Grand Valley State University

ScholarWorks@GVSU

9-30-2022

Using Sentinel-3 Satellite Imagery for Cyanobacteria Development Monitoring

Jillian Greene
Grand Valley State University

Follow this and additional works at: https://scholarworks.gvsu.edu/honorsprojects

Using Sentinel-3 Satellite Imagery for Cyanobacteria Development Monitoring

Jillian Greene[1]

[1] Frederik Meijer Honors College, Annis Water Resource Institute;

Grand Valley State University

September 30, 2022

Abstract

Great Lakes basins and inland waterways experience increased harmful algal bloom (HAB) production as a result of eutrophication– excess nutrient loading into waterways from agriculture and urbanization. Muskegon Lake, an estuary located on the eastern coast of Lake Michigan, was classified as a Great Lakes Area of Concern in 1985 due to severe HABs from several impairments including agricultural runoff and industrial waste dumping. Because of the ecosystem and human health threats posed by HABs such as wildlife and human illness and drinking water contamination, monitoring the abundance of HABs in inland waterways is imperative. The goal of this study was to evaluate the application and validity of remote sensing cyanobacteria abundance in Muskegon Lake using Sentinel-3 Ocean and Land Color Instrument (OLCI) satellite imagery from 2016-2021. Results show that the cyanobacteria quantification algorithm calculated high cell densities up to 1.25 million compared to in situ measurements up to 25,000 from a phycocyanin probe on the Muskegon Lake Buoy Observatory. However, remotely sensed cyanobacteria densities follow the same growth and die-off trends as calculated by the Buoy phycocyanin probe. Moreover, the results of this study display that remote sensing cyanobacteria HABs can supplement field data in flagging high HAB days for public health monitoring as well as documenting historical growth trends.

Introduction

Harmful algal blooms (HABs) have been increasing in their presence in all five of the Great Lakes for several decades (Sayers et al., 2016; Reinl et al., 2020). HABs occur with eutrophication–when there are increased nutrients in the lakes, allowing algae to bloom quickly. To be considered a toxic algal bloom, the blooms contain cyanobacteria producing blue-green algae that can hold many toxins such as microcystin, a liver toxin. HABs will also result in severe oxygen depletion in the lakes (Mancuso et. al., 2021). In recent years, many limnologists have taken a closer look into predicting a season's algal blooms as well as how to control them (Chaffin et al., 2019; Paerl and Huisman, 2008).

Although documented in all five Great Lakes, HABs are most prominent in Lake Erie due to its morphometry and urbanized surroundings. Presence of agriculture influences eutrophication, or development of HABs, because of the high concentrations of nitrogen and phosphorus that are in fertilizers which subsequently enters the lake system through stormwater runoff (EPA, 2021). Similar conditions occur in the Muskegon Lake estuary on the eastern coast of Lake Michigan causing it to receive the Area of Concern designation in 1985 allowing for in-depth studies and remediation techniques.

It is important to study HABs annually and throughout all four seasons to analyze potential implications of eutrophication and HABs. Preexisting known concerns regarding HABs and toxin production include illness in wildlife and humans, unsafe conditions for swimming, and contaminated drinking water (Mancuso et. al., 2021). With these dangers, HABs threaten the ecosystems health through dead zones causing fish mortality, and HABs threaten human health and tourism with swimming and drinking water toxins (Coffer et al., 2021).

With the advent of geospatial technology and remote sensing techniques, numerous studies have been produced using optical satellite imagery and spectral reflectances for monitoring HAB abundance. In 2016, the Sentinel-3 satellite was launched with the Ocean and Land Color Instrument (OLCI) which has the capability to collect spectral reflectance data for 21 different wavelengths (ESA, 2022). The numerous spectral bands provided by the OLCI allow for in-depth geographical analysis including chlorophyll reflectances. This study utilizes the Sentinel-3 OLCI to provide a thorough description of the application and validity of remote sensing cyanobacterial density on the Muskegon Lake Area of Concern from 2016-2021.

Literature Review

History of Harmful Algal Bloom Monitoring

HABs pose threats to an ecosystem's health and functioning as well as human health through drinking water and recreational activities. HABs, being full of cyanobacteria, are capable of threatening people's access to safe drinking water. For example, in the Lake Erie coastal regions of Ohio, thousands of citizens were left without drinking water in September 2013 (Coffer et al., 2021). For those reasons, HABs have been studied and monitored globally to acquire knowledge on their development, causes, and effects. Lake Erie researchers have found that environmental drivers of HABs include both excess nitrogen and phosphorus concentrations, warm waters, and a lack of thermal stratification (Chaffin et al., 2019). These eutrophic conditions not only allow algae to grow at rapid rates, but they also allow for microcystis, a toxic cyanobacteria, to flourish.

The effect of nitrogen on eutrophication and algae-growing conditions as mentioned by Chaffin et al. (2019) is uncertain. However, a long-term study conducted in Iowa lakes concluded that there was strong evidence of nitrogen and phosphorus co-limitation of aquatic

productivity (Filstrup and Downing, 2017). The goal of Filstrup and Downing was to analyze the effects of anthropogenic eutrophication on numerous Iowa lakes, and their data concluded that there was a nitrogen threshold to be met for chlorophyll-a growth. Many types of cyanobacterias are able to fix atmospheric nitrogen and are able to grow in low-nitrogen environments typically provided as a result of algal succession (the process of algal growth starting with diatoms subsequently becoming green algae and cyanobacteria). Although phosphorus concentrations are more accurate predictors of HAB growth, the nitrogen concentrations should not be ruled out as a factor of HAB development.

Another predictor of HAB growth is aquatic temperatures and thermal stratification. In instances such as Lake Erie, where shallow, warm waters cause a lack of thermal stratification, cyanobacterial blooms are regularly severe (Chaffin et al., 2019). Cyanobacterial blooms are known to favor warm water conditions and can adapt to stratification by forming intracellular gas chambers allowing the algae to float to the warmer surface water (Paerl and Huisman, 2008). Therefore, when lakes have weak stratification, HABs are able to flourish throughout the water column, while they are also able to alter their structure to dominate surface waters in strongly stratified lakes.

Atmospheric temperature and precipitation rates are also well known drivers of HAB production. A study including 1200 US Lakes examined the previously mentioned environmental drivers of HABs with an emphasis on atmospheric temperatures and precipitation (Ho and Michalak, 2019). The researchers found that summer temperatures are a driver of total algal abundance while the length of summer is a driving factor for cyanobacterial abundance. Additionally, temperature was found to have a positive correlation with nutrient concentrations (phosphorus) which provides evidence for the link between temperature and HABs. Ho and

Michalak (2019) discuss the implications of precipitation on HABs stating that there is no direct correlation; however, increased precipitation leads to increased nutrient runoff which does contribute to HAB development. A separate study on the Muskegon Lake Area of Concern in 2019 corroborates the ideas presented by Ho and Michalak. Mancuso et al., (2020) conducted biweekly samples of Muskegon Lake in the summer of 2019 and determined that diatoms, the earliest stage of algal succession, occupied the lake in an unseasonably cold and wet summer. Similarly, a secondary long-term study of Muskegon Lake (2003-2016) found that despite a general warming water trend, the restoration efforts and reduction of nutrient concentrations is overruling and causing a decrease in HAB abundance over time (Mancuso et al., 2019). With the looming unpredictability of climate change, HABs will become more difficult to forecast and manage, and the subsequent impact on ecosystems remains unknown.

Satellite Monitoring of HABs

In recent decades, geospatial scientists have defined a way to quantify cyanobacterial blooms with the use of satellite imagery. Specifically, the Medium Resolution Imaging Spectrometer (MERIS; 2002-2012) on Envisat-1 and the Ocean and Land Color Instrument (OLCI; 2016) on Sentinel-3 provide imperative spectral resolutions for the calculation of cyanobacterial blooms (Clark et al., 2017). With the Sentinel-3 satellite, there are the respective spectral bands that allow for calculation of cyanobacterial presence at a spatial resolution of 300m; therefore, only 5.6% of the inland lakes, or 685 drinking water sources, within the continental US are resolvable by the satellite imagery (Clark et al., 2017, Coffer et al., 2021). The prospect of using remote sensing to analyze the state of HABs in inland lakes provides an inexpensive, quick, and convenient way for HABs to be monitored at a large scale.

A cross comparison of satellite image-calculated and in situ chlorophyll-a was conducted for accuracy as well as cost in 2019 (Papenfus et al., 2020). Chlorophyll-a data can be used to further determine the presence of cyanobacteria as cyanobacteria contain high amounts of chlorophyll-a working as a proxy. Satellite sensing of HAB biomass has been proven effective in numerous studies; however, it is seldom employed by state or national monitoring programs as compared to field data (Papenfus et al., 2020). Furthermore, the researchers find that satellite monitoring would be best utilized in collaboration with collection of in situ data. Coffer et al. (2021) corroborates that satellite monitoring provides benefits that cannot be withheld by field data; for example, satellite monitoring programs have a high temporal coverage detailing how HABs develop. With the advent of satellite monitoring programs to complement in situ data, scientists will be able to obtain a more comprehensive view of HAB's development and impact on ecosystems.

HAB Satellite Image Pre-Processing Methods

Prior to running any calculations on the satellite imagery, images must be pre-processed with atmospheric correction algorithms in order to remove light scattering effects from aerosols. Rayleigh scattering, or the scattering of light off of air molecules such as ozone or particulate matter, can alter the data retrieved by passive satellite systems. To ensure accuracy of calculations, an atmospheric correction algorithm must be run on all images; however, there are numerous algorithms that provide differing efficacies based on the scenario they are being used in.

Many studies have been shown comparing different correction algorithms and their varying benefits. In a study using Sentinel-3 OLCI images on the Chesapeake Bay area, researchers found numerous benefits to the C2RCC algorithm over BAC, l2gen, and Polymer

(Windle et al., 2022). It was found that C2RCC performed best correcting water-leaving reflectances and top-of-atmosphere reflectances around the chlorophyll absorption wavelengths (681 nm). Windle et al. (2022) determined that while the other algorithms, BAC, l2gen, and Polymer, excelled in other areas (e.g clear day flagging or shorter wavelengths), C2RCC performed best overall.

