**NASA DEVELOP National Program**



Langley Research Center

*Summer 2013*

New England Water Resources

Multispectral Monitoring of New England Freshwater Resources to Assess Turbidity, Algal Blooms, and Water Quality for Enhanced Natural Resource Management

**Technical Report Final Draft**

August 5, 2013

Tiffani Orne, Liberty University (Project Lead)

Hayley Solak, Clark University

Sam Weber, Virginia Tech

Dr. Kenton Ross, DEVELOP National Program (Science Advisor)

# Abstract

Centered between New York, Vermont, and Quebec, Lake Champlain is a critical water resource for the surrounding area. Approximately 145,000 people rely on the lake for drinking water and it is a major stopping point and breeding ground for migrating birds. Development in the Lake Champlain watershed has led to an increase in nutrients in the lake. Algae in the water thrive on the nutrient flux and reproduce exponentially, causing hazards to both human and environmental health. Interested organizations, including the Lake Champlain Basin Program (LCBP), the Lake Champlain Committee (LCC), and the Vermont Department of Environmental Conservation (VTDEC), mobilize citizen volunteers to collect water samples in various parts of the lake in order to monitor water quality. However, this process requires a large number of volunteers, can be inconsistent, does not account for the quality of the entire lake, and requires the cost of laboratories to test the water samples. Using data from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor aboard the Aqua satellite and from the Enhanced Thematic Mapper Plus (ETM+) and Operational Land Imager (OLI) sensors aboard the Landsat 7 and 8 satellites, this project created a series of maps showing the change in chlorophyll-a, cyanobacteria, phycocyanin, and total suspended sediment (TSS) over time. The methodologies were then transferred to project partners for continued use of remote sensing to monitor water quality with maps were provided as a visual tool to influence public policy.

**Keywords**

Remote Sensing, Water Quality, Algal Blooms, Suspended Sediment, Chlorophyll-a, Cyanobacteria, Phycocyanin, Lake Champlain

# Introduction

# 

# Background Information

# Community Concerns

# During Harmful Algal Blooms (HABs), toxins bioaccumulate in organisms at high trophic levels, such as mussels and fish, which are commonly consumed by humans (Backer & McGillicuddy Jr, 2006). Although the aquatic organisms may or may not show traits of toxic exposure, consumption by humans can result in a wide range of illnesses, including various types of shellfish poisoning. Once toxins in the water build up, water movement can also release those toxins into the air. There are no known remedies for the illnesses that are caused by HABs and must naturally run their course.

# Harmful Algal Blooms also have negative impacts on ecological health. The toxins affect wildlife other than just fish, and pollution sensitive organisms are unable to survive (Glibert, et al., 2005). This affects diversity and ecosystem food webs. Underwater vegetation is starved of sunlight and oxygen. Algae overpower native species, and provide nutrients for invasive species.

# The effects of HABs are particularly great in Lake Champlain because it is a major stopping point and breeding ground for migrating birds. Of the more than 300 species of birds found in the Champlain Basin, 24 are listed as endangered species (Lake Champlain Basin Program [LCBP], 2013b). Lake Champlain is a critical water resource for the surrounding area. Approximately 145,000 people rely on the lake for drinking water (LCBP, 2013a), but there are no sanitizing methods that can remove the HAB toxins from the water (Backer & McGillicuddy Jr, 2006). The local economy, tourism, and recreation are also negatively impacted by millions of dollars a year in losses due to illnesses and beach closures (Glibert et al., 2005).

# Previous Studies

# Previous studies have focused primarily on the detection of HABs through estimation of two optically-active accessory pigments in freshwater cyanobacteria, phycocyanin and chlorophyll-a, both of which have unique spectral properties and can be extracted through application of satellite bio-optical algorithms. Remote sensing techniques have become increasingly accessible in the past two decades for studies concerning biological processes at synoptic scales using sensors such as Advanced Very High Resolution Radiometer (AVHRR), Sea-viewing Wide Field-of-view Sensor (SeaWiFS), and Moderate Resolution Imaging Spectroradiometer (MODIS). The majority of these sensors have products available for download (in particular SEAWIFS) which specifically show chlorophyll-a concentrations. However, studies conducted with these coarser spatial resolution sensors, whose finest spatial resolution reaches 250 meters per pixel, may be inappropriate for the assessment of biophysical properties within comparatively smaller inland water bodies such as Lake Champlain.

