

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Public Health Surveillance of Toxic Cyanobacteria in Freshwater Systems Using Remote Detection Methods

Permalink

<https://escholarship.org/uc/item/1w3686nh>

Author

Mackie, Trina Nicole

Publication Date

2010

Peer reviewed|Thesis/dissertation

Public Health Surveillance of Toxic Cyanobacteria in Freshwater Systems
Using Remote Detection Methods

by

Trina Nicole Mackie

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Health Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Robert Spear, Chair

Professor Michael Jerrett

Professor Peng Gong

Fall 2010

Public Health Surveillance of Toxic Cyanobacteria in Freshwater Systems
Using Remote Detection Methods

© 2010

by Trina Nicole Mackie

ABSTRACT

Public Health Surveillance of Toxic Cyanobacteria in Freshwater Systems Using Remote Detection Methods

by

Trina Nicole Mackie

Doctor of Philosophy in Environmental Health Sciences

University of California, Berkeley

Professor Robert Spear, Chair

Cyanotoxins, the group of toxic chemicals produced by the blue-green algae or cyanobacteria that can proliferate in fresh and salt-water, cause a range of harmful health effects including skin rashes, flu-like symptoms, nausea, diarrhea, tingling and nerve damage, liver damage, tumors, and death. Although cyanobacteria are one of the oldest organisms on the planet, anthropogenic development (e.g. dams, water diversions, nutrient rich runoff from intensive agriculture, decreased impervious area from urbanization, etc.) has caused many watersheds to lose substantial water volume, suffer tremendous inputs of nutrients and other organic and inorganic pollutants, increase in temperature and overall become more suitable for the proliferation of harmful blooms of cyanobacteria. Cyanobacteria blooms are now increasingly prevalent in freshwaters as eutrophication becomes ever more common with human stressors, and climate change is likely to only further exacerbate this problem. Public health professionals are dependant upon early and dependable information on the presence, concentration, and location of cyanobacteria blooms in order to inform the public and reduce potential exposure. This research project evaluated the efficacy of remote sensing data to provide this kind of surveillance and early detection for characterizing the presence of toxic algae in freshwater systems. It explored the relevance of specific remote sensing techniques to freshwater cyanobacteria bloom identification. The various remote sensing platforms available for this kind of research vary in cost, swath coverage, spatial scale and spectral resolution. For this study, three different remote imagery platforms were compared in terms of their ability to identify surface blooms and to distinguish gradients in cell density or bloom intensity. This exploration of the application of remote sensing used a hyperspectral airborne sensor with high spatial resolution (SpecTIR), a multispectral satellite image also with high spatial resolution (IKONOS), and a lower spatial resolution multispectral satellite image (Landsat). Water sampling data (algal pigment concentrations, turbidity, transparency, and temperature) from known blue-green algae blooms dominated by *microcystis aeruginosa* on the Iron Gate and Copco Reservoirs on the Klamath River were used to evaluate and classify these different images.

This research successfully used both satellite and airborne remotely sensed data to visualize where the medium or high density sections of the bloom were located, to quantify the intensity of the bloom in terms of area impacted, and to compare the intensity of the bloom at different dates

in time. Remote sensing can provide a synoptic overview of the entire system, making it possible to truly assess relative bloom intensity. Furthermore the results indicate that when given the choice, the investment in higher spectral resolution should be chosen over higher spatial resolution as the former appears to provide more benefits in cyanobacteria detection.

ACKNOWLEDGEMENTS

I am extremely grateful for the collaboration of Professor Zou Xiaobing at Chongqing University in China and the support of her students Yongyue, Zeng Ting, and Yangfan. I am also especially appreciative of the support and collaboration from Miao Ling He in my first field work on the Yangtze River and in continued communications with collaborators in China.

I am truly grateful to my advisor Robert Spear who was willing to support me in pursuing a new research area and gave me the space to develop my ideas. Bob's subtle positive affirmation was invaluable. He always had an appropriate level of optimism, and though he never told me exactly what to do or how to do it, he let me know that I could do it. He was there as I needed support, skillful in helping me navigate the PhD process, and always let me find my own way.

I owe much to the constant discussion of varied issues and questions related to my own research which occurred within the Klamath Blue-Green Algae Work Group, and I am thankful to the many participants (particularly Gail Louis and Sue Keydel of the U.S. Environmental Protection Agency Region IX, Susan Corum of the Karuk Tribe, Ken Fetcho of the Yurok Tribe, and Linda Prendergast of PacifiCorp) for feedback, encouragement, and loaning boats and equipment for fieldwork on the Klamath River.

Professor Maureen Lahiff (SPH, UC Berkeley), Karen Tuxin-Bettman (Geospatial Innovation Facility, UC Berkeley), and Kevin Koy (Geospatial Innovation Facility, UC Berkeley) all provided invaluable help troubleshooting and addressing software issues and statistical data questions.

Gregg Langlois at the California Department of Public Health and Mark Bettencourt of the California Department of Water Resources, loaned equipment and offered helpful insight on phytoplankton speciation and enumeration during the early stages of my research development. Dr. Bill Draper, also at the California Department of Public Health, provided support and advice on cyanotoxin analyses. Stephen Andrews at the Environmental Sciences Teaching Program of UC Berkeley was also very helpful in loaning and providing guidance on water sampling equipment.

The primary sources of funding for this work came from the CC Chen China Research Fellowship, the Bay Area Water Quality Fellowship, the Center for Occupational and Environmental Health Student Research Award, and the Pacific Rim Research Program Graduate Fellowship.

TABLE OF CONTENTS

Acknowledgements	i
Chapter 1: Introduction and Background	1
1.1 Public Health Risks of Cyanobacteria	1
1.2 Cyanobacteria	2
1.3 Aquatic Remote Sensing	4
1.4 Study Objectives and Rationale	5
1.5 Contributions of this Dissertation	6
1.6 Organization of Dissertation	6
Chapter 2: Cyanobacteria and Water Quality Characterization	12
2.1 Introduction	12
2.1.1 Klamath Field Site	12
2.1.2 Yangtze Field Site	14
2.2 Water Quality	15
2.2.1 Water Quality in the Klamath Field Site	15
2.2.2 Water Quality in the Yangtze Field Site	17
2.3 Water Quality and Field Data Collection	17
2.3.1 Water Data Collection (Klamath Field Site)	17
2.3.2 GPS Data Collection (Klamath Field Site)	18
2.3.3 Water Data Collection (Yangtze Field Site)	19
2.3.4 Water Sampling (Klamath Field Site)	19
2.3.5 Water Sampling (Yangtze Field Site)	21
2.4 Results	23
2.4.1 Results (Klamath Field Site)	23
2.4.2 Results (Yangtze Field Site)	25
2.5 Temporal and Spatial Variability	32
2.6 Discussion	33
Chapter 3: Hyperspectral Image Analysis	40
3.1 Introduction	40
3.1.1 Absorbance and Reflectance Properties of Algal and Plant Pigments	40
3.2 Methods and Results	51
3.2.1 Image Acquisition	51
3.2.2 Image Pre-processing	51
3.2.3 Unsupervised Classification	54
3.2.4 Linear Regression	54
3.2.5 Linear Discriminant Analysis	55
3.2.6 Decision Tree	61
3.3 Discussion	71
Chapter 4: Multispectral and Hyperspectral Image Comparison	74
4.1 Introduction	74

4.2 Methods	74
4.2.1 Image Acquisition	74
4.2.2 Decision Trees	75
4.3 Results	75
4.3.1 Decision Tree for IKONOS	75
4.3.2 Decision Tree for Landsat	78
4.4 Discussion	80
 Chapter 5: Temporal Comparison	 91
5.1 Introduction	91
5.2 Methods	91
5.3 Results	92
5.4 Discussion	93
 Chapter 6 Conclusions and Future Studies	 97
6.1 Conclusions	97
6.2 Policy Framework and Contrast between California and China	99
6.3 Future Studies	101

1 Introduction and Background

1.1 Public Health Risks of Cyanobacteria

Surveillance and early detection are keys to effective public health disease prevention. Whether dealing with chronic heart disease, cancer, or infectious disease, the rate of successful treatment and often prevention increases with a well designed program to educate the public about measures to reduce their risk and to detect early onset symptoms or risk factors in time to take action. Cyanotoxins are no exception to the rule. This group of toxic chemicals produced by the blue-green algae or cyanobacteria that can proliferate in fresh and salt-water can cause a range of harmful health effects including skin rashes, flu-like symptoms, nausea, diarrhea, tingling and nerve damage, liver damage, tumors, and death (Pouria et al. 1998; Chorus and Bartram 1999). Both epidemiologic and toxicological studies have also linked chronic low dose exposures to cyanotoxins with the promotion of liver tumors (Nishiwaki-Matsushima et al. 1992; Yu et al. 2001; Zhou et al. 2002). The potency of cyanotoxins is evident from animal toxicology and the accidents that have led to large human fatality and illness. Most notable are the tragedies in Brazil and Australia. In 1988 in Brazil, 88 deaths occurred from cyanotoxin contamination of a drinking water reservoir (Teixeira Mda et al. 1993). A severe gastroenteritis epidemic was caused by cyanotoxin contamination of the reservoir water used in a dialysis clinic that led to 26 deaths from acute liver failure (Jochimsen et al. 1998; Pouria et al. 1998; Carmichael et al. 2001). In Australia, 100 children were hospitalized for gastroenteritis after a cyanobacterial bloom contaminated drinking water (Falconer et al. 1983; Griffiths and Saker 2003). In addition to consumption of contaminated drinking water, people can be exposed through everyday recreation (swimming, boating, fishing, etc.) or inhalation of aerosolized toxin (Backer et al. 2003; Annadotter et al. 2005). “Healthy people in healthy places”¹ means that poor health can be directly tied to places with toxic exposures. Many people make their living off of the water, live in close proximity to the water, or play and recreate regularly in lakes, rivers and the ocean. For them, chronic exposure, even to low levels of cyanotoxins, is a real risk.

Contamination of drinking water supplies has the greatest potential to create the largest burden of disease. While low levels of microcystin have been detected in treated drinking water on a few occasions in the U.S., large American water treatment plants are largely able to adequately remove (i.e. below detection levels) cyanotoxins through activated carbon filters and ozonation. Smaller treatment plants and those in developing countries however, work with less rigorous treatment technology. Microcystin, the most common cyanotoxin, is also particularly stable and is unaffected by boiling water, a standard approach in the developing world to protect themselves from microorganisms in their drinking water supply. Exposure may occur by water contact, ingestion or inhalation when toxins are aerosolized. The toxins do not penetrate the skin, so strictly contact should result in only minor skin irritations and rashes. Inadvertant ingestion during recreation or ingestion of contaminated drinking water produces an array of gastrointestinal symptoms including nausea, vomiting, diarrhea, as well as other cold and flu symptoms such as fever.

In 1998 the World Health Organization issued recommendations for guidelines for drinking water and recreational water to protect public health from cyanotoxins (Falconer et al. 1999;

¹ CDC’s Health Protection Goals include: “Healthy People in Healthy Places – The places where people live, work, learn, and play will protect and promote their health and safety, especially those at greater risk of health disparities.”

World Health Organization (WHO) 2003). That same year the U.S. Environmental Protection Agency (USEPA) added cyanotoxins to the Contaminant Candidate List, a list of unregulated contaminants which may need regulations under the Safe Drinking Water Act (Environmental Protection Agency 2005). In 2004, Congress also expanded the Harmful Algal Bloom and Hypoxia Research & Control (HABHRCA) Act of 1998 to require interagency research on both freshwater and marine harmful algal blooms. For the last several years the USEPA, the National Oceanic and Atmospheric Administration (NOAA), and the National Aeronautics Space Administration (NASA) have released interagency requests for proposals on the causes, effects, detection and control of harmful algal blooms (USEPA 2005; USEPA 2007). The U.S. Centers for Disease Control (CDC) has been conducting studies to evaluate the public health impact of cyanotoxins and the efficacy of water treatment practices in removing them from the drinking water supply (Centers for Disease Control and Prevention; Backer 2002). The CDC's National Center for Environmental Health (NCER) conducted a pilot study to assess recreational exposure to cyanobacterial toxins in the same two reservoirs of the Klamath River that will be used as a field site in this study (Backer et al. 2010) and a previous study of exposure to a bloom in a Michigan lake (Backer et al. 2008). All of this demonstrates the increasing acknowledgement of the need for research in this area.

1.2 Cyanobacteria

Although cyanobacteria are one of the oldest organisms on the planet, they are increasingly prevalent in freshwaters where they could not have survived 100 years ago. This is because our anthropogenic development (e.g. dams, water diversions, nutrient rich runoff from intensive agriculture, decreased impervious area from urbanization, etc.) that stresses the environment has caused many watersheds to lose substantial water volume, suffer tremendous inputs of nutrients and other organic and inorganic pollutants, increase in temperature and overall become more and more suitable for the proliferation of harmful blooms of cyanobacteria. Climate change has the potential to further exacerbate the issue. In many ways the problem resembles an emerging infectious disease and has been treated as such by numerous local and state health departments.

Cyanobacteria are remarkably adept at growing in many different climates, and they occur worldwide from South America to Asia (Smith 1956). This along with the widespread trend in eutrophication worldwide (fueled by anthropogenic development and climate change) is making cyanotoxins an important emerging public and ecological health concern. With climate change continuing to raise temperatures around the globe, it is reasonable to expect rising water temperatures will play a role in increasing eutrophication and the proliferation of cyanobacteria. This research is therefore relevant to many communities nationally and abroad. When water sources become contaminated with cyanotoxins, the size of the population at risk can be millions of people as was recently the case with a bloom in Taihu Lake near Wuxi, China (Ang 2007; Anonymous 2007; Kun 2007; Xiaofeng 2007).

There are many different types of cyanobacteria and when conditions are right, they may have a competitive advantage over other phytoplankton and can proliferate into dense blooms which move up and down in the epilimnion and are visible from above due to their pigmentation and buoyancy on the water surface. In contrast to true algae, many species of planktonic cyanobacteria possess specialized intracellular gas vesicles, which enables the organism to regulate its buoyancy and thus to actively seek water depths with optimal growth conditions.

Some cyanobacteria can fix nitrogen giving them an important competitive advantage in nitrogen limited systems.

Cyanobacteria have been and continue to be referred to as blue-green algae, a misnomer given that their cellular organization and structure is now known to have more in common with bacteria than algae, but they are also photosynthetic like other algae and the name emerged from their characteristic color from the pigments they contain. These pigment compounds (chlorophylls, carotenoids, and biliproteins/phycoobilins) can be used to distinguish individual phytoplankton genera, and chlorophyll concentrations are often used as a general proxy or phytoplankton population size. This is the basis for the use of remote sensing in the electromagnetic spectrum to identify blooms of different compounds and sizes.

Cyanobacteria produce toxic compounds referred to collectively as cyanotoxins. There are many different cyanobacteria species and many different cyanotoxins. Not all cyanobacteria produce cyanotoxins, but many cyanobacteria produce multiple cyanotoxins. At least 46 species have been shown to cause toxic effects in vertebrates (Sivonen and Jones 1999), but there are thirteen that are most commonly found in freshwater. The most common toxin-producing cyanobacteria are *microcystis* spp., *anabaena* spp., *aphanizomenon* spp. and *nodularia* spp. The cyanotoxins they produce can be grouped into hepatotoxins, neurotoxins, and dermatotoxins.

Microcystis aeruginosa is the type of cyanobacteria found to dominate in the Iron Gate and Copco reservoirs of the Klamath River. It is a unicellular organism that forms blooms in eutrophic fresh and brackish waters. The toxin it most commonly produces is microcystin, a potent liver toxin which inhibits eukaryotic protein phosphatases types 1 and 2A. Microcystin has more than 20 structural variants, with varying toxicities (Sivonen and Jones 1999) and there is evidence that it bioaccumulate at higher trophic levels (e.g. zooplankton, shellfish (Lehman et al. 2005), fish (Xie Liqiang 2005))

Although primary productivity in lakes is often P limited, Upper Klamath Lake is (P)-rich (committee on endangered and Threatened Fishes in the Klamath River Basin, National Research Council, 2004), and the nitrogen fixing *aphanizomenon flos-aquae* regularly blooms there (Carmichael et al. 2000), as it does in Iron Gate and Copco Reservoirs (Kann and Asarian 2007) which are N limited (Moisander et al. 2009). The availability of N may therefore play a role in the dynamics of *Microcystis* blooms which are not capable of nitrogen fixation.

Toxin production may vary between different strains of *Microcystis* and within the same toxin-producing strain depending on environmental conditions. Some evidence of variation in toxin production has been observed with shifts in light (Kaebnick et al. 2000; Kardinaal et al. 2007), nutrients (Sivonen and Jones 1999), and salinity (Tonk et al. 2007). The variant of microcystin toxin has also been shown to vary with environmental conditions (Tonk et al. 2005). Most often cell counts and toxin concentration dynamics are correlated with maximum growth corresponding with maximum toxin production, however the exceptions may relate to changes in growth controlling factors (Downing et al. 2005; Tonk et al. 2009), or an enhanced competitive advantage when light is reduced with increased density of blooms (Kardinaal et al. 2007). In the Copco and Iron Gate Reservoirs nitrogen additions increased the total microcystin production (Moisander et al. 2009) but to a greater degree when biomass was lowest. Runoff (dissolved

inorganic and organic N), NH₄⁺ mixing in from anoxic bottom waters, N₂ fixation, and atmospheric deposition all contribute nitrogen to the Copco and Iron Gate Reservoirs.

Monitoring cyanobacteria in the water can be challenged by the mobility of the blooms with respect to sampling locations, infrastructure for field site access, and the cost and time of field sampling. In some settings, the size of the water body and the logistics of physically collecting samples can be quite prohibitive. Remote sensing offers enormous advantages to address these challenges and is particularly suited to assessments of large geographical areas. While the use of remote detection should be combined with some field sampling, in almost all cases, it serves as a highly effective tool to inform the need for field monitoring and can highlight the most important locations for sample collection. Information that highlights potential algal blooms, even when it is not conclusive that these are specifically toxic cyanobacterial blooms, may still prove enormously useful as the need for field confirmation can be narrowed in time and space.

1.3 Aquatic Remote Sensing

Remotely sensed data from satellites, radar, and airborne sensors have been effectively used to classify terrestrial vegetation and land use (Ustin 2004; Jensen 2007). Although they have been applied with less frequency to aquatic environments, there is still a strong body of literature in which reflectance in the electromagnetic spectrum has been used to detect invasive aquatic vegetation, phytoplankton (open water algae) and other water quality parameters (chlorophyll-a, phaeopigments, aquatic humus, suspended sediment, and temperature (Dekker et al. 1992; Mertes et al. 1993; Jakubauskas et al. 2000; Kallio et al. 2001; Khan and Islam 2003; Li et al. 2003; Sawaya et al. 2003; Williams et al. 2003; Nicandrou et al. 2004; Underwood et al. 2006; Vrieling 2006)). Algae contain a variety of pigments, which give them their own unique spectral signature, in some cases allowing for differentiation between species based on their reflectance and absorption of different wavelengths of electromagnetic energy (Roelfsema et al. 2006). Although some studies have used remotely sensed data (Dekker et al. 1992; Kahru et al. 2000; Vincent et al. 2004; Wynne et al. 2005; Tyler et al. 2006) to detect harmful algal blooms (including some caused specifically by cyanobacteria) in freshwater, most of this research has focused on the ocean (Smith and Baker 1982; Gordon et al. 1983; Carder et al. 1993; Jupp et al. 1994; Stumpf 2001; Tomlinson et al. 2004) for several reasons. First, deep ocean waters are not as highly impacted with turbidity and suspended sediment that can cause near-infrared surface reflection and subsurface volumetric scattering. Therefore there is less interference with the reflectance and absorption from phytoplankton. Secondly, the scale of ocean blooms and red tides is quite large, making their detection possible using sensors with coarse spatial scales (500-1000 m). Satellites such as MODIS (36 bands), SEAWIFS (8 main bands²), and CZCS, which have enough spectral resolution to facilitate distinguishing between the subtle differences in the spectral signatures of planktonic pigments, also have low spatial resolution making it possible to use them for a bloom in the ocean but not in a smaller body of freshwater.

Oki et al. (Oki et al. 1995) developed a model to estimate the concentration of chlorophyll in lakes or inland seas using two wavelengths (675 nm and 700 nm) on a handheld spectroradiometer. Their model attempted to remove the effect of specular reflection at the water surface to improve the accuracy of the chlorophyll estimates. Their algorithms had the

² SEAWIFS has additional bands at 412nm to identify yellow substances through their blue wavelength absorption, at 490 nm to increase sensitivity to chlorophyll concentration, and in the 765-865 nm near-infrared to assist in removing atmospheric attenuation.

highest performance when chlorophyll concentrations were less than 60 mg/L. Roelfsema et al. (Roelfsema et al. 2006) used Landsat 7 TM+ to differentiate the cyanobacteria *L. majuscula* from other phytoplankton and map the extent of the bloom in the clear shallow waters of Moreton Bay, Australia. The method was limited in its average overall accuracy (58% for detecting *L. majuscula*) and it also used only the visible bands (red, green and blue) based on “water penetrating characteristics and preliminary field spectrometer analysis results.” Svab E. et al. (Svab et al. 2005) measured the role of chlorophyll-a and suspended sediment concentrations in the spectral reflectance of Lake Balaton (Europe’s largest shallow lake) in Landsat TM and Landsat ETM+ images. They conclude that while heterogeneous suspended sediment concentrations (SSC) impede direct estimates of the chlorophyll concentration (Chl-a), a spectral linear mixture modeling approach combined with multivariate regression analysis may be able to separate the effect of the two water quality parameters (Chl-a and SSC). This was later done (Tyler et al. 2006) to predict Chl-a in the same lake with an R^2 of 0.95. Vincent et al. (Vincent et al. 2004) was able to predict phycocyanin (a pigment specific to cyanobacteria) concentrations in Lake Erie using a spectral ratio model with Landsat 7 ETM + data. His model had an adjusted R^2 of 73.8% and a standard error (S) of 0.64 ug/l (about 16% of the total phycocyanin range). However, Kutser et al. (Kutser et al. 2006) were unable to distinguish between cyanobacteria and other algae, using data from ALI, Landsat, and MODIS. They suggest that Vincent et al. were actually measuring turbidity not phycocyanin, as Kutser could not detect a unique spectral response for this pigment separate from the suspended sediment response.

While numerous studies have been able to generate strong correlations between satellite spectral bands and chlorophyll-a, to date, none of this work has been done in small freshwater lakes with high spatial resolution IKONOS satellite data nor with high spectral and spatial resolution data as is available through the SpecTIR airborne sensors. Furthermore, very few studies have been able to specifically distinguish cyanobacteria from other classes of phytoplankton.

1.4 Study Objectives and Rational

This research project evaluated the efficacy of remote sensors to assist in characterizing the presence, distribution, and concentration of toxic algae in freshwater systems for surveillance and early detection. A watershed impacted by harmful algal blooms can be managed at a system-wide level looking at the upstream sources of pollution and impairment or locally by directly destabilizing the blooms. Regardless of the ecological management choices, public health officials must respond to the immediate risk by informing the public and helping them to prevent exposure to the algal toxins. Therefore public health professionals are dependant upon early and reliable information on the presence, distribution and concentration of cyanobacterial blooms. This information can also inform appropriate monitoring, research, and resource management for longer-term change. The goal was not to develop new remote sensing classification methods, but to explore the application of existing methods as an accurate, straight forward, accessible tool to semi-automate classification for harmful algal blooms such that it would aid public health agencies and natural resource management with surveillance and early warning detection.

The first specific aim was to select the best classification for the high spatial and spectral resolution remotely sensed data to predict the presence of cyanobacterial blooms. Both

chlorophyll-a and phycocyanin were measured for the purposes of predicting not only relative bloom intensities, but also quantitative concentrations. Chlorophyll-a is a good proxy for total phytoplankton biomass, which is also representative of toxic cyanobacteria biomass during bloom conditions. Phycocyanin is another pigment that can help in distinguishing classes of phytoplankton, because it is specific to cyanobacteria. As is explained later (see Chapter 3), the spatial mobility and temporal variation of the bloom made the pigment concentrations inappropriate for use in the models, but a classification for the relative intensity of the blooms was still generated.

Others (Ustin 2004; Roelfsema et al. 2006; Jensen 2007) have demonstrated the trade-off between accuracy and coverage. While field data collected on the ground manually has high accuracy, it provides coverage of a restricted area. Furthermore, widespread incorporation of remote sensing data into public health management depends on having good access to affordable imagery. A second aim of this study, is to compare the loss in accuracy between different remote sensor image platforms with varied spectral and spatial resolutions (in addition to varied cost and availability) and to identify the technique with the best cost/benefit ratio for generating harmful algal bloom maps for government agencies to generate notices to the public, control access, and plan for longer-term assessment.

Once validated, the application of remote detection of cyanobacterial blooms would ideally be transferable to other water bodies. One of the biggest challenges to large-scale application of these remote detection techniques is that the remote sensing algorithms are often image specific and may not be applicable to images obtained under different conditions (e.g. climate, topography, water depth, etc.). Conducting preliminary exploration of the application of these methods to another field site and images from other dates will provide insight on their limitations and on the opportunities for further research to improve them.

Three classifications were generated, one for each remote sensing data platform. Producer's Accuracy and User's Accuracy were calculated and used to compare them. It was expected that the highest spatial and spectral resolution image would produce the most accurate classification, but comparing the loss of accuracy provided important information to inform the data choice for others. The loss of accuracy with respect to the decrease in cost may be more appropriate in some settings such as surveillance over a large area such as the Yangtze River's Three Gorges Dam Reservoir.

1.5 Contributions of This Dissertation

With climate change and anthropogenic environmental stressors eutrophication is expected to continue to increase, impacting more water bodies and therefore increasing habitat where BGA have a competitive advantage.

Remote sensing offers advantages for public health protection and environmental management of these harmful algal blooms. It provides a way to map water parameters in a consistent way over wide areas at one moment in time. This can make a remarkable difference as even those agencies with regular bimonthly sampling programs are still only collecting samples from a select set of locations, and are unable to know what the full scope of the bloom's impact is. The synoptic overview provided by remote sensing imagery makes it possible to capture the entire water body at one time point. With multiple images from within a bloom season, it is then

possible to learn how the bloom density is changing over time and weather it is better or worse than in previous years. This type of evaluation is just not possible with standard field sampling techniques which allows only a limited number of sites to be sampled and the time it takes to bet between sampling sites is enough time for bloom densities to shift vertically and horizontally.

The investment in remote sensing imagery may ultimately cost less than a routine sampling protocol. Additionally the images are data rich and provide a source of information on much more than harmful algal blooms. They can also be analyzed for questions relating to industrial wastewater discharge, sewage effluent, land use management and change, floods, and erosion.

1.6 Organization of Dissertation

The dissertation consists of six chapters. The current chapter (Chapter 1) lays out the background to the issues of concern regarding cyanobacteria, and an introduction to the remote sensing research tools used to address their management and broad spatial detection. In Chapter 2, I summarize the water sampling I did in the Iron Gate and Copco Reservoirs on the Klamath River and the Three Gorges Dam Reservoir on the Yangtze River. These data describe the current bloom situation in both systems. They support the need for the current research into the applications of remote sensing data to aid in surveillance and management, and they expand knowledge of the overall water quality in the study sites. Because the situation is quite different on the two rivers, there are different rationale for the application of the remote sensing work in each system. Chapter 3 describes the analysis of the high spatial and spectral resolution SpecTIR airborne imagery. Several techniques were applied to the data in the process of selecting the classification decision trees as the most effective approach for the purposes of this study. Chapter 4 uses the same analytical approach with a classification decision tree to analyze the other satellite imagery (Ikonos and Landsat TM). The accuracy and results of the three different remote sensing imagery platforms are compared in Chapter. 4. Since archived Landsat data can be downloaded for free, images from three different years were available and were used to analyze temporal change in a Chapter 5. Chapter 6 summarizes the dissertation, discusses the implications of the findings, future studies needed, and places the findings into the relevant political and economic context.

- Ang, A. (2007). Fast-Growing Algae Smothers Chinese Lake. San Francisco Chronicle. Beijing.
- Annadotter, H., G. Cronberg, et al. (2005). "Endotoxins from Cyanobacteria and Gram-negative Bacteria as the Cause of an Acute Influenza-like Reaction after Inhalation of Aerosols." Ecohealth **2**(3): 209-221.
- Anonymous (2007). South China city hit by toxic "red tide" of algae. Reuters. Beijing.
- Backer, L. C. (2002). "Cyanobacterial Harmful Algal Blooms (CyanoHABs): Developing a Public Health Response." Lake and Reservoir Management **18**(1): 20-31.
- Backer, L. C., W. Carmichael, et al. (2008). "Recreational exposure to low concentrations of microcystins during an algal bloom in a small lake
Characterization of aerosols containing microcystin." Marine Drugs **6**(2): 389-406.
- Backer, L. C., L. E. Fleming, et al. (2003). "Recreational exposure to aerosolized brevetoxins during Florida red tide events." Harmful Algae **2**(1): 19-28.
- Backer, L. C., S. V. McNeel, et al. (2010). "Recreational exposure to microcystins during algal blooms in two California lakes." Toxicon **55**(5): 909-921.

- Carder, K. L., R. G. Steward, et al. (1993). "Avisis Calibration and Application in Coastal Oceanic Environments - Tracers of Soluble and Particulate Constituents of the Tampa Bay Coastal Plume." Photogrammetric Engineering and Remote Sensing **59**(3): 339-344.
- Carmichael, W. W., S. M. Azevedo, et al. (2001). "Human fatalities from cyanobacteria: chemical and biological evidence for cyanotoxins." Environ Health Perspect **109**(7): 663-8.
- Carmichael, W. W., C. Drapeau, et al. (2000). "Harvesting of *Aphanizomenon flos-aquae* Ralfs ex Born. & Flah. var. *flos-aquae* (Cyanobacteria) from Klamath Lake for human dietary use." Journal of Applied Phycology **12**(6): 585-595.
- Centers for Disease Control and Prevention. "Environmental Hazards and Health Effects of Harmful Algal Blooms." Retrieved July 18, 2007, from <http://www.cdc.gov/hab/cyanobacteria/default.htm#cdc>.
- Chorus, I. and J. Bartram, Eds. (1999). Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management. London, E & FN Spon.
- Dekker, A. G., T. J. Malthus, et al. (1992). "Remote sensing as a tool for assessing water quality in Loosdrecht lakes." Hydrobiologia **233**(1 - 3): 137-159.
- Downing, T. G., C. Meyer, et al. (2005). Microcystin content of *Microcystis aeruginosa* is modulated by nitrogen uptake rate relative to specific growth rate or carbon fixation rate, John Wiley & Sons Inc.
- Environmental Protection Agency (2005). Drinking Water Contaminant Candidate List 2; Final Notice, Fed Reg: 70 (36): 9071-9077.
- Falconer, I., J. Bartram, et al. (1999). Safe Levels and Safe Practices. Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management. I. Chorus and J. Bartram. London, E & FN Spon: 155-177.
- Falconer, I. R., A. M. Beresford, et al. (1983). "Evidence of liver damage by toxin from a bloom of the blue-green alga, *Microcystis aeruginosa*." Med J Aust **1**(11): 511-4.
- Gordon, H. R., D. K. Clark, et al. (1983). "Phytoplankton Pigment Concentrations in the Middle Atlantic Bight - Comparison of Ship Determinations and Czcs Estimates." Applied Optics **22**(1): 20-36.
- Griffiths, D. J. and M. L. Saker (2003). "The Palm Island mystery disease 20 years on: A review of research on the cyanotoxin cylindrospermopsin." Environmental Toxicology **18**(2): 78-93.
- Jakubauskas, M., K. Kindscher, et al. (2000). "Close-range remote sensing of aquatic macrophyte vegetation cover." International Journal of Remote Sensing **21**(18): 3533-3538.
- Jensen, J. R. (2007). Remote Sensing of the Environment: An Earth Resource Perspective. Upper Saddle River, NJ, Pearson Education, Inc.
- Jochimsen, E. M., W. W. Carmichael, et al. (1998). "Liver Failure and Death after Exposure to Microcystins at a Hemodialysis Center in Brazil." N Engl J Med **338**(13): 873-878.
- Jupp, D. L. B., J. T. O. Kirk, et al. (1994). "Detection, Identification and Mapping of Cyanobacteria - Using Remote-Sensing to Measure the Optical-Quality of Turbid Inland Waters." Australian Journal of Marine and Freshwater Research **45**(5): 801-828.
- Kaebnick, M., B. A. Neilan, et al. (2000). "Light and the transcriptional response of the microcystin biosynthesis gene cluster." Applied and Environmental Microbiology **66**(8): 3387-3392.

- Kahru, M., J. M. Leppanen, et al. (2000). "Cyanobacteria blooms in the Gulf of Finland triggered by saltwater inflow into the Baltic S." Marine Ecology Progress Series **207**: 13-18.
- Kallio, K., T. Kutser, et al. (2001). "Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons." The Science of The Total Environment **268**(1-3): 59-77.
- Kann, J. and E. Asarian (2007). Nutrient Budgets and Phytoplankton Trends in Iron Gate and Copco Reservoirs, California, May 2005-May 2006, Aquatic Ecosystem Sciences & Kier Associates for the Karuk Tribe of California and State Water Resources Control Board: 81.
- Kardinaal, W. E. A., L. Tonk, et al. (2007). "Competition for light between toxic and nontoxic strains of the harmful cyanobacterium *Microcystis*." Applied and Environmental Microbiology **73**(9): 2939-2946.
- Khan, N. I. and A. Islam (2003). "Quantification of erosion patterns in the Brahmaputra-Jamuna River using geographical information system and remote sensing techniques." Hydrological Processes **17**: 7.
- Kun, Z. (2007). Experts Identify Algae in Taihu Lake. China Daily.
- Kutser, T., L. Metsamaa, et al. (2006). "Monitoring cyanobacterial blooms by satellite remote sensing." Estuarine, Coastal and Shelf Science **67**(1-2): 303-312.
- Lehman, P. W., G. Boyer, et al. (2005). "Distribution and toxicity of a new colonial *Microcystis aeruginosa* bloom in the San Francisco Bay Estuary, California." Hydrobiologia **541**: 87-99.
- Li, R. R., Y. J. Kaufman, et al. (2003). "Remote sensing of suspended sediments and shallow coastal waters." Ieee Transactions on Geoscience and Remote Sensing **41**(3): 559-566.
- Mertes, L. A. K., M. O. Smith, et al. (1993). "ESTIMATING SUSPENDED SEDIMENT CONCENTRATIONS IN SURFACE WATERS OF THE AMAZON RIVER WETLANDS FROM LANDSAT IMAGES." Remote Sensing of Environment **43**(3): 281-301.
- Moisander, P. H., M. Ochiai, et al. (2009). "Nutrient limitation of *Microcystis aeruginosa* in northern California Klamath River reservoirs." Harmful Algae **8**(6): 889-897.
- Nicandrou, A., L. A. Mofor, et al. (2004). Hydrological assessment and modeling of the River Fani catchment in Albania using GIS and remote sensing. Remote Sensing for Agriculture, Ecosystems, and Hydrology V, Barcelona, Spain, SPIE.
- Nishiwaki-Matsushima, R., T. Ohta, et al. (1992). "Liver tumor promotion by the cyanobacterial cyclic peptide toxin microcystin-LR." Journal of Cancer Research and Clinical Oncology (Historical Archive) **118**(6): 420-424.
- Oki, K., Y. Yasuoka, et al. (1995). "Estimation of chlorophyll concentration in lakes and inland seas from near-infrared and red spectral signature." Retrieved July 18, 2007, from <http://www.gisdevelopment.net/aars/acrs/1995/ts4/ts4004pf.htm>.
- Pouria, S., A. de Andrade, et al. (1998). "Fatal microcystin intoxication in haemodialysis unit in Caruaru, Brazil." The Lancet **352**(9121): 21-26.
- Roelfsema, C. M., S. R. Phinn, et al. (2006). "Monitoring toxic cyanobacteria *Lyngbya majuscula* (Gomont) in Moreton Bay, Australia by integrating satellite image data and field mapping." Harmful Algae **5**(1): 45-56.
- Sawaya, K. E., L. G. Olmanson, et al. (2003). "Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery." Remote Sensing of Environment **88**(1-2): 144-156.