Similarly, an assessment of six different atmospheric correction algorithms was executed in order to determine the best fit for inland water and coastal regions. Warren et al. (2019) processed Sentinel-2 imagery covering 13 total water bodies using Acolite, C2RCC, iCOR, l2gen, Polymer and Sen2Cor algorithms to run a cross-comparison of accuracy. C2RCC and Polymer were found to have the best overall performance for inland waters and coastal regions while C2RCC also attained the most accurate spectral shape for inland water regions. Sen2Cor and iCOR were the least accurate in coastal waters specifically, due to their land-based methodology. Warren et al. (2019) and Windle et al. (2022) identify that atmospheric correction algorithm accuracy varies based on the application, and that continuing algorithm development and calibration will be crucial in refining geospatial analysis processes.

Remote Sensing of Cyanobacteria Blooms

With the advent of multispectral satellite instruments, geospatial scientists have explored the numerous possibilities of remote sensing research. In 2008, Wynne et al. published an article relating the spectral shape around 681 nm (chlorophyll pigments) to cyanobacterial blooms. Since then, several case studies have been executed for different satellites and regions allowing for a comprehensive view of the relationship between remotely sensed and in situ cyanobacteria data. SeaWiFS, MODIS, MERIS, OLCI, MSI, Landsat, and VIIRS all collect data that can be used to delineate cyanobacteria HABs (Wynne et al., 2018). Most commonly, the MERIS

(2002-2012), MODIS (1999-present), and OLCI (2016-present) are used for cyanobacteria

calculations with MODIS being used to bridge the data between MERIS and OLCI (Wynne et

al., 2021). With intercalibration of the three satellites over the same region of data, satellite

imagery can provide long-term, historical monitoring of cyanobacteria development in inland

lakes and coastal regions.

To quantify cyanobacteria presence in cells/mL concentrations, a cyanobacteria index

(CI) algorithm was developed using a spectral shape (SS) function used as a decision tree. The

spectral shape centered around 681 nm calculates the cyanobacteria concentration (Wynne et al.,

2008). The presence of cyanobacteria in an algal bloom depresses the chlorophyll reflectance

causing the SS(681) to be negative. Furthermore, a secondary spectral shape around 665 is

applied to exclude false positives (non-cyanobacteria bloom) from phycocyanin light absorption

(Lunetta et al., 2015; Matthews et al., 2012). The CI algorithm has been applied to numerous

scenarios to quantify present and historical HABs and has been validated as an effective

national-level monitoring algorithm (Coffer et al., 2020).

There have been previous efforts to create publicly available CI interfaces that allow for

near real-time monitoring as well as historical archives for research. The Cyanobacteria

Assessment Network (CyAN) (Schaeffer et al., 2018) and the Inland Waters Data (ILW) (Seegers

et al., 2021) were jointly developed as operational CI monitoring programs. Funded by the US

Environmental Protection Agency (EPA), National Aeronautics and Space Administration

(NASA), National Oceanic and Atmospheric Administration (NOAA), and the US Geological

Survey (USGS), CyAN and ILW use MERIS and OLCI to quantify CI data and distribute data to

the public. CyAN and ILW help guide drinking water supply and recreation concerns (Seegers et

al., 2021). Schaeffer et al., (2018) details the software and mobile application configuration

noting resource management as the target users. With continued calibration and validity testing, the CI algorithm holds the potential to eliminate the need for regular field sampling for HAB monitoring.

In Situ Phycocyanin Data

In situ data is important for confirming the validity of remotely sensed data. Phycocyanin, a blue-green pigment found in cyanobacterias, can be utilized as a proxy for measuring HAB abundance in a lake. Phycocyanin probes that read the relative fluorescent units (RFU) converted to an estimated cell density can be utilized in waterways to collect copious amounts of real-time data and estimate HAB abundance (Mchau et al., 2019; Genzoli & Kann, 2016). Muskegon Lake houses a Buoy Observatory Project that has collected phycocyanin RFUs and cells/mL measurements every 15 minutes since 2011 (Biddanda et al., 2021). Genzoli and Kann (2016) note that cell density measurements from phycocyanin probes consistently underestimate cells/mL above 3,500 cells/mL while RFUs provide a valid relationship between probe readings and grab samples. This paper will utilize daily means of the Muskegon Lake Buoy Observatory's phycocyanin cells/mL data as a cross comparison for the remotely sensed CI data.

The purpose of this paper is to summarize the application of the CI algorithm over Muskegon Lake to determine seasonal and annual HAB growth and die-off. It will visualize CI data to demonstrate the impacts of HABs throughout the HAB seasons from 2016-2021 on a Great Lakes estuary. The CI images will subsequently be compared to in situ phycocyanin data from a long-term Muskegon Lake Buoy Observatory Project. The remotely sensed data will be plotted against the in situ data to determine the relationship and validity of quantifying cyanobacteria densities from satellite imagery.

Methods

Sentinel-3 Level 1 Earth Full Resolution (EFR) OLCI images were downloaded from the Copernicus Online Data Hub based on their inclusivity of Muskegon Lake (Figure 1) and lack of clouds over the estuary. With the 300m spatial resolution of the OLCI imagery, for this study, Muskegon Lake is ~18km^2 and comprises about 186 pixels. Images were selected based on the lack of clouds specifically over Muskegon Lake; however, clouds elsewhere were not of concern. The cyanobacteria bloom season (May-October) was examined from the Sentinel-3's launch in 2016 through 2021 averaging 2-3 cloud-free images per month.

Figure 1. Aerial imagery of the study area, Muskegon Lake, MI (Sentinel-2 imagery courtesy of Copernicus Open Access Hub, June 16, 2022).

Based on the findings of Windle et al. (2022) and Warren et al. (2019), the C2RCC atmospheric correction algorithm was chosen for the images. C2RCC outputs the atmospherically corrected angular dependent water leaving reflectances, normalized water leaving reflectances, and the top of atmosphere reflectances at 21 different wavelengths. The

images were masked to only the water pixels of Muskegon Lake in order to remove land mixed

data resulting from the coarse resolution of the Sentinel-3 satellite.

The CI can be calculated using a small decision tree of two spectral shape (SS) derivative

equations outlined in Equation 1:

$$SS(\lambda) = \rho_s(\lambda) - \rho_s(\lambda^-) + [\rho_s(\lambda^-) - \rho_s(\lambda^+)] \, (\lambda - \lambda^- / \lambda^+ - \lambda^-), \qquad\qquad (1)$$

where + and - superscripts are indicating one spectral band (ρ_s) above or below the target

wavelength (Wynne et al., 2008). For example, the SS(681) equation's λ^-, λ, and λ^+ references

665, 681, and 709 nm respectively, and the SS(665) equation's λ^-, λ, and λ^+ references 620, 665,

and 681 nm. Furthermore, a -SS(681) equates to the CI as it corresponds to the chlorophyll's lack

of fluorescence at 681 nm. A +SS(665) is used to discern cyanobacteria blooms against typical

phytoplankton blooms as phycocyanin depresses the spectral shape around 665 nm (Lunetta et

al., 2015; Matthews et al., 2012; Seegers et al., 2021). Additionally, a low-pass filter was applied

to both SS(681) and SS(665) layers to refine any potential outlier pixels of data. Finally, if the

filtered SS(681) < 0 and SS(665) > 0, as previously mentioned, then the -SS(681) can be

multiplied by 10^8 to convert to cells/mL (Wynne et al., 2008; Schaeffer et al., 2018).

In situ Muskegon Lake was collected using a phycocyanin probe on the Muskegon Lake

Buoy Observatory Project (Biddanda et al., 2021). The phycocyanin probe works as a HAB

proxy, measuring RFUs and converting to a cells/mL density. Daily means from 2016-2021 were

calculated from the buoy for ground-truth measurements against the remotely sensed CI data.

The CI data and the Muskegon Lake Buoy Observatory data were subsequently plotted in

RStudio by year to analyze the CI algorithm validity and to determine seasonal and annual

changes of HAB development. Finally, the CI data at a 300m buffer around the buoy coordinates

was retrieved and plotted to determine the relationship between remote sensing and in situ lake

data.

Results

To assess the seasonal and annual changes, the CI algorithm was applied to each year

(2016-2021; see Appendix A for all plots) and plotted by year. 2017 (Figure 2) shows the highest

HAB abundance of any year with maximum cyanobacteria peaking in late August and

September at 1,273,057 cells/mL. The 2017 data also shows the longest lasting consistent bloom,
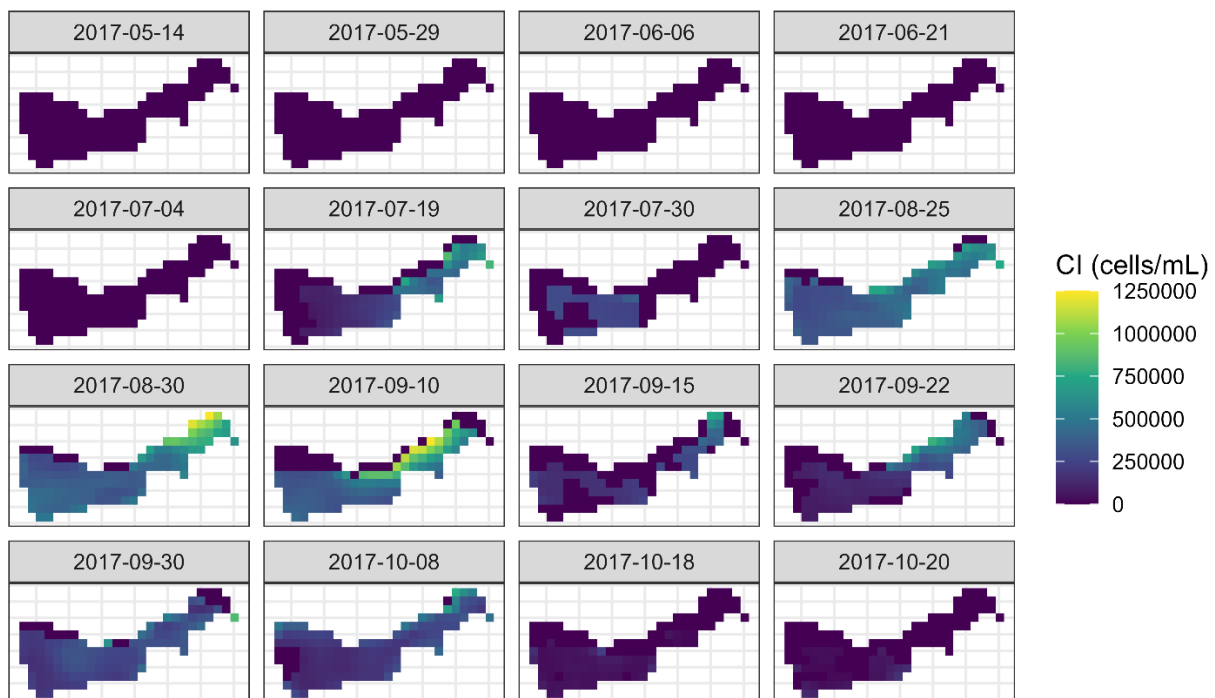
beginning in mid-July through the end of October.

