USING BAND RATIO, SEMI-EMPIRICAL, CURVE FITTING, AND PARTIAL LEAST SQUARES (PLS) MODELS TO ESTIMATE CYANOBACTERIAL PIGMENT CONCENTRATION FROM HYPERSPECTRAL REFLECTANCE

Anthony Lawrence Robertson

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Science
in the Department of Earth Sciences,
Indiana University

July 2009

Accepted by the Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science.

_____

Lin Li, Ph.D., Chair

Master's Thesis
Committee

_____

Lenore P. Tedesco, Ph.D.

_____

Jeffery S. Wilson, Ph.D.

DEDICATION

I would like to dedicate this work to my parents. Both my mom and dad encouraged my love of science.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my fiancée, Karen Nguy. She has always supported my research endeavors and listened when I needed her the most.

I appreciate the assistance and hard work of my graduate committee, especially my graduate chair, Dr. Lin Li. I have learned much from all three members.

I am grateful for the expertise and assistance in the lab from Lani Pascual.

I appreciate the technical support that Bob Hall was able to contribute to the research project. He helped to solve difficulties and assisted in making the field days run as smooth as possible.

I am also thankful to both Kate Randolph and Becky Sengpiel for their hard work in developing the original data set upon which I was able to develop a portion of my models.

I appreciate the assistance of Lara Vallely who made the processing of the 2007 data set much more tolerable, and we were able to teach each other about remote sensing of water.

The assistance provided in the field by Lani Pascual, Bob Hall, Lara Vallely, Vince Hernly, Nicolas Clercin, Zuoning Jiang, Jonathan Bills, Jeremy Webber, Ryan McAfee, Jake Lemon and the other staff, graduate students and interns has been tremendously helpful in facilitating the data collection.

I am also thankful to the interns–Andrew Johnson, Morgan Barrett, Amanda Fulmer, Brian McInerny, Lauren Payne, Michael Turpin, Kendall Sharp–that have assisted me in the processing of the data sets. I hope they enjoyed the experience of learning how to conduct basic research.

I am grateful for having the privilege of using laboratory equipment, greenhouse space and technical assistance provided the Biology staff members Rick Frey, Kim Burrows, and Dawn Bauman.

I appreciate the assistance provided by the Center for Advanced Land Management Information Technologies (CALMIT), specifically Bryan Leavitt, for their technical assistance in operating the CALMIT Data Acquisition Program (CDAP).

I would like to thank the faculty, staff and students in the Department of Earth Science for their assistance in helping me with my presentations.

I appreciate the financial support of Veolia Water Indianapolis, Central Indiana Water Resources Partnership, Department of Earth Science, Graduate Office, Graduate Student Organization, and the School of Science at IUPUI.

ABSTRACT

Anthony Lawrence Robertson

USING BAND RATIO, SEMI-EMPIRICAL, CURVE FITTING, AND PARTIAL LEAST SQUARES (PLS) MODELS TO ESTIMATE CYANOBACTERIAL PIGMENT CONCENTRATION FROM HYPERSPECTRAL REFLECTANCE

This thesis applies several different remote sensing techniques to data collected from 2005 to 2007 on central Indiana reservoirs to determine the best performing algorithms in estimating the cyanobacterial pigments chlorophyll *a* and phycocyanin. This thesis is a set of three scientific papers either in press or review at the time this thesis is published. The first paper describes using a curve fitting model as a novel approach to estimating cyanobacterial pigments from field spectra. The second paper compares the previous method with additional methods, band ratio and semi-empirical algorithms, commonly used in remote sensing. The third paper describes using a partial least squares (PLS) method as a novel approach to estimate cyanobacterial pigments from field spectra. While the three papers had different methodologies and cannot be directly compared, the results from all three studies suggest that no type of algorithm greatly outperformed another in estimating chlorophyll *a* on central Indiana reservoirs. However, algorithms that account for increased complexity, such as the stepwise regression band ratio (also known as 3-band tuning), curve fitting, and PLS, were able to predict phycocyanin with greater confidence.

Lin Li, Ph.D.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AISA	airborne imaging spectrometer for applications sensor
AOP	apparent optical properties
CALMIT	Center for Advanced Land Management Information Technologies
CDAP	CALMIT Data Acquisition Program
CEES	Center for Earth and Environmental Science
CIWRP	Central Indiana Water Resources Partnership
CDOM	color-dissolved organic matter
CHL	chlorophyll *a*
DNR	Department of Natural Resources
ECR	Eagle Creek Reservoir
FWHM	full width half maximum
GFF	glass fiber filters
GR	Geist Reservoir
IOP	inherent optical properties
ISS	inorganic suspended solids
IUPUI	Indiana University – Purdue University Indianapolis
LARE	Lake and River Enhancement
MGM	modified Gaussian model
MNR	Monroe Reservoir
MR	Morse Reservoir
NIR	near infrared
PC	phycocyanin
RMS	root means square
UV	ultraviolet
WHO	World Health Organization

I: INTRODUCTION

In central Indiana, surface drinking waters experience high nutrient input from local agriculture. These nutrient inputs increase the duration and frequency of cyanobacterial, also known as blue-green algae, blooms. Veolia Water Indianapolis, Veolia Environment and the Center for Earth and Environmental Science (CEES) at Indiana University – Purdue University Indianapolis (IUPUI) formed the Central Indiana Water Resources Partnership (CIWRP) to address water quality issues, including these problematic cyanobacterial blooms. Research conducted by CIWRP focuses on three area reservoirs: Geist (GR), Eagle Creek (ECR), and Morse (MR). Studies of these three reservoirs include, but are not limited to, nutrient cycling, ongoing reservoir and watershed monitoring, and remote sensing of algae/cyanobacteria. This thesis will present the latest developments in the remote sensing studies conducted by CIWRP.

The initial remote sensing project conducted through CIWRP began in 2005 and was funded by CIWRP in conjunction with the Indiana Department of Natural Resources (DNR) through a Lake and River Enhancement (LARE) grant. This project collected samples across all three reservoirs: GR, ECR, and MR. The goal was to build models to predict the cyanobacterial pigments chlorophyll *a* (CHL) and phycocyanin (PC) from all three reservoirs with both field (ASD Field Spec) and airborne (AISA-Eagle) spectra. Results from the 2005 study are detailed in Li *et al.* (2006), Li *et al.* (2009), Randolph (2007), Randolph *et al.* (2008), and Sengpiel (2007). Although the data from ECR was unusable due to calibration issues, the initial results were promising for the remote detection of cyanobacterial pigments CHL and PC.

Due to the initial success, CIWRP received funding from Veolia Water Indianapolis to continue the remote sensing research project. For the 2006 field season, temporal differences between the three reservoirs were studied. This temporal variability

added complexity to the models, which was shown to limit the transferability of these models. Vallely (2008) analyzed the 2006 data to indentify factors that reduce the models' performance in estimating CHL and PC from each of the reservoirs. Vallely (2008) found that different water constituents influence the performance of the algorithms on different reservoirs. Vallely (2008) concluded that aggregated data sets provide variability necessary for improved performance of the models.

In 2007 a new sensor, Ocean Optics USB 4000, was purchased. The new sensor was set up in a dual-headed system designed to reduce atmospheric noise. Fewer samples were collected in 2007; however, data were collected on a new reservoir, Monroe (MNR). None of the 2007 data had been published prior to this author's work and will be summarized in this thesis with the exception of the data collected from MNR.

This thesis is a collection of three journal articles regarding the development and application of new and existing approaches to estimating cyanobacterial pigments on GR, ECR, and MR. Each chapter represents a stand-alone article; hence, there is some repetition in the background information and data analysis. A summary of the work presented in the following articles can be found in chapter five. While chapters two, three, and four are to be published in scientific journals, the final published versions may differ from the one found in this thesis.

II: USING A MODIFIED GAUSSIAN MODEL TO PREDICT CYANOBACTERIAL PIGMENT ABUNDANCE IN CENTRAL INDIANA RESERVOIRS

Abstract

Algal pigments phycocyanin (PC) and chlorophyll *a* (CHL) have been used to remotely detect cyanobacteria. This study evaluates an alternative empirical modeling technique using a modified Gaussian model (MGM) to predict the algal pigments PC and CHL in eutrophic reservoirs using field-based spectra. MGM was applied both spatially, using three different reservoirs, and temporally, using data sets from 2005 and 2006. An MGM model based on a 2005 data (PC n = 16; CHL n = 15) from one reservoir in the study region (Morse) resulted in a strong correlation coefficient to both pigments of interest (PC $R^2 = 0.93$; CHL $R^2 = 0.82$). These models, when applied to a multi-temporal data set created in 2006 on the same reservoir, showed reasonable predictive ability (PC: n = 78 $R^2 = 0.77$, RSME = 52.5 μg/L, $p < 0.01$; CHL: n = 79 $R^2 = 0.58$, RSME = 23.5 μg/L, $p < 0.01$). The MGM models for phycocyanin prediction consistently outperformed the traditional spectral indices tested in this paper. These results suggest that MGM can be utilized as an alternative to traditional empirical band-ratio algorithms.

Introduction

Cyanobacteria in surface water systems pose a health concern for humans, livestock, and native wildlife across the globe. Cyanobacteria tend to form blooms in late summer in eutrophic systems during long periods of hot, dry days (Falconer 2005). Although the eutrophication of some water sources is caused by natural processes, this concern is compounded by nutrient loading from sewage, agricultural and other anthropogenic sources (Chorus and Bartram 1999). Some species of cyanobacteria

produce taste-altering chemicals (Chorus and Bartram 1999; Christensen *et al.* 2006) and toxins (Bold and Wynne 1985; Chorus and Bartram 1999) that affect water quality. Due to the degradation of water quality caused by these taste-altering and toxic compounds, surface water managers, health officials, and researchers are interested in rapid identification of cyanobacteria blooms.

Early detection methods, using remotely sensed data, have focused on the estimation of light-harvesting pigments in cyanobacteria. As an indicator of total algal biomass (including both cyanobacteria and algae), chlorophyll *a* is widely used for this purpose due to strong absorption at 660-666 nm (Jeffrey *et al.* 1997; Jeffrey and Wright 2006; Rowan 1989). Despite a strong absorption feature, the spectral range of 428-432 nm is not commonly used in remotely sensed data due to overlapping absorptions of carotenoids in the 400-500 nm range (Jeffrey *et al.* 1997; Rowan 1989). In addition to these two bands of major absorption, chlorophyll *a* also absorbs less strongly around 382.7 nm, 409-417.6 nm, 530-535.5 nm, 575-580.3 nm and 614-618.2 nm (Jeffrey *et al.* 1997).

Another pigment found in cyanobacteria is the phycobiliprotein phycocyanin (Jeffrey and Wright 2006; Rowan 1989). Cyanobacterial phycocyanin has strong absorption in the region of 612-628 nm along with a florescence maximum in the region of 632-651 nm (Rowan 1989). As phycocyanin is not typically found in other algal classes except marine red algae (Rhodophycae) (Delwiche 1999), it is used in inland remote sensing studies to estimate cyanobacteria abundance (Dekker 1993; Gitelson *et al.* 1995; Randolph *et al.* 2008; Schalles and Yacobi 2000; Simis *et al.* 2005; Simis *et al.* 2007).

For studying cyanobacteria, the most commonly used methods are the empirical and semi-empirical approaches that develop spectral indices sensitive to cyanobacterial pigments and examine the correlation between the indices to pigment concentration (e.g., Gitelson *et al.* 1986; Mittenzwey *et al.* 1991; Schalles and Yacobi 2000). These algorithms are effective for open ocean (case I) waters, where turbidity is caused primarily by phytoplankton. When applied to inland and coastal (case II) waters, however, empirical models are influenced by sediments and other non-algal water constituents not commonly found in case I waters (Schalles and Yacobi 2000). Previous research has suggested that differences in absorption and backscattering properties water constituents can limit the transferability of some algorithms across different inland aquatic systems (Dekker 1993; Randolph *et al.* 2008). For example, analyses conducted by Randolph *et al*. (2008) on the reservoirs examined in the current study concluded that band-ratio algorithms had difficulty transferring between two reservoirs located in close proximity with similar characteristics and watersheds; however, there are anthropogenic disturbances (dredging and groundwater inputs) that may cause spectral band shifts and limit transferability.

In this study we used a modified Gaussian model (MGM) as an alternative empirical modeling technique for the estimation of chlorophyll *a* and phycocyanin concentrations in surface water reservoirs. Since MGM is a curve-fitting approach, it can separate overlapping spectral features. The objective of this research was to examine the correlation between MGM parameters (full width half maximum (FWHM), strength, area) and algal pigment (chlorophyll *a* and phycocyanin) concentration. On the basis of

this correlation analysis, MGM transferability was evaluated with datasets collected in multiple years for a single reservoir and datasets collected from multiple reservoirs.

Background

Since the absorption features of many materials are inherently Gaussian in shape, equation 2-1 can be used to describe these absorption spectra (Sunshine *et al.* 1990). The Gaussian distribution (g) of a random variable or, in the case of reflectance spectra, the energy absorbed (x), is expressed in terms of its strength or amplitude (s), center or mean (μ), and width or standard deviation (σ). These sets of Gaussian distributions are overlaid onto continua that are linear functions of energy (Sunshine *et al.* 1990).

$$g(x) = s * \exp\{-(x - \mu)^2 * (2\sigma^2)^{-1}\} \qquad \text{(eq. 2-1)}$$

This model was able to give an estimate of the location of absorption peaks; however, this model is not able to provide an accurate method of identifying and quantifying multiple substances within reflectance spectra. Also it was noted that if multiple initial estimates are entered into the equation, then it is possible to have varying results (Sunshine *et al.* 1990).

In order to improve upon the Gaussian model, Sunshine *et al.* (1990) reported that absorption energy and bond length are related by a power law, thus providing a modified equation relating Gaussian curves of absorption spectra to the bond length of the molecule absorbing energy. Sunshine *et al.* (1990) derived a power function empirically that resulted in the modified Gaussian distribution (m) in equation 2-2.

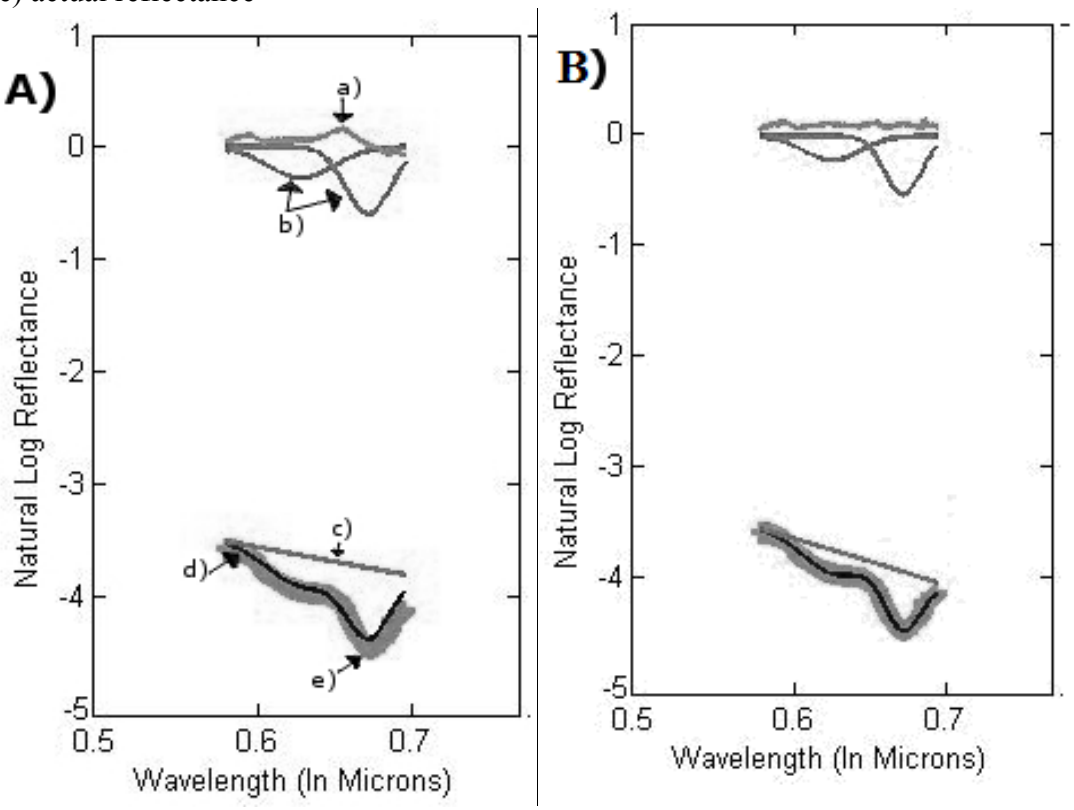
$$m(x) = s * \exp\{-(x^{-1} - \mu^{-1})^2 * (2\sigma^2)^{-1}\} \qquad \text{(eq. 2-2)}$$

The modified Gaussian model (MGM) assumes that each absorption feature of interest can be represented by a Gaussian function. When these curves are combined,

they are matched to a spectral signature provided by the researcher (Sunshine 1999; Sunshine *et al.* 1990). The MGM software presented by Sunshine (1999) is a free download from Brown University (http://www.planetary.brown.edu/mgm/) and is based on the nonlinear least squares algorithm developed by Kaper *et al.* (1966). Users input reflectance along with an initial set of parameters including the width, strength, and centers of Gaussian curves. This set of parameters can be visualized using Matlab (The MathWorks, Inc.) (Figure 2-1). Starting with initial parameters, MGM iteratively adjusts the parameters to best fit the reflectance spectra, until the improvement in root mean square error (RMSE) between the modeled and actual reflectance spectra is less than 1.0e-6 (Figure 2-1b).