- Sivonen, K. and G. J. Jones (1999). Cyanobacterial Toxins. Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management. I. Chorus and J. Bartram. London, E & FN Spon: 41-111.
- Smith, G. M. (1956). The Fresh-Water Algae of The United States. New York, McGraw-Hill Book Company.
- Smith, R. C. and K. S. Baker (1982). "Oceanic chlorophyll concentrations as determined by satellite (Nimbus-7 Coastal Zone Color Scanner)." Marine Biology **66**(3): 269-279.
- Stumpf, R. P. (2001). "Applications of satellite ocean color sensors for monitoring and predicting harmful algal blooms." Human and Ecological Risk Assessment **7**(5): 1363-U15.
- Svab, E., A. N. Tyler, et al. (2005). "Characterizing the spectral reflectance of algae in lake waters with high suspended sediment concentrations." International Journal of Remote Sensing **26**(5): 919-928.
- Teixeira Mda, G., C. Costa Mda, et al. (1993). "Gastroenteritis epidemic in the area of the Itaparica Dam, Bahia, Brazil." Bull Pan Am Health Organ **27**(3): 244-53.
- Tomlinson, M. C., R. P. Stumpf, et al. (2004). "Evaluation of the use of SeaWiFS imagery for detecting *Karenia brevis* harmful algal blooms in the eastern Gulf of Mexico." Remote Sensing of Environment **91**(3-4): 293-303.
- Tonk, L., K. Bosch, et al. (2007). "Salt tolerance of the harmful cyanobacterium *Microcystis aeruginosa*." Aquatic Microbial Ecology **46**(2): 117-123.
- Tonk, L., P. M. Visser, et al. (2005). "The microcystin composition of the cyanobacterium *Planktothrix agardhii* changes toward a more toxic variant with increasing light intensity." Applied and Environmental Microbiology **71**(9): 5177-5181.
- Tonk, L., M. Welker, et al. (2009). "Production of cyanopeptolins, anabaenopeptins, and microcystins by the harmful cyanobacteria *Anabaena* 90 and *Microcystis* PCC 7806." Harmful Algae **8**(2): 219-224.
- Tyler, A. N., E. Svab, et al. (2006). "Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment." International Journal of Remote Sensing **27**(8): 1521-1537.
- Underwood, E. C., M. J. Mulitsch, et al. (2006). "Mapping invasive aquatic vegetation in the Sacramento-San Joaquin Delta using hyperspectral imagery." Environmental Monitoring and Assessment **121**(1-3): 47-64.
- USEPA. (2005). "Ecology and Oceanography of Harmful Algal Blooms, Funding Opportunity Number: EPA-G2006-STAR-B1." Retrieved 11/1/05, 2005, from http://es.epa.gov/ncer/rfa/2005/2005_ecohab.html.
- USEPA. (2007). "Ecology and Oceanography of Harmful Algal Blooms, Funding Opportunity Number: EPA-G2006-STAR-B1." Retrieved 07/19/07, 2007, from http://es.epa.gov/ncer/rfa/2008/2008_star_ecohab.html.
- Ustin, S. L. (2004). Remote Sensing for Natural Resource Management and Environmental Monitoring. Hoboken, NJ, John Wiley & Sons, Inc.
- Vincent, R. K., X. Qin, et al. (2004). "Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie." Remote Sensing of Environment **89**(3): 381-392.
- Vrieling, A. (2006). "Satellite remote sensing for water erosion assessment: A review." CATENA **65**(1): 2-18.

- Williams, D. J., N. B. Rybicki, et al. (2003). "Preliminary investigation of submerged aquatic vegetation mapping using hyperspectral remote sensing." Environmental Monitoring and Assessment **81**(1-3): 383-392.
- World Health Organization (WHO) (2003). Algae and cyanobacteria in fresh water. Guidelines for safe recreational water environments. Geneva, World Health Organization. **1**.
- Wynne, T. T., R. P. Stumpf, et al. (2005). "Detecting *Karenia brevis* blooms and algal resuspension in the western Gulf of Mexico with satellite ocean color imagery." Harmful Algae **4**(6): 992-1003.
- Xiaofeng, G. (2007). Two major lakes again hit by algae. China Daily.
- Xie Liqiang, P. X., Longgen Guo, Li Li, Yuichi Miyabara, Ho-Dong Park, (2005). "Organ distribution and bioaccumulation of microcystins in freshwater fish at different trophic levels from the eutrophic Lake Chaohu, China." Environmental Toxicology **20**(3): 293-300.
- Yu, S., N. Zhao, et al. (2001). "[The relationship between cyanotoxin (microcystin, MC) in pond-ditch water and primary liver cancer in China]." Zhonghua Zhong Liu Za Zhi **23**(2): 96-9.
- Zhou, L., H. Yu, et al. (2002). "Relationship between microcystin in drinking water and colorectal cancer." Biomed. Environ. Sci. **15**(2): 166-171.

2 Cyanobacteria and Water Quality Characterization

2.1 Introduction

This research involved field data collection in two field sites: the Iron Gate and Copco reservoirs in northern California and the Three Gorges Dam reservoir in China. The former are upstream of a dam that was built more than 40 years ago on the Klamath River and the latter is upstream of the newly built Three Gorges Dam. The Iron Gate and Copco Reservoirs have dense cyanobacteria blooms each summer, typically lasting from June through September. This provided an opportunity to evaluate the use of remote sensing in the context of a system clearly impacted by toxic cyanobacteria blooms and for which there is a great deal of data, and to also evaluate shifting dynamics in a newly altered system and the potential benefits offered by the use of remote sensing to better study the changing water quality situation there. This chapter describes the overall water quality situation in these two field sites, the field collection of water samples, and the analyses of the field sample data.

2.1.1 Klamath Field Site

The potential ability of satellite images to detect the color and temperature changes associated with cyanobacteria blooms can be used to better predict the need for further monitoring or management, and its large spatial scale can enhance our understanding of ground sampling. Data from the known and well characterized blue-green algae blooms on the Iron Gate and Copco Reservoirs were used to evaluate satellite image classification as an economic and efficient tool to remotely detect the presence of cyanobacterial growth and to quantify the intensity of a bloom event. An airborne pass-by was scheduled with SpecTIR to collect hyperspectral data which was analyzed for detection and quantification of cyanobacterial bloom intensity in the Iron Gate and Copco Reservoirs of the Klamath River in Northern California. Hyperspectral data contains detailed information on the spectral reflectance of features on the ground at small intervals throughout the electromagnetic spectrum. This level of detail provides the most information available through remote sensing to be used to determine whether blooms are present and to what degree. Ground truth data was also collected for comparison and potential calibration of the model.

The Klamath River runs from southern Oregon across Siskiyou and Humboldt Counties in northern California to its mouth at Klamath, California. The watershed draining into the Klamath River is 40,795 km². (~4 million hectares or 10 million acres) (Board on Environmental Studies and Toxicology and Water Science and Technology Board 2008). The Klamath River is one of the major salmon rivers in the western United States, and is one of the largest river systems in terms of flow, salmon production and economic importance in the western U.S. PacifiCorp Energy owns and operates the 169-megawatt Klamath River Hydroelectric Project regulating flows and generating electricity. The Klamath field site focuses on the Iron Gate and Copco Reservoirs, just two of the five impoundments formed by PacifiCorp's five dams which were built between 1908 and 1962. Iron Gate Reservoir is 944 acres (PacifiCorp 2010). Below the Iron Gate Dam the Klamath river is protected as free flowing under the National Wild and Scenic Rivers System (North Coast Regional Water Quality Control Board 2010b).

The Klamath reservoirs are located in a rural area. The closest population center is Yreka, California (7,290 people in 2000) located about 40 miles to the southwest. Figure 2-1 shows an

overview map of the location of the watershed and Figure 2-2 shows a detailed map of Iron Gate and Copco Reservoirs where remote sensing and field samples were collected for water analysis of nutrients, pigments, temperature, turbidity, depth, and secchi disk transparency.

Figure 2-1: Overview Map of the Iron Gate and Copco Reservoirs on the Klamath River

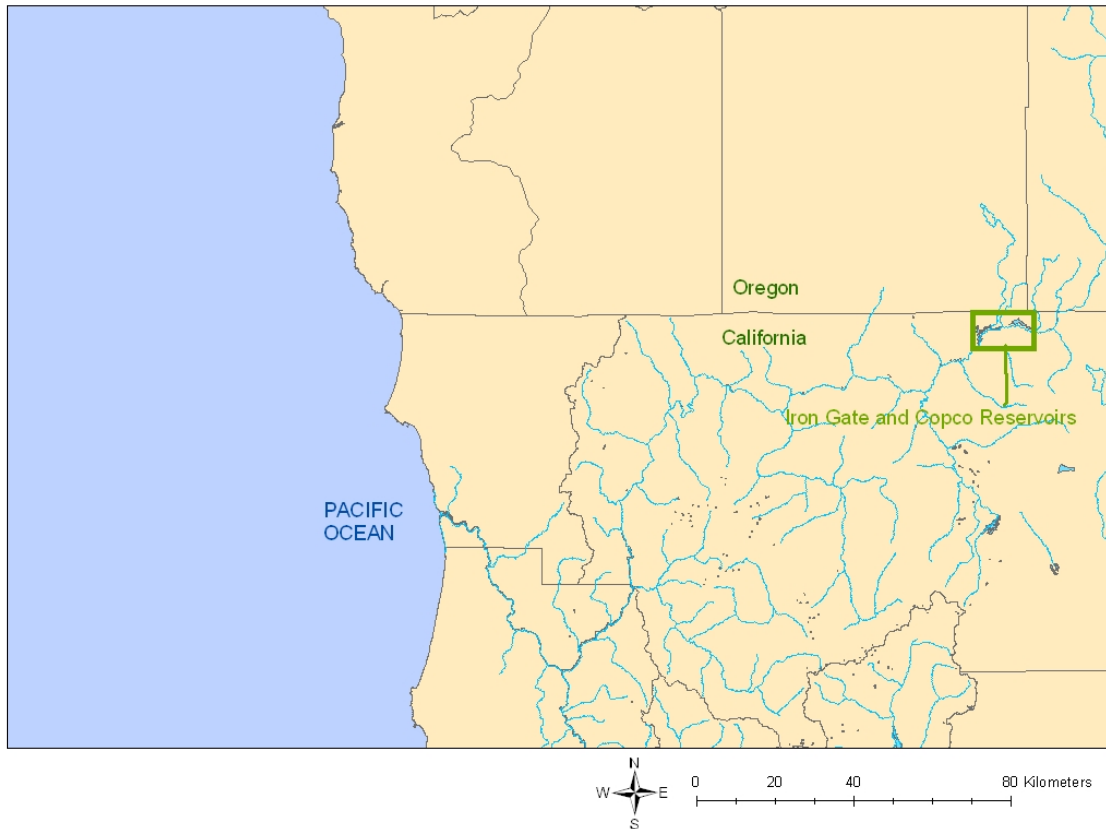


Figure 2-2: Sampling Sites on the Iron Gate and Copco Reservoirs



2.1.2 Yangtze Field Site

The field site in China is the 600-km stretch of the Yangtze River (Chang Jiang or [长江](#)) between Chongqing and the Three Gorges Dam encompassing the newly formed reservoir, whose surface area will eventually reach approximately 1080 km² (Wu et al. 2003). The Yangtze River is the third longest river in the world with its headwaters in the northwest Chinese province of Qinghai and its estuary 6300 km away near the city of Shanghai. It descends more than 5000 m from the Qinghai-Tibet Plateau and drains 1.81x 10⁶ km² (Chen et al. 2001). In Yichang, just downstream of the Three Gorges Dam, the discharge ranges from 2.0x10⁴ - 4.0x10⁴ m³/s during the rainy season (June-August) and 0.4 x 10⁴ m³/s during the dry season (November-February). The construction of the Three Gorges Dam (TGD) on the Yangtze River was completed in 2003, and the flooding to form the reservoir was predicted to continue through 2009.

The construction of the Three Gorges Dam provoked large controversies regarding the trade-offs between facilitating navigation, creating opportunities for south-north water transfer, and flood management versus the dam's potential detrimental impacts on the hydrology of the Yangtze River watershed. Water quality in the reservoir affects the millions of people in cities along its shores whose drinking water is withdrawn from the reservoir. The water body is also used for fishing, washing, recreation and irrigation. Chongqing, at the western and upstream edge of the

reservoir, is the home of scores of industries discharging waste streams directly to the Yangtze which is a vital resource for the 400 million people living in the river basin. Given the water shortages in the North China Plain, its management is pertinent to another 400 million who reside there. China's Yangtze River Reservoir also presents an interesting case to explore the application of remote detection and surveillance since it covers such a large area with poor road access prohibiting adequate field monitoring and making remote detection an even more appropriate alternative. Swimming, fishing, washing clothes and of course drinking water are all examples of the pathways for human exposure to cyanotoxins in China. 1.3 million people in Yunyang county, the location of field sampling sites for this study, receive their drinking water from the reservoir.

Historically the strong stream flow in the Yangtze River has mediated the impacts of extensive water pollution (Maurer et al. 1997). The completion of the Three Gorges Dam in 2003, however, has reduced the river's water velocity. The resulting elevation in water temperature and the concentration of inorganic pollutants from upstream industries as well as local and upstream inputs of organic material and nutrients in the growing reservoir are important variables in the promotion of cyanobacterial growth. Concomitant land use shifts may also result in new agricultural development, which could further raise nutrient levels in the water due to fertilizer input.

2.2 Water Quality

2.2.1 Water Quality in the Klamath Field Site

The water quality of the Klamath River and reservoirs has been well documented through tribal (Kann and Asarian 2005; Kann 2006a; Kann 2006b; Kann 2006c; Kann and Asarian 2006; Kann and Corum 2006; Kann and Asarian 2007; Kann and Corum 2007), state and federal government (North Coast Regional Water Quality Control Board 2010a) monitoring as well as that conducted by PacifiCorp (PacifiCorp 2004; Carlson and Raymond 2008), the company that owns and operates the hydroelectric facilities. Parameters associated with toxic cyanobacteria bloom dynamics that have been monitored include wind, temperature, flow, depth, mixing, turbidity, nutrient concentration, light and residence time.

In 2005 Iron Gate and Copco Reservoirs together retained 11.9% of the total phosphorus inflow, and 3.7% of the total phosphorus was retained during the photoplankton growing season (May-October). This time of year also corresponds with the dry season and the hottest months of the year. For nitrogen inflow the corresponding numbers were 18.1% and 29.8% (Kann and Asarian 2007). The Iron Gate/Copco Reservoir system operates under overall nitrogen limitation for phytoplankton growth during the summer season (Moisander et al. 2009). Both reservoirs are typically thermally stratified during the warm summer months, with low levels of dissolved oxygen and high levels of ammonia and soluble reactive phosphorus in the hypolimnion which correspond with large blooms of phytoplankton in the epilimnion.

These two reservoirs on the Klamath River in Northern California have had documented cyanobacteria blooms between May and October in the last five years (Kann 2006a; Kann 2006b; Kann 2006c; Kann and Corum 2007; Kann and Corum 2009). Blooms have a long duration typically beginning in June, peaking in late August and disappearing mid-October. The blooms are dominated by *microcystis aeruginosa* (MSAE) with cell counts as high as

393,395,000 cells/ml and microcystin concentrations as high 12,176 µg/L (Kann 2006a). Table 2-1 shows the maximum, mean, and median values for all sampling dates and sites within the reservoirs for the last four summers. Each year the sites that are located within the reservoirs have had levels of MSEA and its toxin, microcystin, that far exceed the World Health Organizations recommended guidelines for public health protection (Table 2-2 shows the 3-tiered guidelines from WHO).

TABLE 2-1: Statistics for Cyanobacteria Blooms over last Four Years

	2006	2007	2008	2009
Microcystis aeruginosa (cells/ml)				
Maximum	393,395,000	22,898,635	42,168,500	122,338,260
Mean	12,381,258	1,475,681	3,569,067	9,976,831
Median	108,620	46,979	87,403	928,925
Microcystin Concentration (ug/L)				
Maximum	12,176	30,000	18,000	73,000
Mean	1,449	1,194	1,195	7,221
Median	386	3	10	130

TABLE 2-2: World Health Organization’s Recommended Guidelines for Recreational Water

Guidance level	Health Risks	Typical actions
Low Risk - 20,000 cyanobacterial cells/ml	Short-term adverse health outcomes, e.g., skin irritations, gastrointestinal illness	Post on-site risk advisory signs Inform relevant authorities
Moderate Risk - 100,000 cyanobacterial cells/ml <i>or</i> 50 µg chlorophyll-a/liter with dominance of cyanobacteria	Potential for long-term illness with some cyanobacterial species Short-term adverse health outcomes, e.g., skin irritations, gastrointestinal illness	Watch for scums or conditions conducive to scums Discourage swimming and further investigate hazard Post on-site risk advisory signs Inform relevant authorities
High Risk - Cyanobacterial scum formation in areas where whole-body contact and/or risk of ingestion/aspiration occur	Potential for acute poisoning Potential for long-term illness with some cyanobacterial species Short-term adverse health outcomes, e.g., skin irritations, gastrointestinal illness	Immediate action to control contact with scums; possible prohibition of swimming and other water contact activities Public health follow-up investigation Inform public and relevant authorities

2.2.2 *Water Quality in the Yangtze Field Site*

After the construction of the TGD in 2003 algal blooms were documented in one of the Yangtze River's tributaries, the Xiangxi River which is located 37 km upstream of the TGD, as well as in the confluence of the Xiangxi and the Yangtze (Zeng et al. 2006). In August of 2004, 63 phytoplankton taxa were identified in the Yangtze River. Algal blooms had not been documented in these locations previously. These investigators also monitored phytoplankton populations in the Yangtze River intermittently during a 13 month period (Zeng et al. 2007). They collected data during March, May, July, August and October of 2004 and May of 2005. They found that diatoms dominated the phytoplankton populations and that phytoplankton abundance was determined more by hydrological conditions rather than by nutrient concentrations. Although they intended their study to provide long-term monitoring of fluctuations in TGR water quality, they were only able to capture one snapshot of the spring season (March 2004), three sampling events over one summer (May, July, and August 2004), and one sampling event in the fall (Oct. 2004) followed by one final sampling event the following year in May 2005. Post-impoundment, the Yangtze has continued to undergo change as flooding has moved forward since 2003 in the formation of the vast reservoir stretching 650 kilometers upstream of the dam to Chongqing. During the 6-year process of flooding, hydrological conditions cannot equilibrate, and multiple years of data are necessary to attempt to capture patterns and possibly, reference levels to which to compare data from future time points collected after the completion of the reservoir (estimated for 2009). Therefore the research reported here aimed to continue monitoring where Zeng and colleagues left off (summer 2005) in order to expand the time-series, compare findings, and therefore ultimately enhance the usefulness of the pooled dataset for future planning and understanding of the nearly complete reservoir's biogeochemistry and phytoplankton population diversity and density. Our data added with that of previous studies provides estimated baseline data on water quality and phytoplankton populations in the TGDR.

2.3 Water Quality Field Data Collection

2.3.1 *Water Data Collection (Klamath Field Site)*

Field data collected bi-monthly from the Iron Gate and Copco Reservoirs on phytoplankton counts and species identification, toxin concentrations, and chlorophyll-concentrations have been shared by the Yurok Tribe, the Karuk Tribe and the Klamath Blue-Green Algae Work Group for use in this study. High resolution satellite images are more likely to be useable for this time and region given that the days are typically clear, and cloud-free. The presence of known cyanobacterial blooms, the access to good field data and the potential to obtain clear high-resolution satellite images made this an appropriate case and field site for the development and validation of the remote detection methods.

The water sampling data from known blue-green algae blooms on the Iron Gate and Copco Reservoirs on the Klamath River was intended for the use in evaluation and classification of three different images. Data collection in the Klamath, both remotely by SpecTIR and Ikonos and directly in the field occurred in July 2007. Most water samples were collected within 12-24 hours of the remote sensor data acquisitions. Landsat's orbit brings it over the area of interest (AOI) every 16 days, and the acquisition of IKONOS and SpecTIR airborne imagery were scheduled to occur as close to the day of the Landsat passby that fell during the peak of the toxic BGA bloom in the reservoirs. Landsat's orbit was over the AOI on July 28, 2007, but

unfortunately it could not be purchased as it was never archived and made available by the USGS Center for Earth Resources Observation and Science. 2005 Landsat data was used instead in the image analyses (see Chapter 4). The blooms peak in late July and early August, a time when clear skies predominate, facilitating remote detection of the dense green surface scum. IKONOS satellite acquisition occurred on July 27, 2007 and SpecTIR airborne imagery was collected on July 28, 2007.

Because of the large number of pixels in remotely sensed data, the sample size required to be statistically sound for accuracy assessment based on ground data in a traditional sense can become impractical. Congalton (Congalton 1991) recommends collecting at least 50 samples for each land-cover class. The number can be increased for land classes that are particularly important for the objective of the study (in this case the water in general and the water surface covered by blooms), or similarly decreased for the classes of lesser interest. Greater variation within a particular cover class can also warrant collection of an increased number of reference samples. Because the study area in this project is relatively small ($\sim 7.86 \text{ km}^2$), the goal was to collect at least 50 water samples randomly from the two reservoirs. Up to 20 additional samples were desirable to specifically capture variation within a bloom. In actuality 62 water samples were collected in total. The size of the remote sensing pixels to be analyzed will be 4 m^2 , 1 m^2 , and 900 m^2 for the SpecTIR, IKONOS and Landsat images respectively. 10 samples were collected that were spaced out approximately every 3 meters to capture separate SpecTIR and IKONOS pixels encompassed by one single Landsat pixel. See Chapter 3 for remote sensing data analysis methods.

2.3.2 GPS Data Collection (*Klamath Field Site*)

The geographic coordinates for each sampling site were recorded with a GPS (Global Positioning System), in order to link sampling data with the relevant pixels in the remote sensing data. To help verify the accuracy, coordinates were collected for the corners of a fixed dock and boat ramp located within the study area prior to remote image acquisition.

The open prairie and grassland terrain of the rolling hills surrounding the Copco and Iron Gate Reservoirs offer little if any possible interference for the satellite signal reception, unlike mountainous terrain or densely foliated sites. The GPS unit uses four or more satellites to triangulate the position.

50 random points were generated within the boundaries of the Iron Gate and Copco reservoirs using the Hawth Tools functionality within ArcGIS. The coordinates and map of these locations were used as a rough guide in the field to collect the water samples, however, efficiency required that we abandon the effort to steer our boat to the exact GPS location generated randomly. Rather we moved from east to west over the reservoir collecting samples at approximately the same locations as the randomly generated coordinates with emphasis more on capturing both sides of the reservoir, open water, and coves. Additional deviations were made from the randomly selected points when visual observations indicated stratified concentric layers of different algal densities. Samples were intentionally collected to provide cross-sections of these stratified areas to better capture the variability within the bloom.

A total of 62 sampling sites were collected across the Iron Gate and Copco reservoirs. The location of sample withdraws was recorded using both a differential Trimble GPS (GeoXH Handheld) receiver with the capability of sub-meter accuracy as well as a Garmin GPS receiver (GPSmap 60CSx) with the capability of 2-10 meter accuracy depending on the environment and the position of satellites. Accuracy typically ranged between 2-5 meters for the Garmin GPS receiver because the unit was WAAS (Wide Area Augmentation System) enabled and WAAS reception was available. The Trimble GPS receiver was differentially corrected using Differential Global Positioning Systems to obtain a higher level of accuracy that corrects for atmospheric conditions and irregularities in satellite orbits by referencing a nearby “base station”.

2.3.3 Water Data Collection (Yangtze Field Site)

The Yangtze River field data collection was initiated in July of 2005. The reservoir created by the Three Gorges Dam stretches 700 kilometers upstream of the dam. In 2005, the average width of the reservoir was 300 meters and has gradually widened through 2009 when flooding was estimated to be complete. The expansive reservoir area of approximately 350 square kilometers is difficult to monitor. Adequate characterization of the entire reservoir would require an elaborate sampling plan with well distributed sampling sites throughout the reservoir because wind patterns can create horizontal stratification; because uses, nutrient sources and loads, and other water impacts can vary substantially across such an area; and because other basic water quality parameters are variable in such a large water body. There are several factors that impede such sampling. The first is the challenge of access throughout the reservoir. Much of the surrounding land is rural and difficult to access by land. Boat transport is slow and not organized for such needs. Laboratory facilities are distant and transport of samples in a timely fashion for analysis would be impossible without the investment of significant resources. This project made a compromise in order to provide valuable information to characterize the water quality in the reservoir. Rather than attempt to collect data throughout the reservoir, a preliminary comparative cross-sectional study was conducted to measure the concentration of nitrate, organophosphate, and the other parameters listed in Table 2-3 at seven locations (see map in Figure 2-3) stretched across the full length of the reservoir.

These data from the Three Gorges Dam reservoir detected no algal blooms, however one genus of cyanobacteria (*oscillatoria*) is already present and a dense green algae bloom occurred on a tributary (Xiao Jiang) which joins the Yangtze at the city of Yunyang. It is quite possible that other blooms are being missed due to insufficient sampling locations and/or frequency.

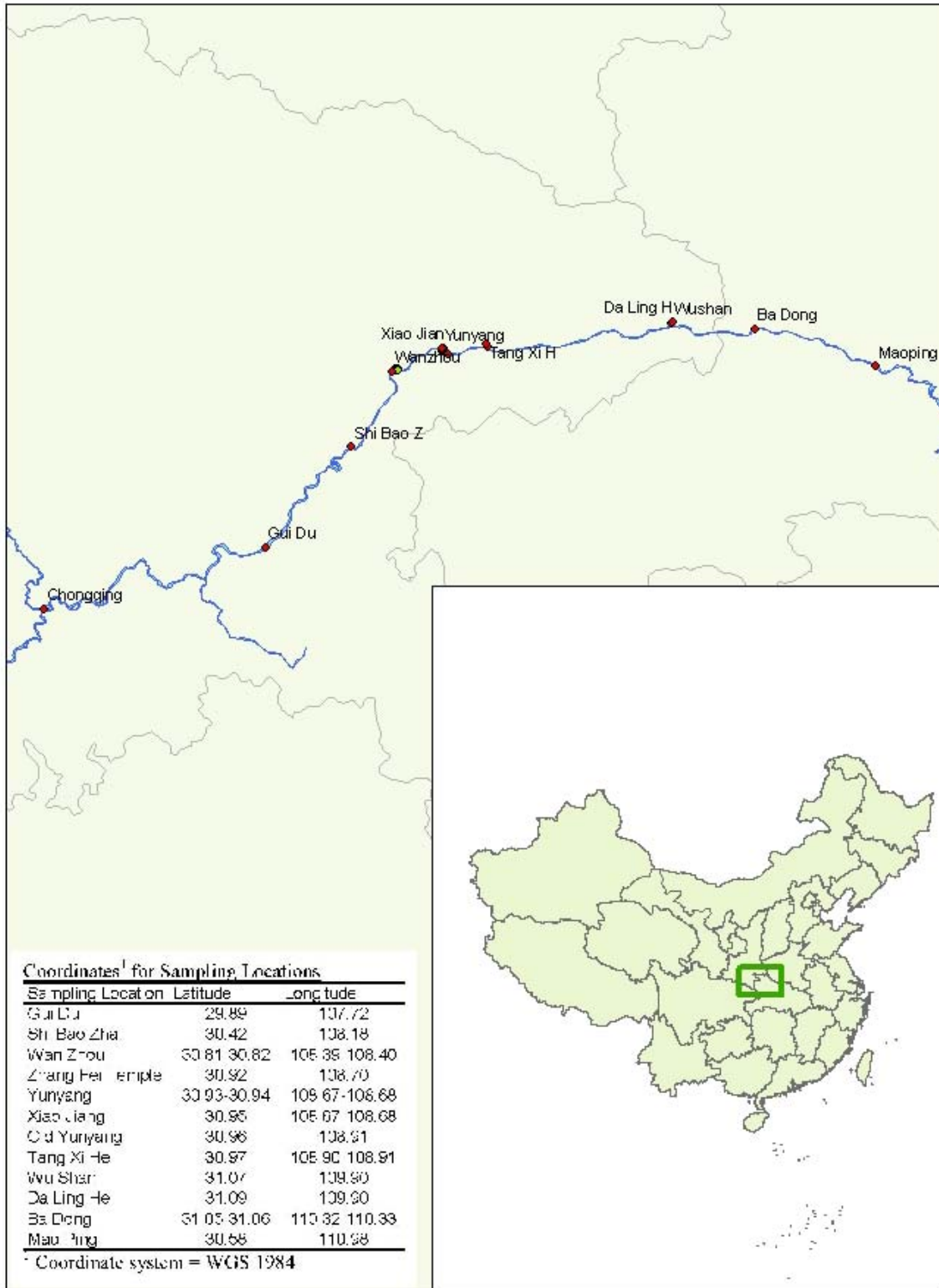
Remote detection would facilitate making a more definitive determination of the presence or absence of harmful algal blooms over the full spatial scope of the reservoir, it would enhance our understanding of existing ground sampling data, and it would inform the need for additional water sampling and the general location where this would be appropriate.

2.3.4 Water Sampling (Klamath Site)

At each sampling site, the water's temperature, depth, conductivity and transparency (via secchi disk) were measured, and 2-3 100-ml surface water samples were collected with a Van Dorn sampler, split with a churn splitter and immediately put on ice in the dark. Several semi-

integrated samples were also collected by combining water from the surface, 1-meter, 5-meters and 10-meters below the surface in the churn splitter. Within twelve hours of collection one

Figure 2-3: Sample Locations in Three Gorges Dam Reservoir for Preliminary Comparative Study



sample from each site was filtered using a glass-fiber filter and frozen to be analyzed later for chlorophyll-*a* and phycocyanin. Phytoplankton identification and enumeration was done for the Yangtze samples but not on the Klamath samples, as this information was obtained for separate samples collected by the Klamath Blue-Green Algae Work Group at 11 sites within these reservoirs (twice a month from May through October) and made available for this research.

Microcystis aeruginosa (a known toxic blue-green algae species) has been identified as the dominant species in all existing bloom datasets for Copco and Iron Gate Reservoirs although *Aphanizomenon flos-aquae*, *Anabaena sp.*, and *Gloeotrichia echinulata*, other species of cyanobacteria were also detected in lower concentrations. Chlorophyll-*a* is often a good proxy of harmful algal biomass in this system, because the phytoplankton population is so uniform. Data from the Klamath system does show an increasing trend in the probability that *microcystis aeruginosa* cell counts and microcystin concentrations will exceed the various recommended WHO guidance levels as chlorophyll-*a* concentrations increase. This is not always the case, however, as many other genera and species of phytoplankton also produce chlorophyll. Therefore the more specific cyanobacterial pigment phycocyanin was also quantified for use in classification of the remote sensing data.

TABLE 2-3: Measured Water Parameters

Dissolved Oxygen
pH
Temperature
Conductivity
Turbidity
Transparency
Depth
Chlorophyll-*a* (Klamath only)
Phycocyanin (Klamath only)
Nitrate (Yangtze only)
Soluble Reactive Phosphorus (Yangtze Only)

All of these parameters were measured in the field using portable probes except nutrients and the pigments, chlorophyll-*a* and phycocyanin, which were measured in the laboratory.

2.3.5 Water Sampling (Yangtze Site)

Visual inspections for discoloration and scum formation, and for transparency with the use of a secchi disc were conducted at each location. Grab samples were collected by hand-dipping 1 liter bottles into the surface of the water. Once every third sampling trip, one site was collected in triplicate to evaluate the variability of the sampling technique.

For phytoplankton enumeration and identification in the Yangtze River, horizontal tows (12-inch diameter opening to a 63 micron mesh plankton net) were used to concentrate phytoplankton in samples. A flow meter attached to the plankton net was used to calculate the exact distance of the tow and later the volume of water filtered by the plankton net. These samples were preserved in the field immediately with Lugol's solution. Phytoplankton genera were identified and samples were counted using a compound light microscope with a Sedgewick-Rafter counting cell

(American Public Health Association (APHA) 2001). 10 randomly selected grid cells were counted to reach a minimum of 100 cells.

Samples to be analyzed for chlorophyll-a were kept on ice and dark for up to 12 hours before they were filtered onto glass fiber filters (Millipore 0.7 micron pore size). Filters were then kept cold ($\leq 4^{\circ}\text{C}$) and dark until analysis. Chlorophyll-a was analyzed spectrophotometrically. The absorbance was measured at 664, 647, and 630 nanometers on a PerkinElmer Lambda 14 UV/Vis Spectrophotometer and applied to equations (1), (2), and (3) of Jeffrey and Humphrey (1975) as recommended in the Standard Methods (American Public Health Association (APHA) 2001).

$$\text{Chlorophyll}(a) = 11.85E_{664} - 1.54E_{647} - 0.08E_{630} \quad (1)$$

$$\text{Chlorophyll}(b) = -5.43E_{664} + 12.03E_{647} - 2.66E_{630} \quad (2)$$

$$\text{Chlorophyll}(c_1 + c_2) = -1.67E_{664} - 7.60E_{647} + 24.52E_{630} \quad (3)$$

Surface grab samples were collected and used for the measurement of nitrate and soluble reactive phosphorous. All samples were kept cold on ice and dark until they were analyzed spectrophotometrically using the HACH cadmium reduction method 8171 and the phosphomolybdate reduction Method 8048 which are equivalent to USEPA Method 365.2 and Standard Method 4500-PE, respectively. Nutrient analyses were conducted no more than 48 hours after sample collection. The analytical detection limits were 0.2 mg/L and 0.05 mg/L respectively. The percent coefficient of variation from replicate analyses of selected samples and the standard was $<15\%$.

Finally temperature, turbidity, conductivity, pH, and dissolved oxygen were measured in the field using a Horiba multi-meter. ELISA (enzyme-linked immunosorbent assay) field kits were purchased to analyze for microcystin, however these analyses were not conducted as large cell counts of cyanobacteria were not detected.

Cross-Sectional Data Sampling Locations

During June 2005 and March 2007 samples were collected from five other locations spread out over the length of the TGRA in addition to Wanzhou. Figure 2-3 shows all the sampling locations on a map and the exact locations are provided as latitude and longitude for each.

The data from the preliminary comparative study, summarized in Table 2-4, indicate minimal spatial variability in basic water quality and current within the reservoir. For this reason and because of the logistic challenges just described, only two locations (the ports of the city of Wanzhou and the city of Yunyang) in the reservoir were selected for longer-term sampling. These two locations were chosen for their accessibility as well as their relevance to public health due to potential exposure pathways (e.g. fishing, washing, swimming, and drinking).

Time Series Data Sampling Locations

Field data were collected in collaboration with Chongqing University in the reservoir along the banks of the cities of Wanzhou (编号) and Yunyang (云阳) seasonally from July 2005 to May 2007 capturing 13 out of 22 months. The data attempts to provide a baseline for water quality parameters most relevant to cyanobacteria population dynamics. At each of these two sampling

TABLE 2-4: Summary Data From Preliminary Comparative Study (7 Sites)

Parameter	Mean	Std Dev.	% Std Dev.
pH	7.44	0.32	4.3
Cond (mC/cm)	0.337	0.019	5.5
Turb (NTU)	453	131.0	28.9
DO (mg/L)	6.0	0.95	15.7
Temp (°C)	25.6	1.8	7.1
Secchi Disk (cm)	11.7	5.6	47.5
N (mg/L)	1.37	0.33	24.3
P (mg/L)	0.15	0.04	25.5

locations (Wanzhou and Yunyang), water grab samples were collected at 9-12 individual sites approximately 100 m apart across a roughly 500x200 m² grid section of the reservoir near the location of drinking water withdrawal pumps and public access points. Thus the sites encompass both banks of the reservoir as well as deep water sites in the middle. Wanzhou and Yunyang were chosen for monitoring because of the extensive use of the reservoir water by the populations of these cities for swimming, fishing, washing, and drinking. Feasible access to the sites for logistical collection and transport of the samples was also a factor.

Sampling Schedule

Sampling occurred between June 2005 and April 2007. The cross-sectional data was collected at the beginning and end of the sampling period (June 2005 and March 2007), while the time series data was collected regularly throughout this time period. Table 2-5 below shows the specific dates for sampling and identifies any exceptions in data collection for that particular sampling event.

2.4 Results

2.4.1 Results (Klamath Site)

Figure 2-4 shows a histogram of the chlorophyll-a concentrations in Copco and Iron Gate Reservoirs. The mean and median concentrations were 2445.7 µg/ml and 116 µg/ml, respectively. These chlorophyll concentrations are indicative of a hypereutrophic water body. There have been numerous indices created to calculate the trophic state of water. The Carlson's Trophic State Index (TSI) uses information on total phosphorus, transparency, and chlorophyll-a to determine where on the continuum from oligotrophic and hypertrophic a particular water body falls. Chlorophyll-a levels above 50 µg/ml are considered hypereutrophic (Carlson 1977).

Table 2-6 shows the mean and median values for the other parameters measured in the field. The average transparency of less than a meter is that of a eutrophic or hypereutrophic water body. Similarly the phosphate concentration in 2007 ranged between 2.5 and 5.1 µmol/L (Moisander et al.; Moisander et al. 2009) or 238-485 ppb, which is also representative of hypertrophic water.

Table 2-5: Sampling Schedule

Date	Locations Sampled		Parameters of the standard suite of 10 not available for sampling event
7/19-20/2005	Shi Bao Zhai Wanzhou Zhang Fei Temple	Wu Shan Da Ning River Mao Ping	
7/28/2005	Wanzhou		
8/15-16/2005	Wanzhou Yunyang		
10/21/2005	Wanzhou Yunyang		
11/07/2005	Wanzhou Yunyang		
11/20/2005	Wanzhou Yunyang		
12/17/2005	Wanzhou Yunyang		
1/11/2006	Wanzhou Yunyang		
3/17/2006	Wanzhou Yunyang		
4/09/2006	Yunyang		pH, cond, turb, DO, temp
4/12/2006	Yunyang		N, P
7/1-2/2006	Wanzhou Yunyang		
10/3/2006	Yunyang		
11/26- 27/2006	Wanzhou Yunyang		Turb
3/23-24/2007	Wanzhou Yunyang Xiao Jiang Old Yunyang	Tang Xi He Wushan Daling He Ba Dong	Turb
4/24/2007	Yunyang		Turb

Figure 2-4: Histogram of July 2007 Chlorophyll-a Concentrations in the Iron Gate and Copco Reservoirs

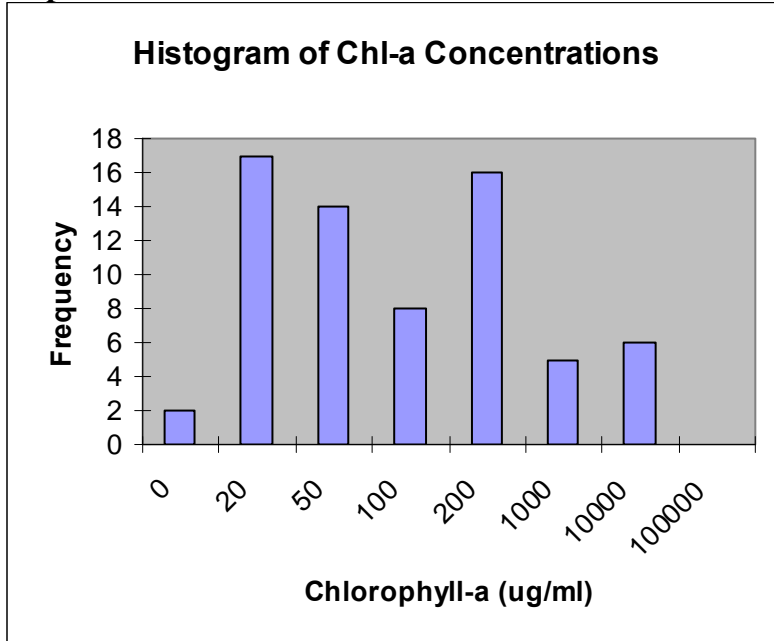


Table 2-6: General Water Data Summary for Klamath Reservoirs

	Surface	Surface	Surface	Surface	Surface	Secchi (m)	Water Depth (feet)
	pH	Cond	Turb	DO	Temp		
MEAN	9.33	0.122	21	13.42	24.98	0.86	46.6
MEDIAN	9.37	0.119	19	13.19	24.95	1	18.8

Figures 2-5 and 2-6 show histograms for the nitrate and soluble reactive phosphorus concentrations in the reservoirs from May-September 2007. Figures 2-7 and 2-8 show the microcystin concentrations and the *Microcystis aeruginosa* cell counts May-September 2007 for eleven sampling sites in the reservoirs.