# In cases where high-resolution imagery is not available, due to inaccessibility or high cost, researchers have successfully developed remote sensing models using freely downloadable moderate-resolution Landsat 7 data to retrieve biophysical information related to HABs. These models often utilize simple band ratios such as (for Landsat 7) band 1 (blue) and band 3 (red) to estimate chlorophyll-a concentrations, and bands 1 (blue), 3 (red), 4 (Near Infrared), and 7 (Shortwave Infrared) to estimate phycocyanin levels. Water turbidity analyses and estimation of Total Suspended Sediment (TSS) have also been useful in examining concentrations of chlorophyll and phycocyanin relative to total water turbidity and the total amount of suspended sediment as cyanobacteria blooms create more turbid water and are often a product of eutrophication due to runoff from nearby agricultural operations which introduce chemically rich sediments into neighboring water bodies.

# Project Objectives

# The purpose of this project was to create maps showing the concentrations of chlorophyll-a, cyanobacteria, phycocyanin, and suspended sediment throughout Lake Champlain in order to demonstrate that water quality can be accurately assessed using NASA Earth observations. Maps of estimated TSS were also created to show that algal blooms can be detected as sediment on a more regular basis. Maps for the same day were compared to show correlation, and maps of the same type were compiled in a time series to show change throughout the study period.

# Study Area

# Lake Champlain is located on the border between the states of New York and Vermont, and some of the northern part of the lake also branches into Quebec, Canada. The lake is 120 miles long, 12 miles wide at its’ widest point, and 400 feet deep at its’ deepest (LCBP, 2013c). The area of Lake Champlain is 435 square miles, and it is subdivided into five segments: South Lake, Main Lake, Malletts Bay, Inland Sea, and Missisquoi Bay.

# Study Period

# Landsat 7 Enhanced Thematic Mapper Plus (ETM+), Landsat 8 Operational Land Imager (OLI), and Aqua MODIS data were downloaded for July 2002 through June 2013. July 2002 was selected as the starting point, because it is when the earliest Aqua data is available. Algal blooms typically occur during the summer, so only data from May through October of each year were used.

# National Application Areas

# This project addresses the NASA national application area of water resources. The goal of the water resources application area is to aid decision makers in areas related to water quality and availability. This project uses NASA Earth observations to monitor water quality in New England, specifically the presence of Harmful Algal Blooms. The project also addresses the issue of availability as the study area provides drinking water to approximately 145,000 people, which is unavailable when water quality is particularly low.

# Project Partners

# The partners for this project were the Lake Champlain Basin Program (LCBP), the Lake Champlain Committee (LCC), and the Vermont Department of Environmental Conservation: Water Management Division. All three organizations also collaborate with each other. Currently, they all primarily use employees and/or volunteers to monitor water quality in Lake Champlain. This project will benefit project partners by providing them with methodologies that can be duplicated and maps that can be used to influence policy and track how policy changes have affected water quality.

# Methodology

**Data Acquisition**

Landsat 7 and Landsat 8 level 1 products were acquired to create chlorophyll-a, cyanobacteria, and phycocyanin maps. Aqua MODIS data were used to create TSS maps. Aqua data acquired were level 2 MYD09GA and MYD09GQ daily surface reflectance products. All data were downloaded from the USGS Global Visualization Viewer (GloVIS) website. Dates were selected manually by clicking through each image and downloading every available day that did not have significant cloud cover over the area. The Landsat data tile used was located at Path 14, Row 29 and the Aqua tile was located at Horizontal 12, Vertical 4. A total of 45 Landsat images (43 from Landsat 7 and 2 from Landsat 8) and 408 Aqua images were acquired.

