospheric correction methods by band, with clouds

<table>
<thead>
<tr>
<th>Method</th>
<th>% MRMSE</th>
<th>% RMSE by OLI Band</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>FLAASH</td>
<td>0.035</td>
<td>0.049</td>
</tr>
<tr>
<td>C2RCC</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>DS Histogram</td>
<td>0.086</td>
<td>0.076</td>
</tr>
<tr>
<td>DS Cloud Shadow</td>
<td>0.021</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Figure 4.C.1: Dark Subtract Histogram method applied to scenes with minimal cloud cover before manual radiometric corrections result in reduction in the Rayleigh-dominated wavelengths.
CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

Freshwater systems are governed by a web of biologic, chemical, physical, and geologic processes. These complex ecosystems have a natural progression from oligotrophy to eutrophy, in which lakes fill with plant life and sediments until a marsh or meadows is formed. Manmade lakes are engineered to slow this process through dredging, ecological management, and controlled releases, in order to act as flood control and the water source for municipal, industrial, and agricultural use in much of the Southeast. An increasing demand on these resources necessitates the careful monitoring and prediction of any changes in water quality, using precise conceptual models.

The literature shows a consensus that when sufficient sunlight, warm temperatures, and limiting nutrients are available, cyanoHABs may appear (Paerl and Huisman, 2008). These drivers for cyanoHABs and other harmful algae are well-described and accepted, however, based on this study we suggest that the conceptual model should also include rainfall duration and stormflow volume as distinct factors driving harmful algal growth (Figure 5.1). Our results at five sites across the two reservoirs show positive relationship between storm duration, stormflow, and algal biomass (Table 3.1). We choose to reject the null hypothesis that storm events are not among the key drivers of phytoplankton growth at these Piedmont reservoirs based on acceptable $R^2$ values, statistical significance, and well-behaved residuals. We characterized the hydrologic drivers using a wide suite of indices to seek agreement and compounding evidence for our case. Many of our statistical models show significant, parallel relationships.
Figure 5.1: An updated conceptual model shows our findings that rainfall duration and stormflow are statistically significant drivers of HABs in freshwater Piedmont reservoirs.

between storm metrics, increasing our confidence that rainfall events are good predictors of HABs.

The most interesting results from our study are the relationship between phytoplankton and storm duration, the difference in biohydrologic model as a function of distance from the tributary mouth, and the temporal scales at which our models showed correlations. We initially hypothesized that storm intensity would provide the kinetic energy to transport nutrients from the entire drainage basin to the reservoir efficiently, without prolonging the mixing of the water column and the disruption to the phytoplankton. What we saw in contrast was a set of models
showing that storm duration and stormflow volume were more significant drivers than event intensity, when the sites were at some distance from the river mouth. In Lake Allatoona, the correlation between storms and phytoplankton increased away from the Etowah, and at Lanier, from the Chattahoochee. At Lake Lanier, there were still positive correlations between stormflow and phytoplankton in the sites near the transition between river and reservoir, but these sites were at greater depths than the Allatoona sites, suggesting a vertical component to the mechanisms. The temporal relationships were interesting as well, as we saw positive relationships between storm events and phytoplankton which occurred up to 30 days after the rainfall, which is a longer increase in phytoplankton outcomes than some studies had previously suggested. Likely factors in these delays and prolonged increase in productivity are the interactions between hydrologic events and the nutrient flux from agricultural land use, which should be investigated in tandem in future studies.

We hypothesize forward that the increased discharge from the large tributaries as well as smaller inflows are interacting with the land use and climate within the Etowah and Chattahoochee River watersheds, driving the nutrient flux which is necessary for phytoplankton growth, which could be studied in the future using event-based, high resolution in-situ monitoring.

**Future Biohydrological Modeling**

In future studies, the conceptual model of these mesotrophic to eutrophic systems and their phytoplanktonic communities could be strengthened with improved computational models of the driving mechanisms. Although many of these drivers are interconnected, there is still uncertainty in the process chain which links climate, land use, and hydrology to bloom events.
The interior of our black-box understanding of biohydrological systems likely contains a vertical component, where depth of the water column, depth of the epilimnion, stability of the thermocline, and DO/pH gradients play a role (Figure 5.2). We see anecdotal evidence in Lakes

Figure 5.2: A view of the watershed includes lowered shoreline due to drought, gravity currents carrying sediments, and upriver storm events all interacting to influence the physicochemical conditions in the reservoir.
Allatoona and Lanier that distance from the tributary mouth and depth of the water at the sampled site are correlated with the relationship between precipitation and algal biomass, although further confirmation is needed. Developing a vertical model of the tributary–bloom hydraulics would be the next step in these biohydrologic models, and should include temperature, dissolved oxygen, Secchi-depth, chlorophyll-\(\alpha\), and velocity profiles during storm and calm conditions as well as during HAB events.