2.4.2 Results (Yangtze Site)

Figure 2-9 shows the individual site values for these seven water quality parameters measured during the June 2005 spatial cross-sectional survey on the central sections of the study area between Shi Bao Zhai and the Da Ling River. The percent coefficient of variation calculated from laboratory replicates indicates that only some sampling sites for some parameters can clearly be distinguished from the rest. For example, after accounting for laboratory variance, Wu Shan has the highest values for DO, temperature, pH and nitrate and Wanzhou has the highest values for phosphate. While variation was observed among these sites, full analysis for trend and spatial pattern was not possible, although it was reported by previous studies (Zeng et al. 2007). Zeng (2007) averaged their nutrient data over all time points (March, May, July, August, and October 2004 and March 2005) for each site in their presentation of the data. Their nitrate values were about an order of magnitude less than the values recorded for this study, while their reported reactive phosphorous values matched with those of this study. The transparency, temperature and reactive phosphorous data reported for the rainy season (August 2004) in Zeng

(2006) matched the ranges measured in this cross-sectional survey, while their reported nitrate values again were an order of magnitude below the values measured here.

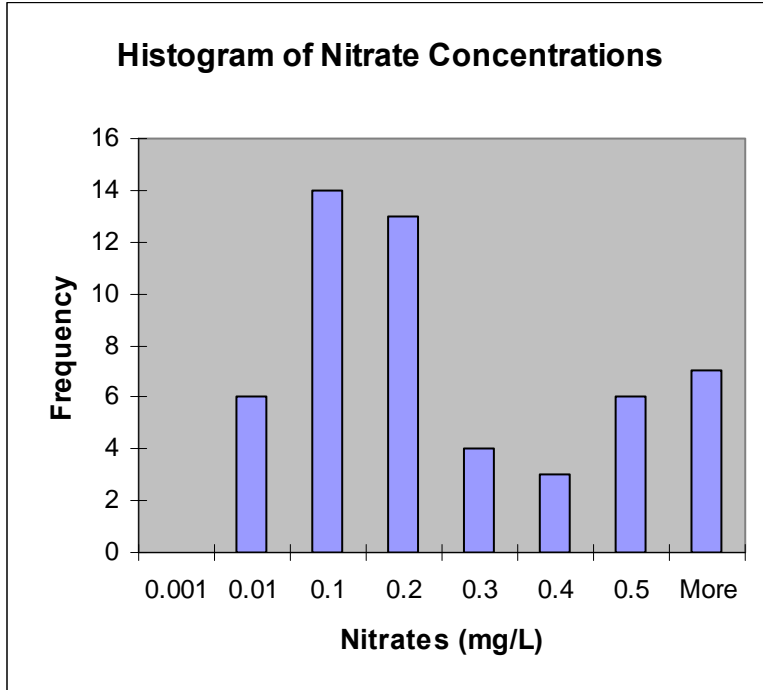
Figure 2-10 shows the expanded spatial cross-sectional survey conducted in March 2007 in which total phytoplankton cell counts were measured in addition to the other parameters and the number of sites was expanded from 5 to 8 on a section of the river shifted to the east of the 2005 survey, but including three of the five sites from 2005. Overall higher nutrient levels than in 2005 were expected due to the spring versus earlier summer sampling season. The drop in DO at the two Wushan sites may be attributable to inflow from the Daling River. Table 2-7 shows the minimum and maximum of the means of the samples taken at each location. Previous findings displayed a spatial trend, but this was not observed in these data. The phytoplankton populations between the seven sites ranged between 5-45 cell units/mL. Cell counts did not increase downstream, but rather were highest at Yunyang and in the tributaries where they join the Yangtze.

Table 2-7: Water Measurements in March 2007 at 8 sites spanning the TGR

	Maximum	Minimum	All Reservoir Mean	All Reservoir Std Dev
Secchi Disk (cm)	347	210	282.5	40.4
pH	7.9	5.6	7.4	0.7
Conductivity (uS/cm)	0.376	0.349	0.363	0.010
DO (mg/l)	10.49	8.88	9.61	0.56
Temperature (°C)	15.9	14.5	15.3	0.5
NO3-2 (mg/l)	1.73	1.14	1.50	0.19
PO4-3 (mg/l)	0.47	0.21	0.29	0.08

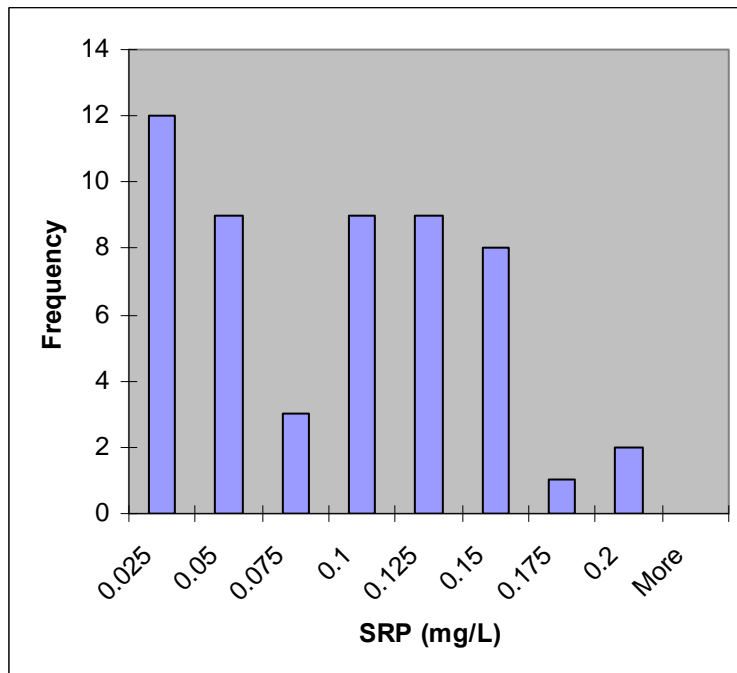
For the longitudinal time series data, mean concentrations of water parameters were compared between the two sampling locations (i.e. Wanzhou and Yunyang) using the Kruskal-Wallis technique (Zar 1996) due to the small sample size with a null hypothesis that the sample means were not different at a significance level of 0.05 (Stata 9.0, StataCorp LP, 2005). There was some variability in nutrient concentrations, temperature and dissolved oxygen between the port at Yunyang and the port at Wanzou. For all sampling events except January 2006 and March 2007 the nitrate concentrations were significantly different between the two cities ($p < 0.05$). For all sampling events except November 2005, January 2006, and November 2006 the reactive phosphorous concentrations were also significantly different between the two cities ($p < 0.05$). Table 2.8 shows the minimum and maximum of the means of the samples taken at each location over the two years. Figure 2-11 shows the mean values (\pm Standard Error) for the different parameters at both Yunyang and Wanzou over the monitoring period. The parameters are affected by seasonal fluctuations in outdoor temperature, precipitation, and flow. An analysis for temporal trend was not conducted since the time-series only spans two years which prevents us

Figure 2-5: Histogram of July 2007 Nitrate Concentrations in the Iron Gate and Copco Reservoirs



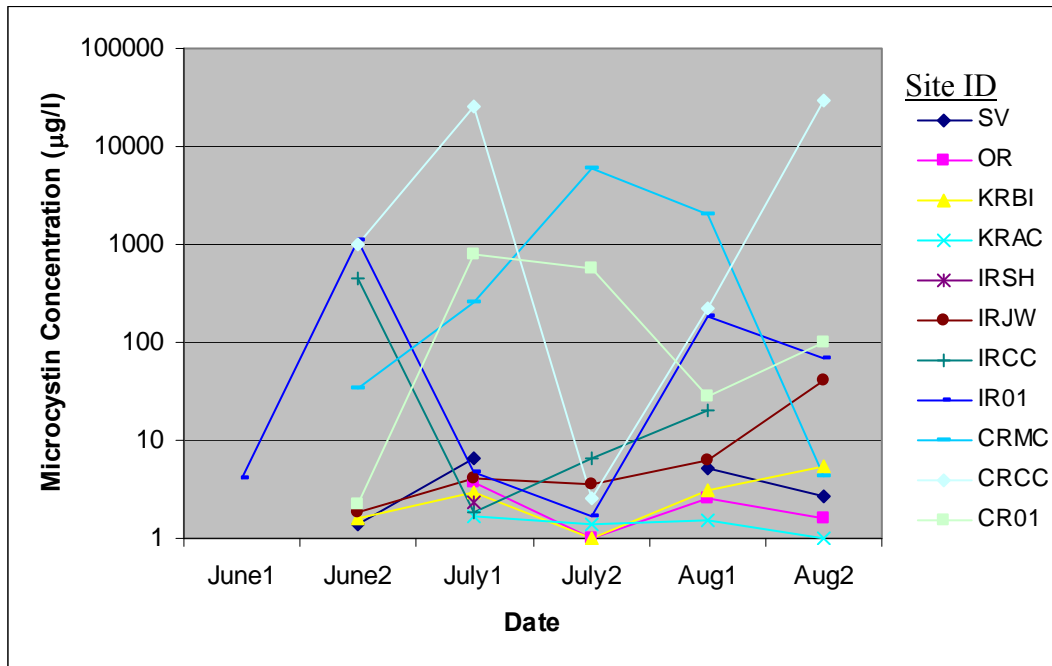
*These data were collected and made available by the Karuk Tribe.

Figure 2-6: Histogram of July 2007 Soluble Reactive Phosphorus Concentrations in the Iron Gate and Copco Reservoirs



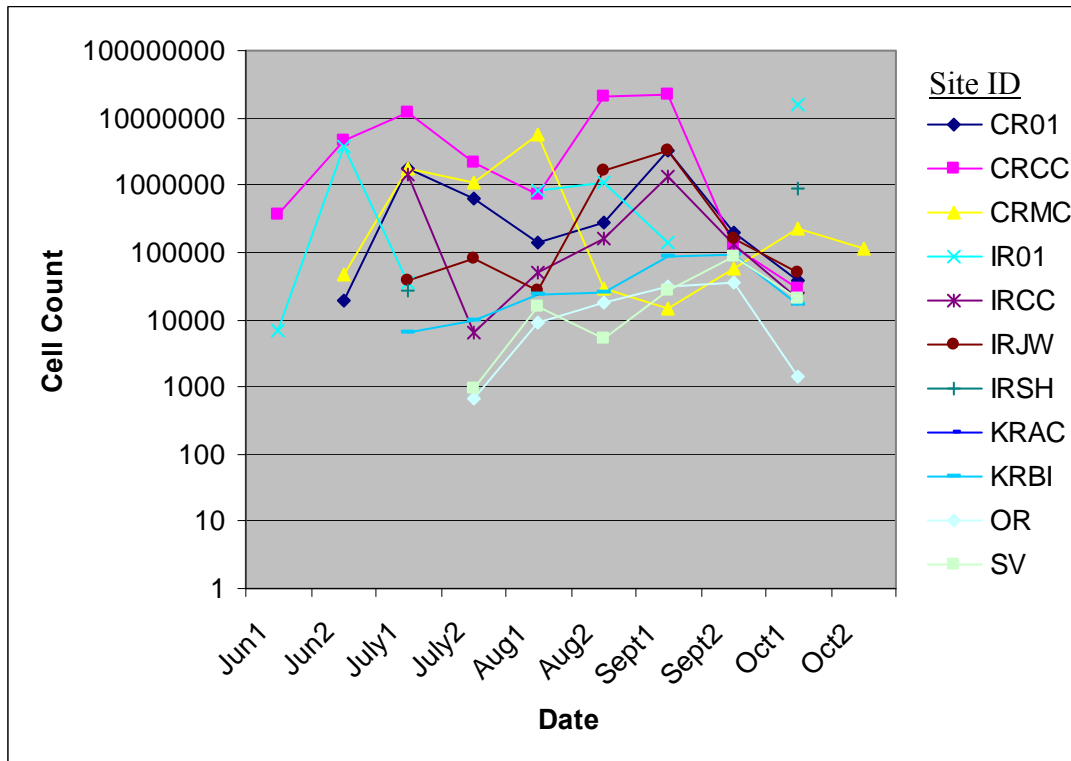
*These data were collected and made available by the Karuk Tribe.

Figure 2-7: Summer 2007 Microcystin Concentrations in Iron Gate and Copco Reservoirs



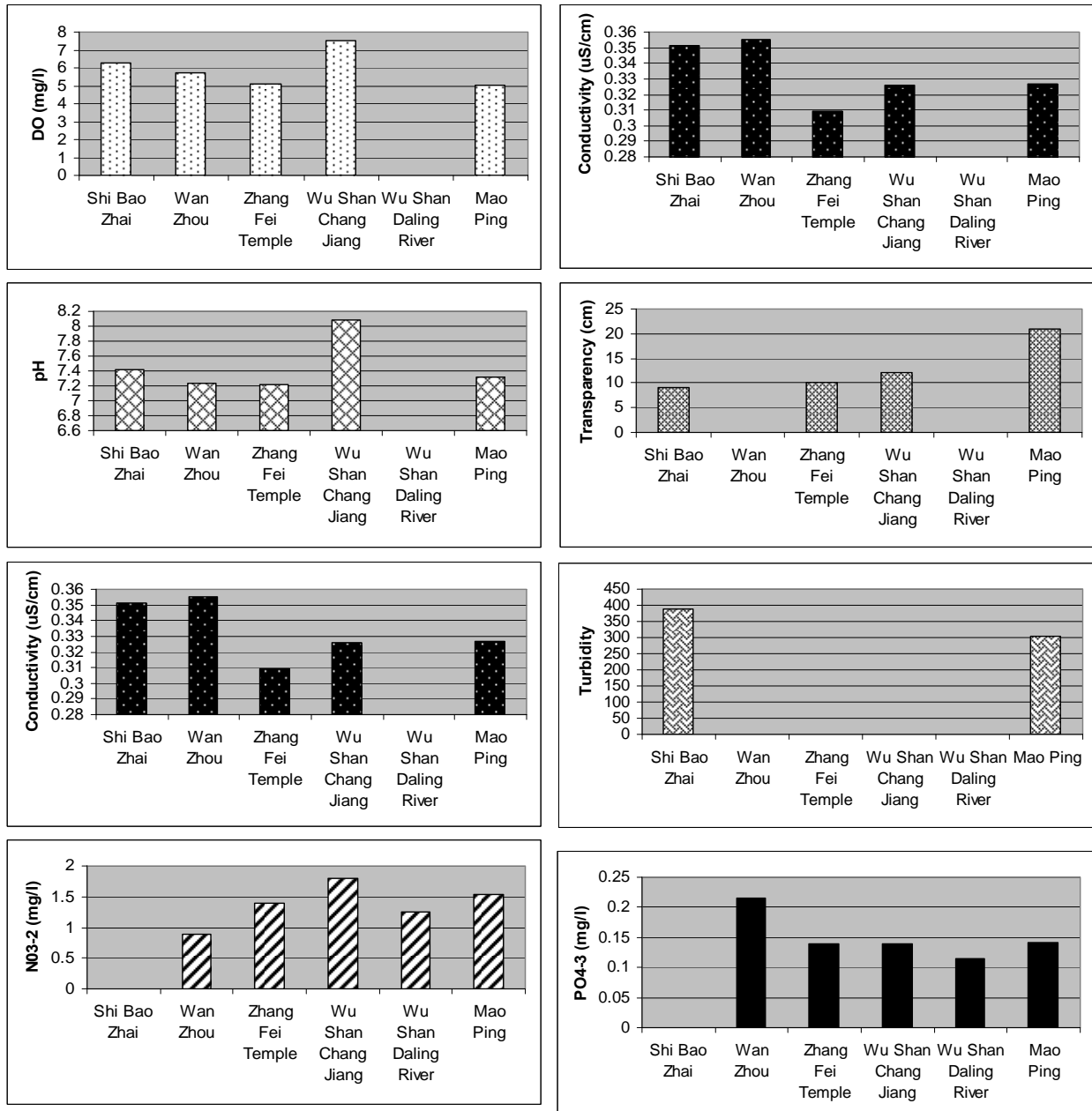
*These data were collected and made available by the Karuk Tribe. The number in the dates refers to the first or second sampling event of the month.

Figure 2-8: Summer 2007 Microcystis aeruginosa Cell Counts in the Iron Gate and Copco Reservoirs



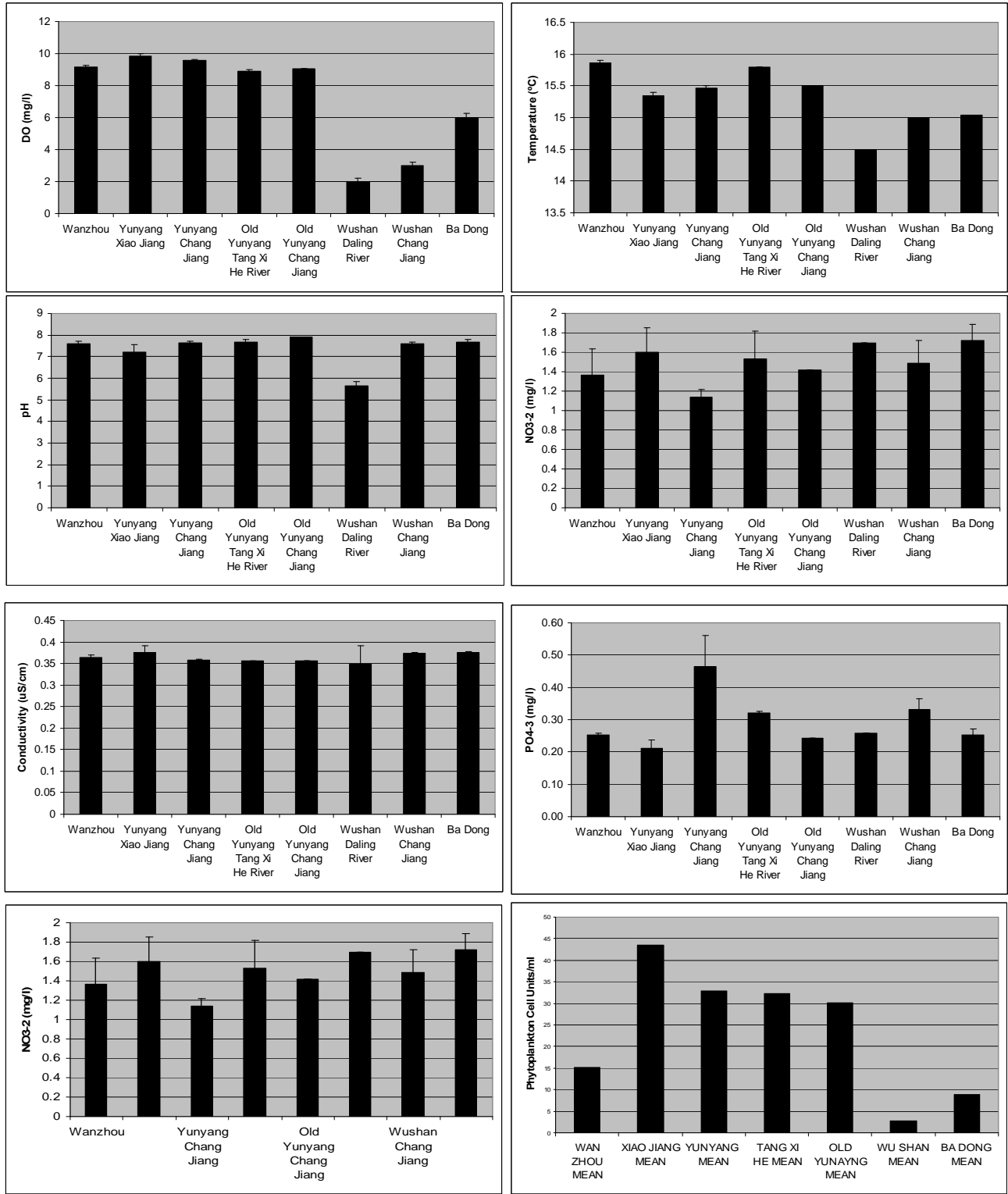
*These data were collected and made available by the Karuk Tribe.

Figure 2-9: June 2005 Spatial Cross-Section on the TGDR. Sites listed from upstream to downstream.³



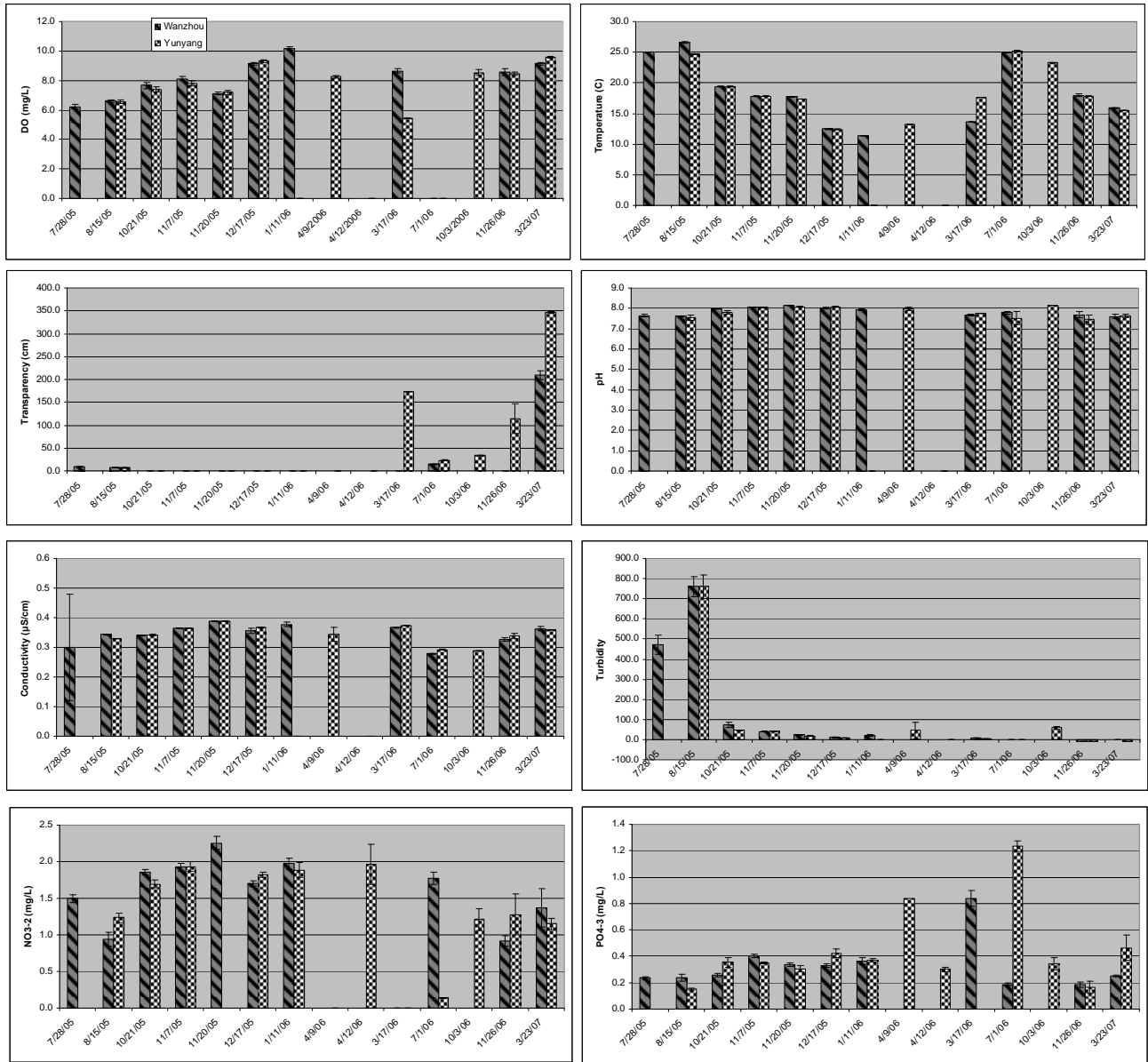
³ Absent columns for any site or date reflects unavailable data, not a null value.

Figure 2-10: March 2007 spatial cross-section on the TGDR. Sites listed from upstream to downstream.⁴



⁴ Absent columns for any site or date reflects unavailable data, not a null value.

Figure 2-11: Longitudinal time-series for Wanzhou and Yunyang on the TGDR.⁵



⁵ Absent columns for any site or date reflects unavailable data, not a null value.

Table 2-8: Summary of Water Quality Parameters in Wanzhou and Yunyang– Maximum and Minimum are taken from the Means for each Sampling Dates

	Secchi Disk (cm)	pH	Cond	Turb	DO	Temp (°C)	N03 ⁻² (mg/l)	PO4 ⁻³ (mg/l)
Y Maximum	347.0	8.1	0.4	759.5	9.6	25.2	1.97	1.24
Y Minimum	7.3	7.4	0.3	11.7	5.4	12.4	0.14	0.15
W Maximum	14.8	8.1	0.4	759.6	10.2	26.6	2.25	0.84
W Minimum	0.0	7.6	0.3	3.0	6.2	11.4	0.91	0.19

from distinguishing temporal trends from seasonal changes especially given irregularities in the sampling frequency.

Chlorophyll-a concentrations ranged between 5 mg/L and 700 mg/L. The highest concentrations were measured in Xiao Jiang River, a tributary that joins the Yangtze just north of Yunyang and at the location of the old city of Yunyang. The chlorophyll concentrations were highest in the spring when levels were usually at least 10 times higher than those of the summer rainy season when flows are highest.

Similarly phytoplankton populations were largest during the spring (March/April), with a second population growth peak in the fall (October/November). Although water temperatures were considerably warmer during the summer months, heavy rainfall at this time causes high suspended sediment levels and increased turbidity. In the spring rainfall had not yet begun meanwhile the water temperatures had risen to levels that foster greater phytoplankton growth. In the fall rainfall had stopped and water temperatures had not cooled to a level no longer conducive for phytoplankton growth. Diatoms dominated the phytoplankton populations for all samples, and asterionella, melosira, synedra, and fragillaria were the most abundant genera. The only cyanobacteria genus detected was oscillatoria. While microcystis colonies were not detected, individual cells may have been present as observed by Zeng (Zeng et al. 2006) but escaped detection in this study due to the counting technique (Sedgwick rafter counting cells under lower magnification) and low population size.

2.5 Temporal and Spatial Variability

There are numerous challenges to field data collection for toxic cyanobacteria surveillance monitoring and quantification. Researchers, managers, and regulators struggle to select sample locations, sampling technique, and sampling depth appropriately to provide good representation of the water body's bloom conditions. It is precisely these challenges that one hopes to overcome using remote detection methods, but the process of collecting field data for ground truthing still requires addressing these issues. Phytoplankton are not evenly distributed throughout the water column, and cell densities can vary 10 to 1000-fold (if not more) between surface samples and depth samples. The depth distribution results from the capacity of some phytoplankton like *Microcystis aeruginosa* to depth regulate using cellular packaging and vacuoles to move within the ideal photic zone that maximizes their ability to photosynthesize during daylight. Their cell densities at any given spot on the water surface can therefore vary across the span of a day due to this vertical variation. Following on the model of other agencies, the Klamath Blue Green Algae Work Group opted to collect surface grab samples to sample blooms and make decisions regarding public health risks. They also collected some samples at depth gradients to help map phytoplankton population trends and some samples integrating surface and a 1 meter depth collection were collected for this study. The more dense a

cyanobacterial bloom is, the less relevant depth samples are in a remote detection context since the surface reflectance will be dominated by the surface bloom. When the phytoplankton are more evenly distributed in the epilimnion of the water column, there will be some contribution, down to a limited depth, to the light reflectance properties at any given point.

Temporal and spatial variation is also created horizontally through ambient wind, water currents and other stimulation (boating, recreation, hydroelectric facility operations, etc.). Therefore the number of samples collected, their geographic spacing, and the location of the each is important. Some monitoring programs have chosen to use just one sample whose measurements for cell densities, chlorophyll-a or toxin concentration will be used to make risk management decisions, but most try to collect more than one sample. Others collect collect samples across transects or collect one shore and one open water site. Nebraska's current beach sampling protocol relies on a single, mid-beach grab sample (Brakhage 2008; Walker et al. 2008; Brakhage 2009). In a past study of Lake Champlain's cyanobacteria, Vermont did one vertical tow at seventeen sites spaced over the lake (Lake Champlain Basin Program 2007). New Zealand's guidelines state that a near-shore sample be collected by compositing five sup-samples collected from a 50 cm depth along a 20-30 m transect parallel to the shore (Wood et al. 2008). USGS recommends selecting a single representative site when conducting a reconnaissance study to determine presence or absence of cyanobacteria, but multiple sites for monitoring or interpretive studies (Graham et al. 2008). Due to the temporal and spatial variation, it is difficult to ensure adequate representation of the true variability. The Klamath Blue Green Algae Work Group selected 11 sites spread across the reservoirs that include both shore/access points as well as open water sites. This decision was partially based on collection time limitations, cost, and consistency with prior work.

In general, it is best to collect samples during the mid-day hours. Ahn, et. al. (Ahn et al. 2008), found cyanobacterial depth modulation resulted in optimal *Microcystis* sampling periods at 12 noon and 3 pm when the biomass is highest and biomass was lowest at the 9 AM and 5 PM sampling times. At noon blooms would be surfacing to maximize photosynthesis.

Environmental monitoring that describes actual variability is regularly addressed by increasing the sample size. Moisander collected data over a 16 hour period on August 26, 2007 at one site in the Iron Gate Reservoir. Her data (Moisander 2008) is replicated in Figure 2-12 which succinctly demonstrates how drastically the density of the bloom can shift over short periods of time at any given site.

2.6 Discussion

The Klamath reservoirs have already displayed a seasonal pattern of toxic cyanobacteria blooms which have been well documented since 2005. Upper Klamath Lake has also had regular algal blooms of non-toxic species. The field data were collected to confirm the trophic status and current situation in the reservoirs and to be used in model development and classification of the remote sensing data, rather than as a way to characterize the water quality of the reservoirs. The Three Gorge Dam Reservoir, on the other hand, as a newly formed water body with changing hydrodynamics was monitored to begin to establish the water quality situation with respect to cyanobacteria and phytoplankton dynamics as an exploration into the need for further research in this area.

Nutrient concentrations in TGDR exceed those measured pre-dam. The dissolved inorganic phosphorous (DIP) concentrations in the stretch of the Yangtze that comprises the new TGDR ranged between 1.5-2.5 μM (~0.14-0.24 mg/L) in April-May 1997 (Liu et al. 2003). In the current study the DIP concentrations had risen to 0.27-0.35 mg/L in April 2006. Zeng et al. (Zeng et al. 2006) found no significant correlation between soluble nutrients and algal abundance in the TGDR during their monitoring in April 2005, however they reported much lower DIP concentrations of 0.06 ± 0.03 mg/l. They reported similar findings in a subsequent paper (Zeng et al. 2007) which also referred to data collected in the spring of 2004 and 2005. The Chinese Ministry of Environmental Protection (MEP) has published an annual Three Gorges Bulletin starting in 2003. The 2007 Three Gorges Bulletin includes a description of 2006 surface water quality in the mainstem of the Yangtze River according to water quality categories (“Grade I-V”) defined by the MEP (Chinese Ministry of Environmental Protection 2008) as listed in Table 2-9. In April of 2006 the Yangtze River sections corresponding with those measured here were classified into Grades III and IV meaning their total phosphorous concentration ranged between 0.1 and 0.3 mg/L which matches with this study’s data. Two years earlier in 2004 the MEP classified all Yangtze River sections as Grade III or better, with none in Grade I, 6.3% in Grade II, and 93.7% in Grade III.

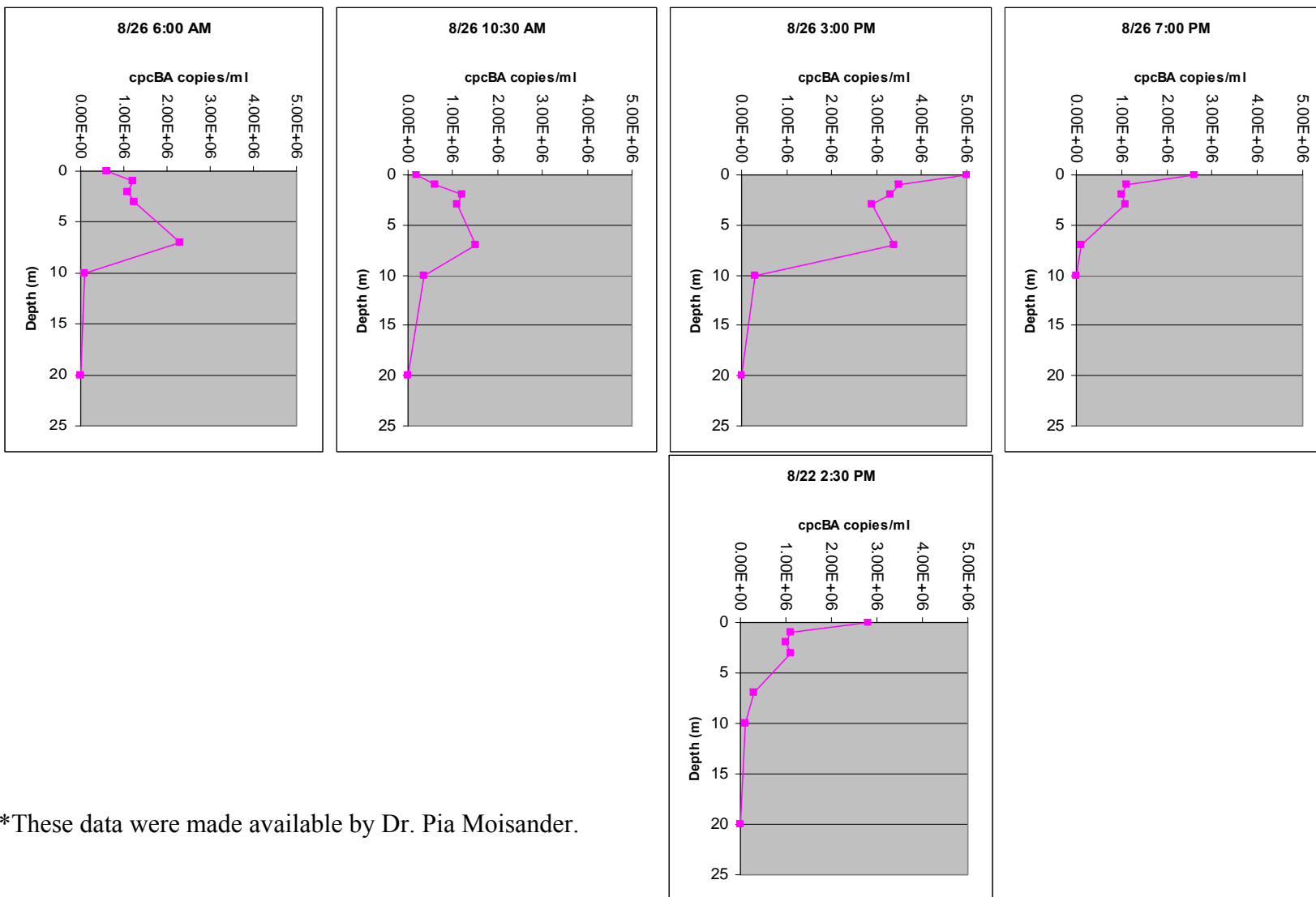
Table 2-9: Chinese MEP Water Quality Categories

Grade	Total Phosphorous	Total Nitrogen
Grade I	<0.02 mg/L	<0.2 mg/L
Grade II	0.02-0.1 mg/L	0.2-0.5 mg/L
Grade III	0.1-0.2 mg/L	0.5-1.0 mg/L
Grade IV	0.2-0.3 mg/L	1.0-1.5 mg/L
Grade V	0.3-0.4 mg/L	1.5-2.0 mg/L

Overall cell counts in the TGDR are much lower than would be predicted by the nutrient concentrations observed between 2005 and 2007. According to the Carlson Trophic State Index (TSI) (Carlson 1977) eutrophic water bodies are characterized by phosphorus concentrations between 0.048 and 0.096 mg/L where blue-green algae begin to dominate and problems with algal scums begin. With increasing phosphorous concentrations the density of algal biomass grows as the water becomes more and more hypereutrophic. The reactive phosphorous concentrations in the TGDR as well as the secchi disk transparency would classify it as a hypereutrophic water body according to the Carlson TSI. The shallow secchi depth reading may be attributable to the high suspended sediment levels in the main stem of the Yangtze, but the phosphorous concentrations are oddly contradictory to the actual phytoplankton populations measured. The reservoir, with the flooding now near completion, shows surprisingly low populations of phytoplankton among all taxa (10-80 cell units/ml), given the nutrient concentrations and water temperature. The sparse populations of phytoplankton may be good news for the large cities that rely on their drinking water from the reservoir which would otherwise be susceptible to toxic cyanobacteria blooms.

