Spring 2021

Using Remote Sensing and Environmental Precursors to Detect and Predict Cyanobacteria Harmful Algal Blooms in Northeastern US Waterbodies

Amanda Craver

Using Remote Sensing and Environmental Precursors to Detect and Predict Cyanobacteria Harmful Algal Blooms in Northeastern US Waterbodies

A Project Presented

by

Amanda Craver

Master of Science in Civil Engineering

Department of Civil and Environmental Engineering
University of Massachusetts
Amherst, MA 01003

May 2021

Using Remote Sensing and Environmental Precursors to Detect and Predict Cyanobacteria Harmful Algal Blooms in Northeastern US Waterbodies

A Masters Project Presented

by

Amanda Craver

Approved as to style and content by:

DocuSigned by:
Kostas Andreadis
1E13F1BC46FF44C...

Konstantinos Andreadis, Chairperson

DocuSigned by:
01F608E2CB7F4FD...

Caitlyn Butler, Member

DocuSigned by:
Colin Gleason
86282A4A8CB34B6...

Colin Gleason, Member

DocuSigned by:
01F608E2CB7F4FD...

Caitlyn Butler
Civil and Environmental Engineering Department

Table of Contents

Abstract

Cyanobacteria harmful algal blooms (cyanoHABs) are a global problem with human health, environmental, and economic concerns. The severity and frequency of toxic cyanoHABs are expected to increase with climate change. Remote sensing has proven to be a useful tool in monitoring cyanoHABs. This study uses remote sensing observations from Sentinel-3 Ocean Land Color Imager (OLCI) combined with the Spectral Shape Algorithm (SSA) to detect the presence of cyanobacteria in numerous waterbodies throughout the Northeast United States over 2016 to 2020. The ACOLITE processor was used for the atmospheric correction of the Sentinel-3 OLCI data, as it has been shown to provide more accurate results over water bodies. Citizen scientist data from two EPA apps, BloomWatch and CyanoScope, were used to guide the cyanoHAB detection results. This study focuses on three different aspects of using the remote sensing methods to detect cyanobacteria harmful algal blooms: (1) examine two different reflectance center band combinations of the SSA at 665 nm and 681 nm, (2) compare SSA controls for non-bloom periods in winter/spring to the summer bloom periods, and (3) examine bloom areas within waterbodies. The SSA method with the center band at 665 nm was found to perform better than the SSA method with the center band at 681 nm. Additionally, cyanoHAB environmental precursors were determined and their effects on the formation of cyanoHABs were assessed, with maximum air temperature having the greatest impact on cyanoHABs formation.

1. Introduction

1.1. CyanoHABs

Harmful algal blooms (HABs) are generally defined as an overgrowth of algae that results in negative environmental and human health consequences. Cyanobacteria harmful algal blooms (cyanoHABs) are the most common form of inland freshwater HABs, inflicting widespread public health, water supply, and economic consequences. Concentrations of cyanobacteria naturally live in fresh waterbodies, in which they play essential roles in primary production, nitrogen fixation, and other nutrient and sediment cycling (Howarth et al., 1988). Nevertheless, cyanobacteria pose major health hazards when the concentration of cyanobacteria increases rapidly creating cyanoHABs. The increase and accumulation of cyanobacteria can be linked to many different factors, including environmental factors, climate change, over-fishing, land-use changes, and agricultural activities (Glibert et al., 2015).

Certain species of cyanobacteria produce highly potent toxins, known as cyanotoxins. These toxins are termed secondary metabolites because they are not essential for their own growth and metabolism (Carmichael, 1992). The compounds produced are considered toxins because they cause health risks and poisoning to humans and animals (Glibert et al., 2005; O'Neil et al., 2012). More than one hundred types of cyanotoxins exist and can be categorized in three different groups: (1) hepatotoxin, that target the liver, (2) neurotoxins that affect the nervous system, and (3) dermatoxins that cause skin irritation (Merel et al., 2013). In a 1999 World Health Organization study, an estimated 60% of global cyanobacteria samples from inland waterbodies contained cyanotoxins (WHO, 1999). The cyanotoxins frequently found throughout U.S. inland waterbodies include microcystins, cylindrospermopsin, anatoxins and saxitoxins (US EPA, 2020). The human health effects include skin and eye irritation, damage to liver and kidney function, and severe flu-like symptoms (Carvalho et al., 2011). Humans can be exposed to cyanotoxins through dermal contact, ingestion and inhalation during recreational activities, contaminated drinking water, consumption of fish and other shellfish, and the use of algal supplements (Loftin et al., 2016). CyanoHABs have caused illness and/or death to humans and animals in almost every US state and many countries (Backer et al., 2015).

The effects of cyanoHABs extend beyond human and animal health, causing great economic burden. The US EPA estimated the costs associated with nutrient pollution and HABs with US data from 2000 to 2012, and found significant costs associated with HABs. The EPA study specifically reviewed Ohio lakes over a two-year period and found associated costs of tourism and recreation and drinking water treatment due to HABs as $47 million and $13 million, respectively. Additionally, the study documented increased emergency room visit costs in Sarasota County, Florida from human contact with algal blooms, which can cost the county more than $130, 000 per bloom season (US EPA, 2015). CyanoHABs deteriorate water quality, threaten aquatic ecosystem health, threaten drinking water supply, and cause human illness (Huisman et al., 2018). CyanoHABs deteriorate water quality by reducing clarity (increasing the turbidity), increasing pH, and depleting dissolved oxygen. Millions of Americans rely on surface waterbodies for drinking water, which have the potential for cyanotoxin contamination. In 2014, more than 500,000 residents in Toledo, Ohio lost access to drinking water due to toxic cyanoHABs near the drinking water supply intake. Toledo reported 2.5 μg/L microcystin in the finished drinking water, which exceeds the World Health Organization's recommended 1.0 μg/L threshold for cyanotoxins in finished drinking water (Steffen et al., 2017).

1.2. Environmental Precursors to CyanoHABs

The severity and frequency of toxic cyanoHABs are increasing globally due to climate change, causing increased threats to public health, increased costs associated with cyanoHABs, and water supply disruption. Inland waterbody phytoplankton communities are greatly influenced spatially and temporally through changes in water quality parameters that include surface water temperature, nutrient concentrations, rainfall, water clarity, wind, and sunlight availability (Watson et al., 1997; Paerl et al.; 2001; Glibert et al., 2005; Heisler et al., 2008; O'Neil, 2015; Beaver et al., 2018; Cermona et al., 2018). Several studies determined that increased nutrients, specifically phosphorus, shifted algal populations towards cyanobacteria domination (Watson et al., 1997; Paerl, 2008; Beaver et al., 2018). Algal populations are found to be dominated by cyanobacteria at total phosphorus concentrations of approximately 100-1000 $\mu g\ L^{-1}$ (Watson et al., 1997; O'Neil et al., 2015). Phosphorus may control the cyanobacteria presence because cyanobacteria have the ability for nitrogen fixation (O'Neil et al., 2015).

The ability for short-term forecasting of future cyanoHABs would be of highest interest to water managers and public health officials. Many studies have tried to predict cyanoHABs with the presence of environmental precursors; the common precursors studied include water temperature, phosphorus concentrations, nitrogen concentrations, sunlight availability, and water clarity (Bukowska et al., 2017; Xiao et al. 2017; Duan et al. 2018; Russo et al., 2020). Russo et al. (2020) found that water temperature, phosphorus concentration and sunlight availability are positively correlated to cyanoHABs through increasing growth rates, whereas, water clarity or Secchi Disk depth was negatively correlated to CyanoHABs. Nitrogen concentrations were found to have mixed correlations to cyanoHABs.

Earlier studies used linear and multiple regressions in single lake or a small number of lakes in the same region to predict cyanoHABs with environmental precursors, such as the Baltic Sea, Lake Taihu, China, and Lake Erie, USA (Kanoshina et al., 2003; Xu et al., 2010; Stumpf et al., 2016; Steffen et al., 2017). Many of these studies use cyanobacteria data at a genus level and disregard differences among species of cyanobacteria, which can be affected differently given varying environmental precursors (Beaver et al., 2018). Recently, studies have expanded the approach to find trends across larger geographic regions in numerous ways: (i) examined how different cyanobacteria react to different environmental precursors; (ii) created frameworks to be applied to lakes that have not yet experienced blooms; (iii) aided analyses with the use of neural networks and models (Bukowska et al., 2017; Beaver et al., 2018; Zhao et al., 2019; Pyo et al., 2020). Beaver et al. (2018) used the USEPA National Lakes Assessment to find regional trends in environmental precursors and cyanobacteria species. The study found correlation between total phosphorus and water clarity with lake phytoplankton composition. Moreover, increased concentrations of the cyanobacteria toxin microsystin strongly correlated with high phosphorus levels and watersheds with low forest cover (Beaver et al. 2018).

1.3. Remote Sensing of CyanoHABs

Comprehensive cyanoHAB trends have not been well documented due to historical paucity of in situ sampling programs and standard analytical methods (Urquhart et al., 2017). Rigorous, widespread in situ sampling efforts to thoroughly document cyanoHAB trends would require significant time, costs, and human-power resources. Remote sensing data has proven to be a useful tool in monitoring cyanoHABs. Many studies have used remote sensing data to identify cyanobacteria throughout inland waterbodies using data from various satellites and sensors, such as Envisat MERIS (Wynne et al., 2008; Wynne et al., 2011; Kahru and Elmgren, 2014; Stumpf et al., 2016), Sentinel-3 OLCI (Urquhart et al., 2017; Mishra et al., 2019; Coffer et al., 2020), Landsat/TM (Oyama et al., 2015), and MODIS/ASTER (Hu et al., 2009; Kudela et. al, 2015; Jia et al., 2019).

Remote sensing methods use surrogate identification methods to detect cyanobacteria presence and estimate biomass. The two primary surrogate identifiers are phycocyanin (PC) and chlorophyll-a (chl-a). According to Stumpf et al. (2016), the decision on which surrogate parameter to choose, should be decided by three factors: availability of data, specificity of the surrogate identifier, and sensitivity regarding its reflectance. Stumpf et al. (2016) establish that chl-a data is more widely available, and PC does not have standard laboratory methods. PC is more specific because it is a pigment indicator found specifically in cyanobacteria, whereas chl-a is found in both cyanobacteria and other algal species. Lastly, Chl-a has an absorbance peak that is better suited for broad remote sensing, falling between 665 and 681 nm, which wavelength is found on many sensors. Specific sensors, including MERIS and OLCI, have available bands near 620 nm, which is advantageous for PC detection (Stumpf et al., 2016).

Previous studies have established many different approaches to measuring these surrogate identifiers, through analytical, semi-analytical and derivative methods. Analytical methods, first outlined in Simis et al. (2005), successfully used in situ radiometry data to extract the spectral absorption and solve for PC. Analytical methods become less accurate when scaled to remote sensing data due to sediment backscatter and atmospheric corrections (Stumpf et al., 2016). Next semi-analytical algorithms, or bio-optical algorithms, commonly use satellite band ratios to isolate algal pigments, from the determined water reflectance. Studies have used two-band and three-band ratio algorithms to calculate the algal surrogate parameters, chl-a (Gilerson et al., 2010; Le et al., 2013; Toming et al., 2017; Pyo et al., 2018; Keith et al., 2019) and PC (Mishra et al., 2009, 2014; Ogashawara et al., 2013; Pyo et al., 2018). Semi-analytical algorithms have been widely used with sensors that have fewer bands, such as Landsat, and are not limited to specific sensors with narrow band wavelengths (MERIS and OLCI).

Lastly, the more complex, derivative, also known as curvature or baseline algorithms, include the Spectral Shape Algorithm, (Wynne et al., 2008; Lunetta et al., 2015) and the Floating Algal Index (Hu et al., 2009; Oyama et al., 2015). Derivative algorithms are used by determining the presence of the surrogate parameter, chl-a, using the extreme changes in the reflectance peaks and valleys. Chl-a is absorbed strongly at 665 and 681 nm, and water is absorbed around 754 nm; thus, the chl-a is estimated using this absorption difference, known as the "red-edge" (Stumpf et al., 2016). The Spectral Shape Algorithm (SSA) is a commonly used derivative algorithm in reviewed literature, which was first done to determine cyanobacteria presence in Lake Erie by Wynne et al. (2008) and later updated by Lunetta et al. (2015). The SSA is a useful method in determining cyanobacteria presence across many different waterbodies, ranging from individual waterbodies to resolvable waterbodies across the entire US (Wynne et al. 2008, 2011, 2013; Kudela et al., 2015; Lunetta et al., 2015; Urquhart et al., 2017; Schaffer et al., 2018; Coffer et al., 2020).

The use of remote sensing data can provide relatively dense observations both temporally and spatially with only minor regional wintertime limitations due to ice and snow cover (Coffer et al., 2020). However, the ability to observe a specific waterbody depends on its size and shape due to the spatial resolution and orbit of the satellite. A 2017 study by Clark et al. used Envisat MERIS and Sentinel-3 OLCI data, with 300 m nadir resolution, to determine resolvable Continental US lakes using the 2012 National Hydrography Dataset. The study found that 5.6% of waterbodies were resolvable using 300 m resolution, and when these data were overlain with the US public water surface intake (PWSI) data, 57% of drinking water supply lakes were resolvable (Clark et al., 2017). Urquhart and Schaeffer's (2019) noted that lakes with a surface area smaller than 27 ha (0.104 square mile) were unresolvable and could not be used in remote sensing analyses. Nevertheless, remotes sensing is a powerful tool that can be used to detect cyanoHABs, determine trends, and forecast future blooms given environmental precursors.

1.4. Citizen Scientist CyanoHAB Data

Previous studies have validated remote sensing methods with in situ samples; however, many waterbodies that experience cyanoHABs lack current standard field sampling campaigns and cyanoHAB trends cannot be assessed in these waterbodies. The use of citizen scientist data is vital to future remote sensing methods to detect cyanoHABs. Current US EPA mobile applications, such as BloomWatch[1] and CyanoScope[2], allow Citizen scientists to document HABs and collect water samples to test for cyanobacteria. Citizen scientists can upload pictures, descriptive bloom information, document location, and even determine the algae species through their water samples. These data help inform water managers of cyanoHABs in order to warn waterbody visitors of potential blooms, enact beach closures at these waterbodies, ensure toxins are not found in finished drinking water. The objective of this study is to examine the presence of cyanoHABs in multiple waterbodies throughout the northeastern region of the US using the SSA with Sentinel-3 OLCI data to detect cyanobacteria and citizen scientist data to direct remote sensing methods to which waterbodies cyanoHABs occurred and knowledge of bloom descriptors. This study uses satellite remote sensing data and the Spectral Shape Algorithm to detect the presence of cyanobacteria across northeastern US waterbodies. In contrast to other studies, we used citizen scientist data for

[1] https://cyanos.org/bloomwatch/
[2] https://cyanos.org/cyanoscope/

waterbody selection and to aid in cyanobacteria detection. The objectives of our research revolve around three questions:

1. Are citizen scientist data an effective tool to guide remote sensing detection of cyanobacteria in waterbodies?
2. Can the Spectral Shape Algorithm be used to detect cyanobacteria presence in these waterbodies?
3. Can meteorological environmental precursors be used to predict the presence of cyanoHABs?

We used two different cyanobacteria detection methods along with a cross-correlation analysis to address the study goals. First, the Wynne et al. (2008) and Lunetta et al. (2015) Spectral Shape Algorithm band combinations are used to determine the best SSA band combinations for detecting cyanobacteria in waterbodies during summer/fall reported bloom events. Next, the SSA was calculated for winter/spring non-bloom periods and compared to the summer/fall bloom periods for each lake. Specific cyanobacteria bloom areas within lakes were examined to better understand the magnitude, duration, and spatial extent of the bloom event. Further, meteorological data was used to understand how environmental precursors correspond to bloom formation in these waterbodies. Cross-correlation was determined between meteorological data and SSA values in waterbodies.

2. Materials and Methods

2.1. Study Area

Sentinel-3 OLCI data acquisitions were compiled for nine resolvable waterbodies across six states in the Northeastern United States, shown in Figure 1. The waterbodies were located in the EPA Level 1 North American Ecoregions Northern Forests and Eastern Temperate Forests. According to the Fourth National Climate Assessment in 2018, the largest land-use sectors in the northeast were 54.8% forests, 14.4% agriculture, 10.2% developed, and 10.2% water (Sleeter et al., 2018). General land-use change in the past decades have observed losses in forests, agricultural, and wetland land covers and increases in developed areas (Sleeter et al., 2018). Runoff from agricultural and developed land-use, containing increased nutrient concentrations of nitrogen and phosphorus, accelerate cyanoHABs in the northeast (Lunetta et al., 2015).

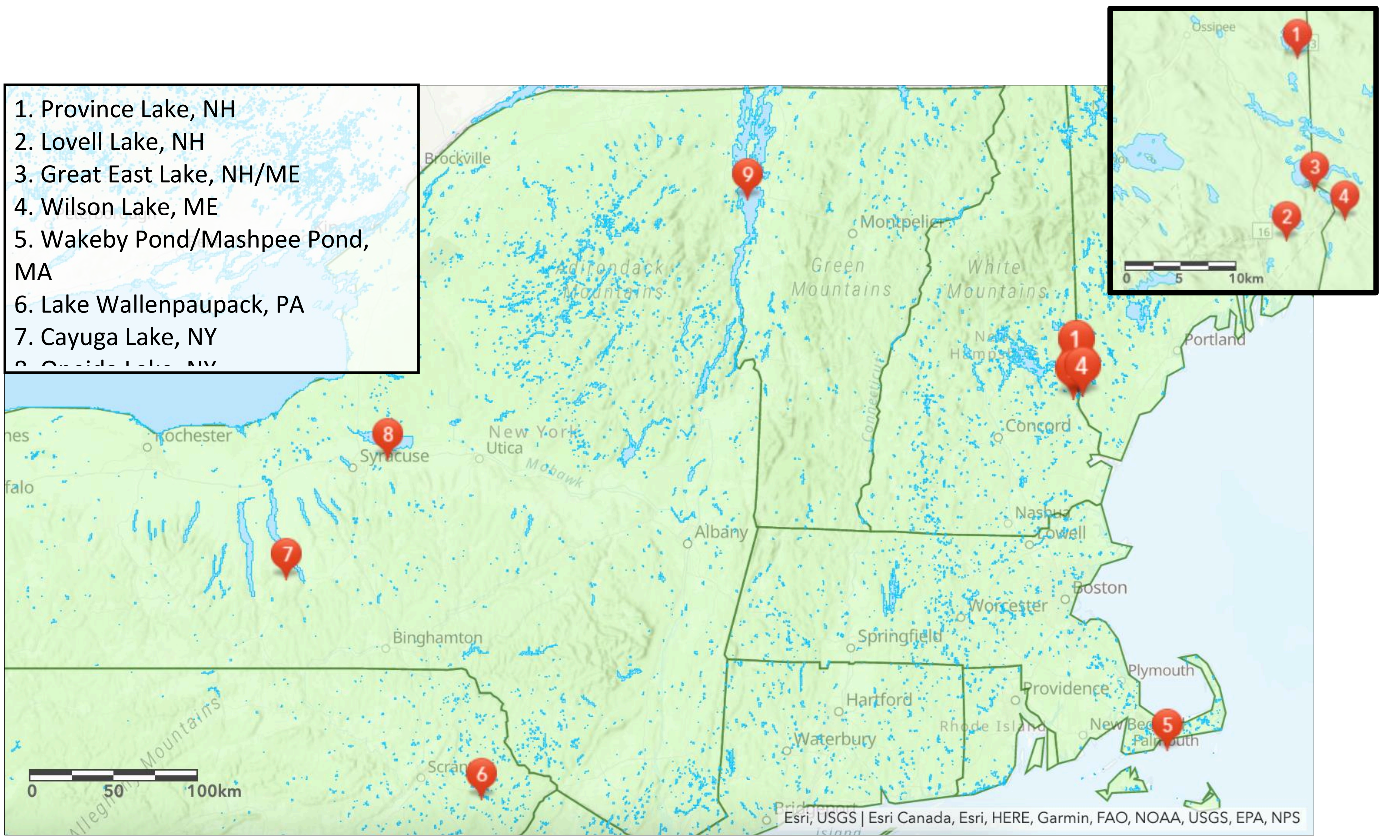


Figure 1. Study Area of 9 Northeastern Waterbodies

The nine waterbodies examined in this study vary in use, that include, water supply, power generation and recreation. The selected waterbodies were all found to be resolvable and have a minimum surface area 0.48 square miles. This study used lakes having 10 pixels or greater to be resolvable. Sentinel-3 OLCI has a resolution of 300 m nadir. The cyanoHAB events in each waterbody varied temporally during the summer/fall months of 2016 to 2020. The characteristics of each waterbody and bloom events can be found in Table 1. The lake area polygons were downloaded from State GIS databases or created with QGIS by manually digitizing the lake extent.

Table 1. Waterbody Surface Area and CyanoHAB Descriptors

	Waterbody	Surface Area (sq. mi.)	Reported CyanoHAB Event Occurrence
1.	Province Lake, NH	1.51	August/September 2019
2.	Lovell Lake, NH	0.84	July-September 2018
3.	Great East Lake, ME/NH	2.86	August/September 2016
4.	Wilson Lake, ME	0.48	August/September 2018
5.	Wakeby Pond, MA	1.14	September 2020
6.	Lake Wallenpaupack, PA	8.91	July-September 2020
7.	Cayuga Lake, NY	66.41	August/September 2019
8.	Oneida Lake, NY	79.80	July-September 2019
9.	Lake Champlain, NY/VT	490	July/August 2019

2.2. Determination of waterbodies experiencing cyanoHABs using citizen scientist apps

Two EPA Mobile Applications, BloomWatch and CyanoScope, were used concurrently to determine waterbodies that experienced cyanoHABs. The BloomWatch App allows citizen scientists to document observed HABs using their smartphones to capture images of the blooms, tag the latitude and longitude location of the bloom, and note additional waterbody and bloom descriptors, including approximate size of bloom and waterbody conditions. The second EPA App, CyanoScope, allows citizen scientists to obtain a water sampling kit, collect water samples of a waterbody experiencing a HAB, and test the sample to identify the cyanobacteria species. The samples species results are then uploaded to an open-source platform with the determined species, location, and time of sample. Both apps encourage interactive communities that inform citizen scientists and stakeholders about cyanoHAB conditions in US waterbodies(US EPA, 2021).

This study used the BloomWatch to first determine which specific waterbodies in the northeastern US were experiencing HABs, the time period of the bloom, and additional bloom descriptors. Then, CyanoScope was used to determine if the algae sample was a cyanobacteria species. The most commonly found cyanobacteria species were Microcystis, Dolichospermum, and Gloeotrichia, which are all capable of producing cyanotoxins. Waterbodies and samples recorded in both apps did not consistently align, and additional resources on specific cyanoHABs were used to fill knowledge gaps, such as State Department of Environmental Protection information on waterbodies experiencing blooms and beach closures due to cyanobacteria in each state. Waterbodies with the most available bloom data and resolvable surface area were subsequently selected to be used with remote sensing methods to detect cyanoHABs (Table 1).

2.3. Remote Sensing Data

This study used the Ocean and Land Color Instrument (OCLI) onboard the Sentinel-3A and Sentinel-3B satellites that were launched in 2016 and 2018, respectively. The sensor has a spatial resolution of 300 m nadir and temporal resolution of 2-3 days. Sentinel-3 OLCI data was downloaded from the Copernicus Open Access Hub (https://scihub.copernicus.eu) provided by the European Space Agency (ESA) Copernicus. The OLCI data were downloaded for weekly observations or finer were used over the entire summer/fall months to capture the broad bloom period. Additional finer-scale time-step data, subject to OLCI data availability, were downloaded for the reported citizen scientist bloom occurrence to capture the exact bloom period.

2.4. ACOLITE Atmospheric Correction

The ACOLITE processor was used for the atmospheric correction of the OLCI data. ACOLITE was developed by the Royal Belgian Institute for Natural Science initially for Landsat and Sentinel-2 imagery, with newer capabilities added to include Sentinel-3 OLCI imagery. ACOLITE was developed to process imagery over coastal areas and inland waterbodies by applying the default dark spectrum (DSF) fitting approach for atmospheric correction (Vanhellemont and Ruddick, 2018, Vanhellemont, 2019). Few studies have used the ACOLITE DSF approach for the atmospheric correction of Sentinel-3 OLCI data, but this atmospheric correction method has been proven effective by Renosh et al. (2020) and Vanhellemont and Ruddick (2021). The DSF approach is an automated process that works by selecting multiple dark spots in the image scene to create a "dark spectrum" that is then used to approximate the atmospheric path reflectance (Vanhellemont, 2019). Python was used to run the ACOLITE processing software and apply the atmospheric correction for the downloaded OLCI imagery.

2.5. Cyanobacteria Detection Algorithm

The Spectral Shape Algorithm (SSA) was used for the detection of cyanobacteria presence in each waterbody, shown in Equation 1,

$$SS(\lambda) = \rho_S(\lambda) - \rho_S(\lambda_-) + \{\rho_S(\lambda_-) - \rho_S(\lambda_+)\} \tag{1}$$

where ρ_S is the top of atmosphere corrected reflectance, λ is the central band, λ_+ is the adjacent upper band, and λ_- is the adjacent lower band. The SSA was first used by Wynne et al. (2008) around the 681 nm fluorescence band, with λ = 681 nm, λ_+ = 709 nm and λ_- = 665 nm. By selecting the center band at 681 nm, the algorithm covered peaks in the chlorophyll-a absorption and fluorescence regions. Water not experiencing blooms has a fluorescence peak at 681 nm, shown in Figure 2; whereas cyanobacteria-laden water has negligible fluorescence. Wynne et al. (2008) defined a baseline between the adjacent upper and lower bands to determine cyanobacteria presence. The presence of cyanobacteria is most likely if the fluorescence band (681 nm) reflectance falls below the baseline between 709 nm and 665 nm, which yields a negative SS(681). Wynne et al. (2008) derived this as the Cyanobacteria Index (CI), so CI = -SS(681). For the instances where cyanobacteria is not present, the reflectance at 681 nm is above the baseline, which yields a positive SS(681). This version of the SSA was found to identify other types of algal blooms, besides cyanobacteria, causing inaccurate SSA values.

