Monitoring Harmful Algal Blooms in Lake Champlain with MODIS and Landsat 8 OLI Remote Sensing Data

Maya Midzik

Advisor: Ron Smith
Second Reader: Pincelli Hull
May 4th, 2016

A Senior Thesis presented to the faculty of the Department of Geology and Geophysics, Yale University, in partial fulfillment of the Bachelor's Degree

CONTENTS

ABSTRACT

Cyanobacteria are a healthy part of many freshwater and marine ecosystems. However, under suitable conditions, these blue-green algae can increase to extreme levels, resulting in visible "blooms" and deteriorating water quality. Furthermore, some of these blooms produce toxins that can kill fish, mammals, and birds, and may even lead to human illness and death. Since 1999, these harmful algal blooms (HABs) have plagued Lake Champlain, the sixth largest freshwater lake in the United States, and a source of drinking water for many communities in the surrounding states of New York and Vermont. However, current monitoring programs rely on infrequent visual identification and limited *in situ* water sampling to initially alert officials to toxic events. This paper presents remote sensing tools for cyanobacteria identification using Landsat 8 OLI and Moderate-Resolution Imaging Spectroradiometer (MODIS) imagery, examining the potential for high spatial and temporal HAB monitoring in Lake Champlain's Missisquoi and St. Albans Bay, two areas highly affected by bloom events. Atmospheric correction algorithms were tested for each sensor and statistical models developed based on three years of *in situ* monitoring data from the Lake Champlain Long-Term Water Quality and Biological Monitoring Program (LTMP). Results for both sensors demonstrate the potential of chlorophyll's green (550 nm) reflectance maximum for HAB identification in shallow and protected areas of Lake Champlain.

SECTION 1. INTRODUCTION

1.1 Harmful Algal Blooms and Lake Champlain Monitoring Protocols

Cyanobacteria, or blue-green algae, are a type of microscopic, algae-like bacteria. These algae are part of many healthy freshwater, coastal and marine ecosystems. However, under suitable conditions cyanobacteria can swell to extreme levels, forming visible "blooms" and deteriorating water quality [Figure 1]. Furthermore, some of these blooms produce toxins that can kill fish, mammals, and birds and may even lead to human illness and death. These harmful algal blooms (HABs) have become increasingly common occurrences worldwide and are of particular concern for inland freshwaters, where concentrations of neurotoxins can quickly multiply, impacting local drinking water supplies, fishery health, and recreation (Paerl, 2011).

Harmful algal blooms were first recorded in Lake Champlain in 1999, after two dogs died from consuming water containing high concentrations of cyanotoxins (Rosen, 2004). Since 2000, HABs have become a major safety concern within the Lake Champlain basin, as studies have demonstrated the dominance of toxin-producing taxa in the Northeastern Arm of the Lake and levels of microcystin and anatoxin-a (two potent cyanotoxins) have regularly exceeded the World Health Organization (WHO) advisory levels of 1μ/L (Rosen, 2001; Rosen, 2004; Boyer, 2004).

Since 2002, HABs in Lake Champlain have been monitored through a tiered alert system based on WHO recommendations [Table 1]. This system relies on volunteer-based visual analyses and infrequent water sampling (about two times a month) for initial alerts. These observations are then reported and displayed through the Vermont Department of Health's 'Algae Tracking Map.'[1] Only once scum, highly discolored water, foul odor, or other potential HAB indicators are observed in these qualitative analyses does the alert system move to a quantitative stage, at which vertical plankton tows are collected for microscopic analysis by experts. If both high cyanobacteria densities and microcystin concentrations exceeding 1μ/L are measured in these quantitative samples, Alert Level 2 is reached, and public health advisories are issued by local agencies (Watzin, 2006).

---

[1] Algae tracking map can be viewed at https://apps.health.vermont.gov/gis/vttracking/BlueGreenAlgae

The efficacy of this alert system is severely limited in 1) its ability to respond rapidly to growing HABs, as initial alerts are dependent upon infrequent and often inaccurate visual observations, and 2) the spatial coverage of monitoring, with *in situ* samples and visual observations only covering a small portion of the 269 square km Northeastern Arm of Lake Champlain (LCBP, 2004). Remote sensing technology offers a solution to both of these limitations, as satellite remote sensors provide regular and synoptic observations of the Lake Champlain Basin. Although airborne and satellite remote sensing cannot replace in situ toxin identification and cell counts, a quantitative remote sensing model for algal blooms in Lake Champlain has the potential to greatly increase the accuracy and frequency of initial HAB alerts, providing an inexpensive and effective addition to existing monitoring protocols.

Figure 1: Blue-green algae blooms at the Route 78 access in Mississquoi Bay and boat-based sampling in Lake Champlain, VT (VT DEC).

Table 1: Outline of current monitoring and alert system for Lake Champlain, based on WHO guidelines (from Watzin, 2006)

Alert Level	Frequency	Samples Collected for	Trigger to Next Level	Public Action
Initial	2/month	Algal identification	Identification of toxin-producing cyanobacteria	None
Quantitative	2/month	Algal enumeration; Chlorophyll a	>2000 cyanobacteria cells/mLa in net samples or lay monitor samples	None
Vigilance	1/week	Algal enumeration; Chlorophyll a	>4000 cyanobacteria cells/mLb in net samples or lay monitor samples	Notify public health officials that cyanobacteria are abundant and blooms could form.
Alert Level 1	1/week	Algal enumeration; Chlorophyll a; Toxin analysis	>1 "g microcystin/L in whole water samples	Notify public health officials of potential risks to humans and animals.
Alert Level 2	1/week	Algal enumeration; Chlorophyll a; Toxin analysis		Notify public health officials that significant risk to humans and animals exists. Public health advisories should be issued by appropriate agencies.

1.2 Remote Sensing of Algal Blooms: Past Research and Applications

1.2a Introduction

Remote sensing of phytoplankton abundance has historically been termed 'ocean color,' as such methods were most primarily developed for synoptic measurement of phytoplankton biomass in the world's oceans. Since the development of the Coastal Zone Color Scanner Experiment (CZCS) in 1978, multiple satellite missions have focused on the measurement of marine phytoplankton biomass. The data from these missions, including the Sea-Viewing Wide Field-of-View Sensor (SeaWIFS), Hyperspectral Imager for the Coastal Ocean (HICO), and Moderate Resolution Imaging Spectroradiometer (MODIS) missions, are then collected, processed, calibrated, archived, and distributed by NASA's Ocean Biology Processing Group (OBPG). Table 2 outlines the specifications of past and currently operational satellites for phytoplankton detection and monitoring. Despite the many ocean color missions of the past three decades, only Landsat 7 and 8 and two MODIS missions remain operational for current phytoplankton monitoring (NASA Ocean Color Data, n.d.).

These missions have been successful at detecting large-scale ocean phytoplankton dynamics. However, many of these 'ocean-color missions' lack the necessary resolutions for coastal or inland lake assessments, as remote sensor design always represents trade-offs between spatial resolution, spectral wavelengths, and temporal frequency (Gordon, 2012). Furthermore, the remote sensing of such turbid, high-sediment, and small-scale water bodies provides unique complications and considerations. Therefore, although many studies have attempted to apply ocean color remote sensing models to inland water assessment, such models must be developed independently for each location and sensor due variations in water quality parameters and atmospheric correction approaches (Lathrop, 1992; Dekker, 1993; Kloiber, 2002; Gons, 2002; Cipman, 2004; Simis, 2007; Becker, 2009; McCullough, 2012; Torbick 2013).

Table 2: Specifications of past and currently operational satellites for phytoplankton detection. Information from NASA Ocean Color Data Website, further individual mission details from NASA satellite websites.

Satellite sensor			Launch Date	End Date	Resolution			
					Spectral Range/Bands	Radiometric	Spatial	Temporal
Coastal Zone Color Scanner Experiment (CZCS)			Oct 1978	Jun 1986	5 Multispectral bands: 1) 433-453 nm; 2) 510-530 nm; 3) 540-560 nm; 4) 660-680 nm; 5) 700-800 nm 1 Thermal band: 10,500-12,500 nm	8-bit	825 m for all bands	26 days
Landsat	1	Multispectral Scanner (MSS)	Jul 1972	Apr 1999	4 Multispectral bands: 1) 500-600 nm; 2) 600-700 nm; 3) 700-800 nm; 4) 800-1,100 nm	6-bit	68 x 83 m	18 days
	2	Multispectral Scanner (MSS)	Jan 1975	Feb 1982		6-bit	68 x 83 m	18 days
	3	Multispectral Scanner (MSS)	Mar 1978	Mar 1983	4 Multispectral bands: 1) 500-600 nm; 2) 600-700 nm; 3) 700-800 nm; 4) 800-1,100 nm. 1 Thermal band: 10,410-12,350 nm	6-bit	68 x 83 m	18 days
	4	Thematic Mapper (TM)	Jul 1982	Jun 2001	6 Multispectral bands: 1) 450-520 nm; 2) 520-600 nm; 3) 630-690 nm; 4) 760-900 nm; 5) 1,550-1,750 nm; 6) 2,080-2,350 nm 1 Thermal band: 10,400-12,500 nm	8-bit	30 m for multispectral; 120 m for thermal	16 days
	5	Thematic Mapper (TM)	Mar 1984	Jan 2013		8-bit	30 m for multispectral; 120 m thermal	16 days
	7	Enhanced Thematic Mapper Plus (ETM+)	Apr 1999	**Operational**	6 Multispectral bands: 1) 450-515 nm; 2) 525-605 nm; 3) 630-690 nm; 4) 750-900 nm; 5) 1,550-1,750 nm; 6) 2,090-2,350 nm 1 Thermal band: 10,400-12,500 nm 1 Panchromatic Band: 520-900 nm	8-bit	30 m for multispectral; 120 m for thermal; 15 m for panchromatic	16 days
	8	Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS)	Feb 2013	**Operational**	7 Multispectral bands: 1) 430-450 nm; 2) 450-510 nm; 3) 530-590 nm; 4) 640-670 nm; 5) 850-880 nm; 6) 1,570-1,650 nm ; 7) 2,110-2,290 nm 2 Thermal bands: 10,600-11,190nm and 11,500-12,510 nm 1 Panchromatic Band: 500-680 nm 1 Cirrus band: 1,360-1,380 nm	12-bit	30 m for multispectral; 120 m for thermal; 15 m for panchromatic	16 days
SeaWiFS			Aug 1997	Dec 2010	8 Multispectral bands: 1) 402-422 nm; 2) 433-453 nm; 3) 480-500 nm; 4) 500-520 nm; 5) 545-565 nm; 6) 660-680 nm; 7) 745-785 nm; 8) 845-885 nm	12-bit	1.1 km for all bands	1 - 2 days
MODIS	Terra		Dec 1999	**Operational**	36 Multispectral bands: 1) 620-670 nm; 2) 841-876 nm; 3) 459-479 nm; 4) 545-565 nm;	12-bit	250 m for bands 1 & 2; 500 m for bands 3 - 7; 1000 m for bands 8 - 36	1-2 days
	Aqua		Mar 2002	**Operational**	5) 1,230-1,250 nm; 6) 1,628-1,652 nm; 7) 2,105-2,155 nm; 8) 405-420 nm; 9) 438-448 nm; 10) 483-493 nm;	12-bit	250 m for bands 1 & 2; 500 m for bands 3 - 7; 1000 m for bands 8 - 36	1-2 days
MERIS (Medium Resolution Imaging Spectrometer)			Mar 2002	Apr 2012	15 Multispectral bands: 1) 407.5-417.5 nm; 2) 437.5-447.5 nm; 3) 485-495 nm; 4) 505-515 nm; 5) 555-565 nm; 6) 615-625 nm; 7) 660-670 nm; 8) 677.5-685 nm; Bands 9-15) 700-905	12-bit	From 300-1,200m	3 days
Hyperspectral Imager for the Coastal Ocean (HICO™)			Sep 2009	Sep 2014	87 Hyperspectral bands from 400 - 900 nm	8-bit	90m	3 days , with pre-scheduled acquisition
Geostationary Ocean Color Imager (GOCI)			Jun 2010	**Operational**	8 Multispectral bands: 1) 402-422 nm; 2)433-453 nm; 3) 480-500 nm; 4) 545-565 nm; 5) 650-670 nm; 6) 675-685 n; 7) 735-755 nm; 8) 845-885nm	12-bit	500m	Constant coverage of Korean Peninsula (2500 x 2500 km2)

1.2b Remote Sensing of Water Properties and Algal Biomass

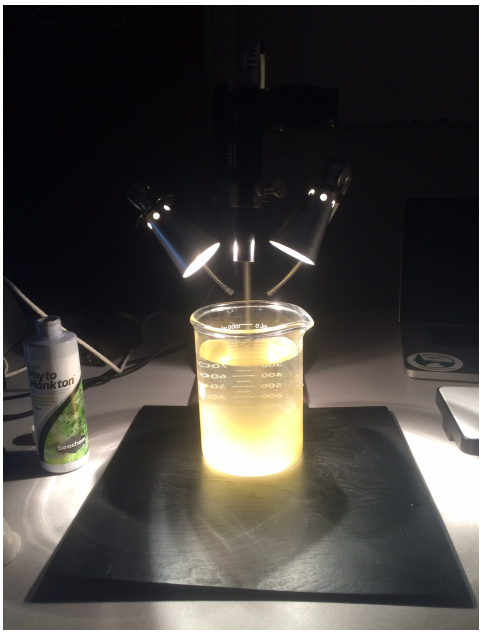
Remote sensing of water quality relies on the development of bio-optical algorithms that relate the reflectance at the surface to water quality parameters. Sunlight entering a clear water body is mostly absorbed within about 2 meters of the surface, depending on the wavelength. Light in the near infrared wavelengths (700 nm-1300 nm) is generally absorbed near the surface of the water. However, the absorption of water in the visible wavelengths varies dramatically based on water column constituents such as phytoplankton, suspended sediments, and dissolved organic matter. These water column constituents each influence the backscattering and absorption (the inherent optical properties, or IOPs) of a water body within the visible range, and therefore the spectral properties of a water body are determined by volume reflection, or the sum of contributions to these IOPs (Gordon, 2012).

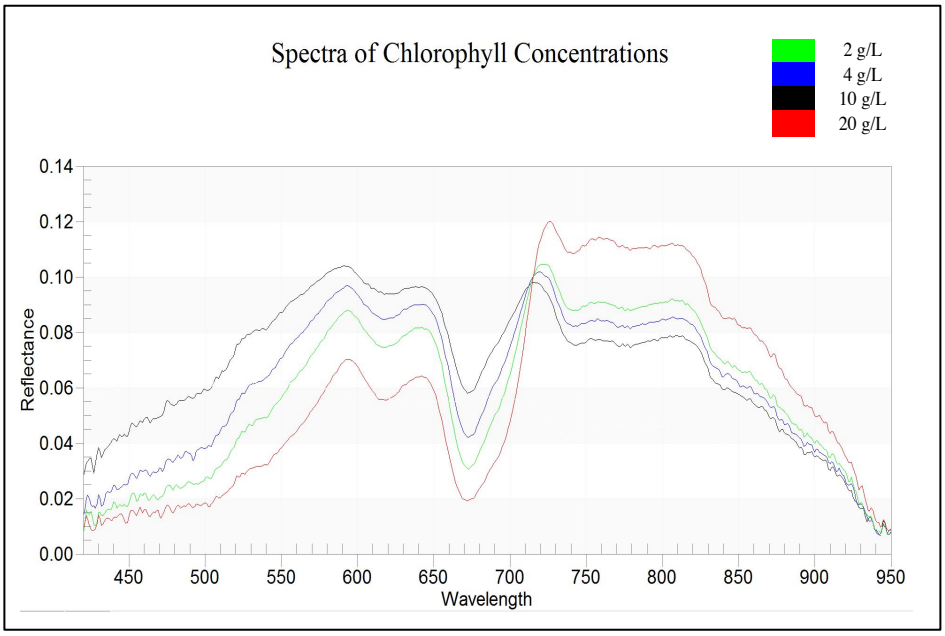
Morel and Prieur (1977) classified ocean water into "Case 1" and "Case 2" waters. The optical properties of Case 1 water are dominated by phytoplankton, with very low sediment and suspended matter (generally open ocean environments). Case 2 waters, on the other hand, are influenced by the optical contributions of other constituents such as suspended sediments or color-dissolved organic matter (CDOM), as well as phytoplankton biomass (Gordon and Morel, 2012). Turbid coastal and inland waters are characterized as Case 2 waters due to land runoff and anthropogenic inputs to the local ecosystem (Li et al., 2003). In both Case 1 and Case 2 waters, algae biomass is generally measured through the concentration of photosynthetic pigment Chlorophyll-a. Chlorophyll-a has two absorbance peaks near 433 nm (blue) and 686 nm (red), a reflectance maximum near 550 nm (green), and a reflectance peak around 690-700 (Cannizzaro, 2006).