**Data Processing and Analysis**

Chlorophyll and Cyanobacteria

Landsat 7 and Landsat 8 data were processed using ArcMap 10.1. Chlorophyll concentration and cyanobacteria concentration were calculated using equations from Adam Trescott’s master’s thesis titled “Remote Sensing Models of Algal Blooms and Cyanobacteria in Lake Champlain.” These algorithms utilized Landsat 7 bands 1, 2 and 3 to produce band 2 divided by band 1 and band 3 divided by band 1 ratios. The algorithms used by the model were:

Chlorophyll-a (μg/L) = -46.51 + 105.30(RB2/RB1) – 40.39(RB3/RB1)

Cyanobacteria (mm3/L) = -0.6858 + 1.616(RB2/RB1) – 0.8025(RB3/RB1)

The equation used for phycocyanin concentration was found in a journal article titled “Phycocyanin Detection from Landsat TM Data for Mapping Cyanobacterial Blooms in Lake Erie” (Vincent et al., 2004). This algorithm used Landsat 7 bands 1, 3, 5, and 7. The algorithm used was:

Phycocyanin (μg/L) = 0.78 – 0.0539(B1) + 0.176(B3) – 0.216(B5) + 0.117(B7)

First, a mask was applied to the Landsat 7 images to isolate Lake Champlain using a shapefile of the lake and the “Extract by Mask” tool in the Spatial Analyst Extraction toolbox. ArcMap Model Builder was used to iterate rasters and automate the masking process. Once the lake was extracted, the digital numbers of the raw Landsat imagery had to be converted to top of the atmosphere (TOA) reflectance values before the algorithms could run. A script written by DEVELOP Geoscience Lead Quinten Geddes was applied for this purpose. The data were then run through the algorithms employing the “Raster Calculator” tool in the Spatial Analyst Map Algebra toolbox. The band ratios were calculated first, then plugged into the chlorophyll-a and cyanobacteria equations. This step was not necessary for phycocyanin calculations because no band ratios are used in the equation.

Except for the phycocyanin concentration, the same methodology was used for the two Landsat 8 images. The chlorophyll-a and cyanobacteria equations used band ratios that could be computed with the comparable Landsat 8 bands. The phycocyanin equation used digital numbers, which are different between Landsat 7 and Landsat 8. Landsat 8 digital numbers were divided by 256 in order to work with the Landsat 7 equation.

Once the chlorophyll-a, cyanobacteria, and phycocyanin values had been calculated, a cloud mask was applied to the newly created maps. Using the “Set Null” tool in the Spatial Analyst Conditional toolbox, value ranges of 28.12 to 28.14 and 0.26 to 0.28 were used to set “NoData” values to mask clouds pixels in the chlorophyll-a and cyanobacteria maps respectively. These thresholds were established based on concise values the algorithms produced for cloud pixellated areas. Clouds in the chlorophyll maps had values of 28.135769 and cyanobacteria maps had values of 0.274720. A range was given to help mask cloud edges. The cloud masking process was also automated using raster iterations in ArcMap Model Builder. Slight inaccuracies in the Landsat sensor created negative numbers, so an iterating “Con” tool in the Spatial Analyst Conditionals toolbox was used to turn any negative numbers into zeroes.

Since the Landsat 7 data after 2003 did not have scan lines corrected, the gaps created erroneous values when the algorithms were calculated. To compensate, another “Extract by Mask” raster iterator was applied to remove values above 150 for the chlorophyll-a maps and above 2 for the cyanobacteria maps. For the phycocyanin maps, values above 6 were removed. These values cleared invalid data without tampering legitimate observations. Finally, an iterative “Filter” tool in the Spatial Analyst Neighborhood toolbox was used to smooth out the final images. This was run on the low filter type setting. Chlorophyll-a, cyanobacteria, and phycocyanin concentration maps were assigned the same color ramps for consistency between the three parameters. Chlorophyll-a map legends were assigned values from 0 to 30, cyanobacteria map legends were assigned values from 0 to 0.35, and phycocyanin map legends were assigned values from 0 to 6 based on upper limit results found by the original sources used to obtain the algorithms.