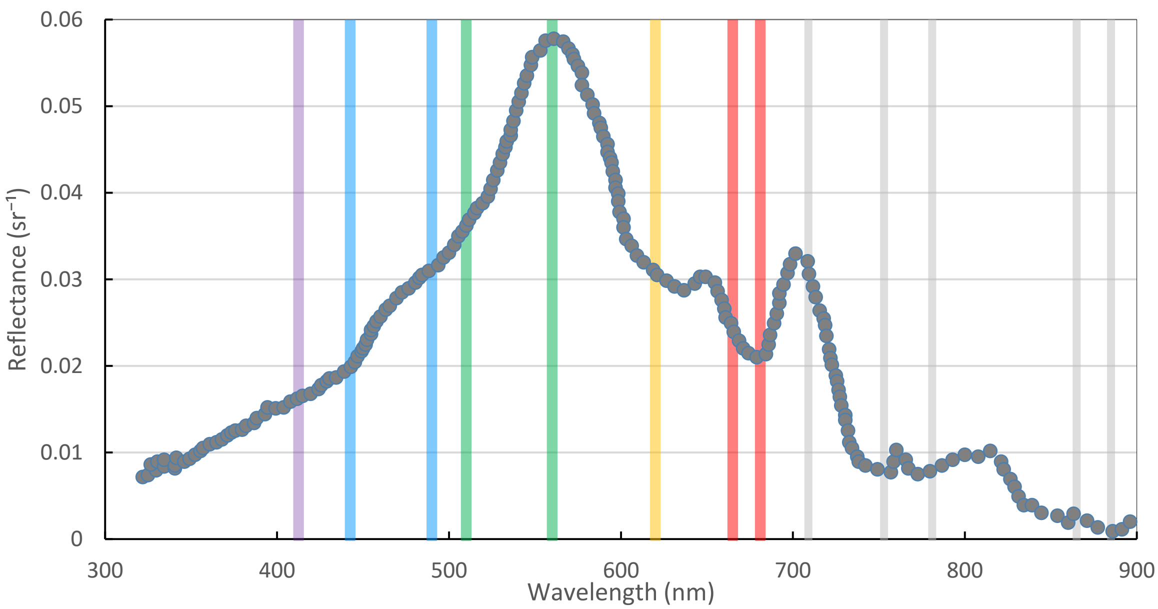


Figure 2. Remote Sensing Reflectance compared to handheld radiometer wavelengths; adapted from Stumpf et al. (2016)

To distinguish cyanobacteria from other algal species, Matthews et al. (2012) first changed the SSA band combinations to include 620 nm, which is sensitive to phycocyanin, in order to separate cyanobacteria from other blooms in African Lakes. Lunetta et al. (2015) is responsible for devising the Matthews et al. (2012) band updates as the CI-multi. Lunetta et al. (2015) used the changed center band at 665 nm and the adjusted baseline between 620 nm and 681 nm, and so λ = 665 nm, λ_+ = 681 nm and $\lambda\_$ = 620 nm. The Lunetta et al. (2015) SSA adjustment accounted for poor atmospheric conditions and was more sensitive to phycocyanin, which strongly absorbs at 620 nm (Figure 2). The increased phycocyanin absorption was found to decrease the reflectance at 620 nm, which changes the SS(665) from negative to positive. (Simis et al., 2005). For the Lunetta et al.

(2015) SSA center band method, cyanobacteria are assumed present when SS(665) is positive, and absent when SS(665) is negative, which is the reverse of the Wynne et al. (2008) SSA band combinations.

2.6. Cyanobacteria Bloom Analyses

Three different analyses were used for cyanobacteria detection in the nine study lakes: (1) SSA center band method comparison; (2) winter/spring non-bloom period comparison; (3) determining specific bloom areas within each lake. First, both SSA center band methods (Wynne et al., 2008 and Lunetta et al., 2015) were used to calculate SSA values for pixels within each lake over the reported summer/fall bloom period. The performance of each method was compared, and SSA results were analyzed to determine trends regarding lake size and location. A winter/spring non-bloom period was selected for each lake, and the SSA was calculated for each pixel within the lake polygon for this non-bloom period. The SSA non-bloom and bloom period values were compared for each lake. For larger lakes, the reported cyanoHAB location was marked and used to determine the exact bloom area. Smaller polygons of these lakes were created in QGIS to better visualize the magnitude and duration of the cyanoHABs.

2.7. Environmental Precursor Methods

Lagged cross-correlations were calculated to determine the relationship of meteorological environmental precursors and cyanobacteria presence/absence within each lake. Maximum temperature, minimum temperature, and precipitation daily summaries were downloaded from the NOAA NCDC database for the closest weather station to each lake. The SSA statistics, namely mean, median, max, and min, were calculated for the lake bloom periods. The meteorological variables were lagged and the cross-correlation was determined between the meteorological environmental precursors and the SSA statistic values. Cross-correlation trends were determined to provide comprehension of how environmental precursors affect cyanoHAB formation.

3. Results

3.1. Cyanobacteria Detection Results

3.1.1. SSA Method Comparisons

The Wynne et al. (2008) and Lunetta et al. (2015) SSA center band methods were calculated for each lake. The Sentinel-3 OLCI spectral band combinations and cyanobacteria detection criteria for each method are shown in Table 2.

Table 2. Sentinel-3 OLCI Spectral Bands and Cyanobacteria Detection Criteria for each SSA Method

SSA Method	Spectral Bands	Cyanobacteria Detection Criteria
Wynne et al., 2008	λ = 681 nm λ_{-} = 665 nm λ_{+} = 709 nm	SS(681) > 0 – cyanobacteria absent SS(681) < 0 – cyanobacteria present
Lunetta et al., 2015	λ = 665 nm λ_{-} = 620 nm λ_{+} = 681 nm	SS(665) > 0 – cyanobacteria present SS(665) < 0 – cyanobacteria absent

Previous studies determined that the Lunetta et al. (2015) SSA method performed better than the Wynne et al. (2008) SSA method. Both SSA methods were used to examine cyanobacteria detection and evaluate lake size and regional trends for lakes in this study (Mishra et al., 2019; Coffer et al., 2020). The Lunetta et al. (2015) SSA method is more sensitive to phycocyanin and distinguishes cyanobacteria from other algal blooms, whereas the Wynne et al. (2008) SSA method can detect other algal species besides cyanobacteria. Both methods were used for cyanobacteria detection in this study to evaluate their cyanobacteria detection ability and confirm the better performance of the Lunetta et al. (2015) algorithm. Because this study used citizen scientist data, not in situ sampling data, to select and guide the cyanobacteria detection in each lake, both SSA methods were used to provide better understanding of the cyanoHABs in the study area lakes.

The SSA center band methods for Oneida Lake and Province Lake are shown in Figures 3 through 6. Both SSA methods were applied for every pixel over the reported cyanoHAB event in each lake. These lakes were selected for comparisons due to their differing sizes, lake characteristics, and geographic locations. Oneida Lake is in Upstate New York northeast of Syracuse, near Lake Ontario. The lake area is 79.8 square miles and has a maximum and average depth of 55 ft and 22 ft, respectively (NY DEC, 2021). Cyanobacteria was reported in the summer/fall month of 2019 in the eastern portion of the lake. Province Lake is located on the border of New Hampshire and Maine, near Effingham, NH. The lake area is 1.51 square miles and has a maximum and average depth of 16 feet and 9 feet, respectively (NHFGD, 2013). Cyanobacteria was reported in Province Lake mid-August 2019.

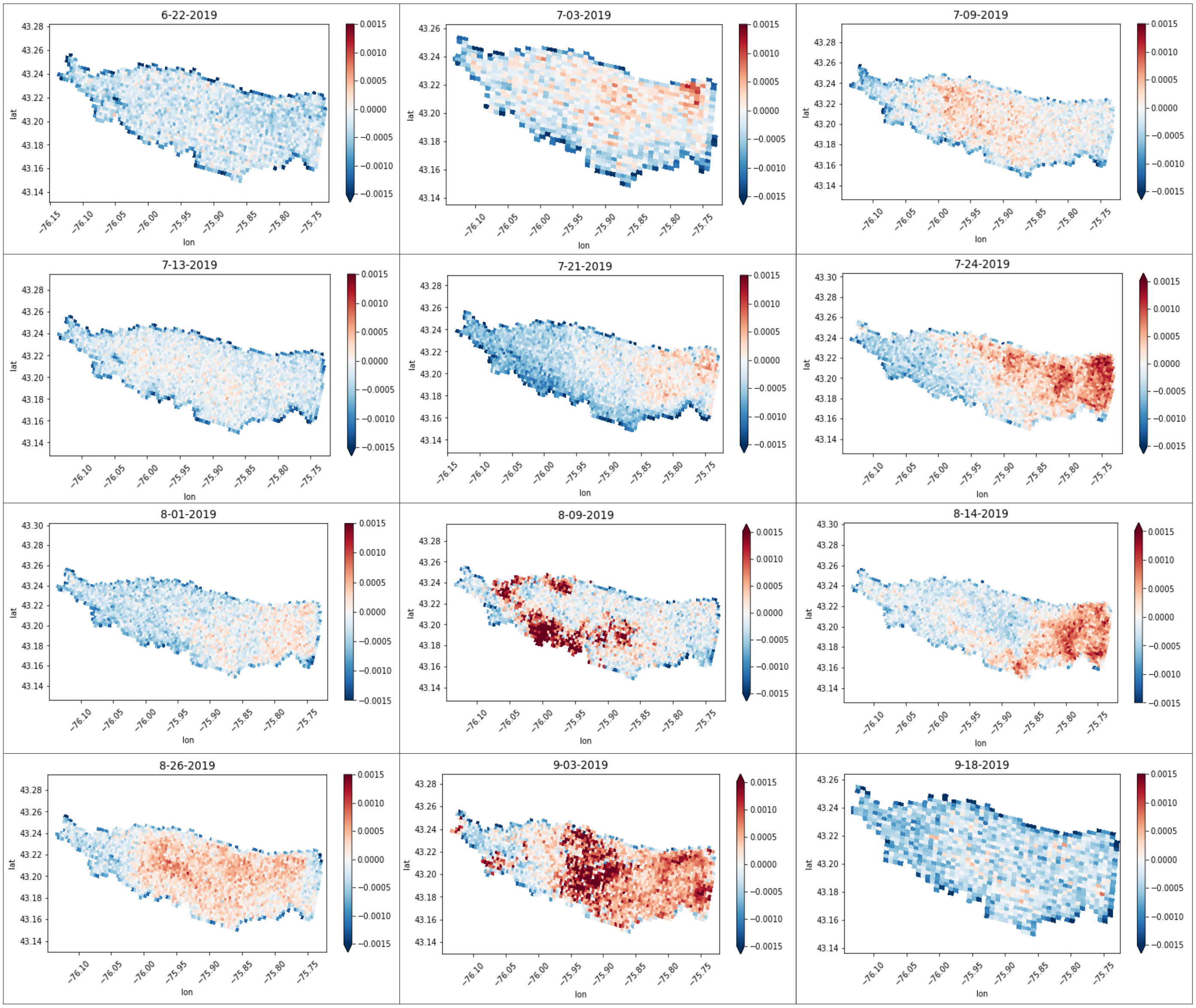


Figure 3. Lunetta et al. (2015) SSA Method for the Oneida Lake, NY 2019 CyanoHAB

Positive SSA values, shown as red pixels on the Lunetta et al. (2015) maps, indicate the presence of cyanobacteria, whereas negative SSA values, shown as blue pixels on the Lunetta et al. (2015) maps, indicate the absence of cyanobacteria. The Lunetta et al. (2015) method indicates cyanobacteria throughout the summer/fall of 2019, with a highly concentrated cyanoHAB in the eastern portion of the lake. The Lunetta et al. (2015) SSA values show the 2019 cyanoHAB intensity peaking on multiple days and moving to the west across the lake throughout different dates of the bloom period. Cyanobacteria was detected in early-July 2019 and remained through early-September 2019. The Oneida Lake 2019 cyanoHAB had multiple peak events in the eastern portion of the lake on July 24th, August 14th, and September 3rd before dissipating by September 18th. The New York Department of Environmental Conservation (NY DEC) reported suspicious, confirmed or confirmed with high cyanotoxin concentrations in Oneida Lake every year since 2013 (NY DEC, 2019). In 2019, the NY DOH reported 32 beach closure days for Oneida Lake due to cyanobacteria (2020).

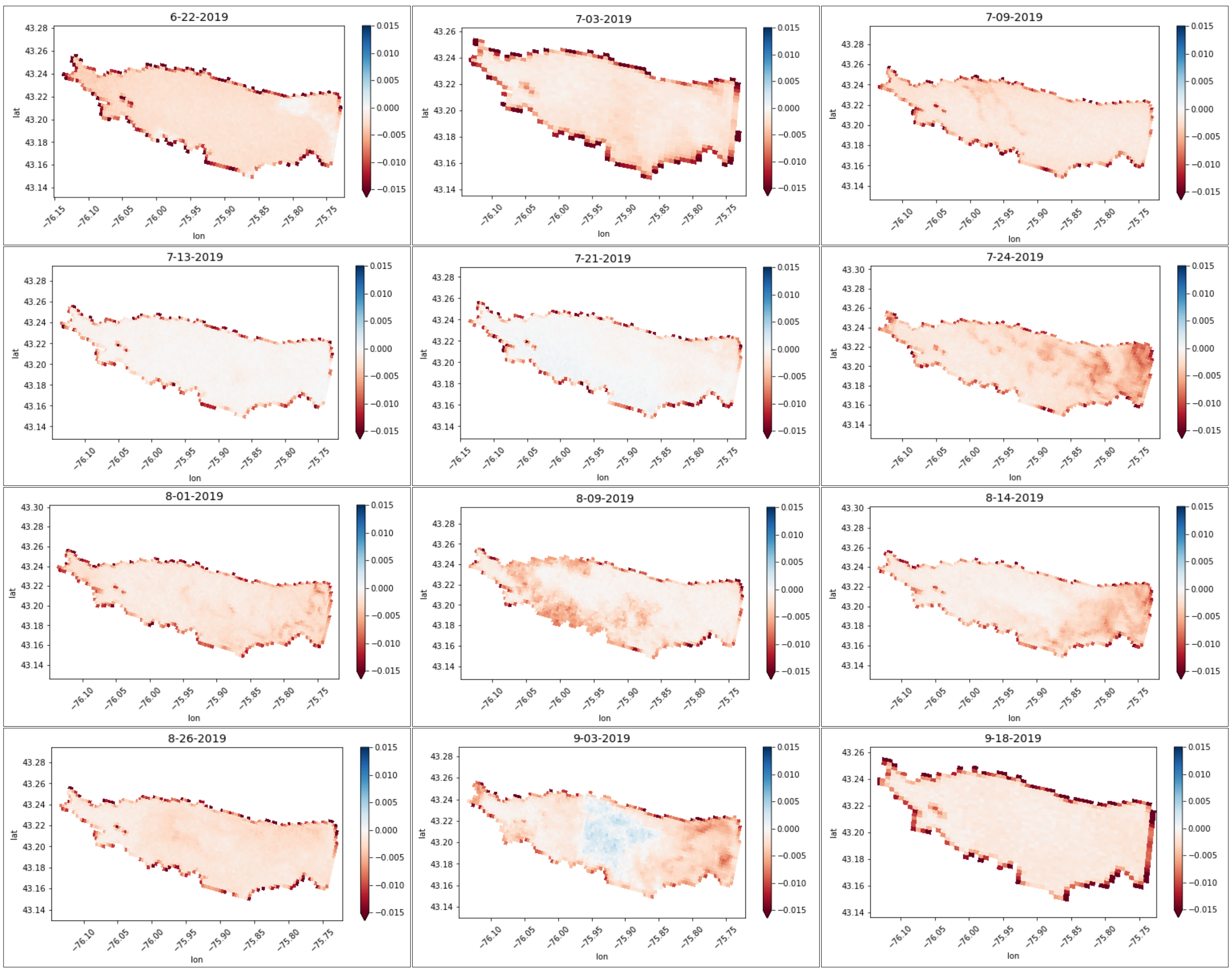


Figure 4. Wynne et al. (2008) SSA Method for the Oneida Lake, NY 2019 CyanoHAB

The Wynne et al. (2008) SSA negative values, shown in red pixels, indicate cyanobacteria presence, whereas the SSA positive values, shown in blue pixels, indicate cyanobacteria absence (Figure 4). For the Oneida Lake 2019 cyanoHAB event, the Lunetta et al. (2015) SSA method performs better than the Wynne et al. (2008) SSA method for cyanobacteria detection. The Wynne et al. (2008) SSA method indicated the same cyanoHAB pattern with peak events in the eastern portion of the lake; however, the Wynne et al. (2008) SSA pixel values over Oneida Lake are generally negative or near zero overpredicting cyanobacteria presence. Additionally, lake pixels along the water coastline interface are largely negative, affecting the SSA scale values of the heatmaps. The Wynne et al. (2008) SSA method calculated false negative values for Oneida Lake, and confirms that the Lunetta et al. (2015) SSA method better predicts cyanobacteria in Oneida Lake.

Consistent with Oneida Lake, calculated Lunetta et al. (2015) SSA positive pixel values in Province Lake, indicate cyanobacteria presence, and the SSA negative pixel values indicate cyanobacteria absence. The Lunetta et al. (2015) SSA method detects cyanobacteria presence in Province Lake starting in July with a significant cyanoHAB peak occurring on September 9, 2019. On the EPA App CyanoScope, a citizen scientist reported the cyanobacteria species *Anabaena* on August 9, 2019 in Province Lake. Additionally, the NH DES issued a cyanobacteria advisory for Province Lake on July 25, 2019 based on collected water samples from the same day. The samples collected contained the cyanobacteria species *Anabaena* with a concentration above the New Hampshire state threshold of 70,000 cells/mL of cyanobacteria. The advisory lasted from July 25 to July 31, 2019 (NH DES, 2019). The Lunetta et al. (2015) SSA map contains positive pixel values during this advisory period, indicating cyanobacteria presence.

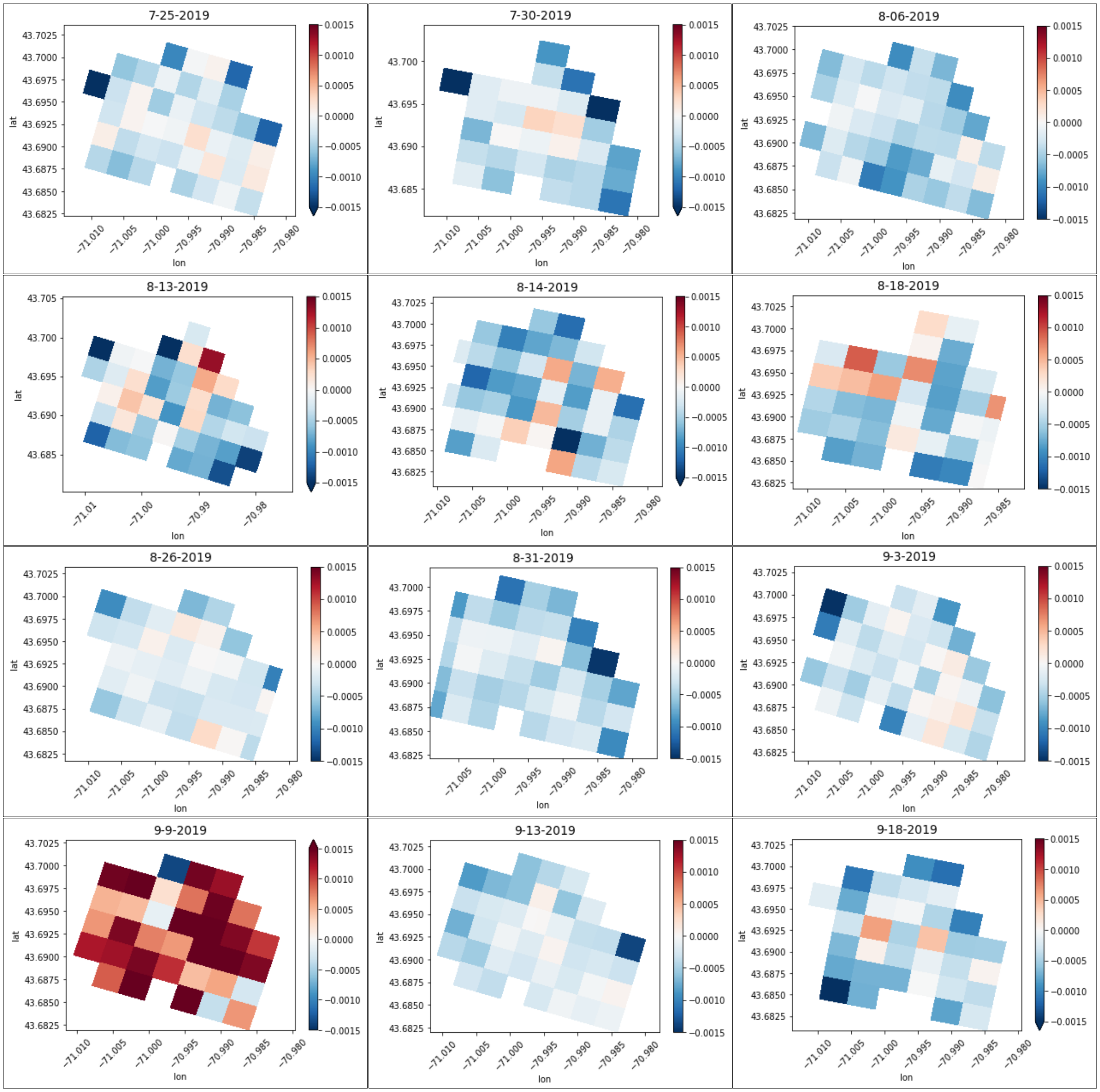


Figure 5. Lunetta et al. (2015) SSA Method for Province Lake, NH 2019 CyanoHAB

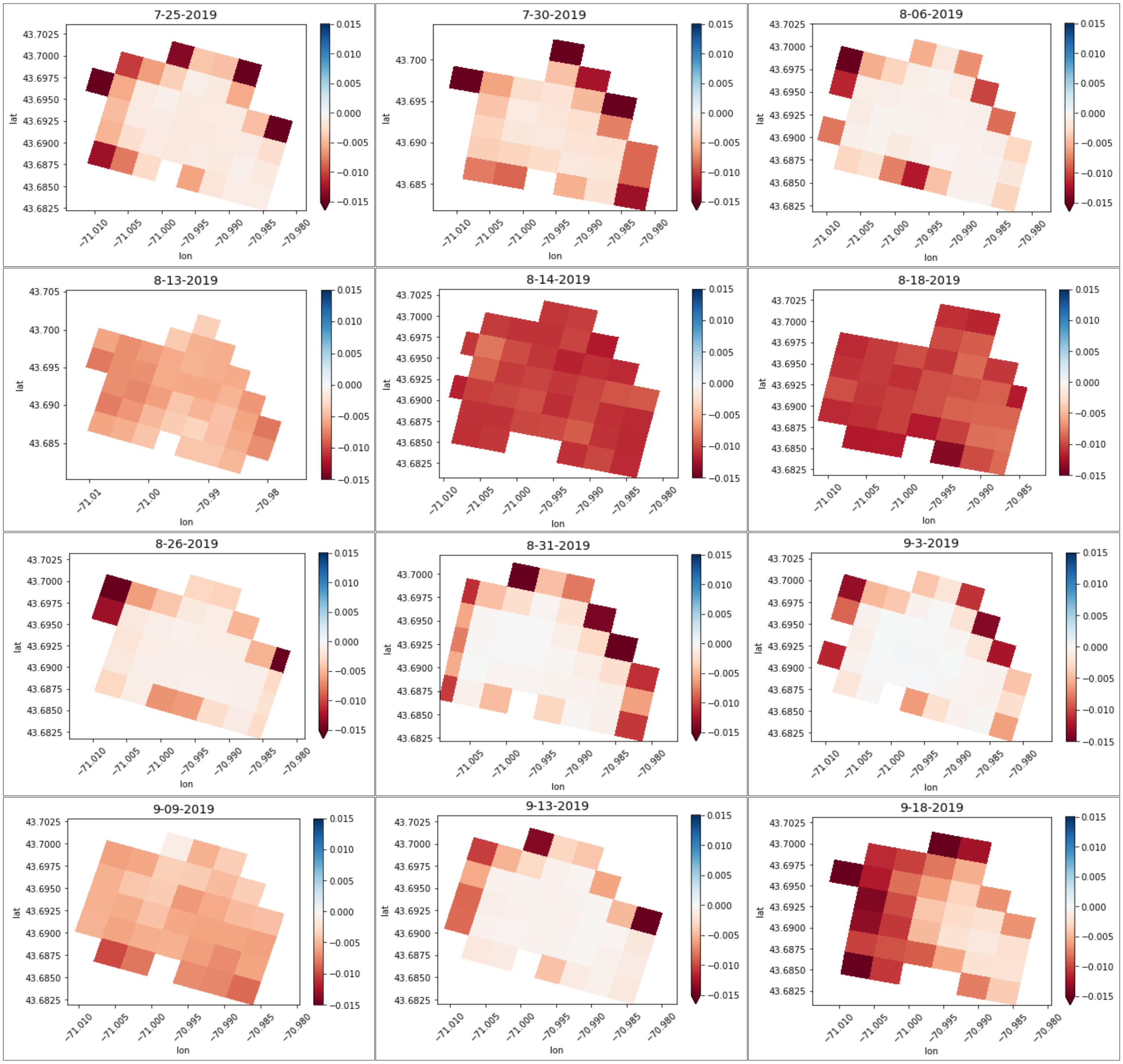


Figure 6. Wynne et al. (2008) SSA Method for Province Lake, NH 2019 CyanoHAB

The negative SSA pixel values are shown in red, indicating cyanobacteria presence for the Wynne et al. (2008) SSA method. The Wynne et al. (2008) SSA values do not match the cyanobacteria detection trends shown in the Lunetta et al. (2015) SSA values (Figure 5 and 6). The Wynne et al.

(2008) method overestimates cyanobacteria presence as all lake pixel values for Province Lake are negative. For both Oneida Lake and Province Lake, which differ in lake size, depth and location, the Lunetta et al. (2015) SSA method performed better than the Wynne et al. (2008) SSA method for cyanobacteria detection. The SSA method comparison results for Oneida Lake and Province Lake are representative for all the lake in this study, with the Lunetta et al. (2015) performing better than the Wynne et al. (2008) method, and the Wynne et al. (2008) method flagging false negative SSA values. Additional SSA method comparisons of Lunetta et al. (2015) and Wynne et al. (2008) can be found in Appendix A and B. For the rest of the analysis, we used the Lunetta et al. (2015) SSA method for the cyanobacteria detection.

3.1.2. SSA Winter/Spring Comparison

The SSA was calculated using the Lunetta et al. (2015) SSA method for a non-bloom period during the winter/springtime to compare with the SSA summer bloom period results. The non-bloom period was selected for each lake from February to May when there was no reported or expected cyanobacteria in these waterbodies. Past ecological trends suggest that cyanoHABs usually occur during warmer months in the summer and early fall. The non-bloom period Lunetta et al. (2015) SSA comparisons for Oneida Lake and Province Lake, the same diverse representative lakes as used in section 3.1.1., are shown in Figures 7 through 9.