In addition to spectral properties recorded in the literature, lab-based spectra were processed as part of this study for empirical measures of chlorophyll-laden water. Phytoplankton samples were taken from commercially available Seachem Reef Phytoplankton, a concentrated blend of green algae with size ranges from 1-20 μm. A 1000mL beaker was filled with 100 mL sand and 800 mL of water, and this sand was allowed to settle to the bottom of the beaker, coating the bottom of the beaker for the remainder of the experiment. Next, 200 mg of additional sand was added to the beaker and, while this sediment remained suspended in the water column, phytoplankton was added to the beaker at increments of 5mL, or 10 mg. Reflectance spectra were measured at 1nm intervals from 350- 2500nm with an ASD FieldSpec Pro spectroradiometer and calibrated with Spectralon as the white reference. Figure 2 shows the

experimental setup and recorded spectra. Lab-based spectra confirmed expected chlorophyll reflectance features, with reflectance minima in the blue (433 nm) and red (678 nm), and maxima in the green (590 nm) and at 722 nm. Overall strength reflectance characteristics increased with increasing chlorophyll concentration as expected, with the magnitude of the 722nm reflectance peak increasing with particularly high concentrations of phytoplankton.

Figure 2. Experimental setup with 200 mg/L of suspended sediment in 1000 mL beaker and ASD FieldSpec Pro spectroradiometer (left). Lab spectra of varying Chlorophyll concentrations, with visible reflectance peaks at 555 and 700nm (right).





The relationships of light reflected or absorbed at specific wavelengths in the visible spectrum, such as those recorded in this study's lab stud measurements, can be used to estimate chlorophyll-a concentrations using satellite bio-optical algorithms (O'Reilly et al. 1998). Bio-optical algorithms have successfully been derived for Case 1 waters, with numerous studies presenting a comprehensive log-linear relationship between chlorophyll-a concentration and selected spectral band ratios (Jensen, 2000). Furthermore, the NASA Ocean Biology Processing Group (OBPG) has developed maximum band ratio algorithms for use in large-scale oceanic and global chlorophyll-a mapping. These "OC" algorithms model a fourth-order polynomial relationship between the ratio of remote sensing reflectance in the blue and green wavelengths and chlorophyll-a, with coefficients varying based on sensor characteristics [Equation 1]:

$$log_{10}(Chl_a) = \beta_0 + \beta_1 X_i + \beta_1 {X_i}^2 + \beta_1 {X_i}^3 + \beta_1 {X_i}^4 + \varepsilon_i \quad \text{Eq. 1}$$

$$X_i = log_{10}(\frac{Band_k}{Band_l})_i$$

Unlike these 'ocean color' algorithms, however, chlorophyll-a abundance in Case 2 waters (turbid coastal or lake waters with high suspended solid concentrations) cannot be described through a single and comprehensive band ratio algorithm due to confounding and complex water column constituents (Nieke, 1997; Doerffer, 2007; Gitelson, 2009). Although bio-optical algorithms may be successfully correlated to water quality parameters in Case 2 waters, these developed algorithms are only viable for a particular location and not applicable on a larger scale (Nieke, 1997; Doerffer, 2007; Gitelson, 2009). Figure 3 illustrates the various water column constituents that may contribute to the water-leaving signal and therefore must be considered when developing bio-optical algorithms, namely surface reflection/win, suspended sediments (both organic and inorganic), bottom reflection, and phytoplankton populations.

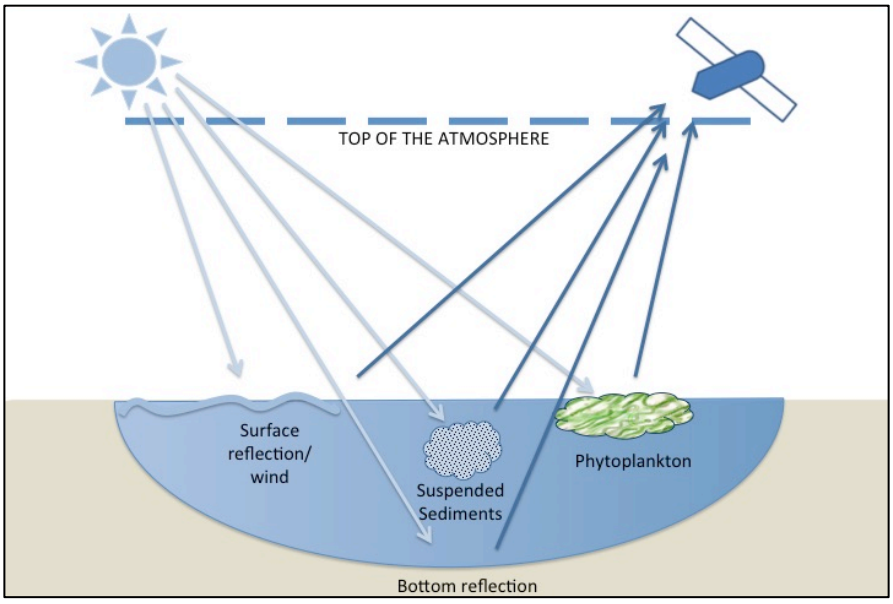


Figure 3: Diagram showing various possible water column contributions to Case 2 water leaving radiance.

1.2c Harmful Algal Bloom Remote Sensing

With the increasing frequency of HABs worldwide, remote sensing classification of these blooms has become an active area of research. Large-scale HAB monitoring programs have been developed by the National Oceanic and Atmospheric Administration (NOAA) and the

National Centers for Coastal Ocean Science (NCCOS), including the application of Medium Resolution Imaging Spectrometer (MERIS) data for HAB monitoring off the Florida Coast and Chesapeake Bay and MODIS-based models monitoring blooms in Southern California oceans (Anderson, 2009).

Because all phytoplankton (both green algae and cyanobacteria) contains chlorophyll-a, remote sensing of chlorophyll alone cannot be used to determine the specific species composition or toxicity of assemblages. Previous studies have attempted to use the accessory pigment phycocyanin, present only in cyanobacteria, to quantify blue-green algae specifically and delineate HABs from healthy phytoplankton growth (Miller, 2006; Marion et al., 2012; Qi et al., 2014). Phycocyanin has a characteristic reflectance minimum at 625 nm, making it possible to separate from a normal chlorophyll spectrum. However, this characteristic feature of phycocyanin is only detectable at high chlorophyll-a concentrations and requires high spectral resolution to distinguish from normal chlorophyll reflectance spectra. Because of the problems with phycocyanin, most studies of HABs remote sensing models rely on chlorophyll-a concentration alone as a bloom indicator. This study in particular will focus on chlorophyll anomalies as HAB indicators, an alert framework designed to work in conjunction with the current monitoring protocol in Lake Champlain, and expanding the potential for initial identification of algal blooms.

A number of studies have successfully modeled recent HABs in the Great Lakes, particularly after the recent toxic events in western Lake Eerie. Vincent 2004 successfully used Landsat TM and ETM+ data to assess the concentration and spatial distribution of cyanobacteria blooms in Lake Erie through step-wise linear regression using *in situ* measurements. Other studies (Pozdnyakov 2005, Becker et al. 2009, Weghorst 2008, Wynne 2008, McCullough 2012, Zhang 2012) applied MODIS and MERIS sensors to derive cyanobacteria abundances at a lower spatial but higher temporal resolution within the Great Lakes.

Within Lake Champlain, a number of studies have attempted to quantify algal bloom growth in recent years. Wheeler (2011) used MERIS observations to test semi-analytical algorithms for phycocyanin and chlorophyll-a concentrations, as well as developing empirical models from a QuickBird (a DigitalGlobe product) image. Isenstein (2014) employed Landsat ETM data for regression modeling of Lake Champlain chlorophyll-a, and Torbick (2015) applied Landsat TM, Rapid Eye, and Proba Compact High Resolution Imaging Spectrometer (CHRIS) data to develop empirical models for chlorophyll-a concentration in Lake Champlain. Despite

the success of these studies, however, all previous remote sensing of Lake Champlain HABs has relied on one or two images for calibration and validation. This study expands upon this previous research, looking at a three-year time interval of *in situ* chlorophyll measurements and satellite imagery, as well as assessing the potential of the currently operational Landsat 8 and MODIS sensors for ongoing HAB monitoring and detection.

1.3 Atmospheric Correction for Water Quality Remote Sensing

1.3a Introduction

Accurate atmospheric correction, critical in all remote sensing studies, is especially important for remote sensing of water quality, as the brightness variation of water bodies is far less than that of land surfaces. Rather than recording water-leaving radiance, space-borne remote sensors observe the total radiance exiting the top of the atmosphere (TOA), at least 90% of which is scattered sunlight from aerosols and air molecules that never penetrated the water surface (Hu et al., 2001) [Figure 3]. In order to retrieve water-leaving radiance from observed TOA images, these overwhelming atmospheric effects must be modeled and removed, accounting for differences in aerosol optical thickness and properties and the way in which they influence the apparent reflectance.

A number of different atmospheric correction methods have been developed to measure and remove heterogeneous atmospheric effects (Gordon and Wang, 1994; Wang et al., 2009). Standard NASA atmospheric correction algorithms are based on Gordon and Wang (1994)'s approach, which models the TOA radiance over water as the sum of atmospheric, surface, and subsurface contributions [Equation 2].

$$Lt(\lambda) = Lr(\lambda) + La(\lambda) + t(\lambda)Lf(\lambda) + T(\lambda)Lg(\lambda) + t(\lambda)Lw(\lambda) \qquad \text{Eq. 2}$$

Where λ denotes the sensor spectral band wavelengths, Lr(λ) is the Raleigh scattering by air molecules in the absence of aerosols, La(λ) is the multiple scattering from aerosols in the absence of Rayleigh as well as Rayleigh–aerosol interactions, t(λ) and T(λ) are the diffuse and direct atmospheric transmittance from surface to sensor, Lf(λ) is the contribution from whitecaps and foam on the surface that is diffusely transmitted to the TOA, Lg(λ) is the specular reflection (glint) from the surface that is directly transmitted to the sensor field of view, and Lw(λ) is the water-leaving radiance that is diffusely transmitted to the TOA. The Raleigh scattering and specular reflectance in a scene are predictable based off solar and sensor characteristics, leaving the water-leaving radiances (Lw(λ)) and the aerosol radiance (La(λ)) as the two primary unknowns in this equation, with the goal of atmospheric correction being to model the latter and retrieve the former.

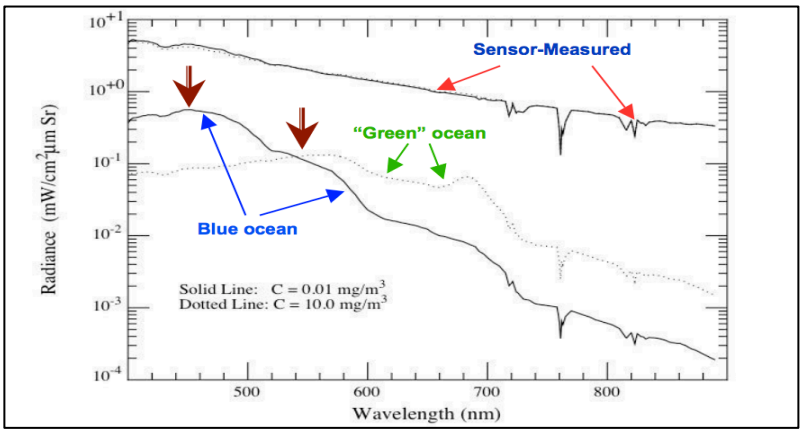


Figure 3: The impact of atmospheric noise on ocean color spectra. A least 90% of the at-sensor radiance (Sensor-Measured) is due to atmospheric noise and must be removed in order to measure ocean/water color (taken from Wang et al., 2001).

The SeaWiFS Data Analysis System (SeaDAS) software package distributed by NASA's Ocean Biology Processing Group has been designed for the processing and viewing of ocean remote sensing data, and, within SeaDAS, the l2gen code can be used to atmospherically correct Level 1 ocean data. Originally designed for the SeaWiFS satellite, over the past few years, the software has been expanded to support the processing of other sensors, including MODIS, VIIRS, HICO, MERIS and, in 2015, Landsat 8 (NASA SeaDAS manual, n.d.)

1.3b Approaches to Modeling Aerosol Radiance

In order to retrieve water-leaving radiance (Lw(λ)), the aerosol radiance (La(λ)) must be modeled based on scene-specific information. This scattering from aerosols can be modeled through multiple approaches, including a suite of atmospheric correction algorithms included within the SeaDAS l2gen package.

The standard aerosol correction algorithm used by NASA's Ocean Biology Processing Group (OBPG) relies on the Gordon and Wang model, with aerosol radiance computed from the ratio of two bands for which the water-leaving radiance is negligible (generally two or more NIR bands). This assumption of complete radiance absorption in the near infrared (NIR) has been demonstrated to prove true over Case 1 (clear ocean) waters (Gordon & Wang, 1994) but can lead to errors when applied to Case 2 (turbid and inland) waters (Ruddick et al., 2000). This concept is illustrated in Figure 3, as the 'Blue Ocean' spectra has a reflectance of close to zero for all wavelengths greater than 800nm. Therefore, aerosol optical thickness (and aerosol radiance) can be modeled by the difference between the blue ocean and sensor-measured spectra

in the NIR. This calculated aerosol optical thickness can then be extrapolated back to lower wavelengths, allowing the atmospheric correction models to remove atmospheric noise from the overall spectral shape. As is evident in Figure 3, the green ocean spectra (that of case 2, high sediment load/chlorophyll water) has non-zero reflectance in the NIR. Therefore, this atmospheric correction approach generally overcorrects for aerosol radiance over inland water bodies, where NIR signals cannot be assumed to be zero, and are affected by in-water constituents, as well as aerosol scattering.

Building off of these NIR models, Wang and Shi developed a short wave infrared (SWIR)-based atmospheric correction. Unlike the NIR wavelengths, SWIR reflectance can be assumed to be near zero for both Case 1 and Case 2 waters. The Wang and Shi model employs a turbid water index to determine which aerosol model (SWIR versus NIR) should be used within a scene. However, due to the low signal-to-noise of most SWIR bands, this method often includes large uncertainties in aerosol radiance and fails over highly productive waters (Wang and Shi, 2007).

Due to these inherent problems in both NIR and SWIR correction, the Remote Sensing and Ecosystem Modeling team developed MUMM (Management Unit of the North Sea Mathematical Models) atmospheric correction method for SeaWiFs application over turbid waters. The method assumes spatial homogeneity of the SeaWiFs 765 nm/865 nm band ratios for aerosol reflectance and for water-leaving reflectance (Ruddick et al., 2000). For MODIS, the two NIR bands centered at 748 nm and 869 nm are used, while in Landsat 8 MUMM correction bands 4 (640 nm) and 5 (850-880 nm) are used.

All three atmospheric correction methods (NIR, SWIR, and MUMM) were examined for application to Landsat 8 and MODIS remote sensing of Lake Champlain, with the best atmospheric correction being used to estimate chlorophyll-a concentration and HAB abundance.

Section 2: SITE DESCRIPTION AND *IN SITU* SAMPLING

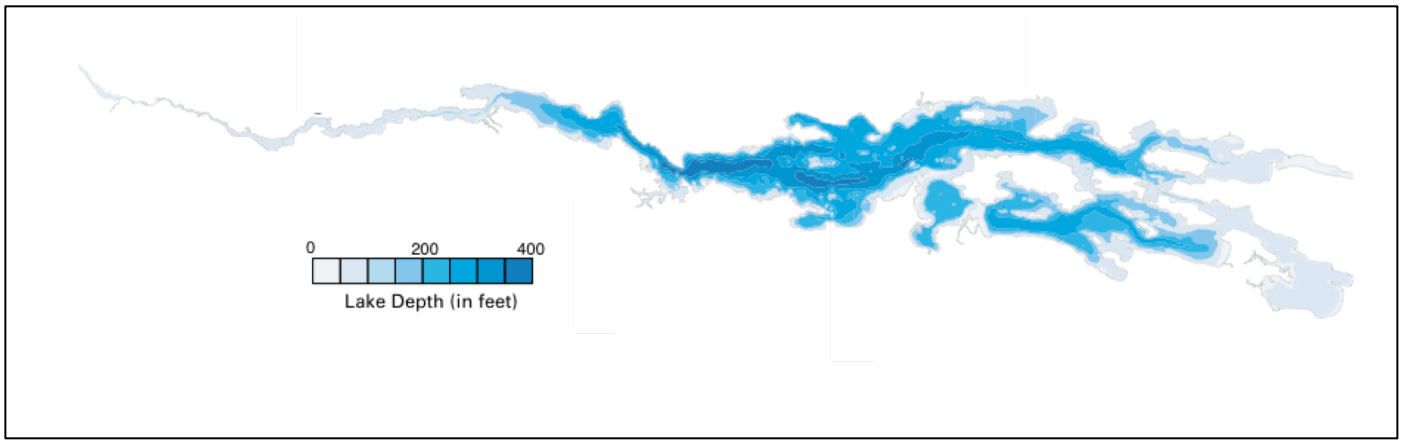
Figure 4: New England Region with Lake Champlain from MODIS imagery.