Figure 2. 2017 Cyanobacteria Index algorithm applied to Muskegon Lake, MI.

HABs in 2019, as calculated by the CI algorithm, were almost non-existent with only 4

dates showing cyanobacteria pixels (Figure 3). May 4, July 12, July 13, and September 23

display CI values. The maximum was on July 12 at 186,991 cells/mL, only 14% of the HAB

calculated in August 2017. The Muskegon Lake Buoy phycocyanin data (Figure 4) corroborates

that 2017 was an exceptionally high HAB year– nearly triple the extent of other years– while
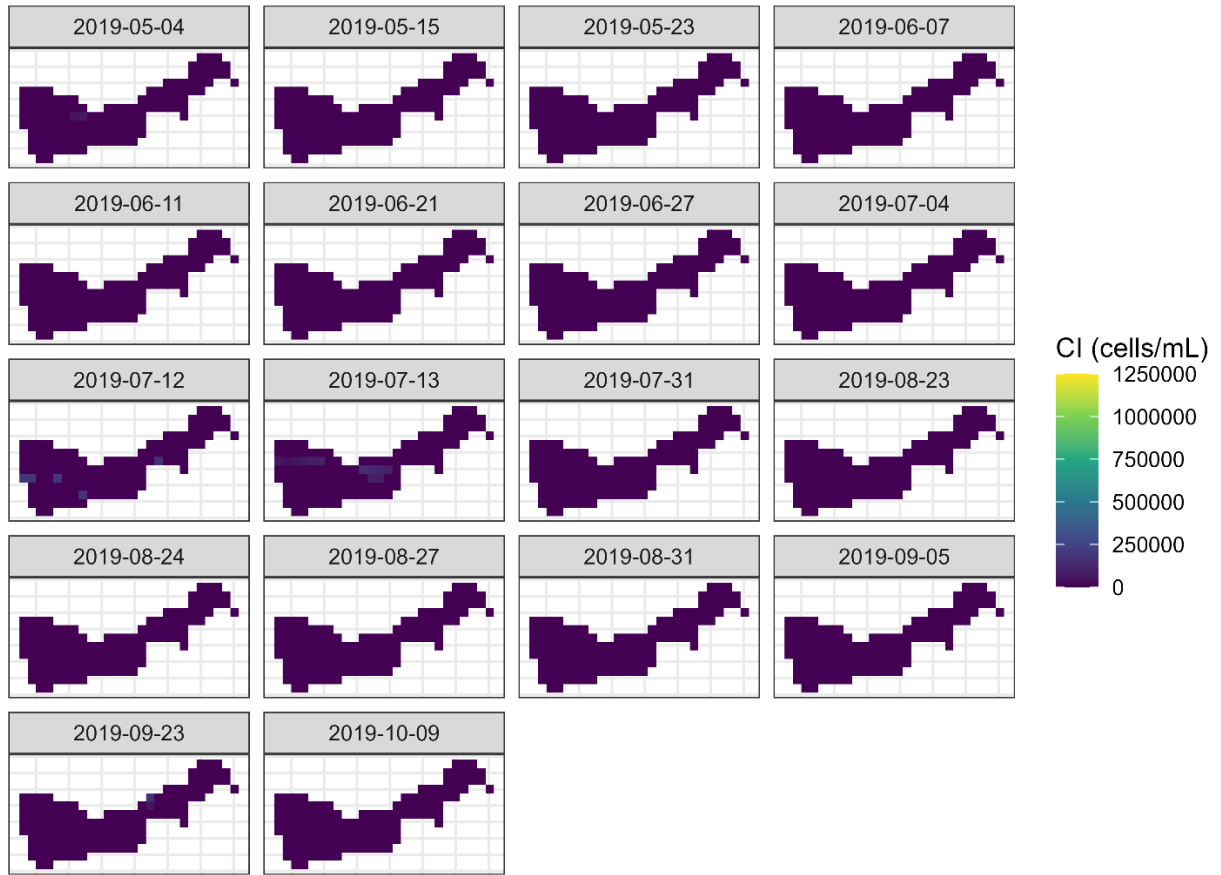
2019 remained lower.

Figure 3. 2019 Cyanobacteria Index algorithm applied to Muskegon Lake, MI.
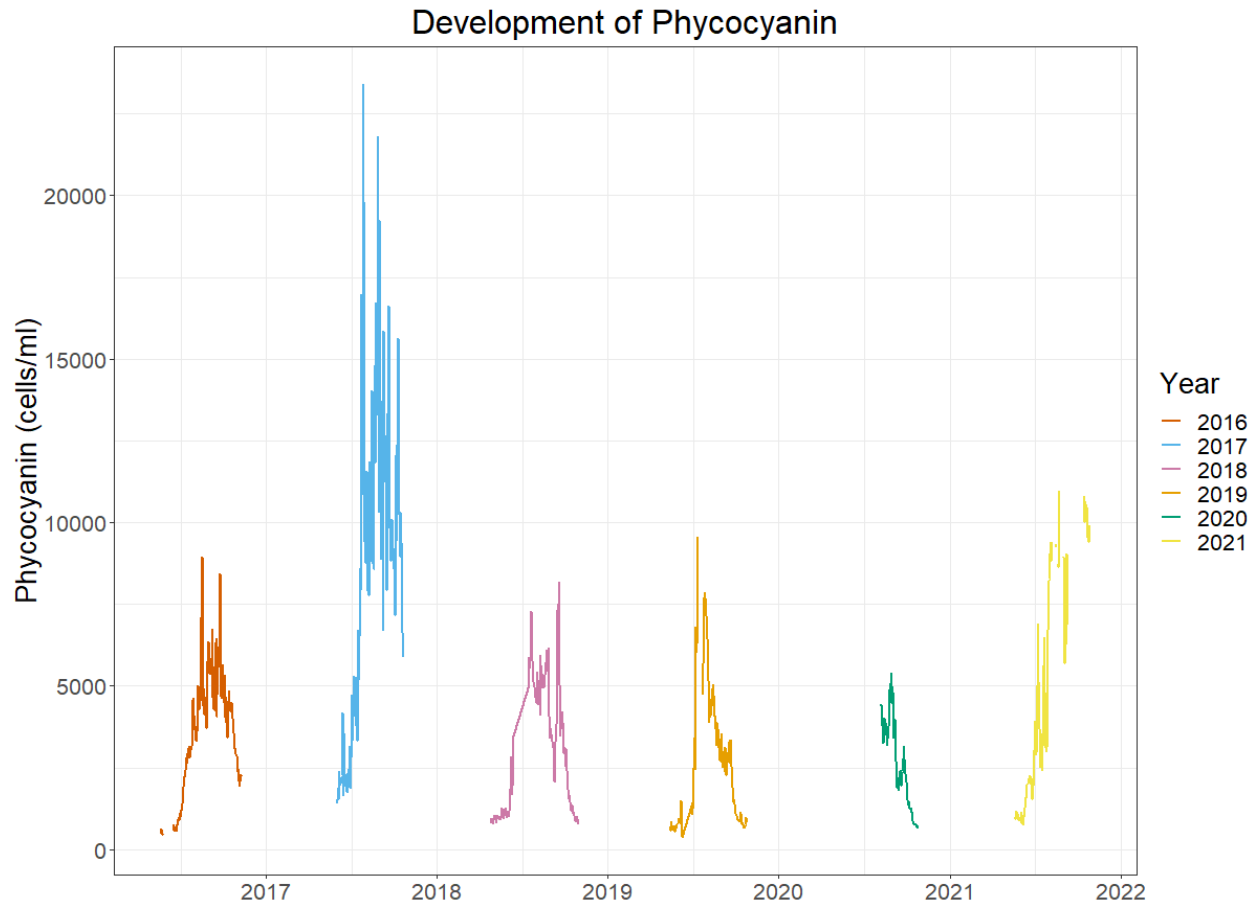
Figure 4. Muskegon Lake Buoy Observatory Phycocyanin data from 2016-2021, (Biddanda et al., 2021). *Note:* 2020 data began on August 6th due to the COVID-19 Pandemic delaying the buoy deployment.