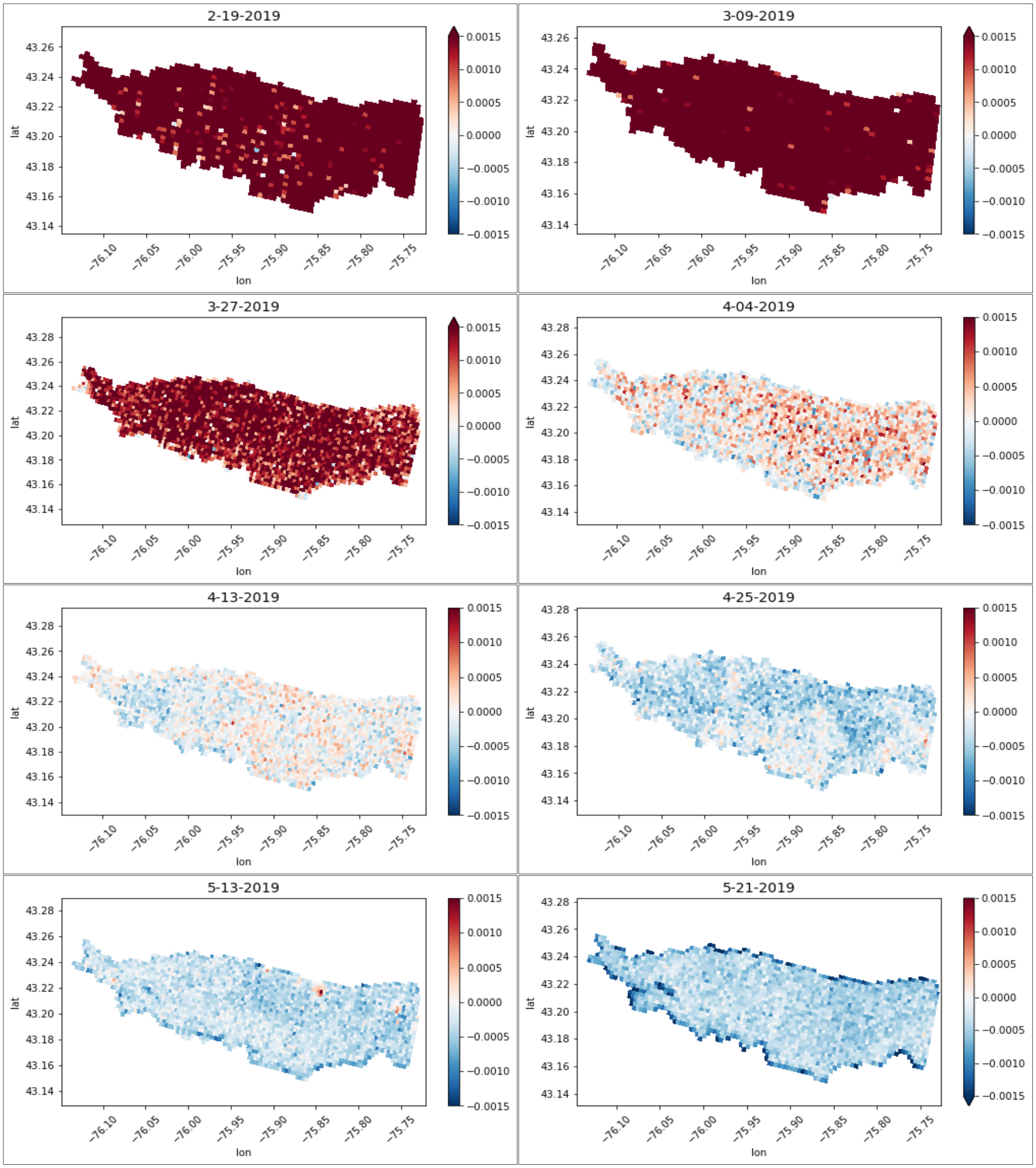


Figure 7. SSA Values over Oneida Lake Non-Bloom Period 2019

Oneida Lake experienced ice and snow cover over the lake surface area into April in 2019. The ice and snow cover result in false positive SSA values, shown in red, for February 19, 2019 through April 4, 2019. As the ice and snow melts, the SSA values return to negative values, indicating clear water, free of ice, snow, and cyanobacteria. The reflectance values for the Sentinel-3 OLCI SSA spectral bands (620, 665, and 681 nm) for snow and ice cover are shown in Figure 8.

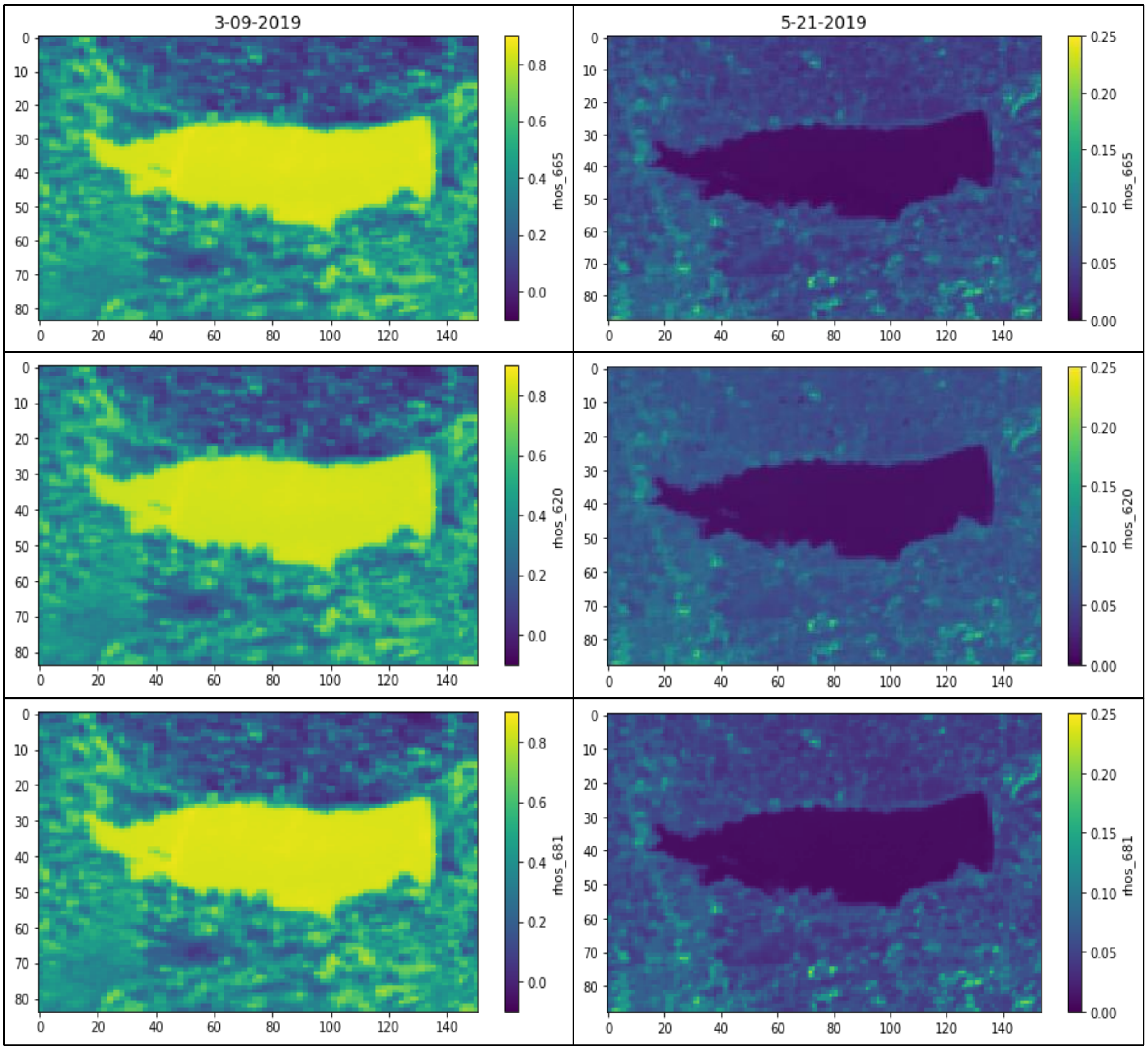


Figure 8. SSA Reflectance Band Values for Oneida Lake Ice and Snow Cover on March 9, 2019 (left side) compared to Oneida Lake Clear Water (no ice and snow cover) on May 21, 2019 (right side) .

The SSA reflectance values for snow and ice cover over Oneida Lake on March 9, 2019 are significantly greater than the SSA reflectance values for clear water over Oneida Lake on May 21, 2019 (Figure 8). Snow and ice cover have much larger reflectance than water in the visible spectrum, and thus, the increased reflectance values for snow and ice cover result in positive SSA values.

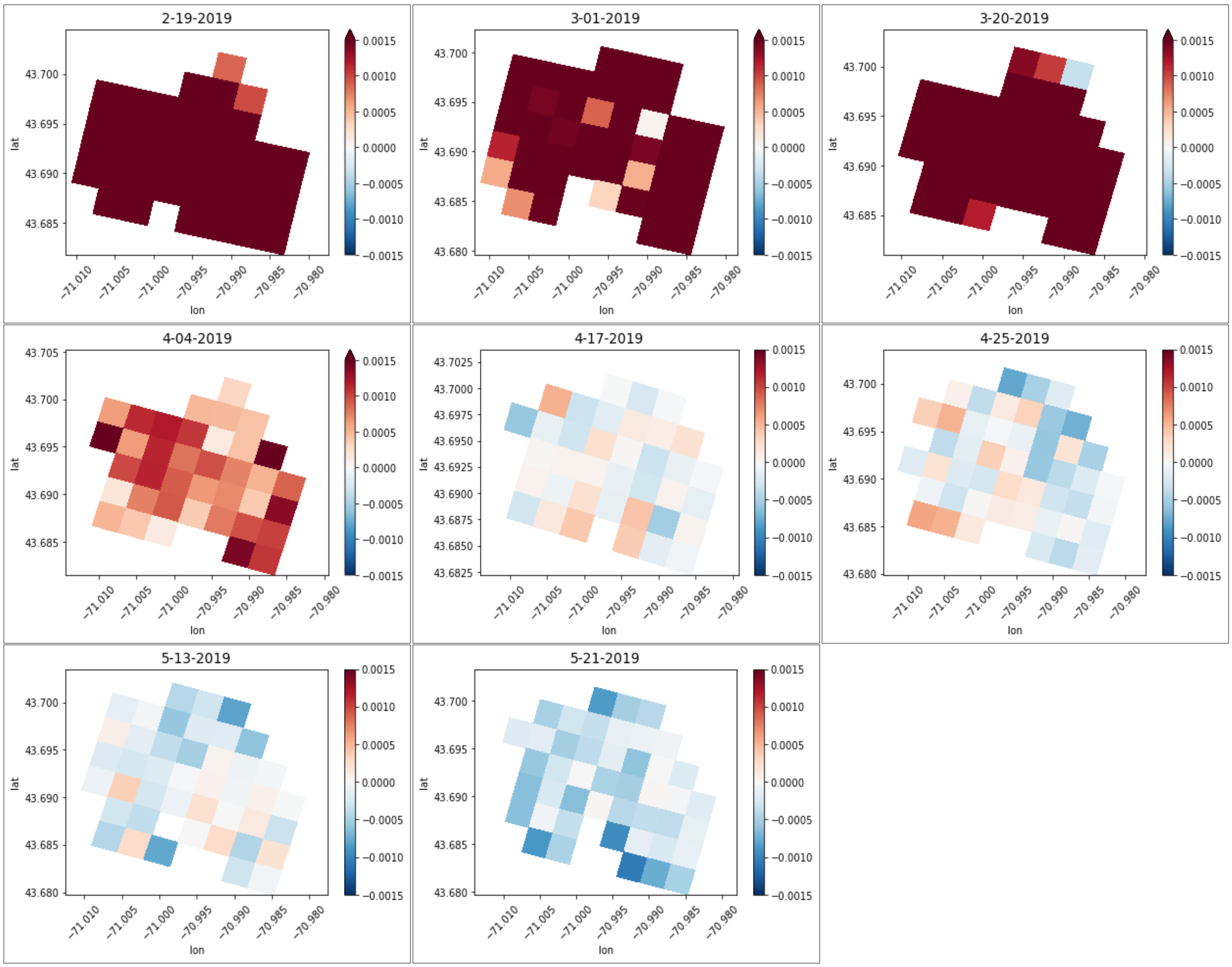


Figure 9. SSA Values over Province Lake Non-Bloom Period 2019

The SSA values for the non-bloom period for Province Lake are shown in Figure 9. Province Lake experiences ice and snow cover into April, similar to Oneida Lake, before returning to clear water in May. Other lakes in this study had similar snow and ice cover patterns for the non-bloom period. The non-bloom period effectively shows SSA trends in the study lakes during the winter and spring months. The lakes show negative SSA values, indicating clear water without cyanobacteria present once the snow and ice have melted. Due to snow and ice cover during winter months, the non-bloom comparison lacks data to be compared cyanoHAB ecological patterns. Ecological patterns indicate cyanoHABs occur in the summer and fall month, disappear over the winter into spring, and do not form again until the next summer in warmer months.

3.1.3. Determining Bloom Areas within Lakes

Larger lakes in the study area, including Lake Champlain and Cayuga Lake, required further analysis to identify the reported cyanoHAB event. Smaller polygons were created to identify and examine cyanobacteria presence using the Lunetta et al. (2015) SSA method. Both the entire lake polygon and the smaller reported bloom area polygons are shown for Lake Champlain and Cayuga Lake (Figures 10 through 13). The SSA was calculated over the entire Lake Champlain surface area (Figure 10). Lake Champlain has experienced an increase in cyanoHABs in recent years (Bockwoldt et al., 2018). Citizen scientist data have been reported in the BloomWatch and CyanoScope apps beginning in 2017. In 2019, cyanobacteria were detected along the beaches near Burlington, VT. The City of Burlington Parks, Recreation and Waterfront reported beach closures for North, Texaco, and Leddy beaches on July 12, 2019 continuing through July 23, 2019 (2021). The Burlington beaches area used for the bloom area analysis is shown in Figure 11. The SSA was calculated over the Burlington, VT beaches to identify cyanoHABs that were associated with the 2019 Burlington beach closures (Figure 12).

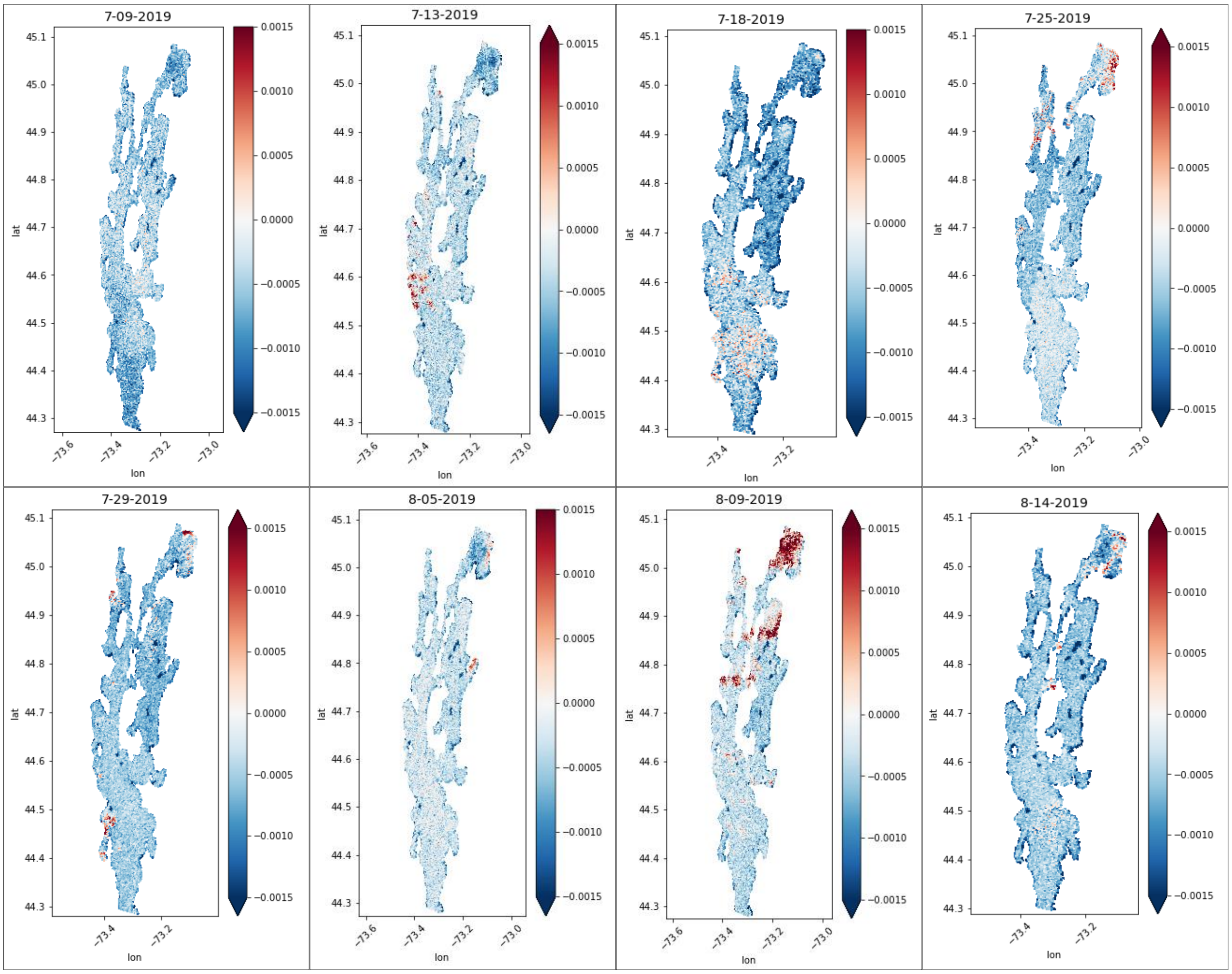


Figure 10. SSA over Lake Champlain (Entire Lake Area) for 2019 CyanoHAB

Figure 11. Burlington Beaches Sub-Area

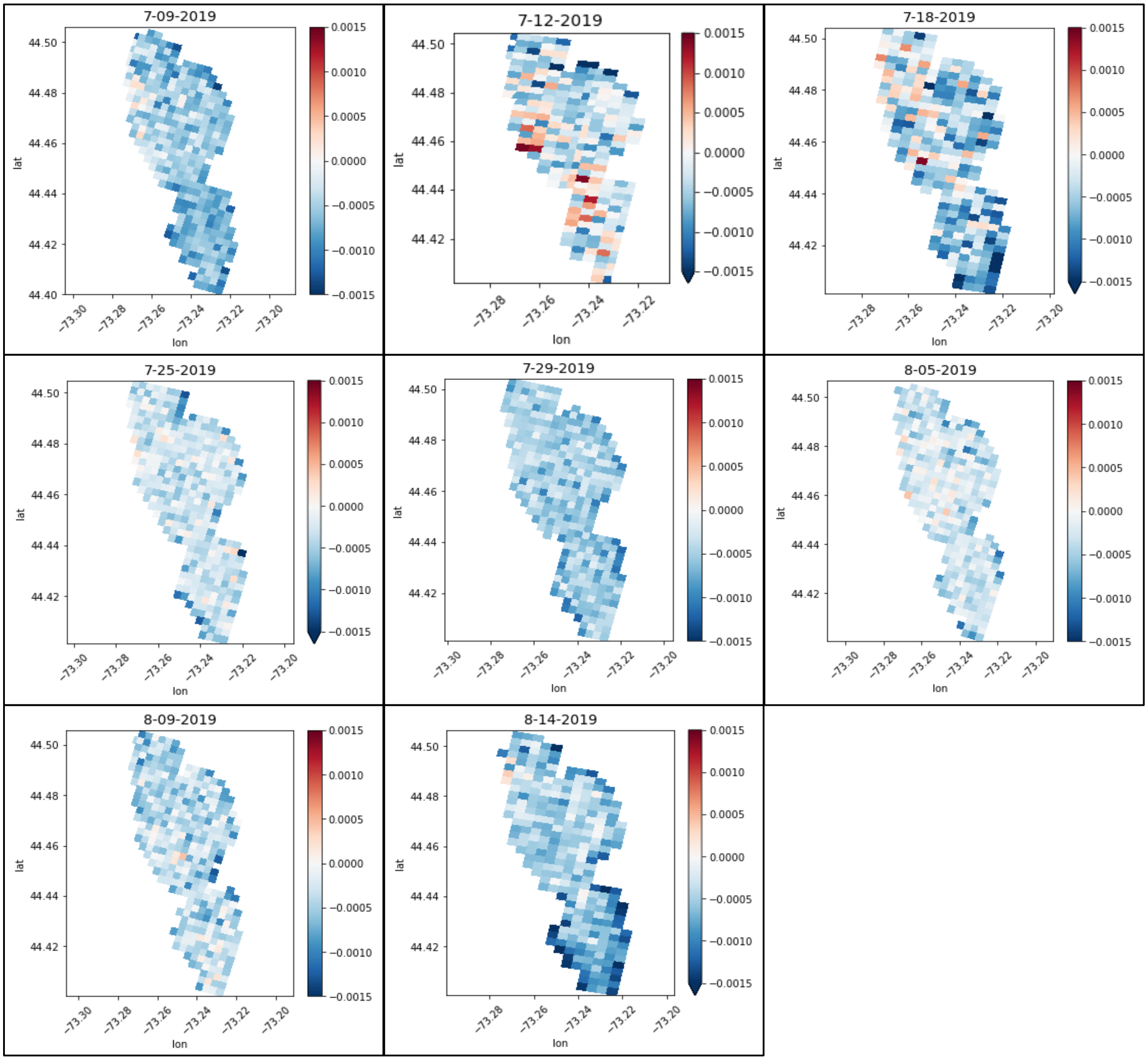


Figure 12. SSA over Lake Champlain (Smaller Reported Bloom Area) for 2019 CyanoHAB

The SSA values for the Burlington bloom area indicate cyanobacteria presence within the timeframe of the Burlington beach closures on July 18th. The cyanobacteria detection did not indicate a large bloom in this area but rather sporadic positive pixels in the SSA results over the reported beach closure period. Lake Champlain is the largest lake in the study, and the smaller Burlington beaches sub-area is still approximately eight miles in length, with a larger surface area than many of the other lakes in this study. Because of the larger size of this sub-area and availability of Sentinel-3 OLCI data, the cyanoHAB event that caused the Burlington beach closures on July 12th may not have been accurately detected in the SSA results.

The SSA was calculated over Cayuga Lake over July 9th to September 6th to capture the 2019 cyanoHAB event (Figure 13). Cyanobacteria was detected in Cayuga Lake beginning August 13th continuing throughout the summer/fall bloom period, before dissipating in early September. The detected cyanobacteria are concentrated mainly in the northern and southern portions of Cayuga Lake with peak bloom events covering much of the lake surface, observed on August 22, 2019. The New York DEC has tracked harmful algal blooms throughout New York waterbodies since 2012 and reports the waterbody bloom status as suspicious bloom, confirmed bloom, or confirmed with high toxin blooms. The New York DEC reported HABs in Cayuga Lake beginning in 2014, and confirmed with high toxin cyanoHABs every year since 2017 in Cayuga Lake (NY DEC, 2019). Citizen scientist data reported cyanobacteria species *Microcystis* and *Dolichospermum* across Cayuga Lake on August 15-16, 2019, with many of the reports in the southern portion of the lake, specifically Taughannock Falls State Park beaches and beaches near Ithaca, NY.

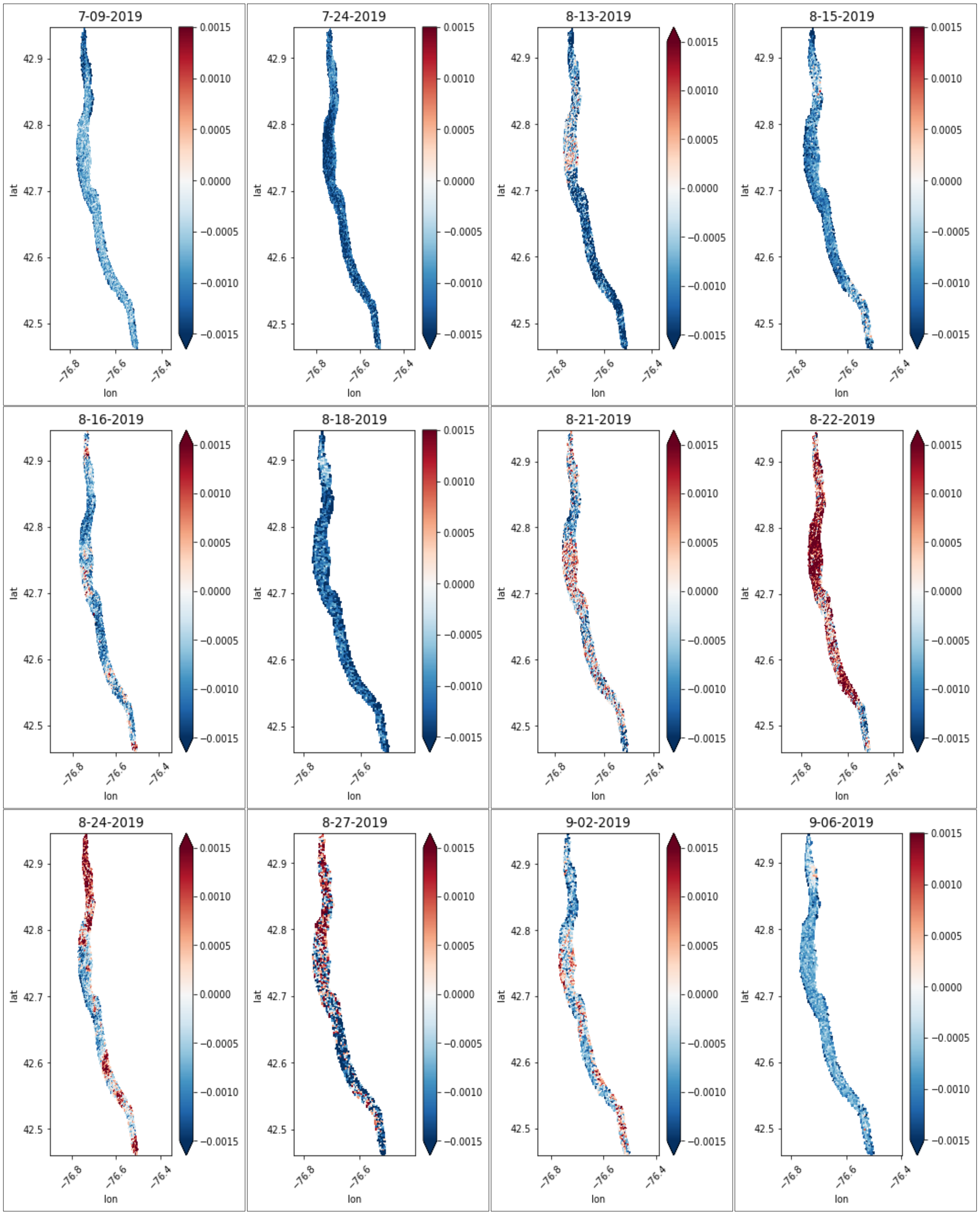


Figure 13. SSA over Cayuga Lake (Entire Lake Area) for 2019 CyanoHAB

The cyanoHAB in the southern portion of Cayuga Lake was isolated for further analysis due to the citizen scientists' cyanobacteria reports and public health concerns; southern Cayuga Lake is a critical drinking water source and has many popular beaches for recreation and tourist attractions.