MGM has been applied to analyzing extraterrestrial mineral composition (Hiroi *et al.* 2000; Mustard *et al.* 2005; Pieters and McFadden 1994), and more recently in analyzing pigments (Combe *et al.* 2005; Lohrenz *et al.* 2003). Combe *et al.* (2005) applied MGM to differentiate a type of algae (microphytobenthos dominated by diatoms) from other features using Digital Airborne Imaging Spectrometer (DAIS) data. The initial parameters were set by evaluating the curve shape of each absorption produced by a batch of accessory pigments (435 nm), diadinoxanthin (500 nm), fucoxanthin (550 nm), chlorophyll *c* (623 nm), chlorophyll *a* (675 nm), cell structure (750 nm, 798 nm, and 886 nm), and water (980 nm). The study was successful in differentiating between the microphytobenthos of interest and macroalgae. Lohrenz *et al.* (2003) used MGM to determine spectral absorption of chlorophyll in the 400-700 nm spectral range for samples of coastal waters collected on glass fiber filters. The strength of the Gaussian peak was then used to quantify chlorophyll *a*, chlorophyll *b* and chlorophyll *c*. The study

Figure 2 - 1. MGM fits to a hyperspectral reflectance spectrum: A) Initial conditions programmed into the MGM software B) Program Output after 257 iterations and an RMS of 1.39 x 10^{-2} and an improvement of < 1.0 x 10^{-6}.
Where:
a) RMS residual error between predicted and actual reflectance
b) predicted Gaussian curves
c) continuum
d) predicted reflectance using current Gaussian parameters
e) actual reflectance



reported a hyperbolic relationship between pigment concentration and Gaussian peak strength attributed to pigment packaging with either an increase in internal pigment concentration or an increase in algal cell size. Quantification of the algal pigments was conducted using reverse phase High Performance Liquid Chromatography (HPLC), with a reported standard error of 8%. The study reported an influence of accessory pigments, specifically carotenoids on the estimation of chlorophyll *b* and *c*. Although both of these studies proved that algal pigment concentrations can be estimated by applying MGM to

hyperspectral reflectance data, no study, as far the authors are aware, have applied MGM to predicting algal pigments from field based spectra.

Methods

Study Site

Three reservoirs located in central Indiana were investigated: Eagle Creek, Geist, and Morse. Figure 2-2 shows the location of each reservoir. These reservoirs provide water for >800,000 residents of the greater Indianapolis area. All three reservoirs have similar characteristics including depth (3.2-4.7 m), surface area (5-7.5 km^2), volume (21-28 million m^3) and residence time (55-70 days) (Li *et al.* 2006; Randolph *et al.* 2008). They are also impaired by high nutrient loads (mean total P = 94-100 $\mu gP\text{-}L^{-1}$, mean total N = 2-4.1 $\mu gP\text{-}L^{-1}$) (Li *et al.* 2006; Randolph *et al.* 2008) which promote the growth of nuisance algae, including cyanobacteria. All three reservoirs have reported taste and odor issues (Li *et al.* 2006; Randolph *et al.* 2008).

<u>In situ</u> Reflectance Measurement

Collection dates, number of samples collected, and spectral signatures are listed in Table 2-1. The sensors used in this study were two ASD Field Spec ultraviolet/visible and near-infrared spectroradiometers (Analytical Devices, Inc., Boulder, CO, USA). Both spectroradiometers have a spectral resolution of 1 nm. The spectral range of the spectroradiometer used on Geist Reservoir 2005 was 350-1050 nm for a total of 701 bands. For all other *in situ* reflectance measurements the ASD spectroradiometer had a spectral range of 350-2500 nm for a total of 2151 bands.

Figure 2 - 2. Central Indiana reservoirs: Map of the three study locations; Eagle Creek, Geist, and Morse Reservoirs. Indianapolis is indicated for reference.

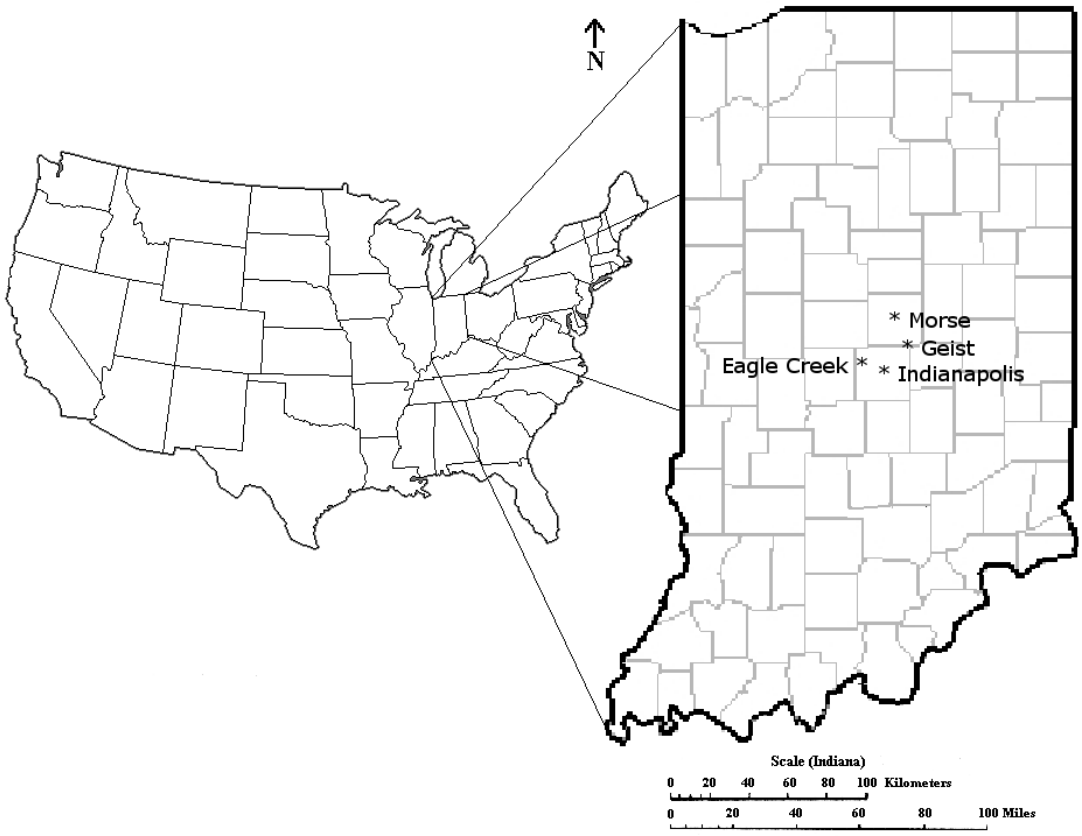


Table 2 - 1. Summary of sample collection: Sampling sites with the date ranges, number of samples, and pigment ranges (μg/L).

Sample Location	Date Range	Sample Number	Chlorophyll *a* Minimum	Chlorophyll *a* Maximum	Phycocyanin Minimum	Phycocyanin Maximum
Eagle Creek	July - Sept. 2006	85	6.3	107.5	0.7	234.3
Geist	Sept. 2005	25	34.7	118.9	30.8	185.1
	July - Sept. 2006	88	23.1	182.6	2.6	210.2
Morse	Sept. 2005	22	18.0	151.7	2.9	135.1
	July - Sept. 2006	91	21.3	128.7	3.3	371.0

Spectra were collected on a boat oriented facing the sun, at least 10 m from shore, and in water depth greater than 2 m to minimized potential effects of bottom reflectance on spectra. The fiber optic cable has an instantaneous-field-of-view (IFOV) of 0.17 rad. The cable was mounted on an extendable pole with a nadir viewing angle and held approximately 0.5 m above the water surface. This produces a measured water surface

area with a diameter of 0.04 m. Calibration for upwelling irradiance was completed by using a white reference panel. In order to reduce noise in the spectra, the reflectance spectra at each site was averaged over 15 readings.

Lab Analysis of Water Samples

Water samples were collected from surface water with 1 L amber HDPE bottles and stored on ice prior to filtration, and analyzed for pigment concentration. Steps for pigment analysis were completed under subdued light conditions. Water samples were filtered in duplicate within 24 hours of the sample collection for later extraction of pigments. Samples for the extraction of chlorophyll *a* were prepared on 0.47 µm acetate filters, and those for the extraction of phycocyanin on 0.47 µm glass fiber filters (GFF). The samples were stored at -20°C for no longer than 6 months before analysis.

Chlorophyll *a* was extracted following the extraction method described in Environmental Protection Agency (EPA) 445 (Arar and Collins 1997). The concentration of chlorophyll *a* was corrected for pheophytin and measured fluormetrically using a TD-700 Fluorometer (Turner Designs, Inc.) fitted with a Daylight White Lamp and Chlorophyll Optical Kit (340-500 nm excitation filter and emission filter > 665 nm). For chlorophyll *a* analysis the fluorometer was calibrated using chlorophyll *a* from a spinach standard (Sigma-Aldrich 10865).

Phycocyanin was extracted based on a method described in Sarada (1999) and presented in Randolph *et al.* (2008). GFFs were suspended in 15 mL of 50 mM phosphate buffer (pH = 6.8). Samples were broken up using a stainless steel spatula and rinsed with 5 mL of 50 mM phosphate buffer. The samples were then homogenized using Teflon coated pestle. Pestles were rinsed with 5 mL of 50 mM phosphate buffer.

Samples were centrifuged at 4°C, 15,000 x g for 25 minutes using a Beckman J2-21M centrifuge. The samples were then stored overnight at 4°C before being homogenized again using the Teflon coated pestle and rinsed with 5 mL of 50 mM phosphate buffer with a total volume of 30 mL. The samples were centrifuged again prior to the collection of the supernatant. The supernatant was measured fluormetrically for phycocyanin using a TD-700 Fluorometer (Turner Designs, Inc.) fitted with a Cool White Mercury Vapor Lamp and a Phycocyanin Optical Kit (630 nm excitation and 660 nm emission filters). For phycocyanin analysis the fluorometer was calibrated using C-phycocyanin from *Spirulina sp.* (Sigma-Aldrich P6161).

If the percent error calculated between a pair of replicates in pigment analysis was larger than 20%, the sample was not used in data analysis.

Estimation of Pigment Abundance with MGM

MATLAB (The MathWorks, Inc.) programs for MGM were modified from Sunshine *et al.* (1999) in order to process multiple files simultaneously. Initial parameters inputted into MGM (Table 2-2) were set based on previous algal pigment studies (Jeffrey *et al.* 1997; Rowan 1989). To minimize influence of the carotenoids, two spectral features were analyzed; phycocyanin (~628 nm) and chlorophyll (~672 nm). Several MGM parameters (FWHM, strength, area) were analyzed for their correlation to the algal pigment (chlorophyll *a* and phycocyanin) concentration.

Table 2 - 2. Initial Parameters used for MGM Modeling: Continuum intercept: 5.00 x 10^{-2}, slope: -1.00 x 10^{-6}.

Parameter	Center of Absorption (nm)	Full-Width Half Maximum (FWHM)	Absorption Strength
Phycocyanin	628	60	-0.4
Chlorophyll *a*	672	40	-0.7

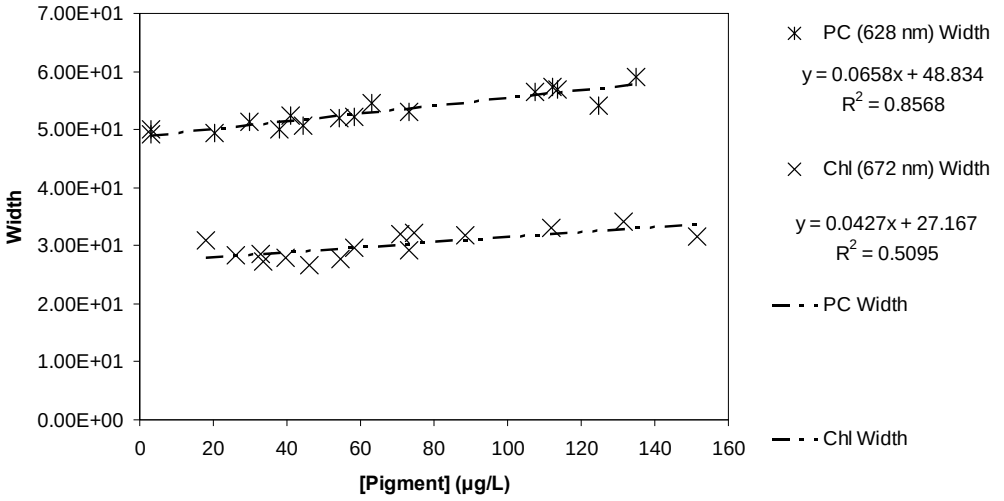
Only the spectra where the gap between the continuum and fitted reflectance is visibly minimal are included in model construction and validation. MGM was used to fit two absorption features between 560 - 710 nm, which are caused by phycocyanin (~628 nm) and chlorophyll *a* (~670 nm). MGM was then used to deconvolve reflectance spectra into individual Gaussian curves. MGM iterations stopped when the root mean square error (RMSE) values or improvement between two iterations was less than 1.0 x 10^{-6}.

Results

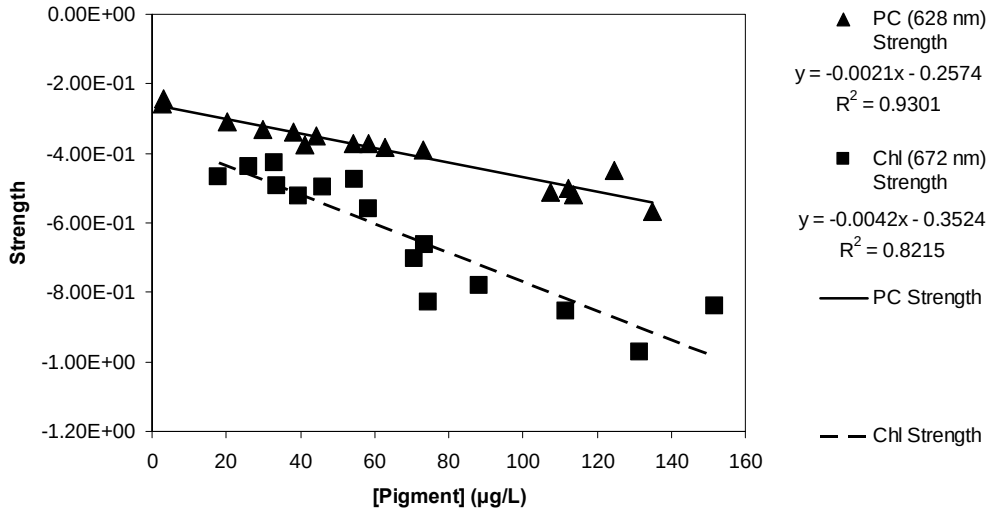
Building Calibration Models with the Morse 2005 Dataset

A series of calibration models were created using the 15 samples collected from Morse Reservoir in 2005 by correlating MGM output parameters (Figure 2-3) to pigment concentration. A summary of the models can be found in Table 2-3. The correlation coefficients of all the models were greater than 0.72 with the exception of the model based on the MGM output parameter for FWHM in predicting chlorophyll *a*.