2021 was a slightly higher than average HAB year; however, the Muskegon Lake Buoy data became biofouled and therefore returned unusable data ("Suspect" readings from phycocyanin probe; see Figure 5 for missing buoy data). On the other hand, the CI algorithm had comparable amounts of data for 2021 versus the other years including the dates that the buoy misses (Figure 6).
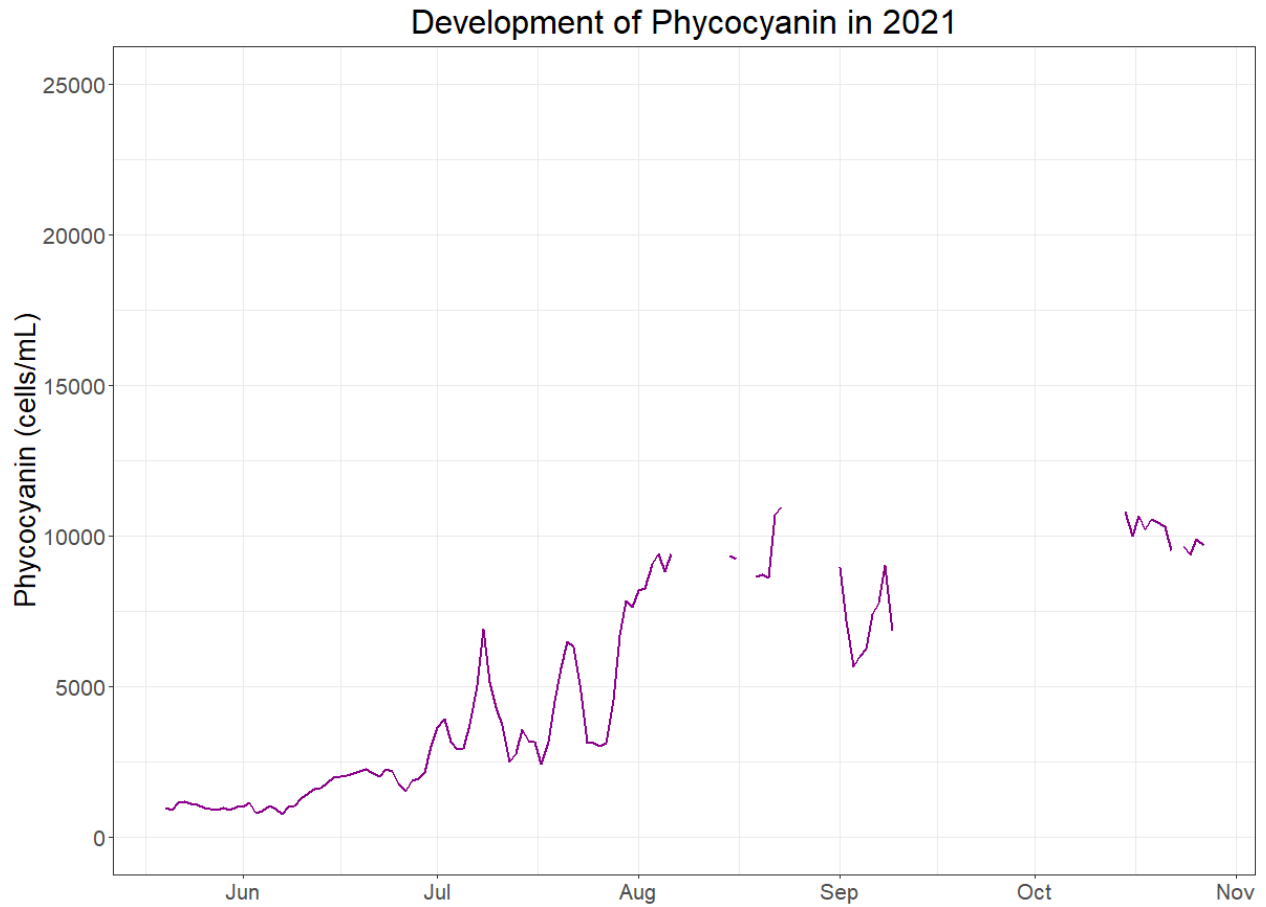
Figure 5. 2021 Muskegon Lake Buoy Observatory phycocyanin data, (Biddanda et al., 2021).

Note: The phycocyanin probe frequently returned "Suspect" or "NaN" leading to gaps in the dataset. The probe was replaced in the 2022 season after further malfunctioning.
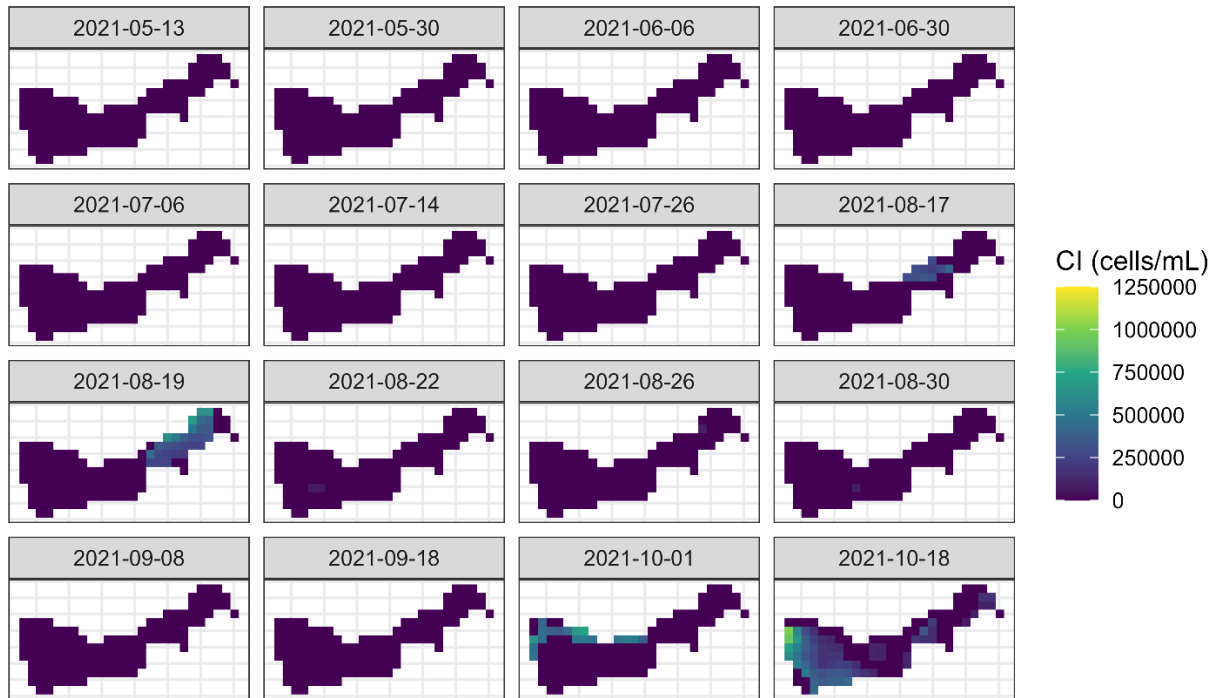
Figure 6. 2021 Cyanobacteria Index algorithm applied to Muskegon Lake, MI.

The CI algorithm from 2021 displayed a maximum value of 984,030 cells/mL on October 18th; 77% of the maximum bloom from 2017. Despite having a high HAB magnitude, 2021's HAB season was not as consistent as the HAB from 2017 that spanned for several months.

The remotely sensed cyanobacteria data was plotted against the in situ Buoy Observatory phycocyanin data in order to determine the correlation despite the difference in magnitude (Figure 7). Remotely sensed data was collected by extracting CI data at the Buoy Observatory coordinates with a 300m buffer in order to include surrounding pixels. The mean of the 300m buffer was calculated as the CI value at the buoy location. Dates with cyanobacteria data both remotely sensed and collected from the buoy returned 12 results: 2 in 2016, 8 in 2017, and 2 in 2021.
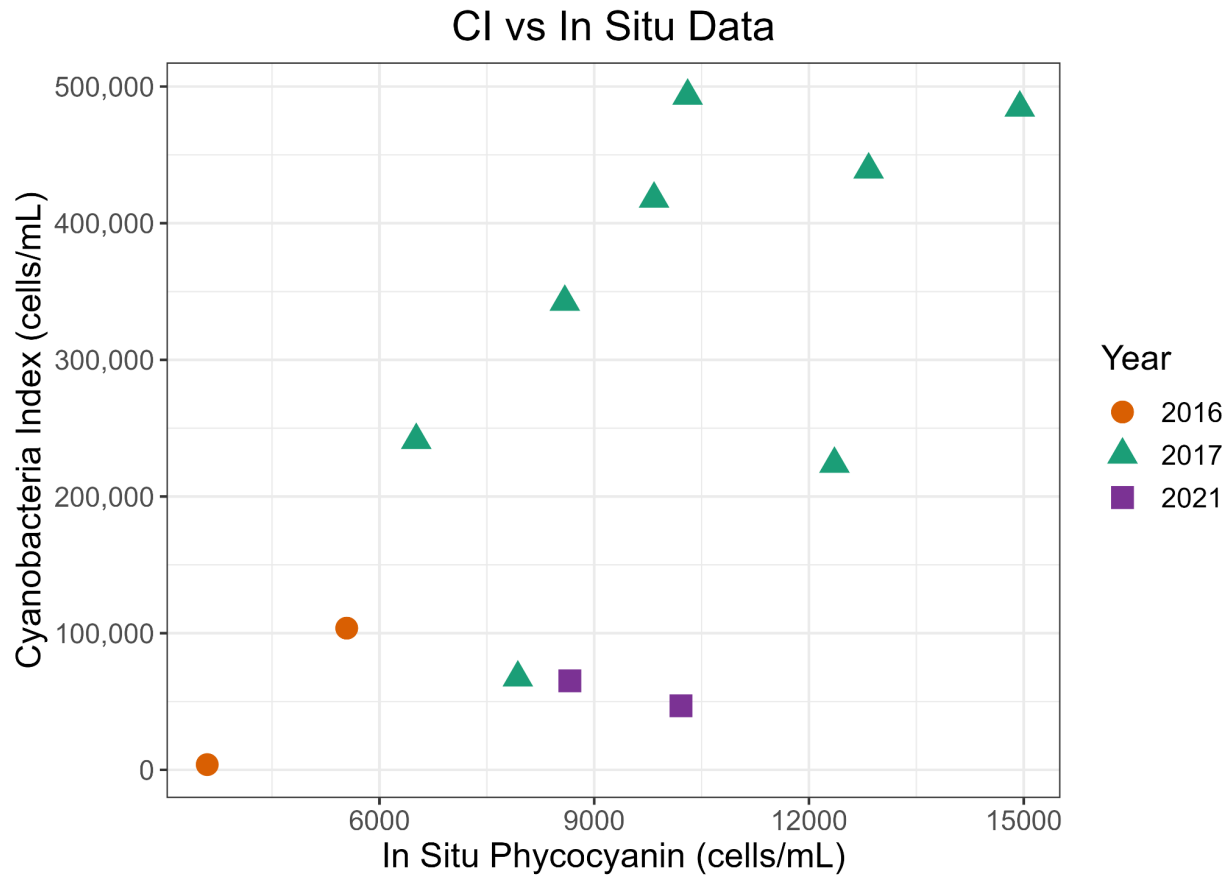
Figure 7. Remotely sensed cyanobacteria (CI) versus in situ cyanobacteria (phycocyanin)

The In Situ Phycocyanin scale ranges from 0-15,000 cells/mL while the CI scale ranges

from 0-500,000 cells/mL. With the exception of 2021, the data generally displays a strong,

positive correlation between the remotely sensed and in situ cyanobacteria cells/mL.