Southern Cayuga Lake supplies drinking water to the majority of Tompkins County, including the city of Ithaca and Cornell University. Currently, there have not been any reports of finished drinking water that have exceeded the New York microcystin health advisory limit of 0.3 µg/L. However, microcystin concentrations have been reported in raw water prior to treatment (NY DEC, 2019). The southern bloom area is shown in Figure 14. The SSA was calculated over the southern Cayuga Lake bloom area (Figure 15).

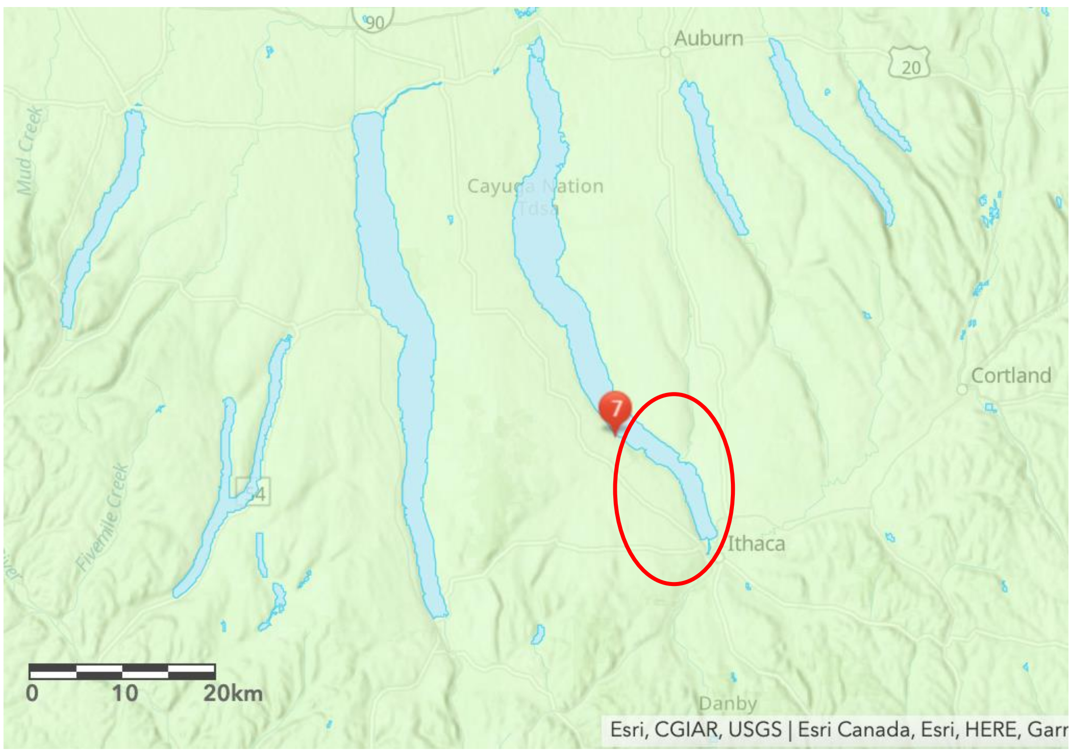


Figure 14. Cayuga Lake Southern Bloom Area

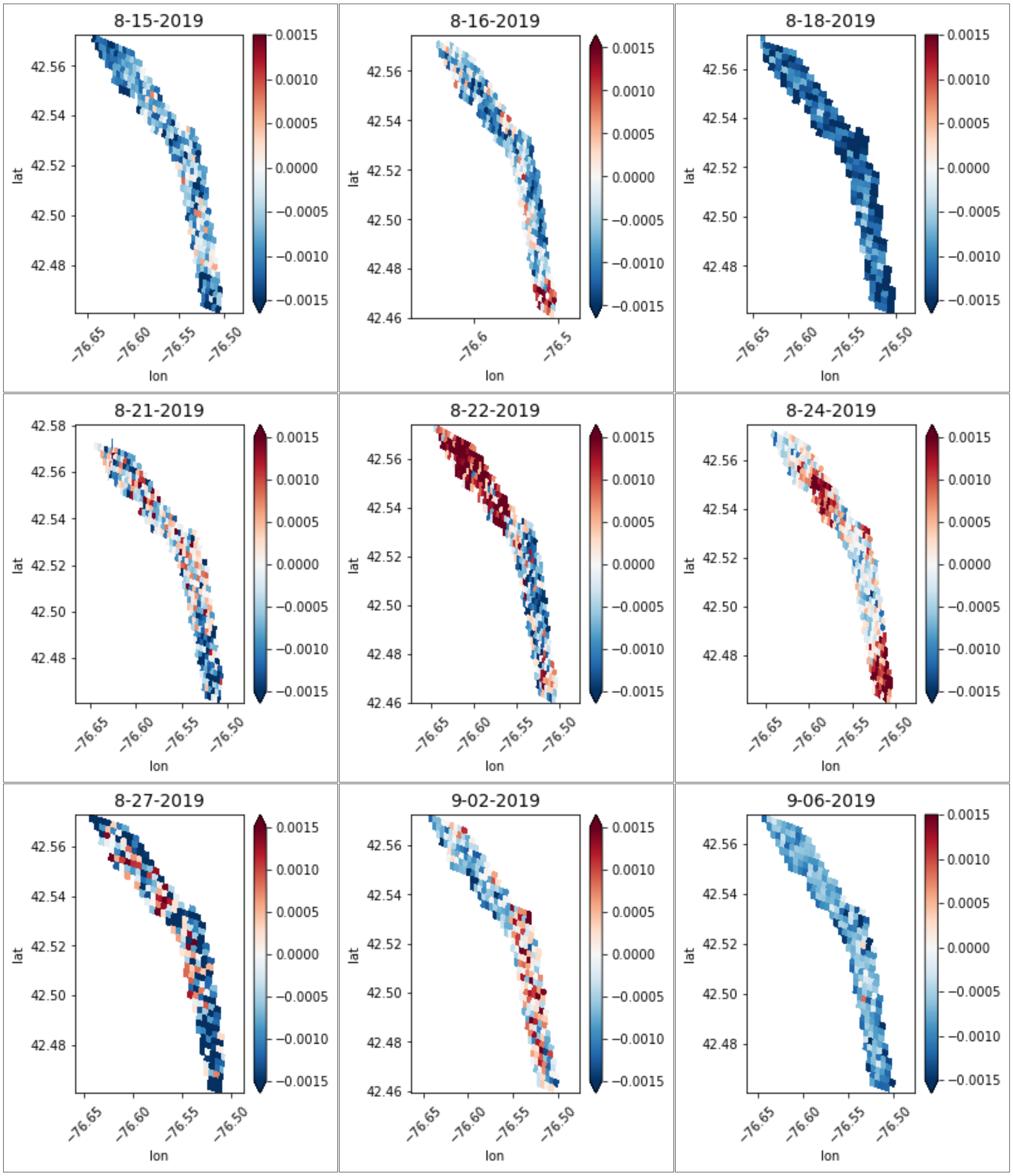


Figure 15. SSA over Cayuga Lake (Smaller Reported Bloom Area) for 2019 CyanoHAB

Cyanobacteria was detected beginning August 15, 2019 and continuing over the southern Cayuga Lake bloom area until early September. Over the bloom period, cyanobacteria were concentrated in the southern parts of the lake near drink water intakes and popular recreation beaches located in the top left of the smaller bloom polygon. The calculated SSA values support the CyanoScope citizen scientists reports of cyanobacteria in southern Cayuga Lake. The bloom area analysis is beneficial in identifying specific cyanoHABs in lager waterbodies in this study area. The specific bloom area analysis allows for more comprehensive documentation of cyanoHABs that can be used to validate citizen scientist reports and inform water managers and public health official of cyanoHABs.

3.2. Cyanobacteria Detection Validation

As validation of the remote sensing cyanobacteria detection performance, the agreement between the citizen scientist cyanoHAB observations, State reports of cyanobacteria, and remote sensing cyanobacteria detection were analyzed for each lake (Table 3). The reported citizen scientist observation date and location were used for the validation, as well as the remote sensing detection for the corresponding date and lake pixels for location. Additionally, resources from specific state warnings and beach closures for individual waterbodies and cyanoHAB events were recorded to provide further descriptors for the reported cyanoHAB events.

Table. 3 Remote Sensing CyanoHAB Detection Validation

Waterbody	CyanoHAB Observation Date(s)	CyanoHAB Observation Location	Did remote sensing detect the cyanoHAB?	Additional Information
Province Lake, NH	July 25, 2019	Center of lake	Yes	NH DES[2] cyanobacteria advisory July 25-30, 2019
Lovell Lake, NH	Aug 20, 2018 Sept 4, 2018	Eastern cove	Yes Unknown[1]	N/A
Great East Lake, ME/NH	July 18, 2016 Aug 17, 2016	Center of lake	Yes Unknown[1]	N/A
Wilson Lake, ME	July 23, 2018 Sept 4, 2018	Center of lake	Yes Unknown[1]	N/A
Wakeby Pond, MA	Sept 29, 2020	N/A	Yes	MA DPH[3] posted lake closure due to toxic cyanoHAB on September 29
Lake Wallenpaupack, PA	July 25, 2020 Aug 22, 2020	Southwestern portion of lake	Yes Yes	PLEON[4] reported lake-wide cyanoHABs throughout July and August 2020
Cayuga Lake, NY	July 11, 2019 Aug 16, 2019	Southern portion	Unknown[1] Yes	NY DEC[5] reported highly toxic cyanoHABs during July-September 2019
Oneida Lake, NY	July 21, 2019 July 26, 2019 Aug 14, 2019	Eastern portion near Verona Beach	Yes Unknown[1] Yes	NY DEC reported highly toxic 2019 cyanoHAB; NY DOH[6] reported 32 beach closure days in 2019 due to cyanobacteria
Lake Champlain, NY/VT	July 12, 2019 July 18, 2019	7/12 – North, Texaco and Leddy Beach; 7/18 – Oakledge Cove and Blanchard Beach	Yes Yes	Burlington Parks, Recreation and Waterfront[7] reported beach closures for North, Texaco and Leddy Beach on July 12; North and Texaco beaches remained closed through July 23; Oakledge Cove and Blanchard Beach closed on July 18

1. No Sentinel-3 OLCI Data available for this exact date.
2. NH DES, 2019
3. Town of Sandwich, MA, 2020
4. LACAWAC, 2020
5. NY DEC, 2019
6. NY DOH, 2020
7. Burlington, VT Parks, Recreation and Waterfront, 2021

The accuracy classification metric was used to evaluate the remote sensing cyanobacteria detection performance. The accuracy can be defined as the fraction of cyanoHAB events that are correctly detected by the remote sensing methods over the total number dates that citizen scientist reported cyanobacteria throughout the study waterbodies. The accuracy of the remote sensing cyanoHAB detection was calculated for each citizen scientist reported observation date and location and the corresponding remote sensing detection for that date and pixels corresponding to the specified location. The accuracy for the remote sensing cyanoHAB detection was 71%. The main limitation affecting the accuracy of the remote sensing cyanobacteria detection methods was the availability of the Sentinel-3 OLCI data, marked as "Unknown" in Table 3. The remote sensing cyanoHAB detection patterns leading up to and following the absent Sentinel-3 OLCI data, when the cyanoHABs were reported, commonly show evidence of cyanobacteria present on the missing data; however, it can not be certain due to the lack of Sentinel-3 OLCI data on specific dates. When excluding the missing data points, the accuracy of the remote sensing cyanobacteria detection was 100%.

3.3. Environmental Precursor Results

Lagged cross-correlations were calculated between three pairs of variables: maximum temperature and maximum SSA values, precipitation and minimum SSA values, and precipitation and maximum SSA values. The Lunetta et al. (2015) method was used to calculate SSA statistics values for every study lake during the selected summer/fall bloom period. The maximum SSA value, or most positive, corresponds to cyanobacteria detection in the study lakes. Theoretically, maximum temperature would result in increased cyanoHAB growth and formation (Kanoshina et al., 2003). The lagged cross-correlation of maximum SSA values and maximum temperatures did not show strong positive relationship (Figure 16). Several lakes show positive cross-correlation after a certain daily lag step, such as Lovell Lake, Great East Lake, Wilson Lake, and Wakeby Pond, suggesting that periods of warmer temperatures influence the presence in cyanobacteria.

However, some study lakes oscillate between negative and positive cross-correlations, which could be due to the small sample size of this study. Additionally, maximum air temperature was used at the nearest NOAA weather station, whereas water surface temperature measurements may be a more accurate indicator for the cross-correlation analysis.

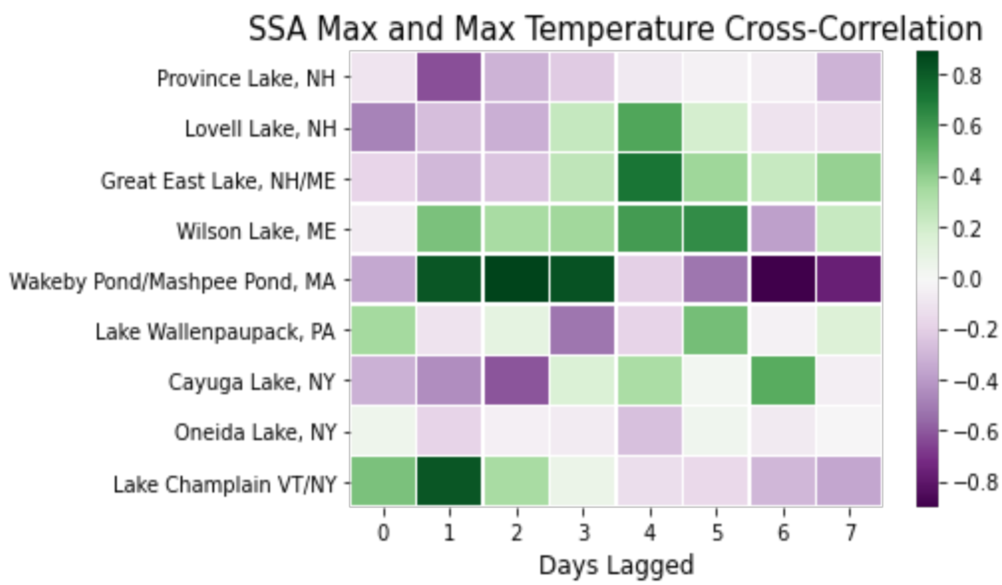


Figure 16. Lagged Cross-Correlation of SSA Maximums and Maximum Temperature

The lagged cross-correlation of precipitation and minimum SSA values and precipitation and maximum SSA values was calculated for the bloom period of each lake (Figures 17 & 18). The impacts of precipitation on cyanoHABs are less understood than temperature (Rousso et al., 2020). Previous studies have observed mixed immediate and long-term effects of precipitation on cyanoHAB growth and formation in waterbodies. Climate change increases the severity and frequency of heavy rainfall events, and heavy rainfall may decrease cyanobacteria biomass through increased hydraulic flushing (Richardson et al., 2019). Conversely, climate change effects the frequency and intensity of rainfall patterns; however, longer dry periods are expected between rainfall events due to climate change. The increased nutrient loading into waterbodies by rainfall runoff combined with the prolonged dry periods are favorable conditions for accelerating cyanobacteria growth (Reichwaldt and Ghadouani, 2012).

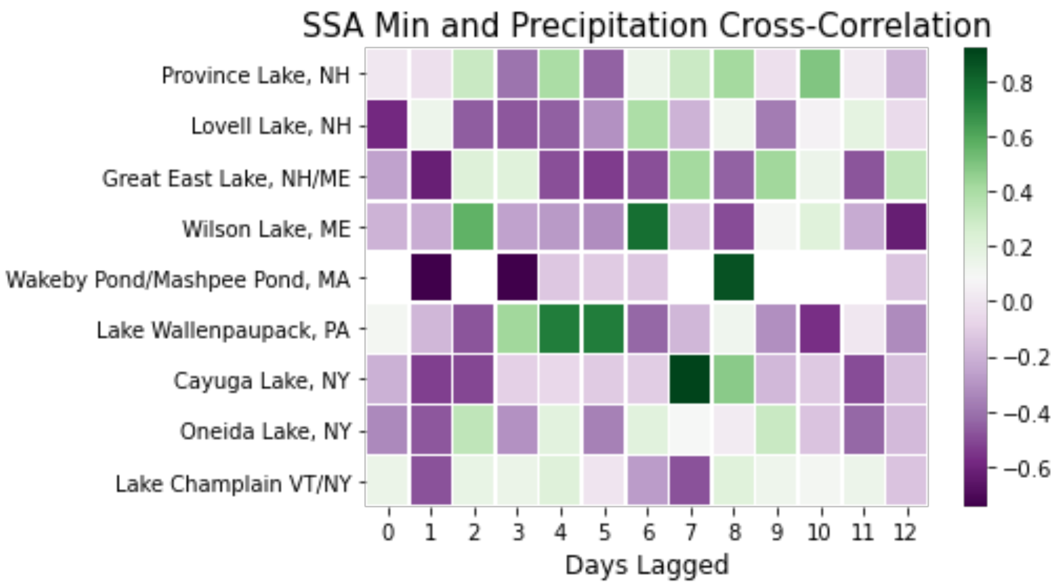


Figure 17. Lagged Cross-Correlation of SSA Minimums and Precipitation

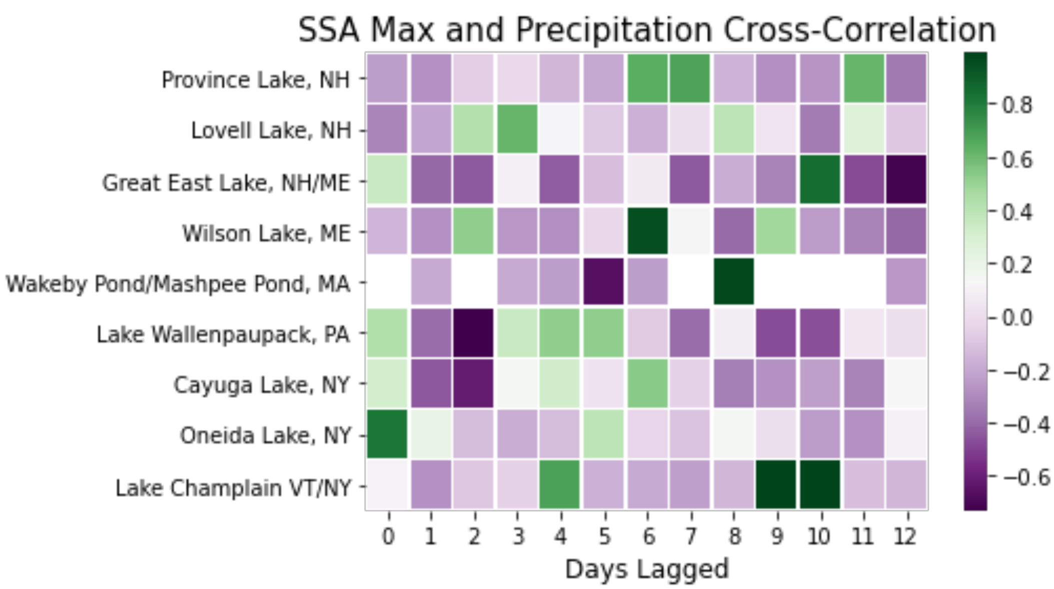


Figure 18. Lagged Cross-Correlation of SSA Maximums and Precipitation

The cross-correlation results for precipitation in this study confirm the mixed effects of precipitation and cyanoHAB growth in waterbodies. The number of lagged days was increased for the precipitation cross-correlation to cover both the immediate and long-term effects of the rainfall events. Both the SSA minimum values and SSA maximum values were used to evaluate the potential effects on cyanoHABs in waterbodies. Precipitation was not found to have strong trends with either SSA maximums or minimums among the study lakes. For both the SSA maximum and minimum with precipitation, the cross-correlation values fluctuate between negative and positive. Few lakes have periods of positive cross-correlation after certain lagged day. Precipitation may have differing effects on cyanoHAB growth and formation depending on the lake, specific rainfall event, and additional weather factors. The cross-correlation of the additional meteorological variables and SSA variables can be found in Appendix C.

4. Discussion

The results of this study demonstrate the ability to use citizen scientist data to select and guide the remote sensing detection of cyanobacteria throughout inland waterbodies. Further, the use of remote sensing was vital to visualizing the magnitude and duration of cyanoHABs in waterbodies. Combined, remote sensing and citizen scientist data were determined to accurately detect and model cyanoHAB events in study lakes. The calculated SSA values over the selected summer/fall bloom period indicated cyanobacteria presence corresponding to citizen scientist data in each lake. This study relied solely on citizen scientist observations for lake selection and to guide remote sensing methods, unlike similar studies that used in situ sampling to validate cyanobacteria remote sensing detection results.

Citizen scientist data has the potential for human error when recording cyanoHAB observations, such as noting the correct observation time, marking the exact location coordinates of the cyanoHAB, bloom characteristics, or identifying the correct cyanobacteria species, that could result in inaccurate cyanobacteria detection. Another consideration that could potentially cause remotes sensing detection inaccuracy is the temporal resolution of the Sentinel-3 OLCI data, causing gaps in available Senteinl-3 OCLI data on dates of reported cyanoHABs. Despite the potential for reported observation errors and Sentinel-3 OLCI data availability, this study carefully selected the study lakes and was highly successful in detecting cyanoHABs that corresponded to citizen scientists' reports. As both the number of users and citizen scientists' observations increase, remote sensing methods utilizing citizen scientist data could become more comprehensive and widespread.

The SSA was used to successfully detect cyanobacteria presence in the study lakes. The Lunetta et al. (2015) SSA method better detected the cyanoHAB events corresponding to the citizen scientist observations than the Wynne et al. (2008) SSA method. The Wynne et al. (2008) SSA method displayed negative SSA pixels, indicating cyanobacteria, that trended with reported cyanoHAB events; however, the Wynne et al. (2008) SSA method flagged numerous invalid pixels that did not correspond with cyanoHAB events. This made it difficult to distinguish between cyanobacteria pixels and clear water pixels throughout the study area lake bloom areas and the winter/springtime non-bloom period comparisons for the Wynne et al. (2008) SSA method. Additionally, this method calculated false negative SSA values for the water-coastline pixels of water bodies in this study, caused by the false identification of the coastline vegetation. Because the Lunetta et al. (2015) SSA method was more sensitive to phycocyanin and could differentiate cyanobacteria from other algal species and vegetation, the water-coastline pixels did not pose significant issues, unlike the Wynne et al. (2008) SSA Method.

The previous confirmed performance of the Lunetta et al. (2015) SSA method at distinguishing cyanobacteria from other erroneous algal was proven in this study. The Lunetta et al. (2015) SSA method clearly distinguished between cyanobacteria pixels and clear water pixels during the bloom periods and the non-bloom comparison period in the study lakes. The Lunetta et al. (2015) SSA method flagged false positive SSA values over wintertime ice cover. This is due to the strong reflectance properties of snow and ice, which caused highly positive SSA values. The ecological patterns of cyanoHABs in the northeastern US seasonal climate were generally found to experience peak cyanoHAB events during the warmer summer and fall months before dissipating in late falls as the temperature and light availability decrease. Cyanobacteria are not commonly present during the colder winter and spring months in the northeastern US (Coffer et al., 2020). However, few studies report cyanobacteria presence in colder winter conditions. Ma et al., (2015) reported winter *Microcystis* cyanoHABs in Lake Taihu, China that persisted in water temperatures below 10°C. Cyanobacteria was also found surviving under ice cover in lakes at higher latitudes, including a study in Lake Nero, Russia that discovered the presence of certain cyanobacteria species under thicker wintertime ice cover with nearly complete lack of oxygen (Babanazarova et al., 2013).

CyanoHAB patterns were similar across the study lakes with cyanoHAB peak events occurring July through September depending on the specific lake, which matches known historical cyanoHAB patterns in this region (Coffer et al., 2020). The magnitude and duration of the cyanoHAB events varied between lakes. The lake size did not significantly influence the magnitude or duration of the cyanoHAB events in the study lakes. Other factors, specific for each lake, most likely controlled the cyanoHAB magnitude and duration, such as, nutrient concentrations, water temperature, lake mixing. Certain study lakes in the same area of the northeast region experienced similar timelines on when blooms began and the duration of the cyanoHAB event. Wilson Lake and Lovell Lake, both located in close proximity in New Hampshire/Maine and have comparable lake sizes, experienced cyanoHABs in the summer of 2018. The cyanoHAB events both peaked mid-August 2018 and lasted until early September 2018. The severity of the Wilson Lake cyanoHAB was significant, covering the entire the water surface, whereas the Lovell Lake cyanoHAB was less severe covering only a portion of the water surface. Lovell Lake has a slightly larger surface area than Wilson Lake, but both lakes have similar water depths.

Cayuga Lake and Oneida Lake both experienced cyanoHAB events during the 2019 summer and fall. Cayuga Lake and Oneida Lake had larger surface areas compared to other study lakes at 66.4 and 79.8 square miles, respectively. The lakes are both located in close proximity in upstate New York. The southern cyanoHAB event in Cayuga Lake peaked at different location over mid-August to early-September, with a large peak occurring August 22-24, 2019. Because of the long shape of Cayuga Lake, the lake experienced several different isolated cyanoHABs across the lake area in this timeframe. Conversely, Oneida Lake had one main cyanoHAB event across much of the lake that experienced several distinct peaks during the lakes longer bloom period of early-July to early-September. The cyanoHAB in Oneida Lake can also be observed moving across the lake surface area during the bloom period (Figure 3). Oneida Lake experienced a longer bloom period than Cayuga Lake, yet both lakes experienced major peaks in mid-August and both cyanoHAB events dissipated in early-September. Additional factors, aside from lake surface area and location, may control cyanoHAB events in these specific lakes, including water temperature, water depth, water column mixing, and nutrient concentrations.

