Using multispectral sensor WASP-Lite to analyze harmful algal blooms

Shari J. McNamara
Using Multispectral Sensor WASP-Lite to Analyze Harmful Algal Blooms

Shari J. McNamara

B.S Rochester Institute of Technology (2005)

A thesis submitted to the
Chester F. Carlson
Center for Imaging Science
in partial fulfillment of the requirements
for the Master of Science Degree
at the Rochester Institute of Technology
Rochester, New York

August 21, 2007
Chester F. Carlson
Center for Imaging Science

USING MULTISPECTRAL SENSOR WASP-LITE
TO ANALYZE HARMFUL ALGAL BLOOMS

By
Shari J. McNamara
B.S Rochester Institute of Technology
(2005)

A thesis submitted in partial fulfillment
of the requirements for the degree of Master of Science
in the Chester F. Carlson Center for Imaging Science
of the College of Science
Rochester Institute of Technology

August 2007

Signature of the Author __________________________________________________________

Accepted by _________________________________________________________________
Coordinator, M.S. Degree Program Date
The M.S. Degree Thesis of Shari J. McNamara has been examined and approved by the thesis committee as satisfactory for the thesis requirement for the Master of Science degree.

Dr. Anthony Vodacek, Thesis Advisor

Dr. Robert Kremens

Dr. Harvey Rhody

Date
Title of Thesis:
USING MULTISPECTRAL SENSOR WASP-LITE TO ANALYZE HARMFUL ALGAL BLOOMS

I, Shari McNamara, hereby grant permission to the Wallace Memorial Library of RIT to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature: _____________________________

Date: _________________________________
ACKNOWLEDGEMENTS

I would like to thank Don McKeown and Dr. Harvey Rhody from the Laboratory for Imaging Algorithms and Systems (LIAS) within the Center for Imaging Science, for hiring me as an undergraduate and exposing me to the great research being performed within LIAS, and continuing to support me through my graduate research by deciding to fund my work through ISSI.

My Advisor Dr. Antony Vodacek has offered great advice and support, his knowledge of water research has been invaluable. He has offered many suggestions and ideas for getting around any roadblocks I had encountered. I especially appreciate that he took me on as an unplanned advisee, and his assistance in developing the application for the basis of my Thesis.

Dr. Robert Kremens was also instrumental in brainstorming for a Thesis research topic. What a great idea to use our in house WASP-Lite system to detect something! (Dr. Vodacek supplied the “something” as harmful algae.) Chats with Dr. Kremens about my research “problem of the day”, along with whatever else happened to come up, including softball of course, helped relieved some stress and frustration and encouraged me to push forward with my work.

Dr. Harvey Rhody was constantly supportive and encouraging. I especially appreciate the image processing assistance that was provided along many steps of the imaging chain. He would always offer advice and observations about the problem at hand and allude to the fact that these are life and career lessons that will never be forgotten. As I start my new job I am sure I will realize more and more of these lessons and think back to my time at LIAS.

I also would like to thank DIRS (Digital Imaging and Remote Sensing Laboratory) staff assistant Cindy Shultz for extending endless help to me as a LIAS student. She always helped me figure out course information, paperwork and especially helped in figuring out the steps needed to defend and submit my thesis. She always seemed to know the answer when no one else did!

I would also like to acknowledge Dr. Gregory Boyer and Mike Satchwell from ESF (SUNY College of Environmental Science and Forestry) for collaborating with RIT, and allowing Yushan Zhu and myself to accompany Mike on the ground collect to perform our measurements. I would also like to thank Jason Faulring for his assistance with using WASP-Lite and for flying the collect needed for this research. Yushan Zhu was also instrumental in assisting me at Lake Champlain, collecting water samples and performing
spectral measurements. And I would like to thank Matt Montanaro and Aaron Gerace for assisting in performing the radiometric calibration necessary for my research.

Last but not least, I would like to thank my family. I could not have done this without the support and encouragement from my parents and my boyfriend, Jeremy. I especially appreciate the phone calls with my dad, discussing all aspects of my research, possible solutions and relating his previous experiences to the issues at hand. My mom was amazing at lifting my spirits and inspired me to become the person I am today. I am grateful for Jeremy’s continuous patience and understanding when it came to late night thesis work and softball games. He was always interested in my progress and kept me focused on my goals.

I would like to acknowledge the funding that has allowed me perform this research, the Integrated Sensing Systems Initiative (ISSI) (NASA contract number NNS06AA84G).
ABSTRACT

Developing methods to monitor harmful algae is a current research “hot-topic.” One type of algae, the blue-green algae or Cyanobacteria, cause blooms that can lead to a health threat to humans and animals. This research will test the use of a cost effective and temporally efficient method using multispectral remote sensing system, WASPLITE, as a monitor of algal blooms. This airborne system will be optimized to the specific application of detecting Cyanobacteria on optically complex waters. Attempts have been made in the past using existing instruments, e.g., SeaWiFS and Landsat, to provide these data, but our solution can provide more information by using optimally selected bands with very high spatial resolution. To analyze these algal blooms, standard multispectral techniques (such as band ratio, spectral curvature and principal component analysis) were used on the airborne data. These results were compared with ground truth collected concurrently with the airborne over flight.

Because of the very high spatial resolution of the system, (0.7 m), compared to many commonly used satellite systems (~30m to 1km), it could be seen that the patchiness of the algae was very high. Difficulties in applying the ground truth were both technical shortcomings and were due to the nature of the algal blooms. Technical issues include the time lag between the ground sample collect and the airborne collect (the water and algae move with time), the drift of the boat during ground sampling (there was no anchor), and the error in the GPS units in both the boat and the plane. The issues due to the nature of water and algae include, sun glint in the imagery, white foam lines created by waves and wind, and most importantly, the patchiness of the algae in the water. Because the ground truth of one sample point per location was not adequate, we could not correlate the ground truth to the imagery. Qualitatively, the images did show a large variation of algae concentration in the water through the principal component analysis. Further, flow-through data from another vessel taken from the same week this research was performed, suggests that the variation that is seen in the imagery is real. Overall, this research shows the difficulties in effectively and accurately performing ground truth measurements to be used to test algorithms and methods that are applied to detecting harmful algae using remotely sensed data. The traditional ground sampling methods failed to capture the spatial variation observed in the image data. With improved techniques we are confident these methods can be used to effectively monitor algal blooms using the high spatial and temporal resolution.
# Table of Contents

ABSTRACT ................................................................................................................................. VII

TABLE OF CONTENTS ............................................................................................................. VIII

LIST OF FIGURES ..................................................................................................................... XI

LIST OF TABLES ......................................................................................................................... XIII

BACKGROUND ........................................................................................................................... 2

The Problem: Cyanobacteria ...................................................................................................... 2

Light and Water ......................................................................................................................... 4

Previous Methods of Identifying Algae .................................................................................... 12

APPROACH ............................................................................................................................... 23

Objectives and Criteria ............................................................................................................ 23

Methods .................................................................................................................................... 25

Experimental Sampling Location and Planning ........................................................................ 26

Remote Sensing and Ground Sampling Equipment Used .......................................................... 31
  Multispectral Instrument WASPLITE .................................................................................... 31
  Ground Sampling Equipment ............................................................................................... 34

Description of Ground Sampled Data ..................................................................................... 34

Optimization of Remote Sensing Equipment ........................................................................... 36

Image Processing Procedures .................................................................................................. 40

Radiometric Calibration: Flat Fielding ....................................................................................... 42

Radiometric Calibration: Performing the Correction ............................................................... 49

Geometric Processing ............................................................................................................... 52
  Lens Distortion Correction ....................................................................................................... 52
Signal to Noise Investigation ..........................................................117
Foam and In-scene Noise Removal Test .................................................122
The Phenomenon of Wave Focusing.......................................................127

CONCLUSIONS & RECOMMENDATIONS ........................................ 128

REFERENCES ................................................................................. 133

APPENDIX A ................................................................................... 139

APPENDIX B ................................................................................... 140
List of Figures

Figure 1: Possible paths of light to the sensor, including: a) reflect of water, scatter in atmosphere, travel to sensor; b) scatter in atmosphere, travel to sensor; c) light reflects off water, travels to sensor; d) light scatters multiple times in atmosphere, reflects of water, travels to sensor; e) light penetrates water, scatters in water, exits water and travels to sensor; and f) light scatters in atmosphere, scatters in water, exits water and travels to sensor. ................................................................. 5
Figure 2: Paths of light due to scattering include: a) light scattering in water and interacting with inorganic suspended material; b) light scattering in the water and travels to sensor; c) light entering water and absorbed by dissolved organic matter; d) light scattering in water off of the ground or bottom of the lake or water source and traveling back to the sensor; and e) light scatters due to phytoplankton in the water and then travels to the sensor. ................................................................. 6
Figure 3: WASP-Lite image of water showing bottom effect .............................................. 11
Figure 11: Reflectance spectra showing chlorophyll and phycocyanin features (Green 2006) .............................................................................................................................. 37
Figure 12: Particle absorption of samples from Missisquoi Bay with the band-centers of WASP-Lite cameras shown as vertical lines ........................................................................ 40
Figure 14: Lens falloff diagram showing the difference in angles from imaging different points on a plane .......................................................................................................................... 44
Figure 33: Spectral response curves for the five WASP-Lite cameras with filters attached ................................................................................................. 50
Figure 34: Calibration curves for the five WASP-Lite cameras ........................................... 51
Figure 42: Three band false color rendition where red is 550nm, green is 630nm and blue is 650nm ......................................................................................................................... 68
Figure 43: Missisquoi Bay, Landsat image with ground sample orthorectified images... 72
Figure 44: Brochet's ground sample start and end locations, and orthorectified image closest to these points................................................................. 73
Figure 45: Brochet's false color three band image, where red is 550nm, green is 630nm and blue is 650nm ................................................................. 74
Figure 46: Center ground sample point locations and orthorectified image .................... 75
Figure 47: Center's false color three band rendition, where red is 550nm, green is 630nm and blue is 650nm ................................................................. 76
Figure 48: Goose Bay's ground sample locations and orthorectified image .................... 77
Figure 49: Goose Bay's false color three band rendition, where red is 550nm, green is 630nm and blue is 650nm. The sampling boat is circled ............... 78
Figure 50: Two orthorectified images that showing the misalignment of the boat ......... 79
Figure 51: Site 50's orthorectified image and ground sample location ......................... 80
Figure 52: Manually rotated color rendition image, showing approximate location of
ground sample point........................................................................................................ 81
Figure 53: Site 50's false color three band rendition, showing radius of error of 30 and 60
meters................................................................................................................................ 82
Figure 54: St Albans Bay, showing inner and outer sampling points and images........... 83
Figure 55: St Albans Inner sampling point location and orthorectified image.......... 84
Figure 56: St Albans inner false color three band rendition................................. 85
Figure 57: St Albans Outer ground sampling point location and orthorectified image... 86
Figure 58: St Albans outer false color three band rendition................................. 87
Figure 59: Cole Bay's ground sample point location and orthorectified image........ 87
Figure 60: Cole Bay false color three band rendition............................................ 88
Figure 61: Missisquoi Bay- Landsat image is the base image, the orthorectified images
for each flight line, the ground sample locations and the flow though fluorometry path are
shown................................................................................................................................. 90
Figure 62: Area near Site 50, showing the orthorectified images and the flow through
data path.......................................................................................................................... 91
Figure 63: Flow through fluorometry data for Site 50 area....................................... 92
Figure 64: Images correspond to above figure, left to left image, right to right image.
These are the three color renditions, showing the variation of content.................... 93
Figure 65: Graph of chlorophyll concentration vs. phycocyanin concentrations collected
while ground sampling..................................................................................................... 97
Figure 66: The 10 bit floating point imperfectly registered image........................... 101
Figure 67: The 8 bit floating point perfectly registered image.................................. 102
Figure 68: The ROIs on the 10 bit floating point imperfectly registered image, with foam
minimized.......................................................................................................................... 103
Figure 69: The PCA image using the 10 bit imperfectly registered data................... 104
Figure 70: The PCA image using the 8 bit perfectly registered data.......................... 104
Figure 71: PCA image of Brochets.............................................................................. 106
Figure 72: ROIs chosen for Brochets, based on PCA............................................... 106
Figure 73: Color rendition image of Brochets, showing patchiness.......................... 107
Figure 74: Spectral curvature results vs. phycocyanin of 10 bit data, foam-minimized.
The lines shows the expected slope of the relationship between the phycocyanin
concentration and the sampling locations................................................................. 109
Figure 75: Spectral curvature results vs. chlorophyll of 10 bit data, foam-minimized. The
line shows the expected slope of the relationship between the chlorophyll concentration
and the sampling locations......................................................................................... 112
Figure 76: Band Ratio (550-630nm) results vs. phycocyanin of 10 bit data, foam-
minimized. The line shows the expected slope of the relationship between the
phycocyanin concentration and the sampling locations............................................. 114
Figure 77: Band Ratio (550-630) results vs. chlorophyll of 10 bit data, foam-minimized. The line shows the expected slope of the relationship between the chlorophyll concentration and the sampling locations.

Figure 78: Band Ratio (650-630) results vs. phycocyanin of 10 bit data, foam-minimized. The line shows the expected slope of the relationship between the phycocyanin concentration and the sampling locations.