**Evaluating the Combined Approach**

The satellite data were used to improve the spatial and temporal resolution of the in-situ measurements. The in-situ water sampling provided biochemical information at 14 discrete sites, which were to be complemented by a characterization using Landsat 5, 7, and 8 surface reflectance data as well as the data for an analysis of bloom size and distribution. The Landsat scenes showed that typical conditions at the two reservoirs were disperse rather than scum-forming, best modeled using a chlorophyll-\(\alpha\) index such as the NIR/Red band ratio (Duan et al., 2007). Phycocyanin was also modeled using an empirical model, but we interpret these data with caution, as the ground-truthing data were below detection limits and the original parametrization took place in a highly eutrophic environment (Vincent et al., 2004). A third model tested was the Floating Algal Index, which did not perform well. The FAI is designed to detect HABs which float on top of the water, while those dispersed through the photic zone will not be well retrieved due to the interaction of the water column, particulates, and dissolved components (Hu et al., 2009). Other dynamics include the highly dendritic nature of the reservoir, which necessitates the highest available spatial resolutions. The combined use of Landsat and Sentinel 2 data is useful
to resolve the shallow and fine features, although in our case we were unable to obtain coincident
data for cross-calibration.

The issues of sensitivity to HAB concentration are more difficult to overcome than
differences in spatial, temporal, or spectral resolution. We recommend that for the remote
sensing of mesotrophic to eutrophic reservoirs, a transect-based ground truthing survey or other
sample to satellite matching be employed to develop local statistical relationships between
blooms and remote signature and compared to global models for increased biochemical
resolution. In some cases, the total variation in chlorophyll-α, phycocyanin, or other
measurements may be narrow, in which case models must be designed for sensitivity to low
concentration data.

**Recommendations**

While uncommon in most of Lake Allatoona and Lake Lanier, the risk of cyanoHABs or
other harmful algal blooms is the highest at the Little River site in eastern Allatoona. The Little
River arm is consistently more turbid, with higher dissolved ions and higher chlorophyll-α
counts (Table 4.1) as well as blooms which are potential hazards (Figure 3.6, bottom left and top
right). The Little River drains an area with primarily low and high intensity urban land use,
which is likely the source of the nutrients and increased productivity within the associated
reservoir arm (UGA CAES, 2009). We recommend increased in-situ monitoring of the Little
River tributary for dissolved oxygen, nutrients, and turbidity in addition to chlorophyll-α,
especially between July and October or during times of prolonged drought and low flows.
REFERENCES


35. Coppin, P. R., and Bauer, M. E., 1994, Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features: IEEE Transactions on Geoscience and remote Sensing, v. 32, no. 4, p. 918-927.


94. Kunkel, K. E., Stevens, L. E., Stevens, S. E., and Sun, L., 2013, Regional climate trends and scenarios for the U.S. National Climate Assessment, Part 2, NOAA.


117. Mayhew, M. C., Rasmussen, T. C., and Parker, A. K., Phosphorus binding by iron-rich soil in the Southeastern Piedmont: Implications for point and non-point sources of phosphorus2001, Georgia Institute of Technology.


132


132. NDMC, National Drought Mitigation Center, University of Nebraska.


189. USACE, 1952, Alabama-Coosa-Tallapoosa (ACT) river basin water control manual.

190. USACE, 1959, Apalachicola-Chattahoochee-Flint (ACF) river basin water control manual.


APPENDIX A

A GUIDE TO FIELD FILTRATION FOR HPLC ANALYSIS

High Performance Liquid Chromatography (HPLC) uses a strained sample of the particulates in water, which may be obtained in the field or in the laboratory. If the water is both collected and filtered in the field, the pigment sample may be stored on dry ice immediately following the steps described below.