We also explored environmental precursors to cyanoHABs to determine trends and potentially predict future cyanoHABs. Because this study did not have in situ measurements for study lakes, meteorological data was used for the environmental precursor analysis. Like the citizen scientist data used in this study, publicly accessible daily meteorological data was used for the environmental precursor analysis. The environmental precursor analysis did not find strong trends in the lagged cross-correlation of the meteorological data and cyanobacteria presence. The cross-correlation of maximum temperature and SSA maximum values resulted in the best trends between study lakes. The results of environmental precursor analysis were highly variable, and trends were not well established between the cyanoHAB events and meteorological data in the study lakes. Further, the environmental precursor analysis could not be used to predict future cyanoHAB events in the study lakes. The small sample size of the bloom area limited the environmental precursor analysis. Additional data over longer time periods could potentially show better environmental precursor cross-correlation. Other environmental precursors data, not meteorological data, would be more appropriate for the environmental precursor analysis, such as water surface temperature or nutrient concentrations.

5. Conclusion

Sentinel-3 OLCI data was used to calculate the SSA over the study lakes to determine cyanobacteria presence. This study used the ACOLITE processor for the atmospheric correction of Sentinel-3 OLCI data. ACOLITE has been widely used for the atmospheric correction of Landsat and Sentinel-2 data; however, few studies have used the ACOLITE DSF approach for Sentinel-3 OLCI atmospheric correction. Vanhellemont and Ruddick (2021) validated the performance of the ACOLITE DSF approach for Sentinel-3 OLCI data, and this study demonstrated the proven performance of this method for Sentinel-3 OLCI atmospheric correction.

Cyanobacteria was detected throughout nine lakes in the northeastern US, which were selected with citizen scientist reported cyanobacteria observations. This study was solely relied on citizen scientist data to guide the cyanobacteria remote sensing methods. Although citizen scientist data can show these cyanoHABs, real-time remote sensing data is critical to show the full extent of cyanoHAB events. With lack of in situ sampling efforts for lakes in this study, citizen scientist data proved to be an effective tool to guide remote sensing methods. Future work should be done to expand the citizen scientist and remote sensing cyanobacteria detection framework, detailed in this study, to include additional waterbodies and cyanoHAB events. With increasing citizen scientist cyanoHAB app users and increasing cyanoHAB reports, this framework could be applied in widespread waterbodies. The remote sensing detection of cyanobacteria and the use of citizen scientist data for guidance framework has the potential for widespread use. This approach provides a solution for cyanobacteria detection in waterbodies that lack in situ sampling efforts. This framework would be highest interest to water manager, public health officials, and other stakeholders to provide valuable information regarding cyanoHABs in widespread waterbodies. Water managers and public health stakeholders must know the limitations of remote sensing methods for effective use.

There are limitations associated with the remote sensing detection of cyanobacteria and the use of citizen scientist data. Remote sensing methods can only be used for waterbodies with sufficient size and shape due to the 300 m spatial resolution of the Sentinel-3 OLCI data. Pixels along the water-coastline interface were used for this study and have the potential to produce false SSA values. Cyanobacteria is commonly found along waterbody edges, and future work should be done to address the water-coastline pixels. Remote sensing cyanobacteria detection methods also have wintertime snow and ice cover limitations, causing gaps in data availability in waterbodies that experience wintertime snow and ice cover. Lastly, well mixed lakes may have cyanobacteria concentrations further down the water column, yet remote sensing measurements can only detect surface cyanobacteria biomass. The citizen scientist observations had the potential for human error that could cause inaccuracies in the reported cyanoHAB events, as discussed in Section 4.

This study did not find strong trends with in the meteorological environmental precursor analysis and future cyanoHABs events could not be predicted. The main limitation of the environmental precursor analysis was the use of meteorological data due to the lack of specific in situ water quality measurements. Water quality parameters, such as nutrient concentrations and water surface temperature, were shown is various studies to be the environmental precursors that most corresponded to cyanoHAB growth and formation. The environmental precursor may be more case dependent between different waterbodies on what controls cyanoHAB formation. Future work could be done using additional in situ water quality measurements for environmental precursor analysis.

References:

Babanazarova, O., Sidelev, S., & Schischeleva, S. (2013). The structure of winter phytoplankton in Lake Nero, Russia, a hypertrophic lake dominated by Planktothrix-like Cyanobacteria. *Aquatic Biosystems*, *9*(1), 18. https://doi.org/10.1186/2046-9063-9-18

Backer, L. C., Manassaram-Baptiste, D., LePrell, R., & Bolton, B. (2015). Cyanobacteria and Algae Blooms: Review of Health and Environmental Data from the Harmful Algal Bloom-Related Illness Surveillance System (HABISS) 2007–2011. *Toxins*, *7*(4), 1048–1064. https://doi.org/10.3390/toxins7041048

Beaver, J. R., Tausz, C. E., Scotese, K. C., Pollard, A. I., & Mitchell, R. M. (2018). Environmental factors influencing the quantitative distribution of microcystin and common potentially toxigenic cyanobacteria in U.S. lakes and reservoirs. *Harmful Algae*, *78*, 118–128. https://doi.org/10.1016/j.hal.2018.08.004

Bockwoldt, K. A., Nodine, E. R., Mihuc, T. B., Shambaugh, A. D., & Stockwell, J. D. (2017). Reduced Phytoplankton and Zooplankton Diversity Associated with Increased Cyanobacteria in Lake Champlain, USA. *Journal of Contemporary Water Research & Education*, *160*(1), 100–118. https://doi.org/10.1111/j.1936-704X.2017.03243.x

Bukowska, A., Kaliński, T., Koper, M., Kostrzewska-Szlakowska, I., Kwiatowski, J., Mazur-Marzec, H., & Jasser, I. (2017). Predicting blooms of toxic cyanobacteria in eutrophic lakes with diverse cyanobacterial communities. *Scientific Reports*, *7*(1), 8342. https://doi.org/10.1038/s41598-017-08701-8

Burlington, VT Parks, Recreation and Waterfront. (2021). *Water Quality Archives*. Burlington Parks, Recreation & Waterfront. https://enjoyburlington.com/water-quality-archives/

Carmichael, W. W. (1992). Cyanobacteria secondary metabolites—The cyanotoxins. *Journal of Applied Bacteriology*, *72*(6), 445–459. https://doi.org/10.1111/j.1365-2672.1992.tb01858.x

Carvalho, L., Miller, C. A., Scott, E. M., Codd, G. A., Davies, P. S., & Tyler, A. N. (2011). Cyanobacterial blooms: Statistical models describing risk factors for national-scale lake assessment and lake management. *Science of The Total Environment*, *409*(24), 5353–5358. https://doi.org/10.1016/j.scitotenv.2011.09.030

Clark, J. M., Schaeffer, B. A., Darling, J. A., Urquhart, E. A., Johnston, J. M., Ignatius, A. R., Myer, M. H., Loftin, K. A., Werdell, P. J., & Stumpf, R. P. (2017). Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. *Ecological Indicators*, *80*, 84–95. https://doi.org/10.1016/j.ecolind.2017.04.046

Coffer, M. M., Schaeffer, B. A., Darling, J. A., Urquhart, E. A., & Salls, W. B. (2020). Quantifying national and regional cyanobacterial occurrence in US lakes using satellite remote sensing. *Ecological Indicators*, *111*, 105976. https://doi.org/10.1016/j.ecolind.2019.105976

Cremona, F., Tuvikene, L., Haberman, J., Nõges, P., & Nõges, T. (2018). Factors controlling the three-decade long rise in cyanobacteria biomass in a eutrophic shallow lake. *Science of The Total Environment*, *621*, 352–359. https://doi.org/10.1016/j.scitotenv.2017.11.250

Duan, Z., Tan, X., Parajuli, K., Upadhyay, S., Zhang, D., Shu, X., & Liu, Q. (2018). Colony formation in two Microcystis morphotypes: Effects of temperature and nutrient availability. *Harmful Algae*, *72*, 14–24. https://doi.org/10.1016/j.hal.2017.12.006

El-Shehawy, R., Gorokhova, E., Fernández-Piñas, F., & del Campo, F. F. (2012). Global warming and hepatotoxin production by cyanobacteria: What can we learn from experiments? Water Research, 46(5), 1420–1429. https://doi.org/10.1016/j.watres.2011.11.021

Glibert, P.M., Anderson, D.A., Gentien, P., Grane´li, E., Sellner, K.G. (2005). The global, complex phenomena of harmful algal blooms. Oceanography 18, 136–147.

Gilerson, A. A., Gitelson, A. A., Zhou, J., Gurlin, D., Moses, W., Ioannou, I., & Ahmed, S. A. (2010). Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Optics Express*, *18*(23), 24109–24125. https://doi.org/10.1364/OE.18.024109

Heisler, J.P., Gilbert, J., Burkholder, J., Anderson, D., Cochlan, W., Dennison, W., Dortch, Q., Gobler, C.J., Heil, C., Humphries, E., Lewitus, A., Magnien, R., Marshall, H., Sellner, K., Stockwell, D., Stoecker, D., Suddleson, M. (2008). Eutrophication and harmful algal blooms: a scientific consensus. Harmful Algae 8, 3–13.

Howarth, R. W., Marino, R., Lane, J., & Cole, J. J. (1988). Nitrogen fixation in freshwater, estuarine, and marine ecosystems. 1. Rates and importance1: Nitrogen fixation: 1. *Limnology and Oceanography*, *33*(4part2), 669–687. https://doi.org/10.4319/lo.1988.33.4part2.0669

Hu, C. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, *113*(10), 2118–2129. https://doi.org/10.1016/j.rse.2009.05.012

Huisman, J., Codd, G.A., Paerl, H.W. et al. Cyanobacterial blooms. Nat Rev Microbiol 16, 471–483 (2018). https://doi.org/10.1038/s41579-018-0040-1

Jia, T., Zhang, X., & Dong, R. (2019). Long-Term Spatial and Temporal Monitoring of Cyanobacteria Blooms Using MODIS on Google Earth Engine: A Case Study in Taihu Lake. Google Earth Engine and Cloud Computing Platforms: Methods and Applications in Big Geo Data Science. doi: https://doi.org/10.3390/rs11192269

Kahru, M., & Elmgren, R. (2014). Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. *Biogeosciences*, *11*(13), 3619–3633. https://doi.org/10.5194/bg-11-3619-2014

Kanoshina, I., Lips, U., & Leppänen, J.-M. (2003). The influence of weather conditions (temperature and wind) on cyanobacterial bloom development in the Gulf of Finland (Baltic Sea). *Harmful Algae*, *2*(1), 29–41. https://doi.org/10.1016/S1568-9883(02)00085-9

Keith, D., Rover, J., Green, J., Zalewsky, B., Charpentier, M., Thursby, G., & Bishop, J. (2018). Monitoring Algal Blooms in drinking water reservoirs using the Landsat 8 Operational Land Imager. *International Journal of Remote Sensing*, *39*(9), 2818–2846. https://doi.org/10.1080/01431161.2018.1430912

Kudela, R. M., Palacios, S. L., Austerberry, D. C., Accorsi, E. K., Guild, L. S., & Torres-Perez, J. (2015). Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. *Remote Sensing of Environment*, *167*, 196–205. https://doi.org/10.1016/j.rse.2015.01.025

LACAWAC. (2020). *LAKE WALLENPAUPACK COMMUNITY-LED MONITORING PROGRAM 2020*. Community Science. https://www.lacawac.org/citizen-science.html.

Le, C., Hu, C., English, D., Cannizzaro, J., & Kovach, C. (2013). Climate-driven chlorophyll-a changes in a turbid estuary: Observations from satellites and implications for management. *Remote Sensing of Environment*, *130*, 11–24. https://doi.org/10.1016/j.rse.2012.11.011

Loftin, K. A., Graham, J. L., Hilborn, E. D., Lehmann, S. C., Meyer, M. T., Dietze, J. E., & Griffith, C. B. (2016). Cyanotoxins in inland lakes of the United States: Occurrence and potential recreational health risks in the EPA National Lakes Assessment 2007. *Harmful Algae*, *56*, 77–90. https://doi.org/10.1016/j.hal.2016.04.001

Lunetta, R. S., Schaeffer, B. A., Stumpf, R. P., Keith, D., Jacobs, S. A., & Murphy, M. S. (2015). Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. *Remote Sensing of Environment*, *157*, 24–34. https://doi.org/10.1016/j.rse.2014.06.008

Ma, J., Qin, B., Paerl, H. W., Brookes, J. D., Hall, N. S., Shi, K., Zhou, Y., Guo, J., Li, Z., Xu, H., Wu, T., & Long, S. (2016). The persistence of cyanobacterial (Microcystis spp.) blooms throughout winter in Lake Taihu, China. *Limnology and Oceanography*, *61*(2), 711–722. https://doi.org/10.1002/lno.10246

Matthews, M. W., Bernard, S., & Robertson, L. (2012). An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment*, *124*, 637–652. https://doi.org/10.1016/j.rse.2012.05.032

Merel, S., Walker, D., Chicana, R., Snyder, S., Baurès, E., & Thomas, O. (2013). State of knowledge and concerns on cyanobacterial blooms and cyanotoxins. *Environment International*, *59*, 303–327. https://doi.org/10.1016/j.envint.2013.06.013

Mishra, S., Mishra, D. R., & Schluchter, W. M. (2009). A Novel Algorithm for Predicting Phycocyanin Concentrations in Cyanobacteria: A Proximal Hyperspectral Remote Sensing Approach. *Remote Sensing*, *1*(4), 758–775. https://doi.org/10.3390/rs1040758

Mishra, S. (2014). A novel remote sensing algorithm to quantify phycocyanin in cyanobacterial algal blooms. Environ. Res. Lett., 10.

Mishra, S., Stumpf, R. P., Schaeffer, B. A., Werdell, P. J., Loftin, K. A., & Meredith, A. (2019). Measurement of Cyanobacterial Bloom Magnitude using Satellite Remote Sensing. *Scientific Reports*, *9*(1), 18310. https://doi.org/10.1038/s41598-019-54453-y

NH DES. (2019). *State Issues Cyanobacteria Advisory for Province Lake in Effingham and Wakefield, New Hampshire*. NH Department of Environmental Services. https://www.des.nh.gov/news-and-media/state-issues-cyanobacteria-advisory-province-lake-effingham-and-wakefield-new

NH FGD. (2013). PROVINCE LAKE Effingham, Wakefield. map. https://wildlife.state.nh.us/maps/bathymetry/province_effingham.pdf.

NY DEC. (2019). *Harmful Algal Bloom Action Plan Cayuga Lake* (p. 124). https://www.dec.ny.gov/docs/water_pdf/cayugahabplan.pdf

NY DEC. (2019). *Harmful Algal Blooms by County 2012-2019* (HARMFUL ALGAL BLOOMS BY COUNTY 2012 - 2019). https://www.dec.ny.gov/docs/water_pdf/habsextentsummary.pdf

NY DEC. (2021). *Oneida Lake (Madison, Onondaga, Oswego, and Oneida)*. Department of Environmental Conservation. https://www.dec.ny.gov/outdoor/41034.html

NY DOH. (2020). *Department of Health*. Harmful Blue-Green Algae Bloom-Related Beach Closure Summary. https://health.ny.gov/environmental/water/drinking/bluegreenalgae/beachdata.htm

Ogashawara, I. (2019). The Use of Sentinel-3 Imagery to Monitor Cyanobacterial Blooms. *Environments*, *6*(6), 60. https://doi.org/10.3390/environments6060060

O'Neil, J. M., Davis, T. W., Burford, M. A., & Gobler, C. J. (2012). The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. *Harmful Algae*, *14*, 313–334. https://doi.org/10.1016/j.hal.2011.10.027

Oyama, Y., Matsushita, B., & Fukushima, T. (2015). Distinguishing surface cyanobacterial blooms and aquatic macrophytes using Landsat/TM and ETM+ shortwave infrared bands. *Remote Sensing of Environment*, *157*, 35–47. https://doi.org/10.1016/j.rse.2014.04.031

Paerl, H. W., Fulton, R. S., Moisander, P. H., & Dyble, J. (2001). Harmful Freshwater Algal Blooms, With an Emphasis on Cyanobacteria. *The Scientific World JOURNAL*, *1*, 76–113. https://doi.org/10.1100/tsw.2001.16

Pyo, J. C., Ligaray, M., Kwon, Y. S., Ahn, M.-H., Kim, K., Lee, H., Kang, T., Cho, S. B., Park, Y., & Cho, K. H. (2018). *High-Spatial Resolution Monitoring of Phycocyanin and Chlorophyll-a Using Airborne Hyperspectral Imagery*. 31.

Pyo, J., Park, L. J., Pachepsky, Y., Baek, S.-S., Kim, K., & Cho, K. H. (2020). Using convolutional neural network for predicting cyanobacteria concentrations in river water. *Water Research*, *186*, 116349. https://doi.org/10.1016/j.watres.2020.116349

Qin, B., Li, W., Zhu, G., Zhang, Y., Wu, T., & Gao, G. (2015). Cyanobacterial bloom management through integrated monitoring and forecasting in large shallow eutrophic Lake Taihu (China). *Journal of Hazardous Materials*, *287*, 356–363. https://doi.org/10.1016/j.jhazmat.2015.01.047

Reichwaldt, E. S., & Ghadouani, A. (2012). Effects of rainfall patterns on toxic cyanobacterial blooms in a changing climate: Between simplistic scenarios and complex dynamics. *Water Research*, *46*(5), 1372–1393. https://doi.org/10.1016/j.watres.2011.11.052

Renosh, P. R., Doxaran, D., Keukelaere, L. D., & Gossn, J. I. (2020). Evaluation of Atmospheric Correction Algorithms for Sentinel-2-MSI and Sentinel-3-OLCI in Highly Turbid Estuarine Waters. *Remote Sensing*, *12*(8), 1285. https://doi.org/10.3390/rs12081285

Richardson, J., Feuchtmayr, H., Miller, C., Hunter, P. D., Maberly, S. C., & Carvalho, L. (2019). Response of cyanobacteria and phytoplankton abundance to warming, extreme rainfall events and nutrient enrichment. *Global Change Biology*, *25*(10), 3365–3380. https://doi.org/10.1111/gcb.14701

Rousso, B. Z., Bertone, E., Stewart, R., & Hamilton, D. P. (2020). A systematic literature review of forecasting and predictive models for cyanobacteria blooms in freshwater lakes. *Water Research*, *182*, 115959. https://doi.org/10.1016/j.watres.2020.115959

Schaeffer, B. A., Bailey, S. W., Conmy, R. N., Galvin, M., Ignatius, A. R., Johnston, J. M., Keith, D. J., Lunetta, R. S., Parmar, R., Stumpf, R. P., Urquhart, E. A., Werdell, P. J., & Wolfe, K. (2018). Mobile device application for monitoring cyanobacteria harmful algal blooms using Sentinel-3 satellite Ocean and Land Colour Instruments. *Environmental Modelling & Software*, *109*, 93–103. https://doi.org/10.1016/j.envsoft.2018.08.015

Simis, S. G. H., Peters, S. W. M., & Gons, H. J. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography*, *50*(1), 237–245. https://doi.org/10.4319/lo.2005.50.1.0237

Sleeter, B.M., T. Loveland, G. Domke, N. Herold, J. Wickham, and N. Wood, 2018: Land Cover and Land-Use Change. In *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II* [Reidmiller, D.R., C.W. Avery, D.R. Easterling, K.E. Kunkel, K.L.M. Lewis, T.K. Maycock, and B.C. Stewart (eds.)]. U.S. Global Change Research Program, Washington, DC, USA, pp. 202–231. doi: 10.7930/NCA4.2018.CH5

Steffen, M. M., Davis, T. W., McKay, R. M. L., Bullerjahn, G. S., Krausfeldt, L. E., Stough, J. M. A., Neitzey, M. L., Gilbert, N. E., Boyer, G. L., Johengen, T. H., Gossiaux, D. C., Burtner, A. M., Palladino, D., Rowe, M. D., Dick, G. J., Meyer, K. A., Levy, S., Boone, B. E., Stumpf, R. P., … Wilhelm, S. W. (2017). Ecophysiological Examination of the Lake Erie Microcystis Bloom in

2014: Linkages between Biology and the Water Supply Shutdown of Toledo, OH. *Environmental Science & Technology*, *51*(12), 6745–6755. https://doi.org/10.1021/acs.est.7b00856

Stumpf, R. P., Davis, T. W., Wynne, T. T., Graham, J. L., Loftin, K. A., Johengen, T. H., Gossiaux, D., Palladino, D., & Burtner, A. (2016). Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. *Harmful Algae*, *54*, 160–173. https://doi.org/10.1016/j.hal.2016.01.005

Stumpf, R. P., Johnson, L. T., Wynne, T. T., & Baker, D. B. (2016). Forecasting annual cyanobacterial bloom biomass to inform management decisions in Lake Erie. *Journal of Great Lakes Research*, *42*(6), 1174–1183. https://doi.org/10.1016/j.jglr.2016.08.006

Toming, K., Kutser, T., Uiboupin, R., Arikas, A., Vahter, K., & Paavel, B. (2017). Mapping Water Quality Parameters with Sentinel-3 Ocean and Land Colour Instrument imagery in the Baltic Sea. *Remote Sensing*, *9*(10), 1070. https://doi.org/10.3390/rs9101070

Town of Sandwich, MA. (2020, September 29). *Sandwich Health Dept Warning: Cyanobacteria Outbreak in Mashpee-Wakeby Pond*. https://www.sandwichmass.org/CivicAlerts.aspx?AID=334&ARC=737.

Urquhart, E. A., & Schaeffer, B. A. (2020). Envisat MERIS and Sentinel-3 OLCI satellite lake biophysical water quality flag dataset for the contiguous United States. *Data in Brief*, *28*, 104826. https://doi.org/10.1016/j.dib.2019.104826

Urquhart, E. A., Schaeffer, B. A., Stumpf, R. P., Loftin, K. A., & Werdell, P. J. (2017). A method for examining temporal changes in cyanobacterial harmful algal bloom spatial extent using satellite remote sensing. *Harmful Algae*, *67*, 144–152. https://doi.org/10.1016/j.hal.2017.06.001

U.S. Environmental Protection Agency (2015). A Compilation of Cost Data Associated With the Impacts and Control of Nutrient Pollution. EPA 820-F-15- 096. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

US EPA. (2018, June 6). *Learn about Cyanobacteria and Cyanotoxins* [Overviews and Factsheets]. US EPA. https://www.epa.gov/cyanohabs/learn-about-cyanobacteria-and-cyanotoxins

US EPA. (2021, April 22). *bloomWatch*. Cyanobacteria Monitoring Collaborative. https://cyanos.org/bloomwatch/.

US EPA. (2021, April 22). *cyanoScope*. Cyanobacteria Monitoring Collaborative. https://cyanos.org/cyanoscope/.

Vanhellemont, Q., & Ruddick, K. (2018). Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications. *Remote Sensing of Environment*, *216*, 586–597. https://doi.org/10.1016/j.rse.2018.07.015

Vanhellemont, Q. (2019). Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. *Remote Sensing of Environment*, *225*, 175–192. https://doi.org/10.1016/j.rse.2019.03.010

Vanhellemont, Q., & Ruddick, K. (2021). Atmospheric correction of Sentinel-3/OLCI data for mapping of suspended particulate matter and chlorophyll-a concentration in Belgian turbid coastal waters. *Remote Sensing of Environment*, *256*, 112284. https://doi.org/10.1016/j.rse.2021.112284

Watson, S.B., McCauley, E., Downing, J.A. (1997). Patterns in phytoplankton taxonomic composition across temperate lakes of differing nutrient status. Limnol. Oceanogr. 42, 487–495.