Figure 79: Site 50, band 2 (550nm)
Figure 80: Horizontal profile of Site 50 (band 2)
Figure 81: Site 50 (band 2) with 21x21 low pass smoothing kernel applied
Figure 82: Horizontal profile of Site 50 image with 21x21 smoothing
Figure 83: Site 50 (band 2) with 41x41 low pass smoothing kernel applied
Figure 84: Horizontal profile of Site 50 with 41x41 smoothing
Figure 85: Site 50 color image, showing foam minimized with ROIs used
Figure 86: SAM image from Site 50
Figure 87: Site 50 Foam minimized using SAM, and showing ROIs used
Figure 88: Site 50 original image showing ROIs location
Figure 89: Zoomed window of Site 50 image, showing the white foam, dark pixels and colored pixels of apparent “blue” water
Figure 90: Zoomed window of Site 50, showing the white foam, dark pixels and colored pixels of apparent “red” water

List of Tables

Table 1: List of data gathered from RIT and ESF ground sampling, and RIT airborne flight
Table 2: Ground sample distances for each camera type at nominal flying altitude of 3000ft.
Table 3: Flying parameters used
Table 4: Bandcenters of interest to water remote sensing and cyanobacteria
Table 5: Arrival times of ground sampling vessel and aircraft to sample locations, and the time difference between them
Table 6: Ground sample locations used
Table 7: Scenarios for image analysis
Table 8: Spectral curvature with corresponding ROI color for Brochets image
Table 9: Spectral curvature results of foam minimized original method vs. method using SAM
Background

The Problem: Cyanobacteria

Cyanobacteria, the blue-green algae, are a problem in our natural water systems. Many species of cyanobacteria produce harmful toxins. Research programs such as MERHAB (Monitoring and Event Response for Harmful Algal Blooms) have been monitoring harmful algae in the Great Lakes and in Lake Champlain for years. In one collaborative document released in 2006, this topic is discussed in depth (Boyer 2006). In this article, Boyer explains that cyanobacteria are ubiquitous in nature and are found in nearly all environments. Cyanobacterial blooms can develop and lead to taste and odor problems in drinking water. The toxins that can be produced are extremely harmful to animals and humans. In Lake Champlain alone, during the summer of 2000 three dogs died from the algal toxin from two different bloom events (Rosen et al. 2001).

Reports have shown that an invasion of zebra mussels into a lake can shift the phytoplankton community, such as increasing the density of Microcystis, a type of cyanobacteria. This is because zebra muscles can ingest all types of cyanobacteria except for Microcystis (Mihuc et al. 2006). Previous research has shown that for the past 80+ years of Lake Champlain’s history, Anabaena usually accounted for 75%+ of the blue-green algae in most of the sites (Mihuc et al. 2006 and within). Current research and
sampling performed by Mihuc, has determined that populations of Microcystis are now the dominant taxon lake wide.

Methods for reducing the harmful algae have been identified. However, direct control is usually only applicable in smaller ponds and embayments. This treatment includes adding algaecides, reducing nutrient inputs, mixing to vertically destratify the water column, reducing retention time by increasing flow rate or flushing, and biological manipulation (Boyer and Dyble circa 2006). The only viable long term solution that has been identified is reduction of nutrient inputs. Overall, the occurrence of the toxic cyanobacteria will remain persistent because these methods will be inadequate and inefficient for large lakes such as Lake Champlain and the Great Lakes. For the treatment of drinking water the toxins can be removed with methods such as chlorination. However, if the pH levels are to high (>8.0) then the effectiveness of the chlorine greatly decreases.

Monitoring harmful algae is a step commonly taken to manage the problem of the cyanobacteria infecting natural water systems. Through monitoring, people can be advised whether water is safe for recreational use, or safe enough for pet access. Ground sampling by boat is currently the most commonly used method. Ground sampling is time consuming and provides limited spatial resolution. When sampling, two algal pigments are usually measured- chlorophyll and phycocyanin. Ground sampling may be
temporally and spatially inefficient, but the pigment measurements are usually very accurate. Airborne imaging has been attempted to accurately map the harmful algal blooms as well. These methods have not been perfected and research is still being done in this field. Algae can currently be mapped with airborne and satellite systems using chlorophyll detection, but the harmful algae detection is more difficult to map because pigment phycocyanin also needs to be detected. If this measurement problem can be solved, the task of monitoring lakes will be much simpler than conventional ground sampling, and will result in higher spatial and temporal resolution, as well.

**Light and Water**

It is important to take a look at how light interacts with water. Figure 1 shows the many paths that light can take before reaching the sensor. These paths include sunlight scattering in the atmosphere and reflecting off the water to the sensor (a), sunlight reflecting directly off the water to the sensor (b), or light emitted from the water going directly to the sensor (c) (IOCCG 2000).
The color of the water is determined by scattering and absorption of visible light by pure water with inorganic particulate, organic particulate and dissolved material present in the water (IOCCG 2000). Remote sensing is useful because it can be used to analyze the variation in magnitude and spectral quality of water leaving radiation, which in turn can
be used to measure the type of substances and concentrations that are present in the water.

Figure 2 shows how the light scatters and reflects off of particles in the water. The three main scattering components consist of phytoplankton, suspended inorganic material and water itself (IOCCG 2000).

Figure 2: Paths of light due to scattering include: a) light scattering in water and interacting with inorganic suspended material; b) light scattering in the water and travels to sensor; c) light entering water and absorbed by dissolved organic matter; d) light scattering in water off of the ground or
bottom of the lake or water source and traveling back to the sensor; and e) light scatters due to phytoplankton in the water and then travels to the sensor.

Phytoplankton are organisms that are found in illuminated surface layers of the water. They are single-celled plants which are at the base of the food web (IOCCG 2000). Many algal species can coexist in the same water volume, and the dominant genera in algal groupings can change spatially, seasonally, and with physical, chemical and biological changes in the water (Wetzel 1983).

The pigments of phytoplankton consist of chlorophylls, carotenoids and biliproteins (Wetzel 1983). Chlorophyll a is the primary photosynthetic pigment of all oxygen-evolving photosynthetic organisms. It is present in all algae and photosynthetic organisms other than photosynthetic bacteria. Variation in phytoplankton densities are responsible for the variation in optical properties of waters (IOCCG 2000). Phytoplankton reflectance is dominated by the two vitro absorptions bands, the red light regions between 660-665nm and near 430nm (Wetzel 1983). Chlorophyll b is a light gathering pigment that transfers absorbed light energy to chlorophyll a for primary photochemistry. The maximum absorption bands for Chlorophyll b peak at 645nm and 435nm respectively. Other chlorophyll pigments include chlorophyll c, d, and e which are all less common. The carotenoids are linear unsaturated hydrocarbons. Like chlorophyll b, light energy absorbed by carotenoids and biliproteins is transferred to
chlorophyll a, leading to fluorescence and excitation of chlorophyll a molecules (Wetzel 1983).

One type of phytoplankton are the cyanobacteria, or blue-green algae. Like bacteria, cyanobacteria have murein in the cell wall, they reproduce by binary fission, and do not divide by mitosis as other algae and higher organisms do (Wetzel 1983). However, unlike bacteria, cyanobacteria have chlorophyll a, which is common to eucaryotic algae and higher plants. Cyanobacteria are also able use water as an electron donor in photosynthesis, which is more advanced than bacterial photosynthesis. Overall, cyanobacteria is structurally and physiologically like bacteria but functionally is like plants in aquatic systems.

Harmful algae, like cyanobacteria, are detrimental not only because they produce toxins, but also because they shade light from other aquatic life. In addition, when a bloom collapses, microbial respiration on the dead and decaying cells can lead to very low oxygen concentrations that can kill fish and other aquatic organisms due to lack of oxygen (Liew et al. 1999).

The pigment that is produced by cyanobacteria, phycocyanin, along with the pigment chlorophyll a, have unique spectral features that are important to the application of using remote sensing for cyanobacterial detection.
Other types of phytoplankton include green algae, yellow-green algae, golden-brown algae, diatoms, cryptomonads, dinoflagellates, euglenoids, brown and red algae (Wetzel 1983).

Suspended material consists of all inorganic and organic particulate that is not phytoplankton (IOCCG 2000). This include bottom sediment that may be in suspension, which alters the color of the water. Sediments strongly influences coastal and inland bodies of water. Many studies have been done using remote sensing for mapping suspended materials. Satellite or airborne remote sensing can be used with water samples to obtain calibration data to create a relationship, which is influenced by the particle size of the sediment and the characteristics of the remote sensor (Knaeps, et al. circa 2005). We recognize that remote sensing can be used for suspended material, but we will not be considering this issue for this paper.

Yellow substances are colored dissolved organic mater (or CDOM), which consist of humic and fulvic acids (IOCCG 2000) which are comprised of hundreds of compounds (Chen et al. 2003). These substances may have a local origin, such as from the degradation of organic particles, or they may have a distant source such as from rivers that flow through heavily wooded regions (IOCCG 2000). CDOM provides increased
absorption of light with decreasing wavelength in the 350-700nm range (Chen et al. 2003, & sources within).

Remote sensing has been used to detect CDOM, as done by Chen et al. (2003). They found a band ratio combination of 670nm to 412nm to be optimal. This is because the reflectance value at 412nm decreases with an increase of DOC, while at 670nm there will be an increase with land originated matter or with phytoplankton. This algorithm was performed on SeaWiFS data and it was concluded that it could be used for water assessment.

Bottom effects are also very important confounding factors to consider when using remote sensing. Bottom effects result when light is reflected off the bottom of a water body. If the water is shallow and clear enough, it can greatly influence the apparent water color (IOCCG 2000).

Figure 3 shows an image taken with WASP-Lite over St. Albans Bay, where the upper part of the image shows the bottom of the lake. Notice that the color is very different from the deeper water on the bottom of the image compared to the shallow ground on the top of the image.
There are two different types of water, Case 1 and Case 2 waters. Case 1 waters are waters in which phytoplankton are the principal agents responsible for the variation in optical properties of water, such as open oceans (IOCCG 2000). Case 2 waters are those that are influenced by phytoplankton and related particles along with other substances such as inorganic particles in suspension and yellow substances, such as coastal regions and lakes.

There are many satellite remote sensing algorithms already in place for analyzing Case 1 waters. Because the smaller scale and complexity of Case 2 waters, these same algorithms are not effective. Jupp et al. (1994) had shown that satellites have been used
to detect algal blooms, but aerial remote sensing provide a spatial scale more suitable for inland waters.

Tassan (1994) also tested and developed algorithms for Case 2 waters, for future use on SeaWiFS data. Previous to his research, SeaWiFS had only been used for Case 1 water algorithms. Tassan showed that by using in situ measurements, the algorithms they used appeared to yield sufficiently accurate results when trying to retrieve phytoplankton, pigments, suspended sediment, and yellow substances in coastal (Case 2) waters.

**Previous Methods of Identifying Algae**

A number of researchers have investigated Chlorophyll a and phycocyanin content in Case 2 waters, with the similar goal of identifying and mapping these pigments using remote sensing. Multispectral and hyperspectral, satellite and airborne systems have all been used to perform this task in the past. A number of researchers have developed concluding results, where the community has learned what algorithms may or may not work, along with specifics of the task, such as ground truth collecting to data acquisition. This section will review a number of these attempts in chronological order.

Dekker et al. (1992) attempted to monitor cyanobacteria using an airborne system called CASI. CASI is a pushbroom system with the capability of being used in spatial mode or a spectral mode. The spatial mode can collect up to 15 spectral bands with 512 pixels
across track. The spectral mode can collect 288 bands at 1.8nm intervals. Dekker made use of the spatial mode and selected band ranges that included 624.4-640.9, and 644.0-651.6nm. These two band ranges were used in a band ratio of 624/648 nm. It was noted that optimally, the bandpasses would be smaller (about 10nm) than what was used on CASI. From the ground truth that was collected, the phycocyanin concentration was calculated using the 624/648nm band ratio. An approximately linear relationship was found between the band ratio used on the ground truth and the band ratio used from the CASI imagery. Therefore a linear interpolation was used. The linear correlation coefficient that was found between the CASI reflectance ratios and the modeled reflectance ratios was 0.70. This preliminary study showed that pigments such as phycocyanin can be detected with remote sensing. Dekker did suggest that the results could have been improved if the ground truth data was collected on the same day as the airborne data instead of two days apart.

CASI was used again by Jupp et al. (1994) to detect, identify and map cyanobacteria. The spatial mode was again used for this study, this time using 12 bands. The bands selected for phycocyanin detection were centers of 618.55, 623.93 and 649.07nm. Chlorophyll \(a\) band centers used were 663.45 and 683.26. Ground truth sampling was performed to be used with the airborne data. This sampling showed that the spatial variation of chlorophyll concentration in the water was significant from sample to sample. This variation makes it is difficult to correlate airborne data with water samples.
taken at the time of the flight. First, a signature referencing method was used to
distinguish characteristics in the water. This is done by using a signature from an area
that appears to look like it has low algal content, and is divided into signatures from other
areas of water. This method did show clear differences between water with and without
algae and cyanobacteria. The other method that was used was a band ratio technique,
with band centers of 710/680nm. Jupp compared the results using this method with a
technique that uses the same ratio, but then normalizes the results. They found that
normalizing did improve results. Jupp used an atmospheric model to aid in estimating
the optical water quality. Then an “inversion” was performed on the data. This inversion
of the spatial data did not seem to distinguish cyanobacteria from green algae when the
chlorophyll concentration was low. Also, when comparing spectral data, it seemed that
the phycocyanin absorption was not detectable when the concentrations were too low.
However, when the concentrations were high, the signatures did show the phycocyanin.
Jupp recommend that in their future work they will try methods using the derivation of
parameters of optical water quality, including looking at the ratio of chlorophyll
concentration to phycocyanin, from airborne imagery. It was concluded that it seems
feasible to measure the amount of phycocyanin relative to the chlorophyll concentration
if atmospheric correction, absorption coefficients and spectral modeling can be improved.

Optical properties of dense algal cultures were analyzed by Gitelson et al. (1995) to
assess the feasibility of making estimates of cyanobacteria concentrations. Reflectance
and vertical attenuation coefficient spectra were collected using a high spectral resolution radiometer for a number of ground samples. Using this data, algorithms were developed for detecting chlorophyll $a$ and phycocyanin. Regression analysis was used on bands 438nm and 676nm to detect chlorophyll $a$. The determination coefficients ($R^2$) were 0.97 and 0.98 respectively for the two bands. Wavelength 624nm was used for phycocyanin detection, where 0.95 determination coefficient was found. This study proved the concept of using remotely sensed data utilizing these specific bands to detect chlorophyll $a$ and phycocyanin. These models were used specifically for detecting a species of cyanobacteria, *Spirulina*. It was suggested that the parameters of the model could be adjusted for other species. These algorithms depend on the diffuse attenuation coefficient measurements, and the optical properties (specific absorption and scattering coefficients) are crucial to these models. Further work by Gitelson et al. (1995) includes testing the potential use and limitations of these models.