Materials

Paper Towels
Aluminum foil squares, roughly 5 inches wide
Permanent marker OR printed sample labels
Glass Microfiber Whatman GF/F 47mm filters (circles) CAT No. 1825-047
Cooler with wet and dry ice
Plastic sandwich bags
Plastic pipettes
DI or distilled water

Equipment

Hand pump with tube attachment
Millipore 250 mL plastic filtration tower (funnel, neck, and chamber)
Plastic beaker, 250mL
2 pair bent-nose tweezers
Figure A.1: Equipment for field filtration.

1. Set up equipment in a cool, dry area with no overhanging foliage, such as a covered picnic table. Samples should remain on wet ice until needed for filtration; after filtration, set samples on dry ice. Clean paper towels or foil can be used to lay equipment when not in use.
2. Rinse filtration towers with DI or distilled water.
3. Unscrew filtration tower funnel from neck and membrane piece (Figure A.1.1).
4. Using tweezers, select a glass Whatman filter circle and place on the white plastic membrane in the neck of the filtration tower. Reattach neck and funnel piece.
5. Attach hand-pump hose to filtration tower’s chamber.
6. Measure known volume of sample water into beaker, and slowly fill filtration funnel.

7. Cover filtration chamber valve with thumb and begin applying a steady vacuum with the hand pump, pulling the water from the funnel to the chamber.
8. During the pumping process, the filter circle will likely darken into a tea-green color and become increasingly difficult to pull water through.
9. When the desired volume of sample water has been filtered, detach hand-pump and filtration funnel and neck.

Figure A.5: The filter will show color from the particulates.
10. Using tweezers, remove the (now colored) filter circle gently from the plastic membrane.

11. On the (untouched) inside of a foil square, lay the filter circle color up, and still using the tweezers, fold in half and half again. Fold foil around the filter to flatten, and secure the foil packet with a sample label. Record sample number and filtered volume in field notes.

Figure A.6: Fold the sample into foil neatly.
Figure A.7: All samples should be neatly and consistently labeled.

12. Place foil-wrapped filter in sandwich bag and place on dry ice.

Figure A.8: In the laboratory, the same procedure may be set up to run several filtration towers simultaneously.
APPENDIX B

STANDARD PRECIPITATION INDEX TUTORIAL

1. Begin by collecting your data. We use NCDC daily observations in this tutorial, but you can use USGS gage data or other precipitation records. The most important qualities for the data are continuity and at least 30 years’ of record.

   a) You can download free daily precipitation from the NCDC at the web address https://gis.ncdc.noaa.gov/maps/ncei. If you do not have a free account, make one now, and log in.

   - Choose Daily.
   - Zoom in to your area of interest.
b) In this tutorial, we use the **GHCN Daily** dataset. To continue, click on the wrench icon near the dataset you choose.

- A toolbox will appear in the upper right of the map. Click on **Identify**.
Select the site you are interested in and note the information which appears in the left hand sidebar. This describes the location and the beginning and end dates for the record.

c) If the record is at least 30 years long, the data may be sufficient for an SPI construction. Click View Station Details. Note the data coverage. 95% coverage is acceptable, so select Add to Cart.
- Select Custom CSV.

- Select the desired date range, preferably the maximum.

- Hit Continue. The units will not matter to the end result.
- You only want precipitation for this application; selecting more parameters will complicate your data preparation.
Accept these selections and review your order on the next page. Select **Submit Order**.

d) Wait for the data to be emailed to your account, which might take several minutes or in some cases several hours. When you receive it, download and save the data to a hard drive, as the link will not be valid after a few days.

2. Download and open in an Excel spreadsheet. Again, you can use R or a different program to restructure the data, but you will want to save the end product as a .COR file.
   a) The data should look similar to the screenshot below. **PRCP** is the depth of precipitation per day in the unit you selected; here, it is in inches.
b) Reformat the data as follows:

```
A      B     C
1      Datetime Precip
2      drop in datetime and precipitation 7/16/2008 0
3
4
5
6
7
8
9
10
11
12
13
```

c) Drop in the Date-time and precipitation, then copy and paste these values into new columns (as in teal instructions).

d) In a new column, get the Year and Month (hereafter known as Yearmo) using the expression:
\[=\text{YEAR} (E3) \& \text{MONTH} (E3)\]

e) Put Yearmo and Precip side by side again by \textbf{Copy->Paste Values}.