WHO. (1999). *Toxic cyanobacteria in water*. Toxic Cyanobacteria in Water; World Health Organization. http://www.who.int/water_sanitation_health/publications/toxicyanobact/en/

Wynne, T., Stumpf, R., Tomlinson, M., Warner, R. A., Tester, P., Dyble, J., & Fahnenstiel, G. L. (2008). Relating spectral shape to cyanobacterial bloom in the Laurentian Great Lakes. *International Journal of Remote Sensing - INT J REMOTE SENS*, *29*, 3665–3672. https://doi.org/10.1080/01431160802007640

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Fahnenstiel, G. L., Dyble, J., Schwab, D. J., & Joshi, S. J. (2013). Evolution of a cyanobacterial bloom forecast system in western Lake Erie: Development and initial evaluation. *Journal of Great Lakes Research*, *39*, 90–99. https://doi.org/10.1016/j.jglr.2012.10.003

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Schwab, D. J., Watabayashi, G. Y., & Christensen, J. D. (2011). Estimating cyanobacterial bloom transport by coupling remotely sensed imagery and a hydrodynamic model. *Ecological Applications*, *21*(7), 2709–2721. https://doi.org/10.1890/10-1454.1

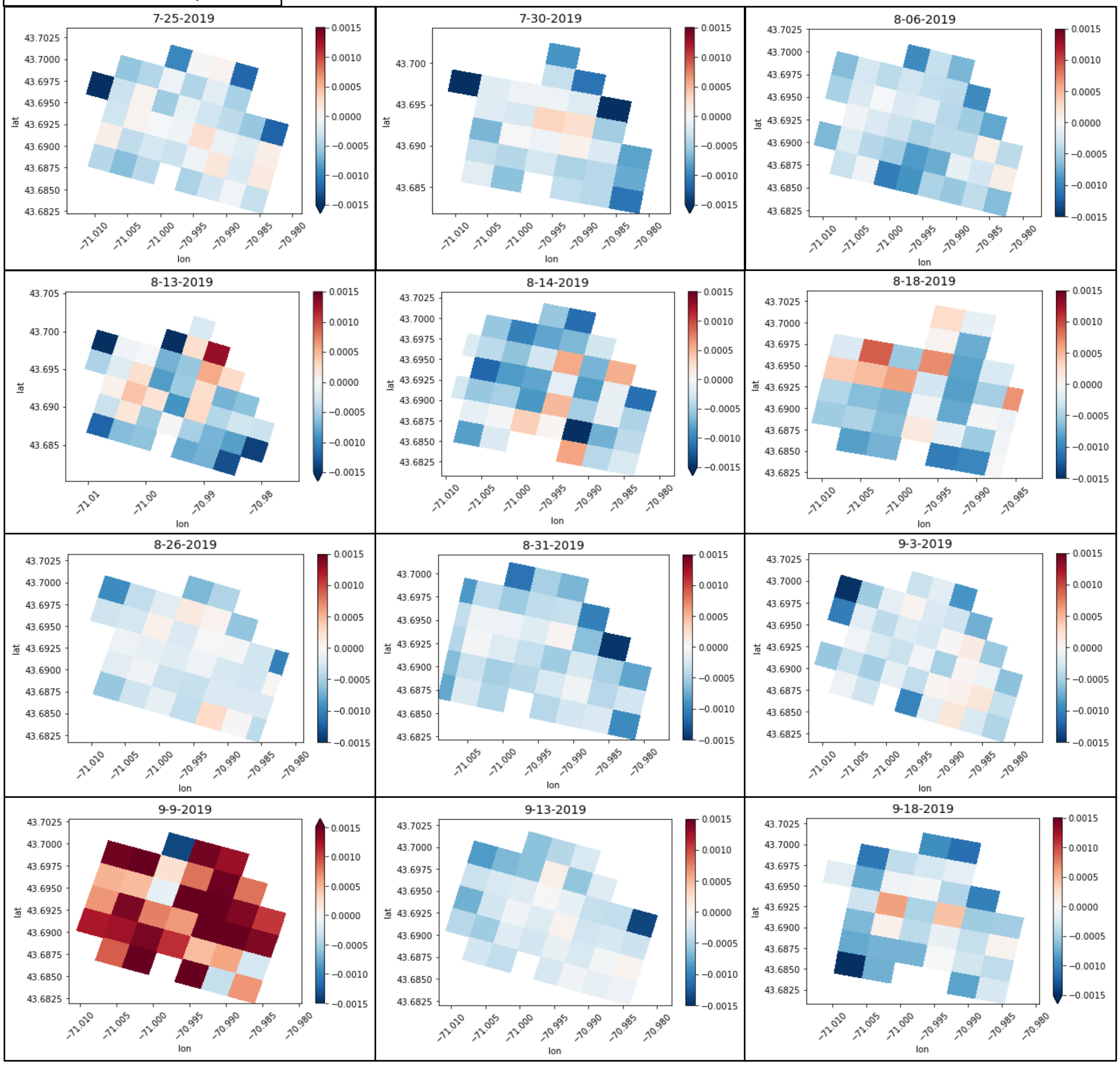
Xiao, X., He, J., Huang, H., Miller, T. R., Christakos, G., Reichwaldt, E. S., Ghadouani, A., Lin, S., Xu, X., & Shi, J. (2017). A novel single-parameter approach for forecasting algal blooms. *Water Research*, *108*, 222–231. https://doi.org/10.1016/j.watres.2016.10.076

Xu, H., Paerl, H. W., Qin, B., Zhu, G., & Gaoa, G. (2010). Nitrogen and phosphorus inputs control phytoplankton growth in eutrophic Lake Taihu, China. *Limnology and Oceanography*, *55*(1), 420–432. https://doi.org/10.4319/lo.2010.55.1.0420

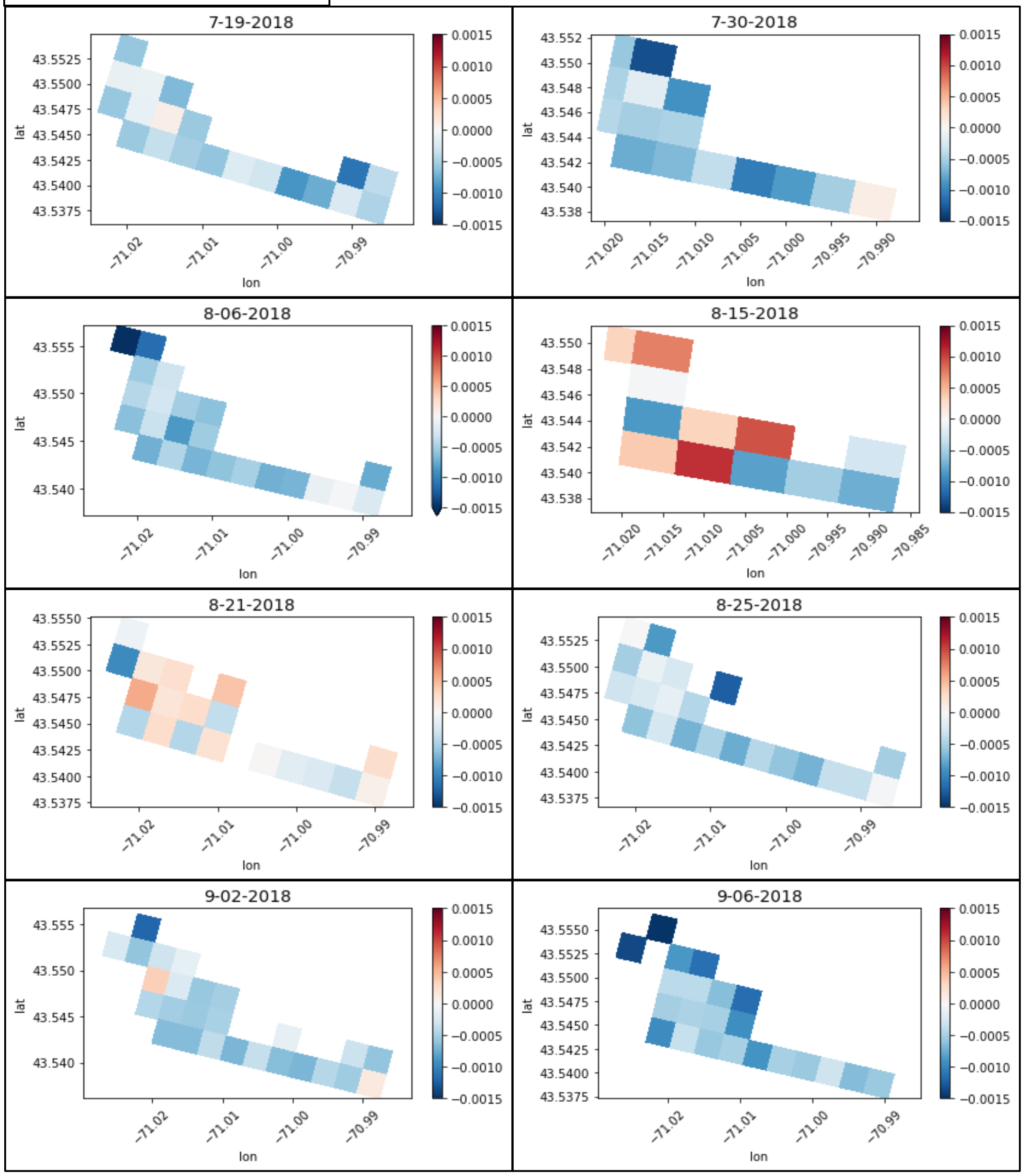
Zhao, C. S., Shao, N. F., Yang, S. T., Ren, H., Ge, Y. R., Feng, P., Dong, B. E., & Zhao, Y. (2019). Predicting cyanobacteria bloom occurrence in lakes and reservoirs before blooms occur. *Science of The Total Environment*, *670*, 837–848. https://doi.org/10.1016/j.scitotenv.2019.03.161

Appendix A – Bloom Period SSA calculated over each Study Lake using the Lunetta et al. (2015) SSA Method