Total Suspended Sediment

Aqua MODIS data were also processed using ArcGIS and Python Programming. The basis for the TSS methodology was an ArcGIS model developed during a previous project by DEVELOP science advisor Dr. Kenton Ross and DEVELOP intern Ande Ehlen. For each date, the model utilized the 500m statefile, the 250m surface reflectance band 1, and the 250m near infrared reflectance (NIR) compilation of bands 1 and 2. The model removes cloud cover and land, calculates suspended sediment, and clips the image to the study area. The algorithm used by the model to calculate TSS was:

TSS=602.63\*(0.5157\*((B1)\*0.0001)-0.0089)+3.1481

Previous use of the model was designed to process one MODIS scene at a time without iteration. Due to the volume of data being processed and the limitations of ArcGIS model builder, the original model was exported to a Python script where it was re-designed to iterate the processes for each of the 408 MODIS scenes collected from the GLOVIS website.

The first step in the process is the extraction of the necessary files: the statefile 500m band from the MYD09GA datafile. The 250m surface reflectance band 1 from the MYD09GQ datafile and the 250m surface reflectance bands 2 and then 1. All files were converted from HDF4 to TIFF format and fit to the extent of the study area using a shapefile of Lake Champlain. The TSS algorithm was then applied to the extracted 250m surface reflectance band 1 and cloud masks were created using the 500m statefile, producing a final TSS image that is cloud free and does not include any unnecessary land, such as coastline.

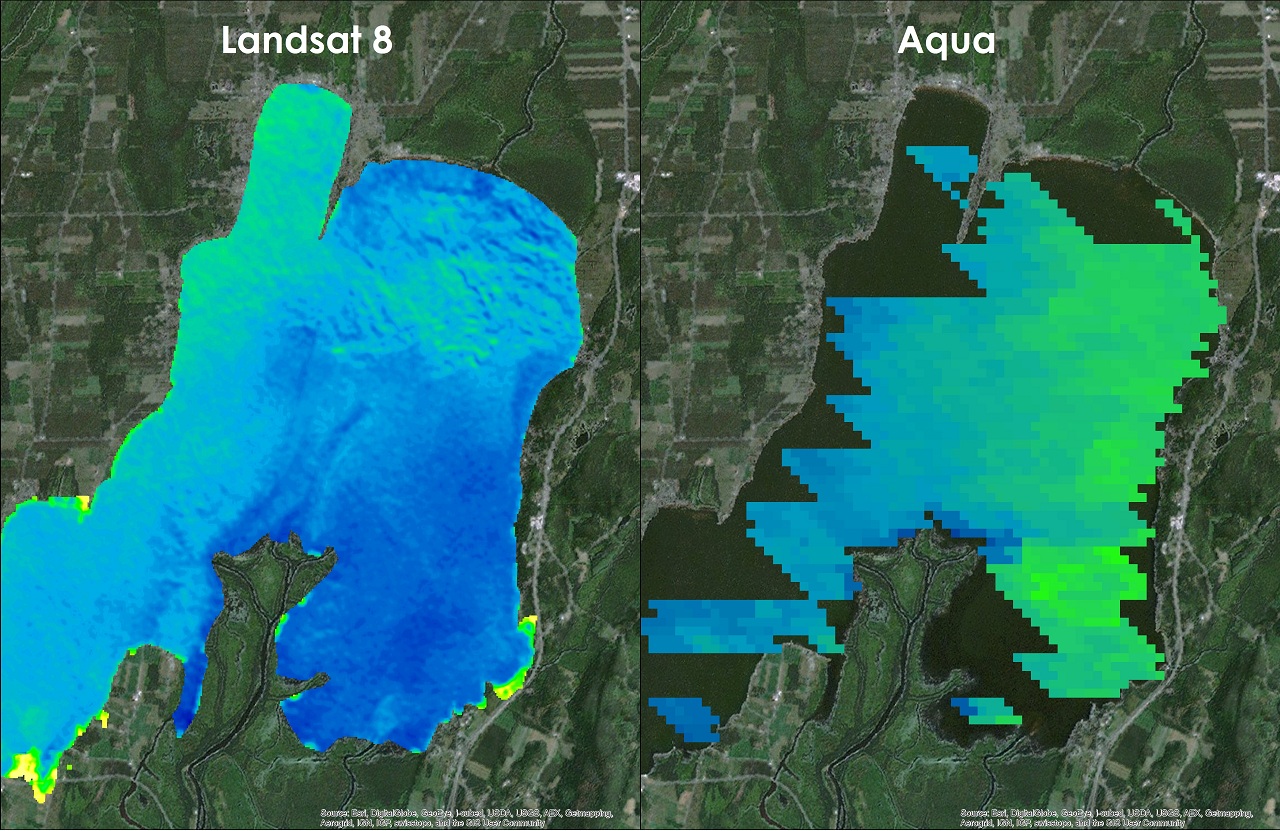
After running the model, the color scheme for each map was changed to go from blue to green so that clear water was blue and sediment was green. The values were scaled from zero to sixteen, creating the same relative scale for sediment across all maps.