The fact that dense algal scums have are not yet being regularly detected in the TGDR raises questions about other dynamics and water pollutants that may be preventing algae proliferation. The unexpected absence of both proliferations of harmful phytoplankton as well as even a healthy natural balance of phytoplankton, the basis of the aquatic ecological foodchain may be a result of the physical hydrology of the reservoirs whose deep, stratified waters may trap

Figure 2.12: Microcystis Spatial Variation Over 16 hours on the Copco Reservoir (Counts measured in gene copies). Data replicated from Moisander 2008.



*These data were made available by Dr. Pia Moisander.

phytoplankton below the photic zone and thereby inhibit the proliferation of their populations. Another explanation may be a disproportionately large population of grazing zooplankton. These observations, however, may also be indicative of unsuitable water quality to foster even healthy phytoplankton populations. Increasing levels of untreated industrial water may be a factor in this and should be evaluated in more depth in future studies. Xiao Jiang and Tang Xi He, two tributaries to the Yangtze both had larger phytoplankton populations than the main stem of the Yangtze, despite having lower nutrient concentrations. These data tend to substantiate the latter speculation that industrial pollutants in the main stem might actually be curbing phytoplankton growth which would otherwise be expected, but more research is needed.

Furthermore, in July 2008 blue-green algae blooms covered a 25-kilometer stretch of the Xiangxi River. This tributary to the Yangtze River also showed blue-green algae blooms during an earlier study (Zeng et al. 2006). A bloom occurred in Xiao Jiang River, a tributary whose confluence with the Yangtze River is at the city of Yunyang where sampling was conducted for this study. A bloom of the green algae, *lemna minor*, occurred in April 2007 in the spring when phytoplankton density is the highest according to our monitoring. Although the blooms were problematic for fishing, boating and other activities in the water they were not toxic cyanobacteria. In May 2007 a large toxic *microcystis aeruginosa* bloom occurred in Lake Tai (Taihu) located near the city of Wuxi which resulted in the contamination of local drinking water. Wuxi is north of the Yangtze River and Tai Hu Lake is connected to the Yangtze Watershed much farther downstream of the Three Gorges Dam.

Prior to the construction of the TGD, others predicted a loss of biodiversity and ecosystem services from habitat fragmentation (e.g. its reduction as well as spatial reconfiguration) (Wu et al. 2003). The reduction in phytoplankton population diversity observed in this study substantiates the occurrence of this fragmentation. However, algal blooms are occurring in the Yangtze Watershed as described above.

The data presented here did not detect the presence of toxic cyanobacteria blooms in the TGDR, however the other water quality data indicate that conditions would be expected to foster harmful infestations of these toxic species of phytoplankton. Due to the level of water treatment in the area, water users are at risk of high exposures in the event of a bloom. The status of flooding and hydrologic modification in the TGDR along with the limited number of studies on water quality and phytoplankton community in this region make it difficult to form conclusions on trends or shifts. However in combination these studies begin to lay the basis for baseline understanding of the ecosystem health and water quality in the TGDR. This study does indicate a need for regular surveillance and monitoring plans in order to protect public and ecological health. Such surveillance may be facilitated through the use of remote sensing. The subsequent chapters evaluate the application of remote sensing for these purposes in the context of the Klamath River reservoir, where large blooms regularly occur. Using the Klamath field site to evaluate the remote sensing applications made sense so that appropriate classification could be developed for actual blooms. The situation on the Yangtze River, however, clearly indicates a water system that may be at risk for future cyanobacterial blooms and a population with regular exposure pathways that would put their health at risk in the event of a major bloom in TGDR. The Klamath reservoirs make good field sites for evaluation of the remote sensing tools, and the

TGDR is an important case for future applications of successful remote sensing tools (see Chapter 3 and 4).

- Ahn, C.-Y., S.-H. Joung, et al. (2008). "Comparison of sampling and analytical methods for monitoring of cyanobacteria-dominated surface waters " *Hydrobiologia* **536**(1): 9.
- American Public Health Association (APHA) (2001). *Standard Methods for the Examination of Water and Wastewater*. 21st edn. Washington DC, American Public Health Association.
- Board on Environmental Studies and Toxicology and Water Science and Technology Board (2008). *Hydrology, Ecology, and Fishes of the Klamath River Basin*. Washington D.C., National Academies Press.
- Brakhage, P. A. (2008). "Cyanobacteria, The Nebraska Experience." Retrieved April 23, 2010, from [http://www.deq.state.ne.us/Publications/a9f87abbcc29fa1f8625687700625436/dc50a548157633e786257590005cec58/\\$FILE/LakeLine-Cyanobacteria.pdf](http://www.deq.state.ne.us/Publications/a9f87abbcc29fa1f8625687700625436/dc50a548157633e786257590005cec58/$FILE/LakeLine-Cyanobacteria.pdf)
- Brakhage, P. A. (2009). "Cyanobacteria, The Nebraska Experience." Retrieved April 23, 2010, from [http://www.deq.state.ne.us/Publications/a9f87abbcc29fa1f8625687700625436/dc50a548157633e786257590005cec58/\\$FILE/LakeLine-Cyanobacteria.pdf](http://www.deq.state.ne.us/Publications/a9f87abbcc29fa1f8625687700625436/dc50a548157633e786257590005cec58/$FILE/LakeLine-Cyanobacteria.pdf)
- Carlson, K. and R. Raymond (2008). Blue-Green Algae (Cyanobacteria) and Microcystin Monitoring Results in the Vicinity of the Klamath Hydroelectric Project: April, May and June 2008.
- Carlson, R. E. (1977). "A Trophic State Index for Lakes." *Limnology and Oceanography* **22**(2): 361-369.
- Chen, Z., L. Yu, et al. (2001). "The Yangtze River: an introduction." *Geomorphology* **41**(2-3): 73-75.
- Chinese Ministry of Environmental Protection. (2008, Feb. 12, 2008). "Three Gorges Bulletin 2007." *Three Gorges Bulletin* Retrieved May 17, 2008, from http://english.sepa.gov.cn/standards_reports/threegorgesbulletin/.
- Congalton, R. G. (1991). "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* **37**(1): 35-46.
- Graham, J. L., K. A. Loftin, et al. (2008). Guidelines for design and sampling for cyanobacterial toxin and taste-and-odor studies in lakes and reservoirs: U.S. Geological Survey Scientific Investigations Report 2008-5038: 39.
- Kann, J. (2006a). Copco/Irongate Reservoir and Klamath River Cyanobacteria Results: Aug 23-24, 2006 -- Technical Memorandum. Ashland, Aquatic Ecosystem Sciences: 9.
- Kann, J. (2006b). Copco/Irongate Reservoir and Klamath River Cyanobacteria Results: July 13th and July 27th, 2006 -- Technical Memorandum. Ashland, Aquatic Ecosystem Sciences: 9.
- Kann, J. (2006c). *Microcystin aeruginosa* Occurrence in the Klamath River System of Southern Oregon and Northern California -- Technical Memorandum prepared for Yurok Tribe Environmental and Fisheries Program. Ashland, Aquatic Ecosystem Sciences: 26.
- Kann, J. and E. Asarian (2005). 2002 Nutrient and hydrologic loading to Iron Gate and Copco Reservoirs, California. Kier Associates Final Report to the Karuk Tribe Department of Natural Resources, Orleans, California: 59 + appendices.

- Kann, J. and E. Asarian (2006). Klamath River Nitrogen Loading and Retention Dynamics, 1996-2004. Kier Associates Final Technical Report to the Yurok Tribe Environmental Program, Klamath, California.: 56 + appendices.
- Kann, J. and E. Asarian (2007). Nutrient Budgets and Phytoplankton Trends in Iron Gate and Copco Reservoirs, California, May 2005-May 2006, Aquatic Ecosystem Sciences & Kier Associates for the Karuk Tribe of California and State Water Resources Control Board: 81.
- Kann, J. and S. Corum (2006). Summary of 2005 Toxic *Microcystis aeruginosa* and Microcystin Trends in Copco and Iron Gate Reservoirs on the Klamath River, CA.
- Kann, J. and S. Corum (2007). Summary of 2006 Toxic *Microcystis aeruginosa* and Microcystin Trends in Copco and Iron Gate Reservoirs, CA.
- Kann, J. and S. Corum (2009). "Toxicogenic *Microcystis aeruginosa* bloom dynamics and cell density/chlorophyll a relationships with microcystin toxin in the Klamath River, 2005-2008."
- Lake Champlain Basin Program. (2007). "Monitoring and Evaluation of Cyanobacteria in Lake Champlain." Retrieved August 3, 2009, from http://www.lcbp.org/techreportPDF/56_BGA_2007.pdf.
- Liu, S. M., J. Zhang, et al. (2003). "Nutrients in the Changjiang and its tributaries." *Biogeochemistry* **62**(1): 1-18.
- Maurer, C., C. Wu, et al. (1997). Water Pollution and Human Health in China, Woodrow Wilson International Center: 28-38.
- Moisander, P. H. (2008). *Microcystis* vertical distribution, University of California, Santa Cruz.
- Moisander, P. H., M. Ochiai, et al. "Nutrient limitation of *Microcystis aeruginosa* in northern California Klamath River reservoirs." **Harmful Algae In Press, Corrected Proof**.
- Moisander, P. H., M. Ochiai, et al. (2009). "Nutrient limitation of *Microcystis aeruginosa* in northern California Klamath River reservoirs." *Harmful Algae* **8**(6): 889-897.
- North Coast Regional Water Quality Control Board (2010a). Final Staff Report for the Klamath River Total Maximum Daily Loads (TMDLs) Addressing Temperature, Dissolved Oxygen, Nutrient, and Microcystin Impairments in California, the Proposed Site Specific Dissolved Oxygen Objectives for the Klamath River in California, and the Klamath River and Lost River Implementation Plans.
- North Coast Regional Water Quality Control Board (2010b). Staff Report for the Klamath River Total Maximum Daily Loads (TMDLs) Addressing Temperature, Dissolved Oxygen, Nutrient, and Microcystin Impairments in California, the Proposed Site Specific Dissolved Oxygen Objectives for the Klamath River in California, and the Klamath and Lost River Implementation Plans.
- PacifiCorp (2004). Final Technical Report, Klamath Hydroelectric Project, Water Resources. Portland, Oregon.
- PacifiCorp. (2010). "Klamath River Hydroelectric Project Overview." Retrieved March 4, 2010, from <http://www.pacificorp.com/es/hydro/hl/kr.html>.
- Walker, S. R., J. C. Lund, et al. (2008). Nebraska Experience. *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs*: 139-152.
- Wood, S., D. P. Hamilton, et al. (2008). "New Zealand Guidelines for Cyanobacteria in Recreational Fresh Waters: Interim Guidelines." Retrieved April 23, 2010, from <http://www.mfe.govt.nz/publications/water/guidelines-for-cyanobacteria/index.html>.

- Wu, J. G., J. H. Huang, et al. (2003). "Three-Gorges Dam - Experiment in habitat fragmentation?" Science **300**(5623): 1239-1240.
- Zar, J. (1996). Biostatistical Analysis. Upper Saddle River, New Jersey, Prentice Hall.
- Zeng, H., L. R. Song, et al. (2006). "Distribution of phytoplankton in the Three-Gorge Reservoir during rainy and dry seasons." Science of the Total Environment **367**(2-3): 999-1009.
- Zeng, H., L. R. Song, et al. (2007). "Post-impoundment biomass and composition of phytoplankton in the Yangtze River." International Review of Hydrobiology **92**(3): 267-280.

3 Hyperspectral Image Analysis

3.1 Introduction

Mapping, monitoring and management of aquatic environments may be enhanced by remote sensing's synoptic capability for attaining spatial and temporal data over expansive areas otherwise logistically unattainable. This is a valuable tool for monitoring toxic cyanobacteria. Image analysis of remotely sensed data may provide quantitative information on water depth, optical water properties and benthic habitat composition. The visible pigmentation of cyanobacteria lend it to remote detection.

This chapter presents the analysis of hyperspectral high spatial resolution (2 meter) imagery from SpecTIR. The company SpecTIR LLC headquartered in Reno, Nevada operates hyperspectral remote sensing airborne missions on demand. They offer visible near infrared/short-wave infrared (VNIR/SWIR), mid-wave infrared (MWIR) and long-wave infrared hyperspectral scanning sensor systems depending on the project needs. Their visible/near-infrared sensor provides 60 bands which are each 9.2 nm in width covering the electromagnetic spectrum from 391-961 nm. SpecTIR imagery was chosen because 60 bands provide sufficient spectral resolution to distinguish spectral signatures for different algal pigments, and the company's physical proximity to the field site made their cost competitive. The spectral resolution of SpecTIR is not quite as high, but still comparable to that of MODIS and SeaWiFS imagery whose bands are 2-4 nm wide in the same range of the electromagnetic spectrum and which have been used to detect, quantify and predict harmful algal blooms in the ocean.

3.1.1 Absorbance and Reflectance Properties of Algal and Plant Pigments

Chlorophyll-a is the primary photosynthetic pigment in phytoplankton. It absorbs more blue and red light than green. Satellite and airborne remote sensing make use of the natural incident sunlight, a portion of which is absorbed and a portion of which is reflected back where it can be detected by the sensors designed for the electromagnetic spectrum. As the concentration of phytoplankton in the water increases, the backscattered light increasingly changes from blue to green (Yentsch 1960).

Chlorophyll-a shows peak absorbance at 433/435 nm and 686 nm. Since it absorbs in the blue and red portion of the electromagnetic spectrum, it appears green. Carotenoids usually absorb high energy light at or below 500 nm. Phycobilins have peak absorbance between 550 and 650 nm. These absorbance features seen as peaks in an absorbance spectrum will appear as troughs, the exact opposite, in a reflectance spectrum. Figures 3-1, 3-2 and 3-3 are reproductions from Rowan 1989 (Rowan 1989) showing the absorbance spectra of pure chlorophylls, carotenoids and biliproteins/phycoobilins including phycocyanin.

Figure 3-1: Absorption spectra of chlorophyllous pigments, (A) Chlorophyll a and b in diethyl ether. (B) Chlorophyll c1 and c2 in acetone containing 2% pyridine. (C) MgDVP in 90% acetone. (Reproduced from Figure 3-3 of Rowan, 1989)

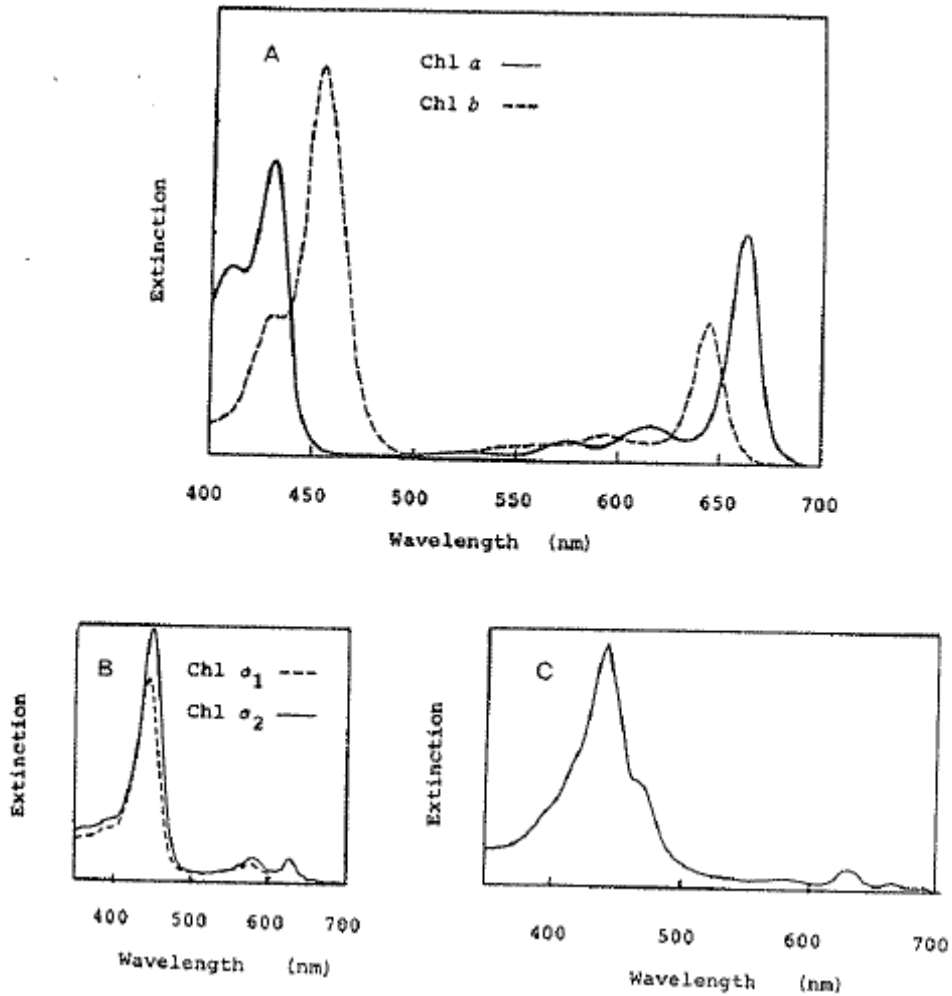


Figure 3-2: The absorption spectra of lycopene (...), a carotene without cyclic end-groups, and β -carotene (—) and echinenone (---), showing the effect of β -cyclization and insertion of a carbonyl group, respectively. (Reproduced from Figure 4-4 of Rowan, 1989)

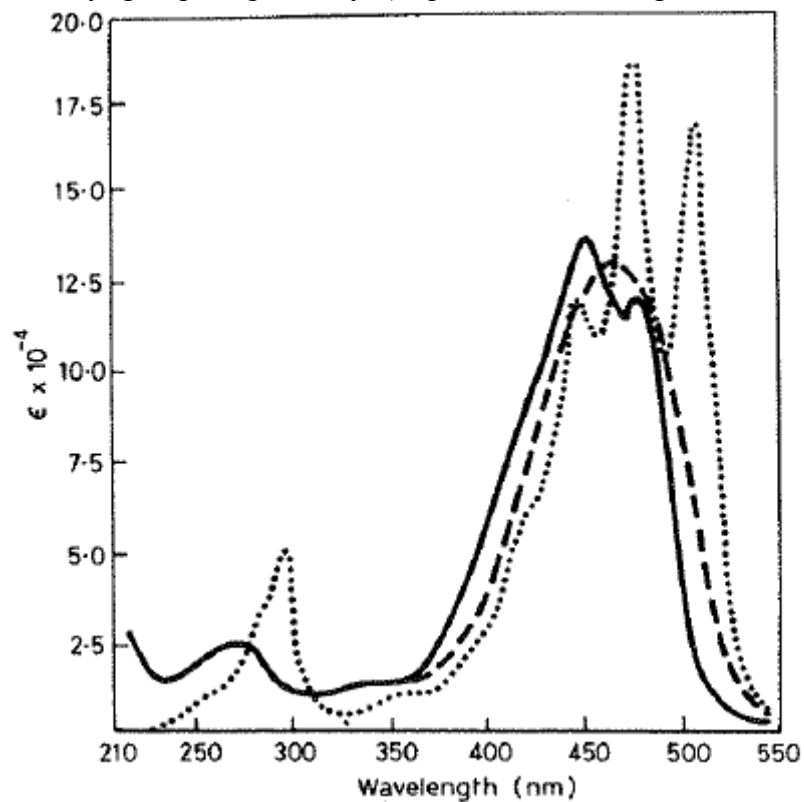
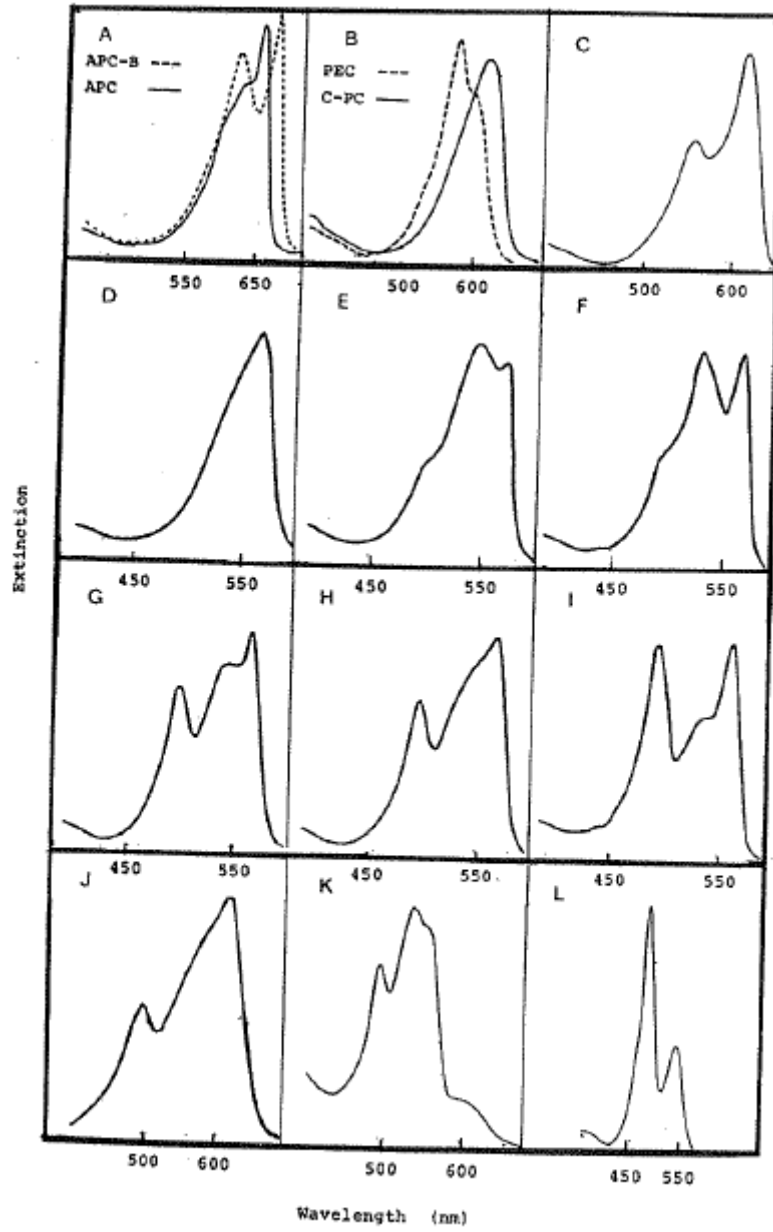


Figure 3-3: Absorption spectra of the biliproteins from the Cyanophyceae and Rhodophyceae. (reproduced from Figure 5-2 of Rowan 1989)



Figures 3-4 shows the reflectance spectra for several algae pixels on the left and several terrestrial vegetation pixels on the right which were visually selected to show the variation within and between these two classes in the SpecTIR image. Figure 3-5 shows the locations of these pixels which include both open water and shoreline sites for algae. As expected, the algae spectra show an absorbance trough at 631 nm characteristic of the bandwidth where phycocyanin absorbs and another absorbance trough 679 nm. Another trough is visible at 499 nm and a very small possible absorbance feature appears at 435 nm. One of these may be the second absorbance peak of chlorophyll typically found at 433 nm and the other may be an absorbance due to one or more carotenoids. In contrast all but the 631 nm absorbance peak are clearly visible in the terrestrial vegetation spectra. The absence of the phycocyanin peak is characteristic of this pigment unique to cyanobacteria.

Figure 3-4: SpecTIR Reflectance Spectra for water surface algae (left) and terrestrial vegetation (right)

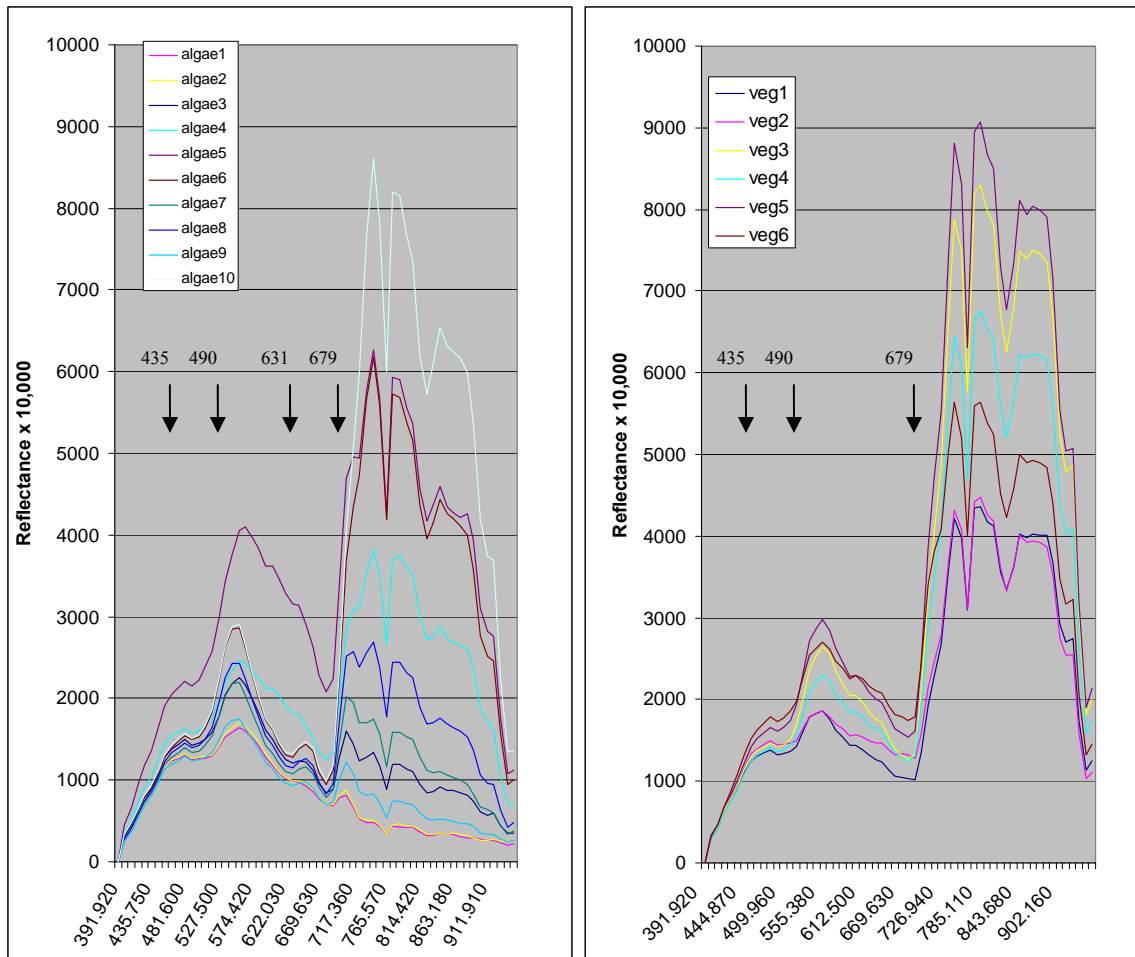
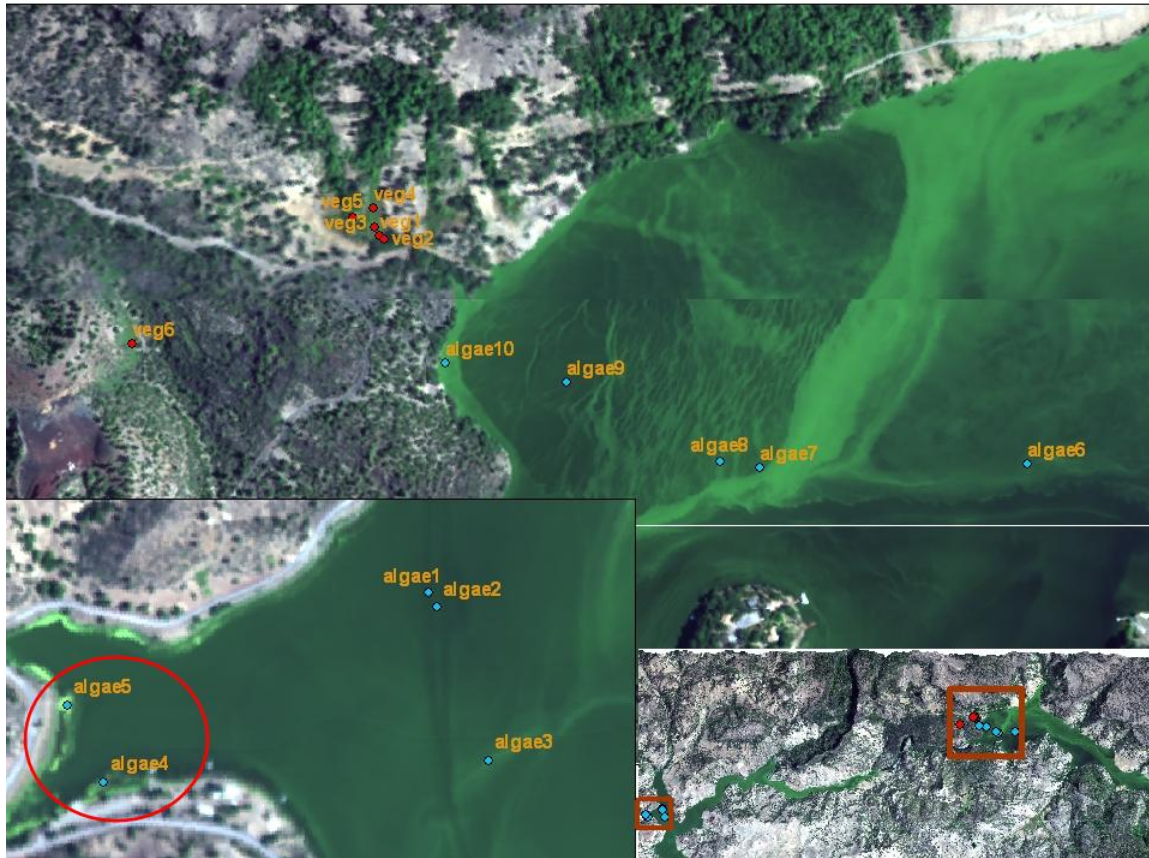


Figure 3-5: Sample Spectra Pixel Locations



The pronounced change in reflectance in the 680-800 nm is called the red edge, and is used in remote sensing of terrestrial environments to map vegetation as it emanates from the re-emission of light from chlorophyll-a (Horler et al. 1983; Jensen 2007). Even just a small amount of water overlying the algae or any vegetation will begin to diminish (at 2 cm) or eliminate (at 2 m) this “red edge”. The spectra for algae on the water surface versus those dispersed in the water column do show such an attenuation in the red edge differential. However, this characteristic appears to be less effective in algae/vegetation discrimination than that of the 631 absorbance trough given that the spectra for algae 10 and algae 6 in Figure 3-4 both show large “red edge” features (e.g. reflectance peaks at 746 nm that far out reach the reflectance peak at 555 by a factor of 2 or 3) comparable to those observed in vegetation spectra. This characteristic may reflect that much of the algae is largely concentrated on the water surface, and therefore its chlorophyll-a re-emits much like terrestrial vegetation.

The water algae and the terrestrial vegetation are visibly different in the near infrared range of the spectrum greater than 775 nm. Both contain reflectance peaks at 785 nm and 843 nm, and an absorbance trough at 824 nm. However, the ratio of the height of the 843 peak (measured from the 824 nm trough) to the height of the 785 nm peak (measured from the 824 nm trough) is larger and closer to one for the terrestrial vegetation than for the water algae. Figure 3-6 shows a sample spectrum and how this ratio (referred to from here on as the NIR ratio) is calculated.

These are just some of the spectral features important in defining and quantifying the presence of freshwater cyanobacteria and the rich spectral dataset offers many potential approaches to this question. Based on these observations of the spectral signatures the following spectral bands were selected for inclusion in the statistical analyses: 435, 481, 499, 555, 631, 650, 679, 746, 785, 824, 843. These eleven wavelengths will be referred to as key wavelengths subsequently.

A simplistic 2-dimensional scatter plot of the red wavelength 650.59 nm versus the blue wavelength 481.60 nm (see Figure 3-7) can be used to distinguish between the spectral features in the image. Figure 3-8 shows patches and swaths within the scatter plot (on the left) that have been highlighted to create different color classes whose corresponding pixels are highlighted to match in a crop of the SpectIR image (on the right). Figures 3-9 and 3-10 show the same thing using the blue wavelength 481.60 nm versus the green wavelength 555.38 nm. This technique is highly subjective because as many or as few classes as desired can be created. Both images however show good distinction between chlorophyll and other algal pigments located over the water surface and the chlorophyll found in land vegetation. The unclassified cropped image is shown in Figure 3-11.