Similarly to the previous researcher, Xiaozhou et al. (1998), developed an algorithm to estimate chlorophyll $a$ from the spectral reflectance of inland water using ground collected reflectance spectra. A regression model using the band ratio of 700/675 nm to detect chlorophyll, in combination with either of the two following ratios of 560/624, or 600/624 was used to detect phycocyanin. This study found that using the chlorophyll ratio in combination with either of the phycocyanin ratios, resulted in good estimation of chlorophyll concentration (of 0.98 and 0.96 correlation coefficients for 600/624 and
560/624nm respectively). When Xiaozhou et al. (1998) used the chlorophyll model developed by Gitelson (1993), the correlation coefficient was found to be 0.88. The Xiaozhou et al. (1998) new model showed a measurably good improvement over previous methods. The algorithms discussed had not been tested on any airborne imagery or data, but their preliminary results were encouraging for using remotely sensed data to detect chlorophyll.

In one attempt to assess the spectral bands of remote sensing satellite instruments for detecting cyanobacteria, Roelfsema et al. (2001) used the set bands of Landsat TM, SeaWiFS, and MODIS to resample data from an ASD Spectrometer that was used to collect spectra of water in a bay in Australia. The specific type of algae in interest had the pigments phycocyanin (with 620nm absorption feature) and phcoerythrin (with a 565nm absorption feature). This study showed “promising” results in separating the cyanobacteria, *Lyngbya majuscula*, from other types of spectra by comparing algae spectra to other spectra, such as clean water, rock, dirt, other organisms and other constituents. It was noted that more endmembers (or spectra) along with further analysis of the influence of the optical water column properties on the spectra was needed to make the study more complete.

Hyperspectral airborne data was used by Galvao et al. (2003), who performed spectral reflectance characterization of shallow lakes. The airborne AVIRIS system was used,
which has 224 bands and 20m spatial resolution. Ground truth was collected, consisting of BRF (Bidirectional Reflectance Factor) measurements, chlorophyll concentrations, along with many other measurements. Principle component analysis (PCA) was used to analyze the ground spectra collected to identify a homogeneous set of lakes. The continuum removal method was also used to normalize the data to isolate the features. The AVIRIS data was converted to reflectance using ATREM (Atmosphere Removal technique) and EFFORT (Empirical Flat Field Optimal Reflectance Transformation). PCA and scatter plots of the AVIRIS data were used to select five classes of endmembers. Linear spectral unmixing and the continuum removal method were both applied to the AVIRIS data. When comparing the spectra of the five endmembers, the features at 630nm (absorption due to phycocyanin) and 667nm (absorption due to chlorophyll) were both present. In conclusion, spectral features (such as 630nm phycocyanin) were observed with ground truth spectra and with the AVIRIS data, so they believe AVIRIS can be used to identify broad algal groups.

In effort to design a potential hyperspectral remote sensing imager for water quality measurements, a research study was performed to determine specifications for this type of satellite (Zur et al. 2003). A spectral resolution of 5 to 10 nm bandwidths would be optimal, along with a minimal spatial resolution of 10m to include use for lake and case 2 waters. The revisit time required would be 3-4 days. Specific band specifications include 624 and 648nm band centers for detection of phycocyanin, 700nm, >700nm, and
670nm band centers were required for chlorophyll a detection, along with other specific bands for use of water quality measuring.

Satellite remote sensing was used by Kutser (2004), to quantitatively detect chlorophyll in cyanobacterial blooms. The first civilian hyperspectral satellite, Hyperion, with its 400-2500nm spectral range, 10nm bandwidths and 30m spatial resolution was used. This data was converted to reflectance using FLAASH (Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes). Another satellite sensor, ALI (Advanced Land Imager), with its 10 bands and 30 meter spatial resolution was also used. As with many of the other studies, the wavelengths 630 and 650nm were of interest when trying to detect phycocyanin. Bio-optical modeling and SAM (Spectral Angle Mapper) were used to produce chlorophyll concentration maps from the images. It was noted that cyanobacterial blooms are extremely patchy in form. The Hyperion images show how patchy the cyanobacterial blooms are, which explains why mapping these blooms have been so difficult in the past using satellite systems with insufficient spatial resolution. Considering that Hyperion has 30 meter spatial resolution, this essay will show that the patchiness is very apparent at even a much higher spatial resolution than 30 meters. The 630nm phycocyanin feature was detected in their data. It was also found that the 650nm reflectance peak was not detected in cyanobacteria spectra when the chlorophyll concentration was less than 10 mg m$^{-3}$. It was found that the peak could be seen when the chlorophyll concentrations were between 30-50 mg m$^{-3}$. 
When considering the ground truth collection, since the concentration of phytoplankton varies with depth, flow-thorough systems used on ships of opportunity cannot provide a reliable estimator of phytoplankton in the water during blooms (Kutser 2004). This problem also occurs in research vessels collecting water samples. The vessel disturbs the water surface and displaces the surface aggregations away from the ship. This suggests that it is practically impossible to collect water samples that would be representative of the natural conditions while surface aggregations of cyanobacteria exist. Thus the amount of chlorophyll seen by the remote sensor will not be the same as measured in the water, even if the spatial resolution of the sensor is equal to the sampling size of the water sample. Thus, as will be shown later in this essay, a single point in situ measurement is inadequate for validation of satellite chlorophyll estimates during cyanobacterial blooms.

LANDSAT TM was used to detect phycocyanin and map cyanobacterial blooms by Vincent et al. (2004). This study used LANDSAT TM data along with in situ water samples collected to develop and algorithm to estimate phycocyanin concentration. The algorithm was developed with one data set, tested and verified with another LANDSAT and ground truth data set, all from 2000. In 2002, three data sets were used to test the algorithm, with only one having any ground truth for comparison. Their results for 2002 were inconclusive, and their results were unexplainable. They tested their algorithm on July 16 2002, Aug 1 2002, and Aug 8 2002. The algorithm used produced a
measurement of a moderate level of phycocyanin on July 16, a decreased level on Aug 1, and high levels on Aug 8. It was also reported in the news on Sept 17 2002 that a large bloom was occurring. It is possible that the Aug 8 data showed the start of this September bloom. It was thought that an analysis error could have occurred on Aug 1 to account for the low concentrations, but multiple testing was performed and that is unlikely. It was also suggested that other type of organism is responsible for the phycocyanin increase in the July 16th data. There were two main algorithms that Vincent compared, both were regressions, one using single bands, and the other using ratios of bands. All of the data was also dark subtracted to reduce the effects of atmospheric haze. From their comparisons, it was found that the regression using the spectral ratio of bands was more robust and reliable than using single band inputs. Even though the unexpected results from 2002 were not validated, it was concluded that LANDSAT TM can be used to evaluate water, including phycocyanin, by using the phycocyanin regression algorithm that uses spectral ratios, and that a spectral resolution of 28.5 m and the LANDSAT bands are adequate to resolve locations for measurement of phycocyanin (Vincent et al. 2004).

In another example of the use of the hyperspectral satellite system, Candiani et al. (2005), used Hyperion for water quality assessment. The Case 2 waters were evaluated using a procedure that maps chlorophyll and tripton concentrations using a direct inversion of a bio-optical model by means of a linear matrix inversion method, as published by Brando
and Dekker (2003). In their processing, to increase signal to noise, a 5x5 low pass filter kernel was used. Their data was also converted to reflectance using a Modtran based algorithm. The bio-optical model computes the spectral subsurface irradiance reflectance using water quality parameters along with parameters output from HYDROLIGHT. The ground truth was collected using a flow through system, the Fluorescence and Turbidity Analyzer, which was mounted to the boat. This data was then used in comparison with the results found from Hyperion. Candiani et al. (2005) showed a good agreement of an $R^2$ of 0.84 for chlorophyll and $R^2$ of 0.78 for tripton estimations.

The multispectral satellite MERIS (Medium Resolution Imaging Spectrometer) was used by Simis et al. (2005) to detect phycocyanin in turbid inland water. MERIS has 300m spatial resolution. Simis et al. (2005) uses the bands 620nm, 665nm, 709nm, and 778nm, absorption and backscattering coefficients, and optical correction factors in a model that retrieved phycocyanin concentration from turbid water reflectance. The phycocyanin concentrations predicted by the algorithm were then compared to the measured phycocyanin concentrations (sample by sample values at 620nm were used for calculating the specific absorption coefficients). The regression analysis showed an excellent agreement with an $R^2$ value of 0.94. When performing this same comparison but using one fixed average specific absorption coefficient for every sample (calculated at 620nm), the linear least-squares fit showed a $R^2$ value of 0.77. This was done to show that the specific absorption coefficient needs to be calculated and used for each sample.
Simis et al. (2005) concluded that this method could be tailored to any sensor that records the reflectance to include the bands used: phycocyanin absorption around 615nm, chlorophyll absorption around 675nm, a far red wavelength greater than 705nm and a near-IR wavelength between 760-800nm. Overall, Simis et al. (2005) believe this algorithm could aid in the monitoring of cyanobacterial populations in turbid, eutrophic lakes and reservoirs.

A study performed by Shuchman et al. (2006), examined the chlorophyll concentration in water using SeaWiFS over a 7 year period to test a bio-optical algorithm. This was specifically a study to test how SeaWiFS (optimally used for Case 1 waters) works with Case 2 waters. The method used here were a fast-operating algorithm that was based on a great lakes hydro-optical model, and a combination of the Levenberg-Marquardt (L-M) multivariate optimization approach and neural network (NN) emulation technique. Overall, the L-M technique provided more accurate results and is more robust for noise contaminated data, but is slower than the other technique.

Most recently, Wheeler (2006), performed an analysis (in a Masters Thesis) of three remote sensing satellite systems, monitoring cyanobacteria in Lake Champlain. Specifically, St Albans Bay and Missisquoi Bay were used as test sites for the experiments. The three systems she used were QuickBird (2.4 m resolution), SPOT (20m resolution), and MERIS (300m resolution). Ground truth consisted of
measurements of concentrations of chlorophyll $a$ and phycocyanin determined from water samples. To detect chlorophyll $a$ or phycocyanin, a number of algorithms were used. The methods used to analyze the SPOT and QuickBird data include comparing ground collected pigment data (Chlorophyll $a$ and phycocyanin) to single bands, band ratios and principal component analysis, using empirically based linear regression analysis. The MERIS data was analyzed with semi-empirical optical models developed by Gons et al. (2005) and Simis (2005). A third model was also used, called the Water Processor, which is an automated chlorophyll model which predicts chlorophyll $a$ for Case II waters. Wheeler obtained this model from the European Space Agency (ESA). Overall, all three systems seemed to have value for detecting and mapping algal blooms in Lake Champlain at various scales, but Wheeler found that the MERIS analysis was most valuable because of the instrument’s spectral resolution, despite having poor spatial resolution.

**Approach**

**Objectives and Criteria**

The goal of this experiment was to determine if we could detect cyanobacteria with the multispectral system WASP-Lite. Other researchers have devised methods to detect cyanobacteria by developing algorithms to use on existing sensors (hyperspectral airborne, or hyperspectral and multispectral satellite), as explained in the previous
section. At times, very intricate algorithms have been developed to use on sensors that do not have optimal bands for cyanobacteria detection. This experiment customizes this sensor to the application, so relatively simple algorithms can be used for analyzing. WASP-Lite and its five bands makes it an ideal platform to test our hypothesis. Other researchers also have not been able to accurately or consistently detect or map cyanobacteria. Using our customized sensor and the high spatial resolution of this airborne system will allow us to investigate why this is the case, and allow us to find out the benefits to having high spatial resolution.

While previous remote sensing methods used multi- and hyperspectral satellite and airborne systems, WASP-Lite can not only offer better spatial resolution, unique spectral band selection, but also good temporal resolution. This system is also inexpensive compared to the previous systems because it is a compact, and can be flown on a single engine aircraft (vs. twin engine aircraft or satellite.)

Tasks to prepare the imagery for exploitation, included radiometric calibration, applying flat field and lens distortion corrections, band to band registration, orthorectification, sun glint minimization, and white foam minimization.

To exploit this imagery, standard multispectral techniques, such as band ratio techniques or spectral ratio methods, were used to identify algal blooms. These results were
compared to ground truth, performed by teams from Rochester Institute of Technology (RIT) and College of Environmental Science and Forestry (ESF) obtained concurrently with the airborne flight.

**Methods**

The following is a brief outline of the Approach section, describing how the goal of detecting harmful algae will be accomplished.

- Experimental Sampling Location
- Remote Sensing and Ground Sampling Equipment Used
- Description of Ground Sampled Data
- Optimization of Remote Sensing Equipment
- Radiometric Calibration including Flat Fielding
- Geometric Processing
- Artifact Removal
- Spectral Methods
  - Band Ratio Technique
  - Spectral Curvature
**Experimental Sampling Location and Planning**

Table 1: List of data gathered from RIT and ESF ground sampling, and RIT airborne flight

<table>
<thead>
<tr>
<th>RIT</th>
<th>ESF</th>
<th>Flight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>Date</td>
<td>Date</td>
</tr>
<tr>
<td>Volume of water used for TSS (l)</td>
<td>Phycocyanin concentration (ug/L)</td>
<td>Time over ground sample point</td>
</tr>
<tr>
<td>TSS concentration (g/m³)</td>
<td>Chlorophyll concentration (ug/L)</td>
<td>Flight Altitude</td>
</tr>
<tr>
<td>Volume of water used for PA (l)</td>
<td>Temp (C)</td>
<td>Ground Sample Distance (m)</td>
</tr>
<tr>
<td>GPS N</td>
<td>pH</td>
<td>Flight GPS N</td>
</tr>
<tr>
<td>GPS W</td>
<td>Secchi Depth (m)</td>
<td>Flight GPS W</td>
</tr>
<tr>
<td>GPS N (drift)</td>
<td>Dissolved Oxygen (mg/L)</td>
<td></td>
</tr>
<tr>
<td>GPS W (drift)</td>
<td>Estimated time of arrival to each site</td>
<td></td>
</tr>
<tr>
<td>Particle Absorption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle – Pigment Absorption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pigment Absorption</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 shows the types of information or data that was collected by each means during the collect. This includes the measurements taken by RIT, ESF and the flight parameters of the RIT system WASP-Lite. The data that was collected by RIT includes the date, the volume of water used for TSS (sediment), the TSS (sediment) concentration, volume of water used for PA (particle absorption), GPS N and W (latitude and longitude), the drift of the GPS (latitude and longitude), the particle absorption (the absorption spectrum due to particles in the water), the particle minus pigment absorption (the absorption spectrum of the particles but with the pigments in the water removed), and the pigment absorption (the absorption spectrum of only the pigment in the water). The measurements recorded by ESF includes the date, the phycocyanin concentration (concentration of the toxic pigment in cyanobacteria), the chlorophyll concentration, the temperature, the pH of the water, the secchi depth (the clarity of the water is when looking straight down through it), dissolved oxygen content, and the estimated time of arrival to each location or site. The data collected by the airborne flight include the date, the time over the ground sample location, the flight altitude, the ground sample distance (the pixel size in meters on the ground), and the GPS latitude and longitude information.