# Results & Discussion

**Analysis of Results**

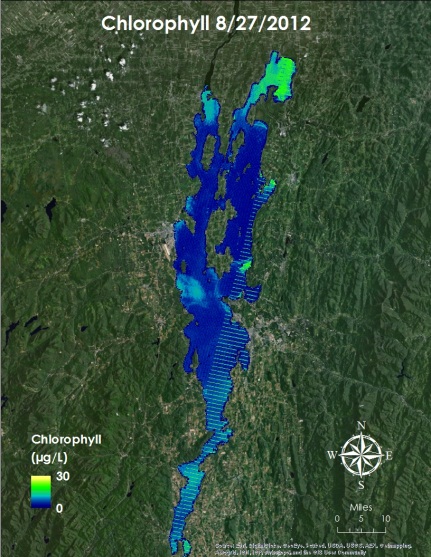
Concentration Maps

Concentration maps for each parameter were created in ArcGIS. Maps for biophysical parameters were much clearer than the TSS maps due to the finer resolution when using Landsat 7 ETM+ and Landsat 8 OLI when compared to Aqua MODIS (Figure 1). Calculated values for all four parameters consistently showed higher levels in Missisquoi Bay than in other regions of the lake, which was expected based on literature review (Figure 2).

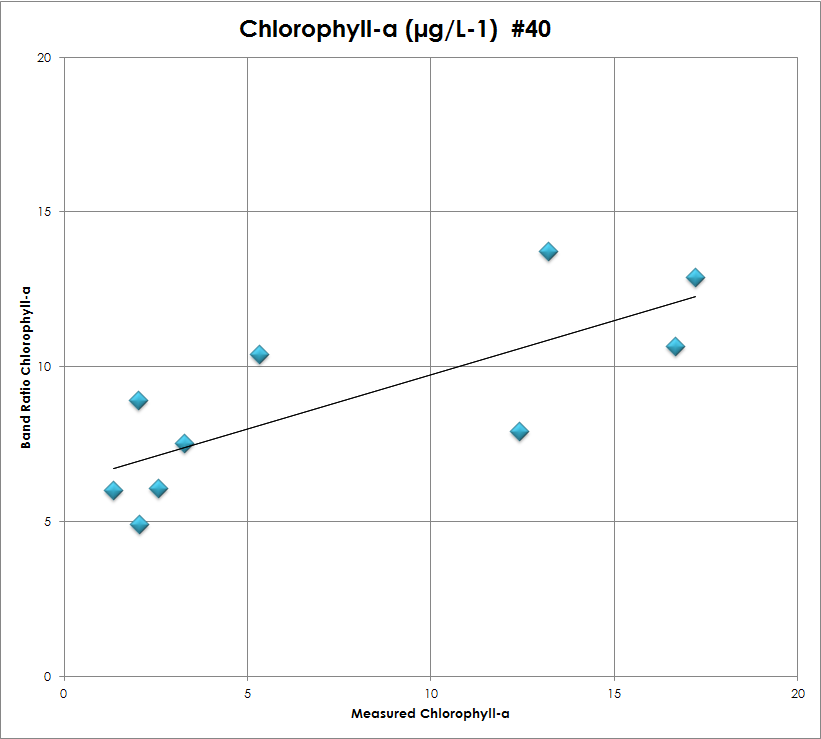


*Figure 1. Map of Missisquoi Bay in the north of Lake Champlain, showing the difference in spatial resolution between Landsat 8 OLI and Aqua MODIS data.*