Figure3-6: Example Spectra Showing Measurements used to Calculate NIR Ratio

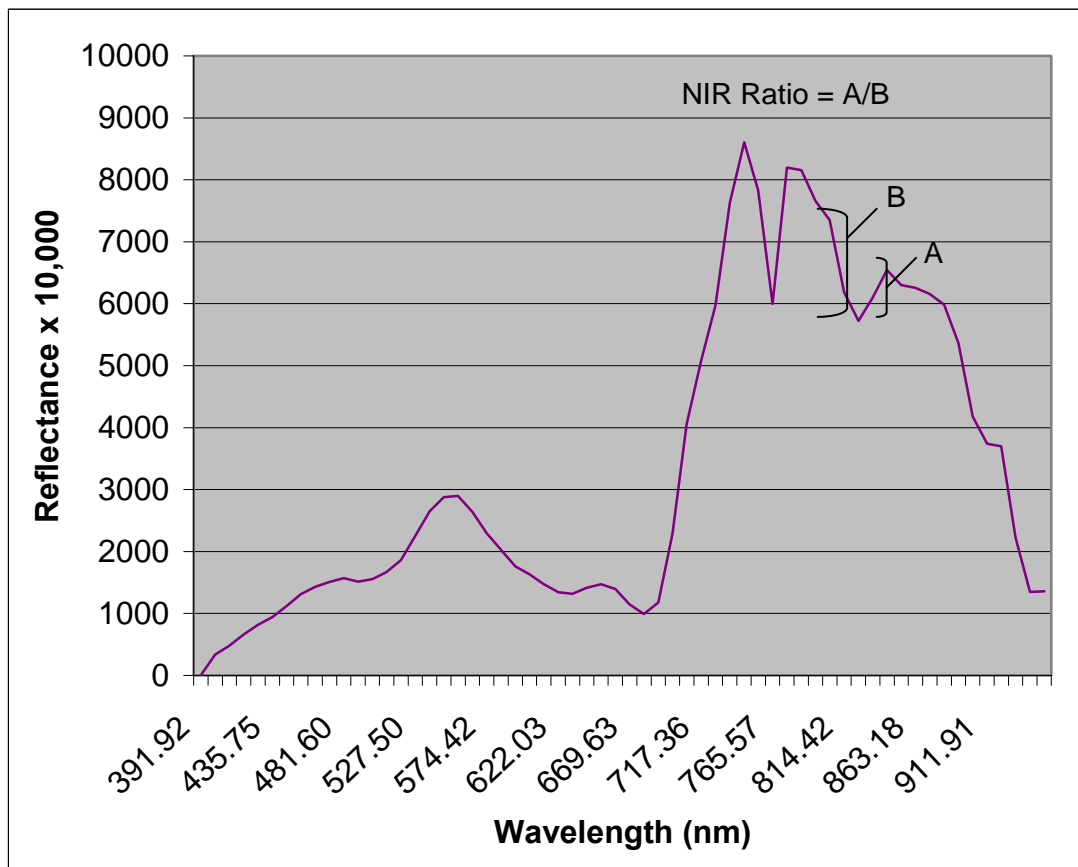


Figure 3-7: 2-Dimensional Scatter plot of red wavelength 650.59 nm versus the blue wavelength 481.60 nm

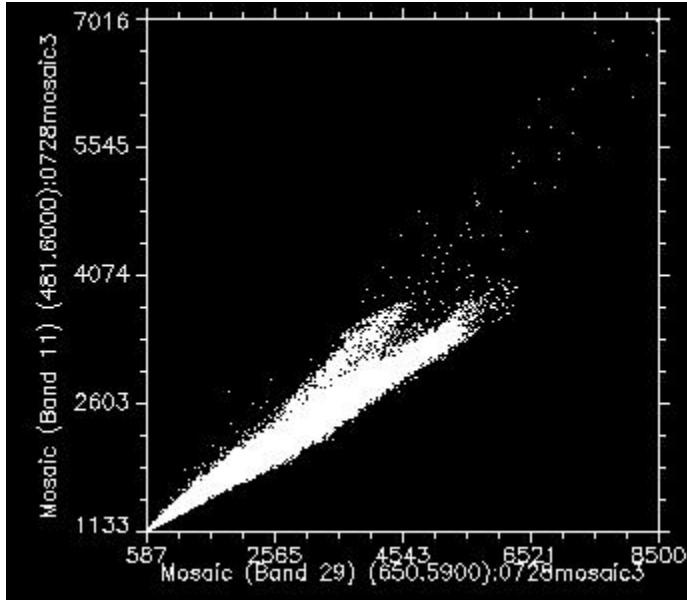


Figure 3-8: Classes created in Red versus Blue 2D scatter plot to distinguish spectral features (Left) and Classes shown in SpecTIR image as created in 2D scatter plot (right)

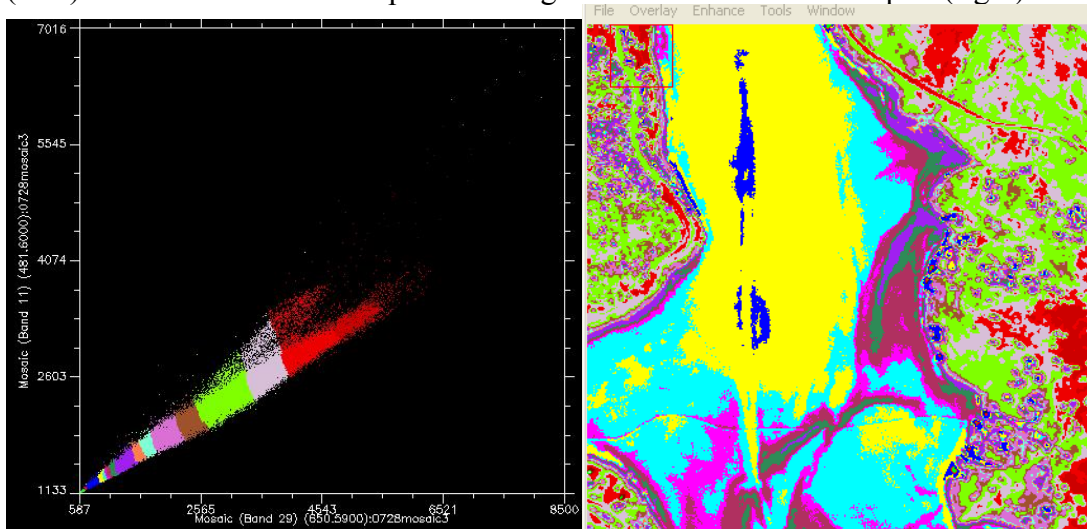


Figure 3-9: 2-Dimensional Scatter plot of blue wavelength 481.60 nm versus the green wavelength 555.38 nm

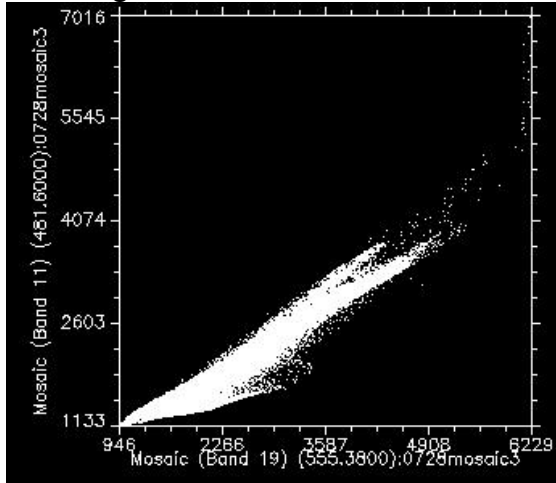


Figure 3-10: Classes created in Blue versus Green 2D scatter plot to distinguish spectral features (Left) and Classes shown in SpecTIR image as created in 2D scatter plot (right)

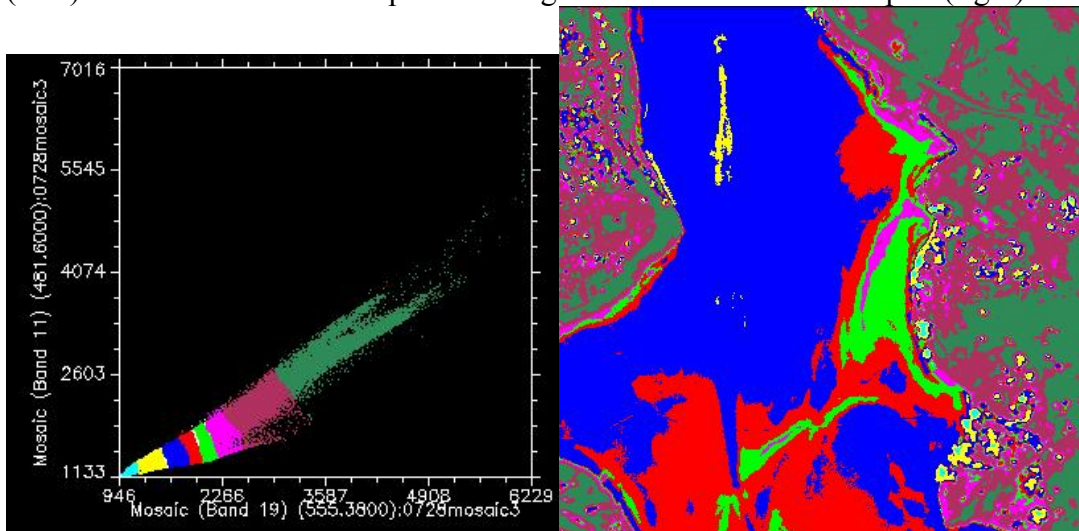
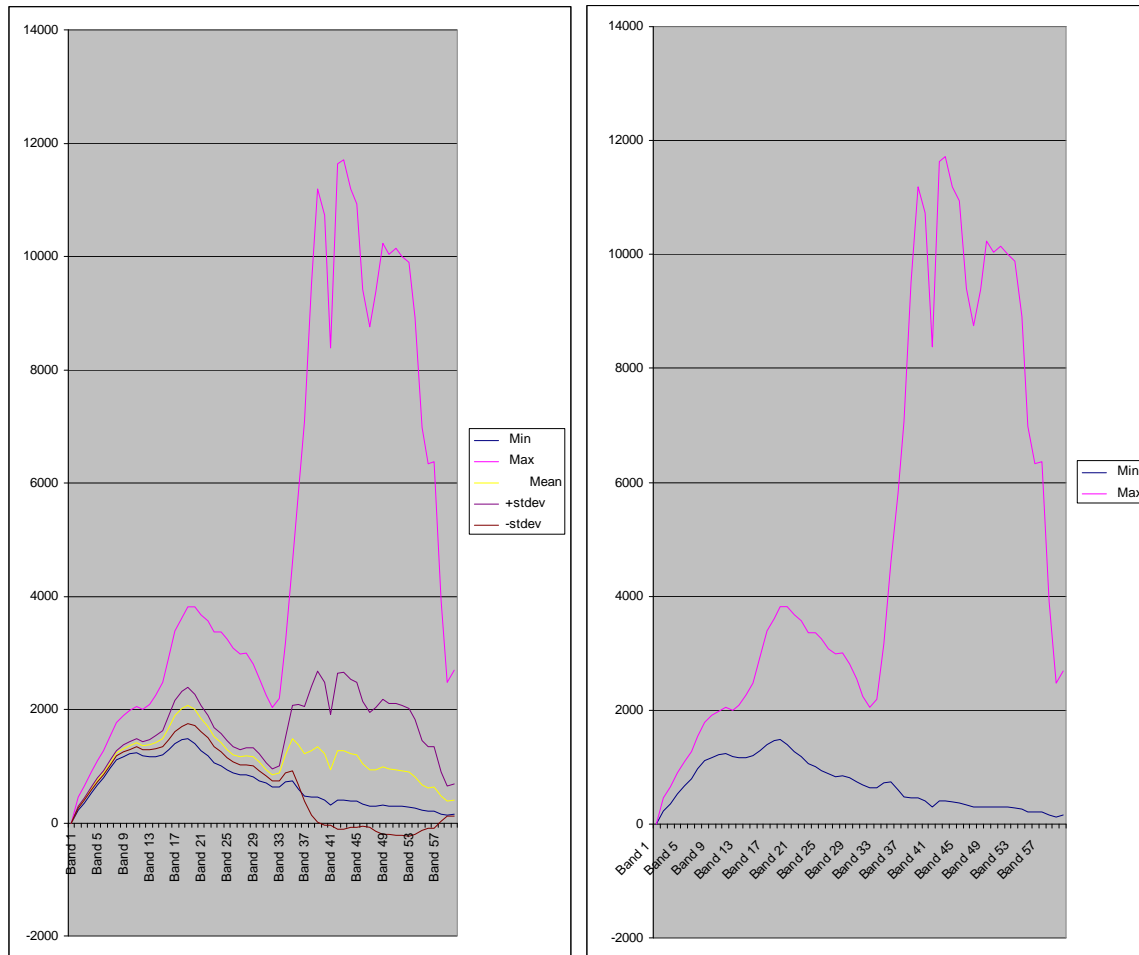


Figure 3-12 shows the same 2-dimensional scatter plot of the blue versus green wavelengths, but this time a single class has been created that mostly encompasses nothing but the water. There is some misclassification, as small patches over the land have also been highlighted. Maximum, minimum, mean and standard deviation of the spectral curves for the pixels encompassed within this single class are shown in Figure 3-13. The maximum and minimum can be roughly interpreted as representative of the range in algal pigment concentrations present in the field: 8-12,805 $\mu\text{g/L}$ chlorophyll-a (see Figure 2-6) and 2-164 $\mu\text{g/L}$ phycocyanin .

Figure 3-13: Blue v. Green 2D Water Class Statistics (Max, Min, Mean, Std. Dev.) and just Max and min to the right.



The spectral number reflectance data (see Section 3.2.2: Image Pre-Processing , p. 3-19) for the pixels corresponding to the locations of five of the water sampling sites were extracted. These five sites had differed greatly in their chlorophyll-a concentrations (See Table 3-1). They were normalized relative to the largest value and their spectral curves are shown in Figure 3-14. As expected the spectra span the range captured in the image and depicted in Figure 3-13, however the order of the spectra does not align exactly with the magnitude of the chlorophyll-a concentration. The highest concentration is in fact the spectra with the largest peaks, but the lowest concentration has the second largest peaks. The other three concentrations fall in the expected order after that. This may reflect the spatial and temporal variability described in Chapter 2. If the water samples could have all been collected at exactly the same instant in time as the imagery was acquired, we would expect that the pigment concentrations would be an effective way to train and/or evaluate the remote sensing data. The same limitations of ground data collection that weaken its explanatory power in toxic cyanobacteria surveillance and compel the need for synoptic techniques also hamper the comparisons between the two types of data.

Table 3-1: Chlorophyll Concentrations for Selected Sites

Site	[Chl-a] (ug/L)	Normalized to Site 036a
010a	8.6	0.001
015a	20.6	0.002
030a	365.6	0.029
004c	3378.9	0.264
036a	12805.3	1.000

In Figure 3-15 the two satellite images of the reservoirs acquired within 24 hours of each other show the differences in the pattern of bloom dispersion just within 1 day due to the vertical and horizontal variability discussed in Section 3.3. This means that the same point at the exact same geographic location can contain varied levels of cyanobacteria if sampled in the morning versus the evening or on two consecutive days as was necessary in order to complete the ground data collection for these reservoirs.

The wavelengths 650, 481, and 555 were chosen for the 2-dimensional scatterplots based on known properties of chlorophyll-a, phycocyanin and carotenoids (Richardson 1996). However suspended sediments, dissolved organic matter, and humic acids as well as atmospheric interference can all combine to complicate the reflectance spectra (as compared to those of pure substances) and therefore their interpretation. The methods of classification described below were used to compensate for these sources of error, noise, and interference and to determine which components of the spectra predict the unique pigment combinations of chlorophyll-a and phycocyanin characteristic of blue-green algae and thereby distinguish them from the spectra of terrestrial vegetation. This study analyzed the image with four different analytical/classification methods: unsupervised classification, linear regression, linear discriminate analysis, and tiered decision tree classification. Each is described below together with the results of the analysis.

3.2 Methods and Results

3.2.1 Image Acquisition

I scheduled an airplane pass-by for July 27, 2007 to acquire the 60 band (each 9.2 nm in width from 391-961 nm) 2-meter resolution image from SpecTIR using a visible and near infrared sensor with a progressive scan CCD (charged couple device) camera. Technical difficulties prevented acquisition on the 27th and the image was collected the following day on July 28 between 9:27 and 10:11 a.m. The image was collected in four flight-lines that were later assembled with the mosaic tools available using ENVI software. The sensor was flown aboard a Cessna 206 at an altitude of approximately 4,000 m above ground level, making the swath width 2720 m, and the survey was completed within an hour time period.

3.2.2 Image Pre-processing

Reflectance and radiance are often used interchangeably in general discussions, but radiance is the true uncorrected light energy reflected by an object or element. It is reported with the units of watts/steradian/square meter and its magnitude is dependent upon the intensity of the light source, the incident angle or direction from which the light source emanates, the object's orientation, and any interference in the light's path through the atmosphere. Therefore the raw radiance numbers are not interpretable until they have been corrected for these influences. Atmospheric correction calibrates the imagery by converting the at-sensor radiance data to unit-less reflectance values that are ratios of the amount of light leaving the target to the amount of

Figure 3-14: Relative Chlorophyll Differences and Spectra

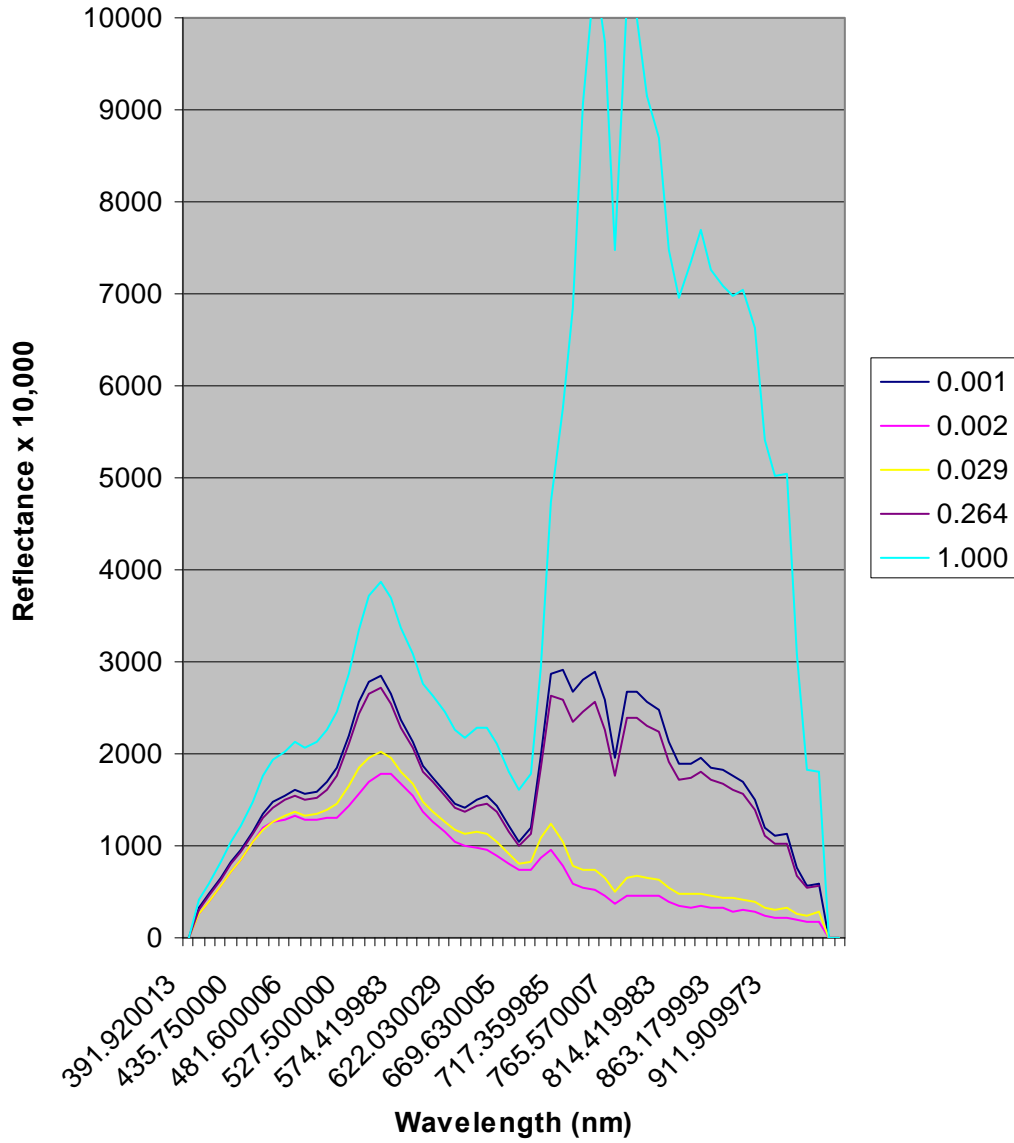
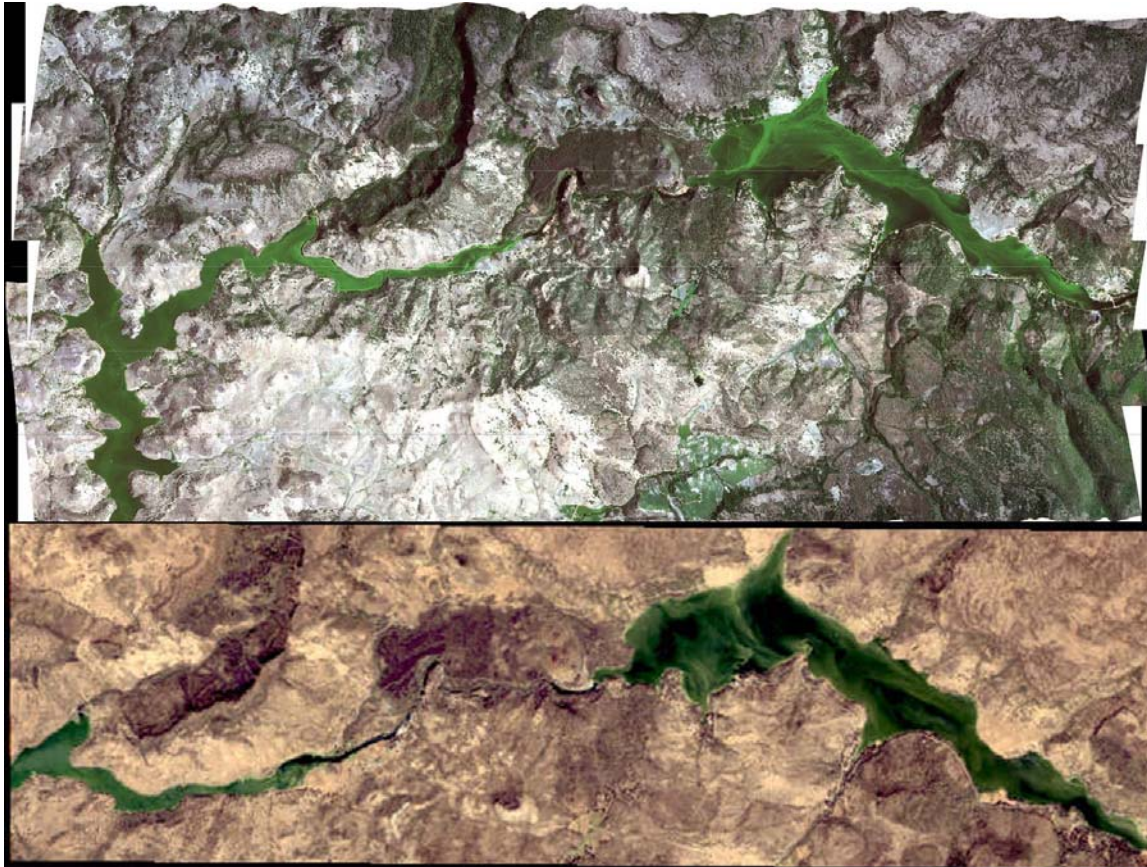


Figure 3-15 Satellite Images on Two Consecutive Days (The top image shows SpecTIR – collected 7/28/07 09:27-10:11 Pacific Time, 2 m resolution. The bottom image shows IKONOS – collected 7/27/07 12:09 Pacific Time, 1 m resolution.)



light striking the target. In any remotely sensed scene, the atmosphere can vary widely in terms of suspended aerosols which both attenuate the light passing through and scatter it. This in turn can result in inaccurate calculations of land use classes due to the altered light detected by the instrument sensor.

SpecTIR data are provided in two separate files. In one the data are in radiance units scaled by a factor of 1000 such that a pixel value 4500 designates an actual radiance value of 4.500 $\text{mW/cm}^2/\text{steradian/m}$. In the other the data are converted to reflectance scaled by a factor of 10,000 so a value of 3000 is actually 0.3000. The reflectance data is corrected using a SpecTIR proprietary program based on a Savitsky-Golay algorithm which handles adjustments of atmospheric absorption features associated with CO_2 and water. The proprietary program implements the industry standard MODTRAN4 radiative transfer code. Despite the many assumptions that are used in these atmospheric corrections as compared to the purer calibrated radiance data, and although Song et. al. (Song et al. 2001) concluded that atmospheric correction does not necessarily improve the classification and is unneeded in many remote sensing applications as long as the training and calibration data come from the same time and place, the

reflectance data corrected for atmospheric interference were used in the subsequent analyses following standard practice in remote sensing science.

The data are georeferenced with either a 3-ring laser gyro-based Inertial Navigation System (INS) or a Fiber Optic Gyro/Mems-based system. SpecTIR has the entire 10 meter resolution NED digital elevation model (DEM) database for the continental United States which is used to provide orthorectified hyperspectral imagery. The signal-to-noise ratio is 500:1 to 1,000:1 in regions away from the water absorption bands near 900 nm. The spatial accuracy of the data is +/- 8 meters (2 pixels) after orthorectification using a USGS digital elevation model.

3.2.3 Unsupervised Classification

Unsupervised classifications were used as the first pass to quickly look at how easily chlorophyll and other pigments of water phytoplankton could be distinguished from those contained in land-based vegetation. K-means is just one of several ENVI incorporated tools for unsupervised classification. It is a clustering algorithm using Euclidean distance as a metric and doesn't generate any kind of model to interpret the contribution of individual bandwidths. Another disadvantage is that it requires that the user input the number of clusters (K) and poor results may be from an inappropriate choice for this parameter. Data ($X^{(n)}$, where n runs from 1 to the number of data points N – in this case the number of pixels in the image) are grouped by minimizing the sum of squares distances (d) between data and the corresponding data centroid. Each vector x has i components (in this case, the 60 bands in SpecTIR), x_i .

$$d(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

The means for the K clusters ($m_1 \dots m_k$) may be specified randomly or by some speculative formulation. The process is iterative by which each data point n is assigned to the cluster with the nearest mean. The cluster means are then recalculated based on the average of all the data points assigned to that cluster. The process repeats until the assignments no longer change (MacKay 2003).

Figure 3-16 below shows a K-means unsupervised classification using 20 clusters. The algorithm is able to separate the water relatively well however the variations in bloom density within the water are not well defined and there is considerable misclassification in large swaths of the water (see large green and purple patches of non-algae classes within the boundaries of the water algae class shown in red in Figure 3-16). Linear regression was applied next to begin to identify the spectral combination and model that could best define the algal blooms.

3.2.4 Linear Regression

Linear regression was used to attempt to develop a model predicting pigment concentrations or cell counts from spectral reflectance data identifying the bandwidths that are most relevant. The chlorophyll-a pigment concentrations (a proxy for total cyanobacteria cell counts) from ground truth field data (see Chapter 2) were used as the dependant variable with the reflectance data from the pixels corresponding to the GPS locations where the field samples were collected used as the independent variables. Despite the large spatial mobility (see section 2.5) over relatively short periods of time (hours), the imagery's pixel reflectance was only moderately correlated (R2

< 0.7) with the actual field measurements of phytoplankton chlorophyll. The reflectance values at the key wavelengths (all eleven are listed in for model 1 in Table 3-2) were included in the first model and removed in a process of stepwise regression.

Figure 3-16

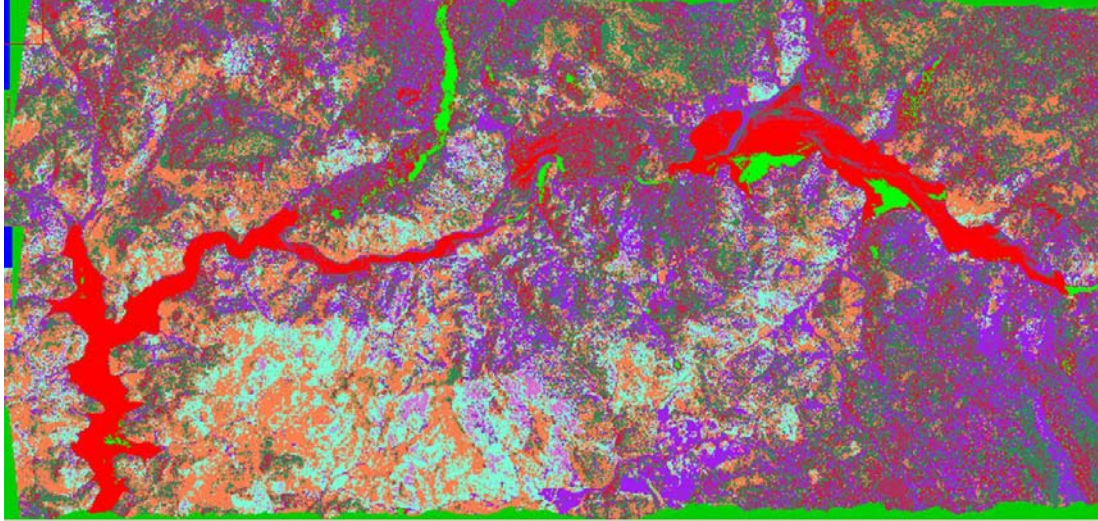


Table 3-2 shows the different regression models, their R^2 , and the number of misclassified pixels in each category. The best fit model included the wavelengths 499, 555, 650, 746, 785, 824, and 843.

The poor relationship between the chlorophyll concentrations and the reflectance peak in Band650 (related to chlorophyll) is shown in Figure 3-17. The field data, however, do provide a guide to the range in pigment concentrations/cell densities present in the field and this range can be linked with the range of relevant spectral data (e.g. wavelength reflectance and/or values calculated from them such as ratios, derivatives, etc.) The linear regression indicates that wavelengths 435, 481, 631, and 679 are less important in predicting presence of cyanobacteria and confirms the usefulness of the others in this task. Since the dependent variable in linear regression must be a continuous/measured variable, linear discriminant analysis was next used to evaluate land use categories as the dependant variable instead of the field sample pigment concentrations which were impacted by temporal variability. Using pixels categorized by land use in linear discriminant analysis also allowed a much larger sample size, since only 62 field samples could be collected and analyzed for chlorophyll-a for this study with the physical constraints involved in collecting them within twelve hours of the imagery acquisition.

3.2.5 Linear Discriminant Analysis

Linear discriminant analysis was used to find a model that identified the bandwidths that contributed the most towards separate variations in density of water phytoplankton and towards differentiating them from other land features (e.g. soil, foliage, roads, grasslands, etc.).

Linear discriminant analysis (LDA), like principle component analysis, classifies data and reduces dimensionality, but LDA uses data classification instead of feature classification. Unlike

PCA, the shape and location of the original data sets is not changed. The separation of any two data points is maximized in LDA by maximizing the ratio of between-class variance to the within-class variance (Balakrishnama and Ganapathiraju). The classification problem has several objects with a set of features (e.g. explanatory variables) measured from those objects and the objects can be put into several different groups based on the measurements of those features. In this case the objects are the individual pixels of the image that I have grouped into land use classes (two classes of algae, three classes of soil, and five classes of vegetation) based on visual assessment, and their features are the reflectance at each of the different band wavelengths. The method then looks for the set of features that can best determine group membership of the object and the classification rule or model to best separate those groups.

Using a LDA model assumes that the groups are linearly separable, that the population covariance matrices are equal for each group, that each group is drawn from a population with a multivariate normal distribution, and that no variable is a linear combination of any other variable (Klecka 1980).

To generate a large sample size of pixels grouped into land use classes, I made use of the unsupervised classification to generate random samples within what could be separated into algae, soil and vegetation groups with my visual assessment as a trained user familiar with the site. The log of the eleven key wavelengths described earlier (in section 3.1.1. and then used in the linear regression) were selected for inclusion in the linear discriminant analysis. The F-tests associated with each dimension are exact and Table 3-3 shows that only 7 of the 9 dimensions created with the discriminant analysis are statistically significant ($p < 0.001$). Dimension 1 alone explains 77.76% of the separability between the groups, and only the first three dimensions are needed to explain 98.84% of the separability. In Table 3-4 the standardized canonical coefficients are summarized for all nine dimensions. The grey highlights the coefficients which make the largest contributions to dimension 1, 2 and 3. For dimension 1 this contribution comes from bands 631, 650 and 679. Absorbance troughs from phycocyanin are seen in the remote sensing spectra at 631 and 679 (see Figure 3-3) and the reflectance peak between them is at 650. These three bands are therefore particularly important in defining the part of the spectra that is uniquely different for the pixels containing cyanobacteria. Bands 746 and 785 make the largest contributions to Dimension 2 while bands 481 (where carotenoids often absorb) and 679 dominate in Dimension 3. Dimensions 8 and 9 are not significant, and this may be because of the tight correlation between the independent variables. Table 3-5 shows the correlation matrix between the natural logs of the eleven key wavelengths showing that all the different bands used as the independent variables have high correlations with one another. All of the correlations are greater than 0.66 and half of them are greater than 0.95.

The accuracy of this classification by LDA was tested using the leave-one-out cross-validation technique which classifies the same observations that were used to estimate the discriminant model holding each data point out separately and then re-generating the model on the remaining observations and then classifying the one that was held out. All the data points are therefore used once as the validation data, and the misclassifications are summed over all the iterations and summarized in Table 3-6. The grey shading highlights where algae were misclassified as vegetation or visa versa. Only 16 pixels of the 2645 were misclassified between algae and vegetation ($< 1\%$ error rate). The groups include several different types of vegetation or several

Figure 3-17: Chlorophyll-a Concentration versus Reflectance at 650 nm

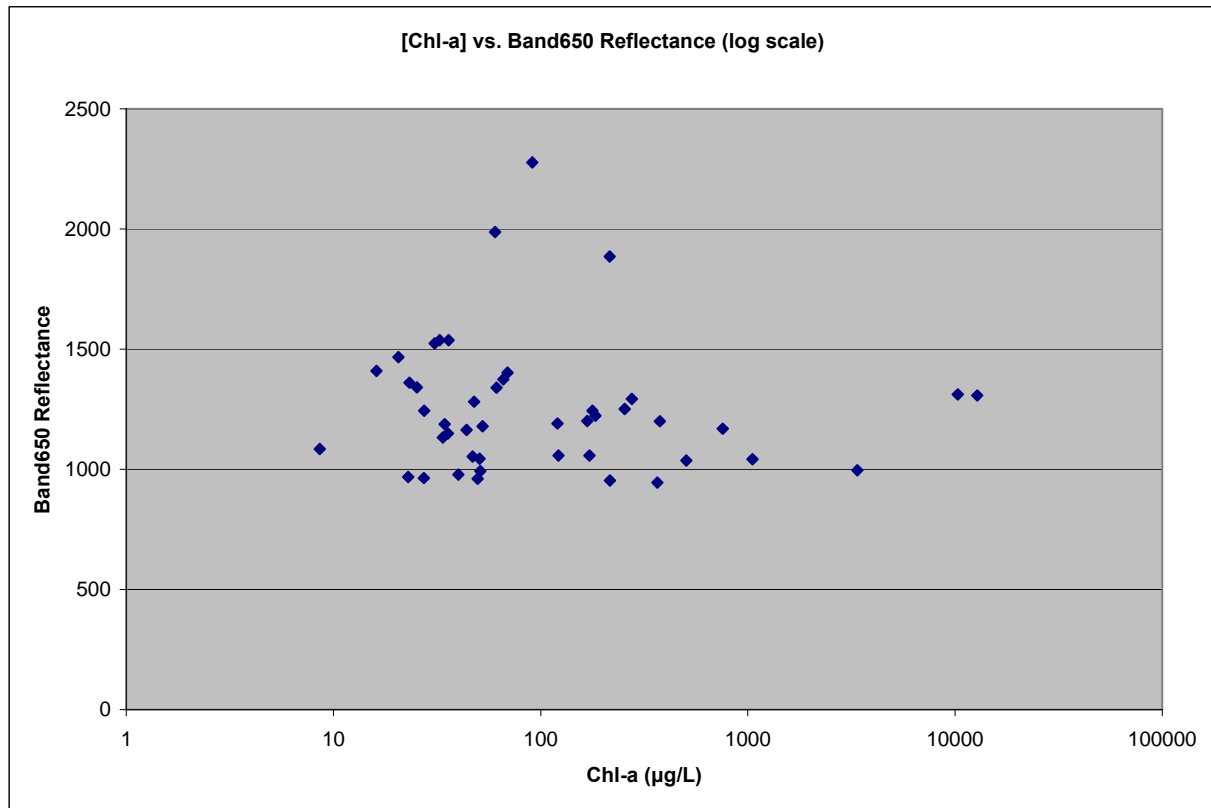


Table 3-2: Linear Regression Model Results

	Dependant Variable	Independent Variables (reflectance at wavelengths listed below)											R2	Adj. R2
Model 1	Continuous ([Chl-a])	435	481	499	555	631	650	679	746	785	824	843	0.7454	0.6606
Model 2	Continuous ([Chl-a])		481	499	555	631	650	679	746	785	824	843	0.7454	0.6705
Model 3	Continuous ([Chl-a])			499	555	631	650	679	746	785	824	843	0.7392	0.6721
Model 4	Continuous ([Chl-a])			499	555		650	679	746	785	824	843	0.7346	0.6757
Model 5	Continuous ([Chl-a])			499	555		650		746	785	824	843	0.7305	0.6795
Model 6	Continuous ([Chl-a])			499			650		746	785	824	843	0.7074	0.6613

Table 3-3: Tests of Discriminant Dimensions: canonical linear discriminant analysis

Dimension	Corr.	Eigen value	Cumulative % of Variance	F	df1	df2	p
1	0.9885	12.6284	77.76	378.01	99	1.90E+04	0.0000
2	0.9510	9.4680	95.03	184.27	80	1.70E+04	0.0000
3	0.8224	2.08896	98.84	81.154	63	1.50E+04	0.0000
4	0.5697	0.48046	99.71	31.455	48	1.30E+04	0.0000
5	0.3135	0.10903	99.91	11.541	35	1.10E+04	0.0000
6	0.1760	0.03198	99.97	5.229	24	9.18E+03	0.0000
7	0.1173	0.01395	100.00	2.7981	15	7.26E+03	0.0002
8	0.0404	0.00164	100.00	0.67458	8	5.26E+03	0.7144
9	0.0203	0.00041	100.00	0.36279	3	2.63E+03	0.7799

Table 3-4: Standardized Discriminant Coefficients

Explanatory Variable (Log of the bandwidth)	Dimension								
	1	2	3	4	5	6	7	8	9
log435	0.639738	-0.3751106	-1.46353	-1.09397	-0.25088	2.28277	1.264451	2.96957	-3.84975
log481	-2.16046	1.700813	3.440948	2.081637	-5.65542	-9.0219	-3.06471	-22.8706	1.679951
log499	0.535206	-0.3679338	1.271371	1.288637	5.731963	7.142347	0.525136	23.34991	3.126471
log555	1.512576	-3.051311	-2.06224	1.64601	-0.99722	1.675152	4.45066	-2.70049	1.721658
log631	-7.63226	2.08656	-0.07736	-0.12113	8.733434	-6.92124	-7.04349	-4.05749	-3.25214
log650	10.86262	0.0113087	2.198957	-7.79437	-7.23247	1.682575	-8.52325	5.61328	-3.15134
log679	-3.89028	-0.9434206	-2.91619	3.519474	-0.93575	3.23302	12.87357	-2.67426	3.567978
log746	-1.90726	3.088629	0.340806	1.514177	9.110604	2.256136	1.040161	5.102164	-1.81945
log785	0.848163	-3.393951	-0.59392	-2.98706	-10.6795	-11.3805	2.937463	-1.95442	-1.12591
log824	0.740849	-0.1701463	1.249756	2.723467	2.70608	6.481963	-6.7003	-1.17803	0.782855
log843	-0.78312	0.6268713	-1.00586	-1.03144	-1.2002	2.311449	1.977214	-1.27898	2.001622

Table 3-5: Correlation Matrix between Log Wavelengths

	log631	log435	log481	log499	log555	log650	log679	log746	log785	log824	log843
log631	1										
log435	0.973	1									
log481	0.9852	0.9945	1								
log499	0.9884	0.9917	0.9995	1							
log555	0.9551	0.9313	0.949	0.955	1						
log650	0.9988	0.9704	0.985	0.9887	0.9609	1					
log679	0.9943	0.9651	0.9755	0.9778	0.9201	0.9913	1				
log746	0.7604	0.6941	0.6911	0.6949	0.6971	0.7443	0.7778	1			
log785	0.7375	0.67	0.6659	0.6695	0.6675	0.7204	0.7581	0.999	1		
log824	0.739	0.672	0.6671	0.6705	0.6632	0.7214	0.7613	0.9979	0.9994	1	
log843	0.7311	0.6658	0.6595	0.6621	0.641	0.712	0.7581	0.994	0.9969	0.9984	1

Table 3-6: Leave One Out Classification Summary

True Group	Classified										
	algae1	algae2	soil	soil2	soil3	veg2	veg3	vegdark	vegshadow	vegsoil	Total
algae1	326	13	0	0	0	0	0	0	2	0	341
algae2	4	68	0	0	0	0	0	6	0	0	78
soil	0	0	317	55	2	0	0	0	0	26	400
soil2	0	0	19	150	30	1	0	0	0	0	200
soil3	0	0	0	15	84	1	0	0	0	0	100
veg2	0	0	12	5	0	258	2	0	0	19	296
veg3	0	0	0	0	0	1	245	0	23	29	298
vegdark	0	8	0	0	0	0	3	178	33	0	222
vegshadow	0	0	0	0	0	0	3	11	245	0	259
vegsoil	0	0	47	0	0	49	25	0	0	330	451
Total	330	89	395	225	116	310	278	195	303	404	2,645

Table 3-7: Accuracy of SpecTIR Classification Decision Tree

	Low	MedLow	MedMed	MedHigh	High	Scum	Land
Commission Error	1.92%	0.38%	2.51%	3.71%	21.54%	27.66%	0.07 %
Area equivalent of pixel error	~64,164 m ²	~13,296 m ²	~112,684 m ²	~90,480 m ²	~287,720 m ²	~250,976 m ²	
User Accuracy	98.08%	99.62%	97.49%	96.29%	78.46%	72.34%	99.93 %

different types of soil and misclassification between those groups is not pertinent to the purpose of this model.