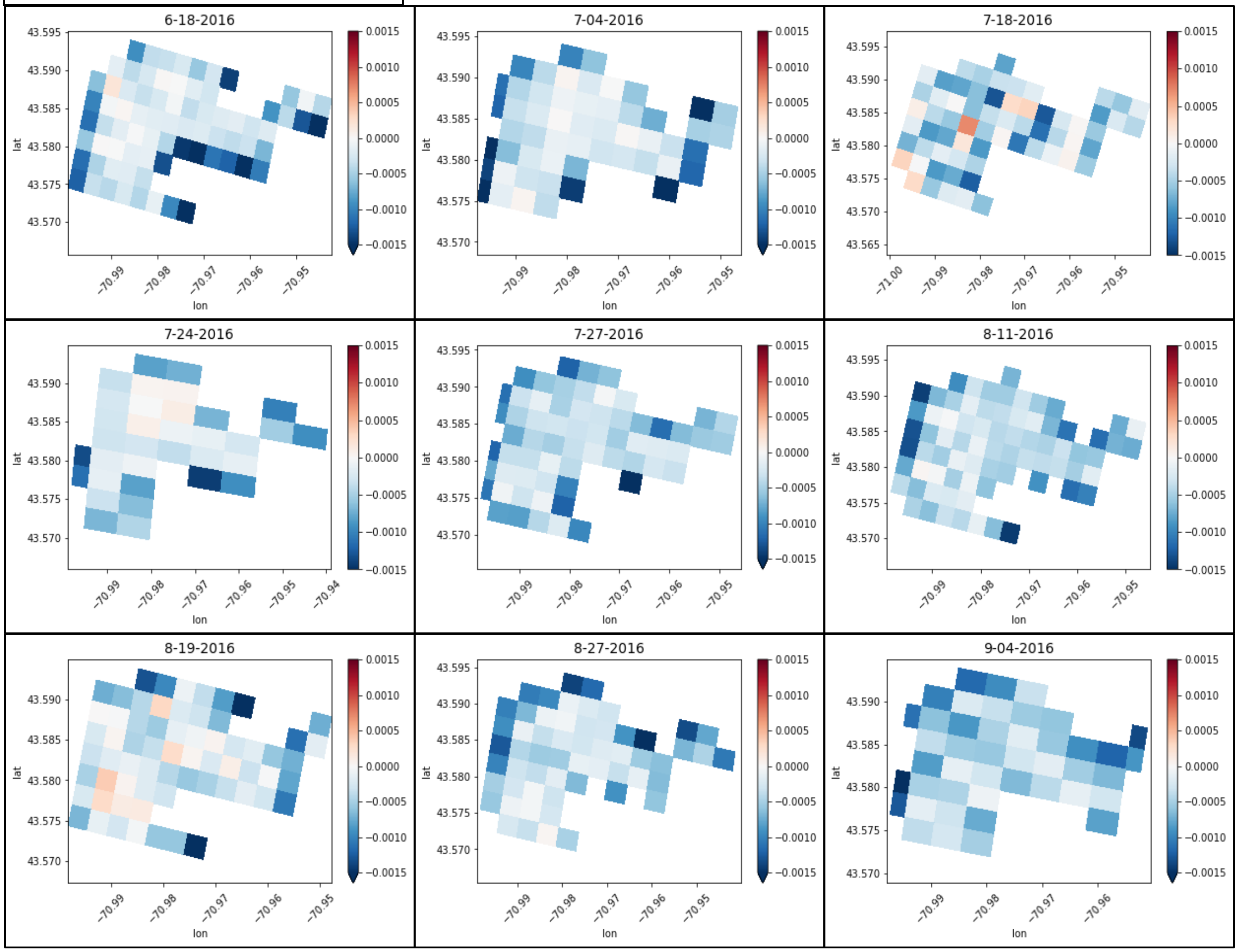
Province Lake, NH



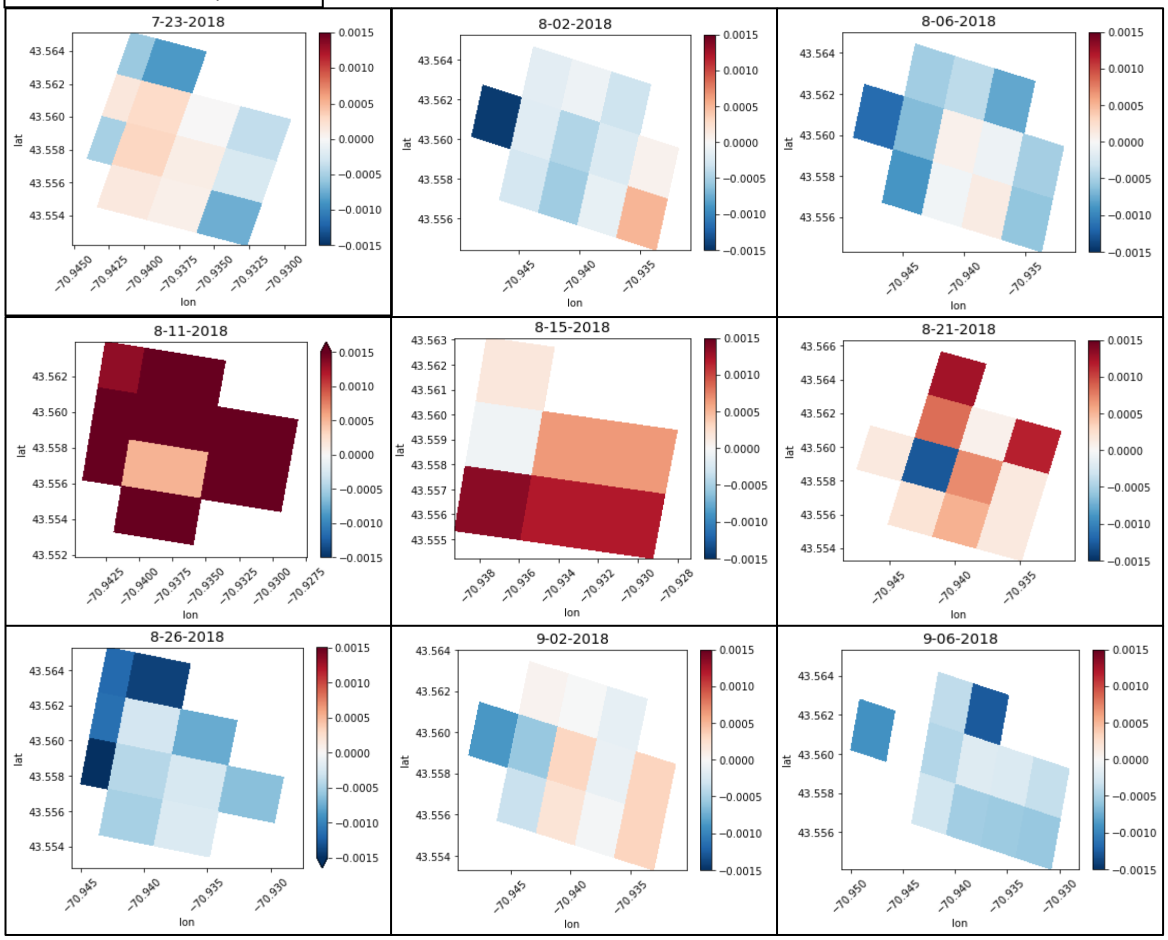
Lovell Lake, NH



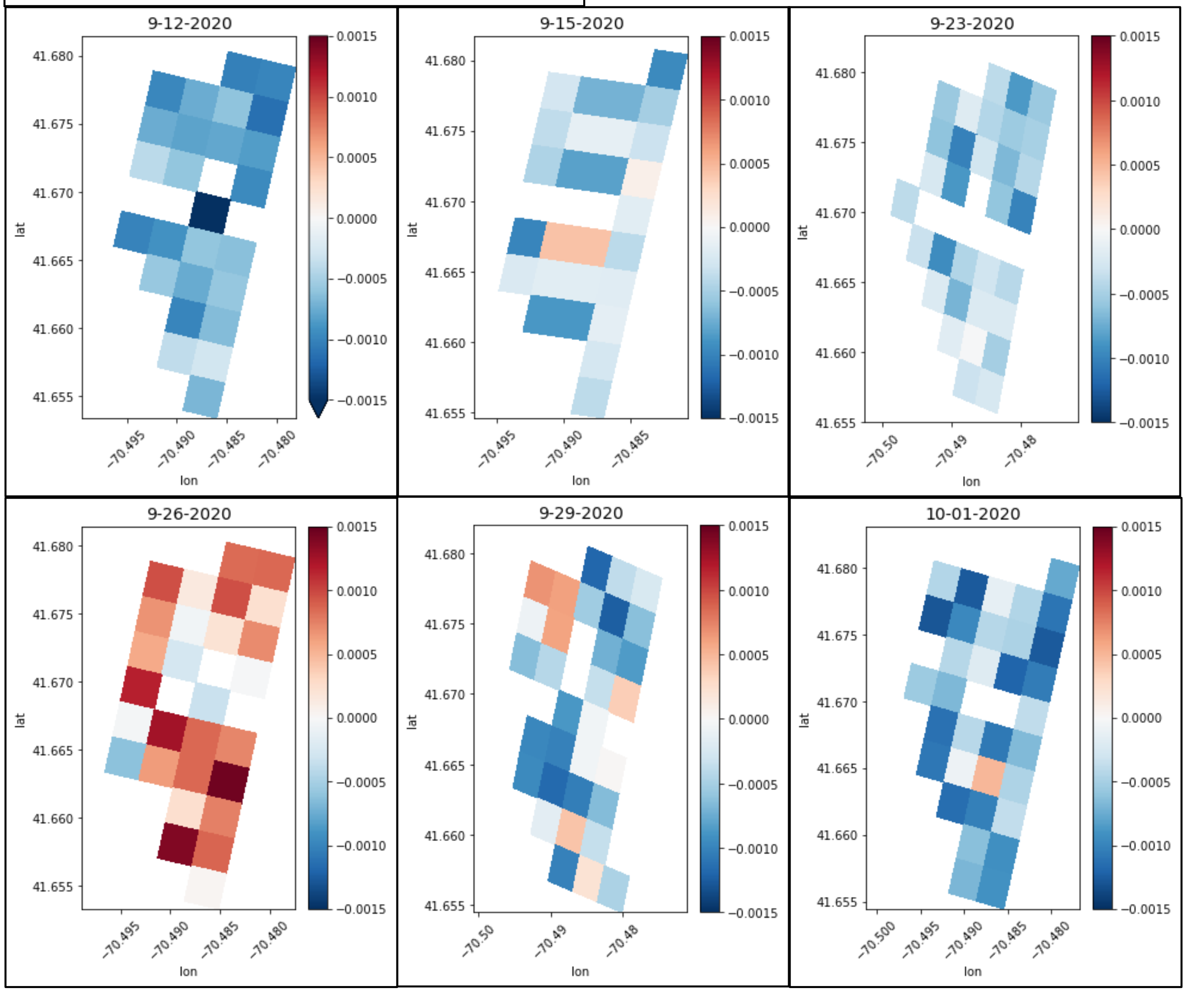
Great East Lake, NH/ME



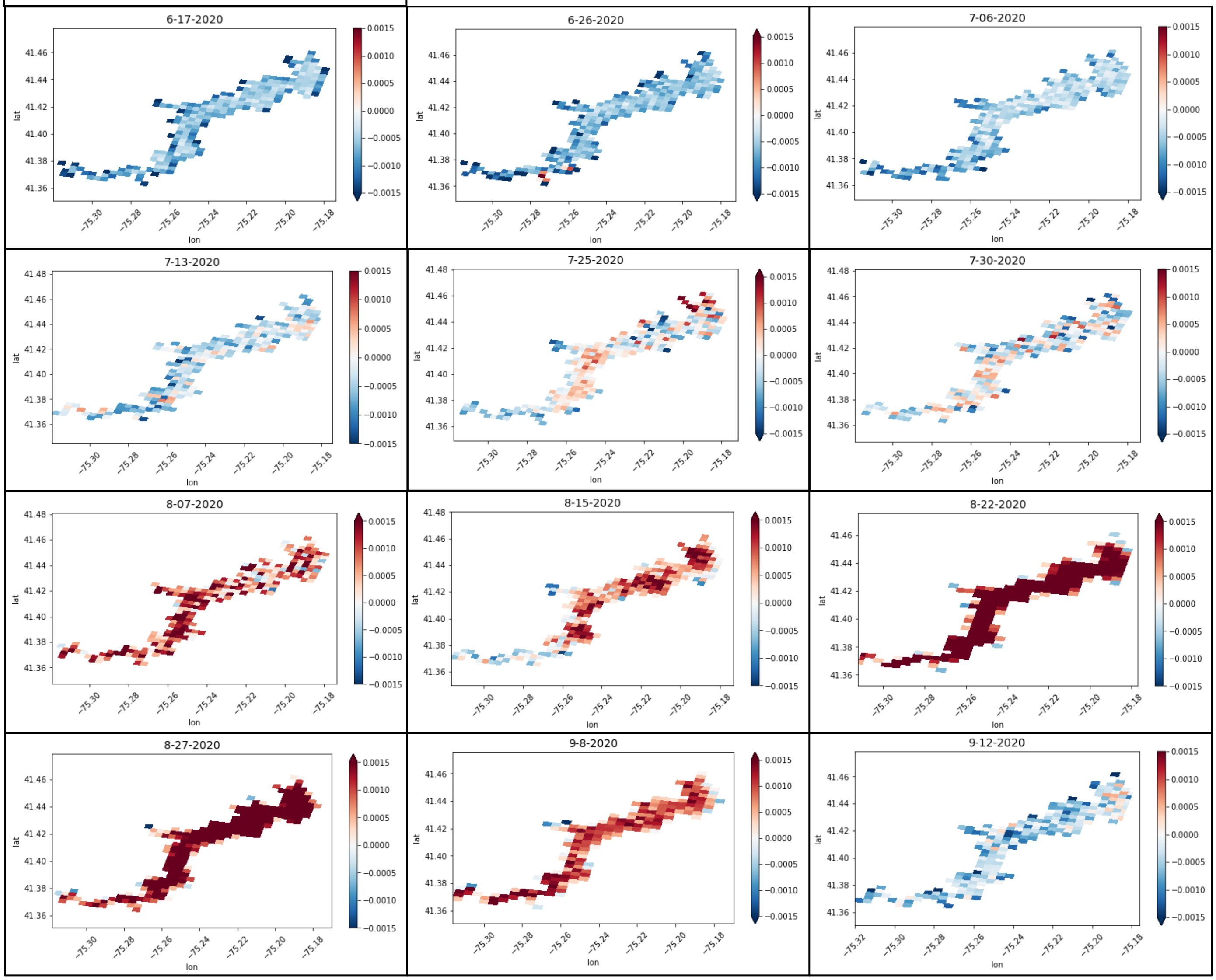
Wilson Lake, ME



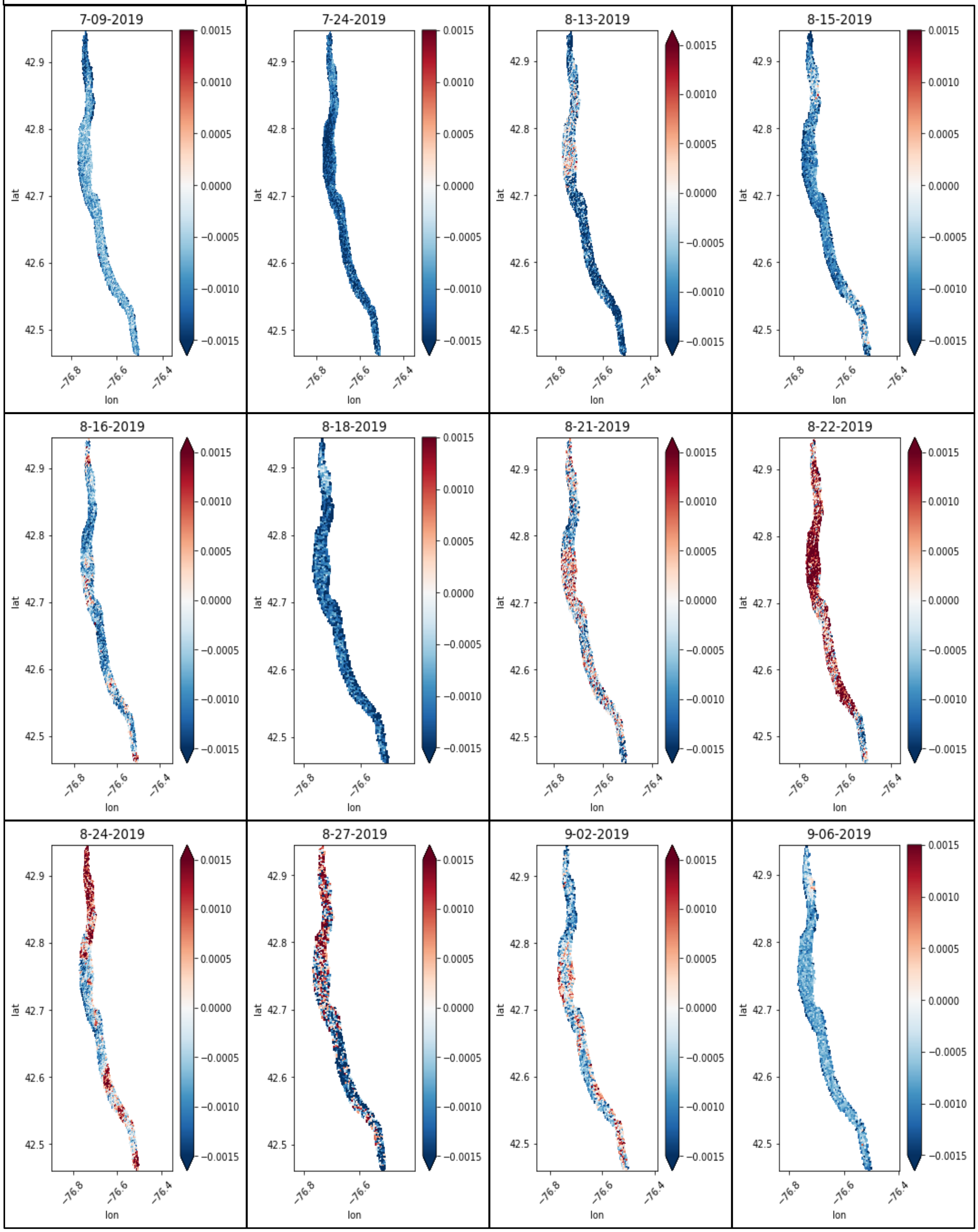
Wakeby Pond/Mashpee Pond, MA



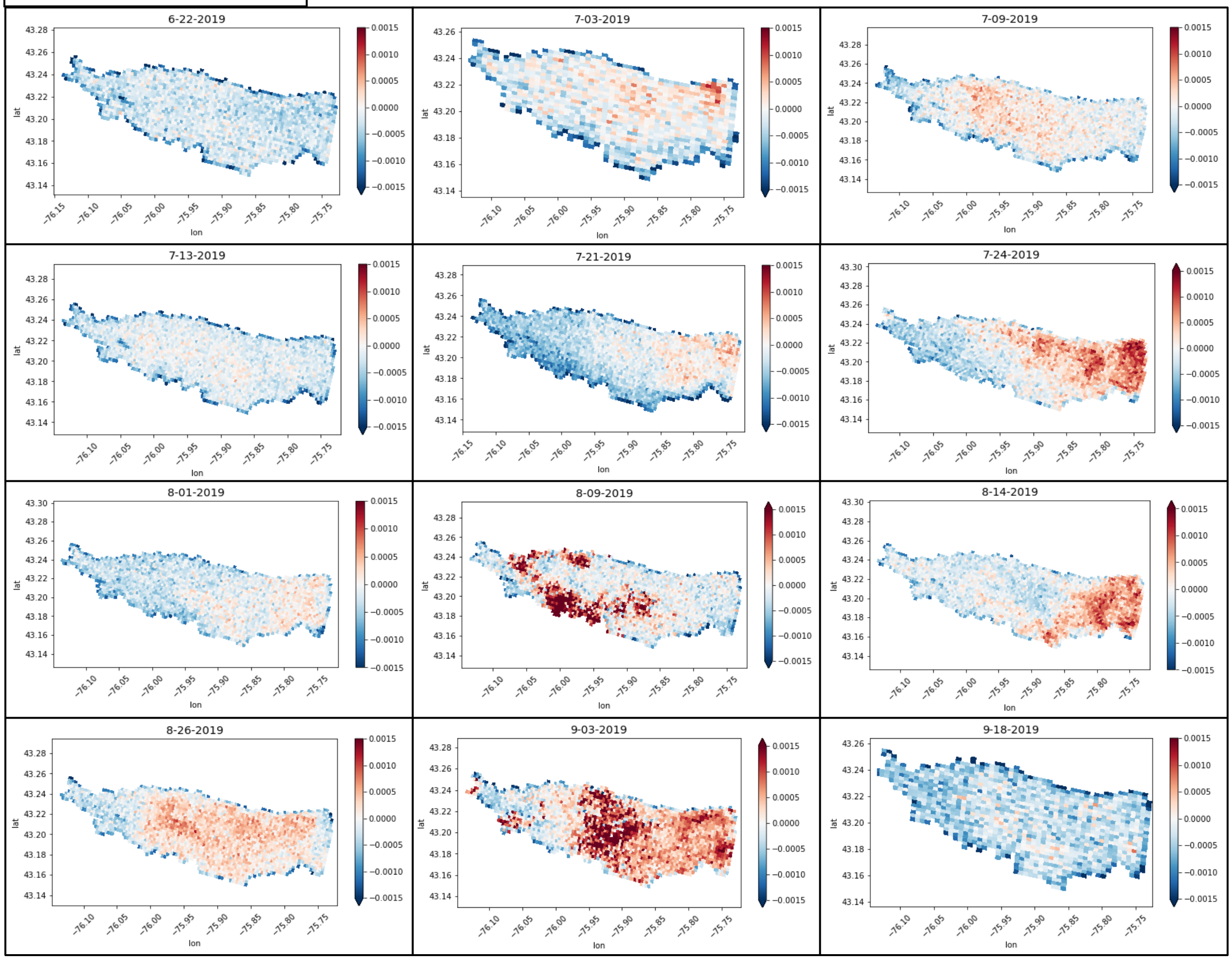
Lake Wallenpaupack, PA



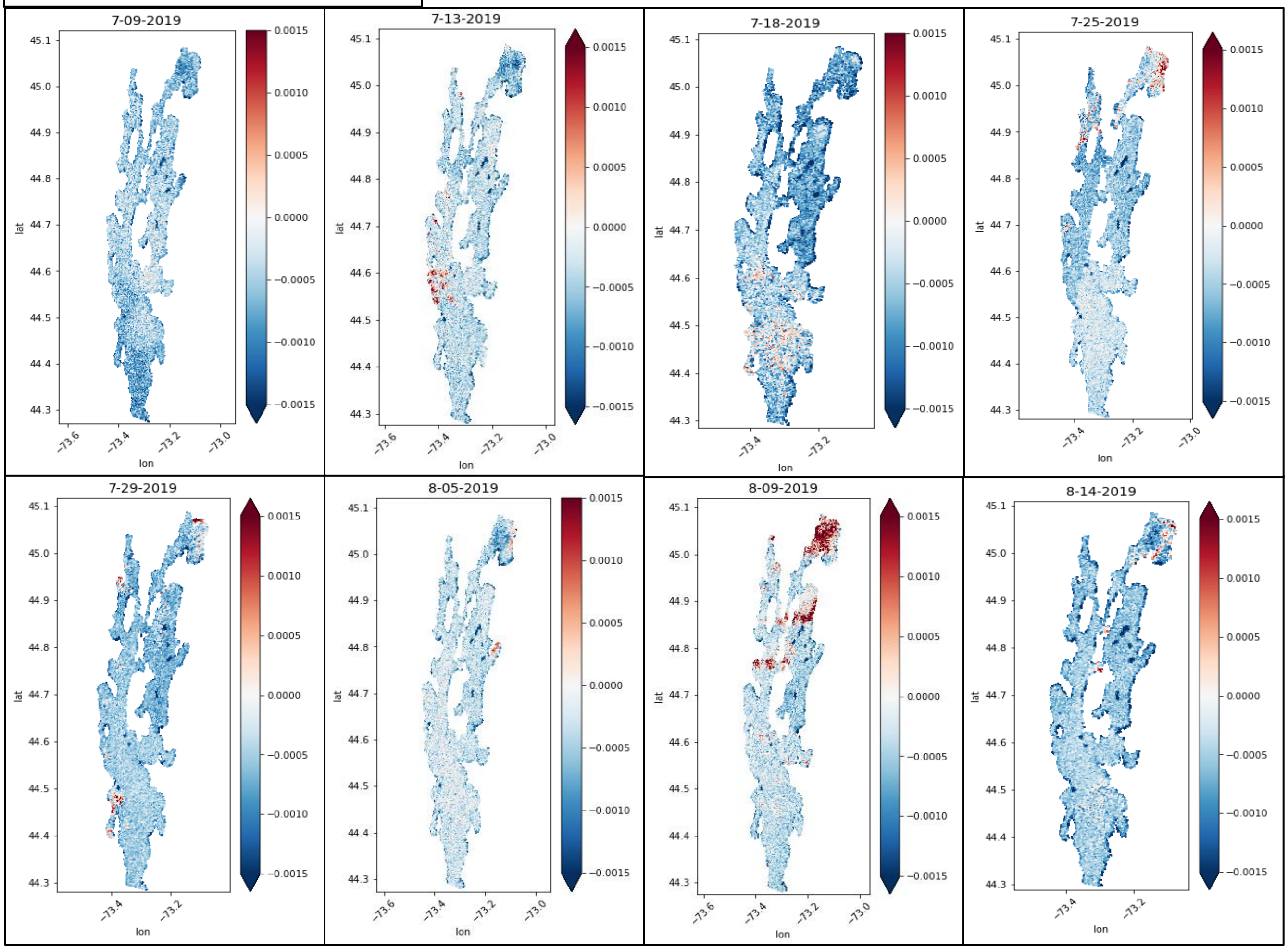
Cayuga Lake, NY



Oneida Lake, NY

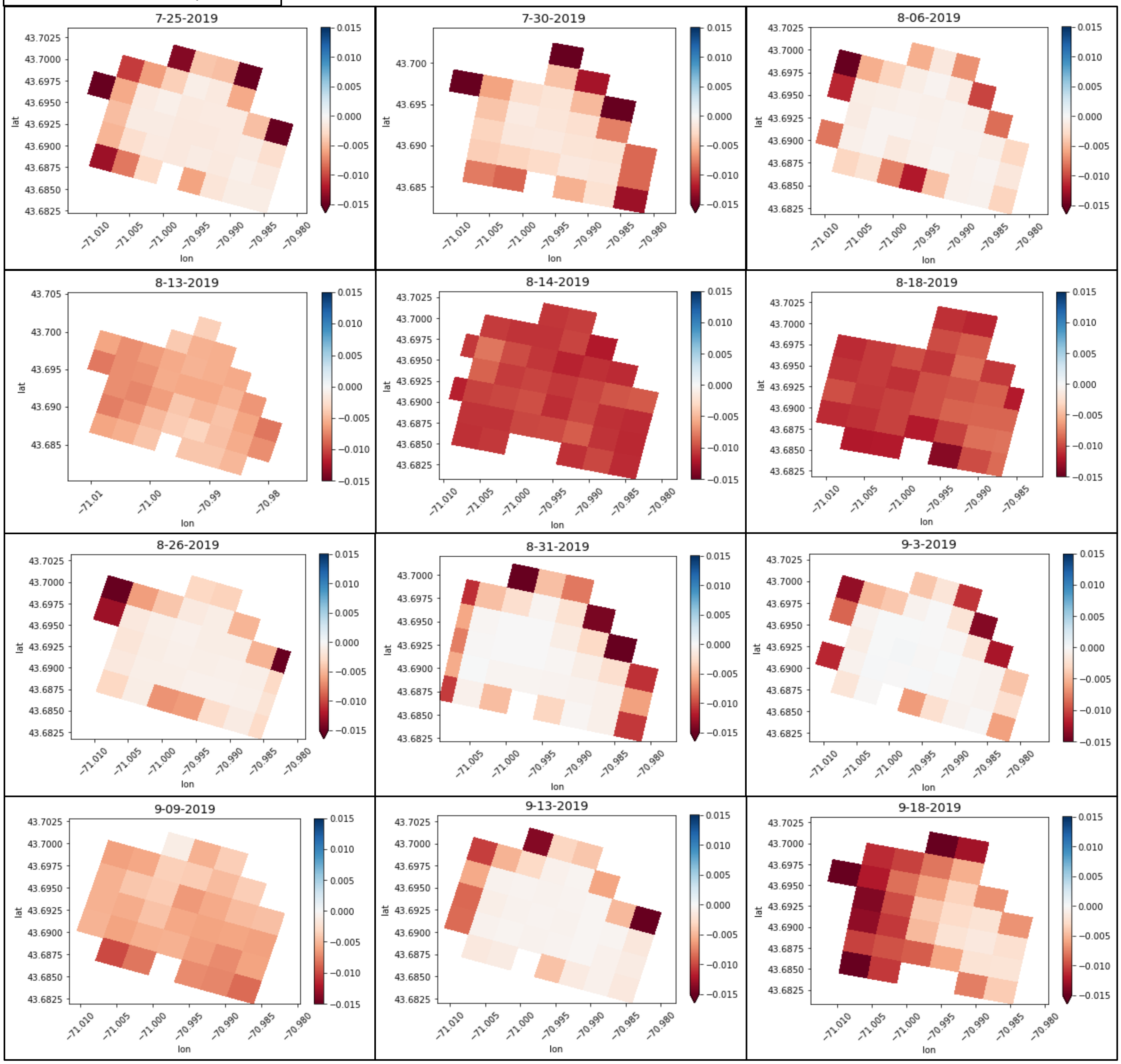


Lake Champlain VT/NY

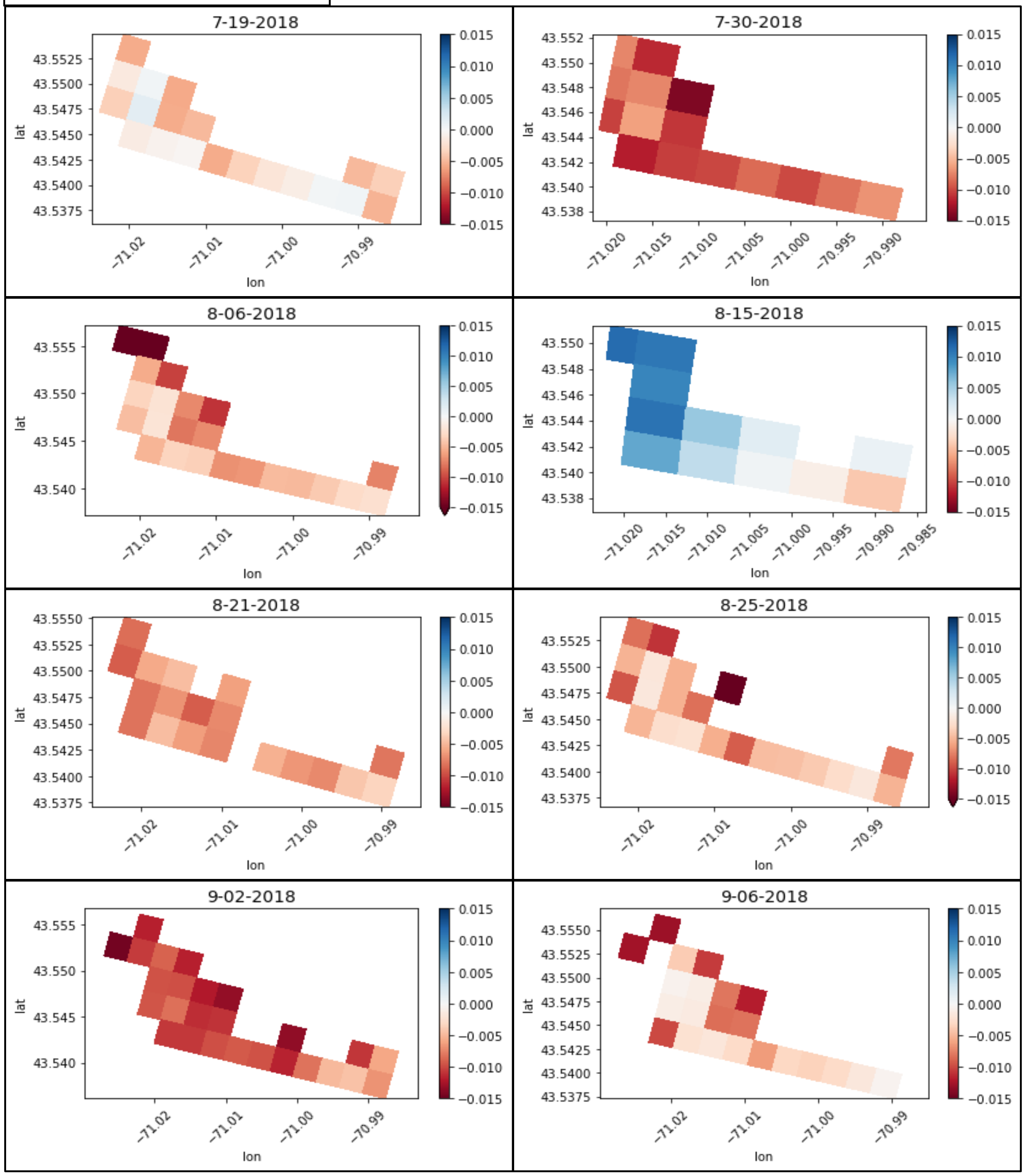


Appendix B – Bloom Period SSA calculated over each Study Lake using the Wynne et al. (2008) SSA Method