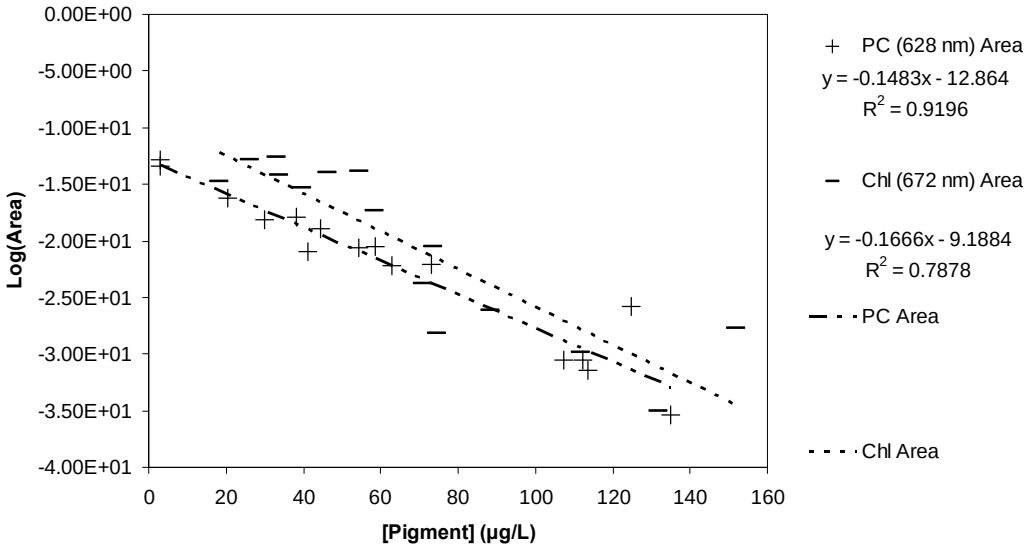
Figure 2 - 3. Building models using the Morse 2005 data set and MGM output: Correlation between pigments of interest, phycocyanin (PC), chlorophyll *a* (CHL), and MGM output a) FWHM b) Strength c) Area.



a)



b)



c)

Table 2 - 3. Models built from MGM output and spectral indices: Where x = spectral parameter, and y = predicted pigment concentration.

Data Set	Pigment	Sample Size (n)	Spectral Parameter	R^2	Model
Morse 2005	Phycocyanin	16	MGM PC FWHM	0.8568	Y = 15.20x-742.16
			MGM PC Strength	0.9301	Y = -238.10 -83.90
			MGM PC Area	0.9196	y = -125x - 1608
			Simis *et al.* 2005 Band Ratio	0.8703	y = 125x - 84.29
	Chlorophyll *a*	15	MGM Chl FWHM	0.5095	Y = 23.42 - 636.23
			MGM Chl Strength	0.8215	y = -476.19 - 122.57
			MGM Chl Area	0.7878	y = -6.00 - 55.15
			Mittenzwey *et al.* 1991	0.7283	y = 188.68 - 5.04
Eagle Creek & Morse 2006	Phycocyanin	127	MGM PC Strength	0.6343	Y = -454.55 - 95.27
			Simis *et al.* 2005 Band Ratio	0.7207	y = -714.29 - 184.93
	Chlorophyll *a*	127	MGM Chl Strength	0.5348	y = -250 - 80.18
			Mittenzwey *et al.* 1991	0.6893	y = 181.82 + 2.55

Pigment Estimation with Morse 2005 Models

Once the models were created, they were applied to additional data sets to test both temporal and spatial transferability. The data is summarized in Table 2-4. Optimal models will have a slope and coefficient of determination (R^2) near one, and a low RMSE

Table 2 - 4. Morse 2005 MGM models predicting additional datasets: Intercept was forced through zero, therefore ideal models will have a 1) slope near one 2) correlation coefficient (R^2) near one and 3) low RMSE.

Data Set	Pigment	Sample Size	Model	Slope	R^2	RMSE (µg/L)	p
Geist 2005	Phycocyanin	27	MGM PC FWHM	1.5820	0.5511	81.39	< 0.01
			MGM PC Strength	1.1088	0.6788	26.84	< 0.01
			MGM PC Area	1.2399	0.6733	39.32	< 0.01
			Simis *et al.* 2005 Band Ratio	0.8401	0.1736	27.81	> 0.1
	Chlorophyll *a*	26	MGM Chl FWHM	0.9222	< 0.05	35.30	> 0.1
			MGM Chl Strength	1.1206	< 0.05	29.07	> 0.1
			MGM Chl Area	1.0557	< 0.05	28.05	> 0.1
			Mittenzwey *et al.* 1991	1.0539	< 0.05	22.87	> 0.1
Eagle Creek 2006	Phycocyanin	52	MGM PC FWHM	0.3760	0.4051	87.82	< 0.01
			MGM PC Strength	0.6459	0.5809	53.15	< 0.01
			MGM PC Area	0.5902	0.6610	56.29	< 0.01
			Simis *et al.* 2005 Band Ratio	0.5022	0.6416	59.74	< 0.01
	Chlorophyll *a*	51	MGM Chl FWHM	0.6101	0.0769	40.86	> 0.1
			MGM Chl Strength	0.9200	0.4429	23.55	< 0.01
			MGM Chl Area	0.8573	0.5065	19.26	< 0.01
			Mittenzwey *et al.* 1991	0.9630	0.5926	20.97	< 0.01
Geist 2006	Phycocyanin	88	MGM PC FWHM	1.0231	0.0946	69.51	> 0.1
			MGM PC Strength	0.2503	0.0795	56.48	> 0.1
			MGM PC Area	0.3754	0.0933	72.02	> 0.1
			Simis *et al.* 2005 Band Ratio	0.5072	< 0.05	53.40	> 0.1
	Chlorophyll *a*	88	MGM Chl FWHM	0.7181	0.2778	35.70	> 0.1
			MGM Chl Strength	0.3026	< 0.05	76.01	> 0.1
			MGM Chl Area	0.3754	< 0.05	57.69	> 0.1
			Mittenzwey *et al.* 1991	0.6116	0.2260	32.59	< 0.05
Morse 2006	Phycocyanin	78	MGM PC FWHM	0.6979	0.6487	71.79	< 0.01
			MGM PC Strength	0.6991	0.7654	52.45	< 0.01
			MGM PC Area	0.7293	0.8196	47.99	< 0.01
			Simis *et al.* 2005 Band Ratio	0.5003	0.3850	71.43	< 0.01
	Chlorophyll *a*	79	MGM Chl FWHM	0.6930	0.2986	40.18	< 0.01
			MGM Chl Strength	0.8137	0.5840	23.47	< 0.01
			MGM Chl Area	0.7854	0.5949	22.80	< 0.01
			Mittenzwey *et al.* 1991	0.8859	0.7410	16.31	< 0.01
Eagle Creek and Morse 2006	Phycocyanin	127	MGM PC FWHM	0.5865	0.5597	78.36	< 0.01
			MGM PC Strength	0.6807	0.7207	52.72	< 0.01
			MGM PC Area	0.6812	0.7738	51.35	< 0.01
			Simis *et al.* 2005 Band Ratio	0.5010	0.4743	66.89	< 0.01
	Chlorophyll *a*	127	MGM Chl FWHM	0.6696	0.2309	40.44	< 0.02
			MGM Chl Strength	0.8438	0.5251	23.50	< 0.01
			MGM Chl Area	0.8058	0.5648	21.53	< 0.01
			Mittenzwey *et al.* 1991	0.9078	0.6775	18.33	< 0.01

When the models based on the Morse 2005 data set were applied to the Geist 2005 data set, all the chlorophyll *a* models performed poorly ($R^2 < 0.05$). The models predicting phycocyanin performed better, based on the coefficient of determination ($0.55 < R^2 < 0.68$).

When these models were applied temporally to the Morse 2006 data set all models under predicted pigments ($0.70 < \text{slope} < 0.89$) but had reasonable coefficients of determination ($0.65 < R^2 < 0.82$) for phycocyanin prediction. However, the models based on MGM output parameters had weaker coefficients of determination ($0.30 < R^2 < 0.59$) for predicting chlorophyll *a*.

When applied both spatially and temporally to the 2006 Eagle Creek data set, only one model based on the MGM output for chlorophyll FWHM was not statistically significant. The models based on the 2005 Morse data which were statistically significant when validated on the 2006 Eagle Creek data set under predict chlorophyll *a* ($0.86 < \text{slope} < 0.92$) and had poor coefficients of determination ($0.44 < R^2 < 51$). Alternatively, the phycocyanin models had higher coefficients of determination when compared to their respective model equivalent ($0.41 < R^2 < 0.66$).

When applied to the Geist 2006 data set, only one model was statistically significant. While the model built from FWHM for predicting chlorophyll *a* was statistically significant ($p < 0.05$), the model performed poorly using coefficients of determination ($R^2 = 0.28$).

Since the goal was to build a model that can be extended spatially and temporally, the models developed with the Morse 2005 data set were applied to a combined Eagle Creek and Morse 2006 data set. Again, all models under predicted pigments ($0.59 <$

slope < 0.84). Similarly to applying the models to the Morse 2006 data set exclusively, the models had reasonable coefficients of determination ($0.59 < R^2 < 0.68$) for estimating phycocyanin. Additionally, the coefficients of determination ($0.23 < R^2 < 0.56$) for predicting chlorophyll *a* were lower.

Building Models with the Eagle Creek and Morse 2006 Dataset

Since the Morse 2005 data set was relatively small (n = 16), and the range of pigments was narrow (Chl: 18.0-151.7; PC: 2.9-135.1 μg/L), models based on a much larger data set using data from both Eagle Creek and Morse 2006 data were created (Table 2-3). The correlations of the pigments of interest to MGM strength (Figure 2-4) were determined. This set of models reduced extrapolation for phycocyanin estimation since there was a greater range of phycocyanin concentrations in the larger calibration data set (0.7-371.0 μg/L). Models based on the MGM output parameters FWHM and area were excluded in this analysis because FWHM does not perform as well as strength in this study. Since the models based on MGM area are a function of FWHM, they were also excluded. Geist was excluded from this model because of inherent differences to the other two reservoirs (discussed below).

Pigment Estimation with Eagle Creek and Morse 2006 Models

The results of applying the models built from the Eagle Creek and Morse 2006 data set are summarized in Table 2-5. When this model was applied spatially over the temporally variable Geist 2006 data set, no model was statistically significant. However, when applied to the Morse 2005 data set, these models performed with higher coefficients of determination ($0.81 < R^2 < 0.93$) than when the Morse 2005 models were applied to the Eagle Creek and Morse 2006 data set. The model for estimating

chlorophyll *a* calibrated with the Eagle Creek and Morse 2006 data set predicted pigments with a near linear correlation (slope = 1.13) when validated to the Morse 2005 data set. However, models for phycocyanin estimation over predicted (slope = 1.49) when applied to the same data set.

Figure 2 - 4. Building models using the Eagle Creek and Morse 2006 data set and MGM output: Correlation between pigments of interest, phycocyanin (PC), chlorophyll *a* (CHL), and MGM output Strength.

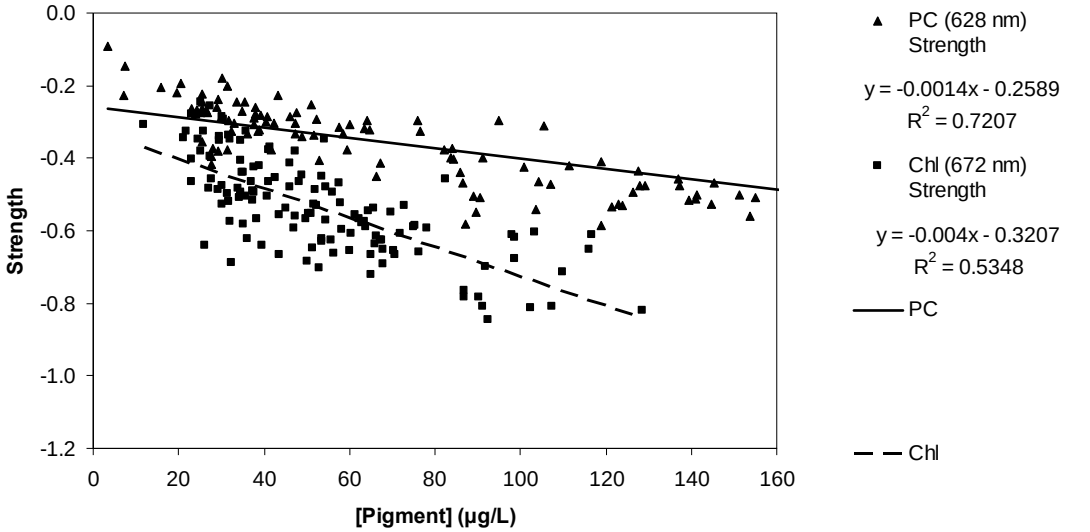


Table 2 - 5. Eagle Creek and Morse 2006 models predicting additional datasets: Intercept was forced through zero, therefore ideal models will have a 1) slope near one 2) correlation coefficient (R^2) near one and 3) low RMSE.

Data Set	Pigment	Sample Size	Model	Slope	R^2	RMSE (µg/L)	p
Geist 2005	Phycocyanin	27	MGM Strength	1.7543	0.5953	88.34	< 0.01
			Simis *et al.* 2005	1.7891	0.5059	97.69	< 0.01
	Chlorophyll *a*	26	MGM Strength	1.315	< 0.05	37.22	> 0.1
			Mittenzwey *et al.* 1991	1.117	< 0.05	25.53	> 0.1
Morse 2005	Phycocyanin	16	MGM Strength	1.4936	0.9301	41.35	< 0.01
			Simis *et al.* 2005	2.0824	0.7793	101.92	< 0.01
	Chlorophyll *a*	15	MGM Strength	1.1343	0.8136	21.59	< 0.01
			Mittenzwey *et al.* 1991	1.0561	0.7211	23.29	< 0.01
Geist 2006	Phycocyanin	88	MGM Strength	0.3660	0.0802	83.33	> 0.1
			Simis *et al.* 2005	0.8307	< 0.05	42.57	> 0.1
	Chlorophyll *a*	88	MGM Strength	0.4171	< 0.05	56.10	> 0.1
			Mittenzwey *et al.* 1991	0.6892	0.1045	28.03	> 0.1

When the Eagle Creek and Morse 2006 models were applied to the Geist 2005 data set, the chlorophyll *a* model was not statistically significant. The phycocyanin model over predicted, but had reasonable coefficient of determination ($R^2 = 0.60$). Oversaturation of the sensor to the sensitivity of phycocyanin concentration may have contributed this observation. When all phycocyanin concentrations above 100 µg/L are removed, as summarized in Table 2-6, the coefficients of determination were high ($R^2 = 0.90$), although the pigments were greatly over predicted (slope = 2.29). Removing all points greater than 100 µg/L in the Geist 2006 data set did not yield statistically significant coefficients of determination.

Table 2 - 6. Performance of phycocyanin models based on the Eagle Creek and Morse 2006 dataset when points greater than 100 µg/L are removed: Intercept was forced through zero, therefore ideal models will have a 1) slope near one 2) correlation coefficient (R^2) near one and 3) low RMSE.

Model	Slope	R^2	RMSE (µg/L)	P
MGM Strength	2.2857	0.9037	87.96	< 0.01
Simis *et al.* 2005	2.3353	0.7382	97.69	< 0.01

Estimation of Pigment Abundance with Band Ratios

While this study demonstrates that MGM can be used in estimating pigments from field spectra, in order to determine if MGM performs at least as well as traditional methods the exact same data sets were also tested with two traditional band ratios. Both ratios presented here were chosen over other ratios due to consistent performance on all three central Indiana reservoirs. Equation 2-3 was developed by Mittenzwey *et al.* (1991) and used to estimate chlorophyll *a*. Equation 2-4 adapted from Simis *et al*. (2005) was used to estimate phycocyanin.

$$[R(705) - R(670)] * [R(550)^{-1}] \qquad \text{(eq. 2-3)}$$

$$[R(709)] * [R(620)^{-1}] \quad \text{(eq. 2-4)}$$

Calibration models using the spectral indices were also built (Figure 2-5 & 2-6; Table 2-3). Results are summarized in Table 2-4, 2-5, & 2-6.