Although the LDA model generated in Stata 9.0 could not be applied to predict other pixels within the spatial framework of ENVI or ERDAS, it serves to further the understanding of the relationship between the different spectral bandwidths and the pigments in phytoplankton. This approach with LDA solved the issue of spatial and temporal inaccuracies in the ground data and gave helpful information about the relevant bandwidths to be included in the model.

Unfortunately it requires that I (or some trained analyst) do a first pass on the imagery to define the categories and select the input pixels. Once generated the ENVI software had no way apply the model to the whole dataset, which is so large with this imagery, and therefore could not be used to classify the whole image, but only an exportable random selection from within the categories. Decision trees were used next as a tool that allows one to incorporate existing knowledge of the different variables, including that generated by the previous analyses.

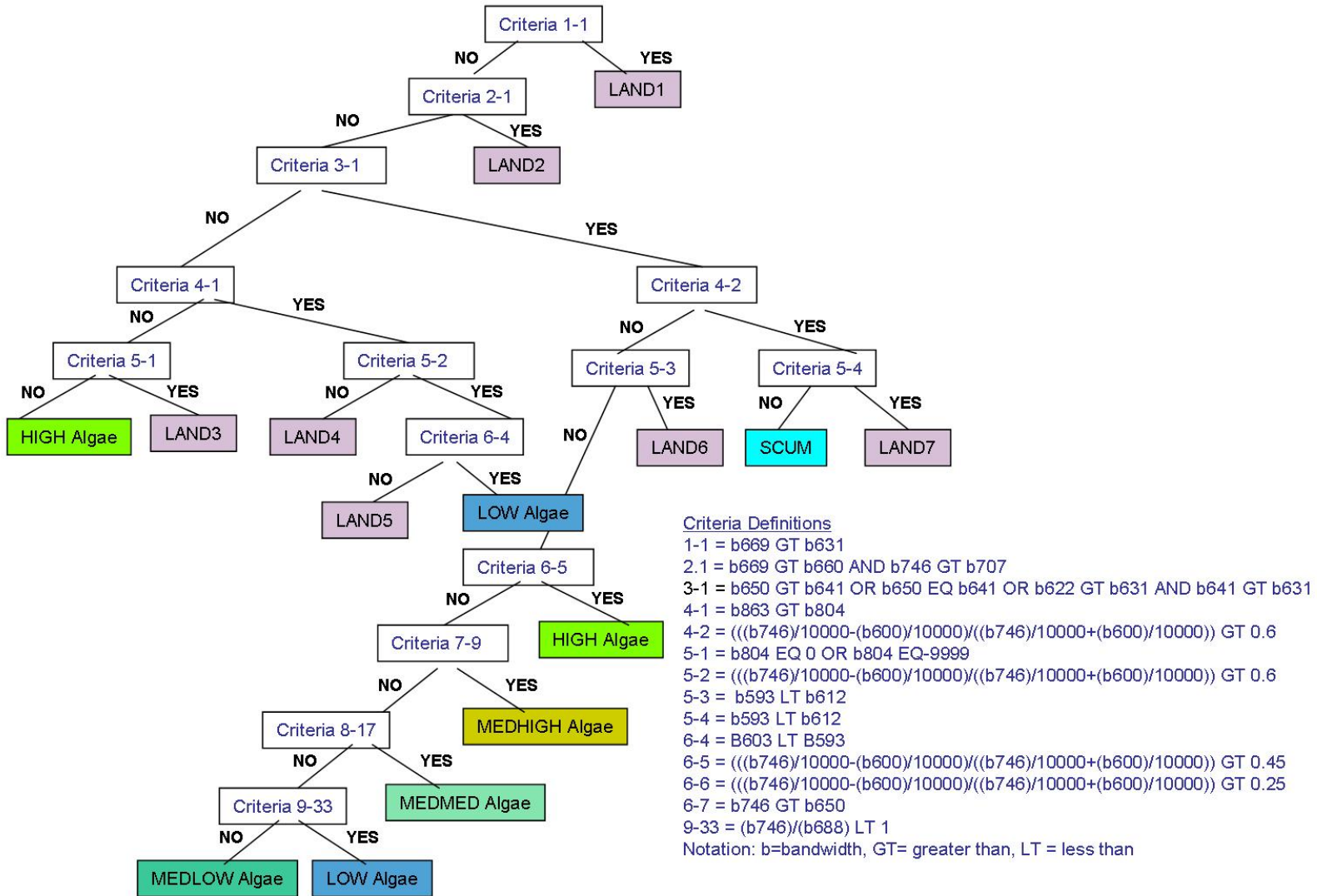
3.2.6 *Decision Trees*

Differentiation between the levels of cyanobacteria bloom density and between all cyanobacteria and vegetation and other land use categories is possible using a tiered decision tree which specifies criteria within the spectra that allow the creation of mutually exclusive classes.

Many others have successfully used classification and regression tree approaches to classify crop type (Yang et al. 2003), land cover (Hansen et al. 1996; Friedl and Brodley 1997), and wetlands (Wright and Gallant 2007) in remote sensing imagery. Decision trees have several advantages over neural networks or other approaches (Hansen et al. 1996; Friedl and Brodley 1997; Yang et al. 2003). They can efficiently handle non-parametric data, nonlinear relationships, missing data, numerical data, categorical data, and non-normal, non-homogeneous data. Decision tree classifiers need not assume the input data displays the central tendency or has a particular distribution. Finally they are computationally efficient and conceptually simplistic such that the rules can be easily interpreted and the relative importance of different variables becomes apparent. The hierarchical nature of the predictor variables makes the interpretation explicit and fairly intuitive to identify those that most contribute to the discrimination of any particular class. Decision tree algorithms exceeded the performance of maximum likelihood and linear discriminant function classifiers (Pal and Mather 2001) in a land use classification of Landsat-7 ETM+ data for an agricultural area. Classification tree theory is described in detail in Breiman et al. (1984)

Fourteen criteria were used in the tiered decision tree. Criteria were selected based on knowledge of the spectral characteristics of the algal pigments, water, vegetation and land which were described in section 3.1.1. This understanding was furthered by the results of the linear regression and the LDA. The spectral characteristics of the output classes were evaluated after each criteria was added to inform the definition of the subsequent criteria.. Based on the known areas of land with vegetation versus water impacted by cyanobacteria, spectra for misclassified and correctly classified pixels were examined in order to select the best next criteria. The process was repeated until visual inspection of the final classified map continually showed no improvement in class area refinement. There may be better decision trees for the data or multiple trees with comparable performance, but the tree captures the spectral features observed by a

Figure 3-18: Tiered Decision Tree for SpecTIR



manual rather than computerized selection and it establishes the feasibility of classifying these data for cyanobacterial bloom surveillance in this way. Figure 3-18 below shows the final decision tree used for the SpecTIR image. All the features on the land (e.g. the groups depicted as LAND 1-7 in Figure 3-18) were combined into one single class and six separate cyanobacteria/algae groups within the water were created corresponding to varying cyanobacteria cell densities: *cyanobacteria scum*, *high*, *medium-high*, *medium-medium*, *medium-low*, *low* and *water/shadow* (water without algae). The selection of each criterion is described below.

Notation: b= bandwidth, GT = greater than, LT = less than, EQ = equal

Criterion 1-1: b669 GT b631

This criterion separates out much of the soil, sand, clay and other ground surface pixels from the vegetation, water and algae. Sample spectra from the image of vegetation pixels, land/soil pixels, and algae pixels are shown in Figure 3-19. Figure 3-20 shows sample spectra for the 6 different categories of algae. Their spectra slope upward in the visible and the near infrared (NIR) range, while vegetation and algae have pigment absorption in the visible red around 669 nm (0.669 μm). Therefore the pixels that are true for this criterion are land.

Criterion 2-1: b669 GT b660 AND b746 GT b707

Some of the land/soil pixels are mixed with vegetation and slight absorption troughs still occur in the red wavelengths around 669 nm. This criteria pulls out a small group of these mixed pixels by shifting the criteria slightly to look for b669 greater than b660 (instead of b631) and includes the additional requirement that their be a steady increase in reflectance in the near infrared between b707 and b746. The latter does not occur in water, but does occur in water covered in algae. Clear water absorbs in the NIR, but the higher the algae density in the water the greater the reflectance in the NIR. The pixels that are true for this criteria are land (some mixed vegetation but no algae or water).

Criterion 3-1: b650 GT b641 OR b650 EQ b641 OR b622 GT b631 AND b641 GT b631

Algae have a peak at 650 nm formed by the trough at 631 nm from phycocyanin absorption and the trough around 669/679 nm from chlorophyll absorption. This criterion highlights the peak at 650 and therefore pulls out the vast majority of the algae for the pixels true for this criteria.

Criterion 4-2: $(b746-b600)/(b746+b600)$ GT 0.6

The pixels that were true for criteria 3-1 are mostly algae and some vegetation. The normalized difference vegetation index (NDVI) is used to separate them. Those pixels with an NDVI greater than 0.6 (e.g. true for this criteria) are mostly vegetation on land with the exception of the extremely thick algal scum that attains a $(\text{NIR}-R)/(\text{NIR}+R)$ reflectance ratio similar to vegetation.

Criterion 5-3: b593 LT b612

Figure 3-19: Sample SpectTIR Spectra

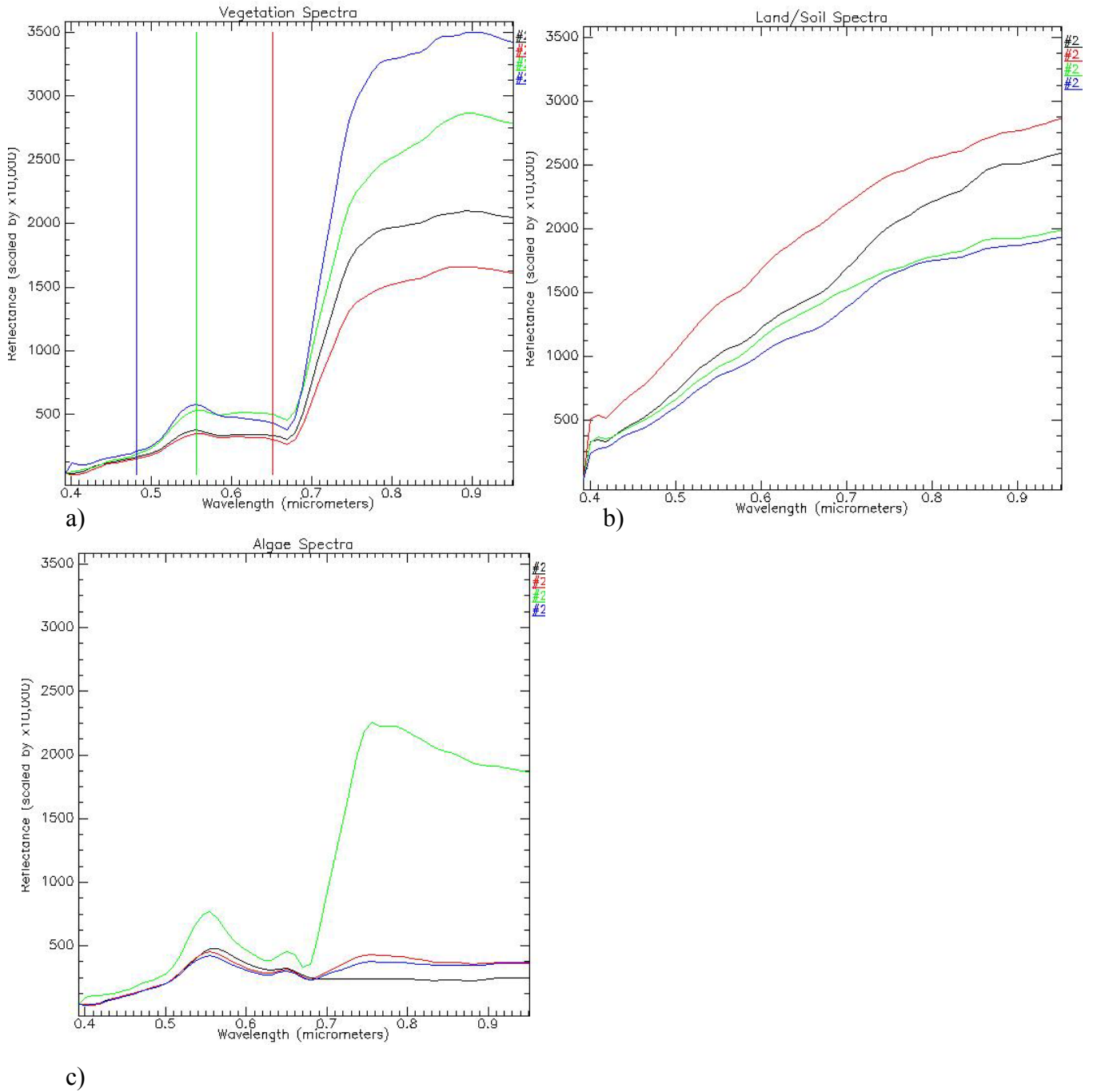
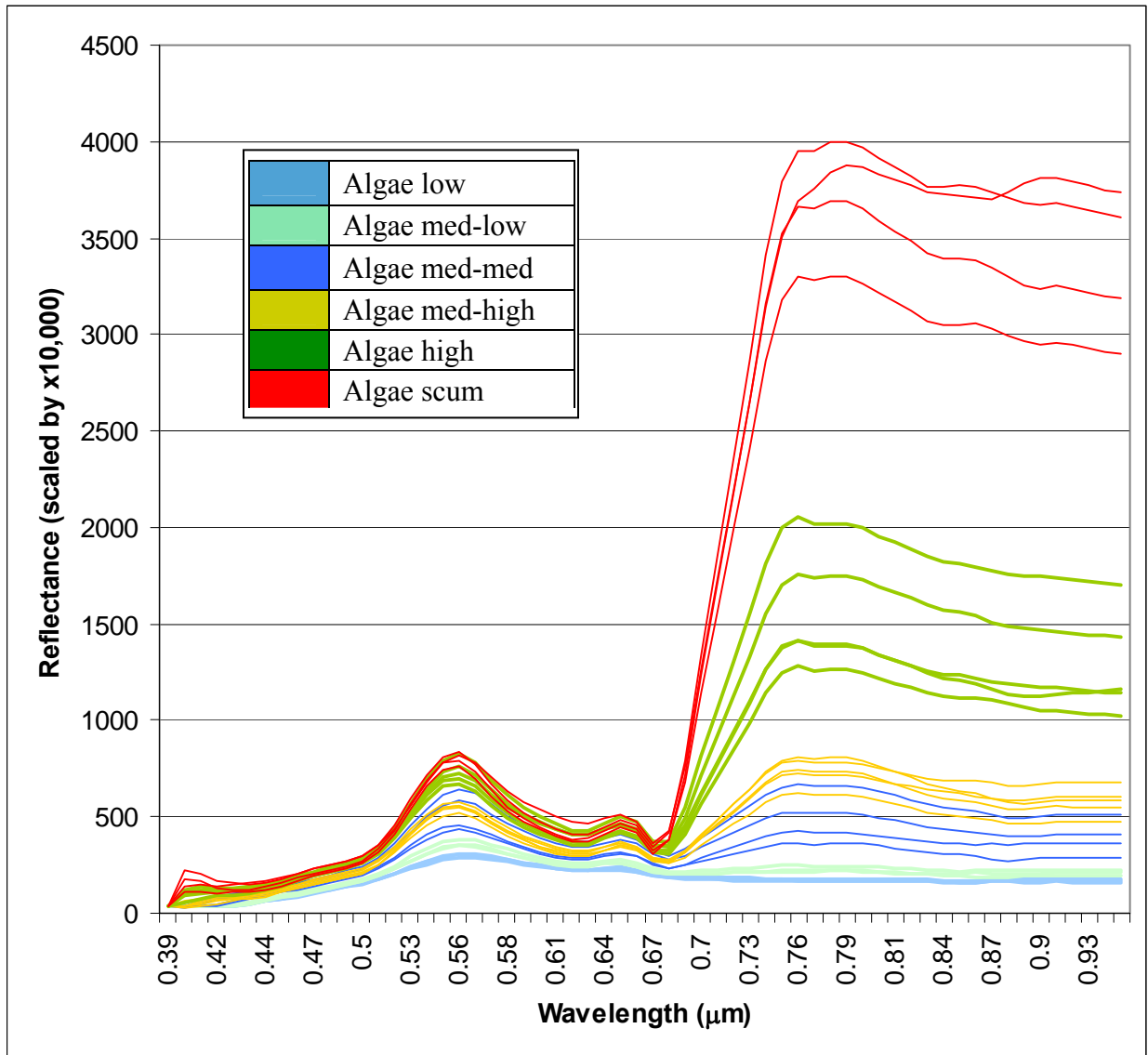


Figure 3-20: Spectra for Categories of Phytoplankton Bloom Density



Criterion 5-4: b593 LT b612

The same criterion is applied to both the true and false pixels from criteria 4-2.

For the pixels that were true for Criteria 4-2, this now separates the scum (false since b612 > b593) from vegetation. Because the algae scum contains phycoyanin which absorbs at 631 creating a trough and making b612 less than b593 (rather than the other way around) for the spectra to slope down into the trough, the scum are false for this criteria.

Criterion 6-5: (b746-b600)/(b746+b600) GT 0.45

All the false pixels for criteria 5-3 are algae. Criteria 6-5 separates the patches of algae with the highest density using the NDVI as in criteria 4-2. This time the criteria looks for pixels with an NDVI greater than 0.45. The pixels that are true are high density algal bloom.

Criterion 7-9: (b746-b600)/(b746+b600) GT 0.25

The next criteria separates out the next level of algal bloom density (Medium-High) using NDVI greater than 0.35.

Criterion 8-17: b746 GT b650

This separates out the next level of algal bloom density (medium-medium) using the fact that the higher densities have reflectance in the NIR (b746) that exceeds their reflectance in the red band (b650). The pixels true for this criterion are the medium-medium algal bloom density. Lower density blooms in the water (pixels false for this criterion) show more of the absorbance in the NIR characteristic of water.

Criterion 9-33: b746 / b688 LT 1

This separates the medium-low algae from the low algae based again on high reflectance in the NIR from those pixels with more algae. The reflectance in b746 in the NIR is greater than reflectance in 688 for the medium-low algae (is it is for all the other higher density algae classes, but for the lowest algae class, b688 is less than b746. The reflectance in the key absorbance troughs at b631 and b679 is incrementally less moving from high algae density to lower algae density, but the variability within each class limited the efficacy of straight values in any band as a criteria.

Criterion 4-1: b863 GT b804

The pixels that were false for criteria 3-1 were mostly land, but some water/algae pixels were also included. Criteria 4-1 uses the peak at the start of the NIR followed by a negative slope down as the wavelength increases. This absorbance in the higher range of the NIR may be a result of water mixed into the pixels, as vegetation does not display this feature. The pixels that are false for this criterion include high scum pixels that may have some aquatic vegetation mixed in as they didn't display the characteristic absorption troughs at 631 and 669/679 that were used to separate the most of the other algae pixels in Criteria 3-1.

Criterion 5-1: b804 EQ 0 or b804 EQ -9999

Also mixed into the pixels that are false for Criteria 4-1 are the empty space pixels from the edge of the image. These are easily removed with Criteria 5-1 as all values for all wavelengths in

those empty pixels are either zero or -9999. b804 was used in this criteria, but any wavelength could be used here without affecting the result.

Criterion 5-2: $(b746-b600)/(b746+b600)$ LT 0.6

Like Criteria 4-2 this criteria is again using the NDVI to separate out vegetation. In this case the false pixels whose NDVI is greater than 0.6 will be vegetation or non-algae land pixels.

Criterion 6-4: b603 LT b593

Some of the nonalgae is still misclassified as algae. The nonalgae dip down at 583/594 and then go up. The algae don't have this trough. The algae trough is at 622/631. Those that are true for this criterion are water with some but very little algae.

Figure 3-21 shows the true color SpecTIR image and Figure 3-22 shows the image classified using the above tiered decision tree criteria. Figures 3-23 and 3-24 show a smaller scale cropped enlargement of the true color and the classified image. Figure 3-25 a-f shows this cropped area with each of the six algae layers overlaid separately.

Image classifications are generally evaluated by measuring the commission and omission error, the producer's and user's accuracy, and a kappa statistic. Commission error measures how many of the pixels within a given generated class are actually known to be part of another feature and therefore have falsely been assigned to the class (e.g. false positive rate). Omission error measures how many pixels known to be part of given feature, have been erroneously left out of the class representing that feature (e.g. false negatives). Producer's Accuracy measures the probability that given pixels known to be in a particular class are actually classified into that class (e.g. sensitivity/specificity). User's accuracy measures the probability that a pixel classified as a given class is truly part of that class (positive/negative predictive value).

All of these except the kappa statistic are calculated for each individual class by comparing the pixels assigned to that class to the pixels known to be a part of that class from ground truth regions of interest (ROIs). These ground truth regions of interest (ROIs) were created manually by visually identifying and assigning the area of the reservoirs to one ROI and the surrounding land area to another ROI. The measures of accuracy for the classification were calculated by comparing the areas assigned to the land and algae classes to these two ROIs. The percent commission errors in Table 3-7, therefore, show the percent of pixels within a given cyanobacteria class that are contained erroneously within the land ground truth ROI or the percent of pixels within the predicted land class that are erroneously assigned to one of the water ground truth ROI. All but that of the cyanobacteria high density and scum have less than 4 % commission error.

Figure 3-21: True-Color SpecTIR Image



Figure 3-22: SpecTIR Image Classified with Tiered Decision Tree

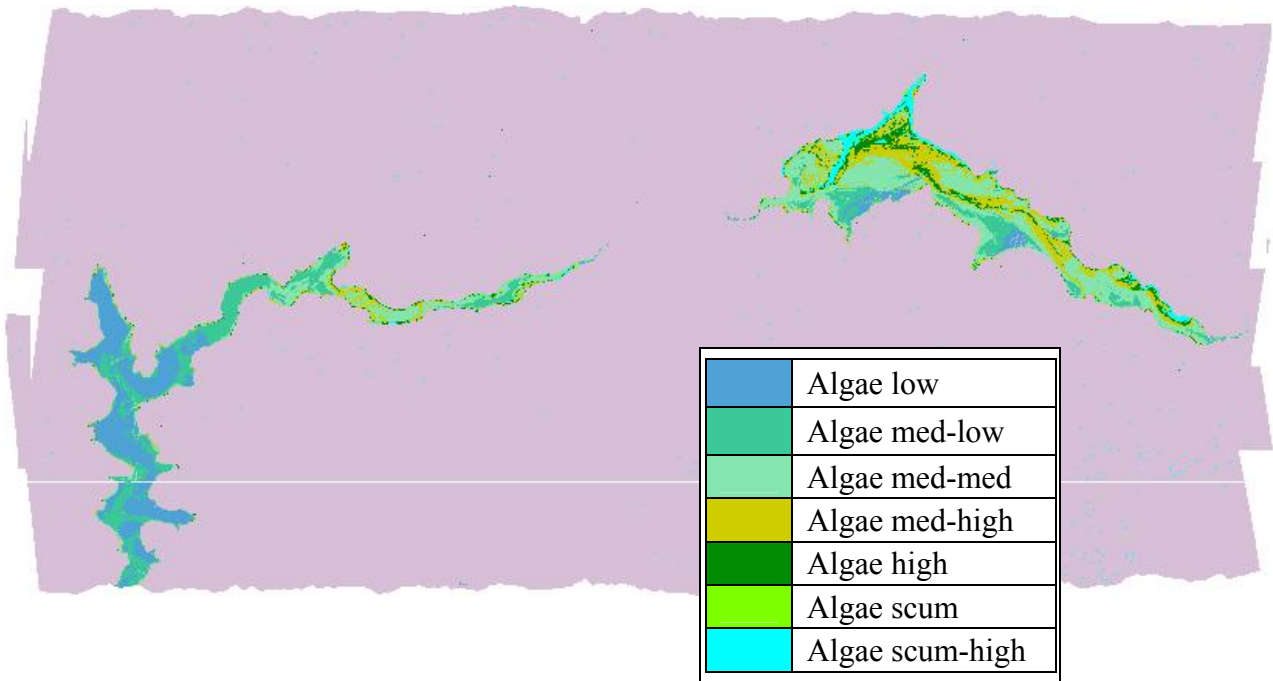


Figure 3-23: Cropped Area True Color

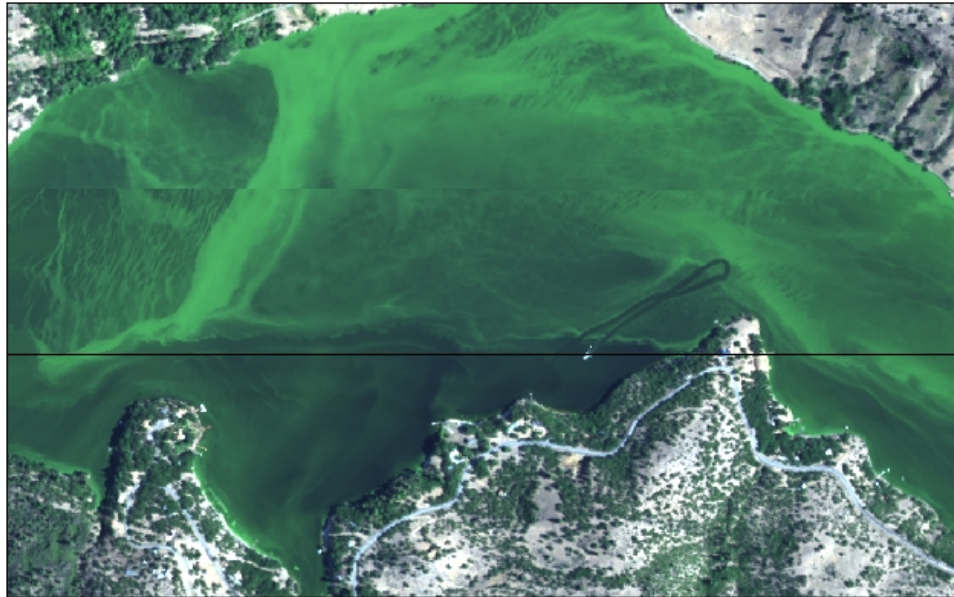


Figure 3-24: Cropped Area Tiered Decision Tree Classification

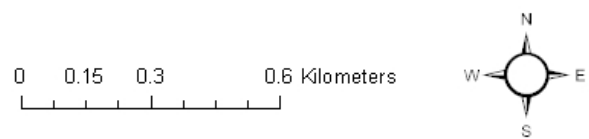
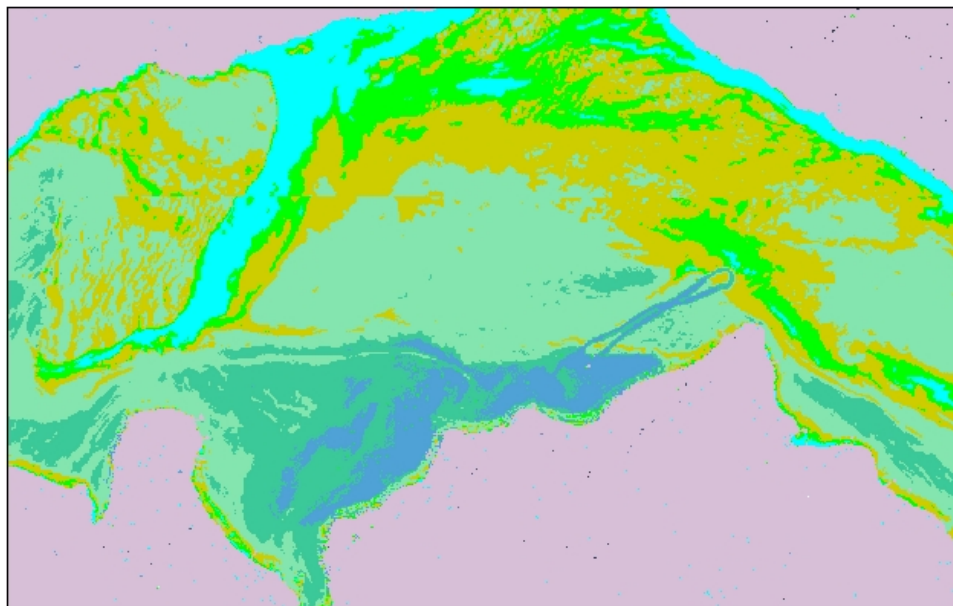
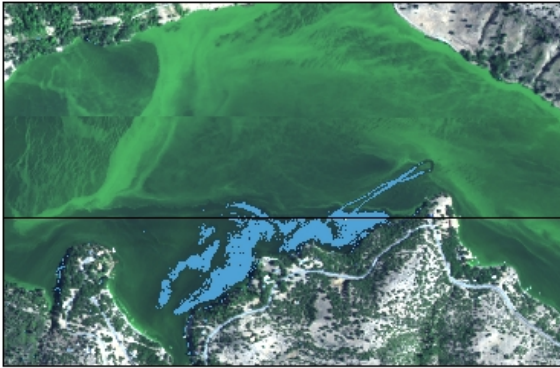


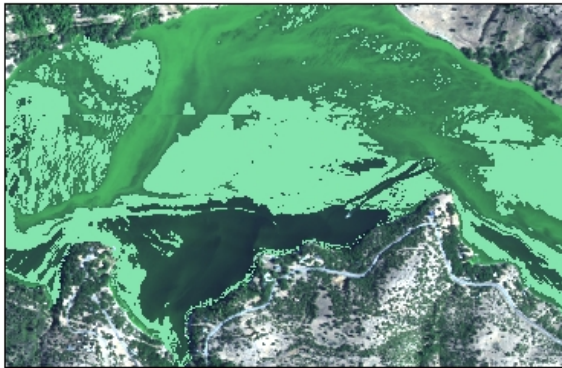
Figure 3-25: Areas Captured by Phytoplankton Bloom Density Categories



a) Low Algae



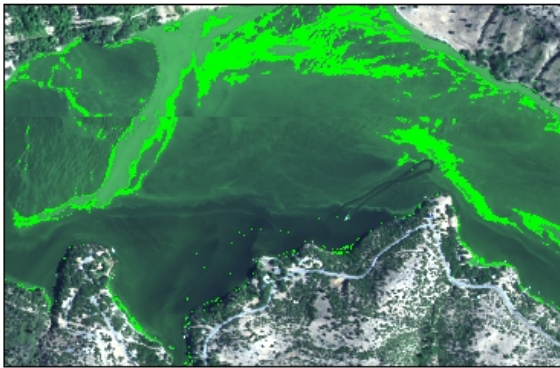
b) Medium-Low Algae



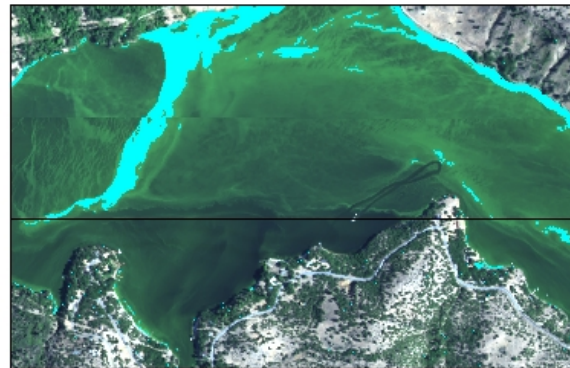
c) Medium-Medium Algae



d) Medium-High Algae



e) High Algae



f) Algae Scum



3.3 Discussion

One key challenge to these data is the large spatial mobility over relatively short periods of time (hours). As a result the time required to mobilize across the field area to collect ground truth data is prohibitive as the system is constantly in flux making the training of any classification using ground truth data potentially flawed due to the shifts in blooms within the water between the time of remote data collection and field data collection for any given site. Jernakoff (Jernakoff et al. 1996) asserts that ground-truthing must be done within 30 minutes of flight times. This limitation is partially remediated in larger systems where the size of the water and the harmful algal bloom are large enough that coarse spatial resolution can be used without compromising the interpretation of variation within the system, as is the case with remote sensing detection and analysis of harmful algal blooms in the ocean. While the blooms are equally mobile in the ocean, their mobility causes shifts within any single pixel (which are typically 0.8-1 km² for the SeaWiFS, MODIS, MERIS or Coastal Zone Color Scanner satellites most commonly used) and therefore should not affect correlations between ground and satellite data. The small size of freshwater algal blooms, however, dictates that they be analyzed using higher spatial resolution which in turn means that the bloom shifts in space move algal densities between pixels rather than just within a single pixel such that a field measurement on the ground several hours after the remote data collection may no longer represent the ground truth for that point.

In future studies the use of multiple stationary field fluorometers that can measure water in real-time may provide suitable ground truth data, but with the present impediments to collecting accurate ground truth data from the algal blooms in small freshwater systems, classification approaches could not benefit from separate training and validation data sets. Rather the full data set is used in generating the classification model. The tiered decision tree was the most effective classification for the SpecTIR image, and visual interpretation as well as understanding of the spectral signatures are used to select the criteria rather than any deviance measure such as the likelihood ratio statistic which can be used to compare all possible splits of the data to find the one that maximizes the dissimilarity between the resulting subsets (Hansen 1996). The absence of ground truth validation data required a modified approach to accuracy assessment. Since the water boundaries are known, the areas classified as algae beyond those boundaries are counted as misclassification and the areas within those boundaries classified as vegetation are also counted as misclassification. In the case of the Klamath reservoirs, we were specifically looking at water bodies known to be impacted by cyanobacteria blooms, and therefore it was not important to be able to separate the bloom's pigment signature from that of land-based vegetation except when there is submerged aquatic vegetation which it is important to be able to distinguish from blooms. The successful separation between blooms and land vegetation is also important for potential future surveillance applications in which images of large areas of land and water are classified to pull out any places where blooms are occurring. This facilitates surveillance and avoids needing to mask images to focus on only water bodies.

Many studies (Friedl and Brodley 1997; Yang et al. 2003; Wright and Gallant 2007; Sesnie et al. 2008; Tooke et al. 2009) have implemented decision trees or classification regression trees by using computer programs such as S-PLUS 'cancor' function (S-Plus, Insightful Corporation, Seattle, Washington), QUEST, and DTREG which divide the training dataset into successively more homogeneous subsets that minimize error and maximize differences between subsets. After

completion, the tree is usually “pruned” by removing branches to avoid over fitting to noise in the training data which may not be perfectly representative of the full dataset. Such software creates a set of hierarchically structured rules or branches based a deviance measure such as the likelihood ratio statistic. Defining the tree with machine learning algorithms by using a statistical procedure is more common, but the manual selection of criteria defined by analyst expert knowledge and qualitative understanding of the class’s spectral properties can sometime produce superior results. This was the case for the present study in which statistical software was unsuccessful in identifying the best nodes and model variables.

The approach for generating the tiered decision tree without training data allows its application as a survey method over areas which have not been visited or tested in the field. Public health officials and natural resource management agencies could apply this decision tree to other imagery for areas with unknown cyanobacteria risks in order to identify sites which warrant the cost of additional field work.

A real value to the classification, in addition to being able to visualize where the medium or high density sections of the bloom are located, is the ability to quantify the intensity of the bloom in terms of area impacted and compare the intensity of the bloom at different dates in time. As is already described, blooms are mobile and one cove or inlet may be densely covered on one day and then remarkably clear another after a shift in wind or other system change. Water managers struggle to identify the best sampling protocol that will facilitate knowing when a particular water body is becoming a health threat and when conditions are resolving. Managers struggle with the number of sites to sample, how many at the surface, and how many at greater depths. Regardless of their decision they are always only able to capture a small snapshot of the system, unlike remote sensing which can provide a synoptic overview of the entire system, making it possible to truly assess relative bloom intensity.

This advantage also extends to temporal comparisons, whereby images from consecutive years can be quantitatively compared to estimate the change in bloom intensity as a result of remediation measures or shifts in environmental conditions. Imagery from multiple dates within on bloom season can further enhance the understanding of the intensity of the bloom and better capture its peak.

Remote sensing for cyanobacteria blooms may be limited in quantification beyond a certain maximum cell density. When scums exist, their thickness may not affect their surface reflectance. A dense scum may accumulate over millimeters of the waters surface or extend several centimeters deep, but both may be identical from above especially with the colony forming properties of cyanobacteria. Both could be capable of fully masking the water surface, but remote sensing would be unable to differentiate between them even though one would obviously have a much higher chlorophyll-a concentration. This would also limit comparisons of overall bloom intensity, but it would be more likely that variations in area covered by the densest bloom material are more relevant to the overall question of bloom intensity.

Balakrishnama, S. and A. Ganapathiraju. "Linear Discriminant Analysis - A Brief Tutorial."
Retrieved August 9, 2009, from

http://www.isip.piconepress.com/publications/reports/isip_internal/1998/linear_discrim_analysis/lda_theory_v1.0.pdf.