\begin{center}
\begin{tabular}{|l|l|}
\hline
\textbf{Date} & \textbf{Yearmo} \\
\hline
7/16/2008 & taking the date and getting yearmo \\
7/17/2008 & 20087 \\
7/18/2008 & 20087 \\
7/19/2008 & 20087 \\
7/20/2008 & 20087 \\
7/21/2008 & 20087 \\
7/22/2008 & 20087 \\
7/23/2008 & 20087 \\
\hline
\end{tabular}
\end{center}

f) Re-paste these columns on a new sheet:

3. Select columns A and B and select \textbf{Data -> Subtotal}.

\begin{center}
\includegraphics[width=0.5\textwidth]{subtotal.png}
\end{center}

a) Set the options to \textbf{At each change in Yearmo, use function sum} and \textbf{Add subtotal to Precip} and \textbf{OK}.
b) Select the small 2 at the far left:

- Select columns A and B, the Select **Home** and **Find and Select**
  and **Go to Special**.
c) Copy and paste these cells into a new sheet; these are your monthly precipitation totals.

- These data need to be in a format with no decimals, so multiply the precipitation totals by 100 and in a new sheet, you should have Yearmo and Precip x 100.
Now you must take Yearmo and retrieve Year and Month as two separate columns.

Format exactly as the screenshot below, with the location in the format: Place, State in A1 and all data starting in Row 2. Columns from left to right will be A2 Years, B2 Months, and C2 Precipitation x 100. Do not add any labels.
d) Copy and paste these data into a new sheet, and save as a text (tab delimited) .COR file as seen below:

![Image of spreadsheet]

```
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allatoona Dam, GA</td>
<td>1952</td>
<td>4</td>
</tr>
<tr>
<td>1952</td>
<td>5</td>
<td>385</td>
</tr>
<tr>
<td>1952</td>
<td>6</td>
<td>165</td>
</tr>
<tr>
<td>1952</td>
<td>7</td>
<td>46</td>
</tr>
<tr>
<td>1952</td>
<td>8</td>
<td>635</td>
</tr>
<tr>
<td>1952</td>
<td>9</td>
<td>403</td>
</tr>
<tr>
<td>1952</td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td>1952</td>
<td>11</td>
<td>299</td>
</tr>
<tr>
<td>1952</td>
<td>12</td>
<td>489</td>
</tr>
<tr>
<td>1953</td>
<td>1</td>
<td>557</td>
</tr>
<tr>
<td>1953</td>
<td>2</td>
<td>534</td>
</tr>
</tbody>
</table>
```

e) To the warning popups, select **OK** and **Yes**.

4. Opening your new file should be possible with Microsoft Word, and look like:
a) You will need to download a program from a website. The address is:

http://drought.unl.edu/MonitoringTools/DownloadableSPIProgram.aspx and the program is called SPI_SL_6.exe.

b) Open the program and agree to run it.

c) Number of scales is how many different SPI scales you want to produce.
I recommend four, namely: 3-Month, 6-Month, 12-Month and 24-Month. For more information on each SPI scale, http://drought.unl.edu/portals/0/docs/spi-program-alternative-method.pdf

- Enter the scales by typing and hitting **Enter** after each entry:

![Screen capture of SPI calculator](image)

d) Make sure that your .COR file is copied into the same folder as the SPI program; you can’t change the working directory. For input file, type the exact file name exactly as shown below.
e) Type the desired output file name and use a .DAT file extension, exactly as shown below:

f) The program will close. Find your .DAT file in the same directory as your .COR and the program. Your file will have data like the following, which you ought to save as a txt or excel file:
The data is month, day, 3-Month SPI, 6-Month SPI, etc. but the -99.00 values represent no data.

a) These entries are a result of the SPI math; you can’t have a 6-Month SPI for a few months into the data. Keep this in mind if you need an exact data period.

b) Turn the SPI into a visual!