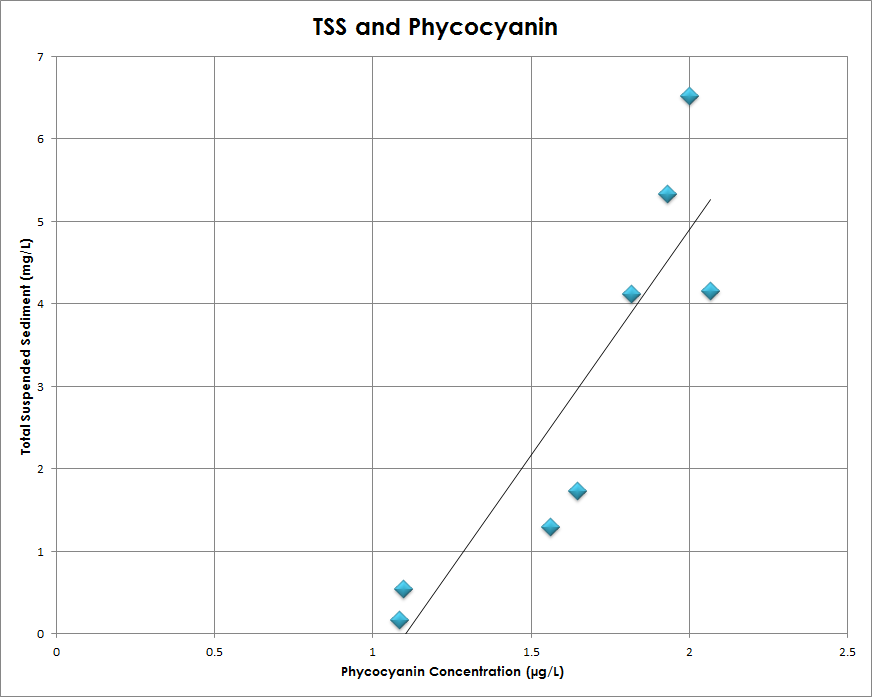
*.*



*Figure 2. Chlorophyll-a concentration map for Lake Champlain on August 27, 2012. This map was created using data obtained by Landsat 7 ETM+ and an algorithm in ArcGIS. This image shows the main concentration of chlorophyll in Missisquoi Bay, and a few smaller locations throughout the lake.*