Discussion

Despite the stark difference in cells/mL magnitudes, the CI algorithm and the buoy data

generally follow the same trends. For example, 2017 was the highest HAB year with peaks in

late-July and August and 2019 was an unusually low HAB year which was corroborated between

the in situ and remotely sensed data. Disregarding the two 2021 data points, Figure 7 displays the

presence of a strong correlation between remotely sensed and field collected data. Therefore,

with further calibration, the remotely sensed data holds the potential to return near-accurate

cyanobacteria values. Furthermore, the two 2021 data points displayed in Figure 7 could be

outliers due to the biofouling and malfunctioning of the phycocyanin probe (see Figure 5 for the

copious amounts of missing data). The phycocyanin probe was subsequently replaced in 2022

after further malfunctioning. For lakes or watersheds without buoy observatories or large

funding, the CI algorithm could be used to flag high-bloom days for increased recreational or

drinking water quality monitoring. Furthermore, the CI algorithm could be used to supplement

days where the Muskegon Lake Buoy does not return phycocyanin values.

Although the remotely sensed CI data has significantly different bounds than the

Muskegon Lake Buoy data, both the CI and the phycocyanin data withstand issues in validity.

The CI is greatly influenced by the presence and type of atmospheric correction algorithm as it

alters the spectral reflectance values. Additionally, the phycocyanin probe is known to

underestimate cells/mL values (Genzoli & Kann, 2016), has missing data due to biofouling or

probe malfunctions– as seen largely in 2021, and only measures one point in the middle of the

lake where HABs are more known to affect shallow areas.

An additional consequence of using Sentinel-3 for remote sensing cyanobacteria in

Muskegon Lake is the low temporal coverage provided by the dataset. Sentinel-3 has a two day

revisit period providing many images per month; however, with a flyover time in the late

afternoon, a majority of images covering Muskegon Lake contained high cloud cover and were

deemed unusable. This study was able to obtain approximately 2-3 cloud-free images per month;

however, other study areas may have more or less success depending on their local climate.

Overall, the CI algorithm provides an efficient and low-cost way to monitor HABs in

waterways globally. Although there are limitations on the CI algorithm's validity and ability to

provide high temporal resolution due to climate conditions, remotely sensing cyanobacteria

blooms stands as an effective way to call attention to possible areas of interest that have

otherwise gone overlooked. If public health managers had access to regularly calculated CI data,

they could verify potential blooms to improve warning time for concerns such as microcystis

poisoning hazards. With higher levels of research and atmospheric correction calibration

techniques, the CI algorithm holds the potential to effectively monitor HAB development over

large areas of interest.

References

Biddanda, B., S. Kendall, A. Weinke, I. Stone, N. Dugener, S. Ruberg, J. Leidig, E. Smith, M.

    Berg, and G. Wolffe. 2021. Muskegon Lake Observatory Buoy Data: Muskegon Lake,

    Michigan: 2011-2019 ver 1. Environmental Data Initiative (Accessed July 2022).

    https://doi.org/10.6073/pasta/d1ef6101a1870a0b0bd54d3914dc13dc

Chaffin, J. D., Mishra, S., Kane, D. D., Bade, D. L., Stanislawczyk, K., Slodysko, K. N., Jones,

    K. W., Parker, E. M., & Fox, E. L. (2019). Cyanobacterial blooms in the central basin of

    Lake Erie: Potentials for cyanotoxins and environmental drivers. *Journal of Great Lakes

    Research*, *45*(2), 277–289. https://doi.org/10.1016/j.jglr.2018.12.006

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

Coffer, M. M., Schaeffer, B. A., Foreman, K., Porteous, A., Loftin, K. A., Stumpf, R. P., Werdell,

    P. J., Urquhart, E., Albert, R. J., & Darling, J. A. (2021). Assessing cyanobacterial

    frequency and abundance at surface waters near drinking water intakes across the United

    States. *Water Research*, *201*, 117377. https://doi.org/10.1016/j.watres.2021.117377

Environmental Protection Agency (EPA). (2021, November 23). Lake Erie. US EPA.

    https://www.epa.gov/greatlakes/lake-erie

European Space Agency (ESA). (2022, January 25). Sentinel-3 - Sentinel Online. Sentinel

Online. Retrieved August 2, 2022, from

https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-3

Filstrup, C. T., & Downing, J. A. (2017). Relationship of chlorophyll to Phosphorus and

Nitrogen in nutrient-rich lakes. *Inland Waters*, *7*(4), 385–400.

https://doi.org/10.1080/20442041.2017.1375176

Genzoli, L., & Kann, J. (2016). *Phycocyanin probes as a monitoring tool for cyanobacteria in

the Klamath River below Iron Gate Dam*. 78.

Ho, J., & Michalak, A. (2019, November 20). *Exploring temperature and precipitation impacts

on harmful algal blooms across continental U.S. lakes—Ho—2020—Limnology and

Oceanography—Wiley Online Library*.

https://aslopubs.onlinelibrary.wiley.com/doi/10.1002/lno.11365

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

Mancuso, J. L., Weinke, A. D., Stone, I. P., Hamsher, S. E., Villar-Argaiz, M., & Biddanda, B. A.

(2021). Cold and wet: Diatoms dominate the phytoplankton community during a year of

anomalous weather in a Great Lakes estuary. *Journal of Great Lakes Research*, *47*(5),

1305–1315. https://doi.org/10.1016/j.jglr.2021.07.003

Mancuso, J. L., Weinke, A. D., Stone, I. P., Hamsher, S. E., Woller-Skar, M. M., Snyder, E. B., &

Biddanda, B. A. (2021). Bloom and bust: Historical trends of harmful algal blooms in

Muskegon Lake, Michigan, a Great Lakes estuary. *Freshwater Science*, *40*(3), 463–477.

https://doi.org/10.1086/716236

Matthews, M., Bernard, S., & Robertson, L. (2012, September). *An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters—ScienceDirect*. https://www.sciencedirect.com/science/article/pii/S0034425712002350

Mchau, G., Makule, E., Machunda, R., Gong, Y. Y., & Kimanya, M. (2019). *Phycocyanin as a proxy for algal blooms in surface waters: Case study of Ukerewe Island, Tanzania | Water Practice & Technology | IWA Publishing*. https://iwaponline.com/wpt/article/14/1/229/65408/Phycocyanin-as-a-proxy-for-algal-blooms-in-surface

Paerl, H. W., & Huisman, J. (2008). Blooms Like It Hot. *Science*, *320*(5872), 57–58. https://doi.org/10.1126/science.1155398

Papenfus, M., Schaeffer, B. A., Pollard, A., & Loftin, K. (2020, December 2). *Exploring the potential value of satellite remote sensing to monitor chlorophyll-a for US lakes and reservoirs | SpringerLink*. https://link.springer.com/article/10.1007/s10661-020-08631-5

Reinl, K. L., Sterner, R. W., Lafrancois, B. M., & Brovold, S. (2020). Fluvial seeding of cyanobacterial blooms in oligotrophic Lake Superior. *Harmful Algae, 100*. https://doi.org/10.1016/j.hal.2020.101941

Sayers, M., Gary L. Fahnenstiel, Matthew Whitley, & Robert A. Shuchman. (2016). Cyanobacteria blooms in three eutrophic basins of the Great Lakes: a comparative analysis using satellite remote sensing. *International Journal of Remote Sensing, 37(17), 4148–4171*. https://doi.org/10.1080/01431161.2016.1207265

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

Seegers, B., Werdell, P. J., Vandermeulen, R., Wilson, S., Stumpf, R. P., Schaeffer, B. A., Owens,

 T., Scott, J., & Loftin, K. A. (n.d.). *Satellites for long-term monitoring of inland U.S.*

 lakes: The MERIS time series and application for chlorophyll-a—ScienceDirect.

 Retrieved May 25, 2022, from

 https://www.sciencedirect.com/science/article/pii/S0034425721004053?via%3Dihub

Warren, M. A., Simis, S. G. H., Martinez-Vicente, V., Poser, K., Bresciani, M., Alikas, K.,

 Spyrakos, E., Giardino, C., & Ansper, A. (2019). Assessment of atmospheric correction

 algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters.

 Remote Sensing of Environment, *225*, 267–289. https://doi.org/10.1016/j.rse.2019.03.018

Windle, A. E., Evers-King, H., Loveday, B. R., Ondrusek, M., & Silsbe, G. M. (2022).

 Evaluating Atmospheric Correction Algorithms Applied to OLCI Sentinel-3 Data of

 Chesapeake Bay Waters. *Remote Sensing*, *14*(8), 1881.

 https://doi.org/10.3390/rs14081881

Wynne, T. T., Meredith, A., Stumpf, R. P., Briggs, T., & Litaker, R. W. (2020). *Harmful Algal*

 Bloom Forecasting Branch Ocean Color Satellite Imagery Processing Guidelines, 2021

 Update. https://doi.org/10.25923/606T-M243

Wynne, T. T., Mishra, S., Meredith, A., Litaker, R. W., & Stumpf, R. P. (2021). Intercalibration

 of MERIS, MODIS, and OLCI Satellite Imagers for Construction of Past, Present, and

 Future Cyanobacterial Biomass Time Series. *Remote Sensing*, *13*(12), 2305.

https://doi.org/10.3390/rs13122305

Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Warner, R. A., Tester, P. A., Dyble, J., &

Fahnenstiel, G. L. (2008). Relating spectral shape to cyanobacterial blooms in the

Laurentian Great Lakes. *International Journal of Remote Sensing, 29*(12), 3665–3672.