- As a next step, try running a Mann-Kendall non-parametric test on the SPI series to identify whether there is a real trend in your SPI values.
Figure B.1: Your SPI should be clearly labeled with dates, location, and an explanation of the index levels. Color coding for the drought and wet regimes makes the trend easier to identify.
APPENDIX C

R FUNCTIONS

#---------------------------------------------------------
#            MULTIPLE REGRESSION CHECKS
#---------------------------------------------------------

#Install packages
install.packages('tree')
install.packages('mgcv')

#Load in packages
library('tree')
library('mgcv')

#Check for model complications
modelcheck <- function(data, y, x1, x2, x3) {
  #check for correlations
  dev.new(height = 7, width = 7)
  pairs(data, gap=0, panel=panel.smooth, cex.labels=0.85,
        labels=c('Label 1','Label 2','Label 3','Label 4'))

  #check for curvature
  library(mgcv)
  dev.new(height = 7, width = 7)
  par(mfrow=c(2,2))
  curvature <- gam(y ~ s(x1) + s(x2) + s(x3))
  plot(curvature, col='burlywood')

  #check for overparametrization
  paramcheck <- function (n, p) {
    indexInitium <- 1
    indexFin <- p
    k <- seq(indexInitium, indexFin, 1) #making a sequence
    #from k=1 to k=p
    combinations <- sum(choose(p, k)) #all combinations of
    #parameter interactions
    lowerlimit <- 3 * combinations #3 sample points per
    #parameter
    if (n < lowerlimit)
      print("Overparameterized!")
    else if (n == lowerlimit)
      print("Ok")
  }
print("Overparameterized!")
else
print("Good to go!"))
}
params <- ncol(data) - 1
paramcheck(n=length(y), p=params)

#check for complex interactions
library(tree)
dev.new(height = 7, width = 7)
par(mfrow=c(1,1))
interaction <- tree(y~., data=data, model=T) #y is #hardcoded here
plot(interaction)
text(interaction, cex=0.65, col='forestgreen')

#------------------------------------------
# TIDY PLOTS
#------------------------------------------

cxlabel <- expression(paste('Log Chl-', alpha, ', ', mu, 'g/L'))

stormplot <- function(x, y, title) {
  #fun(myxvariable, myyvariable, title='mycharttitle')#is how
  #you should do this
  xlims <- c(min(x, na.rm=TRUE), max(x, na.rm=TRUE))
  ylims <- c(min(y, na.rm=TRUE), max(y, na.rm=TRUE))
  rsq <- summary(lm(y~x))$r.squared
  rsq <- round(rsq, 2)
  plot(x, y, xlim=xlims, ylim=ylims, type='p', axes=FALSE, xlab=' ', ylab=' ', main=title, pch=16, col='springgreen4', las=1)
  axis(side=2, col='gray')
  axis(side=1, col='gray')
  abline(lm(y~x), col='black')
  text(x=xlims-2, y=ylims-.5, label=rsq, cex=1.5)
}

stormplot2 <- function(x, y, title) {
  xlims <- c(min(x, na.rm=TRUE), max(x, na.rm=TRUE))
  ylims <- c(min(y, na.rm=TRUE), max(y, na.rm=TRUE))
  rsq <- summary(lm(y~x))$r.squared
  rsq <- round(rsq, 2)
  plot(x, y, xlim=xlims, ylim=ylims, xlab=' ', ylab=' ',

166
stormplot3 <- function(x, y, title) {
  xlimits <- c(min(x, na.rm=TRUE), max(x, na.rm=TRUE))
  ylimits <- c(min(y, na.rm=TRUE), max(y, na.rm=TRUE))
  rsq <- summary(lm(y~x))$r.squared
  rsq <- round(rsq, 2)
  plot(x, y, xlim=xlimits, ylim=ylimits, xlab= '', ylab= '', type='p', axes=FALSE, main=title, pch=16, col='springgreen4', las=1)
  title(ylab='Log Weekly Avg Q, cfs', xlab=' ', cex.lab=2)
  axis(side=2, col='gray')
  axis(side=1, col='gray')
  abline(lm(y~x), col='black')
  text(x=xlimits-2, y=ylimits-.5, label=rsq, cex=1.5)
}

stormplot4 <- function(x, y, title) {
  xlimits <- c(min(x, na.rm=TRUE), max(x, na.rm=TRUE))
  ylimits <- c(min(y, na.rm=TRUE), max(y, na.rm=TRUE))
  rsq <- summary(lm(y~x))$r.squared
  rsq <- round(rsq, 2)
  plot(x, y, xlim=xlimits, ylim=ylimits, xlab= '', ylab= '', type='p', axes=FALSE, main=title, pch=16, col='springgreen4', las=1)
  title(ylab='Log Weekly Avg Q, cfs', xlab=chlabel, cex.lab=2)
  axis(side=2, col='gray')
  axis(side=1, col='gray')
  abline(lm(y~x), col='black')
  text(x=xlimits-2, y=ylimits-.5, label=rsq, cex=1.5)
}