2.1 Site Description: Lake Champlain

2.1a Lake Champlain Atlas Data

Lake Champlain is the sixth largest freshwater lake in the United States. Located between New York and Vermont, and extending into Canada in the North, the lake is 120 miles long with a surface area of 1127 km^2 and a volume of 25.8 km^3 (https://www.epa.gov) [Figure 4]. More than 550,000 people reside in the Lake Champlain watershed, over 200,000 of whom depend on the lake for drinking water. Approximately 4,149 draw water directly from Lake Champlain for individual use, as well as 99 public water systems that source water from the lake. In addition, Lake Champlain is a major recreational area, with 54 public beaches and hundreds of private homes and recreational beaches, which contribute to a fishing and recreation economy of nearly $4 billion annually. Although the average depth of the Lake Champlain is 19.5 meters, the lake reaches depths of over 122 meters in the area between Charlotte, Vermont, and Essex, New York and can be as shallow as 4.5 meters in areas such as Missisquoi and St. Albans Bay [Figure 5] (LCBP Lake Champlain Basin Atlas).

Figure 5: Lake Champlain Bathymetry Map (from Lake Champlain Basin Program Atlas)



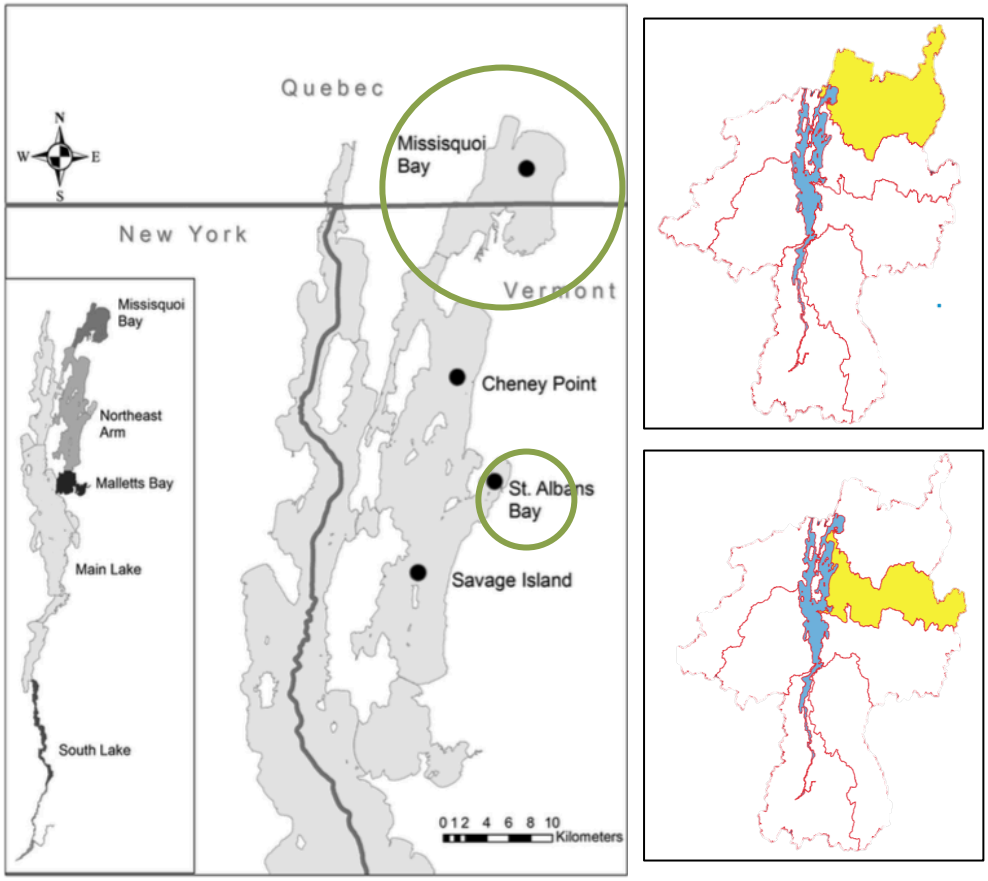
2.1b Spatial Variation in Algal Blooms and Eutrophication

Since the death of several dogs in 1999 and 2000, algal toxins have emerged as an area of study in Lake Champlain. However, these toxic blooms remain concentrated in a small area of the lake, as *in situ* sampling studies have identified dense seasonal blooms predominantly in Missisquoi and St. Albans Bays, the large, shallow embayments of the lake's Northeast Arm (Boyer et al., 2004; Mihuc et al., 2006; Smeltzer et al., 2012). Furthermore, Mihuc (2006) analyzed the species composition of Lake Champlain algal populations and found that all sites in the Northeast Arm of the lake contained primarily blue-green algae species, with *Microcystis* (a toxic algae species) making up the dominant taxa.

The Northeast Arm of Lake Champlain makes up about one-quarter (269km^2) of the lake's surface area, but has a mean depth of only 13m (LCBP Lake Champlain Basin Atlas). Studies of eutrophication in Lake Champlain have historically focused on this area, as eutrophication of the lake first manifested in St. Albans and Mississquoi Bays in the 1960s (Myers and Gruendling, 1979). In fact, between 1979-2009, phosphorus concentration increased by 72% and chlorophyll concentrations doubled in Mississquoi Bay (Smeltzer et al., 2012). Furthermore, based on sediment cores taken from the Northeast Arm of the Lake, Levine et al. (2012) concluded that these patterns of geochemical and biological change were due primarily to eutrophication by nutrients from urban and agricultural sources, as the large catchment area of the Northeast Arm (particularly that of Mississquoi Bay) has seen significant agricultural development over the past century (Levine et at., 2012). Figure 6 illustrates the location of these

two bays, as well as the the watershed areas of Mississquoi and St. Albans Bays, which together make up the catchment of Lake Champlain's Northeast Arm. Based on these characteristics, St. Albans and Mississquoi Bay were selected as areas of focus for this study.

Figure 6: Locations in Lake Champlain Northeast Arm from Levine 2012 (left) and watershed areas of Mississquoi and St. Albans Bays, which combine to make up the entire Northeast Arm drainage area of Lake Champlain (from LCBP Atlas).



2.2 Lake Champlain *In Situ* Sampling

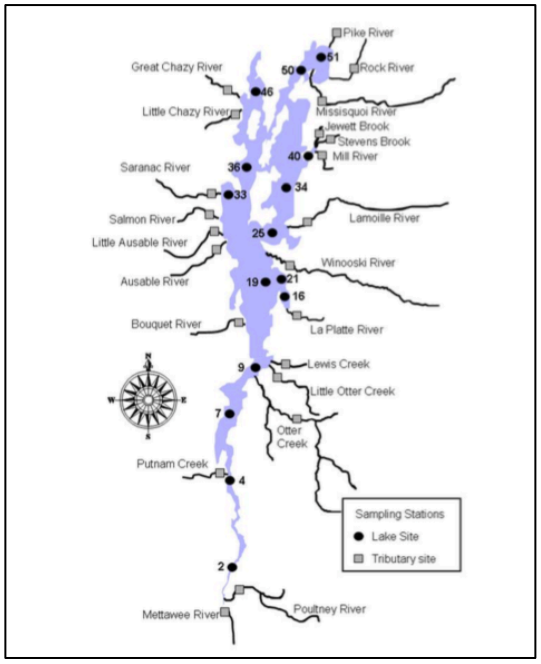
2.2a LTMP Sampling Protocol

Sine 1992, the VT DEC has conducted the Long-Term Water Quality and Biological Monitoring Project (LTMP), focused on providing regular limnological surveys of Lake Champlain and detecting long-term environmental change in the region. The LTMP collects field measurements at 15 lake stations, visiting these stations regularly from late April through October each year [Figure 7]. Sampled parameters include temperature, dissolved oxygen, pH,

alkalinity, total phosphorus, dissolved phosphorus, total nitrogen, dissolved silica, chloride, chlorophyll, phytoplankton, and zooplankton. Three years of biological and water quality measurements (between 2013 and 2015) were used for this study, focusing on the algal bloom indicator of chlorophyll-a.

Within the LTMP protocol, chlorophyll-a samples are collected using a vertically-integrated hose-sampler beginning at the lake surface to a depth representing twice the Secchi depth. Samples are then filtered in the field (100 ml on 47mm diameter GF/A glass fiber filters wrapped in 90 mm No.3 glass fiber filters) and placed in a dark container on ice for transport to the laboratory, where they are measured using a hydrolab fluorometer. These chlorophyll-a samples are reported in concentrations of µg/L (VTDEC, 2006).

Figure 7: Lake Champlain sampling location position and depth (from VT DEC protocol)



Lake Station	Latitude	Longitude	Depth (m)
2	43.71483333	-73.383	5
4	43.95166667	-73.4078333	10
7	44.126	-73.4128333	50
9	44.24216667	-73.3291667	97
16	44.42583333	-73.232	25
19	44.471	-73.2991667	100
21	44.47483333	-73.2316667	15
25	44.582	-73.2811667	32
33	44.70116667	-73.4181667	11
34	44.70816667	-73.2268333	50
36	44.75616667	-73.355	50
40	44.78533333	-73.1621667	7
46	44.94833333	-73.34	7
50	45.01333333	-73.1738333	4
51	45.04166667	-73.1296667	5

		Mississquoi Bay	St. Albans Bay
Bay characteristics			
	Surface area (km2)	77.5	7.2
	Mean depth (m)	2.8	8
	Max depth (m)	4	12
	Volume (km3)	0.22	0.023
Catchment/Watershed			
	Area (km2)	3105	130
	% forest	62	24
	% agriculture	25	56
	%urban	5	14
	P load (kg/km3/yr)	1931	940
Water Chemistry			
	Total P (mg/L)	0.048	0.027
	Total N (mg/L)	0.7	0.44
	TN:TP	15	16
	Dissolved Si	2.07	0.65
	Chlorophull a (µg/L)	14.7	5.4
	Chlorophyll a:total P	0.31	0.37

Table 3: summary of Mississquoi and St. Albans Bay characteristics. Bay characteristics from LCBP Atlas, Catchment/Watershed statistics from Levine 2012, Water Chemistry Calculated from LTMP dataset from 1992-2015.

2.2b Analysis of *in situ* LTMP trends

Despite regular sampling since 1992 by the LTMP, this data has primarily served to monitor local run-off regulation and eutrophication, and few comprehensive analyses have been conducted on the dataset. Smeltzer et al. analyzed LTMP data from 1992-2012 to document water quality and biological changes in the lake. However, no analysis has been documented for the subsequent years of data. Therefore, in order to obtain a more comprehensive understanding of water quality and algal dynamics in the years of study, *in situ* lake data was evaluated for site-based trends.

In order to gain a spatial understanding of water quality parameters, LTMP samples from 1992-2015 were analyzed by site. Analyses were carried out in R statistical software.
The areas of Missisquoi and St. Albans Bay have significantly higher algal biomass than other areas of the lake (mean 16.51 and 13.82 µg/L for all years, as opposed to a lake-wide mean of 6.37 µg/L). The two Missisquoi Bay stations (Station 50 and 51) also had significantly higher total phosphorus and nitrogen measurements across all years. Areas in the southern region of the lake (Stations 2 and 7), also had high chlorophyll-a, phosphorus and nitrogen levels. However

these areas are dominated by river input from the Poultney and Mettawee Rivers and likely reflect more direct runoff sampling than the rest of the stations [Figure 8]. All other measured parameters (Alkalinity, Chloride, Dissolved Inorganic and Organic Carbon, Dissolved Silica, and Total Suspended Solids) did not prove significantly different for the Northeast Arm of Lake Champlain.

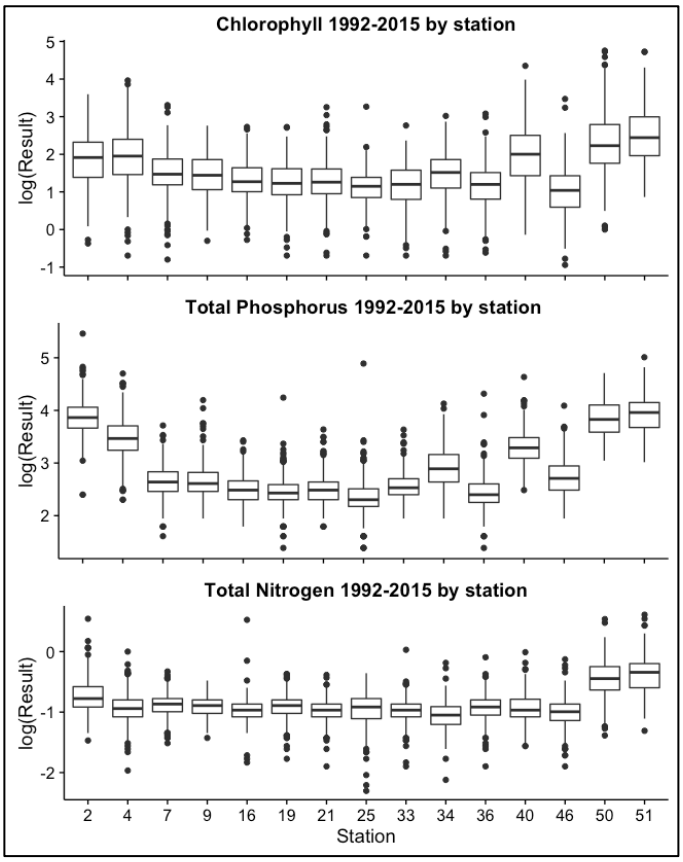


Figure 8: Boxplots for sampled water parameters 1992-2015, showing increased concentrations of Chlorophyll-a, Total Phosphorus and Total Nitrogen at Northeast Arm sampling stations.

Data for 2013-2015 (the years of focus for this study) were then analyzed separately for water quality trends and algal bloom indicators. As mentioned earlier in this paper, chlorophyll content is used as the primary bloom indicator in the current Lake Champlain HAB monitoring protocol. Chlorophyll anomaly "bloom" events showed a clear seasonal trend for all years of study, with chlorophyll concentration reaching extremely high levels in Missisquoi Bay during late August of 2014 and 2015 and September of 2013 (with another apparent "bloom" spike in

early August of 2013. These *in situ* sample trends are consistent with alert statuses from the DEC's Algal tracking map, with "high alert" statuses for St. Albans and Mississquoi Bay (times when high levels of cyanotoxins were quantitatively sampled) closely aligning with peak chlorophyll in LTMP sample results [Figure 9]. This demonstrates a clear relationship between chlorophyll spikes and harmful algal bloom events in both St. Albans and Mississquoi Bay, a trend closely correlated due to the dominance of toxin-producing taxa in these regions. In addition, this relationship suggests that chlorophyll-a remote sensing alone could provide valuable information about potential toxic events in the Northeastern Arm waters.

Figure 9: *In situ* chlorophyll measurements for sites 40, 50 and 51 (St. Albans and Mississquoi Bay) for 2013-2015. Red stars indicate "High Alert" status from Vermont Blue Green Algae Tracker.

SECTION 3: LANDSAT 8

3.1 Methods

3.1a Introduction: Sensor Description and Characteristics

Landsat-8 is part of a lineage of Landsat satellites in operation from 1972 to present, providing the longest record of Earth observation from space. These satellites have been designed almost exclusively for terrestrial remote sensing, with both coverage areas and instrument specifications focusing on geological and ecological terrestrial applications (NASA Landsat, n.d.) However, the Operational Land Imager (OLI) on Landsat-8 offers the potential for expansion of the Landsat suite to aquatic remote sensing applications due to its improved radiometric sensitivity and spectral resolution (Pahlevan, 2012, Gerace, 2013; Vanhellemont, 2014a,b,c; 2015; Franz et al., 2015; Concha, 2015).

Landsat 8 Operational Land Imager (OLI) was launched into orbit February 11, 2013 as the next generation of the Landsat Data Continuity Mission (LDCM). OLI collects seven spectral bands at 30-m spatial resolution between 430 and 2290 nm, one panchromatic (500 to 680 nm) band at 15 m and two thermal infrared (TIRS) channels (10,600 to 11,190; 11,500 to 12,510 nm) at 100-m resolution (Table 4). The increased radiometric sensitivity of Landsat 8 can be defined by its average signal-to-noise-ratio (SNR). This increased SNR is due to the pushbroom sensor design of OLI, in contrast to the whiskbroom design of previous Landsat satellites. This design creates a longer integration time and, therefore, improved sensitivity (USGS, 2013; Schott, 2012). Based on on-orbit characterization and pre-launch simulation, the high SNR of OLI has been shown to be comparable to that of ocean color missions such as SeaWIFs and MODIS (Hu et al., 2013).

Table 4: Landsat 8 band designations and resolutions

Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 - Coastal aerosol	0.43 - 0.45	30
Band 2 - Blue	0.45 - 0.51	30
Band 3 - Green	0.53 - 0.59	30
Band 4 - Red	0.64 - 0.67	30
Band 5 - Near Infrared (NIR)	0.85 - 0.88	30
Band 6 - SWIR 1	1.57 - 1.65	30
Band 7 - SWIR 2	2.11 - 2.29	30
Band 8 - Panchromatic	0.50 - 0.68	15
Band 9 - Cirrus	1.36 - 1.38	30
Band 10 - Thermal Infrared (TIRS) 1	10.60 - 11.19	100 * (30)
Band 11 - Thermal Infrared (TIRS) 2	11.50 - 12.51	100 * (30)

3.1b Acquisition and Atmospheric Correction

14 cloud-free Landsat 8 Level 1T images for path 19, row 29 were obtained from the U.S. Geological Survey Earth Explorer website, spanning a time interval from May 16th, 2013 to September 27th, 2015. Landsat Level 1T (Standard terrain correction) images have been corrected for radiometric and geometric accuracy by incorporating ground control points from the GLS2000 data set and employing a Digital Elevation Model (DEM) for topographic accuracy (http://landsat.usgs.gov/).