https://doi.org/10.1080/01431160802007640

Appendix A: Cyanobacteria Index for all years
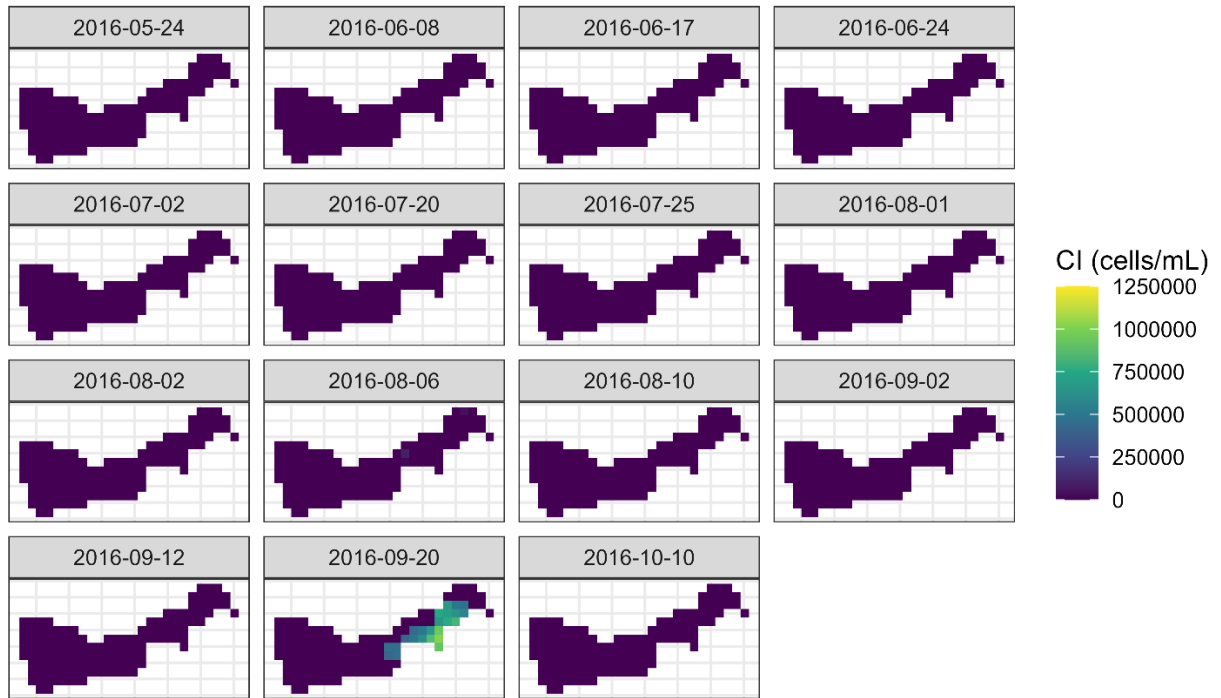
A.1 2016 Cyanobacteria Index of Muskegon Lake, MI
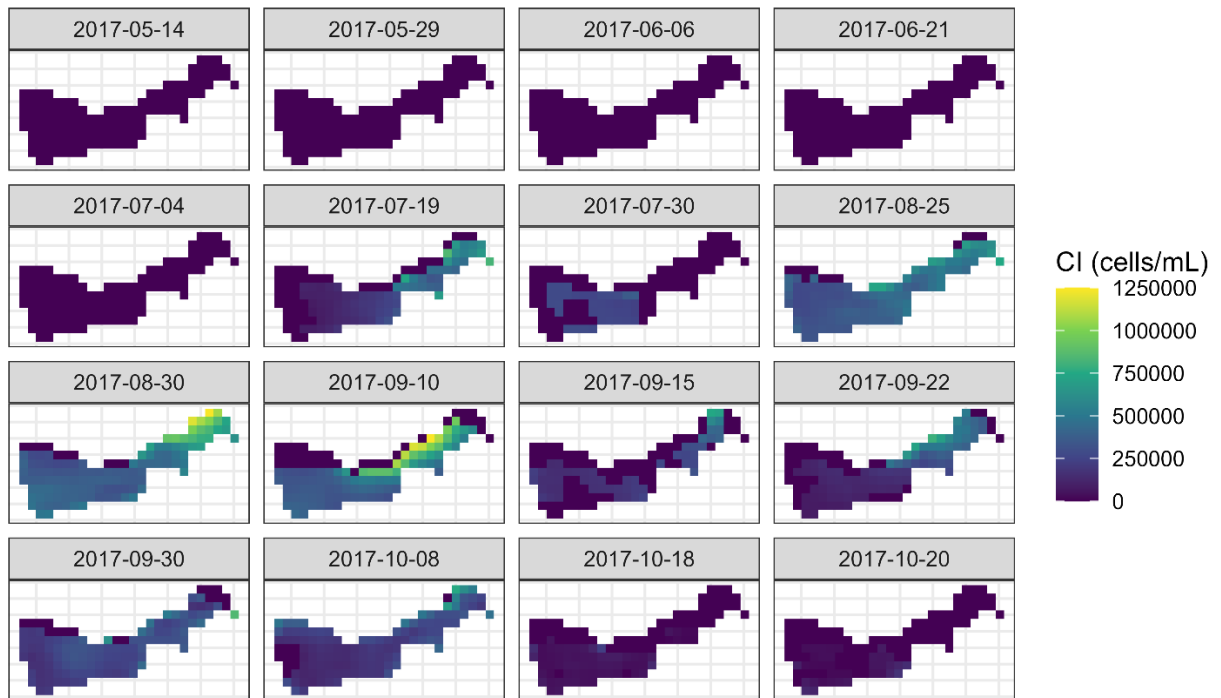
A.2 2017 Cyanobacteria Index of Muskegon Lake, MI