- Friedl, M. A. and C. E. Brodley (1997). "Decision tree classification of land cover from remotely sensed data." Remote Sensing of Environment **61**(3): 399-409.
- Hansen, M., R. Dubayah, et al. (1996). "Classification trees: an alternative to traditional land cover classifiers." International Journal of Remote Sensing **17**(5): 1075 - 1081.
- Horler, D. N. H., M. Dockray, et al. (1983). "Red edge measurements for remotely sensing plant chlorophyll content." Advances in Space Research **3**(2): 273-277.
- Jensen, J. R. (2007). Remote Sensing of the Environment: An Earth Resource Perspective. Upper Saddle River, NJ, Pearson Education, Inc.
- Jernakoff, P., P. Hick, et al. (1996). Remote Sensing of Algal Blooms in the Swan River. C. R. S. U. CSIRO Division of Fisheries, Division of Exploration and Mining, Waters and Rivers Commission, CSIRO Australia.
- Klecka, W. R. (1980). Discriminant Analysis. Beverly Hills, Sage Publications.
- MacKay, D. (2003). Chapter 20. An Example Inference Task: Clustering. Information Theory, Inference and Learning Algorithms. Cambridge, Cambridge University Press: 284-292.
- Pal, M. and P. M. Mather (2001). Decision Tree Based Classification of Remotely sensed Data. 22nd Asian Conference on Remote Sensing. Singapore, Centre for Remote Imaging, Sensing and Processing.
- Richardson, L. L. (1996). "Remote sensing of algal bloom dynamics: New research fuses remote sensing of aquatic ecosystems with algal accessory pigment analysis." Bioscience **46**(7): 492-501.
- Rowan, K. S. (1989). Photosynthetic pigments of algae. New York, Cambridge University Press.
- Sesnie, S. E., P. E. Gessler, et al. (2008). "Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments." Remote Sensing of Environment **112**(5): 2145-2159.
- Song, C., C. E. Woodcock, et al. (2001). "Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects?" Remote Sensing of Environment **75**(2): 230-244.
- Tooke, T. R., N. C. Coops, et al. (2009). "Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications." Remote Sensing of Environment **113**(2): 398-407.
- Wright, C. and A. Gallant (2007). "Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data." Remote Sensing of Environment **107**(4): 582-605.
- Yang, C.-C., S. O. Prasher, et al. (2003). "Application of decision tree technology for image classification using remote sensing data." Agricultural Systems **76**(3): 1101-1117.
- Yentsch, C. S. (1960). "The influence of phytoplankton pigments on the colour of sea water." Deep Sea Research (1953) **7**(1): 1-9.

4 Multispectral and Hyperspectral Image Comparison

4.1 Introduction

Decreases in resolution, both spectral and spatial, are expected to decrease the ability of remotely sensed data to detect the presence of toxic cyanobacteria blooms. The various image platforms available for this kind of research provide a trade-off between high spatial resolution, high spectral resolution, high revisit frequency, cost, image processing time and total area covered. To date other researchers have used imagery from Landsat, AVHRR, MODIS, and SeaWiFS to specifically detect cyanobacterial blooms with varying degrees of success, as discussed in Section 1-Background (Dekker et al. 1992; Jupp et al. 1994; Kahru et al. 2000; Vincent et al. 2004; Tyler et al. 2006). The literature disagrees on the efficacy of Landsat or comparable satellites such as ALI to be able to actually predict the presence and concentration of phycocyanin and chlorophyll concentrations and cyanobacteria cell densities in the presence of suspended sediment spectral interference (Vincent et al. 2004; Kutser et al. 2006). Therefore, three different remote sensing platforms, SpecTIR, Ikonos and Landsat, were compared in terms of their ability to identify surface blooms, cyanobacteria specifically, and reasonable estimates of total phytoplankton cell counts or pigment (chlorophyll-a and phycocyanin) concentrations. As explained in Chapter 2, the inability of field samples to serve as accurate ground truth data made it possible to estimate only relative differences in cyanobacteria density rather than any quantitative cell count and/or pigment concentration.

This methods comparison used a hyperspectral airborne sensor (SpecTIR: 2-m spatial resolution, spectral resolution continuous coverage between 400-1000 nm), a high spatial resolution multispectral satellite image (IKONOS: 4-m spatial resolution for each band, 1-m spatial resolution panchromatic image, 4 separate bands of red, green, blue and IR), and a lower spatial resolution satellite image (LANDSAT: 30-m spatial resolution, 7 bands with 3 in the visible range). Chapter 3 presented the analysis of hyperspectral 2-meter spatial resolution SpecTIR airborne imagery to predict the presence and intensity of toxic cyanobacteria blooms. The decision tree was the preferable classification method for the SpecTIR imagery and this chapter compares those results to that of the IKONOS and Landsat satellite images using the same classification method. To compare the three different images, the same classification method (the decision tree) was used on all three. The decision trees allowed existing knowledge regarding the characteristics individual wavelengths and their relationship to cyanobacteria blooms and the other features in the image to be incorporated into the model. It was also easily applied to the whole image within the ENVI software.

4.2 Methods

4.2.1 Image Acquisitions

IKONOS imagery was acquired on July 27 and July 30, 2007, collections scheduled with GeoEye to coordinate with the scheduled SpecTIR airborne passby. Both images had 0% cloud cover and occurred at 7:09 p.m. and 7:18 p.m. respectively. The resolution of the 4 bands (red, green, blue and near infrared) are 4 meters, but the pan-shaped imagery with 1-meter resolution was also obtained in which the imagery from the 4-meter sensor and the 82-cm black and white panchromatic sensor were merged.

Landsat TM (Path 045, Row 031) imagery was downloaded for August 23, 2005, the most recent summer (i.e. bloom season) dataset available for the Klamath Site. Images can be downloaded on the Earth Science Data Interface of the Global Land Cover Facility (NASA Landsat Program 2005). Landsat passes over the area every 16 days and the image resolution is 30 x 30 meters (Chander et al. 2009).

4.2.2 Decision Trees

See section 3.2.7.

4.3 Results

4.3.1 Decision Tree for IKONOS

Like the analysis of the SpecTIR imagery a decision tree was constructed in ENVI 4.5 (ITT Visual Information Solutions, Boulder, CO). Criteria were selected based on knowledge of the spectral characteristics of the algal pigments, water, vegetation and land which were described in section 3.1.1. The spectral characteristics of the output classes were evaluated after each criteria was added to inform the definition of the subsequent criteria. Based on the known areas of land with vegetation versus water impacted by cyanobacteria, spectra for misclassified and correctly classified pixels were examined in order to select the best next criteria. The process was repeated until visual inspection of the final classified map continually showed no improvement in class area refinement. Figure 4-1 below shows the final decision tree used for the IKONOS image. All the features on the land (e.g. the groups depicted as LAND 1-5 in Figure 4-1) were combined into one single class and three separate cyanobacteria/algae groups within the water were created corresponding to varying cyanobacteria cell densities: *high*, *medium*, and *low*. Ten criteria were used in the tiered decision tree which are described below.

Notation: b= bandwidth, GT = greater than, LT = less than, EQ = equal, Band 1 = NIR (757-853 nm), Band 2 = Blue (445-516 nm), Band 3 = Green (505-595 nm), Band 4 = Red (632-698 nm)

Criterion 1-1: b1 EQ B2 EQ B3 EQ b4 EQ 0

The pixels that are true for this criteria are empty image pixels

Criterion 2-2: (b1-b4)/(b1+b4) GT 0.4

The pixels that are true are mostly high scum and some land vegetation. This is making use of the NDVI, so it separates as the true pixels, those that have particularly strong red edge typical of abundant chlorophyll pigments.

Criterion 3-4: b2/ b3 LT 0.85

Criterion 3-4 then takes the pixels with high NDVI (e.g. high chlorophyll content) that were true for Criteria 2-2 and separates the water based pixels from the land based pixels, by looking for those that have smaller Band 2 peaks with respect to Band 3. The positive slope between Band 2 and Band 3 is greater for the land vegetation pixels and therefore the Band 3 reflectance is greater and the ratio of b2/b3 is smaller. The pixels true for this criterion are high algae content and the false pixels are land based vegetation.

Criterion 3-3: $(b1 - b4) / (b1 + b4) \geq 0.2$ AND $(b1 - b4) / (b1 + b4) < 0.4$

This criterion again uses the NDVI. The true pixels are those whose NDVI falls between 0.2 and 0.4. They are the pixels containing medium to high algae content with some misclassified land pixels which are separated out in the next two criteria (4-6, 5-11 and 5-12). The false pixels are low to medium algae content with some misclassified high algae content. They are separated using the criteria 4-5, 5-9, and 6-17.

Criterion 4-6: $b1 < 340$

This criterion divides the pixels according to the strength of their reflectance in the NIR (band 1).

Criterion 5-11: $b4 / b3 < 0.7$

There is a negative slope between Band 3 and Band 4 indicating that more light is absorbed in the red band than in the green band leaving a peak in the green (Band 3). The true pixels for this criterion are high algae content as opposed to land based pixels which are the false pixels. The true pixels are characterized by a larger peak at Band 3 making their ratios of $B4/B3$ smaller.

Criterion 5-12: $b2 / b3 \geq 0.90$

There is a positive slope between Band 2 and Band 3 indicate that more light is absorbed in the blue band than in the green band leaving a peak in the green (Band 3). The false pixels for this criterion are height algae content as opposed to land based pixels which are the true pixels. The false pixels are characterized by a larger peak at Band 3 making their ratios of $B2/B3$ smaller.

Criterion 4-5: $b1 \geq 100$ AND $b1 < 200$

This criterion uses the numeric values of reflectance in the NIR (band 1) to separate out medium algae content pixels as those who are true for this criterion. Those that are false are separated using Criterion 5-9.

Criterion 5-9: $b1 < b4$

Low algae content results in the high absorbance in the NIR characteristic of water. This is seen in the pixels true for this criterion whose NIR reflectance is less than the reflectance in the red band 4.

Criterion 6-17: $b1 < b2$

Pixels true for this criteria are a few high algae content those that are false are land based pixels.

Figure 4-2 a-c shows the IKONOS image classified by the tiered decision tree. Figure 4.2a is Iron Gate Reservoir where the highest algal density classification falls predominantly along the shoreline, within protected coves, and in the narrow stretch that forms the eastern portion of the reservoir. This matches well with all my field reconnaissance and knowledge of the system. The bloom cell density is less in the open areas of water in the reservoir, which is again accurately represented in the classified image. This is less likely to be a result of physical water parameters such as depth and temperature, and more a feature of the geography of the reservoir and wind patterns which serve to clump and concentrate the blooms within the reservoir. Figure 4.2b shows a similar pattern of the bloom in Copco Reservoir with the densest portions predominating along the outer shores of the reservoir. The water movement and wind patterns do produce swirled patterns of mixed cell density covering the full cross-section of the reservoir. This is

clearly visible in the larger scale cropped enlargement of a portion of Copco Reservoir in Figure 4-2c. Figure 4-3 a-e shows the unclassified true color image along with each of the 3 cyanobacteria/algae layers overlaid separately. These images make it easy to see how well matched the classification is to the image. For example, in true color unclassified cropped image in Figure 4-3b one can see the bright sections of extremely thick scum along the very shore of the reservoir and stretching down in a finger shape along the left half of the image. Figure 4-3e shows the high algae density classification as the green overlaid area which overlaps well with what visual observation (supported by intimate knowledge of the field conditions) shows also to be the higher cell density. Similarly the darker areas in the middle open water sections of the reservoir which are not so visibly green in the true color unclassified cropped image in Figure 4-3b are the areas of the water where there is many fewer cyanobacteria. Although they are still present, they are not present in such thick scum on the water surface. This again matches well with the area classified as low algae density and shown in light blue in figure 4-3c. The remaining intermediate areas are medium algal density and again correctly match with the area classified as such and shown with the green overlay shown in Figure 4-3d. Figure 4-4 shows sample spectra for the 3 different classes of cyanobacteria/algae. The red edge (see Chapter 3) can be clearly seen in the high algae pixels as the NIR band (b1) reflectance is considerably higher than all of the other bands (blue, green and red) and it is particularly minimized or absent in the low algae pixels. The higher reflectance in the green band and more absorbance in the blue and red bands is also apparent as expected with the presence of chlorophyll and phycocyanin.

Again the commission and omission error, and the producer's and user's accuracy Image classifications were calculated for each individual class to evaluate the image classifications in comparison to that of the SpecTIR and the Landsat. Commission error measures how many of the pixels within a given generated class are actually known to be part of another feature and therefore have falsely been assigned to the class (e.g. false positive rate). Omission error measures how many pixels known to be part of given feature have been erroneously left out of the class representing that feature (e.g. false negatives). Producer's Accuracy measures the probability that given pixels known to be in a particular class are actually classified into that class (e.g. sensitivity/specificity). User's accuracy measures the probability that a pixel classified as a given class is truly part of that class (positive/negative predictive value).

Pixels assigned to each class by the decision tree were compared to the pixels known to be a part of that class from ground truth regions of interest (ROIs). These ground truth regions of interest (ROIs) were created manually by visually identifying and assigning the area of the reservoirs to one ROI and the surrounding land area to another ROI. The measures of accuracy for the classification were calculated by comparing the areas assigned to the land and algae classes to these two ROIs. The percent commission errors in Table 3-8 below, therefore, show the percent of pixels within a given cyanobacteria class that are contained erroneously within the land ground truth ROI or the percent of pixels within the predicted land class that are erroneously assigned to one of the water ground truth ROI.

The percent commission errors in Table 4-1 below therefore show the percent of pixels within a given cyanobacteria class that are contained erroneously within the land ground truth region or for the predicted land class the number of pixels erroneously assigned to one of the water/algae

classes. The commission error is only 5.91% for the low cyanobacteria density class, however, it gets increasingly large for the other two class reaching almost 50% for the high cyanobacteria density class. This is significantly worse than for the SpecTIR data where four of the cyanobacteria classes had commission errors less than 4% and the two most dense cyanobacteria classes (high and scum) had commission errors of 21.54% and 27.66% respectively.

TABLE 4-1: Ikonos Decision Tree Classification Accuracy

	Low	Medium	High
Commission Error	5.91%	21.18%	48.64%
Area equivalent of pixel error	~55,816 m ²	~1,982,428 m ²	~741,728 m ²
User Accuracy	94.09%	78.82%	51.36%

There may be better decision trees for the data or multiple trees with comparable performance, but the tree captures the spectral features observed using a manual rather than a computerized selection. The tree criteria were refined until no improvement in classification could be obtained based on visual assessment.

4.3.2 Decision Tree for Landsat

A decision tree was also constructed in ENVI 4.5 (ITT Visual Information Solutions, Boulder, CO) for the 2005 Landsat image. As with the process for the development of the Ikonos and SpecTIR decision trees, criteria were selected based on knowledge of the spectral characteristics of the algal pigments, water, vegetation and land which were described in section 3.1.1. The spectral characteristics of the output classes were evaluated after each criteria was added to inform the definition of the subsequent criteria. Based on the known areas of land with vegetation versus water impacted by cyanobacteria, spectra for misclassified and correctly classified pixels were examined in order to select the best next criteria. The process was repeated until visual inspection of the final classified map continually showed no improvement in class area refinement. Figure 4-5 below shows the final decision tree used for the Landsat image. All the features on the land (e.g. the groups depicted as LAND 1-3 in Figure 4-5) were combined into one single class and three separate cyanobacteria/algae groups within the water were created corresponding to varying cyanobacteria cell densities: *high*, *medium*, and *low*. Five criteria were used in the tiered decision tree which are described below.

Notation: b= bandwidth, GT = greater than, LT = less than, EQ = equal,

b1 = blue = 450-520 nm

b2 = green = 520-600 nm

b3 = red = 630-690 nm

b4 = Near IR = 760-900 nm

b5 = Mid IR = 1550-1750 nm

b6 = Thermal IR = 10400-12500 nm

b7 = Mid IR = 2080-2350 = Mid-IR

Criterion 1-1: b3 GT b2

This criteria separates the land. The vegetation, algae, and water all absorb in the visible red wavelengths, while the land reflects in this range. Land reflects more in B3, the red wavelengths

of 630-690 nm, than for B2, the green wavelengths of 520-600. These land pixels are true for this criterion.

Criterion 2-1: $(b4/1000 - b3/1000)/(b4/1000 + b3/1000) > 0.75$

This is like the NDVI. The pixels with the highest chlorophyll content are true and those with less are false. This criterion separates out the low and medium algae from the vegetation and high algae.

Criterion 3-2: $b5 < b4$

This criterion separates out more of the land-based pixels which are false while the high algae vegetation are true.

Criterion 4-4: $b5 > b3$

This criterion again separates the land-pixels, which are true in this case, from the high algae pixels.

Criterion 3-1: $b4 > 15$

The earlier separated low and medium algae which were both false for Criterion 2-1 are now separated from each other. The medium algae are true for this criterion.

Figure 4-6 a-c shows the Landsat image classified by the tiered decision tree. Figure 4.6a is Iron Gate Reservoir where the highest algal density classification falls predominantly along the shoreline, within protected coves, and in the narrow stretch that forms the eastern portion of the reservoir. This matches well with all my field reconnaissance and knowledge of the system, as well as with the classified and unclassified Ikonos and SpecTIR images. The bloom cell density is less in the open areas of water in the reservoir, which is again accurately represented in the classified image. Figure 4.6b shows a similar pattern of the bloom in Copco Reservoir with the densest portions predominating along the outer shores of the reservoir. The water movement and wind patterns do produce swirled patterns of mixed cell density covering the full cross-section of the reservoir. This is clearly visible in the larger scale cropped enlargement of a portion of Iron Gate Reservoir in Figure 4-6c., including a smaller scale cropped enlargement of a portion of Copco Reservoir. Figure 4-7 a-e shows the unclassified true color image along with each of the 3 cyanobacteria/algae layers overlaid separately. In the unclassified section of Iron Gate Reservoir shown in Figure 4-7b the darker areas without bright green coloring from thick cyanobacteria blooms correspond to the areas classified as low density algal bloom and shown overlaid on the image in Figure 4-7c. The bright green sections along the outer edges of the water which predominate in this narrow and shallow section of the Iron Gate Reservoir correspond well with the area overlaid as the high density algal bloom in Figure 4-7c. Figure 4-8 shows sample spectra for the land, vegetation and cyanobacteria/algae.

Again the commission and omission error, and the producer's and user's accuracy Image classifications were calculated for each individual class to evaluate the image classifications in comparison to that of the SpecTIR and the IKONOS. As before, commission error measures how many of the pixels within a given generated class are actually known to be part of another feature and therefore have falsely been assigned to the class (e.g. false positive rate). Omission error measures how many pixels known to be part of given feature have been erroneously left out

of the class representing that feature (e.g. false negatives). Producer's Accuracy measures the probability that given pixels known to be in a particular class are actually classified into that class (e.g. sensitivity/specificity). User's accuracy measures the probability that a pixel classified as a given class is truly part of that class (positive/negative predictive value).

Pixels assigned to each class by the decision tree were compared to the pixels known to be a part of that class from ground truth regions of interest (ROIs). These ground truth regions of interest (ROIs) were created manually by visually identifying and assigning the area of the reservoirs to one ROI and the surrounding land area to another ROI. The measures of accuracy for the classification were calculated by comparing the areas assigned to the land and algae classes to these two ROIs. The percent commission errors in Table 3-8 below, therefore, show the percent of pixels within a given cyanobacteria class that are contained erroneously within the land ground truth ROI or the percent of pixels within the predicted land class that are erroneously assigned to one of the water ground truth ROI.

The percent commission errors in Table 4-2 below therefore show the percent of pixels within a given cyanobacteria class that are contained erroneously within the land ground truth region or for the predicted land class the number of pixels erroneously assigned to one of the water/algae classes. The commission error is only less than 6% for all 3 cyanobacteria density classes. Consistent with the evaluations of the other images, the accuracy is greatest for the least dense cyanobacterial class, and the accuracy decreases as cyanobacteria density intensifies making the signal overlap more with that of land vegetation both because of the decrease in water's signal and the fact that highest density blooms usually occur closer to the reservoirs banks where water is shallow and aquatic vegetation may be mixed in the pixel.. This is not much worse than for the SpecTIR data where four of the cyanobacteria classes had commission errors less than 4% and the two most dense cyanobacteria classes (high and scum) had commission errors of 21.54% and 27.66% respectively.

TABLE 4-2: Landsat Decision Tree Classification Accuracy

	Low	Medium	High
Commission Error	0.65%	4.63%	5.85%
Area equivalent of pixel error	~28,800 m ²	~63,000 m ²	~87,300 m ²
User Accuracy	99.35%	95.37%	94.15%

As with the classifications for the other images, there may be better decision trees for the data or multiple trees with comparable performance, but the tree captures the spectral features observed using a manual rather than a computerized selection. The tree criteria were refined until no improvement in classification could be obtained based on visual assessment.

4.4 Discussion

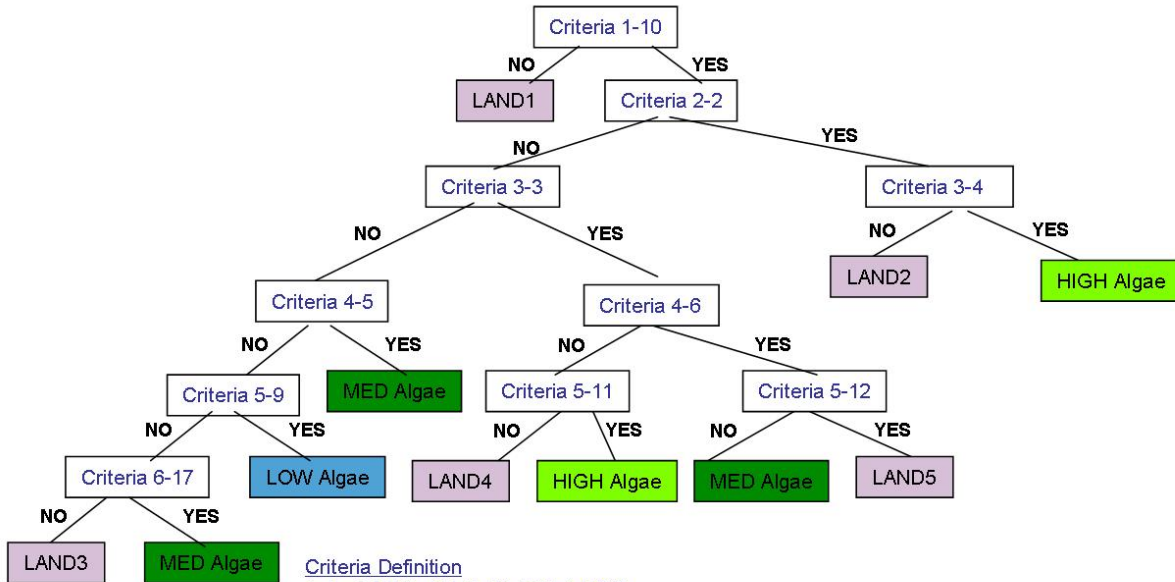
The SpecTIR imagery did have the best accuracy in comparison to the Landsat and the IKONOS. Not only is the SpecTIR data highly accurate, but it provides much more detailed information. Although the Landsat data had less than 6% error for all classes, only three classes of cyanobacteria density could be separated as opposed to the six classes of cyanobacteria separated with the SpecTIR data. The resolution of the SpecTIR data offers a greater opportunity to separate out subtle shifts, early blooms, and blooms of different organisms (not tested here due to the homogeneity of the cyanobacterial blooms in the Klamath). Although the cost of acquiring

these data is higher than the cost of acquiring most types of satellite imagery, much of the cost is related to the fixed cost of transporting the sensor to the study site. This means that the technology to implement this technique may be cost effective if large areas are being surveyed.

Landsat imagery has far better accuracy than the IKONOS despite its much coarser spatial resolution (see Table 4-1 and 4-2). Landsat is currently most accessible and most economical of the three image platforms. The fact that even the highest algae density class still had 94.15% accuracy means that even without the spectral and spatial resolution, valuable surveillance of cyanobacteria blooms is possible. It may be that the panchromatic merging to create the higher spatial resolution of the IKONOS imagery may have a part in the poor accuracy observed. It may also be possible that there is a distinct advantage in separating out the cyanobacteria by having the few additional spectral bands provided in Landsat data. A more likely explanation, however, is the limitations of the spectral scale for the IKONOS data are compounded by having such high spatial resolution. There are more pixels to potentially misclassify and therefore an opportunity for high error.

The results indicate that when given the choice the investment in higher spectral resolution should be chosen over higher spatial resolution as the former appears to provide more benefits in cyanobacteria detection. Ideally high spatial and spectral resolution imagery will become more and more available such that the cost is not prohibitive for this kind of surveillance. In the mean time, there is already a wealth of free or affordable imagery, such as Landsat, of lower resolution which these results indicate will be highly effective in a first round of surveillance for cyanobacteria. Lower spatial resolution imagery also decreases the amount of time invested in image processing.

Figure 4-1: IKONOS Tiered Decision Tree



Criteria Definition
 1-1 = b1 EQ B2 EQ B3 EQ b4 EQ 0
 2-2 = (b1-b4)/(b1+b4) GT 0.4
 3-3 = (b1-b4)/(b1+b4) GT 0.2 AND (b1-b4)/(b1+b4) LT 0.4
 3-4 = b2/b3 LT 0.85
 4-5 = b1 GT 100 AND b1 LT 200
 4-6 = b1 LT 340
 5-9 = b1 LT b4
 5-11 = b4/b3 LT 0.7
 5-12 = b2/b3 GT 0.90
 6-17 = b1 LT b2
 Notation: b=bandwidth, GT= greater than, LT = less than, EQ = equals

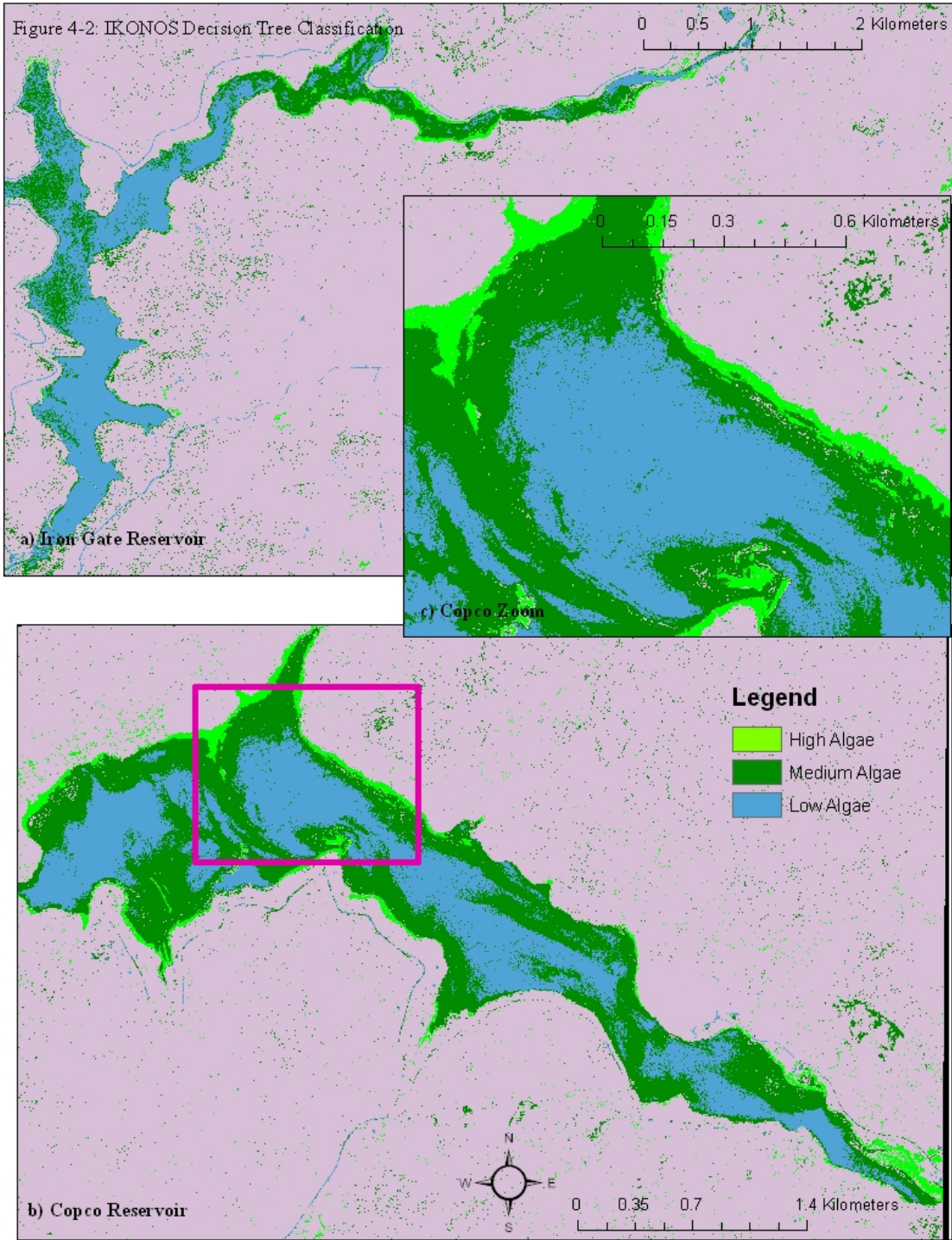


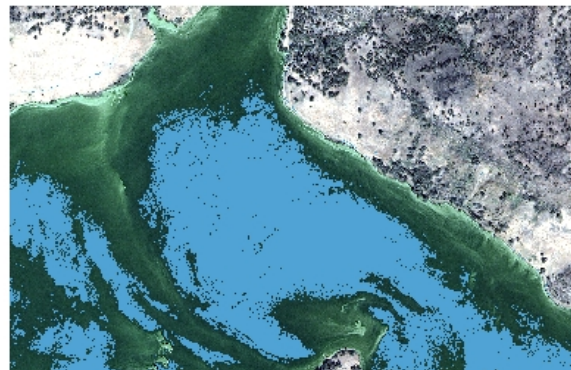
Figure 4-3: Areas of Different Algal Cell Densities in Decision Tree Classification of IKONOS



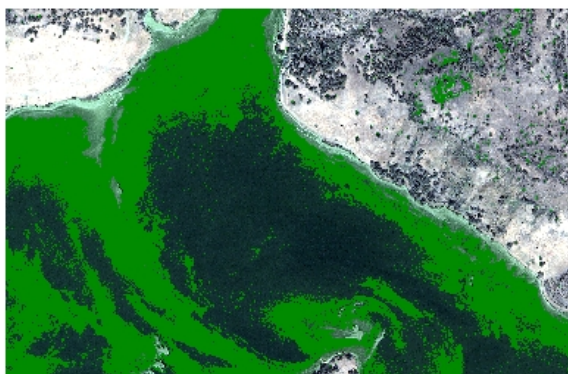
a) True Color Full Area IKONOS Image



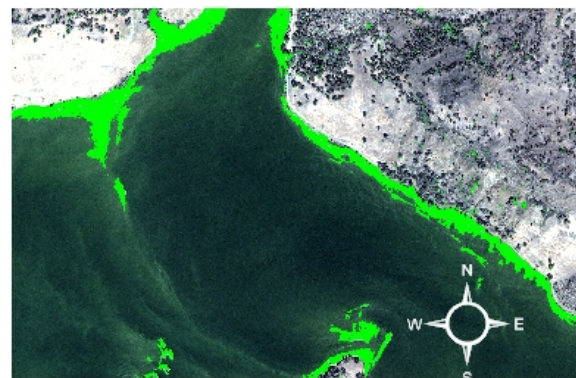
b) True Color Cropped Image



c) Low Algae



d) Medium Algae



e) High Algae 0 0.2 0.4 0.8 Kilometers

Figure 4-4:

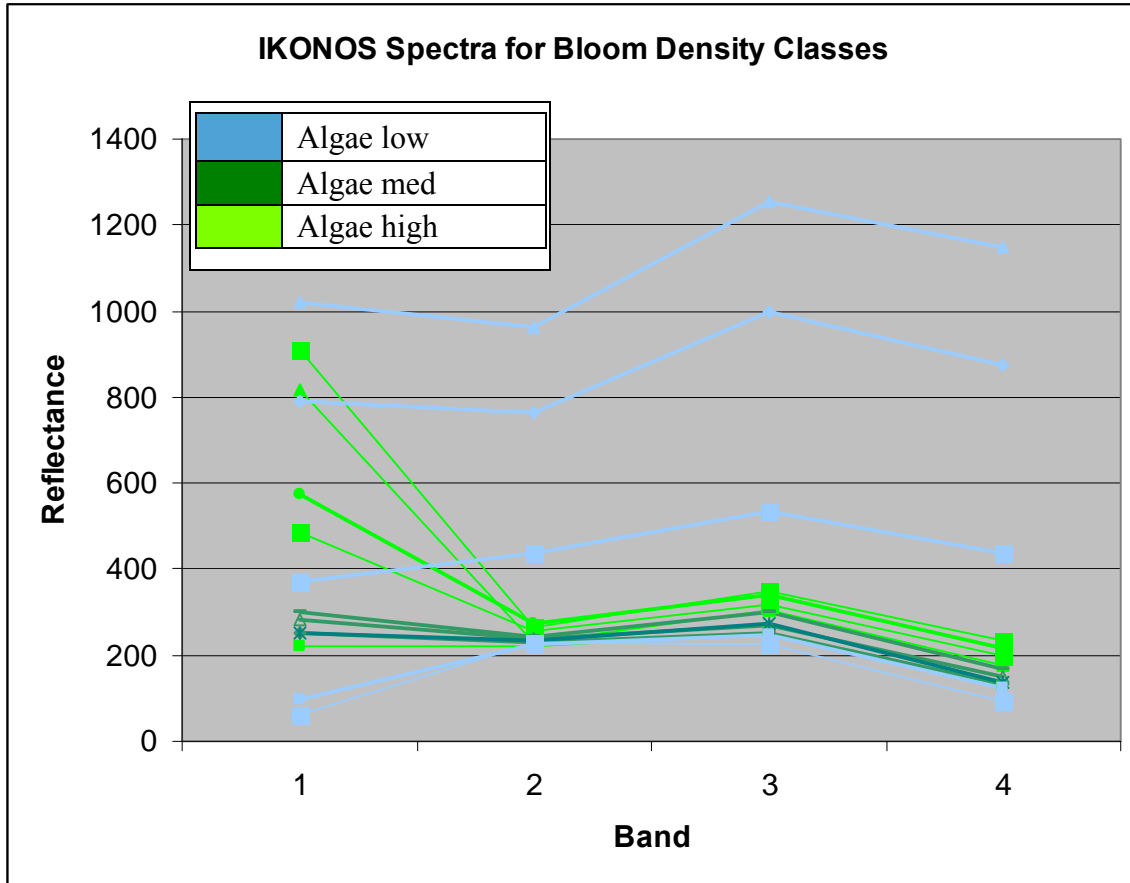
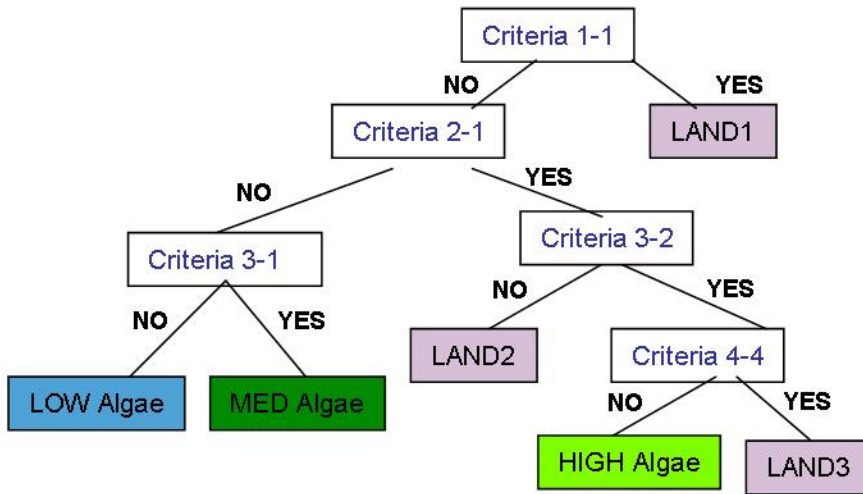