Landsat scenes were subset to an area that includes Missisquoi and St. Albans Bay and converted to netcdf format for processing in SeaDAS. Basic Landsat 8-OLI processing was added to the SeaDAS l2gen code in September, 2015, and since implementation, few studies have investigated the potential of this atmospherically corrected Landsat data, with varying results (Vanhellemont, 2014a,b,c, 2015; Concha, 2015). For this study, the NIR, SWIR, and MUMM algorithms (as described in the introduction) were tested for Lake Champlain Landsat 8-OLI atmospheric correction.

For NIR atmospheric correction of Landsat 8 imagery aer_opt was set to 1- Multi-scattering with 2-band model selection, with bands 4 and 5 (655 and 865 nm) set to the long and short aerosol wavelengths. This model assumes that NIR water-leaving radiance is zero and therefore is expected to overcorrect the atmosphere above turbid lake water (Gordon & Wang,

1994). The Landsat 8 image from September 27th, 2015 (LC80140292015270LGN00) was selected for atmospheric correction evaluation.

Wang and Shi SWIR processing was tested using Landsat 8 bands 6 and 7 (1609 and 2201 nm). The increased radiometric sensitivity of Landsat 8 OLI SWIR bands compared to other Landsat sensors makes this correction method a feasible approach for OLI (Vanhellemont, 2014a).

The MUMM approach was also applied using the SWIR bands (Ruddick et al., 2000). The method was applied using implementation in SeaDAS (aer_opt=-10), using bands 4 and 5 (655 and 865 nm) and setting the calibration parameter $\alpha = 1.95$ for OLI as suggested by Vanhellemont, 2014a.

3.1c Processing and Selection of Sampling Points

Following the evaluation of atmospheric correction methods, all 14 Landsat 8 OLI scenes were processed using the MUMM model. Atmospherically corrected images were converted to hdf format and exported from SeaDAS to ENVI for further data analysis. A land mask was created by buffering SWIR reflectance, as SWIR water-leaving radiance is approximately negligible, and this mask was applied to all images. An xml file was created with locations for stations 40, 46, 50 and 51 and imported into ENVI as regions of interests (ROI). These single-pixel sampling points were then buffered to a 3x3 pixel area (90m x 90m) in order to help account for time lags between image acquisition and *in situ* sampling and heterogeneity of water masses.

MATLAB code was written to export spectra at each sample location, and *in situ* chlorophyll samples from the VT DEC dataset were matched to corresponding images. *In situ* samples were considered matchups within a ± 2 day window, based on time gap recommendations from Stadelmann et al. (2001). A total of 37 sample matchups were used in the development of chlorophyll algorithms.

In order to test the other possible variables affecting water spectral properties, ancillary variables were collected for the Northeastern Arm of Lake Champlain. In order to model bottom-effect water depth (in m) for each sample location, sample depth was obtained from VT DEC sampling protocol, with depth for all locations varying between 4-7 meters. Meteorological data for wind speed data (WSPD, m/s) and wave height (WVHT, m) were acquired from NOAA Buoy Station 45166, located in the center of Mississquoi Bay. All meteorological data was

acquired for the date of overpass at 11:30 EST/15:30 GMT, as Landsat overpass time for Lake Champlain is at approximately 15:38 GMT. In addition, mean discharge rates (in cubic feet per second) for the 7 days prior to image acquisition were calculated for the each image. River discharge rates were obtained for the USGS Gages at Pike River (Gage 04294300) and Mississquoi River (04294000) [Figure 9]. These rates were used as a proxy for riverine sediment discharge and therefore possible suspended sediment loads, in the Northeastern Arm of Lake Champlain, as no coincident measurements of total suspended solids or turbidity exist for LMPT sampling.

Figure 9: Location of USGS Gages at Pike River (Gage 04294300) and Mississquoi River (04294000), tributaries Mississquoi Bay and the entire Northeast Arm of Lake Champlain

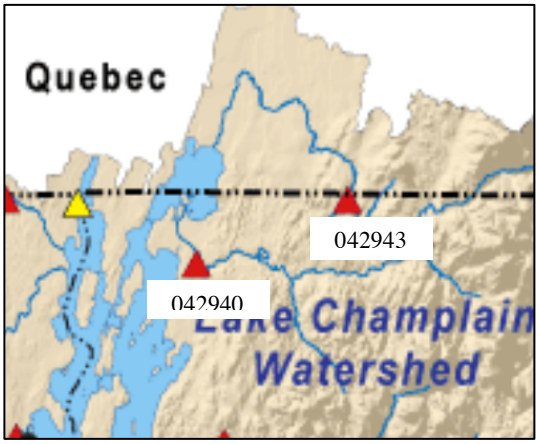


Table 5: Data predictors used for Landsat 8 model development

Sample Chl-a	In situ Chl-a measurements in µg/L from LTMP sampling

Sample Date	Date of in situ sampling MM/DD/YY
Time lag	Time difference: Date of image acquisition - Date of in situ sampling
Wind speed	Wind speed at Buoy Station 45166 (m/s)
Wave height	Average of the highest one-third of all of the wave heights (in m) during the 20-minute sampling period at Buoy Station 45166
Mean river discharge	Discharge rates from USGS gages on Pike and Mississquoi Rivers (in cubic ft/sec), averaged for 7 days prior image acquisition from
Sample location depth	Depth of sample location (m) from LTMP protocol
B1	Landsat 8 Band 1, Coastal Aerosol, 420-450 nm (Rrs)
B2	Landsat 8 Band 2, Blue, 450-510 nm (Rrs)
B3	Landsat 8 Band 3, Green, 530-590 nm (Rrs)
B4	Landsat 8 Band 4, Red, 640-670 nm (Rrs)
B5	Landsat 8 Band 5, NIR, 850-880 nm (Rrs)
B6	Landsat 8 Band 6, SWIR I, 1,570-1,650 nm (Rrs)
B7	Landsat 8 Band 7, SWIR II, 2,110-2,290 nm (Rrs)
B9	Landsat 8 Band 9, Cirrus, 1,360-1,380 nm (Rrs)
B10	Landsat 8 Band 10, TIRS I, 10,600-11,190 nm (BT)

3.1d Analysis

Linear least squares regression analysis was performed for bio-optical algorithm development using Minitab statistical software. Based on existing NASA chlorophyll algorithms and previous research, Landsat bands 1, 2, and 3 were expected to correlate with chlorophyll. Band ratio algorithms, rather than simple band reflectance values, were employed to normalize reflectance spectra between scenes. Best subsets regression was performed between log(Chlorophyll), band ratio combinations, and other predictors [Table 5] to develop a model of chlorophyll remote sensing for the Northeast Arm of Lake Champlain.

3.2 Results

3.2a Evaluation of Atmospheric Correction Models

Fig.1 shows a subset of the Landsat 8 scene for the Northeast corner of Lake Champlain for September 27th, 2015 after processing using NIR (bands 4 and 5), SWIR (bands 6 and 7) and

MUMM atmospheric correction. The relative noise in the SWIR bands of Landsat OLI is apparent in the corrected images, with the SWIR-corrected image appearing speckled and poor-quality [Figure 10].

Spectra were extracted within a 3x3 buffer around each station pixel [Figure 11]. The overcorrection of both NIR and SWIR atmospheric correction models is apparent in these spectral plots, as these atmospheric correction methods produced negative reflectance values for bands 1-3 (wavelengths of 442-561 nm) at most sites. Negative reflectance values are the result of the aerosol modeling approach, as NIR and SWIR correction methods incorrectly recognize the high reflectance values of turbid water as atmospheric effects, over-modeling aerosol contributions to the local atmosphere. The MUMM correction method, on the other hand, replaces assumptions of zero water-leaving radiance with an assumption of spatial homogeneity at bands 4 and 5. The success of this method is clear from this study, as only MUMM correction was capable of retrieving realistic marine reflectance values.

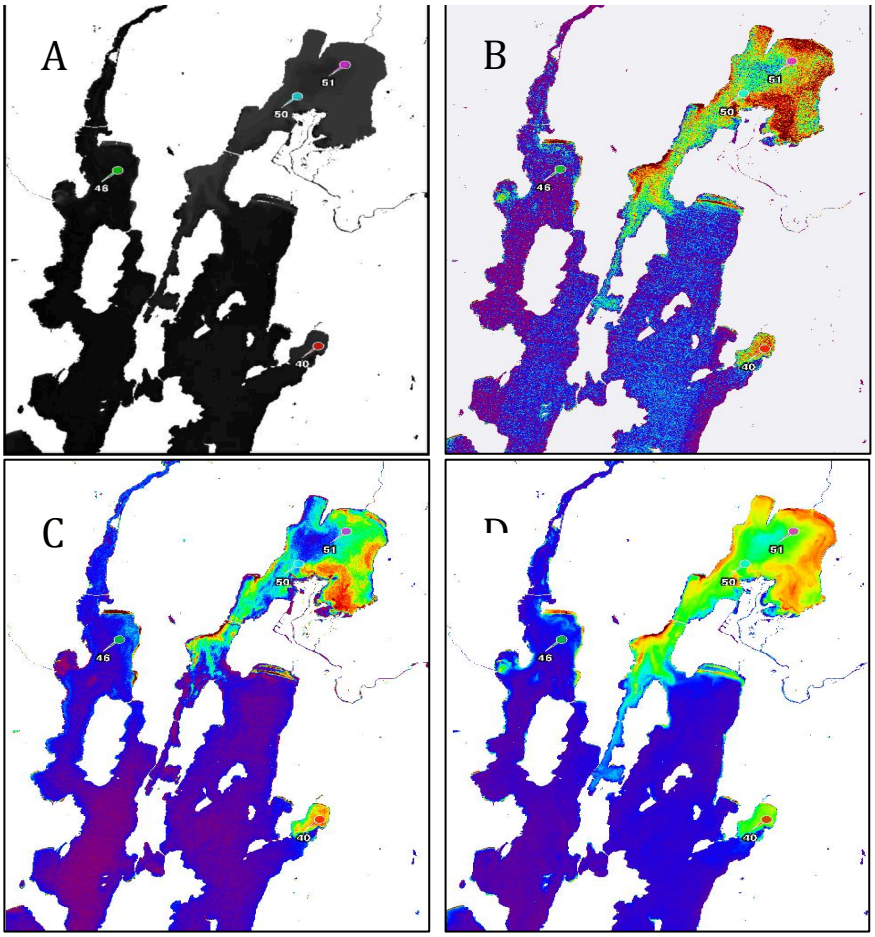


Figure 10: SeaDAS atmospheric correction results: a) location of station locations 40, 46, 50 and 51; b) SeaDAS SWIR correction results for Rrs 561 nm; c) SeaDAS NIR correction results for Rrs 561 n; d) SeaDAS MUMM correction results for Rrs 561 nm

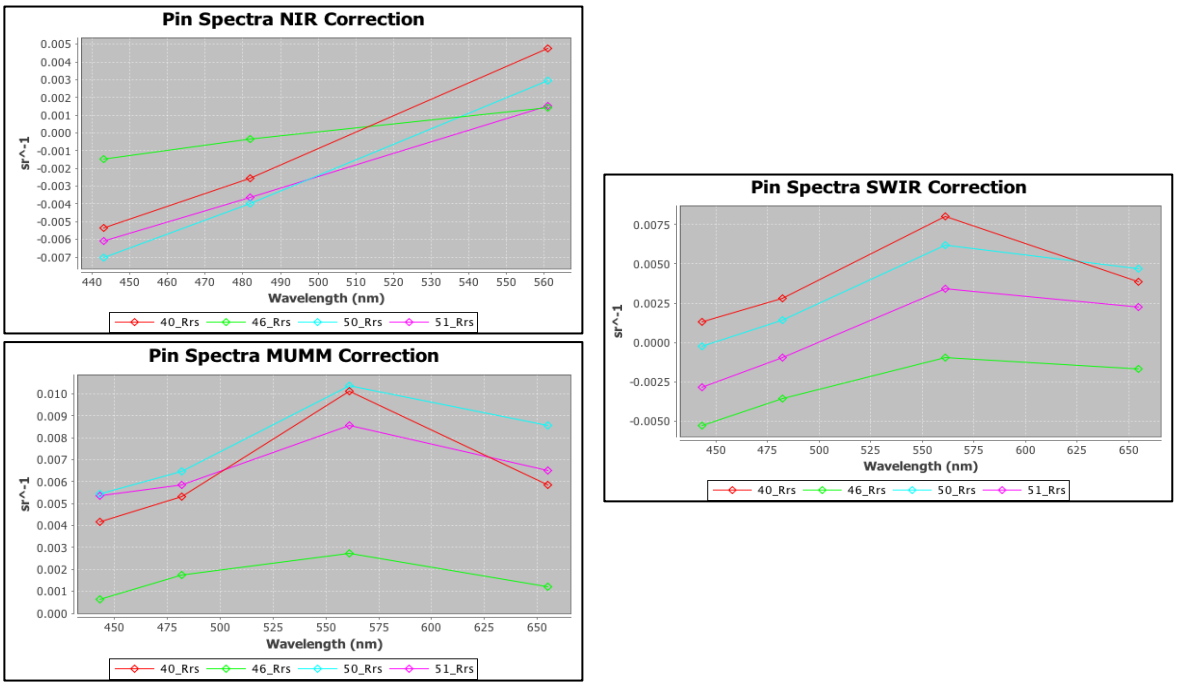


Figure 11: Station pin spectra extracted from SeaDAS for NIR, SWIR and MUMM correction methods

3.2b Regression Analysis

Initial plots of chlorophyll concentration (in µg/L) revealed heteroscedasticity, and therefore log(Chlorophyll) was used as a response variable for all models. All possible band ratios were calculated using Minitab software. Figure 12 is a matrix plot showing potential correlations between logChl and all available band ratios. Red lines represent regression lines for each relationship.

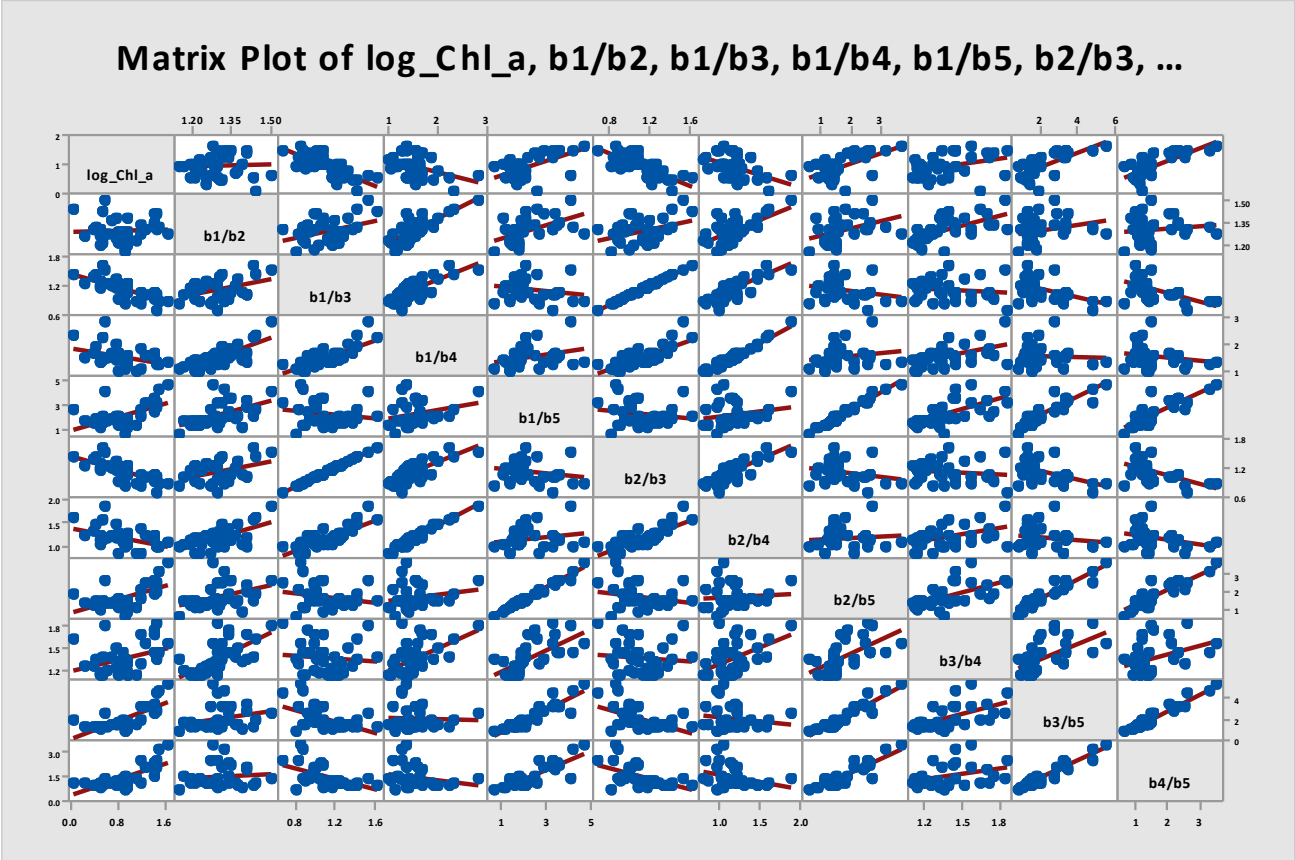


Figure 12: Matrix plot of log(Chl) versus all potential band ratio algorithms for model development. Based on visual evaluation, the matrix plot confirms previous literature, with ratios of b1/b3 (440nm/560nm) and b2/b3 (480nm /560nm) showing strong correlations to measured chlorophyll.