A.3 2018 Cyanobacteria Index of Muskegon Lake, MI

A.4 2019 Cyanobacteria Index of Muskegon Lake, MI

A.5 2020 Cyanobacteria Index of Muskegon Lake, MI

A.6 2021 Cyanobacteria Index of Muskegon Lake, MI

Appendix B: Processing Codes

```
::@echo off
:: enable delayed expansion - used to resolve variable in loop
:: variable has to be used with '!' instead of '%'
setlocal ENABLEDELAYEDEXPANSION


::::::::::::::::::::::::::::::::::::::::::::::
:: User Configuration
::::::::::::::::::::::::::::::::::::::::::::::

:: adapt this path to your needs
set gptPath="C:\Program Files\snap\bin\gpt.exe"

::::::::::::::::::::::::::::::::::::::::::::::
:: Command line handling
::::::::::::::::::::::::::::::::::::::::::::::

:: first parameter is a path to the graph xml
set graphXmlPath=%1

:: second parameter is a path to a parameter file
set parameterFilePath=%2

:: use third parameter for path to source products
set sourceDirectory=%3
:: if sourceDirectory ends with '\' remove it
if %sourceDirectory:~-1%==\ set sourceDirectory=%sourceDirectory:~0,-1%

:: use third parameter for path to target products
set targetDirectory=%4
:: if targetDirectory ends with '\' remove it
if %targetDirectory:~-1%==\ set targetDirectory=%targetDirectory:~0,-1%

:: the fourth parameter is a file prefix for the target product name,
:: typically indicating the type of processing
set targetFilePrefix=%5

:: Create the target directory
md %targetDirectory%


::::::::::::::::::::::::::::::::::::::::::::::
:: Main processing
::::::::::::::::::::::::::::::::::::::::::::::

:: double '%' in batch file and only a single '%' on command line
:: '/D' is for directories like Sentinel data. Remove '/D' when you open files.
for /D /R %sourceDirectory% %%F in (S3*.SEN3) do (
  echo.
  :: '~fF' means abolute path of 'F'
  set sourceFile=%%~fF
  echo Processing !sourceFile!
  :: '~nF' means filename without extension of 'F'
  set targetFile=%targetDirectory%\%targetFilePrefix%_%%~nF.dim
  set procCmd=%gptPath% %graphXmlPath% -e -p %parameterFilePath% -t
"!targetFile!" "!sourceFile!"
```
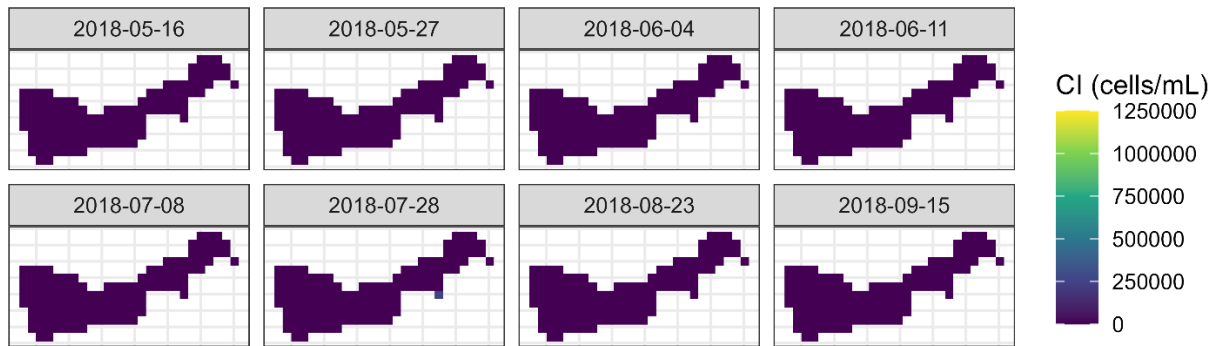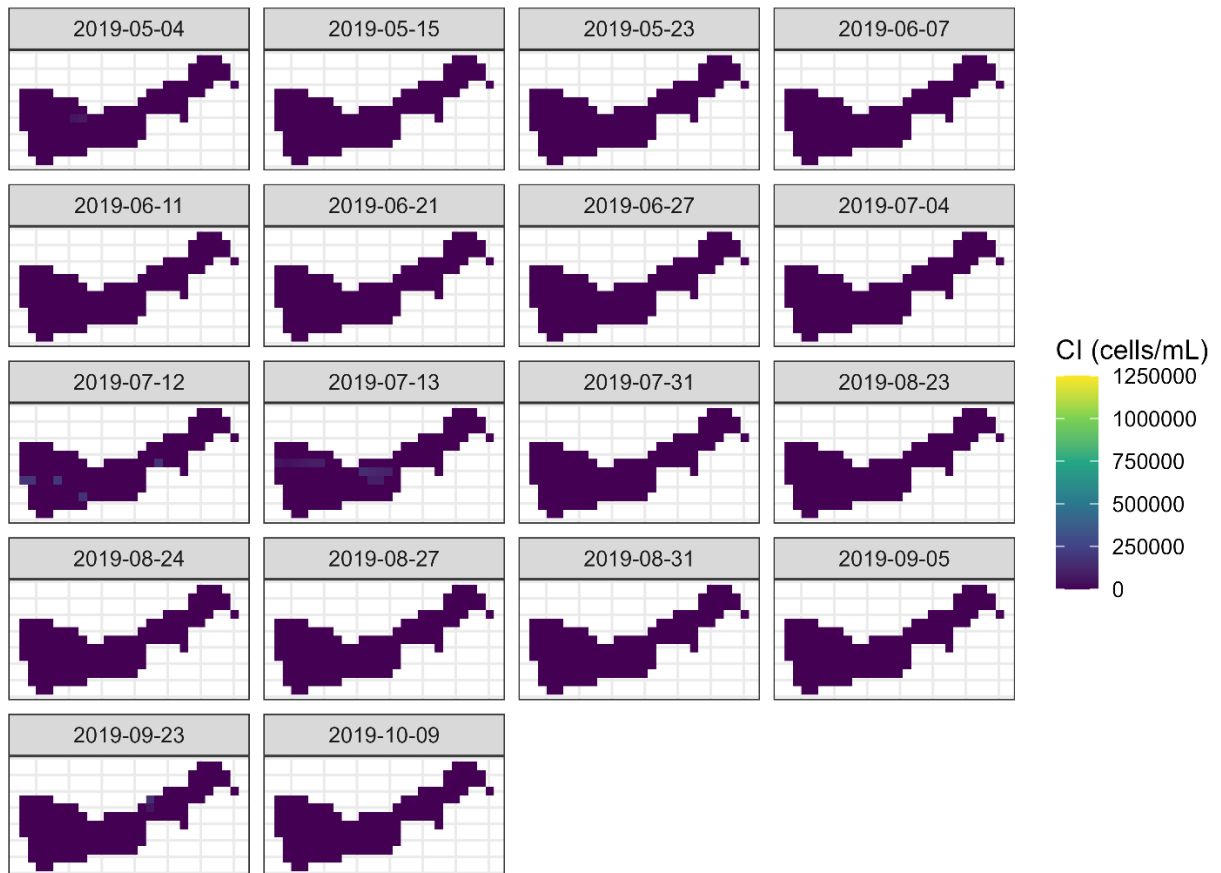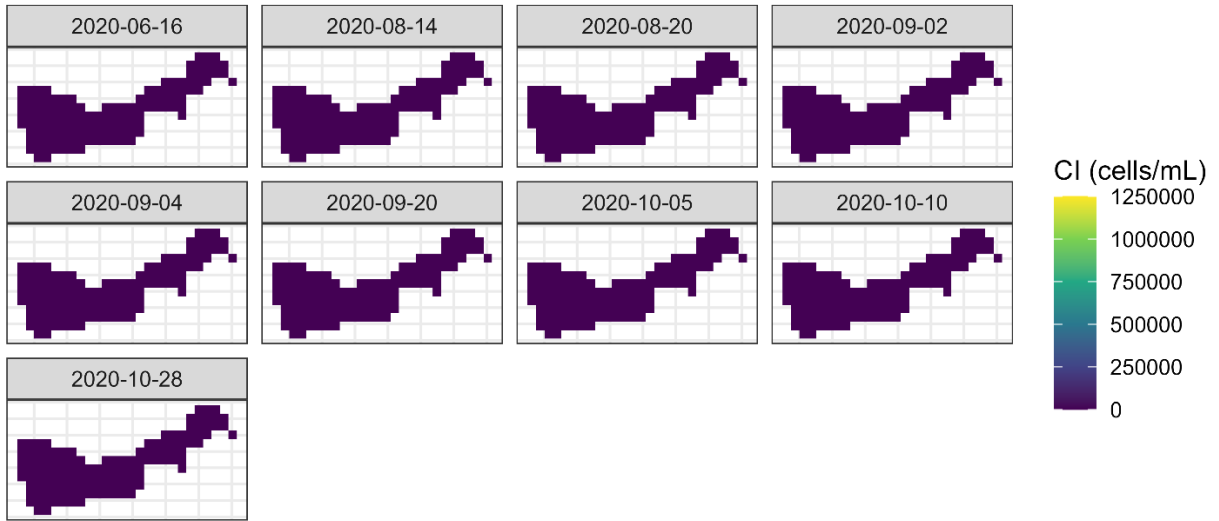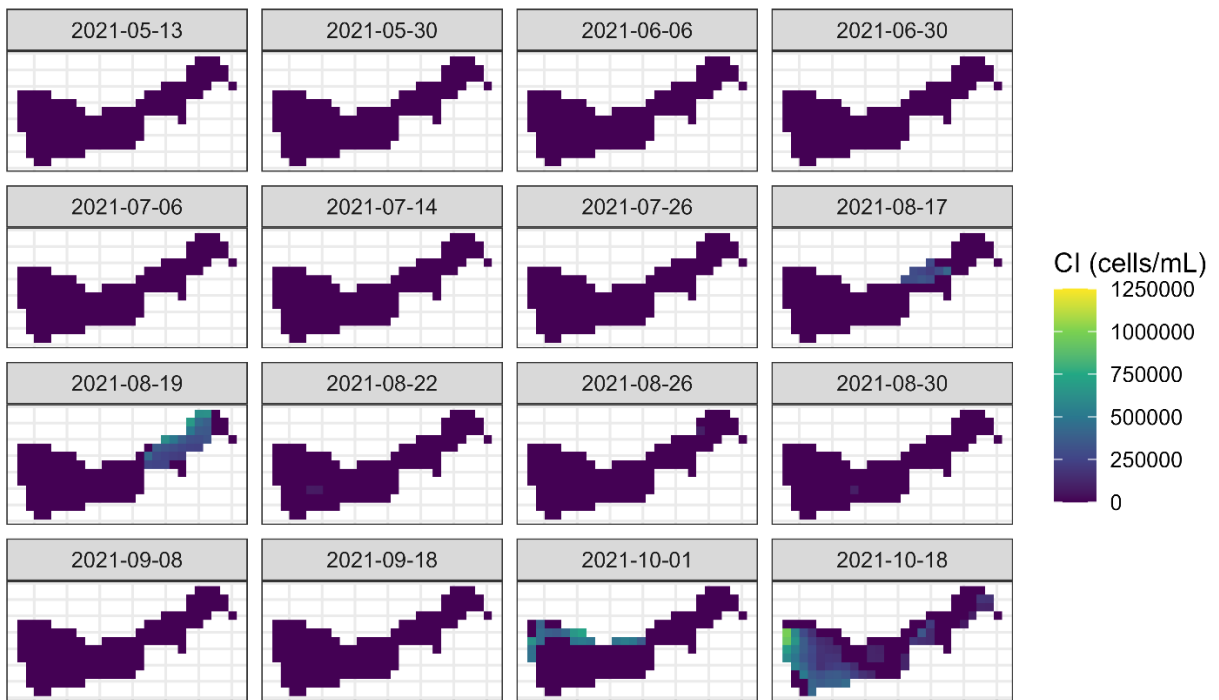
```
  call !procCmd!
)
```

B.1 Batch code for C2RCC atmospheric correction and subset in SNAP

```xml
<graph id="subset_and_apply_c2rcc_to_olci">
  <version>1.0</version>
      <node id="subset">
      <operator>Subset</operator>
      <sources>
            <source>${sourceProduct}</source>
      </sources>
      <parameters>

<sourceBands>Oa01_radiance,Oa02_radiance,Oa03_radiance,Oa04_radiance,Oa05_radi
ance,Oa06_radiance,Oa07_radiance,Oa08_radiance,Oa09_radiance,Oa10_radiance,Oa1
1_radiance,Oa12_radiance,Oa13_radiance,Oa14_radiance,Oa15_radiance,Oa16_radian
ce,Oa17_radiance,Oa18_radiance,Oa19_radiance,Oa20_radiance,Oa21_radiance,altit
ude,latitude,longitude,detector_index,FWHM_band_1,FWHM_band_2,FWHM_band_3,FWHM
_band_4,FWHM_band_5,FWHM_band_6,FWHM_band_7,FWHM_band_8,FWHM_band_9,FWHM_band_
10,FWHM_band_11,FWHM_band_12,FWHM_band_13,FWHM_band_14,FWHM_band_15,FWHM_band_
16,FWHM_band_17,FWHM_band_18,FWHM_band_19,FWHM_band_20,FWHM_band_21,frame_offs
et,lambda0_band_1,lambda0_band_2,lambda0_band_3,lambda0_band_4,lambda0_band_5,
lambda0_band_6,lambda0_band_7,lambda0_band_8,lambda0_band_9,lambda0_band_10,la
mbda0_band_11,lambda0_band_12,lambda0_band_13,lambda0_band_14,lambda0_band_15,
lambda0_band_16,lambda0_band_17,lambda0_band_18,lambda0_band_19,lambda0_band_2
0,lambda0_band_21,solar_flux_band_1,solar_flux_band_2,solar_flux_band_3,solar_
flux_band_4,solar_flux_band_5,solar_flux_band_6,solar_flux_band_7,solar_flux_b
and_8,solar_flux_band_9,solar_flux_band_10,solar_flux_band_11,solar_flux_band_
12,solar_flux_band_13,solar_flux_band_14,solar_flux_band_15,solar_flux_band_16
,solar_flux_band_17,solar_flux_band_18,solar_flux_band_19,solar_flux_band_20,s
olar_flux_band_21,quality_flags</sourceBands>
            <geoRegion>${POLYGON}</geoRegion>
            <subSamplingX>1</subSamplingX>
            <subSamplingY>1</subSamplingY>
            <fullSwath>false</fullSwath>
              <tiePointGridNames/>
            <copyMetadata>true</copyMetadata>
      </parameters>
      </node>
  <node id="c2rcc">
      <operator>c2rcc.olci</operator>
      <sources>
      <sourceProduct>subset</sourceProduct>
        <!--
     <tomsomiStartProduct>${tomsomiStartProduct}</tomsomiStartProduct>
     <tomsomiEndProduct>${tomsomiEndProduct}</tomsomiEndProduct>
     <ncepStartProduct>${ncepStartProduct}</ncepStartProduct>
     <ncepEndProduct>${ncepEndProduct}</ncepEndProduct>
        -->
```

```xml
        </sources>
        <parameters>

        <salinity>0.0001</salinity>
        <temperature>27</temperature>
        <ozone>330</ozone>
        <press>1000</press>
        <TSMfakBpart>1.72</TSMfakBpart>
        <TSMfakBwit>3.1</TSMfakBwit>
        <CHLexp>1.04</CHLexp>
        <CHLfak>21</CHLfak>
        <thresholdRtosaOOS>0.005</thresholdRtosaOOS>
        <thresholdAcReflecOos>0.1</thresholdAcReflecOos>
        <thresholdCloudTDown865>0.955</thresholdCloudTDown865>
        <outputAsRrs>false</outputAsRrs>

<deriveRwFromPathAndTransmittance>false</deriveRwFromPathAndTransmittance>
        <useEcmwfAuxData>true</useEcmwfAuxData>
        <outputRtoa>true</outputRtoa>
        <outputRtosaGc>false</outputRtosaGc>
        <outputRtosaGcAann>false</outputRtosaGcAann>
        <outputRpath>true</outputRpath>
        <outputTdown>false</outputTdown>
        <outputTup>false</outputTup>
        <outputAcReflectance>true</outputAcReflectance>
        <outputRhown>true</outputRhown>
        <outputOos>false</outputOos>
        <outputKd>true</outputKd>
        <outputUncertainties>true</outputUncertainties>
        </parameters>
    </node>
</graph>
```