Lake Champlain was chosen as the target site because of the collaborative effort with ESF and their scheduled water sampling and experiments. Three areas were chosen based on the history of algal content as reported by ESF personnel. Missisquoi Bay was chosen as an area that in the past has been dominated by Microcystis and historically, is
the most toxic area of the lake. There also may be some species of algae as well. St. Albans Bay was chosen as an area with mixed Anabeana and Microcystis concentrations, both of which are toxic. Areas around Cole Bay were chosen as areas with low algal population and no toxicity.

Lake Champlain is 120 miles long, and the widest point is 12 miles wide. The greatest depth is 400 ft, while the average depth is about 64 feet (Lake Champlain Basin Program, 2007).

Figure 4 shows all of Lake Champlain, with the three general areas that were used for the ground sample collection; Missisquoi Bay, St Albans Bay, and Northwest/Cole Bays. Figure 5 shows the sampling stations for Missisquoi Bay which include: Brochets River, Goose Bay, Center of Missisquoi Bay, Site 50, Venise, and added point A. The sampling stations for St. Albans Bay (shown in Figure 6) includes St Albans Inner, St Albans Outer, Lapans Bay, and added point St Albans. The southern bay area (shown in Figure 7) includes Northwest Bay, Cole Bay and Site 7.
Figure 4: Lake Champlain with the sample locations.

Figure 5: Missisquoi Bay with all possible sample locations
Six locations in Missisquoi Bay were planned as sampling locations. In practice, only five locations were both ground sampled and over flown. On the morning of the flight we heard that one particular location had high amounts of algae (from boaters on the lake), and we tried to change one sampling point to this new location. Because the flight navigation software already had the original locations programmed, the flight missed the new point. Because the ground crew was unaware of this, only five out of the six locations have corroborating data. This also occurred in St Albans Bay, where 2 of 3
ground points have collaborating data. All three of the points around Cole Bay are collaborating. Thus, 10 points in all were over flown and ground sampled.

Because of time restraints, the ground crew collected the area around Cole Bay the previous day, while Missisquoi Bay and St Albans ground samples were collected on the same day as the flight.

**Remote Sensing and Ground Sampling Equipment Used**

**Multispectral Instrument WASPLITE**

The multispectral imaging system, WASP-Lite, was designed and built in the Laboratory for Imaging Algorithms and Systems, (LIAS) at Rochester Institute of Technology’s (RIT) Center for Imaging Science. The objective of this system was to offer a relatively cheap and compact system, built with off the shelf parts, for the initial use of detecting and mapping wildfires and other environmental phenomenon. This system can be flown aboard a single engine aircraft (Cessna 172), at a nominal speed of 90 knots and a nominal operational altitude of 3000 feet.

The system includes the computer acquisition software and components, the sensor head and a monitor for controlling the system in the aircraft. There are a total of seven cameras in the sensor head. An oblique view of the sensor head is shown in Figure 8, a face-on view of the cameras is shown in Figure 9, and a labeled schematic is shown in
Figure 10  One camera is a longwave microbolometer for infrared applications. A second camera is a high resolution panchromatic camera, to be used for sharpening. The last five cameras are identical panchromatic cameras which use optical filters for band selection. The ground sample distance at the nominal speed and flying altitude are shown in Table 2, and the flying parameters that were used for this experiment are shown in Table 3.
Table 2: Ground sample distances for each camera type at nominal flying altitude of 3000ft.

<table>
<thead>
<tr>
<th>Camera</th>
<th>Ground Sample Distance (GSD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panchromatic High Resolution</td>
<td>0.34 m</td>
</tr>
<tr>
<td>The five panchromatic cameras</td>
<td>0.84 m</td>
</tr>
<tr>
<td>Infrared longwave</td>
<td>2.1 m</td>
</tr>
</tbody>
</table>

Table 3: Flying parameters used

<table>
<thead>
<tr>
<th>Parameters Used</th>
<th>Meters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flying altitude</td>
<td>~650 m (~2100 ft)</td>
</tr>
<tr>
<td>Ground sample distance (GSD)</td>
<td>~0.6 m</td>
</tr>
<tr>
<td>Footprint on the ground</td>
<td>~388 m x 296 m</td>
</tr>
</tbody>
</table>

The cameras of interest for this application are the five panchromatic cameras. These cameras are Sony XCL-V500 cameras. The sensor is a progressive scan, interline CCD. The pixel size is 7.4 x 7.4 um. The pixel array is 648 x 494. The maximum frame rate is 60 Hz, and the shutter speed can be ¼ to 1/100,000 sec. The dynamic range of these cameras are 10 bits. The nominal lens focal length is 8 mm. The spectral pass band of these cameras are 0.4 to 1.0 um. Using the nominal settings, the image size would be 548 x 418 meters on the ground (McKeown 2006).
Ground Sampling Equipment

RIT’s ground sample equipment included plastic water bottles, water filtration pumps, bottles and filters, and liquid nitrogen cooler to hold samples until return to lab. A dual beam spectrophotometer, (Shimadzu UV2100U) was used to record the absorption measurements in the laboratory.

ESF’s ground sample equipment included the research vessel, and the various equipment used on the vessel, including a water pump filtration system, a YSI 6600 Sonde, Secchi Disk, filter paper, and dry ice. In the laboratory, the equipment included a Turner Designs 700 Fluorometer, a Turner Designs 10AU Fluorometer, a centrifuge, acetone, ice and a freezer.

Description of Ground Sampled Data

Ground sample data was collected by RIT graduate students. From one liter water sample bottles, the water was filtered using a pump and filtration system with glass fiber filters in the evening of each collect. The samples were then placed into a liquid nitrogen tank for preservation until returning to the lab. Once returning to the lab, spectrophotometer measurements were taken. The measurements results included CDOM absorption on water samples that were filtered through 0.2 um pore size nylon syringe filters. The CDOM absorption (a_y) was calculated by the equation

\[ a_y = \frac{2.303 A_y}{r} \]

where \( A_y \) is the CDOM absorbance, and \( r \) is the optical path length.
(DeGrandpre et al. 1996). The particle absorption of the sample collected on a glass fiber filter was measured with the spectrophotometer. The sample was then rinsed with methanol leaving the particle minus pigment. The absorption of the particle minus pigment was then measured. The absorption for both were measured over the spectral range of about 400 to 700 nm. The pigment absorption was then calculated by subtracting the particle minus pigment absorption from the particle absorption. All of this particle data was corrected for scattering and converted to absorption coefficient using a technique from Cleveland & Weidermann (1993).

The ground sampled data that was analyzed by ESF was collected by ESF and RIT personnel. The final data set from ESF, included collection date, temperature of water, pH, secchi depth (m), dissolved oxygen (D.O.) (mg/L), extracted phycocyanin (ug/L), and extracted chlorophyll a (ug/L).

The secchi depth was collected using a secchi disk using standard procedures. The extracted phycocyanin was analyzed by collecting 300 ml of water through a pump filtration system, through 47mm polycarbonate filters, with 1 um pore size. The samples were stored in cryogen tubes in a dry ice cooler. The samples used to measure extracted chlorophyll samples required 1 liter of water to be pumped through the filtration system, through 47mm glass fiber filters. These samples were also stored in cryogen tubes in the dry ice cooler. In the laboratory, the phycocyanin extraction protocol, (a detailed
description shown in Appendix A requires freezing and thawing the sample a number of times, centrifuging the samples, and using the 10-AU fluorometer to record fluorescence. The following equation was used to determine the phycocyanin concentration of each sample:

\[ PC(\text{mg/dL}) = \frac{10 \times \text{AUPC} \times \text{VolumeExtracted} \times \text{DilutionFactor}}{\text{VolumeFiltered}} \]

where 10AUPC is the fluorescent concentration obtained from the 10AU Flurometer. was used to determine the phycocyanin concentrations. The method for determining the extracted chlorophyll is also shown in detail in Appendix B, requires the sample to be sonicated, froze, and measured with the TD 700 fluorometer. The concentrations are calculated by

\[ \frac{\text{CorrectedChl(\text{mg/L})}}{\text{VolumeExtracted} \times \text{DilutionFactor}} = \frac{\text{TD700Reading} \times \text{VolumeFiltered}}{\text{VolumeExtracted} \times \text{DilutionFactor}} \]

The dissolved oxygen was read in the field using a YSI 6600 sonde.

**Optimization of Remote Sensing Equipment**

Past research has shown many that bands are of interest to water remote sensing. Some of these bands are shown in Table 4. Band of interest, centered at 405nm and 865nm were chosen because they could be used, if necessary for atmospheric correction. Band 480 was chosen because it shows a carotnoid absorption feature. Bandcenter 520nm was interesting because the slope of the reflectance at this point is different when comparing Chlorophyll to Chlorophyll with Phycocyanin. This can be seen in Figure 11. The bandcenter at 550nm was important because it is a common feature for plant material that
contains chlorophyll a, as a reference point. The features at 630nm and 650nm are due to the absorption and reflectance (respectively) of the toxin Phycocyanin. The feature at 670nm is also an absorption feature for Chlorophyll a. The two peaks at 710 and 750nm were wavelengths that were suggested to be of interest to water research (Vodacek 2006)

![Reflectance spectra showing chlorophyll and phycocyanin features](https://example.com/figure11.png)

Figure 11: Reflectance spectra showing chlorophyll and phycocyanin features (Green 2006)
### Table 4: Bandcenters of interest to water remote sensing and cyanobacteria

<table>
<thead>
<tr>
<th>BandCenter (nm)</th>
<th>Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>405</td>
<td>For Atmosphere Correction</td>
</tr>
<tr>
<td>480</td>
<td>Carotnoid Absorption Feature</td>
</tr>
<tr>
<td>520</td>
<td>Shows Slope of Reflectance Feature</td>
</tr>
<tr>
<td>550</td>
<td>Reflectance due to Chlorophyll a</td>
</tr>
<tr>
<td>630</td>
<td>Phycocyanin Absorption</td>
</tr>
<tr>
<td>650</td>
<td>Reflection due to Phycocyanin</td>
</tr>
<tr>
<td>670</td>
<td>Chlorophyll a Absorption</td>
</tr>
<tr>
<td>710</td>
<td>Interesting Water Band</td>
</tr>
<tr>
<td>750</td>
<td>Interesting Water Band</td>
</tr>
<tr>
<td>865</td>
<td>For Atmosphere Correction</td>
</tr>
</tbody>
</table>

As the summer ended, and fall began, which is the end of the algal bloom season, it appeared that due to time constraints only one flight and sampling event was going to take place. This meant that the “best” bandpasses had to be chosen for the five available cameras. Considering that band ratio and spectral curvature techniques were going to be used, bandpasses 550nm, 630nm and 650nm were chosen. The features at 630nm and 650nm are due to the pigment, phycocyanin, and the feature at 550nm was due to chlorophyll. Bandpasses 405 and 865 were chosen for the possible use in atmosphere correction. If methods were used that require absolute pixel values, then the atmospheric
correction would need to be performed. Even though these methods were not expected to be used, it was important to make sure we had as many options available before we collected the data. The methods used did not depend on absolute pixel values but did depend on band ratios, which are relative comparisons, so that these atmospheric bands were not used for atmospheric correction.

Figure 12 shows the particle absorption spectra of algae samples from the Lake Champlain ground sampling collect. The vertical lines show where the band centers are located on the spectral curve. Note that this absorption spectra was obtained during the data collection and was not used in the band selection process. This graph illustrates that the bands chosen were indeed a good set to use for the algae that was present in the water at the time.
Figure 12: Particle absorption of samples from Missisquoi Bay with the band-centers of WASP-Lite cameras shown as vertical lines

**Image Processing Procedures**

The image analysis process consists of two steps. The first step is orthorectification of one camera (band 1) that has been boresighted. Boresighting allows the coordinate system of the camera focal plane to be transformed into the coordinate system of the inertial management unit. By knowing the location of the aircraft (using standard navigation methods and the data form the IMU) and the pointing of the camera (also from the IMU), we can orthorectify the image, that is, we can transform the focal plane image
into a geographically correct image. For my experiment, only camera one was boresighted, so a method had to be devised to register the other bands to this geographically correct image. This turned out to be a difficult problem that was not soluble in the time allowed, so instead, we registered the five bands to each other in an arbitrary coordinate space, and then used the result of this transformation to visually locate the new, arbitrary coordinates space with respect to the geographically correct image from camera one. To obtain the best image quality, several different analysis processes were used, which are shown graphically in Figure 13. We used two different band-to-band registration methods and several methods for artifact removal, which will be explained in detail later. Only after this processing is done, can spectral methods be used to analyze the images.
Flat Fielding

Radiometric Calibration

Lens Distortion Correction

Orthorectification

Map Ground Sampled Locations

Band to Band Registration

Artifact Removal

Spectral Methods

GDBICP

IDL Program Based on GDBICP

Glint Minimization

White Foam Minimization

Figure 13: Flow diagram of image processing

**Radiometric Calibration: Flat Fielding**

Flat fielding is a process that corrects an image for lens falloff. Ideally, the camera sensitivity should be uniform over the whole image plane. In practice, this is not the
Lens falloff, or natural vignetting, refers to the radial falloff due to geometric optics (van Walree 2007). The illumination varies from the center of the image plane to the corner. Figure 14 shows the difference in angles from imaging a point in the center of the plane to imaging a point in the corner of the plane, shown as angle b. There are three cosine effects that result in the $\cos^4(b)$ illumination falloff factor. The first is the $\cos^2(b)$ factor due to the inverse square law. The light has a longer path to travel to the image corner. Second, the pupil seen by the off-axis point is elliptical, not round as seen by the on-axis point, and has a smaller capture area than the round pupil. This yields another $\cos(b)$ factor. Third, while the light hits the center of the image at normal incidence, it hits the corner of the image at angle b. This results in another $\cos(b)$ factor. The combined effects result in the $\cos^4(b)$ illumination factor. This factor does vary with focus distance, lens configuration and design, but is a good approximate for many lenses.
First, the $\cos^4$ term was calculated for the WASPLITE system. This factor was applied to each image by an IDL program which divided this radially dependent term into each pixel of each image. The original image from Camera 1 is shown in Figure 16. The image shown in Figure 15 is Figure 16 after the $\cos^4$ correction has been applied.
Because lens falloff was not completely corrected for, an additional technique was used. Using the images that have had the $\cos^4$ term removed, the next step is to separate the images by flight line (which is inherently done by the data acquisition software.) The flight lines over the same locations can be used in one processing step. This means that if there are two flight lines over one area, the images in both flight lines will be used together. Next, one representative image is picked out of this set of images. Then all of the images in this set are scaled to this one representative image, so the data values are all on the same scale. An average is then taken of this set of images. This average image is shown in Figure 17. This process averages out the details of the scene and results in one image that shows only the lens falloff due to the camera system.