A backwards-stepwise regression model could not be initially developed due to the high correlations between various band ratios (those with common numerator or denominators). Therefore, as a first step, linear regressions were performed between each band ratio and log(chl) to quantitatively assess which bands proved to be the best predictors of chlorophyll concentration. Band ratios b2/b3 (480nm /560nm) and b1/b3 (440nm/560nm) proved significant predictors of chlorophyll concentration (R-squared=64.64% and 54.14%, with p=0) [Table 6].

Next, backward stepwise regressions were performed for the band ratio that showed the most promise as a remote sensing predictor of chlorophyll concentration (b2/b3 and b1/b3). Wind speed, wave height, sample depth, time difference, and river discharge were included as continuous covariates, and sample date was included as a categorical factor. For each step of the regression, the variable with the highest p-value was removed until all variables were significant predictors ($p<.05$) of measured chlorophyll concentration. After completing backwards-stepwise regression, only band ratio proved to be a significant predictor of chlorophyll concentration, as all other variables were removed due to insufficient correlation. Predictors were removed in order as follows: wind speed, wave height, time difference, sample date river discharge, and depth. In addition, polynomial regressions were tested for band ratio b2/b3, none of which resulted in a more significant regression fit. Although the quadratic model resulted in a slightly higher R-squared (64.67%) than the linear model, neither predictors were significant ($p= 0.616$, 0.873) and visual analysis of the data did not reveal a quadratic trend. Finally, an indicator variable was created for the data for site 40 on 6/4/14 due to the extremely low chlorophyll content of the sample (the lowest of all samples 2013-15 for sites 40, 50 and 51). With the inclusion of the indicator variable, the final model achieved in this study [Equation 3] has a R-squared of 71.30% [Figure 13]. Residual plots [Figure 14] demonstrate a normal distribution of residuals, indicating that the model meets regression assumptions. Full regression results are included in the appendix of this paper. Figure 15 demonstrates the success of this model in monitoring bloom events, with the 2015 HAB event clearly visible in the modeled Landsat imagery.

$$\text{Log(Chlorophyll-a)} = 2.738 - 2.084 \log(Rrs_{490nm} / Rrs_{555nm}) \quad \text{Eq. 3}$$

Table 6: Initial linear regression results between Landsat band ratios and log(chl). Band2/Band3 ratio was the most significant predictor of chlorophyll concentration.

Band Ratio	R-sq	R-sq (adj)	*p*
b2/b3	64.64%	63.63%	0
b1/b3	54.14%	52.83%	0
b2/b4	21.86%	19.63%	0.004
b1/b4	14.11%	11.66%	0.026
b1/b5	14.07%	12.18%	0.032
b3/b4	13.76%	11.30%	0.085
b4/b5	10.21%	7.00%	0.264
b2/b5	8.97%	7.23%	0.532
b3/b5	7.12%	5.89%	0.809
b1/b2	0.17%	0.00%	0.828

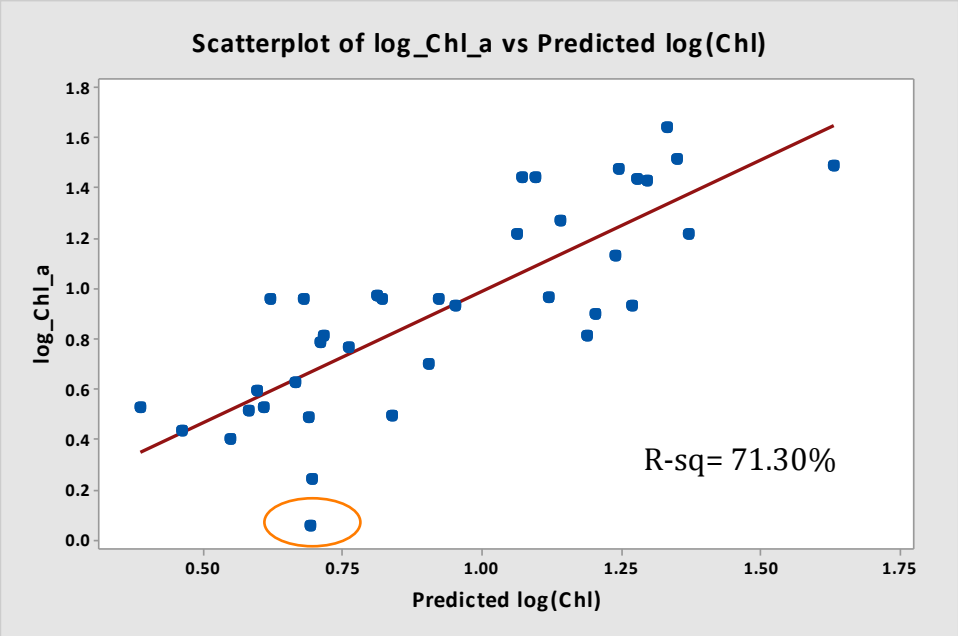


Figure 13: Model-predicted log(Chlorophyll) using Equation 3 and all available sample matchups versus measured log(Chlorophyll) concentration. 6/4/14 Measurement is shown in red circle and given indicator variable in final model.

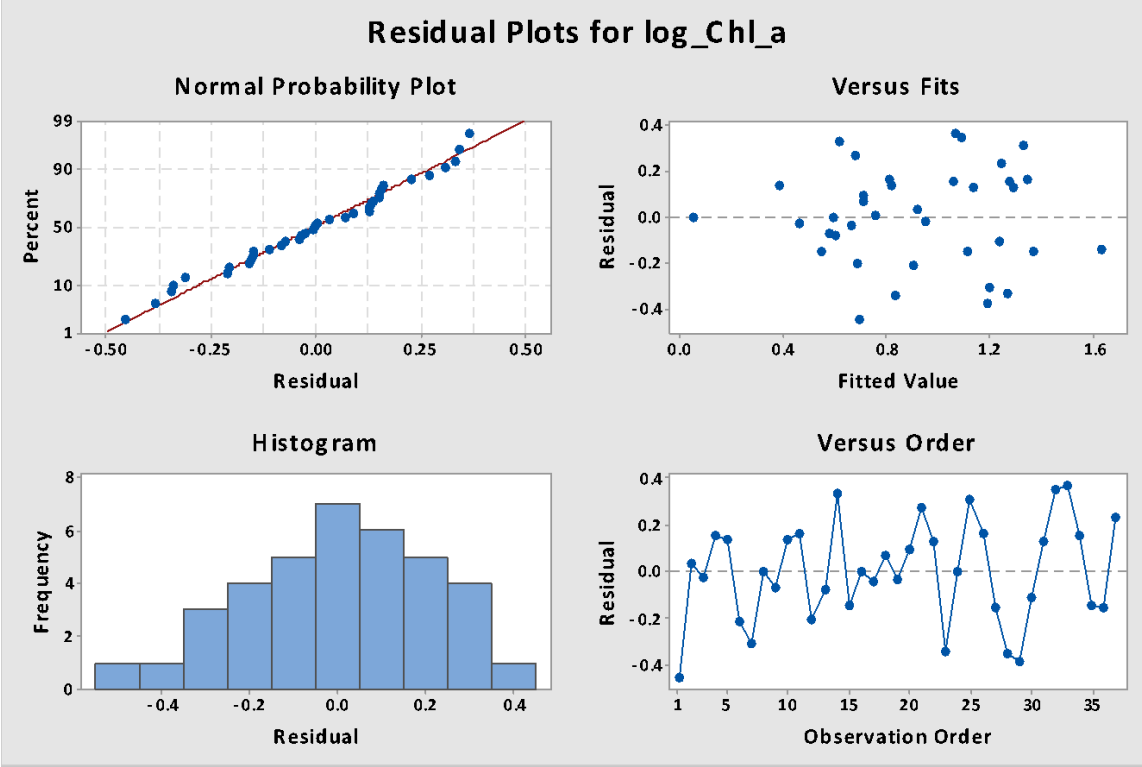


Figure 14: Residual plots for final model, demonstrating normality of residuals and glm assumptions.

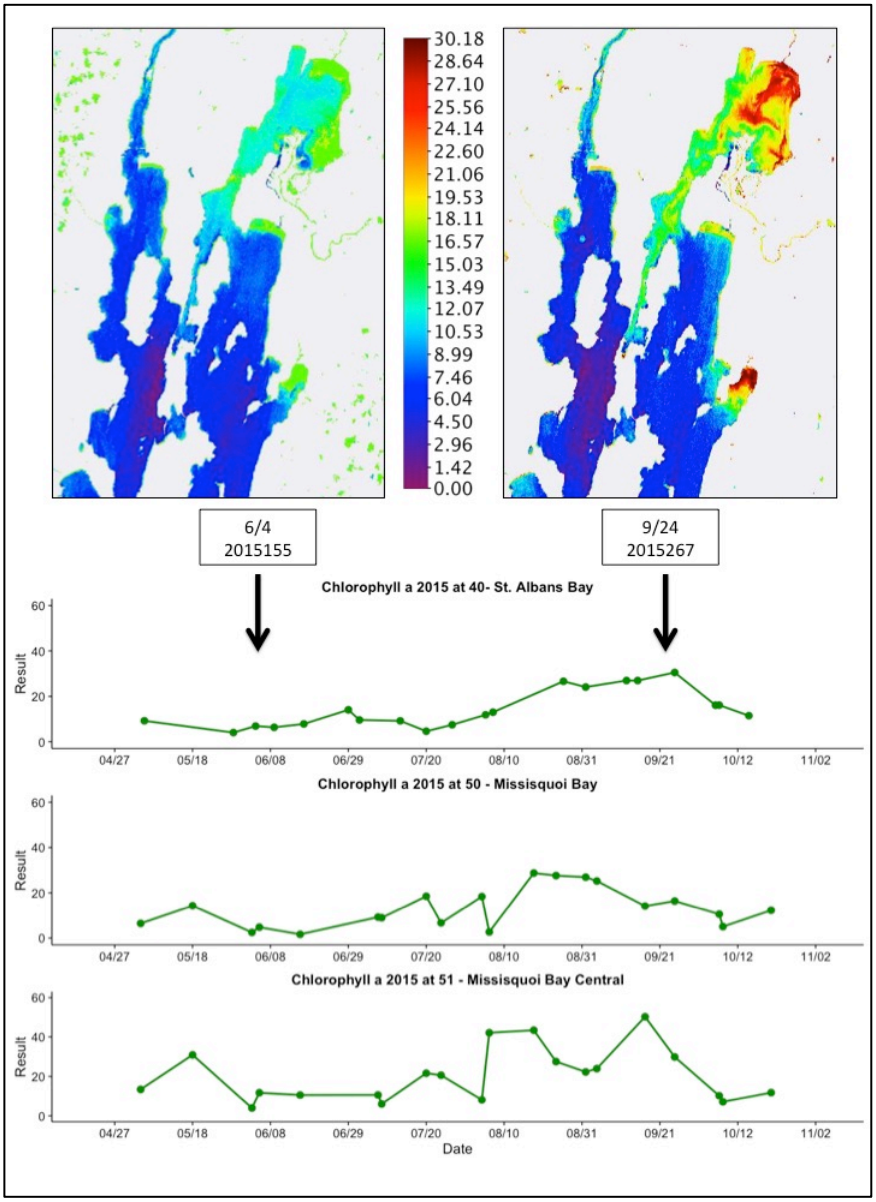


Figure 15: Landsat chlorophyll model applied to June 4th and September 24th, 2014 images, taken during non-bloom and bloom conditions (as defined by *in situ* chlorophyll sampling and VT DEC Algal Tracker "High Alert" status.

3.3 Conclusion

A band ratio bio-optical algorithm was successfully developed for Landsat 8 OLI remote sensing of harmful algal blooms in the Northeastern Arm of Lake Champlain (R-squared =71.3%). The linear model developed in this section was accurate for all 14 Landsat images and unaffected by changes in wind speed, suspended sediment, or other temporally variable factors, demonstrating the strength of the model for continuous and accurate HAB monitoring. This study contributes to the recent developments in Landsat 8 OLI aquatic remote sensing, further illustrating the increased potential of the next generation of Landsat sensors for water quality applications and the benefits of increased radiometric sensitivity in aquatic remote sensing. Furthermore, the atmospheric corrections tested within this study demonstrate the success of the MUMM approach to atmospheric correction over the Case 2 waters of Lake Champlain and the success of Landsat 8 OLI MUMM implementation in NASA's SeaDAS software. With Landsat's relatively high spatial resolution (30m), this model has the potential to dramatically increase the extent of local HAB monitoring, providing a spatial understanding harmful algal bloom extent with each cloud-free Landsat overpass.

SECTION 4: MODIS

4.1 Methods

4.1a Introduction: Sensor Description and Characteristics

The Moderate-Resolution Imaging Spectroradiometer (MODIS) instrument was launched in 1999 on the Terra satellite and in 2002 on the Aqua satellite. MODIS was designed to provide improved monitoring for land, ocean, and atmospheric research, with spectral measurements in 36 bands between 0.405 μm and 14.385 μm, including seven bands designed for land remote sensing, seven bands designed for ocean color, and the other 22 bands designed for the lower atmosphere remote sensing. These bands vary in spatial resolution. Bands 1 and 2 have 250 m resolution, bands 3-7 have 500 m resolution, and bands 8-36 have 1 km resolution [Table 7]. MODIS Terra and Aqua have a viewing swath width of 2,330 km and view the entire surface of the Earth every one to two days. In addition, MODIS has high radiometric sensitivity, with 12-bit quantization of both MODIS sensors. The spectral band range and primary uses of the MODIS data are shown in Table 2.2. MODIS Aqua images were selected for this study due to the degraded response functions (as much as 40%) of the MODIS Terra sensors since initial launch calibrations (NASA MODIS, n.d.).

MODIS's high temporal resolution makes it an ideal sensor for algal bloom events, as MODIS remote sensing has the potential for daily HAB evaluation. MODIS remote sensing models have been applied to marine HABs along both the Florida and Southern California Coast (Hu, 2005; Kahru, 2005; Anderson, 2009). However, perhaps because of its low spatial resolution, applications of MODIS to inland water quality have been limited. A few studies have evaluated the performance of MODIS data to mapping cyanobacteria blooms in the Great Lakes (Weghorst, 2008: Becker, 2009; Zhang, 2012; McCullough, 2012). However this study represents a novel attempt to apply MODIS water quality remote sensing to smaller water bodies and, to this author's knowledge, the smallest scale application of MODIS to aquatic remote sensing to date.

Table 7: MODIS Band Specifications: bandwidth (in nm) and resolution. Bands used in this study highlighted in grey.

Band	Bandwidth	Resolution
1	620 - 670	250 m
2	841 - 876	250 m
3	459 - 479	500 m
4	545 - 565	500 m
5	1230 - 1250	500 m
6	1628 - 1652	500 m
7	2105 - 2155	500 m
8-19	Multispectral	1 km
20-36	Thermal	1 km

4.1b Acquisition, Atmospheric Correction, and Processing:

MODIS Aqua Level 1 images data from May 2013 to October 2015 were downloaded from the Level 1 and Atmosphere Archive and Distribution System (LAADS). These images were visually evaluated for clouds, and 66 cloud free images were retained for processing. MODIS Level 1A data were processed to Level 2 water-leaving radiance using the SeaDAS Multi-Level processor and batch processed on a Linux computer system, with Unix computer code adapted from Zhang (2012) and SeaDAS online materials (NASA SeaDAS, n.d.). Final processing code is included in the Appendix and an outline of processing steps is provided in Figure 16.

Level 1A MODIS images contain raw radiance data for 36 MODIS spectral bands. Prior to processing, all data was subset to a rectangular area covering the Lake Champlain Area defined by the coordinates 45ºN, 43.5ºS, -72ºE and -75ºW in order to reduce runtime. A geolocation data file was then generated for each Level 1A image.