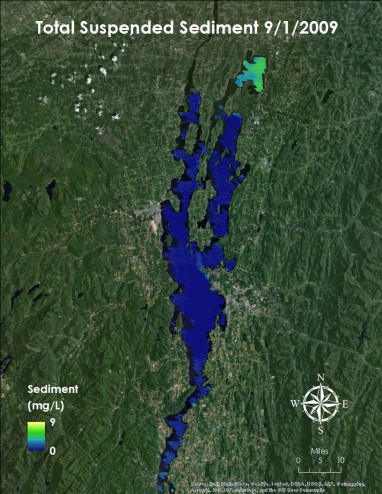
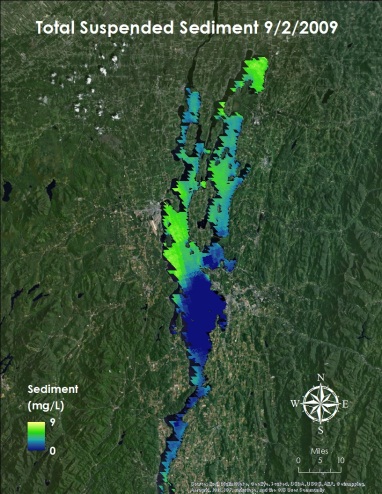
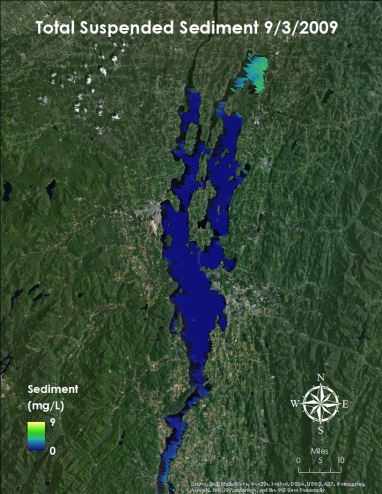
 Statistical analyses showed that calculated values were positively correlated with in situ measurements, so our algorithms were accurate in calculating the amount of each parameter that was measured to be present in the lake (Figure 3). This allowed us to apply the same algorithms to dates for which there were no in situ measurements to compare. TSS concentrations were also correlated with biophysical parameters, which shows that TSS can be used to accurately identify algae within the lake (Figure 4). This is important because, while Landsat imagery has a finer spatial resolution, Aqua has a finer temporal resolution, allowing end users to monitor algae on a daily basis.

*Figure 3. Scatterplot showing the correlation between calculated values of Chlorophyll and in situ measurements (N=10, r=0.76715).*



*Figure 4 Scatterplot showing the correlation between calculated values of total suspended sediment and phycocyanin (N=8, r=0.89144).*

Time Series

 Time series videos were created to show change in each parameter over time. The time series for TSS provided an excellent illustration of the change that can occur in a single day (Figure 5). On September 1, sediment is concentrated in Missisquoi Bay. The next day, there is a high concentration throughout the lake, and on the third, the main concentration is again limited to the north. There is also a noticeable increase in the frequency of algal blooms over time.

*Figure 5. Time series showing the change in TSS from September 1, 2009 through September 3, 2009.*

Phosphorus

One additional analysis compared Phosphorus levels in the lake over time with biophysical parameters to assess the time lag between phosphorus runoff and the occurrence of algal blooms. Plots showed a pattern in peak times, but the pattern was difficult to quantify due to a lack of data (Figure 6).

*Figure 6. Chart showing lag time and concurrency between in situ phosphorus measurements and calculated biophysical parameter values.*

**Errors & Uncertainty**

Earth science research is inherently complicated as it encompasses a broad spectrum of disciplines, all of which interact and complicate simple questions. Some environmental factors that could potentially skew observed values include the difficult task of masking high altitude cirrus clouds, cloud edges, and shadows, climactic anomalies, lake wave activity, and land usage change over time.

More technical potential issues include difficulty establishing an initial Lake Champlain shapefile to accurately depict land/water boundaries for masking purposes, missing data due to the inactive Landsat 7 scan line corrector, translation of Landsat 8 to Landsat 7 digital numbers, and scattered data pixels for cloudy Aqua MODIS days. Validating the remotely sensed data was also a slight concern as in situ measurement dates often did not correspond with Landsat/Aqua acquisition dates. Having more overlapping dates would have strengthened statistical significance and been more symbolic of the true rapport between the two observing methods.

**Future Work**

The nature of frequent algal bloom events in inland water bodies is often related to an increase in anthropogenic influence, such as urban development and agriculture leading to eutrophication. Additional analysis of the relationship between such activities and the occurrence of algal blooms would be beneficial to organizations involved with Lake Champlain for the purpose of influencing public policy related to human development around the Lake.

Temperature increases in regions surrounding the Lake also have an impact on the occurrence of algal blooms, which tend to be more frequent during warmer months when water temperature is higher and there is increased light (New South Wales Department of Primary Industries, 2009). The time series format of this project makes it possible to transition from observing biophysical processes within the water such as chlorophyll-a, phycocyanin and cyanobacteria, to looking at temperature trends in the water as well as over surrounding land, and determining if a relationship exists with the trends in the biophysical data.