B.2 XML for C2RCC atmospheric correction and subset in SNAP

```
validPixelExpression=!quality_flags.invalid && (!quality_flags.land ||
quality_flags.fresh_inland_water)

POLYGON=POLYGON((-86.57388641903981 44.16769790459829, -85.57899248327368
44.05690914210022, -85.95584939799701 42.35032915933289, -86.92306751241911
42.46246449598452, -86.57388641903981 44.16769790459829))
```

B.3 Properties file for C2RCC atmospheric correction and subset in SNAP

```python
#Name - RasterCalcCode_v3.py
#Descripton - Automating CI calculations and extract values
#Requirements - Spatial Analyst Extension
#Author - Jillian Greene
#Date - 7/13/22

import arcpy
from arcpy.sa import *
import os
import csv

rootFolder = "D:/Greene/SNAPSentinelData/2016"
ctr = 1


#get list of folders starting with "projected"
folders = [filename for filename in os.listdir(rootFolder) if
filename.startswith("projected")]

#Check out ArcGIS Spatial Analyst extension
arcpy.CheckOutExtension("Spatial")
arcpy.env.overwriteOutput = True
arcpy.env.snapRaster = "D:/Greene/ML Data/ML_raster.tif"

#name csv header and data
header =['Date', 'Time', 'CI']
data = []

#open csv file in write mode
with open ("MLBuoy_CIData.csv", 'w') as f:
      writer = csv.writer(f)
      writer.writerow(header)


#loop through folders and set workspace each time
for i in folders:

      date = i.split("_")[9].split("T")[0]
      time = i.split("_")[9].split("T")[1]

      #set arcpy workspace to folder using arcpy.env.workspace =
      arcpy.env.workspace = "D:/Greene/SNAPSentinelData/2016/" + i

      #set local variables as rasters
      rtoa7 = Raster("rtoa_7.img")
      rtoa8 = Raster("rtoa_8.img")
      rtoa10 = Raster("rtoa_10.img")
      rtoa11 = Raster("rtoa_11.img")
```

```python
    #SS raster calculator
    SS_665 = rtoa8 - rtoa7 + (rtoa7 - rtoa10) * ((665.0-620.0)/(681.0-620.0))
    SS_665.save("D:/Greene/SNAPSentinelData/2016/SS_665_" + date + ".tif")


    SS_681 = rtoa10 - rtoa8+ (rtoa8 - rtoa11) * ((681.0-665.0)/(709.0-665.0))
    SS_681.save("D:/Greene/SNAPSentinelData/2016/SS_681_" + date + ".tif")


    #Set SS equations as rasters
    ss_665 = Raster("D:/Greene/SNAPSentinelData/2016/SS_665_" + date +
".tif")
    ss_681 = Raster("D:/Greene/SNAPSentinelData/2016/SS_681_" + date +
".tif")


    #Low pass filter bands
    filter_665 = Filter(ss_665, "LOW", "DATA")
    filter_665.save("D:/Greene/SNAPSentinelData/2016/filter_665_" + date +
".tif")


    filter_681 = Filter(ss_681, "LOW", "DATA")
    filter_681.save("D:/Greene/SNAPSentinelData/2016/filter_681_" + date +
".tif")


    #CI Equation, take out the ", 0" to make ND transparent
    CI = Abs(Con((filter_681 < 0) & (filter_665 > 0), (filter_681), 0)) *
100000000

    CI.save("D:/Greene/SNAPSentinelData/2016/CI_" + date + ".tif")


    #clip to ML
     CI_Extract = ExtractByMask("D:/Greene/SNAPSentinelData/2016/CI_" + date +
".tif", "D:/Greene/ML Data/ML_WaterOnly_raster.tif")
    CI_Extract.save("D:/Greene/SNAPSentinelData/2016/CI_ML_" + date + ".tif")
    print("D:/Greene/SNAPSentinelData/2016/CI_ML_" + date + ".tif")
```

B.4 Python code for raster calculations via ArcGIS

Reading in CI rasters

After importing raster packages, set working directory, load in rasters and stack *Rasters must be clipped to the same extent for the "stack" function, raster layers needed to be aligned

```
library(purrr)

setwd("D:/Greene/CI_codes/data")

rasters <- list.files("D:/Greene/CI_codes/data/allyears", pattern =
("*aligned.tif"), full.names=TRUE)
rasters2 <- list.files("D:/Greene/CI_codes/data/allyears", pattern =
("*aligned.tif"), full.names=FALSE)

date <- map(strsplit(rasters2, "_"), 1)

rasStack = stack(rasters)
```

Cropping Rasters

Crop single files or raster stacks to a shape file

```
library(stars)

library(sf)
library(ggplot2)

aligned <- read_stars("D:/Greene/CI_codes/data/2021/20210819_aligned.tif")
mlpoly <- st_read("D:/Greene/ML Data/ML_water_shape.shp")

#one file crop
cropped_20210819 <- st_crop(aligned, mlpoly)

#raster stack crop
crop_stack <- stack(raster::crop(rasStack, mlpoly))
crop_stack <- stack(raster::mask(crop_stack, mlpoly))
```

Plotting Single Rasters with a Basemap

Create a plot function to use in a for loop through the raster stack

```
library(ggplot2)
library(ggspatial)
library(stringr)

plot_fun <- function(image){

CI_plot <- ggplot() +
```

```
  ggtitle(label = "2021 Muskegon Lake Cyanobacteria Index",
          subtitle = paste0("Date: ", str_sub(names(image)[1], 6, 7), "-",
str_sub(names(image)[1], 8, 9))) +
  annotation_map_tile(type = "cartolight") +
  layer_spatial(data = image) +
 scale_fill_viridis_c(na.value = "transparent",
                      name = "CI (cells/mL)",
                      limits = c(0, 1250000)) +
  theme_bw() +
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank(),
        axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank())
  ggsave(paste0(names(image)[1], ".png"), CI_plot, dpi=600, width = 10,
height = 7, units="in")
  return(CI_plot)
}

for (i in 1:nlayers(crop_stack)) {
  plot_fun(crop_stack[[i]])
}
```

Making the GIF

Set the list of plots as a magick list using image_read(), string together with image_append(), and animate using image_animate() and image_write()

```
library(magick)

#Read file names based on extension
saved_plots <- list.files("D:/Greene/CI_codes/R", pattern = ("*.png"),
full.names=TRUE)
saved_plots <- image_read(saved_plots)

image_append(image_scale(saved_plots, "x200"))
```

```
animation.2017 <- image_animate(image_scale(saved_plots, "800x800"), fps = 1,
dispose = "previous")
image_write(animation.2017, "2019Cyanobacteria.gif")
```

Facet_wrap Plots

Set the raster stack as a data frame and use the melt() function to gather all values into one column

```
library(reshape)
```

```r
library(tidyverse)

rasstack_2017_df <- as.data.frame(crop_stack, xy = TRUE) %>%
  melt(id.vars = c('x', 'y')) %>%
  mutate(Date = str_sub(variable, 2, 9))

rasstack_2017_df$Date <- as.Date(rasstack_2017_df$Date, format =  "%Y%m%d")

#plot all dates/facet
ggplot() +

  coord_sf(crs = 3160) +
  geom_raster(data = rasstack_2017_df,
              aes(x = x, y = y, fill = value)) +
   facet_wrap(~Date, ncol = 4, scales = "fixed") +
  scale_fill_viridis_c(na.value = "transparent",
                       name = "CI (cells/mL)",
                       limits = c(0, 1250000)) +
  theme_bw() +
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank(),
        axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank())
```

```r
ggsave("CI_2017.png", last_plot(), dpi=600, width = 7, height = 10,
units="in")
```

Load in buoy data

```r
library(dplyr)
library(ggplot2)
library(scales)

setwd("D:/Greene/CI_codes/data")

df <- read.csv("ML Phyco and Chlor.csv")

names (df) <- c("Datetime", "Chla", "Phyco")

df <- df %>%
  replace(.=="-Suspect-", NA) %>%
  replace(.=="", NA)

df$Datetime <- as.Date(df$Datetime, format =  "%m/%d/%Y %H:%M")

df$Chla <- as.numeric(df$Chla)
```

```r
df$Phyco <- as.numeric(df$Phyco)

df <- df %>%
  group_by(Datetime) %>%
  summarise(mean_chla = mean(Chla, na.rm = TRUE),mean_phyco = mean(Phyco,
na.rm = TRUE)) %>%
  mutate(across(2:3, round, 2))
```

Extract data at buoy

```r
pointCoordinates = read.csv("D:/Greene/ML Data/MLBuoyCoordinates.csv")

coordinates(pointCoordinates) <- ~Latitude + Longitude

rasValue = raster::extract(rasStack, pointCoordinates)
#print(rasValue)

mean_mlbuoy_value <- raster::extract(x = rasStack,
                        y = pointCoordinates,
                        buffer = 300,
                        fun = mean, na.rm = TRUE)
mean_buoy_value <- t(as.data.frame(mean_mlbuoy_value))
mean_buoy_value <- as.data.frame(cbind(mean_buoy_value, date))

CI_buoy <- read.csv("D:/Greene/CI_codes/data/Buoy_v_CI.csv")

CI_buoy$Date <- as.Date(CI_buoy$Date, format =  "%m/%d/%Y")

CI_buoy$Year <- as.factor(format(CI_buoy$Date, "%Y"))

CI_buoy <- CI_buoy[-c(13),]
```

Plot the buoy data vs Sentinel-3

```r
library(ggplot2)

ggplot(data = CI_buoy, aes(x=Buoy, y=CI, group=Year)) +
  geom_point(aes(shape=Year, color=Year), size =4) +
  scale_color_manual(values=c('#d95f02','#1b9e77', '#7b3294'))+
  labs(title="CI vs In Situ Data",
       x="In Situ Phycocyanin (cells/mL)",
       y="Cyanobacteria Index (cells/mL)") +
    scale_y_continuous(labels = comma) +
  theme_bw() +
  theme(text = element_text(size = 14)) +
  theme(plot.title = element_text(hjust = 0.5))

ggsave("buoy_CI.png", last_plot(), dpi=600, width = 7, height = 5,
units="in")
```

B.5 R Studio code for analyzing data and creating graphics