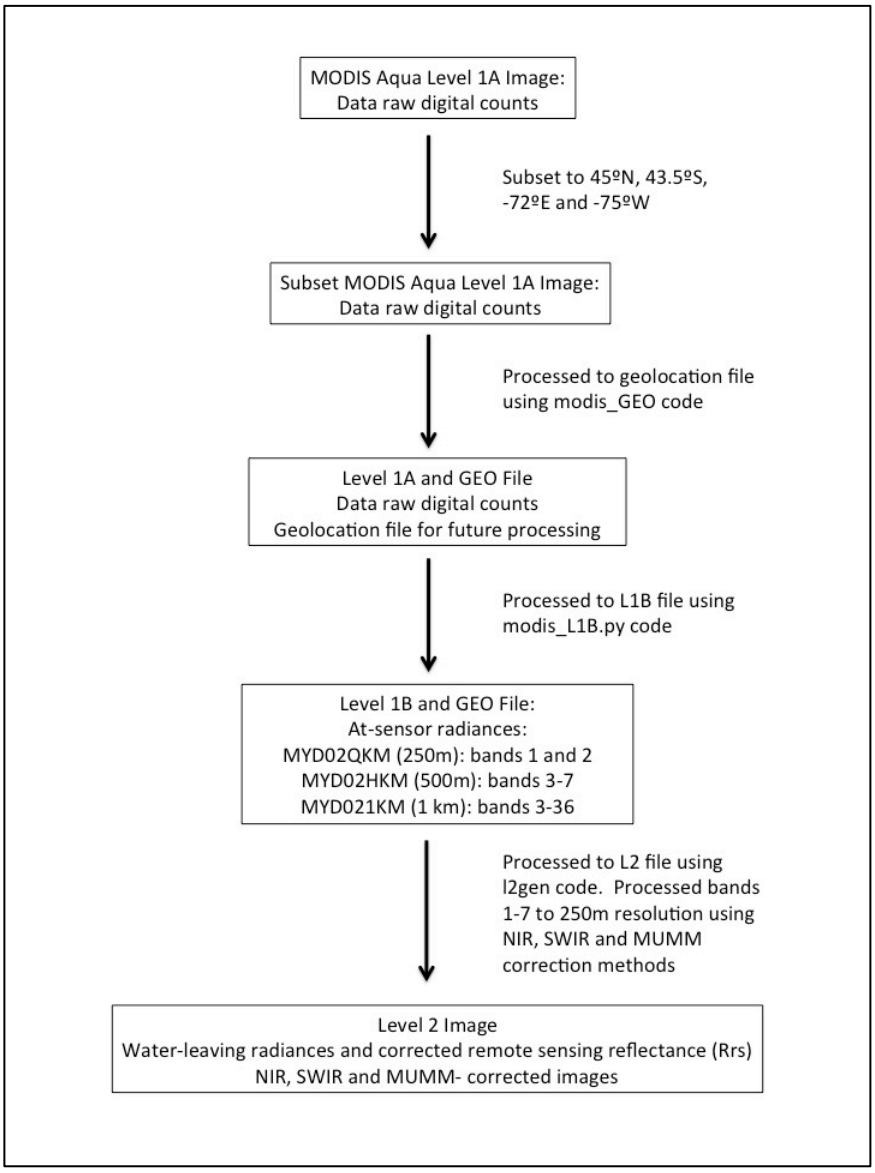
The 66 Level 1A images were then processed to generate geolocation files and Level 1B images using the SeaDAS software. Level 1B images contain at-sensor radiances, calibrated and geolocated based on Level 1A data raw digital counts and calibration look up tables (LUTs) included in the SeaDAS program. Although Level 1B images contain all 36 MODIS bands, the data is split into 3 files based on spatial resolution: MYD02QKM, MYD02HKM and MYD021KM, with resolutions of 250 m, 500 m, and 1 km respectively.

Level 1B at-sensor radiances were then processed to generate Level 2 water-leaving radiances and atmospherically corrected remote sensing reflectance (Rrs) using the SeaDAS l2gen code. The l2gen atmospheric correction module has been recently extended to support not

only the traditional 'ocean color' bands of MODIS, which have 1km resolutions, but also the lower spatial resolution 'land bands' at resolutions of 500 and 250m. SeaDAS' l2gen code has the ability to process images to a 1km, 500m or 250m resolutions, using bilinear interpolation of the lower resolution bands (Franz, 2006).

Because of the small scale of Lake Champlain, images were processed to a 250m resolution. In addition, only bands 1-7 were used in MSL12 processing, as interpolation of higher 1km bands to 250m resolutions resulted in high spectral mixing from shore and lake pixels and lowered spectral purity. MODIS data was processed in the l2gen using NIR, SWIR, and MUMM atmospheric correction models, and results were analyzed to determine the most effective approach.

Figure 16: Flowchart of MODIS processing methods in SeaDAS software.



4.1c *In Situ* Correlation and Analysis

Due to the larger spatial resolution of the MODIS imagery (as compared to Landsat), the embayment of St. Albans Bay (site 40) could not be resolved, and therefore only sampling sites 50 and 51 were included in the model. In addition, only MODIS bands 3 (469 nm), 4 (555 nm), and 1 (645 nm) were included in analysis, as these are the only visible bands with original resolutions of 250 or 500m (the rest of MODIS' multispectral bands having a 1km resolution).

MODIS images were matched to corresponding (within a ± 2 day window) chlorophyll samples from the VT DEC dataset, following the methods used for Landsat algorithm

development in Section 3. In addition, water depth (from VT DEC sampling protocol), wind speed, wave height (from Buoy Station 45166), and river discharge rate (from USGS Gages 04294300 and 04294000) were included in the model. A total of 73 matchups were recorded for the three years of MODIS imagery. The number of matchups was limited by the frequency of *in situ* sampling (rather than the acquisition of MODIS imagery), as only 106 total *in situ* samples were taken at sites 46, 50, and 51 combined over the three years of study. This demonstrates the potential benefit of MODIS algal monitoring, as the rapid revisit time of MODIS has the potential for frequent and continuous monitoring of chlorophyll anomalies.

The quality of data for all 73 in situ matchups was examined prior to regression modeling, and four sample location-date points removed due to contamination from cloud or cloud shadows in the image. MODIS band ratios for the remaining 69 match-ups were calculated in Minitab, and backwards-stepwise regression was run to determine the best model for chlorophyll concentration from MODIS imagery.

Table 8: Data predictors used for MODIS model development

Sample Chl-a	In situ Chl-a measurements in µg/L from LTMP sampling
Sample Date	Date of in situ sampling MM/DD/YY
Time lag	Time difference: Date of image acquisition - Date of in situ sampling
Wind speed	Wind speed at Bouy Station 45166 (m/s)
Wave height	Average of the highest one-third of all of the wave heights (in m) during the 20-minute sampling period at Bouy Station 45166
Mean river discharge	Discharge rates from USGS gages on Pike and Mississquoi Rivers (in cubic ft/sec), averaged for 7 days prior image acquisition from
Sample location depth	Depth of sample location (m) from LTMP protocol
B1	MODIS Band 1, 645 nm (Rrs)
B3	MODIS Band 3, 469 nm (Rrs)
B4	MODIS Band 1, 555 nm (Rrs)

4.2 RESULTS

4.2a Evaluation of Atmospheric Correction Models

Figure 17 shows a subset of the MODIS scene for the Northeast corner of Lake Champlain for July 21, 2013 after processing using NIR (bands 4 and 5), SWIR (bands 6 and 7), and MUMM atmospheric correction. Spectra were extracted within a 3x3 buffer around each station pixel [Figure 17]. Results were consistent with Landsat 8 atmospheric correction testing, with both NIR and SWIR methods resulting in an overcorrection of aerosol radiances, and therefore negative reflectance values within the visible range.

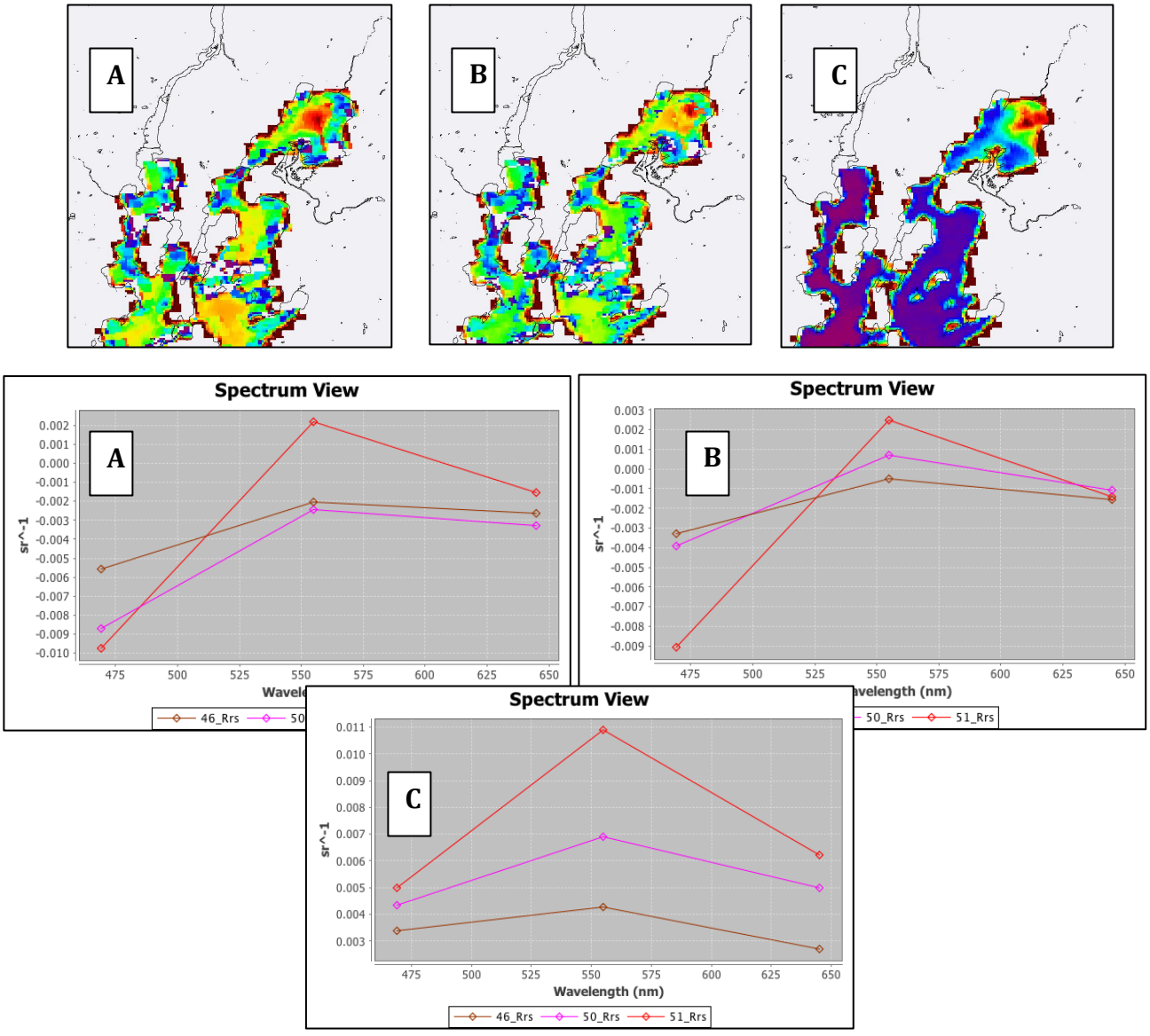


Figure 17: SeaDAS MODIS atmospheric correction results: A) SeaDAS NIR correction results for Rrs 555 nm and spectra; b) SeaDAS SWIR correction results for Rrs 555 nm and spectra; d) SeaDAS MUMM correction results for Rrs 561 nm and spectra

4.2b Regression Analysis

Band ratios were calculated in Minitab to normalize scene-dependent reflectance. Linear regressions were then performed between each band ratio and log(chlorophyll concentration) in order to assess the relative predictive power of each band combination. The ratio of band 4/band 3 (555/469) had the highest correlation to *in situ* measurements and was used in further model development [Table 9].

Table 9: Initial MODIS band ratio regression results

Band Ratio	R-sq	R-sq (adj)	*p*
555/469	34.65%	33.47%	0.000
555/645	15.67%	13.95%	0.005
645/469	13.86%	12.29%	0.004

Backwards-stepwise regression was run, with band ratio 555/469, wind speed, wave height, river discharge, and depth as continuous predictors, and sample date and sample site as categorical factors. After removing each variable with the highest p-value, only band ratio was found to be a significant predictor of chlorophyll-a, with other variables removed stepwise as follows: wave height, sample date, wind speed, sample site, depth, and river discharge.

The band ratio (555/469) was then log transformed in order to increase the spread of the data, and backwards stepwise regression was re-run with all included predictors. Within this model, band ratio still proved the only significant predictor (p=0), and the R-squared increased slightly (R-sq = 35.27). The log-transform also revealed a trend between wind speed and model predictions, and the six points with negative log(555/469) also had the highest recorded *in situ* wind speeds. An indicator variable was created for all observations with a wind speed larger than 7 m/s. The inclusion of this indicator variable in the model increased R-squared to 38.34%, with both log(555/469) and the wind speed indicator as significant predictors. Finally, a quadratic model was tested [Equation 4] and found to slightly increase the predictive power of the model to an R-sq of 39.44%. This model was used in final image analyses [Figure 18]. Residual plots [Figure 19] demonstrate a normal distribution of residuals, indicating that the model meets regression assumptions. Full regression results are included in the appendix of this

paper. Figure 20 demonstrates the success of this model in monitoring bloom events, with the 2015 HAB event (for comparison, the same dates as the Landsat example in Section 2), clearly visible.

$$\text{Log(Chlorophyll-a)} = 0.663 + 3.41 \log(Rrs_{555nm}/Rrs_{469nm}) - 2.56 (\log(Rrs_{555nm}/Rrs_{469nm}))^2 \quad \text{Eq. 4}$$

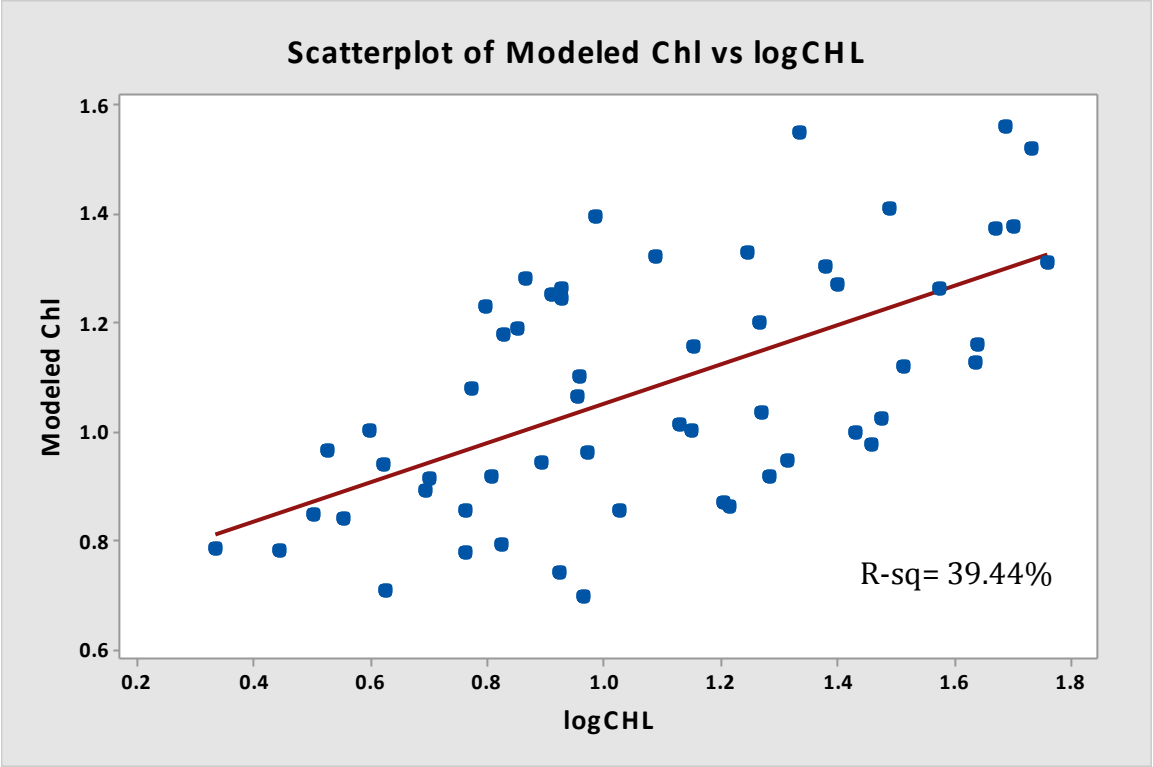


Figure 18: Model-predicted log(Chlorophyll) using Equation 4 for MODIS and all available sample matchups versus measured log(Chlorophyll) concentration.