Discussion

Performance of MGM

Overall, the MGM models were able to estimate both phycocyanin and chlorophyll *a*. The MGM based models yielded higher coefficients of determination (R^2) in predicting phycocyanin over chlorophyll *a*. Using coefficients of determination and slope as an indicator of performance, MGM based models consistently outperformed the spectral index for predicting phycocyanin. There was no consistent difference between MGM based models and the spectral index for chlorophyll *a*.

One reason MGM models do not outperform the spectral index for predicting chlorophyll *a* may be related to how MGM fits spectral curves. Although this study only examined the concentration of chlorophyll *a*, there are additional forms of chlorophyll that were not quantified. MGM output is likely correlated to all forms of chlorophyll since it has the capability to adjust its center, FWHM, and strength. Due to this additional variability, the accuracy of the models in predicting chlorophyll *a* could be reduced. Although the absorption of phycocyanin and chlorophyll overlap, (Jeffrey *et al.*1997; Rowan 1989) as shown in Figure 2-7, its accuracy was minimally affected due to MGM's ability in unmixing spectra.

Figure 2 - 5. Building models using the Morse 2005 data set and spectral indices: Correlation between pigments of interest, phycocyanin (PC), chlorophyll *a* (CHL), and spectral indices

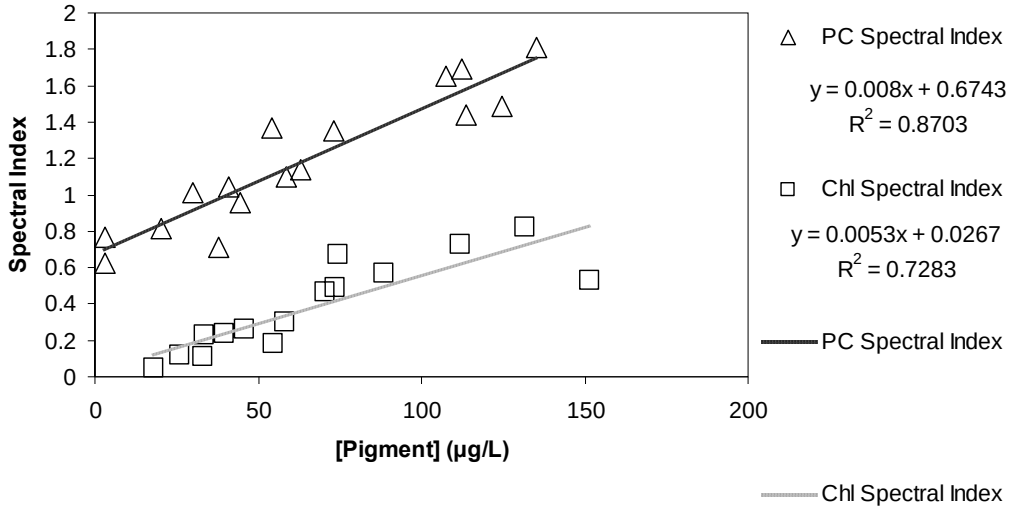


Figure 2 - 6. Building models using the Eagle Creek and Morse 2006 data set and spectral indices: Correlation between pigments of interest, phycocyanin (PC), chlorophyll *a* (CHL), and spectral indices

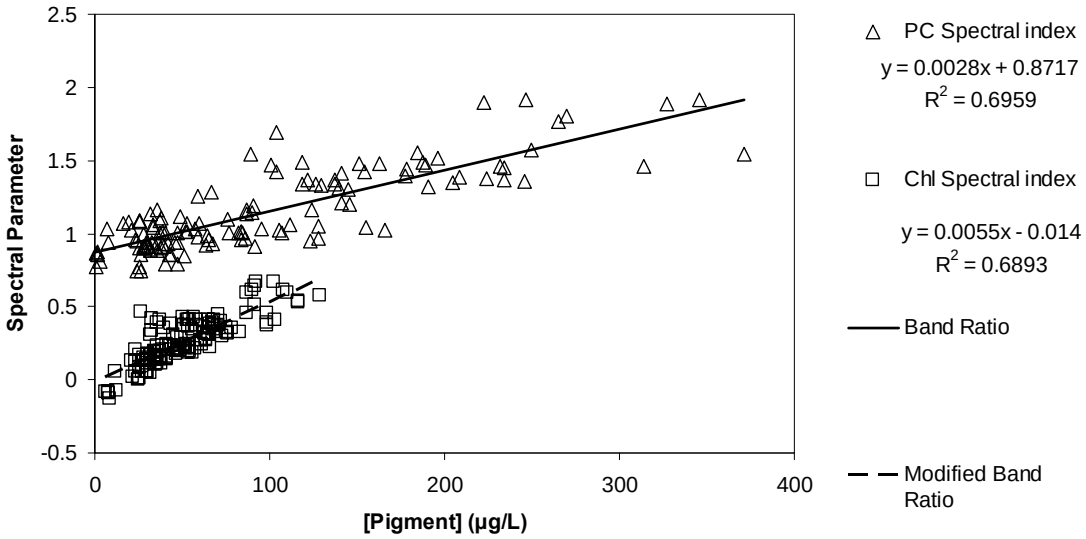
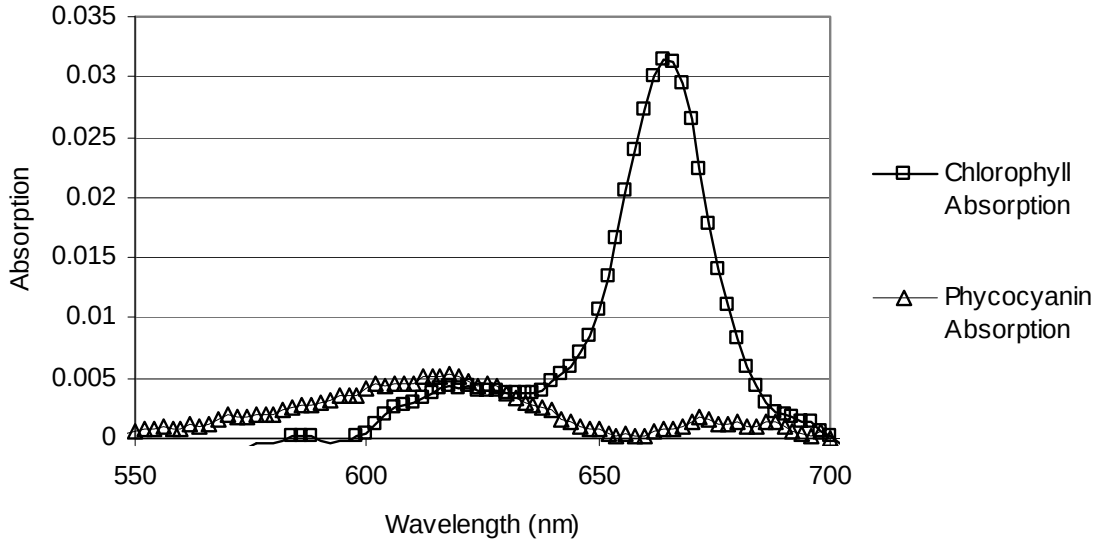


Figure 2 - 7. Absorption of extracted pigments: Absorption measured on a Shimadzu spectrophotometer. Chlorophyll *a* (32.1 µg/L) was extracted in 90% buffered acetone and phycocyanin (188.3 µg/L) was extracted in 50 mM phosphate buffer.



The models based on MGM strength output consistently had higher coefficients of determination and slopes closer to one than the output FWHM. Neither the models based on MGM strength nor the models built on MGM area consistently outperformed the other. There are two reasons why MGM area occasionally outperformed MGM strength. First, while MGM strength is typically correlated to increases in pigment concentration, this correlation may be reduced at increased pigment concentrations due to spectral saturation. At these higher pigment concentrations MGM strength would increase only slightly compared to increases in FWHM. Alternatively, MGM area is a function of both MGM FWHM and MGM strength, the occasional outperformance of MGM area could be due to correlation with the MGM strength output. Either models using MGM strength or MGM area are recommended for further study.

The models based on the Eagle Creek and Morse 2006 data set outperformed the models built with the Morse 2005 data set. For phycocyanin estimation, this was likely

due to extrapolation of the models created with the Morse 2005 data set. The range of pigments found in the Eagle Creek and Morse 2006 data set was greater than most of the other data sets, the exception being higher chlorophyll *a* concentrations. This suggests that the calibration data set should include a larger number of samples in order to better represent the natural variation in the reservoirs and minimize extrapolation of the model.

Limitations using MGM

From this study we found that there are two major limitations with using MGM based models. First, MGM models are currently more time consuming to implement when compared to traditional spectral indices. While this may limit MGM based models, this issue can be overcome with more efficient programming and faster computers.

Second, there are some concerns with the transferability of these models based on Eagle Creek and Morse data sets to Geist data sets. While all three reservoirs are drinking water reservoirs, Geist has additional anthropogenic disturbances. A local aggregate mining company routinely dredges Geist for silt and sand particles and an estimated ten-million gallons of water is pumped back into the reservoir (Constantino 1999). Because of the increased suspended sediment from dredging operations it is possible that this is causing saturation of the pigment signal at lower concentrations, meaning that as pigment concentration increases, reflectance spectra remains minimally changed. Additionally, the water input back into the reservoir may be causing a localized dilution effect which locally disrupts the correlation of pigment to reflectance. This not only causes problems for developing a MGM output-based model on Geist Reservoir, but also for developing a spectra index-based model.

Conclusions

Summary of MGM modeling

This paper shows that MGM is applicable to field based spectra. In general, MGM based models predicted phycocyanin more accurately than chlorophyll *a*. Phycocyanin is more commonly used as an indicator of cyanobacterial species. Since water managers are more interested in identifying cyanobacteria from all other algal species, MGM based models may be a better tool over spectral indices. There are still some limitations on how MGM based models can be applied across spatially distinct data sets. However, since this is an issue with both spectral indices and MGM based models, additional approaches should be considered.

Future Work

To address transferability, additional models, such as stepwise regression, bio-optical and semi-empirical models, should be tested on these data sets. MGM should also be compared to a larger series of spectral indices, since different spectral indices may work better in removing influences from confounding factors (such as sediments). Also these models should be tested for transferability across sensors, such as additional field-based, aerial, and satellite sensors. If there is interest in improving chlorophyll estimation, all forms of chlorophyll should be quantified.

Acknowledgements

We would like to thank Veolia Water Indianapolis and the Central Indiana Water Resources Partnership for financial support. We would also like to thank Lani Pascual, Kate Randolph, Becky Sengpiel, Lara Vallely, Bob Hall, Vince Hernly, and other staff, graduate students and interns involved with the collection and processing of the data sets.

References

Arar, E.J., & Collins, G.B. (1997). In Vitro Determination of Chlorophyll *a* and Pheophytin a in Marine and Freshwater Algae by Fluorescence. In. Cincinnati, OH.

Bold, H.C., & Wynne, M.J. (1985). *Introduction to the Algae: structure and reproduction*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Chorus, I., & Bartram, J. (1999). *Toxic Cyanobacteria in Water: A guide to their public health consequences, monitoring and management*. Bury St Edmunds, Suffolk: St. Edmundsbury Press.

Christensen, V.G., Graham, J.L., Milligan, C.R., Pope, L.M., & Ziegler, A.C. (2006). Water quality and relation to taste and odor compounds in the North Fork Ninnescah River, and Cheney Reservoir, south-central Kansas. In, *U.S. Geologcial Survey Scientific Investigations Report 2006.*

Combe, J.-P., Launeau, P., Carrere, V., Despan, D., Meleder, V., Barille, L., & Sotin, C. (2005). Mapping microphytobenthos biomass by non-linear inversion of visible-infrared hyperspectral images. *Remote Sensing of Environment, 98*, 371-387.

Constantino (1999). In the Lap of Luxury. *Pit and Quarry* Cleveland, OH.

Dekker, A.G. (1993). Detection of the optical water quality parameters for eutrophic waters by high resolution remote sensing. In. Ansterdam: Free University

Delwiche, C.F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *The American Naturalist, 154*, S164-177.

Falconer, I.R. (2005). *Cyanobacterial Toxins of Drinking Water Supplies*. Boca Raton, FL: CRC Press.

Gitelson, A.A. (1992). The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. *International Journal of Remote Sensing, 13*, 3367 – 3373.

Gitelson, A.A., Laorawat, S., Keydan, G.P., & Vonshak, A. (1995). Optical properties of dense algal cultures outdoors and their application to remote estimation of biomass and pigment concentration in *Spirulina Platensis* (cyanobacteria). *Journal of Phycology, 31*, 828-834.

Gitelson, A., Nikanorov, A., Sabo, G., & Szilagyi, F. (1986). Etude de la qualite des eaux de surface par teledetection. *IAHS Publications, 157*, 111-121.

Hiroi, T., Taylor, L.A., Morris, R.V., Keller, L.P., McKay, D.S., & Wentworth, S. (2000). Initial Results of MGM Analysis on Apollo 17 Soil Suite. In, *Lunar and Planetary Science XXXI(CD_ROM), abstract 1880, Houston, Texas.*

Jeffrey, S.W., Mantoura, R.F.C., & Wright, S.W. (1997). *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods*. Paris: UNSECO.

Jeffrey, S.W., & Wright, S.W. (2006). *Photosynthetic Pigments in Marine Microalgae: Insights from Cultures and the Sea*. Enfield, NH: Science Publishers.

Kaper, H.G., Smits, D.W., Schwarz, U., Takakuro, K., & Woerden, H.V. (1966). Computer Analysis of Observed Distributions into Gaussian Components. *Bull. Astr. Inst. Netherlands, 18*, 465-487.

Li, L., Pascual, D.L., Tedesco, L.P., Randolph, K.L., Sengpiel, R.E., Hall, B.E., & Wilson, J.S. (2006). Developing a survey tool for the rapid assessment of blue-green algae in central Indiana's reservoirs. In, *2005 LARE Grant Report*: Indiana Department of Natural Resources.

Lohrenz,S. E., Weidemann, A.D., & Tuel, M. (2003). Phytoplankton spectral absorption as influenced by community size structure and pigment composition. *Journal of Plankton Research, 25*, 35-61.

Mittenzwey, K.-H., Breitwieser, S., Penig, J., Gitelson, A.A., Dubovitzkii, G., Garbusov, G., Ullrich, S., Vobach, V., & Müller, A. (1991). Fluorescence and Reflectance for the *in-situ* Determination of Some Quality Parameters of Surface Waters. *Acta Hydrochimica et Hydrobiologica, 19*, 1-15.

Mustard, J.F., Poulet, F., Gendrin, A., Bibring, J.-P., Langevin, Y., Gondet, B., Mangold, N., Bellucci, G., & Altieri, F. (2005). Olivine and Pyroxene Diversity in the Crust of Mars. *Science, 307*, 1594-1597.

Pieters, C.M., & McFadden, L.A. (1994). Meteorite and Asteroid Reflectance Spectroscopy: Clues to Early Solar System Processes. *Annual Review of Earth and Planetary Sciences, 22*, 457-497.

Randolph, K.L., Wilson, J.S., Tedesco, L.P., Lin, L., Pascual, D.L., & Soyeux, E. (2008). Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll *a* and phycocyanin *Remote Sensing of Environment, 112*, 4009-4019.

Rowan, K.S. (1989). *Photosynthetic pigments of algae*. New York, New York: Cambridge University Press.

Sarada, R.M., Pillai, G., & Ravishankar, G.A. (1999). Phycocyanin from *Spirulina sp.*: influence of processing of biomass on phycocyanin yield, analysis of efficacy of extraction methods and stability studies on phycocyanin. *Process Biochemistry, 34*, 795-801.

Schalles, J.F., & Yacobi, Y.Z. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid, and chlorophyll pigments in eutrophic waters. *Archives Hydrobiologica Special Issues on Advanced Limnology, 55*, 153-168.

Simis, S.G.H., Peters, S.W.M., & Gons, H.J. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography, 50*, 237-245.

Simis, S.G.H., Ruiz-Verdu, A., Dominguez, J.A., Pena-Martinez, R., Peters, S.W.M., & Gons, H.J. (2007). Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sensing of Environment, 106*, 414-427.

Sunshine, J.M., C.M. Pieters, S.F. Pratt and K.S. McNaron-Brown (1999). Absorption Band Modeling in Reflectance Spectra: Availability of the Modified Gaussian Model. *30th Annual Lunar and Planetary Science Conference. Abstract 1306*. Houston, TX.

Sunshine, J.M., Pieters, C.M., & Pratt, S.F. (1990). Deconvolution of Mineral Absorption Bands: An Improved Approach. *Journal of Geophysical Research, 95*, 6955-6966.