Figure 17: Average of all images in a set

Figure 18: Final flat fielded image
Using IDL, a surface fit is calculated from this average image. The surface fit shows the variance in the image. Using IDL, the surface fit is then removed from each individual image in the set of images being used (the flight lines that were selected over one common location.) This step was performed using logarithms, so when we “subtract” the variance from the images, it is a multiplicative process, to ensure image integrity. The final flat fielded image is shown in Figure 18.

Taking a look at one single image and comparing the data values before and after the flat fielding, it can be seen that the average values were adjusted upwards to the values near the center of the original image. This was done because the radiometric calibration (which will be discussed later) was done using data values from the center of the images. Figure 19 and Figure 21 show the images before and after for a Camera 1 image, while Figure 20 and Figure 22 show the horizontal profile of each image, with “value” being pixel value, and “sample” being pixel location.

Figure 19: Raw image from camera 1
Figure 20: Horizontal profile of raw image from camera 1
The next 10 images show the before and after results of the flat fielding process for one image. It can be seen that there is still a small peak or bright circle in some of the images, but this effect is much smaller than before flat fielding.
Figure 25: Camera 2 raw image  
Figure 26: Camera 2 flat fielded image

Figure 27: Camera 3 raw image  
Figure 28: Camera 3 flat fielded image
Radiometric Calibration: Performing the Correction

The radiometric calibration converts the pixel values from raw digital counts (10bit) to radiance units. The calibration is carried out by imaging an integrating sphere where the output spectral radiance of the sphere is known absolutely. The output of the sphere is multiplied by the system response curve of the camera. This is then numerically
integrated over the entire composite curve to produce an integrated radiance in the pass band of the camera. To back out the radiance incident on the front of the camera, the integrated radiance from above is divided by the integrated normalized response curve of the camera. In other words, the integrated normalized camera response is divided into the integrated radiance to recover the “per nm” unit in spectral radiance. The equation

\[ L_\lambda = \frac{\int L(\lambda)R'(\lambda)d\lambda}{\int R'(\lambda)d\lambda} \]

describes the above, where \( R'(\lambda) \) is the normalized system response, \( L_\lambda \) is the radiance incident at the front of the camera, and \( L(\lambda) \) is the output radiance of the integrating sphere.

![Response Curves for the Five Cameras](image)

**Figure 33:** Spectral response curves for the five WASP-Lite cameras with filters attached

The normalized systems response \( R'(\lambda) \) is obtained by first measuring the spectral response. The spectral response of the instrument is obtained by measuring the camera
output from each camera over a range of discrete wavelengths. A monochromator was used to produce a monochromatic beam incident on the camera. The digital count from the region of the image where the beam is incident was measured. The monochromator wavelength is changed in discrete steps over a wavelength range large enough to accurately map each filter/camera combination. For example, for a 550nm filter (10nm bandwidth), images should be obtained for wavelengths of about 540nm to 560nm. A dark frame should also be obtained to subtract out the dark current noise. The noise-subtracted digital counts from each wavelength are then plotted to give the shape of the response curve of the sensor. The spectral responses for the cameras are shown in Figure 33. Finally, the curve is normalized to a maximum value of unity to produce the normalized response curve for the system.

Figure 34: Calibration curves for the five WASP-Lite cameras
The radiance, $L_\lambda$, is then associated with the mean digital count value of the image obtained with the camera. The mean digital count in a dark frame captured with the sphere’s shutter closed will provide another data point on the radiance versus digital count calibration curve. Assuming the response is linear, the calibration curve will be a straight line connecting the two data points from above. The curves for each camera are shown in Figure 34. The “dark” digital counts are the same for each camera, because they are identical cameras. The radiance “light” value is different for each camera because each camera has a different interference filter, changing the integrated spectral response of each camera. These filters will be discussed at a later time.

**Geometric Processing**

**Lens Distortion Correction**

Lens distortion was corrected for the lenses by other RIT graduate students and personnel. This is done by imaging a fixed target consisting of a set of point targets and using a program to determine the distortion. The correction is applied by using an IDL program.

**Band to Band Registration**

Because this system did not yet have a band to band registration technique developed, one needed to be developed. A technique called Generalized Dual Bootstrap- Iterative
Closest Point (GDBICP) was used (Yang et al, 2007). It is a fully automated 2D image registration algorithm designed to register two images taken of the same scene. “Fully automated” means that it includes an initialization technique, and estimation algorithm and a decision step. It can also handle substantial illumination and spatial differences in the scene as well. The program can use a variety of transforms, but here the homography transform is used. A homography transforms simply means that one point on one plane is mapped onto another plane. However, there are a few notes to be made about this technique. This program runs in a windows DOS environment, and cannot be run from IDL. The program can register images for two cases- single band to single band images, or three band to three band images (such as RGB.) This is issue for our case, considering that WASPLITE has five bands. Additionally, the program only works with byte scaled images (8 bit), and the images used here are in floating point format (because of the radiometric calibration.) Because this method uses a homography transform, the images are transformed to a new space different from either of the original input spaces (it does not anchor one image and register image two to image one.)

A process was developed to use this program, and “trick” it into working for the five bands of WASPLITE. Figure 35 shows a flow diagram of the processing used to register one five band set of images. Camera 1 image and Camera 2 images are first registered together. Then using IDL, these two images were stacked along with a “blank” band (zeros), creating a three band image. This was also done for the Camera 3 and Camera 4
images. The fifth band, was stacked into a three band image, where, this time, bands 2 and 3 were blank. This Camera 5 image was then registered to the Camera 4 & 5 registered image. IDL was used to re-stack the images, remove the blank bands, and stack them in the correct order (3,4 and 5). This three band image was then registered with the Camera 1 and 2 image. IDL was used again, to remove the blank band and to restack the images in the correct band order. This is a long and tedious process. Considering there are hundreds of images, this process is at the limit of practicality.

Because this program requires each image band to start out as byte format, a precaution needs to be taken to ensure the spectral integrity of the image. When the 10 bit floating point, individual bands are byte scaled to 8 bit, they are not scaled with respect to the other four bands they correspond to because they are not yet registered. This means that all of the radiometric calibration is lost and the spectral content is not accurate. To deal with this issue, a program was written to manually byte-scale the individual bands to be registered. When doing this, the scaling factors are stored. After the five bands have been registered with the GDBICP program, one can then manually scale each band back into floating point, and maintain the spectral integrity. It is also important to note that once the image is converted to 8 bit, even when we convert the image back to floating point, the image remains 8 bit. Two bits of data have been lost in this process.
The results of this process were amazing. Considering the spatial complexity of the images (lake water), algae and waves line up perfectly in almost all of the images. Figure 36 shows a registered image, displaying bands 2, 3 and 4 as red, green and blue. The red box outlined in this image is shown in Figure 37 as a zoomed in window of the area.

Figure 35: Flow diagram for image registration process
Notice that the white foam in the image is a pure white, and does not show misregistered pixels.

Figure 36: Perfectly registered image using GDBICP. Red box represents the "zoom" window in the figure below

Figure 37: Perfectly registered image, zoomed in to see how the waves and white foam are aligned
Even though this registration process worked so well, the fact remains, it is not efficient to run. It can be run only on byte scaled images. To preserve the 10 bit radiometric calibration by keeping the pixel values in floating point, and to streamline the process, IDL was used. The registration program does not only produce the two registered images, but it also produces the Homography transform matrix that was used on the two images. A program was developed in IDL to use this transform matrix and apply it to the two images (Rhody 2007). This program was developed further to automatically register and re-stack all five bands in one image, using the same basic data flow shown in Figure 35.

The results of this process were very good, but not as perfect as the original GDPCIP registration. Figure 38 shows the resulting registered image where bands 2, 3 and 4 are displayed as red, green and blue. The red box in Figure 38 is a zoomed in window of the image, shown in Figure 39. In this zoomed in image, it can be seen that the white foam does not have clean white edges, but instead has a red and blue shift of pixels. The pixel shift is about 1 to 4 pixels, depending on the location in the image, and the bands observed.
Figure 38: Imperfectly registered image using program based on GBICP

Figure 39: Imperfectly registered image, zoomed in to show the misregistration in the waves and white foam
Both methods were used - the IDL program that results in imperfectly registered data, and using the tedious DOS GDBICP program for select images that provides perfectly registered data in 8 bit format. From this point, the data resulting from these two methods will be referred to as the imperfect registration or the perfect registration.

**Orthorectification**

The orthorectification process maps the images to the earth’s surface. For our application, this orthorectification is required for both the mosaicing of images and to allow correct spatial comparison with ground truth.

The first step of this process is to calculate boresight angles. The boresight angles are the calculated angles between the focal plane of the camera and the internal navigation system of WASPLITE by RIT personnel. This has been performed only for Camera 1. Camera’s 2-5 boresight angles do not need to be calculated because they will be registered to Camera 1’s image space. Camera 1’s images were orthoreciftied using Lieca Inc. Photogrammetry suite in Leica’s Imagine software (Leica Geosystems 2005).

**Artifact Removal**

**Glint & White Foam**

Figure 40 and Figure 41 represent images from the collect that show either sun glint or white foam (respectively). Sun glint and white foam are optically different; where an
image with sun glint may still have usable spectral information, white foam is optically opaque, and no information about the underlying water can be obtained from these pixels. Because glint and foam are optically different, different techniques were attempted to reduce or minimize the effects of each of the issues.

To minimize the sun glint, a de-glinting algorithm from Hedley et al. (2005) was attempted. This de-glinting program was used first on the imperfectly registered data, which did not provide a satisfactory result. It is believed the spatial registration was not good enough for the program to work adequately. The algorithm was then attempted on the perfectly registered data, and the resulting images were better. The slight registration improvement resulted in improved de-glinted data. It is important to note that this program not only removes sun glint, but because the white foam pixels are so bright, and spatially uniform, it reduces some of these artifacts as well.

The algorithm that Hedley et al. (2005) uses is shown in the equation below. This method works by establishing a linear relationship between the NIR and visible bands through a linear regression based on the sample pixels. For this application, the NIR band had a 865nm band center, and the visible bands were centered at 405nm, 550nm, 630nm, and 650nm. Also, instead of using one region from band 865nm as the “sample” the whole image was used (as long as nothing especially unique was visible, like a boat, land or the bottom of the lake). Each visible band is included in a linear regression of
NIR brightness against the visible band brightness. The slope for band \( i \) is \( b_i \), and the pixels in the image that can be de-glinted follow the equation,

\[
R'_{i} = R_{i} - b_i (R_{\text{NIR}} - Min_{\text{NIR}})
\]

where the sun-glint removed pixel brightness in band \( i \) is the pixel value in band \( i \) minus the regression slope times the difference between the pixel NIR value and the ambient NIR level.

White foam was actually a much bigger problem than the sun glint, as seen in Figure 41. It was necessary to develop a method to correct for the foam, especially when the images were processed using the program that does not allow for perfectly registered images. The best method that was developed involved masking out the white foam by setting a pixel value threshold and setting those pixels to “not a number.” This worked quite well. Other methods that could be performed are a nearest neighbor technique, or an averaging kernel to remove the white foam pixels.
Figure 40: Image showing sunglint in the upper lefthand part of the image

Figure 41: Image showing white foam lines throughout the image
Spectral Methods

Band Ratio Technique

The first identification method is a normalized band ratio technique. Two different band ratio combinations were used, based on the bands chosen for the WASPLITE system. The first method uses the Phycocyanin absorption peak at 630nm in a normalized ratio with the Chlorophyll a reflectance peak at 550nm. The second uses the Phycocyanin reflectance peak at 650nm in a normalized ratio with the Phycocyanin absorption peak at 630nm.

\[
\text{Ratio}_{630,550} = \frac{\text{PixelValue}_{630} - \text{PixelValue}_{550}}{\text{PixelValue}_{630} + \text{PixelValue}_{550}}
\]

\[
\text{Ratio}_{650,630} = \frac{\text{PixelValue}_{650} - \text{PixelValue}_{630}}{\text{PixelValue}_{650} + \text{PixelValue}_{630}}
\]

This band ratio technique was used in two ways- one way was by choosing regions of interest (ROIs) for an image, taking the mean of each ROI, and performing the band ratio on each ROI mean. The band ratio values for the ROIs with in an image would then be compared, as well as comparing the band ratio values from image to image, to see if they were correlated. Another way this method was used, is by calculating the band ratio on the entire image using ENVI and visible looking at the differences in the image and then analyzing the values within that image.
Spectral Curvature Technique

Because the three bands of interest (550, 630 and 650nm) do make a spectral “curve” when the particle absorption is plotted out, spectral curvature technique can be used (Campbell and Esaias 1983). The spectra in Figure 12 show particle absorption measured from ground sampled water during the collect. All four of the samples here do show the Phycocyanin absorption feature around 630nm. If there was no Phycocyanin, and only algae with Chlorophyll a was shown as particle absorption, the “spectral curve” would be flatter than it is shown here, because the absorption feature at 630nm would not be present. This is why it is expected that this method can produce satisfactory results.

\[
G_{550,650}(630) = \frac{S(630)^2}{S(550)S(650)}
\]

As with the band ratio method, the spectral curvature method was also used in two ways. The ROIs were chosen, the means of the ROIs were calculated, then the spectral curvature was calculated on each ROI mean and compared within each image and from image to image. The other way spectral curvature was used is by calculating the spectral curvature on the whole image and comparing images in that manner.