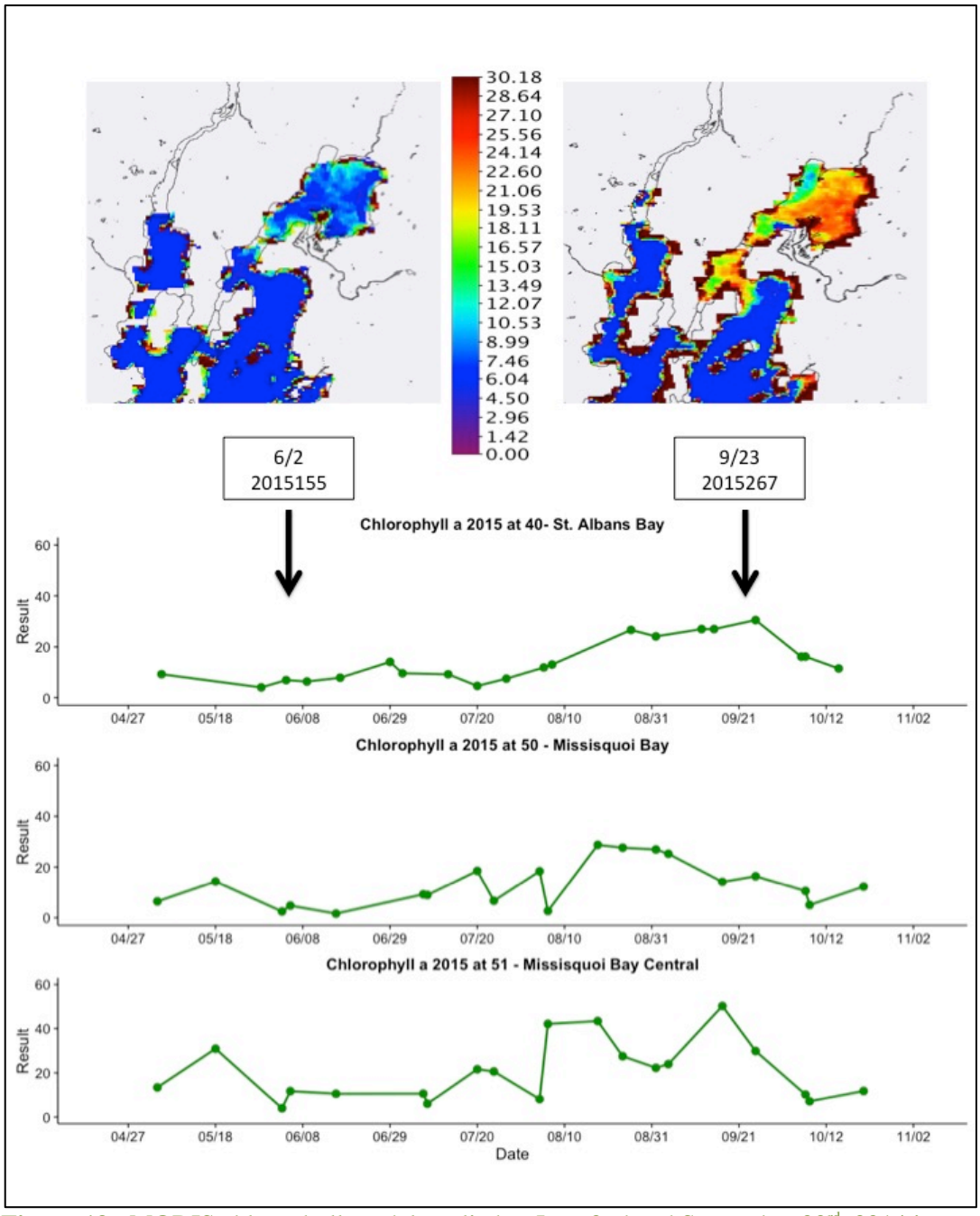


Figure 19: MODIS chlorophyll model applied to June 2nd and September 23rd, 2014 images, taken during non-bloom and bloom conditions (as defined by *in situ* chlorophyll sampling and VT DEC Algal Tracker "High Alert" status). Bottom: plots of *in situ* chlorophyll samples from LTMP sampling showing seasonal variability.

4.3 Conclusion

A band ratio bio-optical algorithm was successfully developed for MODIS remote sensing of harmful algal blooms in the Northeastern Arm of Lake Champlain (R-squared 39.44%). The quadratic model developed in this section was accurate for nearly all 69 MODIS images, however under high wind speeds (over 7 m/s) the model failed, suggesting that specular reflection from waves at high winds may result in poor chlorophyll retrieval. The model was by all other temporally variable factors, however demonstrating the strength of the model for continuous HAB monitoring. This study demonstrates the successful application of MUMM atmospheric correction for MODIS Case 2 aquatic remote sensing. Furthermore, the application of MODIS imagery for HAB remote sensing represents the smallest scale application of MODIS imagery for aquatic remote sensing to date. The success of MODIS for Lake Champlain remote sensing has important implications for future research, suggesting that, even with its limited spectral resolution at 250 and 500m, MODIS has exciting potential for small-scale aquatic remote sensing. This model has the potential to dramatically increase the frequency of HAB monitoring in Lake Champlain. With MODIS' revisit time of 1-2 days, this model suggests the potential application of MODIS data to rapid response monitoring of HAB events.

Summary and Future Work

This project demonstrates the ability of both Landsat 8 and MODIS Aqua for remote sensing of Harmful Algal Blooms in Lake Champlain, VT. The application of Landsat 8 to turbid, inland water bodies is a recent development in remote sensing, as Landsat 8 was launched in 2013 and atmospheric correction wasn't implemented in NASA processing software until September of 2015. The linear Landsat 8 models developed in this study represent a robust and accurate (Rsq= 71.3%) application of Landsat data for harmful algal bloom remote sensing. Furthermore, this study represents the smallest scale aquatic application of MODIS remote sensing. The MODIS band ratio algorithm developed in this study further contributes to current HAB alert systems, with the potential for near-daily monitoring of chlorophyll concentration and bloom events within turbid lake waters.

Although the general linear modeling techniques applied in this study proved successful, a number of other statistical techniques could be applied to the data for further analysis. Principle component analysis (PCA), Bayesian hierarchical regression, and neural network analysis have all proved successful in past aquatic remote sensing studies. A comparison between these various techniques could prove fruitful in continuing to develop accurate algal bloom remote sensing models.

Furthermore, although this study attempted to address the confounding variable of suspended solids through the quantification of local river discharge, more precise measurements of total suspended solids (TSS) could allow for a better understanding of the various factors influencing observed remote sensing reflectance. For future model development, it would be beneficial to conduct a cruise aligned with Landsat and MODIS overpass dates, measuring TSS and turbidity, as well as chlorophyll, for model development. Given the promising results of this study, such a campaign has great promise for further Lake Champlain HAB remote sensing.

The success of MODIS models in this study suggests that future work should continue to apply this sensor in small-scale aquatic remote sensing. Given the long time series available for MODIS data (2002-2016), this study has exciting implications for future work. For example, future work could expand the results in this study to all available MODIS dates, allowing for a more comprehensive time series of HAB events in Lake Champlain and other inland water bodies affected by the toxic events.

Acknowledgements

I am grateful to my advisor Ron Smith, not just for his help throughout this project but for originally getting me hooked on remote sensing. Thank you to both Ron and Larry Bonneau for your energy and enthusiasm for remote sensing here at Yale, for keeping the Yale Center for Earth Observation up and running so smoothly, and for you patience teaching every student in Observing the Earth from Space everything from the basics of spectral analysis to the quirks of ENVI. Larry, thank you for always helping me trouble-shoot any tech issue and supporting my use of the Spectrometer. Your enthusiasm for remote sensing made long hours in the CEO easy and worthwhile. I would also like to thank Dana Tomlin for his courses in geospatial software design and all that he has taught me about GIS and geospatial data, and NASA SARP for continuing my remote sensing education over the past summer. Last but not least, thank you to Cody Kahoe for supporting me through long hours in lab and helping me organize spreadsheets for hours on end.

Works Cited

Anderson, Clarissa R., David A. Siegel, Raphael M. Kudela, and Mark A. Brzezinski. "Empirical models of toxigenic Pseudo-nitzschia blooms: potential use as a remote detection tool in the Santa Barbara Channel." *Harmful Algae* 8, no. 3 (2009): 478-492.

Becker, Richard H., Mohamed I. Sultan, Gregory L. Boyer, Michael R. Twiss, and Elizabeth Konopko. "Mapping cyanobacterial blooms in the Great Lakes using MODIS." *Journal of Great Lakes Research* 35, no. 3 (2009): 447-453.

Berkman, J., and M. Canova. "Algal Biomass Indicators: US Geological Survey National Field Manual for the Collection of Water-Quality Data, book 9, chap." *A7* 5 (2007).

Bonansea, Matias, María Claudia Rodriguez, Lucio Pinotti, and Susana Ferrero. "Using multi-temporal Landsat imagery and linear mixed models for assessing water quality parameters in Río Tercero reservoir (Argentina)." *Remote Sensing of Environment* 158 (2015): 28-41.

Boyer, Gregory L., Mary C. Watzin, Angela D. Shambaugh, Michael F. Satchwell, Barry H. Rosen, and Timothy Mihuc. "The occurrence of cyanobacterial toxins in Lake Champlain." In *Lake Champlain: Partnerships and research in the new millennium*, pp. 241-257. Springer US, 2004.

Cannizzaro, Jennifer Patch, and Kendall L. Carder. "Estimating chlorophyll a concentrations from remote-sensing reflectance in optically shallow waters." *Remote Sensing of Environment* 101, no. 1 (2006): 13-24.

Chipman, Jonathan W., Thomas M. Lillesand, Jeffrey E. Schmaltz, Jill E. Leale, and Mark J. Nordheim. "Mapping lake water clarity with Landsat images in Wisconsin, USA." *Canadian journal of remote sensing* 30, no. 1 (2004): 1-7.

Clarke, George L., Gifford C. Ewing, and Carl J. Lorenzen. "Spectra of backscattered light from the sea obtained from aircraft as a measure of chlorophyll concentration." *Science* 167, no. 3921 (1970): 1119-1121.

Codd, Geoffrey A., Louise F. Morrison, and James S. Metcalf. "Cyanobacterial toxins: risk management for health protection." *Toxicology and applied pharmacology* 203, no. 3 (2005): 264-272.

Concha, Javier A., and John R. Schott. "Atmospheric correction for Landsat 8 over case 2 waters." *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2015.

Dekker, A. G., and S. W. M. Peters. "The use of the Thematic Mapper for the analysis of eutrophic lakes: a case study in the Netherlands." *International Journal of Remote Sensing* 14, no. 5 (1993): 799-821.

Doerffer, R., and H. Schiller. "The MERIS Case 2 water algorithm." *International Journal of Remote Sensing* 28, no. 3-4 (2007): 517-535.

Franz, Bryan A., P. Jeremy Werdell, Gerhard Meister, Ewa J. Kwiatkowska, Sean W. Bailey, Ziauddin Ahmad, and Charles R. McClain. "MODIS land bands for ocean remote sensing applications." In *Proc. Ocean Optics XVIII, Montreal, Canada*, vol. 10. 2006.

Franz, Bryan Alden, et al. "Ocean color measurements from Landsat-8 OLI using SeaDAS." (2014).

Gerace, Aaron, and John Schott. "The increased potential for the Landsat Data Continuity Mission to contribute to case 2 water quality studies." *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2009.

Gitelson, Anatoly A., Daniela Gurlin, Wesley J. Moses, and Tadd Barrow. "A bio-optical algorithm for the remote estimation of the chlorophyll-a concentration in case 2 waters." *Environmental Research Letters* 4, no. 4 (2009): 045003.

Gitelson, Anatoly A., John F. Schalles, and Christine M. Hladik. "Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study." *Remote Sensing of Environment* 109, no. 4 (2007): 464-472.

Gons, Herman J., Machteld Rijkeboer, and Kevin G. Ruddick. "A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters." *Journal of Plankton Research* 24, no. 9 (2002): 947-951.

Gordon, Howard R., and André Y. Morel. *Remote assessment of ocean color for interpretation of satellite visible imagery: A review*. Vol. 4. Springer Science & Business Media, 2012.

Gordon, Howard R., and Kenneth J. Voss. "MODIS Normalized Water-leaving Radiance, Algorithm Theoretical Basis Document (MOD 18), version 4." *NASA Contract Number NAS503163* (1999).

Gordon, Howard R., and Menghua Wang. "Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm." *Applied optics* 33, no. 3 (1994): 443-452.

Han, Luoheng, and Donald C. Rundquist. "Comparison of NIR/RED ratio and first derivative of reflectance in estimating algal-chlorophyll concentration: A case study in a turbid reservoir." *Remote sensing of Environment* 62, no. 3 (1997): 253-261.

Hu, Chuanmin, Frank E. Muller-Karger, Charles Judd Taylor, Kendall L. Carder, Christopher Kelble, Elizabeth Johns, and Cynthia A. Heil. "Red tide detection and tracing using MODIS fluorescence data: A regional example in SW Florida coastal waters." *Remote Sensing of Environment* 97, no. 3 (2005): 311-321.

Hu, Chuanmin, Frank E. Muller-Karger, Serge Andrefouet, and Kendall L. Carder. "Atmospheric correction and cross-calibration of LANDSAT-7/ETM+ imagery over aquatic environments: A multiplatform approach using SeaWiFS/MODIS." *Remote Sensing of Environment* 78, no. 1 (2001): 99-107.

Isenstein, Elizabeth M., and Mi-Hyun Park. "Assessment of nutrient distributions in Lake Champlain using satellite remote sensing." *Journal of Environmental Sciences* 26, no. 9 (2014): 1831-1836.

Jensen, John R. *Remote sensing of the environment: An earth resource perspective 2/e*. Pearson Education India, 2009.

Kahru, Mati, B. Greg Mitchell, and Anibal Diaz. "Using MODIS medium-resolution bands to monitor harmful algal blooms." In *Optics & Photonics 2005*, pp. 58850K-58850K. International Society for Optics and Photonics, 2005.

Kaufman, Y. J., D. Tanré, H. R. Gordon, T. Nakajima, J. Lenoble, R. Frouin, H. Grassl, B. M. Herman, M. D. King, and P. M. Teillet. "Passive remote sensing of tropospheric aerosol and atmospheric correction for the aerosol effect." *Journal of Geophysical Research: Atmospheres* 102, no. D14 (1997): 16815-16830.

Kloiber, Steven M., Patrick L. Brezonik, and Marvin E. Bauer. "Application of Landsat imagery to regional-scale assessments of lake clarity." *Water Research* 36, no. 17 (2002): 4330-4340.

Kutser, Tiit, Liisa Metsamaa, Ele Vahtmäe, and Niklas Strömbeck. "Suitability of MODIS 250 m resolution band data for quantitative mapping of cyanobacterial blooms." *Proc. Estonian Acad. Sci. Biol. Ecol* 55, no. 4 (2006): 318-328.

Lake Champlain Basin Program. 2004. "Lake Champlain Atlas." < http://atlas.lcbp.org>

Lake Champlain Long-Term Water Quality and Biological Monitoring Program. 2015. "Project Description"
Lake Champlain Long-Term Water Quality and Biological Monitoring Program. 2015. "2015 Quality Assurance Project Plan"
Lathrop, R. G. "Landsat Thematic Mapper monitoring of turbid inland water quality." *Photogrammetric Engineering and Remote Sensing;(United States)* 58 (1992).
Lévesque, Benoît, Marie-Christine Gervais, Pierre Chevalier, Denis Gauvin, Elhadji Anassour-Laouan-Sidi, Suzanne Gingras, Nathalie Fortin, Geneviève Brisson, Charles Greer, and David Bird. "Prospective study of acute health effects in relation to exposure to cyanobacteria." *Science of the Total Environment* 466 (2014): 397-403.
Levine, Suzanne N., et al. "The eutrophication of Lake Champlain's northeastern arm: Insights from paleolimnological analyses." *Journal of Great Lakes Research* 38 (2012): 35-48.
Li, Rong-Rong, Yoram J. Kaufman, Bo-Cai Gao, and Curtiss O. Davis. "Remote sensing of suspended sediments and shallow coastal waters." *Geoscience and Remote Sensing, IEEE Transactions on* 41, no. 3 (2003): 559-566.
Ma, Ronghua, and Jinfang Dai. "Investigation of chlorophyll-a and total suspended matter concentrations using Landsat ETM and field spectral measurement in Taihu Lake, China." *International Journal of Remote Sensing* 26, no. 13 (2005): 2779-2795.
Marion, Jason W., Jiyoung Lee, J. R. Wilkins III, Stanley Lemeshow, Cheonghoon Lee, Evan J. Waletzko, and Timothy J. Buckley. "In vivo phycocyanin flourometry as a potential rapid screening tool for predicting elevated microcystin concentrations at eutrophic lakes." *Environmental science & technology* 46, no. 8 (2012): 4523-4531.
McCullough, Ian M., Cynthia S. Loftin, and Steven A. Sader. "High-frequency remote monitoring of large lakes with MODIS 500m imagery." *Remote Sensing of Environment* 124 (2012): 234-241.
Mihuc, T. B., G. L. Boyer, M. F. Satchwell, M. Pellam, J. Jones, J. Vasile, A. Bouchard, and R. Bonham. "2002 phytoplankton community composition and cyanobacterial toxins in Lake Champlain, USA." *Internationale Vereinigung fur Theoretische und Angewandte Limnologie Verhandlungen* 29, no. 1 (2005): 328-333.
Miller, Peter I., Jamie D. Shutler, Gerald F. Moore, and Steve B. Groom. "SeaWiFS discrimination of harmful algal bloom evolution." *International Journal of Remote Sensing* 27, no. 11 (2006): 2287-2301.
Morel, Anclré, and Louis Prieur. "Analysis of variations in ocean color." *Limnology and oceanography* 22, no. 4 (1977): 709-722.
Moses, Wesley J., Anatoly A. Gitelson, Sergey Berdnikov, and V. Povazhnyy. "Estimation of chlorophyll-a concentration in case II waters using MODIS and MERIS data—successes and challenges." *Environmental Research Letters* 4, no. 4 (2009): 045005.
Myer, Glenn E., and Gerhard K. Gruendling. *Limnology of Lake Champlain*. Vol. 30. Lake Champlain Basin Study, New England River Basins Commission, 1979.
Myers, G.E., Gruendling, G.K., 1979. Limnology of Lake Champlain. Champlain Basin

NASA Ocean Color Data. NASA. http://oceandata.sci.gsfc.nasa.gov
Nieke, B., R. Reuter, R. Heuermann, H. Wang, M. Babin, and J. C. Therriault. "Light absorption and fluorescence properties of chromophoric dissolved organic matter (CDOM), in the St. Lawrence Estuary (Case 2 waters)." *Continental Shelf Research* 17, no. 3 (1997): 235-252.
O'Reilly, John E., Stephane Maritorena, B. Greg Mitchell, David A. Siegel, Kendall L. Carder, Sara A. Garver, Mati Kahru, and Charles McClain. "Ocean color chlorophyll algorithms

for SeaWiFS." *Journal of Geophysical Research: Oceans* 103, no. C11 (1998): 24937-24953.