III: COMPARISON BETWEEN EMPIRICAL, SEMI-EMPIRICAL, AND CURVE FITTING MODELS IN PREDICTING CYANOBACTERIAL PIGMENTS

Abstract

This study presents a comparative analysis of several algorithms for the estimation of cyanobacterial pigments chlorophyll *a* (CHL) and phycocyanin (PC). This study seeks to provide a consistent basis for determining the best performing models in predicting both CHL and PC concentrations on three eutrophic central Indiana reservoirs from data collected in 2005-2007. Simple band ratio algorithms, band tuning methods, semi-empirical algorithms, and the modified Gaussian model (MGM) parameters were used to estimate CHL and PC from multiple source spectral datasets of three eutrophic central Indiana reservoirs: Eagle Creek, Geist and Morse. The spectral datasets were collected over a three year period (2005-2007) using two field-based (ASD Field Spec, Ocean Optics USB4000) and the sensor Airborne Imaging Spectrometer for Application (AISA-Eagle). Spectral parameters resulting from these mapping algorithms were examined for their correlation to the CHL and PC concentrations. The results demonstrate no major performance difference between most of the CHL estimating algorithms and a simple band ratio approach should be favored to use in most cases. However, for PC estimation, complex models such as 3-band tuning model and the MGM model consistently outperformed band ratio and semi-empirical algorithms tested in this study. While the MGM model performed equivalently well as the band tuning method when applied to the field measured spectral data, the latter is favorable when applied to the AISA imaging spectra. This recommendation is based on the observation that the

MGM models did not perform well for estimating PC when applied to the AISA-Eagle spectra due to the coarser spectral resolution than the field measured spectra.

Introduction

Surface water managers are interested in identifying cyanobacterial blooms since their occurrence in surface drinking water supplies and recreational waters pose a public health threat (Chorus and Bartram 1999; Christensen *et al.* 2006) and toxins (Bold and Wynne 1985; Chorus and Bartram 1999). While cyanobacteria naturally develop in late summer due to long periods of hot, dry days (Falconer 2005), anthropogenic eutrophication of surface waters has increased the frequency and duration of cyanobacterial blooms (Chorus and Bartram 1999). This has led to increased closures to surface waters for recreational access and increased costs for removing and neutralizing taste/odor and toxic compounds produced by cyanobacteria. To minimize the effects of cyanobacterial blooms water managers are interested in a rapid identification system to monitor these blooms. Remote sensing provides an effective approach to this monitoring.

Cyanobacterial pigments, chlorophyll *a* and phycocyanin, have been used to evaluate inland water quality and identify cyanobacterial blooms. (Dekker 1993; Gitelson *et al.* 1995; Li *et al.* in press; Randolph *et al.* 2008; Richardson 1996; Schalles 2006; Schalles and Yacobi 2000; Simis *et al.* 2005; Simis *et al.* 2007). Commonly scientists use chlorophyll *a* to estimate total algal biomass and phycocyanin is used as an indicator of cyanobacteria biomass. Remote sensing is capable of monitoring chlorophyll a and phycocyanin by utilizing a strong absorption of chlorophyll *a* at 660-666 nm that is not influenced by accessory pigments (Jeffrey *et al.* 1997; Jeffrey and Wright 2006; Rowan 1989) and the strong absorption of phycocyanin in the region of 612-628 nm

(Rowan 1989). Among numerous remote sensing algorithms for estimating chlorophyll *a* and phycocyanin concentrations, empirical methods using spectral ratios, semi-empirical using the absorption coefficient of the pigments and curve-fitting using the Gaussian function are commonly used and reported in literature.

Simple Band Ratio

Simple band ratios are one of the most common empirical methods used in remotely mapping algal pigments (Dekker 1993; Gitelson *et al.* 1995; Randolph *et al.* 2008; Schalles 2006; Schalles and Yacobi 2000; Simis *et al.* 2005; Simis *et al.* 2007). Simple band ratios are created by dividing a band that is sensitive to changes in pigment concentration (band 1) by a spectral band that is insensitive to changes in pigment concentration (band 2) (equation 3-1).

$$\text{Simple band ratio} = R_{band1}/R_{band2}. \quad \text{(eq. 3-1)}$$

The main reason why band ratios are favored is their ease of use. However, as pointed out by both Strombeck and Pierson (2001) and Simis *et al.* (2005) these ratios are based on several assumptions. Ratios commonly use reflectance near 705-709 nm since this wavelength range is more sensitive to changes in pigment concentrations and it is assumed that the increase in scattering caused by cell walls represents an increase in the algal population (Simis *et al.* 2005; Strombeck and Pierson 2001). Also assumptions are made that the absorption of other water constituents such as color dissolved organic matter (CDOM) and non-algal suspended matter (tripton) does not contribute significantly to reflectance (Simis *et al.* 2005; Strombeck and Pierson 2001). However, studies have shown that varying concentrations of tripton, CDOM, and algae can cause

band shifts and interfere with the performance of band ratio algorithms (Han 1997; Han and Rundquist 1994; Strombeck and Pierson 2001; Vallely 2008).

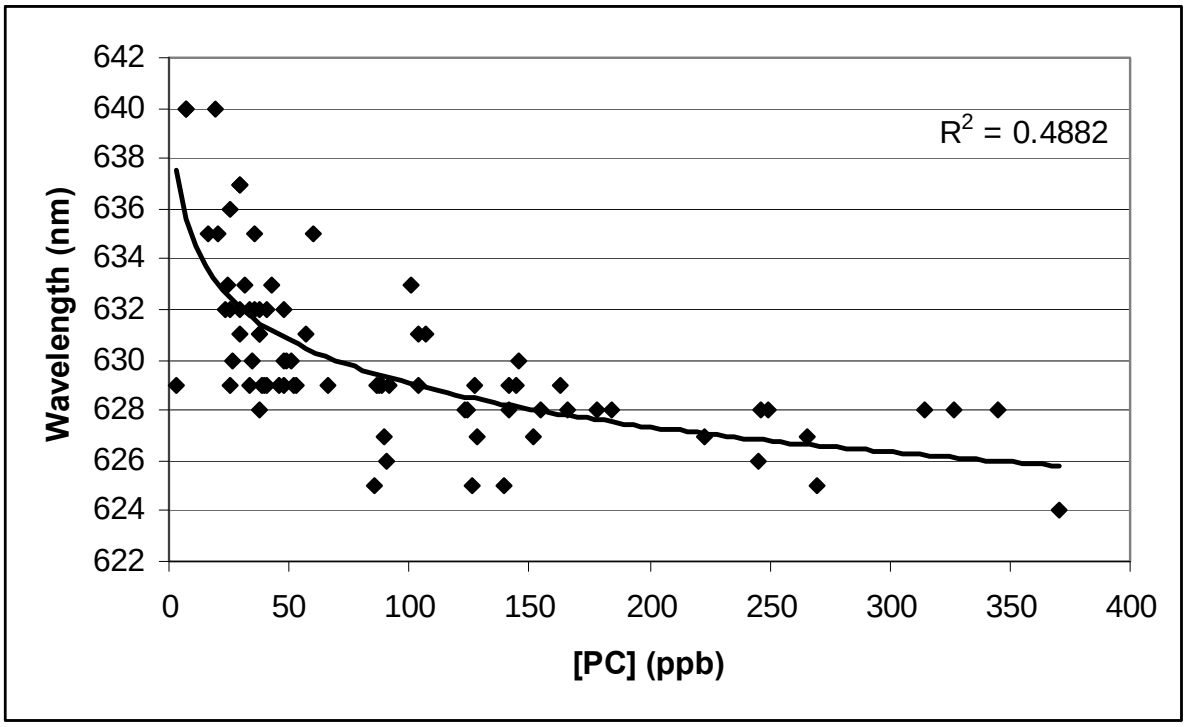
Maximum Peak/Troughs

Gitelson *et al.* (2008) discussed many band ratio algorithms for estimating cyanobacterial pigments are ratios of near infrared (NIR) and red reflectance and stated that the success of these algorithms depend on the selection of the NIR wavelength. In an attempt to address the effects of band shifts, simple band ratios created to detect chlorophyll *a* have been modified to identify peaks within a specific wavelength range, specifically the NIR peak near 700 nm (Gitelson 1992; Schalles *et al.* 1998; Yacobi *et al.* 1995), while the wavelength corresponding to chlorophyll *a* absorption (670 nm) was fixed resulting in equation 3-2 which is correlated to chlorophyll *a* concentrations (Gitelson 1992; Yacobi *et al.* 1995). The reflectance trough for chlorophyll *a* does not have to remain fixed and can also be adjusted (Randolph 2007, Vallely 2008).

$$R_{max(680\text{-}720)}/R_{670} \quad \text{(eq. 3-2)}$$

In addition to band shifts caused by chlorophyll *a* in the 700 nm range, research conducted by the authors suggests that phycocyanin concentration can also cause band shifts in the reflectance trough near 620 nm (Figure 3-1). Band ratio algorithms can easily be adjusted to account for these spectral band shifts by locating the maximum peak and trough values in ranges of interest such as the reflectance troughs for phycocyanin (~620 nm), and reflectance peak near 700 nm corresponding to cell scattering (Randolph 2007; Vallely 2008).

Figure 3 - 1. Phycocyanin concentration compared to central wavelength of reflectance trough near 628 nm: n = 78



Band Tuning

One weakness with the approach that utilizes maximum peak and trough reflectance values is the possibility of saturation of the reflectance signal at the wavelengths at relatively low pigment concentrations, resulting in a constant reflectance with increasing pigment concentrations as reported with chlorophyll *a* (Dekker 1993; Yacobi *et al.* 1995). To address this issue and the concern with the sensitivity of simple band ratios to the wavelength selection in the NIR region, Dall'Olmo *et al.* (2003) applied step-wise regression to select the wavelengths of highest correlation to chlorophyll *a* with reasonable accuracy in preliminary results (RMSE < 13 µg/L) and the stepwise regression or band tuning model was validated in Dall'Olmo and Gitelson (2005) and Gitelson *et al*. (2008). Gitelson *et al.* (2005) proposed a similar three-band

tuning model with slightly different starting and stepwise regression conditions. Both methods use the spectra index described by equation 3-3 after determining the three wavelengths with the highest correlation.

$(1/R_{\lambda 1} - 1/R_{\lambda 2})*R_{\lambda 3}$ (Dall'Olmo *et al.* 2003, Gitelson *et al.* 2005) (eq. 3-3)

Gitelson *et al.* (2008) explains that $R_{\lambda 1}$ should be maximally sensitive to the absorption of the pigment of interest. However, $R_{\lambda 1}$ is affected by the influence of the absorption properties of tripton, color dissolved organic matter (CDOM), and water, and the scattering properties of particulate matter (Gitelson 2008). Thus, $R_{\lambda 2}$ should be a wavelength which has minimum sensitivity to the pigment of interest and contain similar absorption values of tripton and CDOM found in $R_{\lambda 1}$ (Gitelson 2008). $R_{\lambda 3}$ should be in a region that is sensitive to the backscattering primarily caused by tripton (Gitelson 2008).

Semi-Empirical Models

In another attempt to address the weakness of band ratios, researchers developed semi-empirical models which are based on band ratios; however, the models also include inherent optical properties such as absorption and backscattering coefficients since these properties are neglected in traditional band ratios (Simis *et al.* 2005; Simis *et al.* 2007). Simis *et al.* (2005) suggests that the traditional band ratio algorithm for phycocyanin estimation would lead to an underestimation of absorption caused by phycocyanin with an increase in pigment concentration. Alternatively, since tripton and CDOM are considered negligible, an overestimation of the absorption caused by phycocyanin can occur with increasing tripton and CDOM (Simis *et al.* 2005). The model proposed by Simis *et al.* (2005) performs well when cyanobacteria are the dominant algal species and

the absorption coefficient for phycocyanin is calculated and applied for each sample (R^2 = 0.94); however, when a fixed absorption coefficient is used, overestimation occurred in several of the cruises and only one subset performed with reasonable confidence (R^2 = 0.77). Since phycocyanin estimation is also influenced by absorption of chlorophyll that results in over estimation of phycocyanin, Simis *et al.* (2007) applied a semi-empirical algorithm that can correct for absorptions caused by chlorophyll *a*, in addition to CDOM, and tripton. Simis *et al.* (2007) validated their earlier model for phycocyanin extraction by improving extraction methods to reduce overestimation and most of the cruise trips were able to predict phycocyanin reasonably well ($R^2 > 0.7$). By including the inherent optical properties of these water constituents into the semi-empirical model these effects can be reduced or eliminated. A simpler equation representing the semi-empirical model is displayed in equation 3-4, where the absorption of a pigment of interest at wavelength λ is divided by the specific absorption coefficient for that same pigment at the same wavelength.

$$[\text{Pigment}] = [a_{\text{pigment}(\lambda)}] / [a^*_{\text{pigment}(\lambda)}] \text{ (Simis } \textit{et al.}\text{2005) (eq. 3-4)}$$

Curve Fitting Models

Another model that can adjust for shifts in spectral bands is the modified Gaussian model (MGM) proposed by Sunshine *et al.* (1990). Since the absorption features of algal pigments are inherently Gaussian in shape, equation 3-5 can be used to describe them (Sunshine *et al.* 1990). The rationale for using a modified Gaussian function over the traditional Gaussian function is described in both Sunshine *et al.* (1990) and Robertson *et al.* (in review). The Gaussian distribution (m) of energy absorbed is

expressed in terms of its strength (s), center (μ), and width (σ) and a continuum is superimposed onto these Gaussian functions. (Sunshine *et al.* 1990).

$$m(x) = s * \exp\{-(x^{-1} - \mu^{-1})^2 * (2\sigma^2)^{-1}\}$$ (Sunshine *et al.* 1990) (eq. 3-5)

Curve fitting models, such as MGM, have several advantages over traditional band ratios since they can adjust their strength (or height), center, and width according to the spectra being analyzed. These Gaussian parameters can be correlated to pigment concentration to develop relationships for estimating pigment concentration. The MGM recently has been used to analyze algal pigments (Combe *et al.* 2005; Lohrenz *et al.* 2003; Robertson *et al.* in review). Combe *et al.* (2005) applied MGM to differentiate a microphytobenthos from other features using pigment data from Digital Airborne Imaging Spectrometer (DAIS) imagery. The study was successful in differentiating between the microphytobenthos of interest and macroalgae. Lohrenz *et al.* (2003) used MGM to quantify chlorophyll *a*, chlorophyll *b* and chlorophyll *c* collected on glass fiber filters. The study reported an influence of accessory pigments, specifically carotenoids on the estimation of chlorophyll *b* and *c*. Robertson *et al.* (in review) applied MGM to field measured and airborne hyperspectral spectra to estimate both phycocyanin and chlorophyll *a* concentration.

Although many of these models perform well on datasets collected in specific studies, this study seeks to compare simple band ratio, band ratios set to maximum peaks/troughs, empirical band tuning, semi-empirical, and curve-fitting models. Since these algorithms are dataset-dependant, this study will provide a consistent basis for determining the best performing models in predicting both chlorophyll *a* and

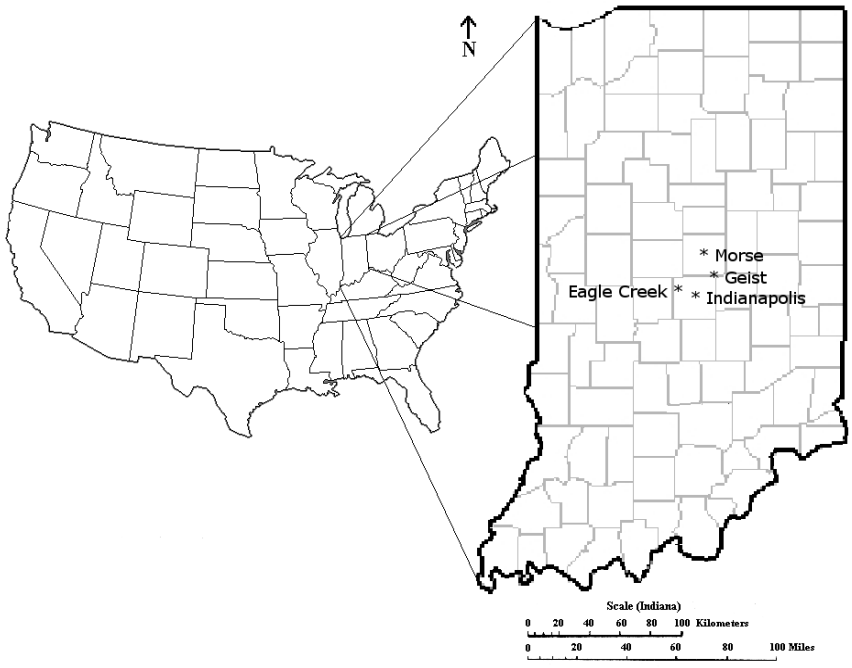
phycocyanin concentrations on three eutrophic central Indiana reservoirs from data collected in 2005-2007.