Principal Component Analysis (PCA)

Principal component analysis was not initially a method that was going to be used because this algorithm needs perfectly registered images, as with the de-glinting algorithm. This analysis is an image transform that is designed to decorrelate the data.
and maximize the variability in a reduced number of features. Each feature value in the transformed data set is a linear combination of the features in the input data set. The equation below shows the output of one principle component of a set, where $x$ is the vector comprised of $l$ digital count values corresponding to the $l$ features. $PC_1$ is the brightness values of the first principal component feature, $e_1$ is the first principle component vector (eigenvector) composed of $l$ weights.

$$PC_1 = e_1^T x = e_{i1}DC_1 + e_{i2}DC_2 + \ldots + e_{il}DC_l$$

For the current analysis, the program ENVI was used to apply the principal component analysis.

This analysis was used to aid in discriminating what was really in the water. For example, when algae was present, the images showed outlines of the algal boundaries, and it reinforced any questions about the variability that was seen. PCA also aided in selecting the ROIs for the above band ratio and spectral curvature techniques.
Results

Overview

- Interpreting the False Color Display of Multispectral Algal Data
- Temporal Correlation of Ground Sampled Points and Airborne Data
- Spatial Correlation of Ground Sampled Points and Airborne Data
- Patchiness of the Algae and Flow Through Data Analysis
- Expected Results: Spectral Curvature and Band Ratio
- The Analysis Procedure
- Laboratory Analysis of the Ground Truth Samples
- Image Analysis Scenarios
- Discussion
- Results from Spectral Method Analysis
- Signal to Noise Investigation
- Foam and In-scene Noise Removal Test
- The Phenomenon of Wave Focusing

Interpreting the False Color Display of Multispectral Algal Data

It is important to consider how a five band image is being displayed for viewing. Because an image can only be displayed in three bands, red, green and blue, we have to carefully assign bands with colors. Considering that the 405nm and 865nm bands are
were not chosen for water analysis (they were chosen for atmospheric correction), we will only visually look at 550nm, 630nm, and 650nm. When assigning red, green, and blue to these three bands, respectively, we would expect that where algae exist, green would be reflected, and hence show more red in the image. Because the feature of phycocyanin at 630nm is an absorption feature, we would not expect to see much green where phycocyanin exists. Because of this, we would assume that where the image is red, there is algae- which may or may not contain phycocyanin.

Figure 42 shows an example of broad range of “red” to “blue” water, in which we expect that the red colored water would have higher concentrations of algae and the bluer water would have lower concentrations of algae.
Figure 42: Three band false color rendition where red is 550nm, green is 630nm and blue is 650nm.
Temporal Correlation of Ground Sampled Points and Airborne Data

Table 5: Arrival times of ground sampling vessel and aircraft to sample locations, and the time difference between them

<table>
<thead>
<tr>
<th>Location</th>
<th>Approximate Ground Sample Arrival Time (stayed for 20min)</th>
<th>Exact Fly-Over Time</th>
<th>Time Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 50</td>
<td>9:45 am</td>
<td>11:52 am</td>
<td>2 hr 7 min</td>
</tr>
<tr>
<td>Center</td>
<td>10:15 am</td>
<td>11:49 am</td>
<td>1 hr 34 min</td>
</tr>
<tr>
<td>Brochets</td>
<td>11:15 am</td>
<td>11:46 am</td>
<td>31 min</td>
</tr>
<tr>
<td>Goose Bay</td>
<td>11:55 am</td>
<td>11:55 am</td>
<td>0 min</td>
</tr>
<tr>
<td>St Albans Inner</td>
<td>2:30 pm</td>
<td>12:11 pm</td>
<td>2 hr 19 min</td>
</tr>
<tr>
<td>St Albans Outer</td>
<td>3:00 pm</td>
<td>12:15 pm</td>
<td>2 hr 45 min</td>
</tr>
<tr>
<td>Cole Bay</td>
<td>Previous day</td>
<td>12:45 pm</td>
<td>~24 hours</td>
</tr>
</tbody>
</table>

One of the issues with trying to correlate ground samples to airborne data, is the time difference between when the sample was collected and when the fly-over occurred. For this collect, the whole flight took one hour. The ground sampling took many hours, spanned over the time of two days. Missisquoi Bay’s ground sampling is the closest (in time) to when the plane flew overhead. The last sample point in that bay occurred exactly when the plane was flying overhead (Goose Bay). The plane took 15 minutes to
cover all of Missisquoi Bay’s sample points, the boat took about 2 hours. Referencing Table 5, it can be seen that the lag time between the first ground sampled point and the fly over was about 2 hours, and the shortest lag time was 0 minutes—when the fly over occurred just as the boat had stopped for sampling. The 2 hour time difference can allow for the algae to drift and move considerably, especially considering that the algae is extremely patchy and can vary within tens to hundreds of meters. The potential time lag between the over head flight of the St Albans sampled points and the ground sampling was about 2.5 to 3 hours.

Though this is not as critical as the time lag between ground and airborne sampling, the drift of the boat is also important. The start and end locations were collected with the GPS in the boat for each sample location. When these points were mapped to a geo-referenced image, there were locations where there was very little drift—less than 30 meters, and locations where there was more drift, about 30 – 60 meters. Even if this was the only source of error, it would still be influential, considering that the algae itself can vary with in tens to hundreds of meters.
Spatial Correlation of Ground Sampled Points and Airborne Data

Table 6: Ground sample locations used

<table>
<thead>
<tr>
<th>Location</th>
<th>Sub-location Used</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missisquoi Bay</td>
<td>Site 50</td>
<td>45.013333</td>
<td>-73.173833</td>
</tr>
<tr>
<td></td>
<td>Center</td>
<td>45.039167</td>
<td>-73.141683</td>
</tr>
<tr>
<td></td>
<td>Brochets Bay</td>
<td>45.064433</td>
<td>-73.104</td>
</tr>
<tr>
<td></td>
<td>Goose Bay</td>
<td>44.9864</td>
<td>-73.120083</td>
</tr>
<tr>
<td>St. Albans Bay</td>
<td>St. Albans Inner</td>
<td>44.785333</td>
<td>-73.162167</td>
</tr>
<tr>
<td></td>
<td>St. Albans Outer</td>
<td>44.76765</td>
<td>-73.186483</td>
</tr>
<tr>
<td>Southern Area</td>
<td>Cole Bay</td>
<td>44.138083</td>
<td>-73.42055</td>
</tr>
<tr>
<td></td>
<td>Northwest Bay</td>
<td>44.183783</td>
<td>-73.417283</td>
</tr>
</tbody>
</table>

As discussed earlier, there were a few locations that were planned as sample location but were not used. Table 6 shows breaks down the locations that were used, from the main location to the sub-locations. These sites will be discussed in further detail below. For each site, a base orthorectified image (like Landsat) will be used as a base map for the ground sample points and the orthorectified images that correspond to these points. A false three color rendition for each sample location is also shown.
Missisquoi Bay

The overview image in Figure 43 shows the four orthorectified images that corresponded to ground sample locations, along with the name of the area. Each of the areas will be examined.
Figure 44: Brochets ground sample start and end locations, and orthorectified image closest to these points

Figure 44 shows the closest image to the ground sampled points for this site. The drift of the boat was also significant here. This will be discussed in further details later. The drift is about 180 meters. This distance was estimated by using the pixel size of the background image, which is 30 meters (Landsat).

The image in Figure 45 shows the three band false color rendition of this image that is closest to the ground sample points. Notice that the image shows a great deal of variation in color, from blue-green, to blue, to red. This variation will also be discussed later in further detail.
Figure 45: Brochet's false color three band image, where red is 550nm, green is 630nm and blue is 650nm.
Center

Figure 46: Center ground sample point locations and orthorectified image

This location, in Figure 46, shows the ground sample points within the image, but there was some drift of the boat. It seems the drift is just over 30 meters. Figure 47 shows the false color rendition. It appears there is a bit of sun glint in the upper portion of the image.
Figure 47: Center's false color three band rendition, where red is 550nm, green is 630nm and blue is 650nm.
Goose Bay

This location is the only one where the plane flew over the boat as it was taking ground samples. From the image in Figure 48, you can see that there is about 60 meter drift in the boat, according to the GPS unit on the boat, and that it does not line up perfectly with the actual location of the boat. When looking at the boat from image to image, it should ideally line up perfectly, and it does not. This suggests there is error in the GPS of the WASP-lite system, which was expected because of what was known about the internal navigation system. Figure 49 shows the false color rendition of the image, and shows the
boat circled. The line through the center and upper right side of the image is the boat’s path. There does seem to be some variation in color as well.

Figure 49: Goose Bay's false color three band rendition, where red is 550nm, green is 630nm and blue is 650nm. The sampling boat is circled
Figure 50: Two orthorectified images that showing the misalignment of the boat

Figure 50 shows that the contributing error is not only limited to the GPS system of the boat, and the drift of the boat but also the orthorectification of the images. The internal navigation system used for WASP-Lite is not considered very accurate, and other analysis is being done to improve this aspect of the system. In this image, the error seems to be about 30 meters.
Site 50

Figure 51: Site 50's orthorectified image and ground sample location

Figure 51 shows one of the few locations where the ground sample point is in the middle of the orthorectified image. This image and location will be used as an example to help illustrate how the ground errors affect interpreting the results. Figure 52 shows the manually rotated three band color image with the circle representing where the ground sample point should be. Figure 53 then shows the same three band image, (not rotated), with the ground sample location and two circles with radiuses of 30 and 60 meters. From the previous analysis at Goose Bay, we know there is at least a 30 meter error. Now
looking at this figure, we can see that if the ground sample point is anywhere in the 30 meters, it could be red colored or blue colored water. This means we have no way of determining which type of water the ground sample was taken from.

Figure 52: Manually rotated color rendition image, showing approximate location of ground sample point
Figure 53: Site 50's false color three band rendition, showing radius of error of 30 and 60 meters.
St Albans (Inner and Outer)

Figure 54: St Albans Bay, showing inner and outer sampling points and images

The overall image of St Albans, shown in Figure 54, shows where in the bay the inner and outer points were.
Figure 55: St Albans Inner sampling point location and orthorectified image

Figure 55, shows the St. Albans Inner orthorectified image, and the ground sampled point which is off of the image. Figure 56 shows false three color rendition and it can be noticed that the bottom is visible in the upper part of the image. When looking at Figure 54, the overview of St Albans Bay, it can be seen that the Inner image is close to a small island, so it is not unexpected that the Bay is shallow in that area.
Figure 56: St Albans inner false color three band rendition
St Albans Outer

Figure 57: St Albans Outer ground sampling point location and orthorectified image

The ground sample point shown in Figure 57, for St Albans Outer, lines up very well with the image. Figure 58 shows the three band false color rendition, and it seems that there is very little variation of water color in the image.
Figure 58: St Albans outer false color three band rendition

Cole Bay image 764

Figure 59: Cole Bay's ground sample point location and orthorectified image
The Cole Bay ground point also lines up very well with the image. From the image above, Figure 59, you can see that the orthorectified image and ground sample point are close to the shore and to an island. Figure 60, below shows the false three color rendition of the image, and we may be seeing shallow bottom on the right.

![Figure 60: Cole Bay false color three band rendition](image)

**Patchiness of the Algae and Flow Through Data Analysis**

Looking at the water color within one image, it appears that the patchiness of the algae is very dynamic. There are areas where the algae looks to have a swirling pattern, or a striped pattern - where it can vary from non-algal water to algal water within meters. This
will be important to consider when comparing the images to the ground sampled data when performing spectral method analysis.

Additional data was obtained from other research vessels from (SUNY Plattsburgh) that were sampling the same week as ESF and RIT. They were collecting flow through fluorometry data the day after the ESF/RIT sampling. The two pigment concentrations they were measuring are chlorophyll and phycocyanin.

Due to a problem with calibrating the fluorometer, these data can only be used in a qualitative sense. However, the fluorometer data provides the information needed to help show the patchiness of the algae.

Overlaying all of the data obtained (orthoregistered images, ESF ground sample locations, and flow through locations) we can see the overlap of the flow though points in Figure 61.
Figure 61: Missisquoi Bay- Landsat image is the base image, the orthorectified images for each flight line, the ground sample locations and the flow though fluorometry path are shown.
Figure 62: Area near Site 50, showing the orthorectified images and the flow through data path
Figure 63: Flow through fluorometry data for Site 50 area

Figure 62 shows the orthorectified images and the path of the flow through system. The flow though fluorometry values shown in Figure 63. The variability of the values shown here, with a difference of about 5:1, show that the variation that is seen in the images is really in the water. Even though we can not get concentration values for the fluorometry data, we can recall that the concentration of the ESF sample for phycocyanin was relatively high (20s).
The rotated color rendition images in Figure 64, show the varying color of the images that correlate to the varying flow though data. Even though this data is taken a day apart, this example does show that the algae is very patchy by nature.

Figure 64: Images correspond to above figure, left to left image, right to right image. These are the three color renditions, showing the variation of content
**Expected Results**

**Spectral Curvature**

It is a good idea to figure out what kind of results should be expected for varying degrees of spectral curvature. We are measuring how curved the spectrum is between 550, 630 and 650nm. An example with arbitrary radiance numbers, where the value at 550nm is 4, the value at 630nm is 2 and the value at 650nm is 4, then the spectral curvature would be 0.25. Performing the same calculation on a more “curved” spectrum, where the value at 630nm is 1, then the spectral curvature would be 0.0625. Therefore when the spectrum is more concave upward, the spectral curvature is smaller. Because a stronger absorption at the 630nm band, and stronger reflectance in both the 550nm and 650nm bands suggest an increase in cyanobacteria, this will result in a more curved spectrum for which the spectral curvature will be smaller. In other words, when there is more cyanobacteria, the spectral curvature value of the reflected radiance spectrum will be smaller.

**Band Ratios**

The relationship that the radiance band ratios have to the concentration level varies inversely to that of the spectral curvature relationship. Using an example where the value at 550nm is 3 and the value at 630 is 2, compared to when the values are 3 and 1 respectively (more absorption at 630 means more phycocyanin content and lower radiance), then the first scenario would result in a normalized radiance band ratio value of
1/5, and the second scenario (with more phycocyanin) results in a normalized radiance band ratio value of 1/2. This means that more phycocyanin results in a larger band ratio value.