Paerl, Hans W., Nathan S. Hall, and Elizabeth S. Calandrino. "Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change." *Science of the Total Environment* 409.10 (2011): 1739-1745.

Pahlevan, Nima, and John R. Schott. "Characterizing the relative calibration of Landsat-7 (ETM+) visible bands with Terra (MODIS) over clear waters: The implications for monitoring water resources." *Remote Sensing of Environment* 125 (2012): 167-180.

Pouria, Shideh, Al de Andrade, J. Barbosa, R. L. Cavalcanti, V. T. S. Barreto, C. J. Ward, W. Preiser, Grace K. Poon, G. H. Neild, and G. A. Codd. "Fatal microcystin intoxication in haemodialysis unit in Caruaru, Brazil." *The Lancet* 352, no. 9121 (1998): 21-26.

Pozdnyakov, D. V., et al. "MODIS evidences the river run-off impact on the Kara Sea trophy." *International Journal of Remote Sensing* 26.17 (2005): 3641-3648.

Qi, Lin, Chuanmin Hu, Hongtao Duan, Jennifer Cannizzaro, and Ronghua Ma. "A novel MERIS algorithm to derive cyanobacterial phycocyanin pigment concentrations in a eutrophic lake: Theoretical basis and practical considerations." *Remote Sensing of Environment* 154 (2014): 298-317.

Rao, P. V., N. Gupta, A. S. Bhaskar, and R. Jayaraj. "Toxins and bioactive compounds from cyanobacteria and their implications on human health." *Journal of environmental biology/Academy of Environmental Biology, India* 23, no. 3 (2002): 215-224.

Rinta-Kanto, J. M., A. J. A. Ouellette, G. L. Boyer, M. R. Twiss, T. B. Bridgeman, and S. W. Wilhelm. "Quantification of toxic Microcystis spp. during the 2003 and 2004 blooms in western Lake Erie using quantitative real-time PCR." *Environmental science & technology* 39, no. 11 (2005): 4198-4205.

Rosen, B. H., A. Shambaugh, M. Watzin, G. Boyer, F. Smith, L. Ferber, C. Eliopoulos, and P. Stangel. "Evaluation of potential blue-green algal toxins in Lake Champlain." *Lake Champlain Basin Program, Grand Isle, Vermont* (2001).

Ruddick, Kevin George, Fabrice Ovidio, and Machteld Rijkeboer. "Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters." *Applied optics* 39, no. 6 (2000): 897-912.

Ruddick, Kevin George, Fabrice Ovidio, and Machteld Rijkeboer. "Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters." *Applied optics* 39, no. 6 (2000): 897-912.

Simis, Stefan GH, Antonio Ruiz-Verdú, Jose Antonio Domínguez-Gómez, Ramón Peña-Martinez, Steef WM Peters, and Herman J. Gons. "Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass." *Remote Sensing of Environment* 106, no. 4 (2007): 414-427.

Smeltzer, Eric, Angela d Shambaugh, and Pete Stangel. "Environmental change in Lake Champlain revealed by long-term monitoring." *Journal of Great Lakes Research* 38 (2012): 6-18.

Stadelmann, Teresa H., Patrick L. Brezonik, and Steven Kloiber. "Seasonal patterns of chlorophyll a and Secchi disk transparency in lakes of East-Central Minnesota: Implications for design of ground-and satellite-based monitoring programs." *Lake and Reservoir Management* 17, no. 4 (2001): 299-314.

Study. New England River Basins Commission, Burlington, VT, p. 407.

Torbick, Nathan, and Megan Corbiere. "A Multiscale Mapping Assessment of Lake Champlain Cyanobacterial Harmful Algal Blooms." *International journal of environmental research and public health* 12, no. 9 (2015): 11560-11578.

Torbick, Nathan, Sarah Hession, Stephen Hagen, Narumon Wiangwang, Brian Becker, and Jiaguo Qi. "Mapping inland lake water quality across the Lower Peninsula of Michigan using Landsat TM imagery." *International journal of remote sensing* 34, no. 21 (2013): 7607-7624.

Trescott, Adam, Elizabeth Isenstein, and Mi-Hyun Park. "Remote sensing of cyanobacterial blooms in Lake Champlain, USA." *Water Science and Technology: Water Supply* 13, no. 5 (2013): 1402-1409.

U.S. Geological Survey. "Explore Map of USGS Gages in the Lake Champlain Watershed." <http://vt.water.usgs.gov/echo_gage/basin_map.htm>

Vanhellemont, Quinten, and Kevin Ruddick. "Advantages of high quality SWIR bands for ocean colour processing: Examples from Landsat-8." *Remote Sensing of Environment* 161 (2015): 89-106.

Vanhellemont, Quinten, and Kevin Ruddick. "Landsat-8 as a precursor to Sentinel-2: Observations of human impacts in coastal waters." *ESA Special Publication* 726 (2014).

Vanhellemont, Quinten, and Kevin Ruddick. "Turbid wakes associated with offshore wind turbines observed with Landsat 8." *Remote Sensing of Environment* 145 (2014): 105-115.

Vanhellemont, Quinten, et al. "Atmospheric correction of Landsat-8 imagery using SeaDAS." *ESA Special Publication* 726 (2014).

Vermote, E. F., N. El Saleous, C. O. Justice, Y. J. Kaufman, J. L. Privette, L. Remer, J. C. Roger, and D. Tanre. "Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation." *Journal of Geophysical Research: Atmospheres* 102, no. D14 (1997): 17131-17141.

Vincent, Robert K., Xiaoming Qin, R. Michael L. McKay, Jeffrey Miner, Kevin Czajkowski, Jeffrey Savino, and Thomas Bridgeman. "Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie." *Remote Sensing of Environment* 89, no. 3 (2004): 381-392.

Wang, Menghua, and Wei Shi. "The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing." *Optics Express* 15, no. 24 (2007): 15722-15733.

Watzin, Mary C., Emily Brines Miller, Angela D. Shambaugh, and Meghan A. Kreider. "Application of the WHO alert level framework to cyanobacterial monitoring of Lake Champlain, Vermont." *Environmental toxicology* 21, no. 3 (2006): 278-288.

Weghorst, Pamela Leigh. "MODIS algorithm assessment and principal component analysis of chlorophyll concentration in Lake Erie." PhD diss., Kent State University, 2008.

Wheeler, Sarah M., Leslie A. Morrissey, Suzanne N. Levine, Gerald P. Livingston, and Warwick F. Vincent. "Mapping cyanobacterial blooms in Lake Champlain's Missisquoi Bay using QuickBird and MERIS satellite data." *Journal of Great Lakes Research* 38 (2012): 68-75.

Wynne, T. T., et al. "Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes." *International Journal of Remote Sensing* 29.12 (2008): 3665-3672.

Yacobi, Yosef Z., Anatoly Gitelson, and Meir Mayo. "Remote sensing of chlorophyll in Lake Kinneret using highspectral-resolution radiometer and Landsat TM: spectral features of reflectance and algorithm development." *Journal of Plankton Research* 17, no. 11 (1995): 2155-2173.

Zhang, Li. "Remote sensing of water quality in Lake Erie using MODIS imagery data." PhD diss., The Ohio State University, 2012.

Appendix 1: Landsat Regression Results

Regression Analysis: log_Chl_a versus b2/b3, 6/4/14

Method

Categorical predictor coding (1, 0)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	4.1238	2.06188	42.23	0.000
b2/b3	1	3.3636	3.36355	68.90	0.000
6/4/14	1	0.3853	0.38528	7.89	0.008
Error	34	1.6599	0.04882		
Total	36	5.7837			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.220954	71.30%	69.61%	*

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	2.738	0.220	12.47	0.000	
b2/b3	-2.084	0.251	-8.30	0.000	1.02
6/4/14					
1	-0.635	0.226	-2.81	0.008	1.02

Regression Equation

6/4/14

0 log_Chl_a = 2.738 - 2.084 b2/b3

1 log_Chl_a = 2.103 - 2.084 b2/b3

Fits and Diagnostics for Unusual Observations

Obs	log_Chl_a	Fit	Resid	Std Resid	
1	0.241	0.695	-0.454	-2.10	R
8	0.057	0.057	0.000	*	X

R Large residual
X Unusual X

Appendix 2: MODIS SeaDAS Processing Code

Adapted from Zhang (2012) and SeaDAS online materials (NASA SeaDAS, n.d.).

```
########################################
###Spatially subset MODIS files and generate associated geo files for
MODIS L1A images with MODIS_geo code
########################################
datDir=“/Desktop/Lake_Champlain_water_quality/L1ArawAqua”
datoutDir="/Desktop/Lake_Champlain_water_quality/L1AsubAqua”
cd $datDir
for i in *.L1A_LAC
do
a="$i"
echo $a
b="${a:0:23}"".geo"
export geofile="$b"
echo $geofile
#generate L1A *.geo file into the datDir folder
$seadas/ocssw/run/scripts/modis_GEO.py “$i” -o
"${datDir}/${geofile}"
#generate *sub.hdf and *sub.geo files
SWlon=-75
SWlat=43.5
NElon=-72
NElat=45
OutL1Asub="${a:0:23}""sub.L1A_LAC"
OurL1Asubgeo=“${a:0:23}""sub.geo”
echo $OutL1Asub
echo $OutL1Asubgeo
$seadas/ocssw/run/scripts/modis_L1A_extract.py “$i” “$geofile”
“$SWlon”
“$SWlat” “$NElon” “$NElat” “${datoutDir}/${OutL1Asub}”
“${datoutDir}/${OutL1Asubgeo}”
done
########################################


########################################
###Generate MODIS L1B files for 1KM HKM QKM Resolutions using
modis_L1B.py code
########################################

datDir="/Desktop/Lake_Champlain_water_quality/L1AsubAqua"
outDir=“/Desktop/Lake_Champlain_water_quality/L1BAqua"
cd $datDir
for i in *.hdf
do
 a="$i"
 echo $a
```

```
 L1Asub="${a:0:23}""sub.hdf"
 L1Asubgeo="${a:0:23}""sub.geo"
 OutL1B1KM="${a:0:23}"".L1B.1KM"
 OutL1BHKM="${a:0:23}"".L1B.HKM"
 OutL1BQKM="${a:0:23}"".L1B.QKM"
 echo $L1Asub
 echo $L1Asubgeo
 echo $OutL1B1KM
 echo $OutL1BHKM
 echo $OutL1BQKM
 $seadas/ocssw/run/scripts/modis_L1B.py "$L1Asub" "$L1Asubgeo" -o
"${outDir}/${OutL1B1KM}" -h "${outDir}/${OutL1BHKM}" -q
"${outDir}/${OutL1BQKM}"
done

########################################

########################################
###Generate MODIS L2 using l2gen code
########################################

### using MUMM Atmospheric Correction

datDir=/Desktop/Lake_Champlain_water_quality/L1BAqua
geoDir=/Desktop/Lake_Champlain_water_quality/L1AsubAqua
outDir=/Desktop/Lake_Champlain_water_quality/L2Aqua.MUMM.cldmsk
cd $datDir
for i in *.1KM
do
 a="$i"   #a is the file for L1B 1KM
 echo $a
 b="${a:0:23}""sub.geo"
export geofile="$b"
 echo $geofile
 OutL2NWLR="${a:0:23}"".nwlr.hdf"
 OutL2CREF="${a:0:23}"".cref.hdf"
 OutL2RHOM="${a:0:23}"".rhom.hdf"
 OutL2CHLA="${a:0:23}"".chla.hdf"
 echo $OutL2chl
 echo $OutL2ncref
 echo $OutL2ncrad
 echo "Determining ancillary data for Level-2 processing.."
 echo "ms_met.csh $a"
 ms_met.csh   $a
 echo "ms_zone.csh $a"
 ms_ozone.csh $a
 ms_oisst.csh $a
 parfile=MODISL1BtoL2seadas.par
 echo "Generating L2 $a"
 l2gen par=$parfile par=$a.met_list par=$a.ozone_list
par=$a.sst_list ifile=${datDir}/$a geofile=${geoDir}/$b
ofile1=${outDir}/${OutL2NWLR} ofile2=${outDir}/${OutL2CREF}
```

```
ofile3=${outDir}/${OutL2RHOM} ofile4=${outDir}/${OutL2CHLA}
resolution=250 aer_opt=-10 maskcloud=1 mumm_alpha=1.94500
mumm_gamma=1.00000 mumm_epsilon=1.00000
done|tee VMODISL1BtoL2seadas.MUMM.cldmsk.log
rm *_list
echo "Done"
########################################

########################################
###MODIS.par file for l2gen specs for converting MODIS L1B to L2
########################################

l2prod1=Rrs_469,Rrs_555,Rrs_645,l2_flags
spixl= 1
epixl= -1
dpixl= 1
sline= 1
eline= -1
dline= 1
ctl_pt_incr= 1
proc_ocean= 2
atmocor= 1
proc_land= 0
proc_sst= 1
resolution= 250
gas_opt= 11
pol_opt= 3
aer_opt= -3
aermodmin= 0
aermodmax= 0
aermodrat= 0.00000
mumm_alpha= 1.94500
mumm_gamma= 1.00000
mumm_epsilon= 1.00000
aer_rrs_short= -1.00000
aer_rrs_long= -1.00000
aer_swir_short= 1240
aer_swir_long= 2130
aer_wave_short= 748
aer_wave_long= 896
aer_iter_max= 10
brdf_opt= 7
iop_opt= 0
qaa_opt= 1
glint_opt= 1
outband_opt= 2
filter_opt= 1
filter_file=$OCDATAROOT/modisa/msl12_filter.dat
no2file=$OCDATAROOT/common/no2_climatology.hdf
land=$OCDATAROOT/common/landmask.dat
water=$OCDATAROOT/common/watermask.dat
icefile=$OCDATAROOT/common/ice_mask.hdf
```

```
gain=[0.9710,0.9848,1.0020,0.9795,0.9870,0.9850,0.9842,1.0049,0.9
797,0.9776,0.9855,1.0304,1.000,1.055,1.000,1.115]
offset=[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
.0,0.0]
albedo=    0.0180000
rhoamin=  0.000100000
qaa_adg_s=    0.0150000
chloc2_wave=[469,555]
chloc2_coef=[0.1543,-1.9764,1.0704,-0.2327,-1.1404]
chloc3_wave=[443,489,550]
chloc3_coef=[0.283,-2.753, 1.457, 0.659,-1.403]
chloc4_wave=[]
chloc4_coef=[]
chlclark_wave=[443,488,551]
chlclark_coef=[0.789273,-3.925523,11.637764,-
27.157997,27.936958,-10.398587]
nlwmin= 0.150000
wsmax= 8.00000
tauamax= 0.300000
epsmin= 0.850000
epsmax=      1.35000
glint=   0.00500000
windspeed= -1000
windangle= -1000
pressure= -1000
ozone=   -1000
watervapor= -1000
relhumid= -1000
sunzen= 70.0000
satzen= 60.0000
maskland= 0
maskcloud= 0
maskglint= 0
maskbath= 0
masksunzen= 0
masksatzen=0
maskhilt=0
maskstlight=0
```

Appendix 3: MODIS Regression Results

Regression Analysis: logCHL versus log(555/469), 555/469sq, High_wind

Method

Categorical predictor coding (1, 0)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	3.09738	1.03246	11.59	0.000
log(555/469)	1	0.28151	0.28151	3.16	0.081
555/469sq	1	0.01282	0.01282	0.14	0.706
High_wind	1	0.00415	0.00415	0.05	0.830
Error	53	4.72300	0.08911		
Total	56	7.82039			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.298518	39.61%	36.19%	31.99%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.722	0.131	5.50	0.000	
log(555/469)	3.01	1.70	1.78	0.081	17.15
555/469sq	-1.78	4.70	-0.38	0.706	13.69
High_wind					
1	0.047	0.218	0.22	0.830	2.44

Regression Equation

High_wind

0 logCHL = 0.722 + 3.01 log(555/469) - 1.78 555/469sq

1 logCHL = 0.769 + 3.01 log(555/469) - 1.78 555/469sq

Fits and Diagnostics for Unusual Observations

Obs	logCHL	Fit	Resid	Std Resid	
2	0.622	0.744	-0.122	-0.46	X
3	0.529	0.762	-0.234	-0.89	X
14	1.688	1.579	0.109	0.44	X
21	0.765	0.643	0.122	0.46	X
42	1.335	1.567	-0.231	-0.91	X

X Unusual X