Methods

Study Site

Three reservoirs located in central Indiana were investigated: Eagle Creek, Geist, and Morse. Figure 3-2 shows the location of each reservoir. Eagle Creek, Geist, and Morse Reservoirs provide water for > 800,000 residents of the greater Indianapolis area. All three reservoirs have similar characteristics including depth (3.2-4.7 m), surface area (5-7.5 km^2), volume (21-28 million m^3), watershed area (420-590 km^2) and residence time (55-70 days) (Li *et al.* 2006). They are also impaired by high nutrient loads (mean total P = 94-100 $\mu gP\text{-}L^{-1}$, mean total N = 2-4.1 $\mu gP\text{-}L^{-1}$) (Li *et al.* 2006) which promote the growth of nuisance algae, including cyanobacteria. All three reservoirs have reported taste and odor issues (IDEM 2006; Li *et al.* 2006).

Figure 3 - 2. Central Indiana reservoirs: Map of the three study locations; Eagle Creek, Geist, and Morse Reservoirs. Indianapolis is indicated for reference.



In situ Reflectance and Water Samples

Sensor used, collection dates, number of samples for water samples and spectra signatures, and the pigment ranges for these samples are listed in Table 3-1. For the 2005 field season all samples were concurrently taken on the same day for both Geist and Morse Reservoirs. For the 2006 season, samples were collected throughout the field season and no two reservoirs were sampled concurrently. For the 2007 field season samples were collected on each reservoir only once and were not taken with any other reservoirs concurrently. Water samples for pigment analysis were collected from surface water with 1 L amber HDPE bottles and stored on ice prior to filtration. The boat-based sensors used in this study include two ASD Field Spec ultraviolet/visible and near-infrared spectroradiometers (Analytical Devices, Inc., Boulder, CO, USA). Both spectroradiometers have a spectral resolution of 1 nm. The spectral range of the spectroradiometer used on Geist Reservoir in 2005 was 350-1050 nm for a total of 701 bands. A second ASD spectroradiometer was used for all other *in situ* reflectance measurements collected in 2005 and 2006 with a spectral range of 350-2500 nm for a total of 2151 bands.

The fiber optic cable for the ASD spectroradiometers has an instantaneous-field-of-view (IFOV) of 0.17 rad. The cable was mounted on an extendable pole with a nadir viewing angle and held approximately 0.5 m above the water surface. This produces a measured water surface area with a diameter of 0.08 m. Calibration for upwelling irradiance for the ASD spectroradiometers was completed by using a white reference panel. In order to reduce noise in the spectra, the reflectance spectrum at each site was averaged over 15 readings.

Table 3 - 1. Summary of sample collection: Includes the year samples were taken, reservoir, number of samples and sensor used.

Reservoir		Year/Sensor							
		2005 ASD FieldSpec		2005 AISA Eagle		2006 ASD FieldSpec		2007 OO USB4000	
		Chl-*a*	PC	Chl-*a*	PC	Chl-*a*	PC	Chl-*a*	PC
Eagle Creek	Sample Number (n)					53	54	22	22
	Minimum (µg/L)					6.3	0.7	16.9	30.9
	Maximum (µg/L)					107.5	234.3	255.0	114.1
Geist	Sample Number (n)	26	27	26	27	88	88	16	16
	Minimum (µg/L)	34.7	25.2	34.7	25.2	23.1	2.6	14.5	0.9
	Maximum (µg/L)	118.9	185.1	118.9	185.1	182.6	210.2	193.2	149.0
Morse	Sample Number (n)	15	16	26	27	79	78	14	15
	Minimum (µg/L)	18.0	2.9	18.0	2.0	21.3	3.3	26.9	41.2
	Maximum (µg/L)	151.7	135.1	168.6	135.1	128.7	371.0	203.8	136.3

An additional boat based sensor includes a pair of Ocean Optics (OO) USB4000 visible and near infrared (VIS/NIR) spectroradiometers (Ocean Optics, Inc., Dunedin, FL, USA) in a dual head system. One USB4000 spectroradiometer was set with a cosine corrector to determine down-welling radiance, thus reducing atmospheric effects. The second USB 4000 spectroradiometer was used to measure up-welling radiance from the target of interest. The USB 4000 spectroradiometer has a spectral resolution of approximately 0.2 nm and a spectra range of 351-1047 nm for a total of 3645 bands.

The fiber optic cable for the Ocean Optics USB4000 spectroradiometers has an instantaneous field of view (IFOV) of 0.14 rad. The cable was mounted on an extendable pole with a nadir viewing angle and held approximately 0.5 m above the water surface.

This produces a measured water surface area with a diameter of 0.07 m. Calibration for upwelling irradiance for the USB4000 spectroradiometers was completed by using a 20% Spectrolon reflectance panel. In order to reduce noise in the spectra, the reflectance spectrum at each site was averaged over 8 readings.

An airborne imaging spectrometer for applications sensor (AISA), model "AISA-Eagle" (Spectral Imaging Ltd., Oulu, Finland), was used to acquire airborne hyperspectral imagery of Geist and Morse reservoirs in 2005. This sensor was flown by the Center for Advanced Land Management Information Technologies (CALMIT) at the University of Nebraska-Lincoln and fitted onboard a Piper-Saratoga airplane. The AISA-Eagle was set to collect the images with 62 bands in the spectral region of approximately 392-982 nm with a spectral range of 7- 9 nm. The IFOV of the AISA sensor across the track is 1 mrad, resulting in 1 m wide pixels and 1000 m wide swath from an altitude of 1000 m. The entirety of Geist and Morse Reservoirs were covered with four and five swaths respectively. Each set of swaths were geo-referenced using 2003 aerial photograph of Marion and Hamilton counties as the base map and the mosaicked image of each reservoir was created using ENVI 4.2 (ITT VIS) mosaicking tool as described in Sengpiel (2007). Atmospheric effects were removed by using an empirical line calibration method where the ASD field spectra were used as reference (Sengpiel, 2007). For this study reflectance spectra for water sample stations were extracted from the image on the basis of their GPS coordinates.

Lab Analysis

Steps for pigment analysis were completed under subdued light conditions. Water samples were filtered in duplicate within 24 hours of the sample collection for later

extraction of pigments. Samples for the extraction of chlorophyll *a* were prepared on 0.47 µm acetate filters, and those for the extraction of phycocyanin on 0.47 µm glass fiber filters (GFF). The samples were stored at -20°C for no longer than 6 months before analysis.

Chlorophyll *a* was extracted following the extraction method described in Environmental Protection Agency (EPA) 445 (Arar and Collins 1997). The concentration of chlorophyll *a* was corrected for pheophytin and measured fluorometrically using a TD-700 Fluorometer (Turner Designs, Inc.) fitted with a Daylight White Lamp and Chlorophyll Optical Kit (340-500 nm excitation filter and emission filter > 665 nm). For chlorophyll *a* analysis the fluorometer was calibrated using chlorophyll *a* from a spinach standard (Sigma-Aldrich 10865).

Phycocyanin was extracted based on method 4 described in Sarada (1999) and Randolph *et al.* (2008). GFFs were suspended in 15 mL of 50 mM phosphate buffer (pH = 6.8). Samples were broken up using a stainless steel spatula and rinsed with 5 mL of 50 mM phosphate buffer. The samples were then homogenized using Teflon coated pestle. Pestles were rinsed with 5 mL of 50 mM phosphate buffer. Samples were centrifuged at 4°C, 15,000 x g for 25 minutes using a Beckman J2-21M centrifuge. The samples were then stored overnight at 4°C before being homogenized again using the Teflon coated pestle and rinsed with 5 mL of 50 mM phosphate buffer with a total volume of 30 mL. The samples were centrifuged again prior to the collection of the supernatant. The supernatant was measured fluorometrically for phycocyanin using a TD-700 Fluorometer (Turner Designs, Inc.) fitted with a Cool White Mercury Vapor Lamp and a Phycocyanin Optical Kit (630 nm excitation and 660 nm emission filters).

For phycocyanin analysis the fluorometer was calibrated using C-phycocyanin from *Spirulina sp.* (Sigma-Aldrich P6161).

If the percent error calculated between replicates in pigment analysis was larger than 20%, the sample was not used in data analysis.

Application of Spectral Indices and Parameters to Mapping Pigments

Band ratios: In this study the spectral indices shown in equations 3-6, 3-7, 3-8, 3-9 and 3-10 were used for estimating chlorophyll *a* and phycocyanin:

Chlorophyll *a*: R_{700} / R_{670} (Gitelson *et al.* 1986) (eq. 3-6)

Chlorophyll *a*: $(R_{705} - R_{670})/ R_{670}$ (Mittenzwey *et al.* 1991) (eq. 3-7)

Phycocyanin: R_{709} / R_{620} (modified from Simis *et al.* 2005) (eq. 3-8)

Chlorophyll *a*: $R_{max(680\text{-}720)}/R_{min(650\text{-}690)}$ (modified from Gitelson *et al.* 1986) (eq. 3-9)

Phycocyanin: $R_{max(680\text{-}720)}/R_{min(600\text{-}640)}$ (modified from Simis *et al.* 2005) (eq. 3-10)

Band tuning: For this study band tuning models presented by Dall'Olmo *et al.* (2003) and Gitelson *et al.* (2005) were evaluated and the corresponding stepwise regression methods were applied for determining optimal three wavelength locations and creating the spectral parameter shown in equation 3-3 with the highest correlation to the pigment of interest. Given $R_{\lambda 1}$, $R_{\lambda 2}$, and $R_{\lambda 3}$ representing three optimal wavelength locations, the stepwise regression procedure introduced by Dall'Olmo *et al*. (2003) and Gitelson *et al*. (2005) finds three optimized wavelength locations in specified spectral regions; $R_{\lambda 1(Chlorophyll)}$ = 660-690 nm, $R_{\lambda 1(Phycocyanin)}$ = 610-640, $R_{\lambda 2}$ = 690-730 nm, and $R_{\lambda 3}$ = 740-800 nm. The Dall'Olmo method estimates $R_{\lambda 1}$ ($R_{est\lambda 1}$) and then solves for a final $R_{\lambda 1}$ value after determining $R_{\lambda 2}$ and $R_{\lambda 3}$. The two different stepwise regression procedures are described below:

Dall'Olmo *et al.* (2003)

a) Solve for $R_{est.\lambda 1}$: $1/R_{est.\lambda 1} - 1/R_{720}$

b) Solve for $R_{\lambda 2}$: $1/R_{est.\lambda 1} - 1/R_{\lambda 2}$

c) Solve for $R_{\lambda 3}$: $(1/R_{est.\lambda 1} - 1/R_{\lambda 2}) * R_{\lambda 3}$

d) Solve for $R_{\lambda 1}$: $(1/R_{\lambda 1} - 1/R_{\lambda 2}) * R_{\lambda 3}$

Gitelson *et al.* (2005)

Solve for $R_{\lambda 2}$: $(1/R_{675} - 1/R_{\lambda 2}) * R_{800}$

Solve for $R_{\lambda 3}$: $(1/R_{675} - 1/R_{\lambda 2}) * R_{\lambda 3}$

Solve for $R_{\lambda 1}$: $(1/R_{\lambda 1} - 1/R_{\lambda 2}) * R_{\lambda 3}$

The final selected wavelengths are presented in Table 3-2.

Table 3 - 2. Wavelengths used in band tuning methods: Final wavelengths are determined by stepwise regression methods described in Dall'Olmo *et al.* 2003 and Gitelson *et al.* 2005 for each calibration data set.

Calibration Dataset	Wavelengths	Dall'Olmo *et al.*2003		Gitelson *et al.*2005	
		Chl-*a*	PC	Chl-*a*	PC
Morse 2005 ASD Field Spec	R1	686	636	688	615
	R2	721	699	697	691
	R3	771	760	797	760
Eagle Creek and Morse 2006 ASD Field Spec	R1	676	628	661	628
	R2	713	703	709	693
	R3	761	760	790	760
Random	R1	676	629	676	629
	R2	723	704	723	965
	R3	761	751	761	770

Semi-empirical models: The semi-empirical models shown in equations 3-11 and 3-12 were tested. All absorption and backscattering coefficients used in this study for testing semi-empirical models were adopted from Simis *et al.* (2005).

Chlorophyll *a*: $[a_{chl(665)}] / [a^*_{chl(665)}]$ (Simis *et al.* 2005) (eq. 3-11)

Phycocyanin: $[a_{pc(620)}] / [a^*_{pc(620)}]$ (Simis *et al.* 2005) (eq. 3-12)

Spectral Curve Fitting with MGM: MATLAB (The MathWorks, Inc.) programs for MGM were modified from Sunshine *et al*. (1999) in order to process multiple files simultaneously. Initial parameters inputted into MGM (Table 3-3) were set based on previous algal pigment studies (Jeffrey *et al*. 1997; Robertson *et al.* in review; Rowan 1989). These include MGM parameters for only describing the absorptions of phycocyanin (628 nm) and chlorophyll (672 nm). MGM iterations stopped when the root mean square error (RMSE) values or improvement between two iterations was less than 1.0×10^{-6}. The resultant MGM parameters, FWHM (Full-Width Half Maximum), strength, area were analyzed for their correlation to the algal pigment (chlorophyll *a* and phycocyanin) concentration.

Table 3 - 3. Initial parameters used for MGM modeling: Continuum intercept: 5.00×10^{-2}, slope: -1.00×10^{-6}.

Parameter	Center of Absorption (nm)	Full-Width Half Maximum (FWHM)	Absorption Strength
Phycocyanin	628	60	-0.4
Chlorophyll *a*	672	40	-0.7

Calibration and Validation Data Sets

Three data sets were used in model calibration. The first calibration data set was the Morse 2005 ASD Field Spec data set. Calibration models derived from this data set was validated with the 2005 Morse AISA-Eagle data set and 2007 Morse OO USB4000 data set. This calibration data set was selected to determine instrument transferability issues with the models when applied to the Morse 2005 AISA-Eagle data set without temporal or spatial variability so MGM output parameters with the highest correlation could be progressed. Validation with the Morse 2007 OO USB4000 data set was to test the instrument transferability of the models to another sensor to support the conclusions

of the best performing MGM output parameters. Models built from the Morse 2005 ASD Field Spec data set were not tested on additional data sets due to concerns that extrapolating the models may result in large errors for estimating phycocyanin (see Table 3-1 for pigment ranges).

The second calibration data set is the Eagle Creek and Morse 2006 ASD Field Spec data set. This data set was chosen because of its high spatial and temporal variability within the summer season (June-September). Models built from this data set will be validated with the Geist and Morse 2005 ASD Field Spec and AISA-Eagle data sets and the Eagle Creek, Morse, and Geist 2007 data sets.

The final calibration data set is a collection of the Eagle Creek and Morse data from all three sensors (ASD Field Spec, AISA-Eagle, and OO USB4000) in 2005, 2006, and 2007. This data set was selected because of the high spatial and temporal variability. Geist was excluded in this data set due to limitations of applying the models from the two previous calibration data sets to Geist Reservoir (see results). The Eagle Creek and Morse 2005-2007 data set will then be divided into a calibration (CHL n = 151, PC n = 154) and validation (CHL n = 51, PC n = 51) data set using a random sequence generator from http://www.random.org.