**The Analysis Procedure**

To analyze the images and the multiple methods, as shown in Table 7: Scenarios for image analysis, a standard procedure was used. First, the images to be analyzed were chosen by comparing the ground sampled locations to the orthorectified images. Then, regions of interest (ROIs) were located in the image. The ROI procedure was done for every registered image, the 8 bit registered, the 10 bit registered and these images with or without foam. (every registered image; 8bit register, 10 bit registered, and with or without foam). This was done first without the aid of the PCA image, so regions that appeared to look similar were put into the same region. When PCA was used, the spatial differences seen in the image were used to define the ROIs. The means of the ROIs were then calculated and stored. The different metrics (spectral curvature, the various band ratios) were then performed on the means of the ROIs for each image.

Considering the error in the GPS of both the plane and the boat, the drift of the boat, and the drift of the algae, the exact location of the ground truth can not be determined. When the ground truth point was found to be in an image, it was assumed that it really was somewhere in the image, but not exactly in that point (and this maybe a poor
(assumption). When the ground truth point was outside the image (to the side perhaps), it was assumed that the image may or may not contain the ground truth values. The time difference between the fly over and the ground sample is different for every sample, so this was also considered. There was one image where there boat was in the image sampling at the time (the perfect coordination) but there are some locations with hours of difference.

**Laboratory Analysis of the Ground Truth Samples**

Because we can not visually tell in an image which algae has phycocyanin or not, we analyzed the ground truth obtained by ESF; the chlorophyll $a$ and phycocyanin concentrations. It was expected that there would be a linear relationship between the phycocyanin and chlorophyll concentrations when the algae was cyanobacteria. It was also expected that if there were some points where there was more non-cyanobacteria in the sample, then these points would be off of the linear relationship because they would contain more chlorophyll. Figure 65 shows this is exactly the case. There are two data points that are off of the linear curve, and suggest that these two points contain more non-cyanobacteria type of algae, like green algae, than the other samples.
Figure 65: Graph of chlorophyll concentration vs. phycocyanin concentrations collected while ground sampling

It is important to note that two of these ground sampled points do not have imagery data associated with them. One of the “high chlorophyll” points is one of these samples. Also, there are two sampled locations that do not have chlorophyll concentrations because they were not collected in the field. However, phycocyanin was collected for these two points, and these two samples happen have very high phycocyanin concentrations. This means that we don’t know how the relationship responds with higher phycocyanin concentrations.
Considering the error in the GPS of both the plane and the boat, the drift of the boat, and the lake currents, the exact location of the ground truth can not be determined. When the ground truth point was found to be in an image, it was assumed that it really was somewhere in the image, but not exactly in that point (and this maybe a poor assumption). When the ground truth point was outside the image (to the side perhaps), it was assumed that the image may or may not contain the ground truth values. The time difference between the fly over and the ground sample is different for every sample, so this was also considered. There was one image where there boat was in the image sampling at the time (the perfect coordination) but there are some locations with hours of difference.

**Image Analysis Scenarios**

For the image analysis shown in this section, it is assumed that the images have been flat fielded, radiometrically calibrated, and corrected for lens distortion. This section will compare the results from the two different registration methods, the different spectral analysis techniques, and one artifact removal method. Table 7 shows the possible processing situations that will be discussed here.
Table 7: Scenarios for image analysis

<table>
<thead>
<tr>
<th>Imperfect Registration (10 bit)</th>
<th>Perfect Registration (8 bit) (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No PCA</td>
<td>PCA (2)</td>
</tr>
<tr>
<td>Spectral Curvature (4)</td>
<td>Spectral Curvature</td>
</tr>
<tr>
<td>Minimized Foam (3)</td>
<td>Minimized Foam</td>
</tr>
<tr>
<td>Did Not Minimize Foam</td>
<td>Did Not Minimize Foam</td>
</tr>
<tr>
<td>Band Ratio 550-630 (4)</td>
<td>Band Ratio 550-630</td>
</tr>
<tr>
<td>Minimized Foam (3)</td>
<td>Minimized Foam</td>
</tr>
<tr>
<td>Did Not Minimize Foam</td>
<td>Did Not Minimize Foam</td>
</tr>
<tr>
<td>Band Ratio 650-630 (4)</td>
<td>Band Ratio 650-630</td>
</tr>
<tr>
<td>Minimized Foam (3)</td>
<td>Minimized Foam</td>
</tr>
<tr>
<td>Did Not Minimize Foam</td>
<td>Did Not Minimize Foam</td>
</tr>
</tbody>
</table>

1) The 8 bit data was found to produce similar or less consistent results because of the loss of 2 bits. Even though this data is perfectly registered, the spectral methods used require regions of interest, and are not pixel by pixel operations. The spectral method results from this 8 bit data will not be discussed further.

2) The PCA analysis of the 8 bit data did prove to work well. Because PCA is a pixel by pixel operation, and the 8 bit data is perfectly registered, better results were obtained than with the 10 bit data.

3) The foam-minimized data did provide slightly altered results from the original, foam containing data. Because the foam is optically opaque, the extreme pixels were minimized because they will not offer any spectral content to the imagery. Truer results will be obtained with the foam-minimized data.

4) All three of the spectral methods provided similar results.
5) The discussion of the three spectral method analysis will continue with the 10 bit data. Even though this data is not perfectly registered, the results are acceptable because regions of interest were used, meaning the operations are not pixel by pixel. The ROIs can then be compared to the PCA image produced by the 10 bit data, to check that the ROIs do not overlap varying regions of the image.

**Discussion**

Figure 66 and Figure 67 show an image over Site 50, where the first image is the 10 bit, imperfectly registered data, and the second image is the 8 bit perfectly registered data. The only major difference that can be seen from this view is that the white foam in the 8 bit image is more pure or white because it is better registered. These images show a lot of white foam lines, some wave formation, and not very much glint, if any at all. The false color interpretation shows that the red areas have more algae than the non-red areas. The “redness” of the image also varies, as does the “greenness” and “blueness” throughout the image. This alone suggests that the algae is highly variable.
Figure 66: The 10 bit floating point imperfectly registered image
Figure 67: The 8 bit floating point perfectly registered image

Figure 68 shows a foam-minimized image with a number of ROIs chosen for image analysis. Many regions were chosen to cover the whole spectral range of the image, which may show more differences than visually looking at the color rendition alone.
The PCA image developed from the 10 bit data is shown in Figure 69. Because of the mis-registration, the foam lines and waves are very apparent. This makes it difficult to see the spatial content that PCA is supposed to separate. Figure 70 shows the PCA image from the 8 bit data, which is perfectly registered. The foam lines and waves are not an issue when visually inspecting this image. The borders and edges of the algae can clearly be seen.
Figure 69: The PCA image using the 10 bit imperfectly registered data

Figure 70: The PCA image using the 8 bit perfectly registered data.
Because the PCA image only shows differences in the water spatially, it does not provide any information about the algae itself. This analysis only helps show where the outlines of the algae are, so when comparing ROIs within an image, they can be grouped according to how similar they visually look in the PCA.

**Results from Spectral Method Analysis**

**Spectral Curvature Results with in an image**

First we are going to perform an analysis with an image to illustrate how the process works, and to show the results when only comparing with one image. Figure 71 shows the PCA image for the Brochets location (within Missisquoi Bay). The ROIs that are shown in Figure 72 were chosen using this PCA image. The means of these ROIs were then recorded, from which the spectral curvature was then calculated.
Figure 71: PCA image of Brochets

Figure 72: ROIs chosen for Brochets, based on PCA
The spectral curvature was then ordered from low values to high values with the corresponding ROI color. From previous analysis, we expect that with lower spectral curvature values, there would be more algae and the water would appear red. We would also expect higher spectral curvature values when there is less algae and the water would appear blue. Looking at the ROIs that correspond to the low spectral curvature values, using Figure 73 as a guide (because you can easily see the color and patchiness of the water in this image), that the maroon, yellow, magenta, and cyan ROIs are clearly over red colored water. The green, red and blue ROIs, which correspond to higher spectral curvature values, clearly correspond to bluer colored water. This suggests that within a scene the spectral curvature analysis provides consistent qualitative results when comparing the spectral curvature values to the apparent color of the water. Similar
results were obtained with the band ratio methods. The only time consistent results were not clearly obtained was when the variation in the water was not obviously patchy or variable. In that case, it is difficult to verify the method.

Table 8: Spectral curvature with corresponding ROI color for Brochets image

<table>
<thead>
<tr>
<th>ROI Color</th>
<th>Spectral Curvature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maroon</td>
<td>0.563</td>
</tr>
<tr>
<td>Yellow</td>
<td>0.563</td>
</tr>
<tr>
<td>Magenta</td>
<td>0.563</td>
</tr>
<tr>
<td>Cyan</td>
<td>0.572</td>
</tr>
<tr>
<td>Green</td>
<td>0.581</td>
</tr>
<tr>
<td>Red</td>
<td>0.599</td>
</tr>
<tr>
<td>Blue</td>
<td>0.608</td>
</tr>
</tbody>
</table>
Results Comparing Location to Location (or Image to Image)

Figure 74: Spectral curvature results vs. phycocyanin of 10 bit data, foam-minimized. The lines show the expected slope of the relationship between the phycocyanin concentration and the sampling locations.

Figure 74 shows the spectral curvature vs. phycocyanin for the 10 bit data, foam-minimized. The previous example reviewed the procedure that was used on each image. Here we will see the comparison from image to image (or location to location). To reiterate this procedure, the ROIs were defined for each image that was closest to the ground sampled location. The mean of the ROIs were then used to calculate the spectral curvature (or band ratio) for each image. For example, the data points for CenterMB
represent the spectral curvature for the 12 or so ROIs chosen for that image. Because we don’t know exactly where in the image (if it is at all) the ground sample point is, we had to take ROIs from the whole image. This is then plotted with respect to the concentration found when performing ground sampling in the water. Overall, we can only look at the trend of the values, keeping in mind the many issues that may alter the accuracy of this analysis, such as time lags between ground sampling and the airborne collection.

Remembering the review of what is expected from the spectral curvature results, as there is more phycocyanin, we should expect to see a lower spectral curvature value. Because we assumed that the water in the image appears red (when displaying in false color), because the chlorophyll in the algae is reflecting at 550nm, (which is displayed as red), we can try to correlate with in an image how “red” the water looks to how “curved” the spectra is. Looking within each of these images, the assumption holds true. The ROIs with “redder” water does have a lower spectral curvature value. Now comparing location to location, we will first only consider the four locations within Missisquoi Bay (Center, Brochets, Site50, and Goose). Recalling that the time lag varies- there is a zero time difference between the ground sample and the airborne collect for Goose, but there is about 2 hours (the longest difference) for Site 50. Keeping this in mind, when looking at the four locations, there seems to be a slight trend- as the concentration goes from low to high, the spectral curvature goes from high to low. Again, because the ground samples were not all taken exactly when the airborne data was collected, we can not be exactly sure the data contains water that is represented by the ground sample.
Now looking at the other two general locations, the two sample locations from St. Albans Bay, and the one sample from the southern bay area - Cole Bay, it can be seen that they are on a different trend than the Missisquoi Bay data points. The first consideration is the time difference between the when the ground sample was taken and the airborne collect occurred. For these two points the time difference is about 2.5 and 3 hours. The Cole ground sampling occurred the previous day. Even though water content can change within hours much less days, the Cole Bay location is historically known for very low algal content. Because of the 2.5-3 hour time difference for St. Albans Bay, it is difficult to explain why the spectral curvature is what it is. The ground truth is not adequate.
Figure 75: Spectral curvature results vs. chlorophyll of 10 bit data, foam-minimized. The line shows the expected slope of the relationship between the chlorophyll concentration and the sampling locations.

Figure 75 shows a plot of spectral curvature vs. chlorophyll concentration. Again, as with the phycocyanin concentration, we expect that the spectral curvature to have lower values with a higher concentration. Notice that the Brochets data points do not come close to following the “correct” trend. From our previous analysis of chlorophyll concentrations vs. phycocyanin concentration, this site contains more chlorophyll than the others with respect to phycocyanin, and therefore contains more non-cyanobacteria.
type of algae. Therefore, the spectral ratio value will be skewed from the expected trend. (the chlorophyll maybe higher values, but the phycocyanin is not, so the curviness of the spectrum is skewed.)

Notice that this graph does not have any data points for the Center and Site50 locations. This is because, as mentioned before, samples were not collected in the field for chlorophyll - only phycocyanin.

Recall there is a time difference between when the ground sampling was performed to when the airborne collect occurred. For Brochets and Goose Bay the time difference was 25 and 0 minutes respectively, and at the two St Albans points, the time difference was 2.5-3 hours, and the difference for Cole Bay was one day.
Figure 76: Band Ratio (550-630nm) results vs. phycocyanin of 10 bit data, foam-minimized. The line shows the expected slope of the relationship between the phycocyanin concentration and the sampling locations.

The band ratio combination of 550-630nm vs. phycocyanin, (Figure 76), shows very similar results when compared to the spectral curvature vs. phycocyanin concentration. The overall trend of the data is inverted- as the band ratio increases, the concentration increases. Regardless of this, the discussion for the spectral curvature results (in Figure 74) hold true for this result as well.
Figure 77: Band Ratio (550-630) results vs. chlorophyll of 10 bit data, foam-minimized. The line shows the expected slope of the relationship between the chlorophyll concentration and the sampling locations.

The Band Ratio of 550-630nm vs. chlorophyll, shown in Figure 77, also had very similar results when compared to the spectral curvature vs. chlorophyll. The only difference is that the trend is inverted, just as the previous graph (vs. phycocyanin).
Figure 78: Band Ratio (650-630) results vs. phycocyanin of 10 bit data, foam-minimized. The line shows the expected slope of the relationship between the phycocyanin concentration and the sampling locations.

The band ratio analysis of 650-630nm vs. phycocyanin, shown in Figure 78, shows a decent trend. It is expected that as the phycocyanin concentration increases, the band ratio increases. Most of the data points do this except for the Cole Bay data points.

A plot of this band ratio vs. chlorophyll is not analyzed because we do not expect to find any correlation of either of these bands, 650 or 630nm to chlorophyll. The other two
methods are useful to analyze because the 550nm band is included, which reflects chlorophyll.