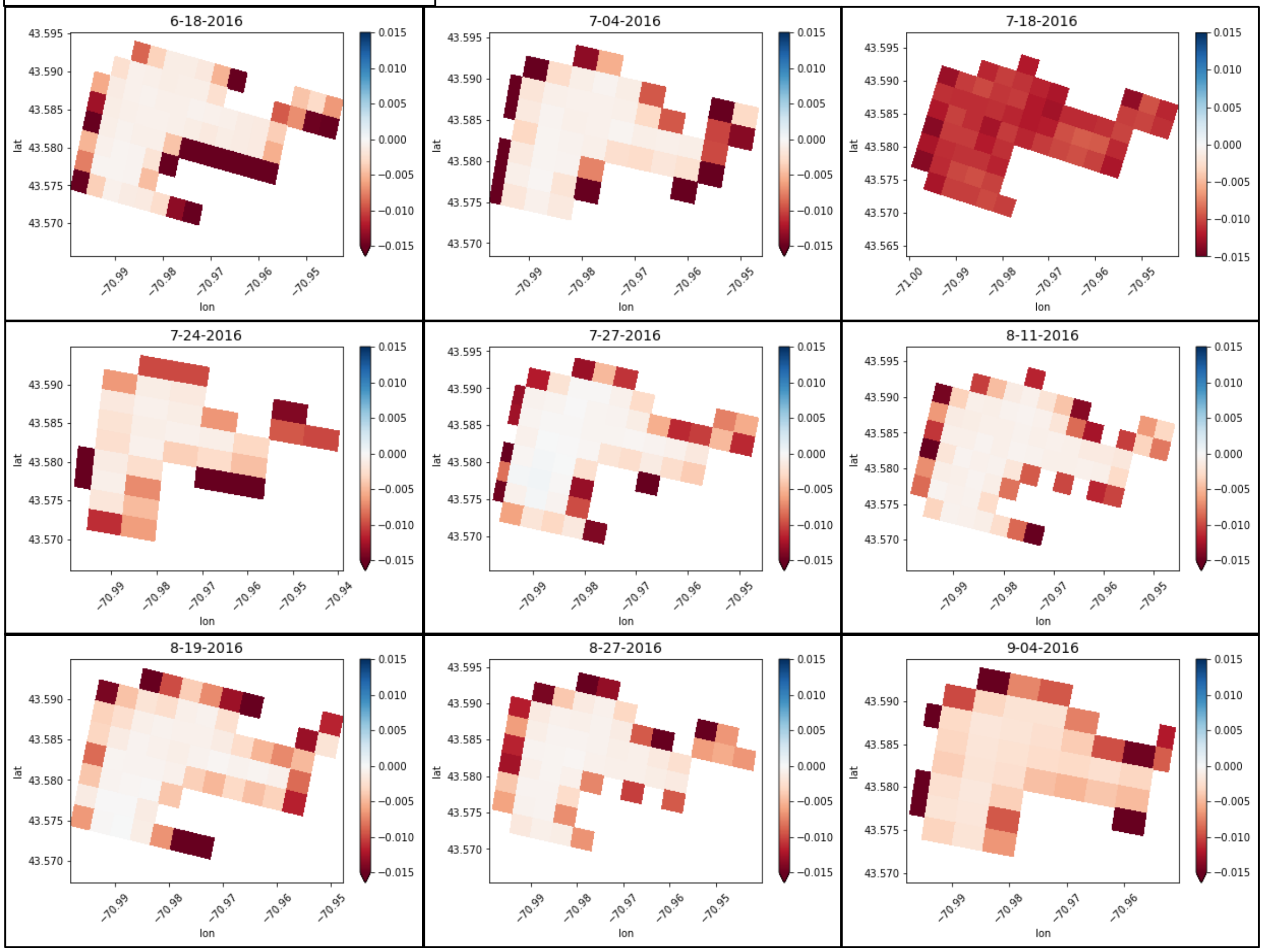
Province Lake, NH



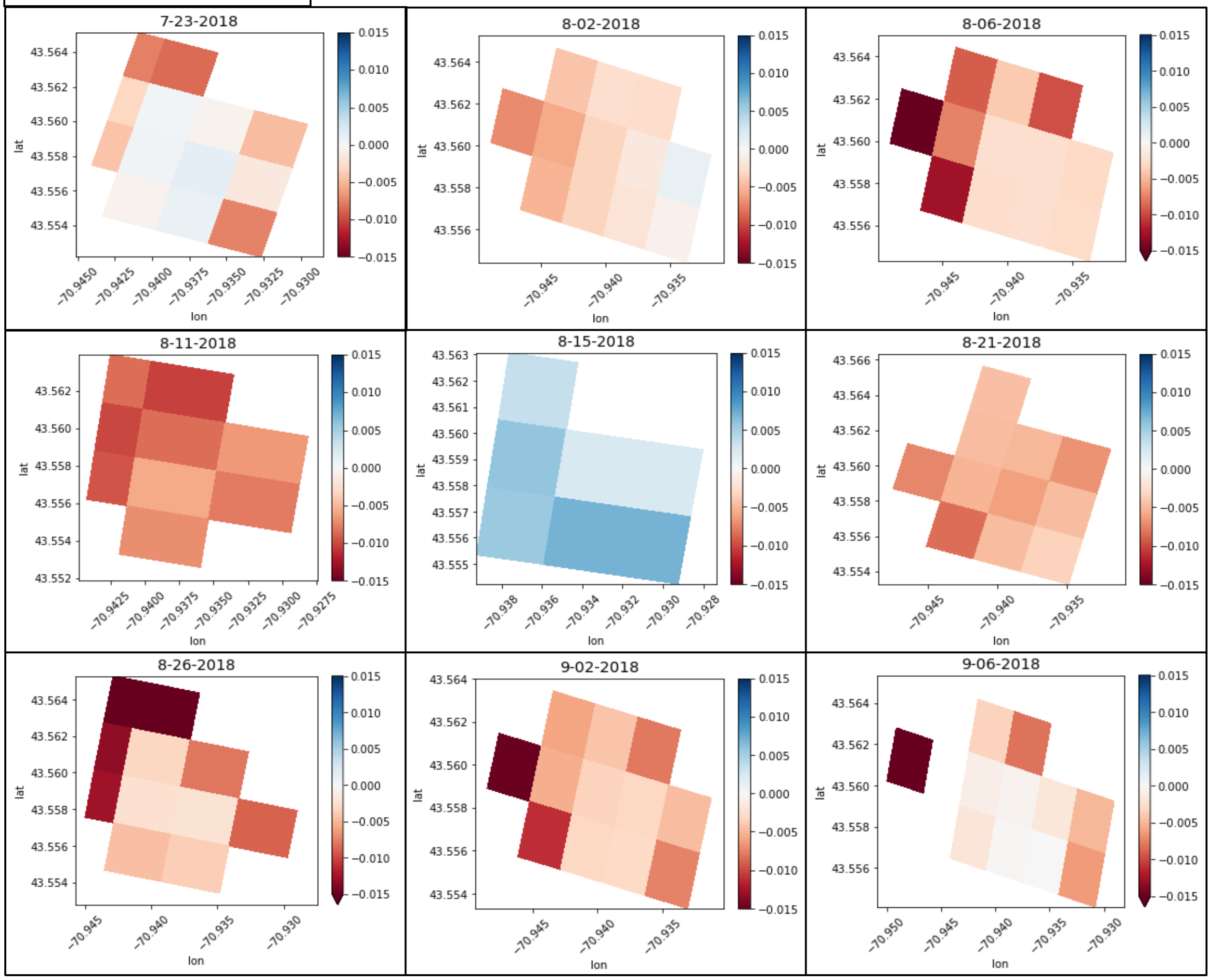
Lovell Lake, NH



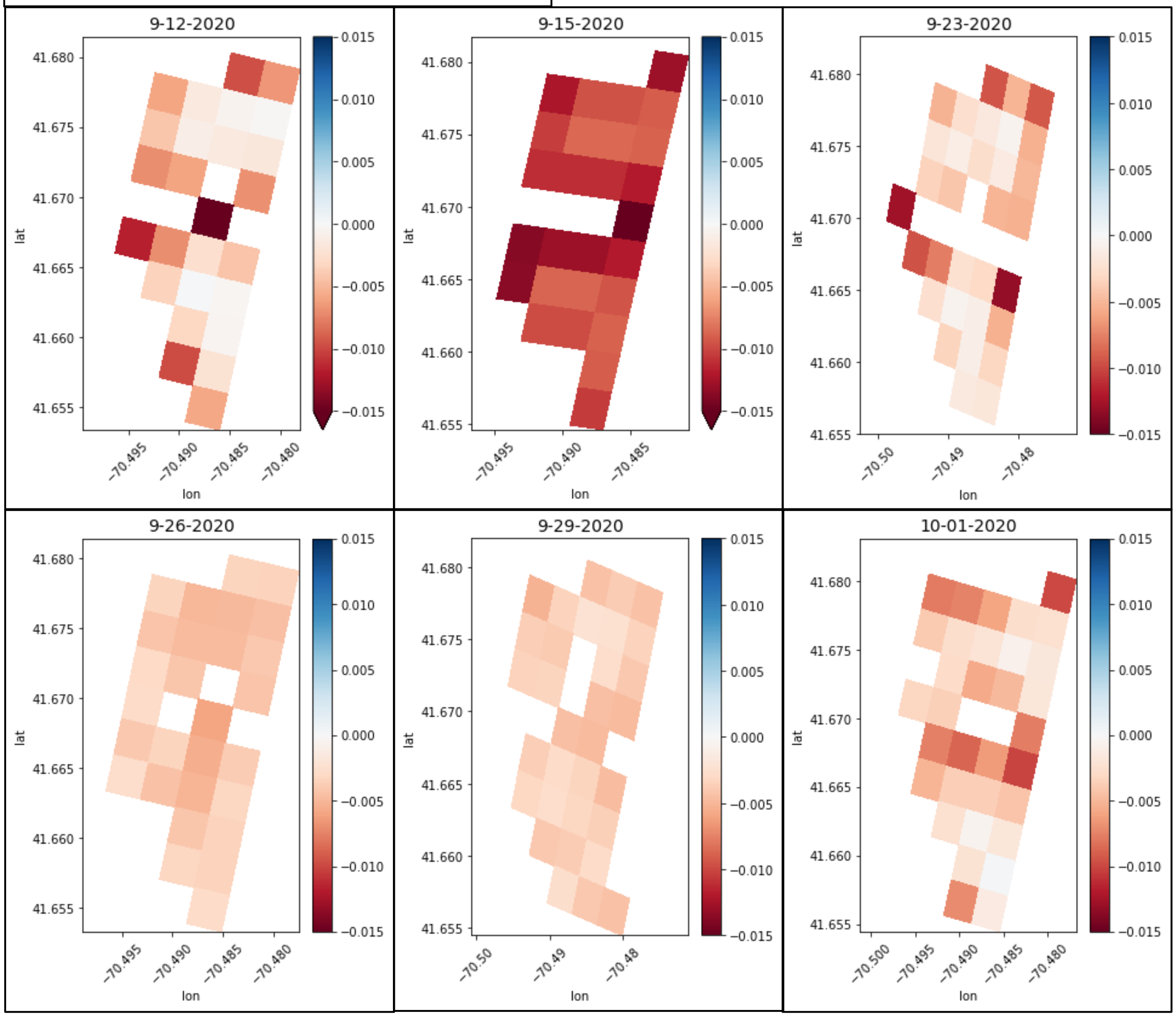
Great East Lake, NH/ME



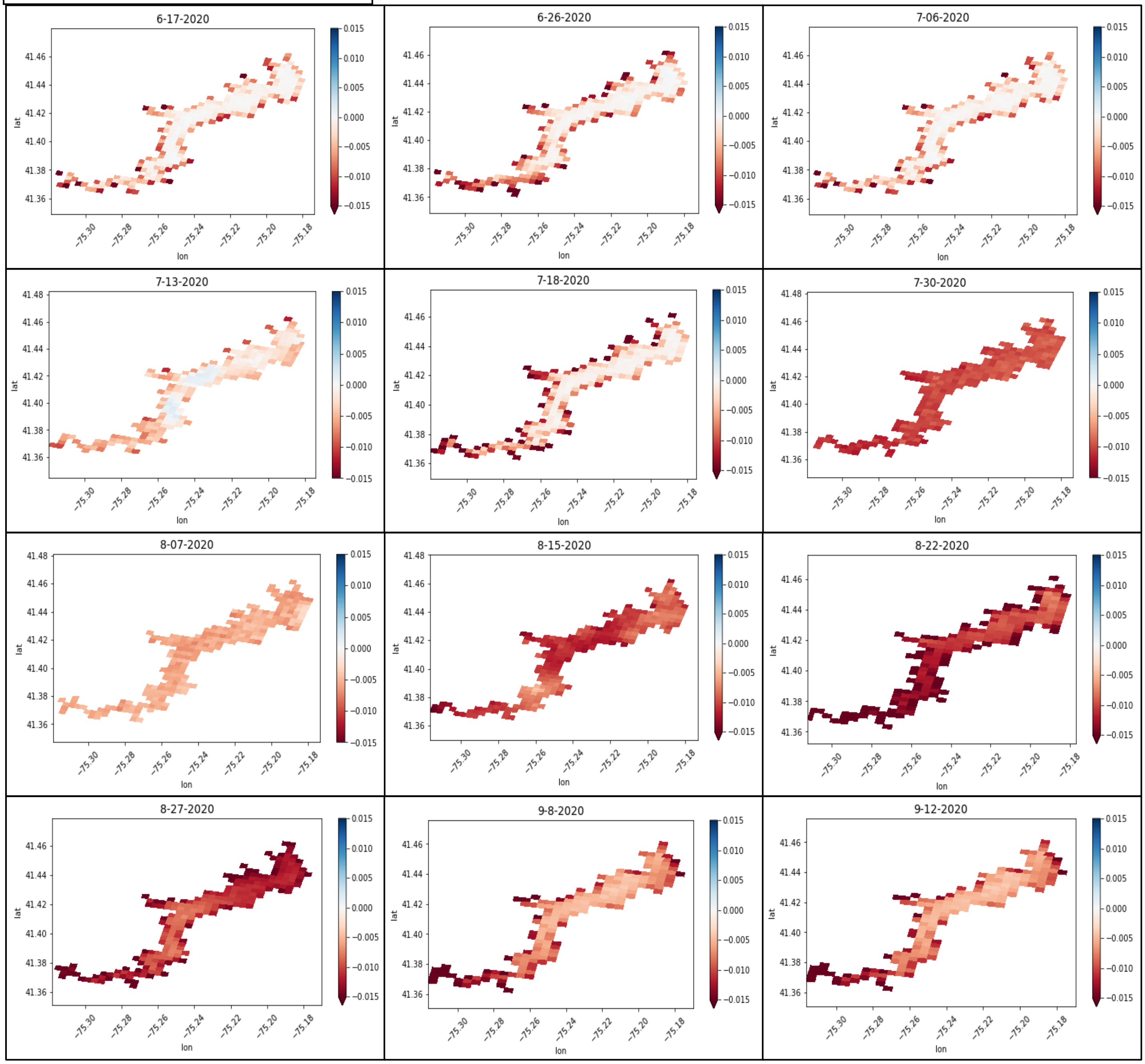
Wilson Lake, ME



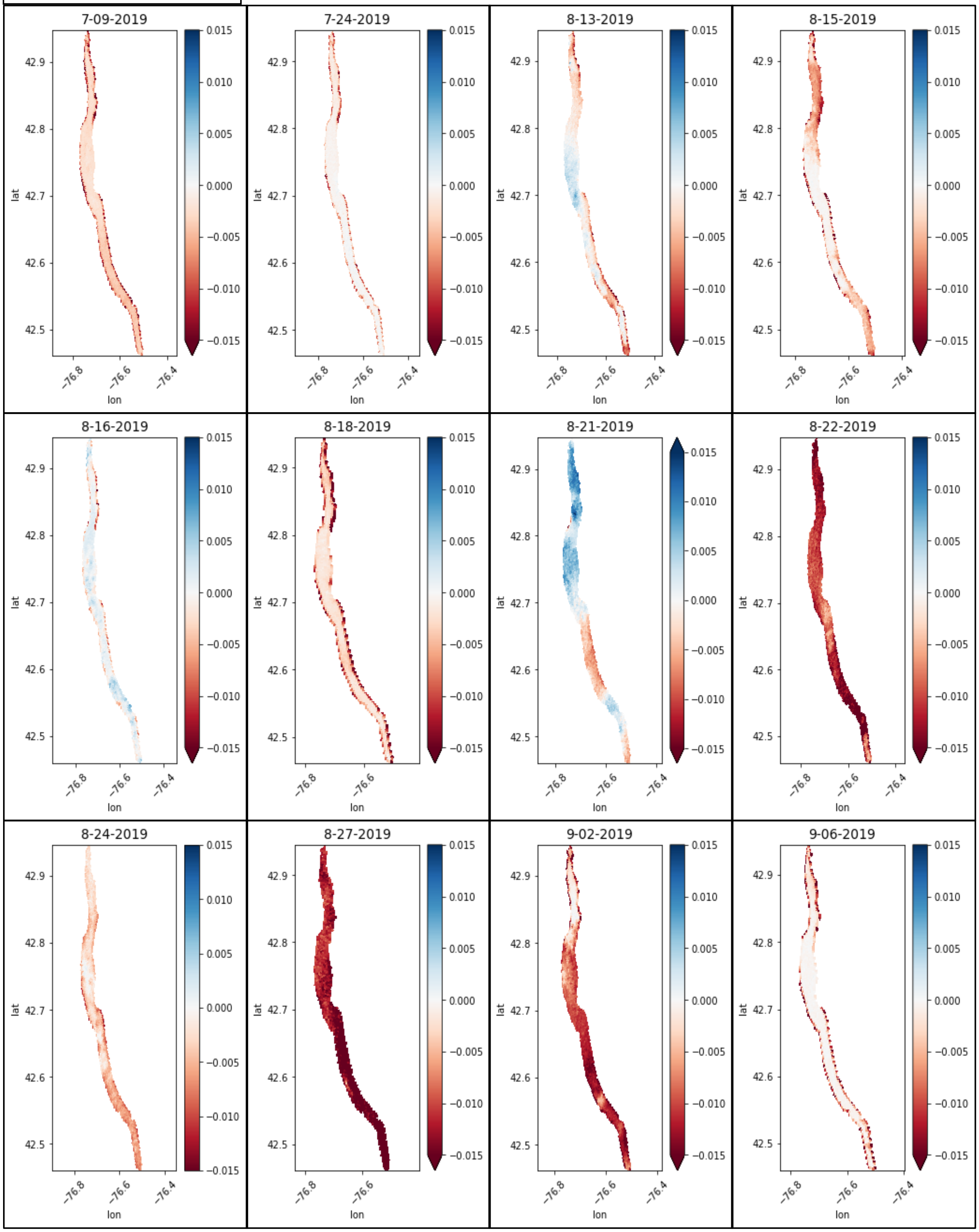
Wakeby Pond/Mashpee Pond, MA



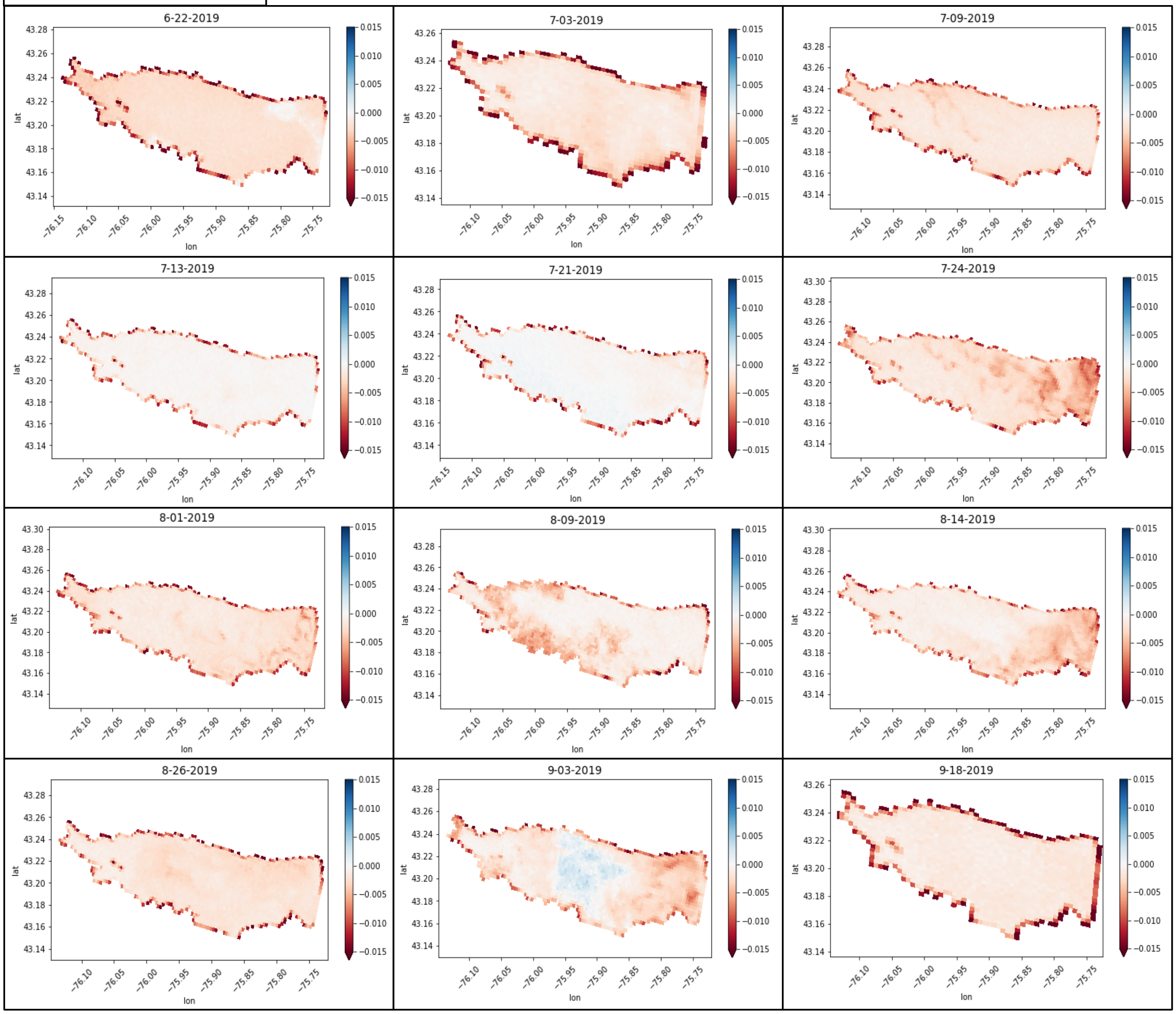
Lake Wallenpaupack, PA



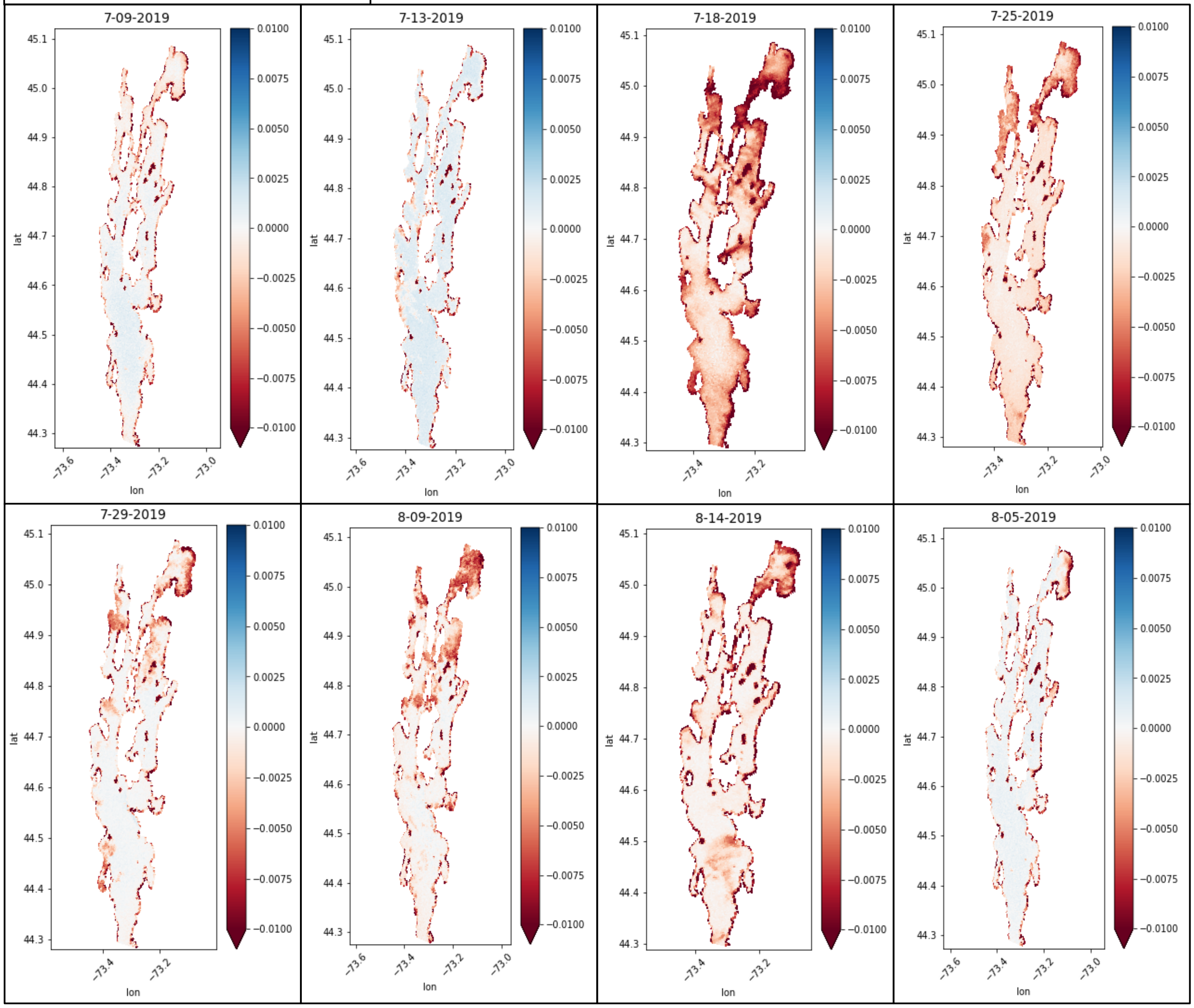
Cayuga Lake, NY



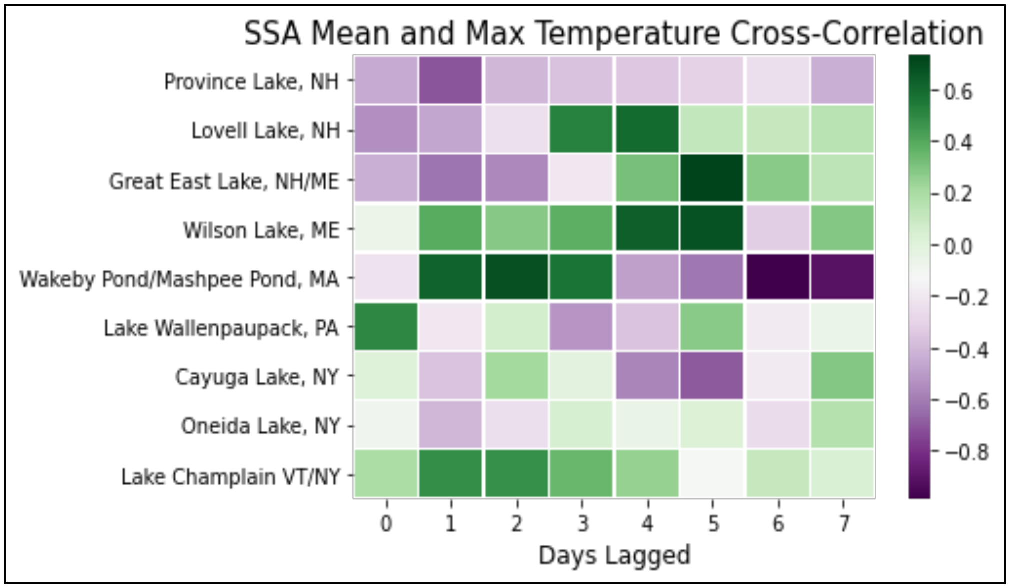
Oneida Lake, NY

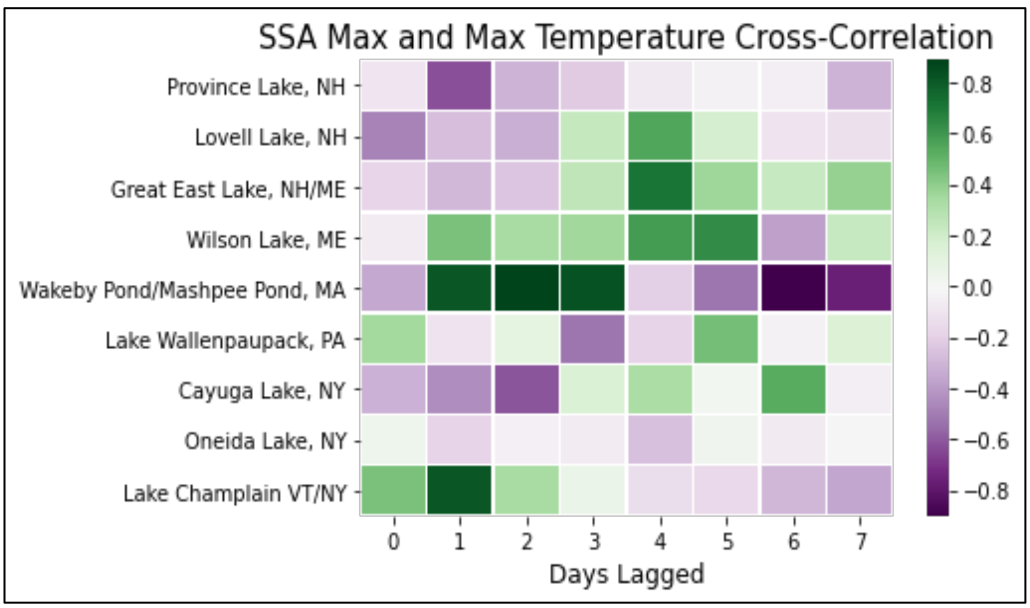


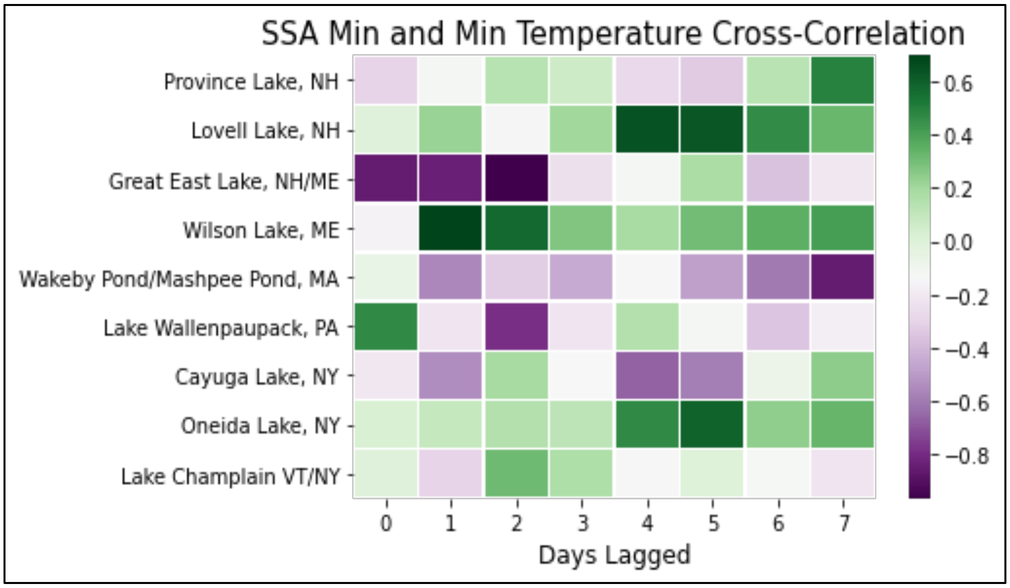
Lake Champlain VT/NY

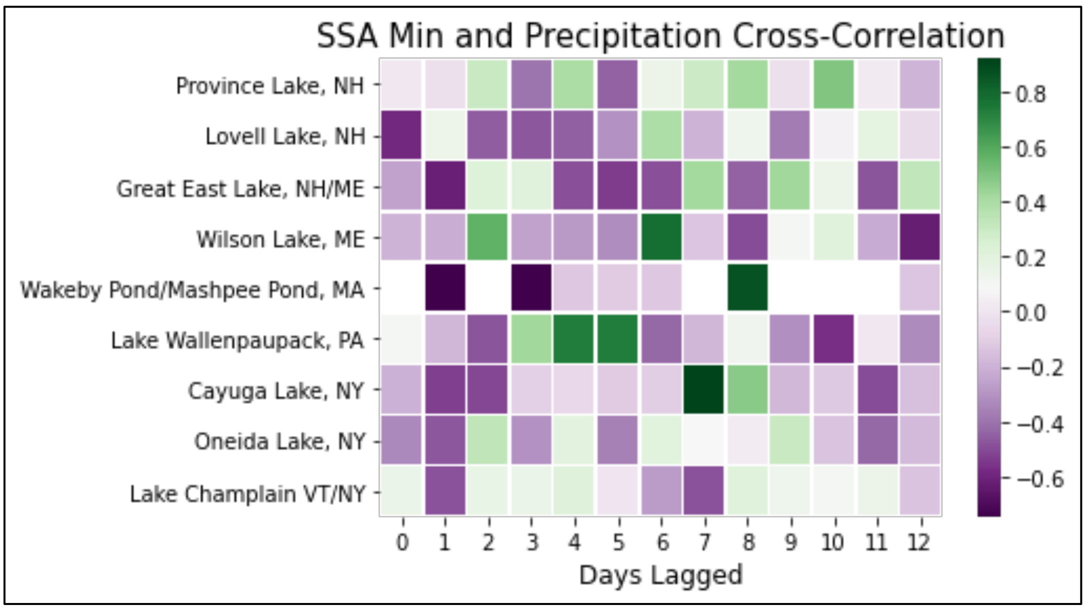


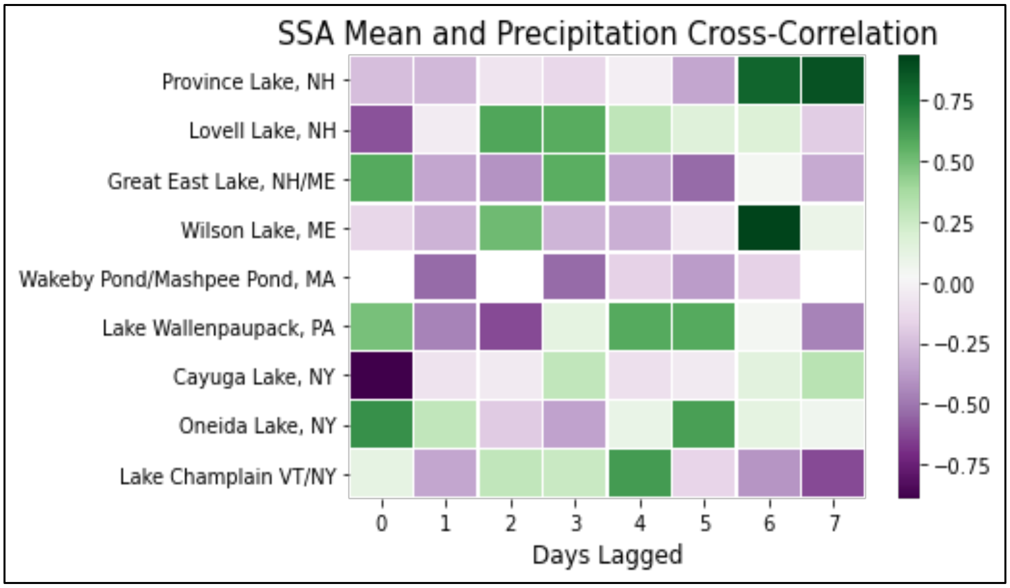
Appendix C – Cross-Correlation of Meteorological Environmental Precursors and SSA Statistical Values for each Study Lake

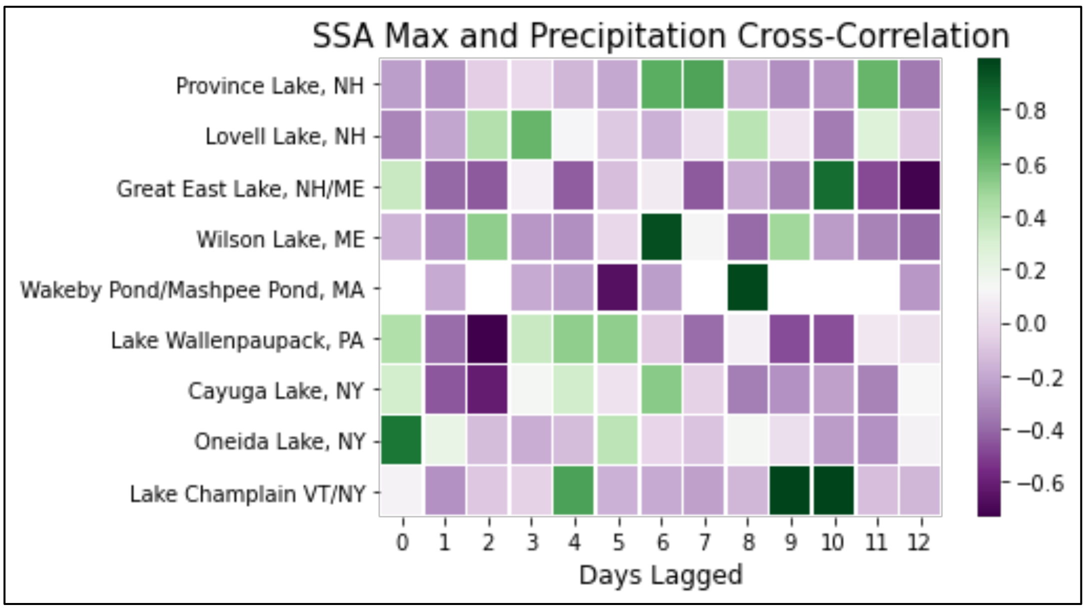





SSA Min and Min Temperature Cross-Correlation
Province Lake, NH
Lovell Lake, NH
Great East Lake, NH/ME
Wilson Lake, ME
Wakeby Pond/Mashpee Pond, MA
Lake Wallenpaupack, PA
Cayuga Lake, NY
Oneida Lake, NY
Lake Champlain VT/NY
0
1
2
3
4
5
6
7
Days Lagged
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8


SSA Min and Precipitation Cross-Correlation
Province Lake, NH
Lovell Lake, NH
Great East Lake, NH/ME
Wilson Lake, ME
Wakeby Pond/Mashpee Pond, MA
Lake Wallenpaupack, PA
Cayuga Lake, NY
Oneida Lake, NY
Lake Champlain VT/NY
0
1
2
3
4
5
6
7
8
9
10
11
12
Days Lagged
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6


SSA Mean and Precipitation Cross-Correlation
Province Lake, NH
Lovell Lake, NH
Great East Lake, NH/ME
Wilson Lake, ME
Wakeby Pond/Mashpee Pond, MA
Lake Wallenpaupack, PA
Cayuga Lake, NY
Oneida Lake, NY
Lake Champlain VT/NY
0
1
2
3
4
5
6
7
Days Lagged
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75


SSA Max and Precipitation Cross-Correlation
Province Lake, NH
Lovell Lake, NH
Great East Lake, NH/ME
Wilson Lake, ME
Wakeby Pond/Mashpee Pond, MA
Lake Wallenpaupack, PA
Cayuga Lake, NY
Oneida Lake, NY
Lake Champlain VT/NY
0
1
2
3
4
5
6
7
8
9
10
11
12
Days Lagged
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6