dev.new(height = 7, width = 7)
par(mfrow=c(3,3), mar=c(5, 5, 4, 4))
stormplot(log(chla8), log(setoWAQ), title='Etowah, Site 8')
stormplot(log(chla9), log(setoWAQ), title='Etowah, Site 9')
stormplot(log(chla10), log(setoWAQ), title='Etowah, Site 10')
stormplot2(log(chla11), log(setoWAQ), title='Etowah, Site 11')
stormplot(log(chla12), log(setoWAQ), title='Etowah, Site 12')
OLI TO SENTINEL 2B VICARIOUS CALIBRATION

1. Sentinel 2 processing

Surface reflectance scenes from Sentinel 2B and Landsat 8 OLI were obtained from August 18, 2017 and were processed into Normalized Difference Chlorophyll Index and Floating Algal Index products. The clouds were identified using an unsupervised K-means classification in ENVI, and cloud-free water was identified.
2. Selection of Deep Water

Areas of deep water were selected to avoid data with bottom-sediment interaction. These two areas in eastern Lake Allatoona each correspond to a point of ground chlorophyll concentration data.
3. Sampling

A regular grid of points within the deep water was created and the NDCI and FAI values at each point were sampled for comparison.

4. Comparison of Indices

The Normalized Difference Chlorophyll Index was modeled as a function of the Floating Algal Index, in order to determine the slope and intercept of their relationship. The relationship was expected to be linear and have low RMSE, with a strong correlation. The resulting $r^2$ of 0.2 is likely due to the presence of unmasked sun glint, which has a different effect on each band’s reflectance values, and reducing the correlation between these two different indices.
5. Estimating $\text{NDCI}_{\text{OLI}}$

In the case that the two indices had a stronger linear relationship and the confidence in the reflectance values was higher, the relationship could be applied to the Landsat 8 OLI Floating Algal Index, as shown below.

$$\text{NDCI}_{\text{OLI}} = 0.0002 \times \text{FAI}_{\text{OLI}} + 0.0055$$

The OLI Floating Algal Index:
FAI_{OLI}, clouds masked from scene:
Determination of Chlorophyll-a concentration from NDCI$_{\text{Sen2}}$

6. Chlorophyll-a concentration may be estimated using the methodology in Mishra and Mishra, 2012:

$$\text{Chl-a} = 14.039 + 86.115(\text{NDCI}_{\text{Sen2}}) + 194.325(\text{NDCI}_{\text{Sen2}})^2$$
However, the issues resulting in negative values in the original Sentinel indices have propagated through, resulting in much higher and lower modeled concentrations than are actually observed in the lake, seen below:

<table>
<thead>
<tr>
<th>Lake Allatoona Site</th>
<th>Mean chl-a, µg/L</th>
<th>Bloom appearance?</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 11, Mid-lake</td>
<td>1.29</td>
<td>--</td>
<td>8/18/17</td>
</tr>
<tr>
<td>Site 10, Little River</td>
<td>3.31</td>
<td>Subsurface, possible</td>
<td>8/18/17</td>
</tr>
<tr>
<td>Site 8, Etowah River</td>
<td>3.41</td>
<td>Subsurface, possible</td>
<td>8/18/17</td>
</tr>
</tbody>
</table>

I believe that the primary source of error in this vicarious calibration is the sun glint in the Sentinel 2B data, seen below:

[Image of lake with satellite data and glint annotation]

7. Check Outlier Glint
A. Before looking at the data table, I visually identified where in the water the sun glint was maximal and minimal.

3. I selected the minimal glint pixel subset and reran the linear model.

\[ y = 0.0005x + 0.0234 \]
\[ R^2 = 0.7098 \]
4. Using the derived relationship, I obtained a revised NDCI\textsubscript{OLI}. 
8. Unfortunately, likely through error accumulated during atmospheric correction or from sun glint, the values obtained for the NDCI OLI using the linear relationship are out of bounds for this normalized-difference model. A cloud-free coincident scene will likely be needed to complete the cross-sensor calibration at Lake Allatoona.