**Signal to Noise Investigation**

Other than the spatial errors, another consideration for possible error could be the signal to noise of the detector or of the scene. The detector signal to noise was approximated using two different methods. One method included using an in-scene procedure over land, where a pond and other targets of different radiance values were visible. The signal to noise was plotted vs. the radiance values of the scene (as radiance increases, so does signal to noise). The worst case signal to noise was about 40:1. The second method was performed by using a dark image (taken with a lens cap fixed on the lens), where the worse case ratio is expected. The signal to noise of that image was also about 40:1. Because we are analyzing water in the images, we will assume that the worst signal to noise of the water in the image could be 40:1. The next step was to compare this signal to noise ratio to other satellite systems that were designed and are used for water analysis. The SNR was found for both SeaWiFS and MODIS for the same radiance values that were found in our water images. The SNR values were 670 and 1077 respectively. (Esais et al. 1998). We then wanted to find out how big we would have to make a WASP-Lite pixel to obtain a SNR of about 800 (a value between the SeaWiFS and MODIS systems). Because 800 is larger than 40 by a factor of 20, a 20 by 20 pixel (or 400 total pixels) of
WASP-Lite would be required. When comparing this value to the ROI sizes used for the analysis of the images, the ROIs sizes are much large (from 1200 to 3000 pixels). This means that the detector signal to noise is adequate when using ROIs that contain at least 400 pixels.

The next analysis was to determine the in-scene signal to noise. Figure 79 shows the original image from Site 50, showing band 2 (550nm). Figure 80 shows the horizontal profile of this image. The foam and sun glint create in-scene noise. This effect can be seen in the horizontal profile. The slight variation of the bright and dark thick algal “stripes” in the image, are barely noticeable in the profile. The signal to noise appears to be about 1:1. The variation caused by the foam is just as much as the actual variation in the water due to algae.
A low pass smoothing kernel was applied to this image, resulting in Figure 81. This kernel 21x21, which should result in a signal to noise of the detector of about 800. The horizontal profile of this image is shown in Figure 82, where the signal to noise appears to be about 5:1. The variation due to the algal content in the water is more noticeable.
Figure 81: Site 50 (band 2) with 21x21 low pass smoothing kernel applied

Figure 82: Horizontal profile of Site 50 image with 21x21 smoothing

In attempt to morph the image to obtain a spatial resolution that represents Landsat (30 meters), a 41x41 sized low pass filter was applied to the original image. As shown in Figure 83, the lines from the waves and foam are less noticeable, and the variation due to
the algal content is more noticeable. The horizontal profile shown in Figure 84, shows that the in-scene signal to noise has greatly increased, to about 20 or 30:1.

Figure 83: Site 50 (band 2) with 41x41 low pass smoothing kernel applied

Figure 84: Horizontal profile of Site 50 with 41x41 smoothing
Most satellite systems designed for monitoring the oceans do not have this in-scene noise problem that is due to the very high spatial resolution of these airborne images. Satellite images are processed in a way that does account for the white caps, or foam in an image by a value that is a function of wind speed. This value is then removed from the whole image. Because we can actually resolve the white foam, we need to devise a different processing method to remove this. Because we are examining fresh water lakes with high organic matter concentration, the white foam is different than the white caps that are in the salt water open oceans. This white foam in the lakes requires its own research project to characterize it over different weather patterns and algae conditions. No research has been found that analyzes this white foam phenomenon that causes such problems with in scene noise of high spatial resolution systems.

**Foam and In-scene Noise Removal Test**

One additional test was performed to test the application of spectral angle mapper (SAM) to the white foam images, to assess its success in removing the foam. Figure 85 shows the foam minimized image that was used for the processing for this experiment. In attempt to remove the foam completely, spectra of the water were collected as endmembers for the spectral angle mapper tool. Figure 86 shows the SAM image, where the red pixels represent the “red colored water” endmember, and the white pixels represent the “blue colored water” endmember. The black pixels represent the white foam areas. This SAM image was then applied to the original image as a mask, resulting
in Figure 87. Looking closely at this image, it did seem to remove all of the foam. There may be some spots where waves and sun glint still show through, so the sun glint algorithm could be tested on this image.

Figure 85: Site 50 color image, showing foam minimized with ROIs used
Figure 86: SAM image from Site 50

Figure 87: Site 50 Foam minimized using SAM, and showing ROIs used
The spectral curvature results for Figure 85 (original method with foam minimized) and Figure 87 (new method with SAM removing foam) are shown in Table 9. The ROIs are compared for the two methods, and they are grouped according to different types of font. For example, the blue font ROI colors represent four ROIs that are grouped together when comparing the two methods. Overall, these in scene results show that the spectral curvature values using the two methods do not seem to produce significantly different results. The spectral curvature values for both methods show the correlation of apparent red colored water to low values, and blue colored water to high values.
<table>
<thead>
<tr>
<th>ROI Color</th>
<th>Spectral Curvature (original)</th>
<th>ROI Color</th>
<th>Spectral Curvature (using SAM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>magenta</td>
<td>0.575397</td>
<td>magenta</td>
<td>0.552728</td>
</tr>
<tr>
<td>sienna</td>
<td>0.557049</td>
<td>sienna</td>
<td>0.537706</td>
</tr>
<tr>
<td>sea-green</td>
<td>0.556355</td>
<td>sea-green</td>
<td>0.536584</td>
</tr>
<tr>
<td>thistle</td>
<td>0.554726</td>
<td>cyan</td>
<td>0.535292</td>
</tr>
<tr>
<td>orchid</td>
<td>0.55018</td>
<td>thistle</td>
<td>0.534805</td>
</tr>
<tr>
<td>cyan</td>
<td>0.549151</td>
<td>blue</td>
<td>0.529862</td>
</tr>
<tr>
<td>blue</td>
<td>0.546911</td>
<td>orchid</td>
<td>0.529578</td>
</tr>
<tr>
<td>chartreuse</td>
<td>0.539823</td>
<td>maroon</td>
<td>0.526523</td>
</tr>
<tr>
<td>green</td>
<td>0.537795</td>
<td>red</td>
<td>0.526187</td>
</tr>
<tr>
<td>red</td>
<td>0.537476</td>
<td>green</td>
<td>0.523644</td>
</tr>
<tr>
<td>purple</td>
<td>0.536499</td>
<td>purple</td>
<td>0.521257</td>
</tr>
<tr>
<td>maroon</td>
<td>0.533164</td>
<td>chartreuse</td>
<td>0.521071</td>
</tr>
<tr>
<td>yellow</td>
<td>0.527873</td>
<td>aqua</td>
<td>0.519529</td>
</tr>
<tr>
<td>aqua</td>
<td>0.526911</td>
<td>yellow</td>
<td>0.517384</td>
</tr>
<tr>
<td>coral</td>
<td>0.516224</td>
<td>coral</td>
<td>0.517234</td>
</tr>
</tbody>
</table>
**The Phenomenon of Wave Focusing**

When closely analyzing the imagery and looking at the white foam, it was noticed that a dark line of pixels would be between the white line of white foam and colored water pixels. This is shown in both Figure 89 and Figure 90. One possible explanation of this could be wave focusing, where some areas of the water allow the sun light to penetrate the water and the colored water to be visible. The areas where the sun light does not “focus” could be where there are dark pixels in the image.

![Figure 89: Zoomed window of Site 50 image, showing the white foam, dark pixels and colored pixels of apparent “blue” water](image)
As discussed in the objectives section, there are two main benefits to our system over previous systems and methods used: unique band selection capability to allow the sensor to be customized to this application, and very high spatial resolution, which had not been used by previous researchers. The high spatial resolution offered more insight than was expected, allowing us to understand why monitoring cyanobacteria with airborne and satellite data has been so difficult. The use of correct spectral bands allowed variation in the water to be observed, and the high spatial resolution allowed the scale of the algal variation to be observed. Because this spatial variation was unexpected, the experimental technique used was not optimal for testing algorithms (matching image analysis to ground truth values). The experimental techniques that need to be improved upon include increasing the number of ground sampling points used, decreasing the time delay between ground sampling and the airborne collect, reducing boat drift, minimizing sun
glint and white foam, improving the signal to noise ratio of the cameras and the orthorectification of all the images.

Specifically, the number of ground truth samples that were collected were insufficient for the tests done here. Even though standard sampling procedures were followed, the patchiness of the algae was unknown, and it was unknown that because of this patchiness, many more sampling points would be needed. The patchiness was observed in the imagery, and was correlated to flow through fluorometry data that was collected one day later. Because of the time difference, the exact location of the patches were not correlated. The quantitative observation of variation in phycocyanin and chlorophyll concentrations in the general area were correlated. To accurately capture the patchiness of the algae, it would be recommended in the future, that a flow through system would be used concurrently with an airborne data collect with WASP-Lite. The high spatial resolution of WASP-Lite would provide a good match with fluorometry data, because specific algal bloom lines and boundaries could then be compared and correlated.

The temporal resolution of the ground truth was that of standard sampling techniques, but again, because of the patchiness, the temporal resolution needs to be improved. The overall time it took the airborne system to collect all of the data points was about one hour. The ground sampling spanned over about 5 hours. This could be improved with the use of the flow through fluorometry data, only stopping to collect calibration samples
for the fluorometer. In other words, only necessary stopping and sampling for this experiment should be performed, to keep the ground sampling as short as possible. It would also be suggested that the plane fly in the same path of the sampling boat (or as close as possible), and that it fly two times- one when the boat sampling starts and once when the boat sampling is finishing. This will help to help correlate ground truth because the boat should be in two images while it was sampling, the time lag would be zero, and it would allow the drift of the algae to be better understood and seen in the imagery.

The other temporal issues involved the drift of the boat while it was stopped for about 20 minutes to perform sampling. This would be avoided if the only stops were to quickly collect water samples for later analysis.

The detector signal to noise ratio was calculated to be adequate when using a ROI of at least 400 pixels, so it would be recommended to continue using ROIs of this size or larger.

The sun glint and white foam issues that were problematic here, were not understood as well as they should be. With this high resolution imagery, these two different phenomenon could be resolved with better detail then usual. The white foam minimization seemed to work adequately here, but only for the extremely bright foam lines. The test using the spectral angle mapper tool did seem to completely remove the
white foam. It would be recommended that the sun deglinting algorithm should then be run on the resulting image. This should then dramatically decrease the in scene noise. It would be recommended that the in scene signal to noise should be calculated after this processing was performed.

One task that was not performed here, that would definitely help with the suggested experiment of using flow through fluorometry data, would be to orthorectify all of the cameras- not just one. This would allow for the images to be mosaiced, and the fluorometry data could then be overlaid on top of the mosaiced imagery. This would allow much better correlation with the ground sampled data.

The impact that this experiment could have on the scientific community is crucial because the variation and patchiness of the algae that was observed shows that the standard sampling techniques and methods are not adequate for scientific research. Using the flow through fluorometry system, as suggested earlier, would dramatically improve scientific experiments on cyanobacteria. Using an airborne sensor could provide even more information if the previous suggestions were applied to the experiment, so the scientific community could use this airborne data in collaboration with their own cyanobacterial research.
Another way this work would create an impact on the scientific community if the recommendations were followed, the final product of a concentration map of cyanobacteria would allow satellites to have a “truth” that could be tested against, to determine if the satellite and its algorithms work successfully on broad scale detection of the cyanobacteria.

The greatest impact this work would could provide, if all of the recommendations are performed and the cyanobacteria concentration map is produced, would be for the monitoring of lakes for notifying the public of when it is safe to use for recreational purposes.

Overall, the experiment did not allow the detection and mapping of cyanobacteria but did give insight to how the algae exists in lakes and how to improve the methods to result in cyanobacterial detection. Through the high spatial resolution and unique band selection, the high patchiness, or variability of the algae, has been shown to be a crucial issue when attempting to map cyanobacteria by correlating the imagery to ground truth. Through making the suggested improvements to the techniques, there would be a great improvement in the results obtained.
References


Leica Geosystems, ERDAS IMAGINE Tour Guides 9 Dec 2005


Appendix A

Phycocyanin Extraction Protocol (Performed by ESF)

For Glass Fiber Filters (GF/F):

- Take the GF/F filter and unroll it carefully from the 13x100 test tube and place it flat (with cells facing up) at the bottom of a 100 mL beaker.
- To the beaker, add 10 mL of a 10 mM phosphate buffer (pH 6.8).
- Place the beakers into the freezer (-20°C) for approximately 35 minutes, or until the samples are completely frozen.
- Remove the beakers from the freezer after the allotted time and place them into the bath sonicator found in the cold room (4°C). In the bath sonicator, place ice packs around the samples to keep the water cold during sonication. Sonicate the samples for approximately 35 minutes, or until the samples are completely thawed. Put the samples back into the freezer.
- Repeat this freeze/thaw method for a total of 3 cycles.
- After the last thawing cycle, remove the filters from the beakers and transfer the samples to a plastic extraction tube. (Perform this in the dark)
- Before centrifuging, turn on the 10-AU, which has been calibrated for extracted phycocyanin.
- Centrifuge the samples at 10,000xg for 15 minutes for clarification.
- Transfer the supernatant into 13x100 test tubes and place the tubes on ice and temporarily store them in the dark.
- Read the supernatant on the 10-AU Fluorometer and record the fluorescence (Perform this in the cold room in the dark with only the red light bulb on).
- To determine the amount of phycocyanin in the lake, use the following correction equation:

\[
PC (\mu g/mL) = \frac{10\text{-AU PC} \times \text{Volume Extracted} \times \text{Dilution Factor}}{\text{Volume Filtered}}
\]

Where 10-AU PC is the fluorescent concentration obtained from the Turner Designs 10-AU Fluorometer.
Appendix B

Chlorophyll a Extraction Protocol (Performed by ESF)

1. Add 8 mL of 90% cold acetone to the 13x100 test tube containing the chlorophyll filter.
2. Bath sonicate the sample for 1 hour with ice packs to keep the water cold.
3. After sonication, remove the filter from the test tube. Make sure to keep the lights off while performing this step.
4. Place the sample in the freezer (-20°C) for 2 hours.
5. Turn on the TD700 fluorometer which has been calibrated for extracted chlorophyll a. Allow the fluorometer to warm up for 10 minutes before calibrating with the solid standard and 90% acetone as a blank.
6. Read the samples in the TD700 fluorometer. Dilute as needed with 90% acetone.
7. Calculate the amount of chlorophyll a in the sample with the following formula:

\[
\text{Corrected Chl } a \text{ (ug/L)} = \frac{\text{TD700 Reading} \times \text{Volume Extracted} \times \text{Dilution Factor}}{\text{Volume Filtered}}
\]

In this formula, volume extracted is 8 mL, and the volume filtered is 15 mL.