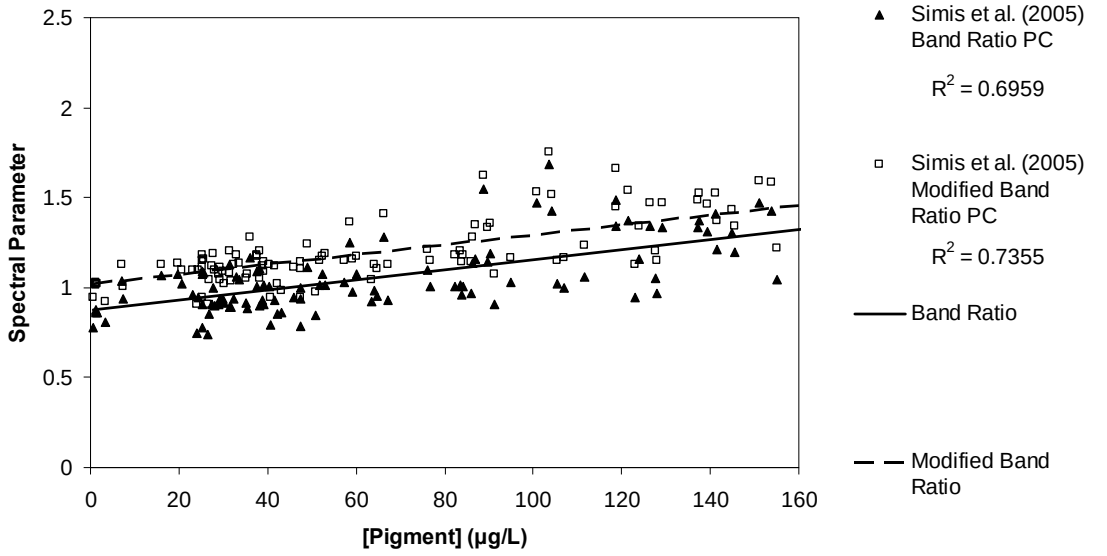
Results

Calibration Models

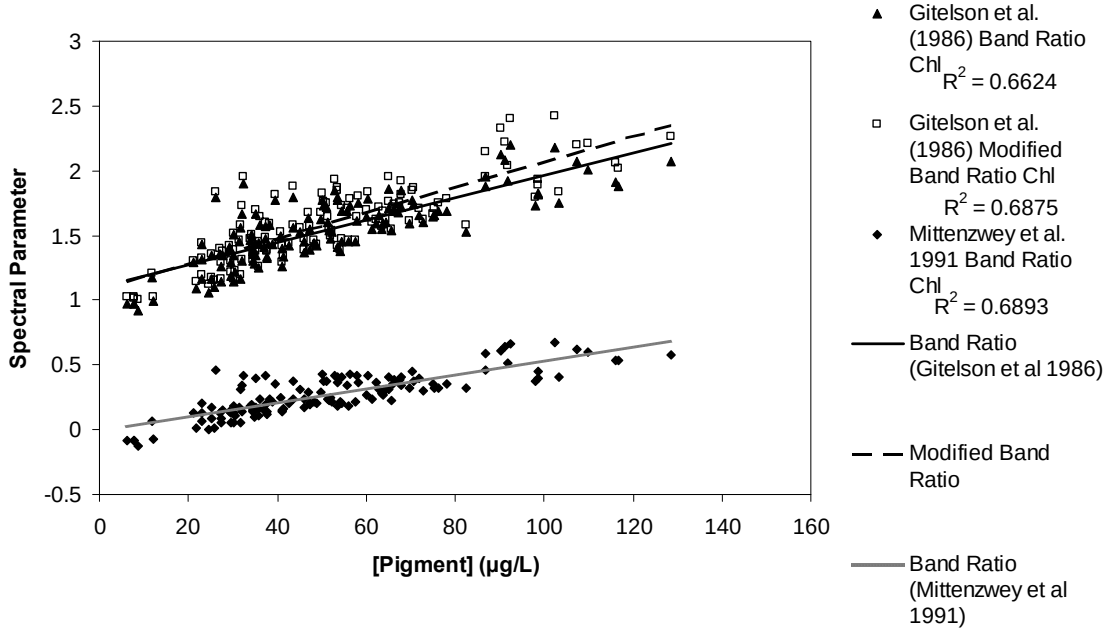
When building the models with individual calibration data sets, spectral parameters were correlated to pigment concentrations. Figure 3-3 shows these models built with the Eagle Creek and Morse 2006 ASD Field Spec calibration data set. A summary of the model created from each calibration data set is presented in Table 3-4.

Figure 3 - 3. Relationship between algal pigment concentration and spectral parameters using the Eagle Creek and Morse 2006 Data ASD Fieldspec Data set: Correlations between pigment concentration (µg/L) and spectral parameters from a) band ratio/modified band ratio algorithms for phycocyanin estimation, b) band ratio/modified band ratio algorithms for chlorophyll *a* estimation, c) Dall'Olmo *et al.* (2003) 3-band tuning, d) Gitelson *et al.* (2005) 3-band tuning, e) Simis *et al.* (2005) semi-empirical, and f) MGM strength.

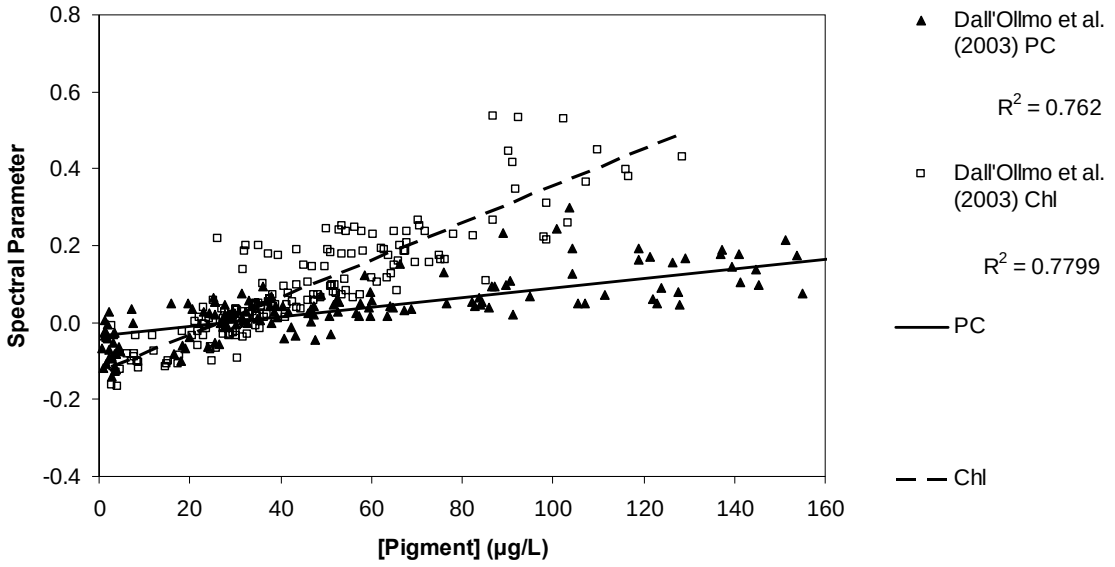
a)



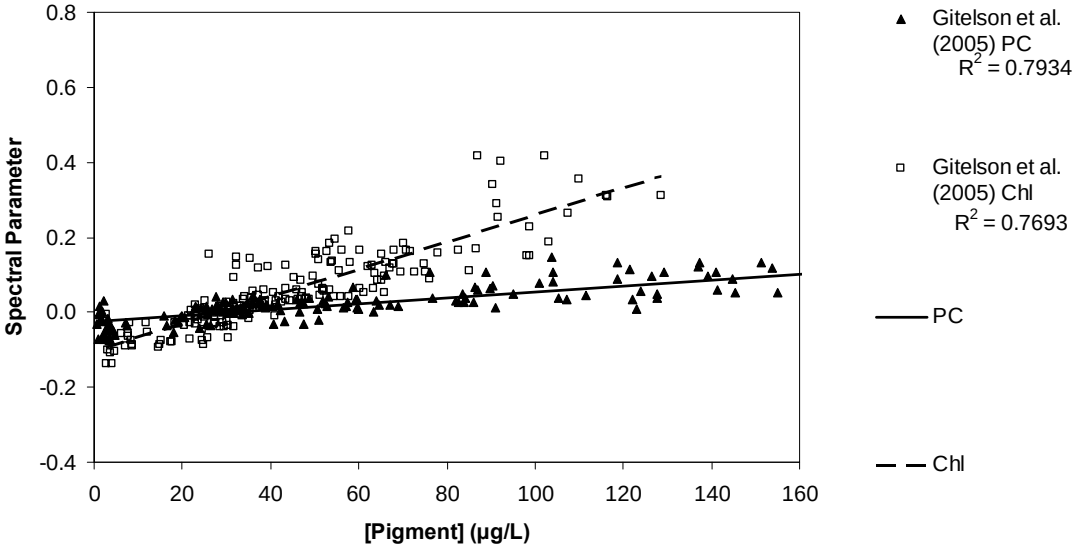
b)



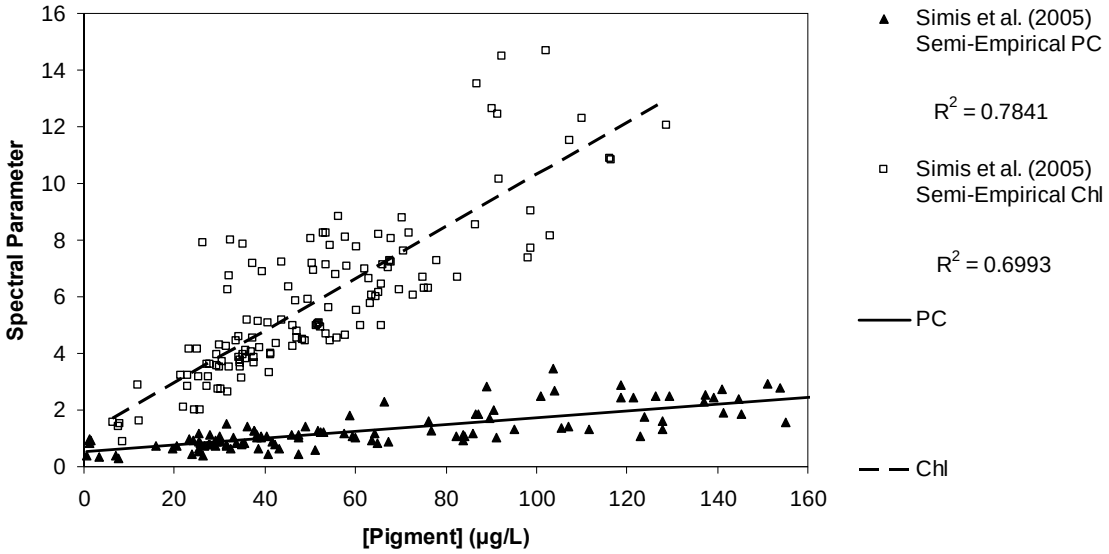
c)


Spectral Parameter
[Pigment] (µg/L)
Dall'Ollmo et al. (2003) PC
R2 = 0.762
Dall'Ollmo et al. (2003) Chl
R2 = 0.7799
PC
Chl

d)


Spectral Parameter
[Pigment] (µg/L)
Gitelson et al. (2005) PC
R2 = 0.7934
Gitelson et al. (2005) Chl
R2 = 0.7693
PC
Chl

e)


Spectral Parameter
[Pigment] (µg/L)
Simis et al. (2005) Semi-Empirical PC
R2 = 0.7841
Simis et al. (2005) Semi-Empirical Chl
R2 = 0.6993
PC
Chl

f)

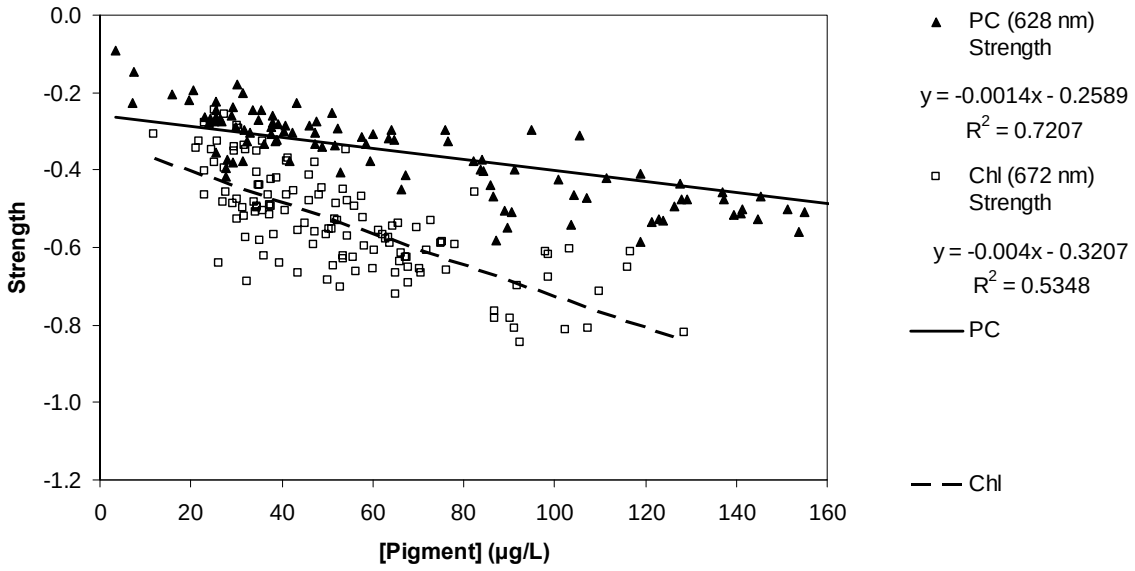

Strength
[Pigment] (µg/L)
PC (628 nm) Strength
y = -0.0014x - 0.2589
R2 = 0.7207
Chl (672 nm) Strength
y = -0.004x - 0.3207
R2 = 0.5348
PC
Chl

Table 3 - 4. Building models with the calibration data sets: Models based on correlation between pigments and spectral indices/parameters. Where x = spectral index/parameter and y = pigment concentration.

Data Set	Pigment	Sample Size (n)	Spectral Parameter	R^2	Model
Morse 2005	Phycocyanin	16	Simis *et al.* 2005 Band Ratio	0.8703	y = 125.00x - 84.29
			Simis *et al.* 2005 Modified Band Ratio	0.8766	y = 136.99x - 116.33
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.8037	y = 714.29x + 103.43
			Gitelson *et al.* 2005 3 Band Tuning	0.7903	y = 38.91x + 909.09
			Simis *et al.* 2005 Semi-Empirical	0.8729	y = 43.29x + 0.31
			Robertson *et al.* MGM PC Width	0.8568	y = 15.20x - 742.16
			Robertson *et al.* MGM PC Strength	0.9301	y = -238.10x - 83.90
			Robertson *et al.* MGM PC Area	0.9196	y = -125.00x - 1608.00
	Chlorophyll *a*	15	Gitelson *et al.* 1986 Band Ratio	0.7736	y = 108.70x - 126.54
			Mittenzwey *et al.* 1991 Band Ratio	0.7283	y = 188.68x - 5.04
			Gitelson *et al.* 1986 Modified Band Ratio	0.7643	y = 89.29x - 99.76
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.7618	y = 333.33x + 95.23
			Gitelson *et al.* 2005 3 Band Tuning	0.7795	y = 588.24x + 17
			Simis *et al.* 2005 Semi-Empirical	0.7048	y = 8.62x - 7.80
			Robertson *et al.* MGM Chl Width	0.5095	y = 23.42x - 636.23
			Robertson *et al.* MGM Chl Strength	0.8215	y = -476.19x - 122.57
			Robertson *et al.* MGM Chl Area	0.7878	y = -6.00x - 55.15
Eagle Creek & Morse 2006	Phycocyanin	127	Simis *et al.* 2005 Band Ratio	0.6959	y = 357.14x - 311.32
			Simis *et al.* 2005 Modified Band Ratio	0.7355	y = 370.37x - 377.22
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.7620	y = 801.28x + 28.02
			Gitelson *et al.* 2005 3 Band Tuning	0.7934	y = 127.714 + 3.07
			Simis *et al.* 2005 Semi-Empirical	0.7841	y = 83.33x - 43.71
			Robertson *et al.* MGM PC Strength	0.7207	y = -714.29x - 184.929
	Chlorophyll *a*	127	Gitelson *et al.* 1986 Band Ratio	0.6624	y = 114.94x - 125.95
			Mittenzwey *et al.* 1991 Band Ratio	0.6893	y = 181.82x + 2.55
			Gitelson *et al.* 1986 Modified Band Ratio	0.6875	y = 101.01x - 109.38
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.7799	y = 208.42x + 26.36
			Gitelson *et al.* 2005 3 Band Tuning	0.7693	y = 275.41x + 28.72
			Simis *et al* 2005 Semi-Empircal	0.8346	y = 27.93x + 33.53
			Robertson *et al.* MGM Chl Strength	0.5348	y =250.00x - 80.18
Eagle Creek & Morse 2005, 2006, 2007 Random	Phycocyanin	154	Simis *et al.* 2005 Band Ratio	0.6322	y = 285.71x – 237.20
			Simis *et al.* 2005 Modified Band Ratio	0.6363	y = 294.12x – 282.15
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.6692	y = 759.30x + 21.83
			Gitelson *et al.* 2005 3 Band Tuning	0.7483	y = 1129.94x + 21.02
			Simis *et al.* 2005 Semi-Empirical	0.6249	y = 69.44x – 37.65
			Robertson *et al.* MGM PC Strength	0.6813	y = -714.29x – 191.07
	Chlorophyll *a*	151	Gitelson *et al.* 1986 Band Ratio	0.6334	y = 175.44x – 214.68
			Mittenzwey *et al.* 1991 Band Ratio	0.7033	y = 131.58x - 156.24
			Gitelson *et al.* 1986 Modified Band Ratio	0.7210	y = 263.16x – 14.89
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.7575	y = 258.73x + 61.54
			Gitelson *et al.* 2005 3 Band Tuning	0.7575	y = 258.73x + 61.54
			Simis *et al.* 2005 Semi-Empirical	0.1889	y = 27.70 - 92.21
			Robertson *et al.* MGM Chl Strength	0.6739	y = 303.03x – 45.36

Models calibrated with the Morse 2005 ASD Field Spec data set had reasonable coefficients of determination ($0.70 < R^2 < 0.93$) with the exception of the model built using MGM width for predicting chlorophyll ($R^2 = 0.51$). Models based on MGM output for FWHM and area were only built using the Morse 2005 ASD Field Spec data set due to issues involving transferability to data sets collected with different sensors (see discussion for details). The models calibrated with the Eagle Creek and Morse 2006 ASD Field spec data set had weaker coefficients of determination ($0.53 < R^2 < 0.83$), likely due to increase in sample size. The models calibrated with the random Eagle Creek and Morse 2005-2007 data set had reasonable coefficients of determination ($0.62 < R^2 < 0.76$) with the exception of the model built using the Simis *et al.* (2005) semi-empirical model ($R^2 = 0.19$).