# Conclusions

After performing visual and statistical comparisons between recorded harmful algal blooms and measured in situ concentration levels of chlorophyll-a, cyanobacteria, phycocyanin, and total suspended sediment, NASA Earth observations were shown to accurately monitor biophysical parameters associated with algal bloom outbreaks. Not only do remotely sensed values correlate well with in situ values, but biophysical parameters showed relation to one another as well. A 0.89 R-value was recorded between phycocyanin and total suspended sediment, indicating a strong linear relationship between the two. Just as it is important to understand the significance of remotely sensed data compared with ground truth, it is equally important to recognize the effects one biophysical parameter may have on another.

Unlike localized in situ measurements, remote sensing allows end user to observe water quality throughout the entire lake. In situ measurements are only representative of one specific location, while remote sensing allows a simultaneous view of the entire body of water. Water quality can change sharply over short distances, so having coverage of an entire body of water is advantageous. Water quality can also vary drastically on small time scales. The Vermont Department of Environmental Conservation samples the lake approximately ten times per year from May to October, with a majority of the samples coming once a bloom has been detected. This leaves plenty of time between measurements for harmful algal blooms to peak and wane without ever being noticed.

High resolution Landsat data are only available every 15 days, and cloud cover can easily mask desired study areas. More frequent data available from Aqua MODIS allows greater temporal resolution, but at a much coarser resolution. Compared to in situ measurements, remote sensing has the ability to observe potential bloom conditions at a lower cost to end users. Ground measurements require a larger workforce to take water samples, and consistent funding to conduct lab testing for biophysical parameters. Landsat and Aqua MODIS data are available for free from the USGS through their GLOVIS website. Both observing methods have advantages and disadvantages, but if used together they can create a valuable water monitoring tool. Continued in situ measurements would help improve and polish existing algorithms while remote sensing provides additional data between ground measurements.

# Acknowledgments

Dr. Kenton Ross (NASA DEVELOP National Science Advisor)

Quinten Geddes (NASA DEVELOP Center Lead, Goddard Space Flight Center)

Bill Howland (Lake Champlain Basin Program)

Lori Fisher (Lake Champlain Committee)

Mike Winslow (Lake Champlain Committee, Staff Scientist)

Angela Shambaugh (Vermont Department of Environmental Conservation)

# References

Backer, L. C. & McGillicuddy Jr, D. J. (2006). Harmful algal blooms at the interface between coastal oceanography and human health. *Oceanography, 19*(2), 94-106. doi:http://dx.doi.org/10.5670/oceanog.2006.72

Glibert, P. M., Anderson, D. M., Gentien, P., Graneli, E., & Sellner, K. G. (2005). The global, complex phenomena of harmful algal blooms. *Oceanography, 18*(2), 136-147. doi:http://dx.doi.org/10.5670/oceanog/2005.49

Lake Champlain Basin Program (2013a). Drinking water. Retrieved from http://www.lcbp.org/water-environment/human-health/drinking-water/

Lake Champlain Basin Program (2013b). Fish and wildlife. Retreived from http://www.lcbp.org/water-environment/ecosystem-healt/fish-and-wildlife/

Lake Champlain Basin Program (2013c). Lake and basin facts. Retrieved from http://www.lcbp.org/about-the-basin/facts/

New South Wales Department of Primary Industries (2009). What causes algal blooms. Retrieved from http://www.water.nsw.gov.au/Water-Management/Water-quality/Algal-information/What-causes-algal-blooms/

Trescott, A. (2012). *Remote sensing models of algal blooms and cyanobacteria in Lake Champlain.* (Master’s Thesis). Retrieved from scholarworks.umass.edu/cee\_ewre/ 48/

Vincent, R. K., Qin, X., McKay, R. M. L., Miner, J., Czajkowski, K., Savino, J. & Bridgeman, T. (2004). Phycocyanin detection from Landsat TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment, 89*(3), 381-392. doi:http://dx.doi.org/10.1016/j.rse.2003.