Figure 4-5: LANDSAT Tiered Decision Tree



Criteria Definition

1-1 = b3 GT b2

2-1 = $(b4/1000 - b3/1000) / (b4/1000 + b3/1000)$ GT 0.75

3-1 = b4 GT 15

3-2 = b5 LT b4

4-4 = b5 GT b3

Notation: b=bandwidth, GT= greater than, LT = less than, EQ = equals

Figure 4-6: Landsat Decision Tree

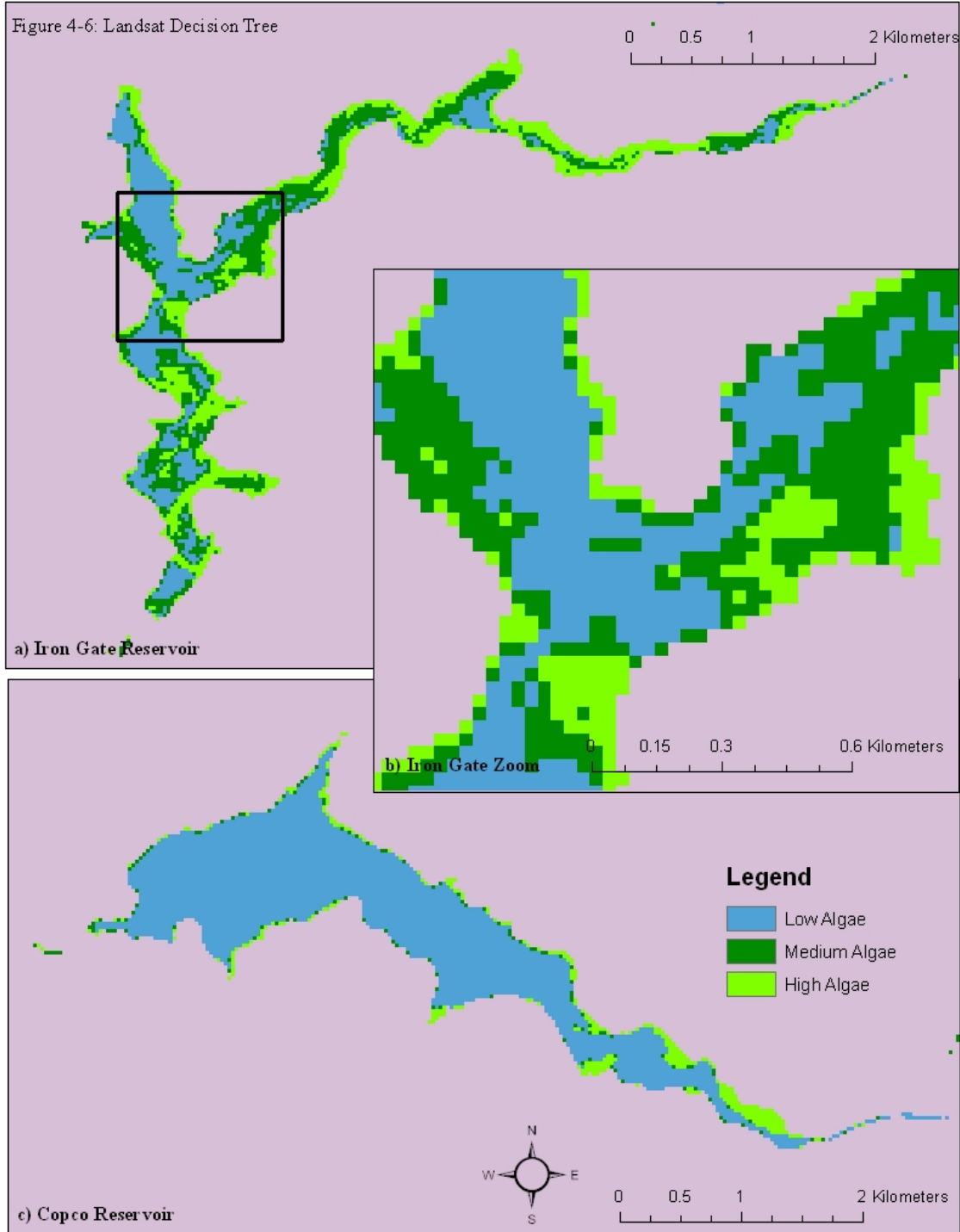
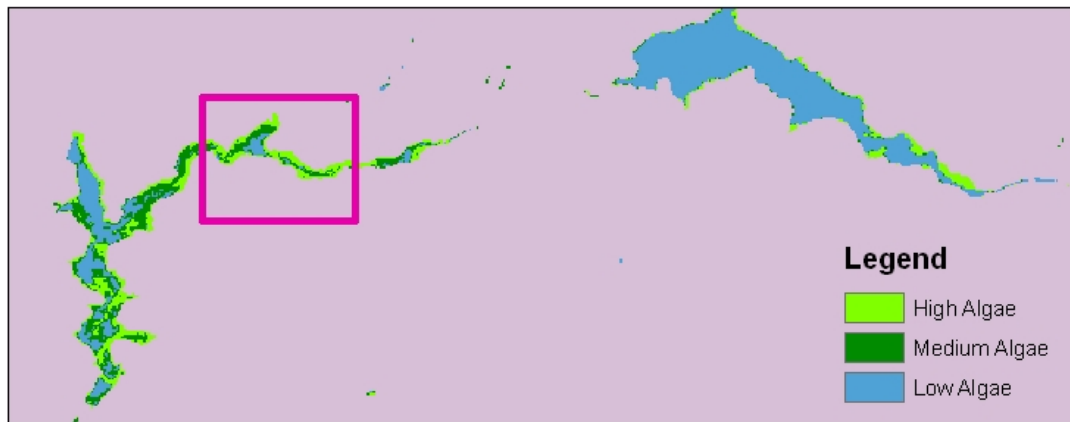


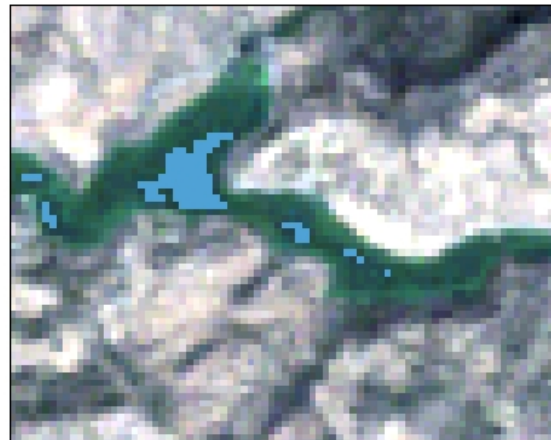
Figure 4-7: Areas Captured by Phytoplankton Bloom Density Classes in Decision Tree Classification of Landsat



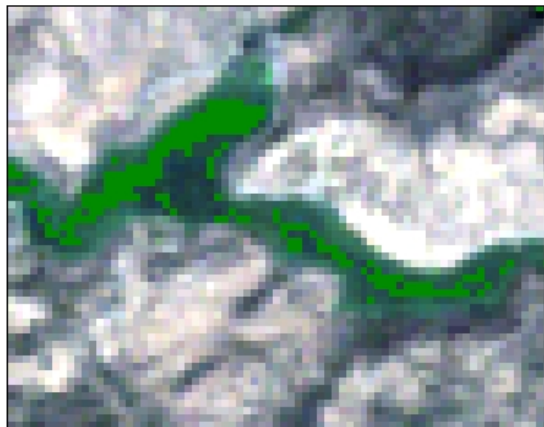
a) Landsat Decision Tree Classification



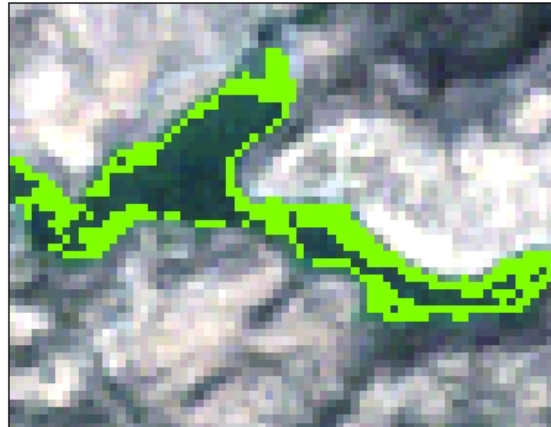
b) Unclassified Landsat Image



c) Low Algae

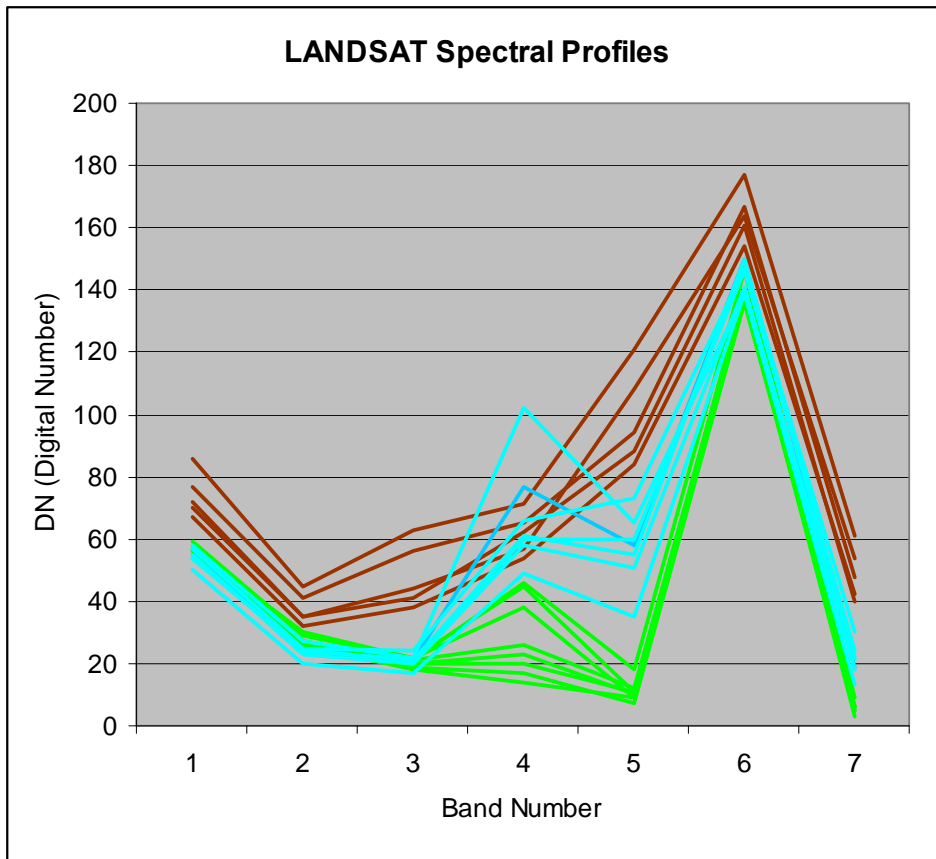


d) Medium Algae



e) High Algae

Figure 4-8: Sample Spectral Profiles for Three Classes in Landsat Image



Soil/Ground
Vegetation
Algae

- Chander, G., B. L. Markham, et al. (2009). "Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors." Remote Sensing of Environment **113**: 893-903.
- Dekker, A. G., T. J. Malthus, et al. (1992). "Remote sensing as a tool for assessing water quality in Loosdrecht lakes." Hydrobiologia **233**(1 - 3): 137-159.
- Jupp, D. L. B., J. T. O. Kirk, et al. (1994). "Detection, Identification and Mapping of Cyanobacteria - Using Remote-Sensing to Measure the Optical-Quality of Turbid Inland Waters." Australian Journal of Marine and Freshwater Research **45**(5): 801-828.
- Kahru, M., J. M. Leppanen, et al. (2000). "Cyanobacteria blooms in the Gulf of Finland triggered by saltwater inflow into the Baltic S." Marine Ecology Progress Series **207**: 13-18.
- Kutser, T., L. Metsamaa, et al. (2006). "Monitoring cyanobacterial blooms by satellite remote sensing." Estuarine, Coastal and Shelf Science **67**(1-2): 303-312.
- NASA Landsat Program (2005). Landsat TM Scene L5CPF20050701_20050930_04. Sioux Falls, USGS.
- Tyler, A. N., E. Svab, et al. (2006). "Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment." International Journal of Remote Sensing **27**(8): 1521-1537.
- Vincent, R. K., X. Qin, et al. (2004). "Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie." Remote Sensing of Environment **89**(3): 381-392.

5 Temporal Comparison

5.1 Introduction

Remote sensing has shown great potential in agricultural mapping and monitoring due to its advantages over traditional procedures in terms of cost effectiveness and timeliness in the availability of information over larger areas (Murthy et al, 1998).

A large advantage of monitoring cyanobacteria blooms using remote sensing data is the synoptic overview provided of the whole system at a single point in time. Images from multiple points in time facilitate analyzing temporal shifts and following the results of environmental changes or remediation steps. Remote sensing has shown great potential for mapping and monitoring of other land features such as agriculture with its availability of information over larger areas and temporal change of forest cover and vegetation patterns has also been successfully detected,

Change detection techniques can calculate new maps highlighting those pixels that are assigned to different classes in the two images. These techniques have been used in many different contexts including the quantification of impacts from wildfires (Richards 1984), grazing (Blanco et al. 2009), re-vegetation programs, agriculture (R.S. Lunetta and C.D. Elvidge 1999), urban and regional planning (Li and Yeh 1998; Masek et al. 2000), and other land use and land cover changes (Ghioca-Robrecht et al. 2008). Karfs and Wallace (Karfs and Wallace 2001) found that information on long-term trends and changes over time are best provided by an integrated method that uses both the extrapolative capacity of remote sensing and on-ground field data. Most change detection studies start with a modeling phase in which some combination of algorithms are used to classify the spectral data in both images followed by a subtraction phase in which the change is calculated by the difference from subtracting one map's values from the other. The difference itself can then be visualized on a new "change" map (Kennedy et al. 2009).

Landsat TM data of the Iron Gate and Copco reservoirs is available free for the end of the bloom season (late August/early September) over four years (2005, 2006, 2008, and 2009). Given the high accuracy achieved with the decision tree for Landsat data that was developed and described in Chapter 4 it was used to classify these images from four different years.

5.2 Methods

Landsat TM (Path 045, Row 031) imagery was downloaded on the Earth Science Data Interface of the Global Land Cover Facility (NASA Landsat Program 2005). Landsat passes over the area every 16 days and the image resolution is 30 x 30 meters (Chander et al. 2009).

The tiered decision tree was used with criteria specifying the spectra that allow the creation of mutually exclusive classes for degrees of cyanobacteria density and non-bloom materials. The same decision tree which is described in section 4.3.2 was applied to all the Landsat images. The number of pixels within the water boundary assigned to each class in all images was calculated and compared. Unlike most other change detection studies, it was not important for these data to know the specific locations in the reservoir that had changed in bloom density since the blooms

are already mobile both vertically and horizontally. Rather the total area impacted by each bloom density level was important, and therefore no subtraction phase was necessary.

5.3 Results

Table 5-1 shows the number of pixels correctly classified within each of the three cyanobacteria classes for each of the images. Three images were downloaded from 2005 from three dates at the end of the bloom season (August 23, August 31, and September 24). The first image collected on 8/23/2005 has the largest high density bloom area. Eight days later on 8/31/05 the amount of high density bloom decreased by more than 75%, the medium density bloom decreased by 5% and the low density bloom increased. Almost a month later at 9/24/05 the high density bloom had increased, the medium density bloom had decreased by 80% and the low density bloom had increased slightly (~9%). As expected the three images confirm the seasonal decrease in the overall bloom intensity with the high and medium density decreasing from roughly 20% of the water surface on August 23 to roughly 3-5% of the water surface on September 24.

The 2006 image was acquired on August 26. The 2006 bloom is smaller than the 2005 bloom on both August 23 and August 31, but the September 2005 bloom and the August 2006 bloom are roughly equivalent in size. It may be that the 2006 bloom was quite large, but diminished earlier in the summer than occurred in 2005, or the bloom intensity overall may just have been less in 2006. This uncertainty in drawing conclusions from the data highlights a challenge for temporal trend analysis. Even with synoptic data for single points in time, drawing conclusions about the whole season requires an image from more than one time point. With only one image for each of the other 3 years, it is not appropriate to conclude which year had a more or less intense bloom. Although the bloom in the August 26, 2006 image is smaller than those in the 2008 and 2009 images, and the 2009 image bloom is smaller than the 2008 image bloom, multiple images from each year are needed to compare the blooms from these three years. The regular bi-monthly monitoring of the bloom dynamics within the Iron Gate and Copco reservoirs indicates that 3-4 images collected between June 1 and August 31 at fairly equal intervals should be adequate to allow rough comparisons in bloom intensity between years.

Table 5-2 shows the commission error and user accuracy for the six different Landsat images. Although the accuracy does vary by image, the pixels within each class in Table 5-1 are all correctly classified within the water body and are unlikely affected by the range in commission errors for each image.

Table 5-1: Per Cent of Pixels in each Class

Date	Low	Med	High
8/23/2005	62%	18%	20%
8/31/2005	80%	16%	3%
9/24/2005	92%	3%	5%
8/26/2006	93%	3%	3%
8/31/2008	78%	15%	7%
9/3/2009	83%	11%	6%

Table 5-2: Image Accuracy

LANDSAT Acquisition Date	8/23/2005	8/31/2005	9/24/2005	8/26/2006	8/31/2008	9/3/2009
	Landsat 7	Landsat 5	Landsat 5	Landsat 5	Landsat 5	Landsat 5
Commission Error						
Low	0.65	16.34	3.53	1.5	0.4	0.75
Medium	4.63	51.86	42.74	35.78	11.38	9.74
High	5.85	38.35	7.07	11.74	24.93	9.94
User Accuracy						
Low	99.35	83.66	96.47	98.5	99.6	99.25
Medium	95.37	48.14	57.26	64.22	88.62	90.26

Figure 5-1 shows the true color images for each time point and Figure 5-2 shows the classified images depicting locations of blooms with algal density varying between low, medium and high.

5.4 Discussion

In any given water system the specific objectives of the monitoring and knowledge of the individual system will help determine the necessary frequency for image acquisition in order to do temporal trend analysis. Comparing the intensity of bloom seasons over consecutive years will require several images over the course of the bloom to capture it during the peak and during its decline. If the blooms last several months, as they do in the Klamath reservoirs, every 2 weeks throughout the bloom season might be suitable, making the Landsat's 17 day pass-by a possibility. Shorter blooms would need more acquisitions more closely spaced. Since Landsat 5 and Landsat 7 each have a separate 17 day pass-by cycle, using both jointly would provide closer spacing time between the images acquisitions.

The Klamath reservoirs have now been monitored in the field for five consecutive years and it is clear that once the cyanobacteria blooms have reached a certain extent, their overall intensity only fluctuates modestly for a period of weeks despite potentially large spatial shifts in the spread of bloom density within the reservoirs. The blooms do appear to have a peak in intensity followed by a gradual dissipation. One single image might actually prove quite informative regarding changes in intensity over years if that image were acquired close to the seasonal bloom peak. The seasonal peak seems to occur towards the end of July to the middle of August, and unfortunately none of the available images were captured around this time. The review of these images therefore confirmed my intuition that they would not be effective at determining which year's bloom overall was more or less intense. The exercise does confirm the utility of the classification decision tree to be applied to different images and the ease with the results can be compared.

If the monitoring is aimed at determining when blooms have subsided, less frequent image acquisitions might be suitable, especially if the exact date at which the water body returns to non-bloom conditions is not needed. Looking at temporal trends within a season is also possible and the more dynamic the system the more frequent the image acquisition should be.

If the data are to be used for immediate decisions there will be at least a 2-5 day delay to allow time for processing and image delivery and then for image analysis once the decision trees have been developed.

Figure 5-1: True Color Landsat Images of the Copco and Iron Gate Reservoirs on Multiple Dates



August 23, 2005



August 31, 2005



Sept. 24, 2005



August 26, 2006

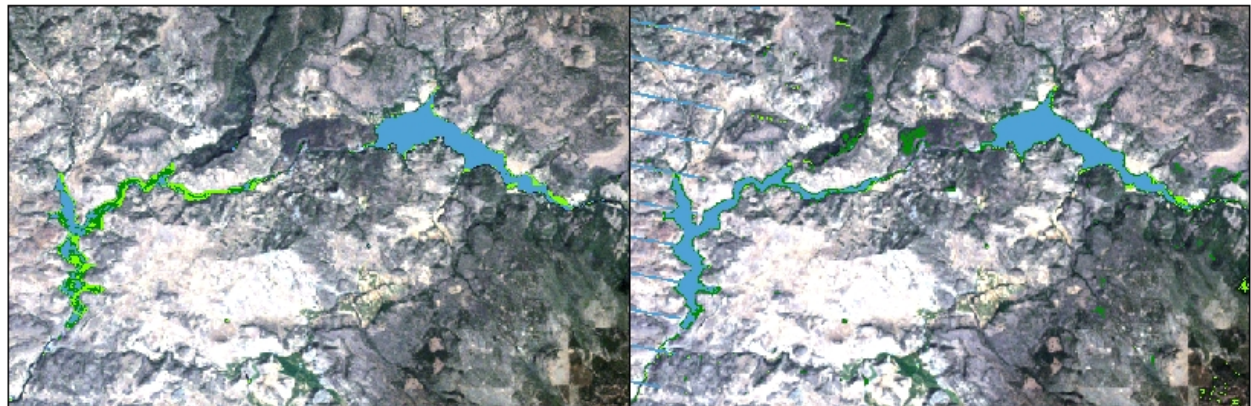


August 31, 2008



Sept. 3, 2009

Figure 5-2: Classified Landsat Images Showing Areas impacted by Low, Medium and High Density Cyanobacteria



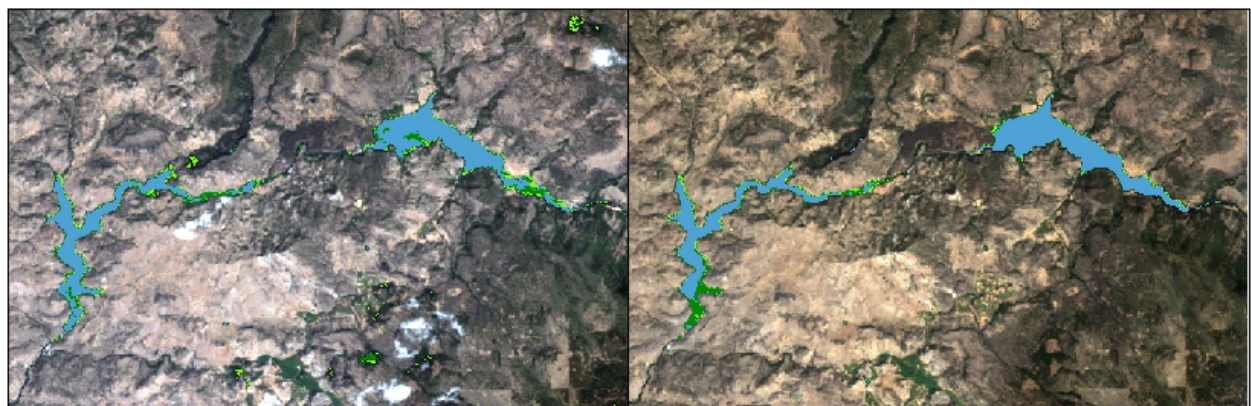
August 23, 2005

August 31, 2005






Sept. 24, 2005

August 26, 2006



August 31, 2008

Sept. 3, 2009

	Algae low
	Algae med
	Algae high

- Blanco, L. J., C. A. Ferrando, et al. (2009). "Remote Sensing of Spatial and Temporal Vegetation Patterns in Two Grazing Systems." Rangeland Ecology & Management **62**(5): 445-451.
- Chander, G., B. L. Markham, et al. (2009). "Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors." Remote Sensing of Environment **113**: 893-903.
- Ghioca-Robrecht, D. M., C. A. Johnston, et al. (2008). "ASSESSING THE USE OF MULTISEASON QUICKBIRD IMAGERY FOR MAPPING INVASIVE SPECIES IN A LAKE ERIE COASTAL MARSH." Wetlands **28**(4): 1028-1039.
- Karfs, R. A. and J. F. Wallace (2001). "An analysis of temporal change at rangeland monitoring sites using remote sensing in Northwest Australia." IEEE Transactions on Geoscience and Remote Sensing.
- Kennedy, R. E., P. A. Townsend, et al. (2009). "Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects." Remote Sensing of Environment **113**(7): 1382-1396.
- Li, X. and A. G. O. Yeh (1998). "Principal component analysis of stacked multi-temporal images for the monitoring of rapid urban expansion in the Pearl River Delta." International Journal of Remote Sensing **19**(8): 1501-1518.
- Masek, J. G., F. E. Lindsay, et al. (2000). "Dynamics of urban growth in the Washington DC metropolitan area, 1973-1996, from Landsat observations." International Journal of Remote Sensing **21**(18): 3473-3486.
- NASA Landsat Program (2005). Landsat TM Scene L5CPF20050701_20050930_04. Sioux Falls, USGS.
- R.S. Lunetta and C.D. Elvidge, Eds. (1999). Remote Sensing change detection: environmental monitoring, methods and applications Michigan, Taylor and Francis.
- Richards, J. A. (1984). "THEMATIC MAPPING FROM MULTITEMPORAL IMAGE DATA USING THE PRINCIPAL COMPONENTS TRANSFORMATION." Remote Sensing of Environment **16**(1): 35-46.

6 Conclusions and Future Studies

6.1 Conclusions

One of the objectives of this research was to evaluate the potential of remote sensing as a tool to be used in surveillance of toxic cyanobacteria blooms in freshwater. Remote sensing has been used extensively to predict, detect and study blooms in the ocean, but the relatively smaller size of blooms in freshwater has made it impractical to use many of the satellite systems that are suitable for ocean blooms. The results of this research show that even relatively coarse spectral resolution (4-7 different bands dividing the electromagnetic spectrum) is capable of providing information on the presence and density of cyanobacteria.

Those focused on the public health management of Klamath BGA, have relied on field sample collection twice a month over the bloom season (May – October). Samples are analyzed for microcystins and phytoplankton species are identified and quantified, with 48 hour turn around request once surface scums are visible in the reservoirs. Informational signs are posted at the start of the season as a precaution, and warning signs are posted as soon as scums are visible, or data on microcystin concentrations or microcystis cell counts exceed recommended guidelines of 8 ppb and 40,000 cells/l respectively. This approach has several limitations. First, is the question of whether the site measured is representative of the conditions in the whole water body. Since the blooms are moved by the wind and water currents, any particular site may be clear one day, and covered in dense scum the next. The public health sampling was therefore not limited to any fixed site (as opposed to the ecological trend monitoring sites), but rather samplers were instructed to visually assess the shores by foot or boat within reasonable time limits and collect samples within what appears to be most impacted location. Secondly, the samples were collected from the surface scum, but cyanobacteria can depth regulate and what is in the uppermost meter of the water column at any given time in the day, may be quite different 3, 6, or 12 hours later. To understand changes in the water column over time and to monitor trends, water over the whole photic zone should be collected, but this does not overcome the challenge of spatial variability. The use of satellite remote sensing data would provide synoptic coverage for a given point in time, but it could be cost prohibitive to have this data collected twice a month. Landsat satellites do pass by approximately twice a month (every 17 days), but currently there is often a long delay between collection and when the images are made available and not all images are ever even made available. Furthermore, shifts in population structure (between species or strains) or in toxin production would not necessarily be captured by the remote sensing data.

Even when the image is correctly classified and bloom material is correctly identified, it is possible that the bloom is either not cyanobacteria or that it is cyanobacteria but is nontoxic. As mentioned earlier (see Chapter 1), not all cyanobacteria are toxic and even those that do produce toxins can vary their level of toxin production in response to poorly understood environmental triggers. Though less likely there could be a situation where cell counts are high, but toxin production is low, or visa versa. In such a scenario the conclusions from remote data would err on the side of safety and precaution. This possibility affirms the need to verify remote classification results in the field. The spectral data interpretation should be used for making management decisions regarding the need for more monitoring and data collection but not the specific need for posting or press releases.

While the remote sensing cannot replace the field sampling in the Klamath given the way the data is used, it can be used to augment and direct the sampling that happens in the field. Specifically the synoptic data from remote sensing may be used to help better select sites on the ground where samples should be collected. It allows better overall understanding of the intensity of the bloom, and can be a good way of confirming when the bloom has fully subsided. In 2008, PacifiCorp worked with the company SolarBee to test the use of solar powered machines that use circulation of the water to try to reduce cyanobacteria density. Twelve SolarBee circulators (model SB10000v12-HW) were placed on the upstream end of one of the reservoirs and left in place for six months (April 17, 2008 through October 22, 2008) and an attempt was made to evaluate the effects of this remediation through the use of one single open water sampling site (Eilers and Walker 2010). The manufacturers of the circulators concluded that their installation resulted in a 99 percent reduction in *Microcystis aeruginosa* cell density and in total microcystin concentrations based on a comparison of the data from one sampling location in 2007 season versus 2008 season. The type of data available from remote sensing is well-suited for this type of research question. Remote sensing images of the whole water body from both seasons can be a more accurate way to evaluate the change in bloom intensity whether for the evaluation of the efficacy of remediation, shifts from other natural or anthropogenic environmental stressors, or the assessment of trends. The remote sensing data is particularly adept at monitoring the efficacy of longer-term management measures. 3-6 images within each season could effectively demonstrate bloom size, duration within a season, and an increase or decrease in bloom density between seasons.

Jernakoff 1996 (Jernakoff et al. 1996) also did a comparison of two data types. They used two airborne imaging systems: the CASI (Compact Airborne Spectrographic Imager), which has 288 bands in imaging spectrometer mode and 14 bands in spatial imaging mode and the DMSV (Digital Multi-Spectral Video) which has less spectral resolution but is less expensive and more available at short notice. Both imagers are carried aboard aircraft and therefore the spatial resolution is adjustable depending on the flight height. In this study the height of the aircraft allowed for 2 m x 2 m resolution. The CASI with higher spectral resolution provided greater accuracy and sensitivity to several different algal bloom types at low and moderate concentrations, but the 3 DMSV bands were still capable of separating the sites with high reliability.

In the research reported here, I was not able to test the efficacy of these remotely sensed data sets in detecting the presence of cyanobacteria in very low concentrations, as would be expected in pre-bloom conditions. Jupp (Jupp et al. 1994) asserts that a remote sensing system must be able to measure accurately down to 10 µg/l in turbid waters in order to be practical for this very purpose. Here it was not possible to look at the ability of the imagery to classify such low concentrations of algae because of a lack of ground truth data in these low ranges and a lack of low bloom conditions in the images.

From a broader perspective, the Intergovernmental Panel on Climate Change (IPCC) reported that the average global temperature will rise by an amount within the range of 1.1°-6.4° through the end of the 21st century depending on the emission scenario of greenhouse gases (Intergovernmental Panel on Climate Change (IPCC) 2007). Continuing releases of greenhouse

gases from human activities will perpetuate the decline in snow cover, shifts in plant and animal distributions, and ultimately changes in algal and fish communities and foodwebs. The predicted rise in average global temperature will bring about changes in the range for habitat and the migration patterns of many organisms. As this occurs the structure and function of ecosystems will be altered and so to the phytoplankton dynamics and trophic interactions – hence a trend for increasingly abundant cyanobacteria blooms which is already being seen. Given the rising abundance of cyanobacteria blooms and the myriad of both human health and ecological impacts from their proliferation, this research adds valuable contributions for future efforts to detect and manage them.

In California cyanobacterial blooms have been detected in non-marine water bodies, including Los Vaqueros and Mallard Reservoirs, the Sacramento River, the San Joaquin River, the Old River, Crowley Lake, Black Butte Lake, Clear Lake, the South Fork Eel River, Lake Hensley, Lake Isabella, Big Bear Lake, Perris Lake, Lake Elsinore, Canyon Lake, Pinto Lake, Lake Hennessey, Lake Britton, the Klamath River and its reservoirs, and in surface water components of the Metropolitan Water District of Southern California. Cyanobacterial blooms also occur in Big Lagoon, an estuary, and in the Salton Sea, an inland salt-water lake. To date the specific cyanotoxins identified in California include microcystins and anatoxin-a. A growing list of waterbodies in California are impacted by toxic BGA. Given the size of California, the list, long as it is, is still most certainly an underestimate of the true impact of toxic BGA. Without a formal monitoring and surveillance program in place nor specific requirements for drinking water districts the state functions at a reactionary level when blooms are detected unintentionally.

The results of this research support the use of remote sensing as a tool to detect the presence and intensity of cyanobacteria and to monitor temporal change in these blooms in response to natural or anthropogenic changes in the system. This in turn facilitates better public health protection and watershed management. Land uses and planning choices have the potential to impact water quality in a way that alters the ecosystem and phytoplankton community, and promotes cyanobacteria proliferation. Indeed, the research reported here supported the original premise that remote sensing is a tool to assess the risks to public health associated with land use and ecological change such as that produced by the construction of the Three Gorges Dam.

6.2 Policy Framework Contrast between California and China

The National Academy of Sciences published a book on the Klamath Basin in which they stated that less than 1% of the energy needs of PacifiCorp's customers is actually generated by the Klamath dams. Additionally the Klamath Reclamation area has low electricity rates so there is no motivation to reduce the pumping of irrigation water as a way to conserve both energy and water (Board on Environmental Studies and Toxicology and Water Science and Technology Board 2008). The Klamath River is listed as an impaired water body on the Clean Water Act (CWA) section 303(d) list for temperature, nutrients, dissolved oxygen and microcystin. The North Coast Regional Water Quality control Board (NCRWQCB) is in the process of finalizing and adopting a Total Maximum Daily Load (TMDL) for the Klamath River. The revised Staff Report and TMDL Action Plan were released in December 2009. PacifiCorp Energy applied for relicensing with the Federal Energy Regulation Commission on February 23, 2004 which included documentation on section 401 of the Clean Water Act. The State Water Board reviewed this and responded in February 2007 by stating that it was not clear that the relicensing

of PacifiCorp's hydroelectric PacifiCorp would protect water quality and additional studies were requested. An Agreement in Principle (AIP) to remove four Klamath River Dams was announced on November 13, 2008.

Hydroelectric dams often fail to meet expectations. While they do provide the benefits they were intended for, these are often at a much lower level than was anticipated (International River Network 2008). In 1996 for the first time the Federal Energy Regulatory Commission recommended the removal of a dam (Newport Dam #11 on the Clyde River in Vermont) for environmental reasons (Platt 2003). Increasingly in the U.S., where there are few unappropriated water supplies, dam decommissioning and dam removal are happening to benefit fish population recovery and to counteract any number of environmental impacts. In contrast, many other countries in the developing world are following the past model of the U.S. by building many dams on most of their rivers.

Many major tributaries along the Yangtze now contain high volumes of algae (Yan 2007) The first annual report on the Yangtze River protection was released in 2007 by the Nanjing Institute of Geography and Limnology, the Yangtze River Water Resources Commission, and the WWF, and it showed continued inflows of billions of tons of waste leaving more than 600 kilometers of the river in critical condition. The reservoir is polluted with pesticides, fertilizers and sewage from passenger boats (Xinhua China News Agency 2007). Varis reports that severe eutrophication problems are regularly observed in all major Chinese river systems, including the Yangtze (Varis and Vakkilainen 2001). China, however, lacks a unified water administration and management. Their transition to a market-based economy has also involved some decentralization of their government, but the water institutions are not necessarily all keeping up (Varis and Vakkilainen 2001).

When I contacted officials at Chinese environmental agencies for collaboration on water monitoring of harmful algal blooms and related dynamics in the Yangtze River, they communicated that their hands were tied when it came to research on issues that could be viewed as highlighting potential negative repercussions of the Three Gorges Dam construction. Nonetheless in October of 2006 a delegation of officials from China, who were all part of the Three Gorges Project Construction Committee and working on its management, came to California where they scheduled meetings with many state and federal agencies to learn about management and planning for U.S. dams. In a meeting with the U.S. EPA where I presented information regarding cyanobacteria in the Klamath reservoirs, they were particularly interested in learning about the risks and management of cyanobacteria. Although the sampling conducted as a part of this study did not detect cyanobacteria blooms in the Three Gorges Dam reservoir, harmful algal blooms have been reported on other parts of the Yangtze watershed and the delegation's keen interest in this issue may be indicative of their awareness of the growing issue on the Yangtze.

The findings from this research show that remote sensing data, even with coarse spectral resolution, can be used successfully to map the freshwater area impacted by blooms of various intensity. Periodic collection and classification of remote sensing can provide valuable information over the large area of the TGDR regarding where blooms may be occurring that is

otherwise unattainable in a site like the TGDR where the reservoir is so large and logistically difficult to access.

6.3 Future Studies

The results of this work lead to a number of other possible research questions that it would be helpful to explore in the future. One of these is working to develop a way to have real ground truth data collected that can truly represent field conditions at multiple sites throughout the water at the instant when the satellite image was collected. Part of the issue with correlating the field pigment concentration data with the pixel reflectance data is the diurnal temporality. Blooms rise to the water surface during the day and depending on local winds, water currents and their location in the water column they can be moved within the reservoir. The remote sensing imagery is collected for the entire area in a near instant in time, while 2 days were needed to collect the field samples. GPS coordinates were collected for the field sample sites (see Chapter 2), but as they were not at the exact hour when the imagery was collected large discrepancies exist between the concentrations of algae observed in the field versus those observed in the imagery. One way around this issue might be the use of real time monitoring probes placed throughout the reservoir. Those probes (available through YSI and Turner Designs) measure fluorescence in the field in real time, providing field data that can be assessed at precisely the time corresponding to image acquisition.

Other types of remotely sensed information might also be quite beneficial. For example the use of sensors (even simple cameras), mounted on posts at fixed heights over the water so that the field of view encompasses as much approximately 10 sq. meters (depending on the mounted height) can be programmed to capture images daily, weekly, monthly (or any other fixed interval) at the same time each day. With some automation, these images could quickly be used to evaluate impact, and to dictate when to collect field data for verification or further information such as speciation and toxin concentration.

Board on Environmental Studies and Toxicology and Water Science and Technology Board (2008). Hydrology, Ecology, and Fishes of the Klamath River Basin. Washington D.C., National Academies Press.

Eilers, J. and S. Walker (2010). An assessment of Circulation as a Means to Suppress Cyanobacteria and Cyanotoxins in Copco Reservoir, CA: Technical Memorandum, SolarBee.

Intergovernmental Panel on Climate Change (IPCC), P. M. I., Canziani O.F., Patulikof J. P., van der Linden P. J., Hanson C. E., (2007). Climate Change 2007—Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge, UK, Cambridge University Press: 976.

International River Network (2008). Dammed Rivers, Damned Lives: The Case Against Large Dams.

Jernakoff, P., P. Hick, et al. (1996). Remote Sensing of Algal Blooms in the Swan River. C. R. S. U. CSIRO Division of Fisheries, Division of Exploration and Mining, Waters and Rivers Commission, CSIRO Australia.

- Jupp, D. L. B., J. T. O. Kirk, et al. (1994). "Detection, Identification and Mapping of Cyanobacteria - Using Remote-Sensing to Measure the Optical-Quality of Turbid Inland Waters." Australian Journal of Marine and Freshwater Research **45**(5): 801-828.
- Pilotto, L. S., R. M. Douglas, et al. (1997). "Health effects of exposure to cyanobacteria (blue-green algae) during recreational water-related activities." Australian and New Zealand Journal of Public Health **21**(6): 562-566.
- Platt, J. (2003). Economic Analysis of Dam Decommissioning, U.S. Bureau of Reclamation.
- Pouria, S., A. de Andrade, et al. (1998). "Fatal microcystin intoxication in haemodialysis unit in Caruaru, Brazil." The Lancet **352**(9121): 21-26.
- Rabalais, N. N., R. E. Turner, et al. (2009). "Global change and eutrophication of coastal waters." ICES J. Mar. Sci. **66**(7): 1528-1537.
- Varis, O. and P. Vakkilainen (2001). "China's 8 challenges to water resources management in the first quarter of the 21st Century." Geomorphology **41**(2-3): 93-104.
- Xinhua China News Agency (2007). China report shows Yangtze water environment deteriorating. BBC Monitoring Asia Pacific. Changsha.
- Yan, A. (2007). Yangtze's overall water quality worsening, first -time report says. South China Morning Post.