Validation of the Calibration Models for Estimating Phycocyanin

Calibration Models Based on the Morse 2005 ASD Field Spec Data Set

The validation of the models based on the Morse 2005 ASD Field Spec calibration data set is presented in Table 3-5. These models were applied to data sets collected with the AISA-Eagle, and OO USB4000. Models based on the MGM output for FWHM and Area performed with minimal statistical significance on the two data sets tested with the models calibrated with the Morse 2005 ASD Field Spec data set. For estimating phycocyanin on the Morse 2005 AISA Eagle data set, the models built using the band ratio, modified band ratio and semi-empirical algorithms resulted in the highest correlation coefficients ($0.80 < R^2 < 0.81$). The only other model that had a reasonable correlation coefficient ($R^2 = 0.63$) was that built from the Gitelson *et al.* (2005) 3-band tuning algorithm.

Table 3 - 5. Applying models calibrated with the Morse 2005 ASD Field Spec data set: Models are sorted with descending correlation coefficient (R2). Intercept was forced through zero, therefore ideal models will have a 1) slope near one 2) correlation coefficient (R^2) near one and 3) low RMSE. Caution should be taken in interpreting the Morse 2007 OO USB4000 validation data set for chlorophyll *a* estimation due to extrapolation of the models.

Data Set	Sensor Used	Pigment	Model	Slope	R^2	RMSE (µg/L)	p
Morse 2005	AISA-Eagle	Phycocyanin	Simis *et al.* 2005 Modified Band Ratio	0.9068	0.8147	18.40	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.8620	0.8044	20.03	< 0.01
			Simis *et al.* 2005 Band Ratio	0.8964	0.8022	19.52	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.6617	0.6326	36.11	< 0.01
			Robertson *et al.* MGM PC Strength	0.7992	0.4850	25.44	< 0.02
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.5622	< 0.05	78.90	> 0.1
			Robertson *et al.* MGM PC Width	0.9506	< 0.05	194.98	> 0.1
			Robertson *et al.* MGM PC Area	-0.9360	< 0.05	112.62	> 0.1
		Chlorophyll *a*	Mittenzwey *et al.* 1991 Band Ratio	0.7646	0.8931	20.72	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	0.7937	0.8306	22.94	< 0.01
			Gitelson *et al.* 1986 Band Ratio	0.4698	0.7798	38.62	< 0.01
			Simis *et al.* 2005 Semi-Empirical	1.0398	0.7582	20.99	< 0.01
			Robertson *et al.* MGM Chl Strength	0.7107	0.6056	30.80	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.2188	0.4292	40.23	< 0.05
			Gitelson *et al.* 2005 3 Band Tuning	1.2198	0.3399	40.25	< 0.05
			Robertson *et al.* MGM Chl Width	0.6593	0.1851	76.76	> 0.1
			Robertson *et al.* MGM Chl Area	-0.6733	< 0.05	95.00	> 0.1
Morse 2007	Ocean Optics USB 4000	Phycocyanin	Simis *et al.* 2005 Semi-Empirical	1.4079	0.8822	41.15	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.7570	0.7808	17.54	< 0.01
			Simis *et al.* 2005 Band Ratio	0.6890	0.7757	32.58	< 0.01
			Robertson *et al.* MGM PC Strength	0.8648	0.6787	17.92	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	0.4848	0.4975	54.05	< 0.10
			Robertson *et al.* MGM PC Width	0.4402	0.0929	51.33	> 0.1
			Gitelson *et al.* 2005 3 Band Tuning	-0.1903	< 0.05	10.56	> 0.1
			Robertson *et al.* MGM PC Area	7.5062	< 0.05	591.74	> 0.1
		Chlorophyll *a*	Mittenzwey *et al.* 1991 Band Ratio	0.7680	0.9435	24.69	< 0.01
			Gitelson *et al.* 1986 Band Ratio	0.6994	0.9413	30.95	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	0.7119	0.9288	30.47	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	1.1190	0.8917	17.94	< 0.01
			Robertson *et al.* MGM Chl Strength	0.8688	0.8684	19.39	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.9985	0.8624	18.29	< 0.01
			Robertson *et al.* MGM Chl Width	0.7542	0.5436	30.88	< 0.05
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.1846	0.4861	82.03	> 0.1
			Robertson *et al.* MGM Chl Area	6.2266	0.2251	542.08	> 0.1

Care must be taken in interpreting the results for estimating phycocyanin with the models built from Morse 2005 ASD field Spec data set and validated with the Morse 2007 OO USB4000 due to extrapolation. The two models, semi-empirical and Dall'Olmo *et al.* (2003) 3-band tuning, with the highest correlation coefficient ($0.78 < R^2 < 0.88$) also over predicted ($1.41 < \text{slope} < 1.76$). The other two models with reasonable correlation coefficients ($0.68 < R^2 < 0.78$), based on the band ratio and curve fitting using strength algorithms, under predicted ($0.69 < \text{slope} < 0.86$).

Calibration Models Based on the Eagle Creek and Morse 2006 ASD Field Spec Data Set

The validation of the models based on the Eagle Creek and Morse 2006 ASD Field Spec calibration data set is presented in Table 3-6. The validation on the Geist 2005 ASD Field Spec data set indicates that five models had reasonable correlation coefficients ($0.51 < R^2 < 0.68$). Only the model using the Dall'Olmo *et al*. (2003) 3-band tuning algorithm performed with weak correlation ($R^2 = 0.46$). However, all models over estimated phycocyanin ($1.39 < \text{slope} < 1.79$), and the models with the highest slope resulted in higher errors ($52.19 < \text{RMSE} < 97.69$). When applied to the Geist 2005 AISA Eagle data set, none of the models had reasonable correlation coefficients ($R^2 < 0.05$). When applied to the Geist 2007 OO USB4000 data set, four models based on the band tuning, band ratio and semi-empirical algorithms, had strong correlation coefficients ($0.77 < R^2 < 0.82$). However, all four models grossly over predict phycocyanin ($2.01 < \text{slope} < 3.14$).

Table 3 - 6. Applying models calibrated with the Eagle Creek and Morse 2006 ASD Field Spec data set: Models are sorted with descending correlation coefficient (R2). Intercept was forced through zero, therefore ideal models will have a 1) slope near one 2) correlation coefficient (R^2) near one and 3) low RMSE.

Data Set	Sensor Used	Pigment	Model	Slope	R^2	RMSE (µg/L)	p
Geist 2005	ASD Field Spec	Phycocyanin	Gitelson *et al.* 2005 3 Band Tuning	1.3900	0.6755	52.19	< 0.01
			Robertson *et al.* MGM PC Strength	1.7543	0.5953	88.34	< 0.01
			Simis *et al.* 2005 Semi-Empirical	1.4070	0.5835	58.14	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	1.7160	0.5065	89.87	< 0.01
			Simis *et al.* 2005 Band Ratio	1.7891	0.5059	97.69	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.6339	0.4637	78.54	< 0.02
		Chlorophyll *a*	Dall'Olmo *et al.* 2003 3 Band Tuning	1.1844	0.4464	27.17	< 0.05
			Simis *et al.* 2005 Semi-Empirical	1.1832	0.2571	18.38	> 0.1
			Gitelson *et al.* 1986 Band Ratio	1.2623	< 0.05	32.93	> 0.1
			Gitelson *et al.* 1986 Modified Band Ratio	1.2465	< 0.05	31.81	> 0.1
			Mittenzwey *et al.* 1991 Band Ratio	1.1170	< 0.05	25.53	> 0.1
			Gitelson *et al.* 2005 3 Band Tuning	1.1811	< 0.05	28.34	> 0.1
			Robertson *et al.* MGM Chl Strength	1.3150	< 0.05	32.22	> 0.1
Morse 2005	ASD Field Spec	Phycocyanin	Robertson *et al.* MGM PC Strength	1.4936	0.9301	41.35	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.9111	0.8372	58.65	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	1.4961	0.8040	54.41	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	1.9637	0.7944	90.71	< 0.01
			Simis *et al.* 2005 Semi-Empirical	1.4329	0.7931	50.99	< 0.01
			Simis *et al.* 2005 Band Ratio	2.0824	0.7793	101.92	< 0.01
		Chlorophyll *a*	Robertson *et al.* MGM Chl Strength	1.1343	0.8136	21.59	< 0.01
			Gitelson *et al.* 1986 Band Ratio	1.1482	0.7664	24.92	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	1.1712	0.7631	27.38	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.0793	0.7386	18.87	< 0.01
			Mittenzwey *et al.* 1991 Band Ratio	1.0561	0.7211	23.29	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	1.1646	0.7010	31.17	< 0.01
			Simis *et al.* 2005 Semi-Empirical	1.1921	0.5857	53.88	< 0.05
Geist 2005	AISA-Eagle	Phycocyanin	Simis *et al.* 2005 Band Ratio	1.4649	< 0.05	72.19	> 0.1
			Simis *et al.* 2005 Modified Band Ratio	1.5220	< 0.05	74.72	> 0.1
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.3567	< 0.05	92.50	> 0.1
			Gitelson *et al.* 2005 3 Band Tuning	1.3690	< 0.05	82.33	> 0.1
			Simis *et al.* 2005 Semi-Empirical	1.0029	< 0.05	37.62	> 0.1
			Robertson *et al.* MGM PC Strength	1.4120	< 0.05	58.23	> 0.1
		Chlorophyll *a*	Simis *et al.* 2005 Semi-Empirical	0.9179	0.0505	105.33	> 0.1
			Gitelson *et al.* 1986 Band Ratio	0.7554	< 0.05	26.18	> 0.1
			Gitelson *et al.* 1986 Modified Band Ratio	1.1878	< 0.05	32.31	> 0.1
			Mittenzwey *et al.* 1991 Band Ratio	1.0236	< 0.05	25.12	> 0.1
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.0876	< 0.05	41.31	> 0.1
			Gitelson *et al.* 2005 3 Band Tuning	0.7389	< 0.05	31.02	> 0.1
			Robertson *et al.* MGM Chl Strength	1.2339	< 0.05	35.84	> 0.1
Morse 2005	AISA-Eagle	Phycocyanin	Gitelson *et al.* 2005 3 Band Tuning	1.62	0.8011	54.41	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.6115	0.7626	58.65	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	1.6771	0.6993	71.3	< 0.01
			Simis *et al.* 2005 Band Ratio	1.7403	0.6492	82.94	< 0.01
			Simis *et al.* 2005 Semi-Empirical	1.1391	0.6356	48.64	< 0.01
			Robertson *et al.* MGM PC Strength	1.2249	0.5129	40.7	< 0.01
		Chlorophyll *a*	Gitelson *et al.* 1986 Modified Band Ratio	0.9056	0.9076	14.94	< 0.01
			Mittenzwey *et al.* 1991 Band Ratio	0.8574	0.8658	18.41	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.8535	0.8353	18.9	< 0.01
			Robertson *et al.* MGM Chl Strength	0.9266	0.7988	25.18	< 0.01
			Gitelson *et al.* 1986 Band Ratio	0.5885	0.7985	30.95	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.668	0.7282	31.17	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.8021	0.3940	49.66	< 0.05

Table 3 -6 (cont.)

Data Set	Sensor Used	Pigment	Model	Slope	R^2	RMSE (µg/L)	p
Eagle Creek 2007	Ocean Optics USB 4000	Phycocyanin	Simis *et al.* 2005 Semi-Empirical	2.6312	0.6982	136.58	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.6629	0.6318	64.66	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	1.0379	0.6106	21.84	< 0.01
			Simis *et al.* 2005 Band Ratio	1.7244	0.5765	76.07	< 0.01
			Robertson *et al.* MGM PC Strength	1.1042	0.5398	42.70	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	1.0625	0.3150	56.98	> 0.1
		Chlorophyll *a*	Mittenzwey *et al.* 1991 Band Ratio	0.6808	0.9067	45.36	< 0.01
			Robertson *et al.* MGM Chl Strength	0.7124	0.9062	41.73	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	0.6881	0.9059	44.71	< 0.01
			Gitelson *et al.* 1986 Band Ratio	0.6422	0.9007	50.58	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.7275	0.8930	39.71	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.2098	0.7538	109.05	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.3759	0.6513	89.73	< 0.01
Geist 2007	Ocean Optics USB 4000	Phycocyanin	Dall'Olmo *et al.* 2003 3 Band Tuning	2.2549	0.8214	136.44	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	2.0057	0.8032	112.95	< 0.01
			Simis *et al.* 2005 Band Ratio	2.0318	0.7803	114.62	< 0.01
			Simis *et al.* 2005 Semi-Empirical	3.1426	0.7673	229.97	< 0.01
			Robertson *et al.* MGM PC Strength	1.6482	0.5199	91.39	< 0.05
			Simis *et al.* 2005 Modified Band Ratio	0.9310	0.2975	33.43	> 0.1
		Chlorophyll *a*	Dall'Olmo *et al.* 2003 3 Band Tuning	0.6713	0.9026	50.51	< 0.01
			Mittenzwey *et al.* 1991 Band Ratio	0.5421	0.8816	69.62	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	0.5626	0.7297	68.05	< 0.01
			Gitelson *et al.* 1986 Band Ratio	0.5059	0.7085	76.28	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.3677	0.6312	97.71	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.1756	0.5732	124.52	< 0.05
			Robertson *et al.* MGM Chl Strength	0.5105	0.3102	79.40	> 0.1
Morse 2007	Ocean Optics USB 4000	Phycocyanin	Simis *et al.* 2005 Semi-Empirical	2.2301	0.7967	119.17	< 0.01
			Robertson *et al.* MGM PC Strength	1.2858	0.6857	31.94	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.3382	0.6723	59.80	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.7518	0.6047	42.97	< 0.01
			Simis *et al.* 2005 Band Ratio	1.2169	0.5140	78.68	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	0.6229	0.2100	105.41	< 0.1
		Chlorophyll *a*	Dall'Olmo *et al.* 2003 3 Band Tuning	0.9446	0.9422	13.31	< 0.01
			Mittenzwey *et al.* 1991 Band Ratio	0.7565	0.9388	25.99	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	0.8994	0.9381	14.88	< 0.01
			Gitelson *et al.* 1986 Band Ratio	0.8090	0.9309	21.50	< 0.01
			Robertson *et al.* MGM Chl Strength	0.9806	0.8170	18.01	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.4234	0.4785	71.65	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.2124	0.4296	80.75	< 0.01
Eagle Creek and Morse 2007	Ocean Optics USB 4000	Phycocyanin	Simis *et al.* 2005 Semi-Empirical	2.4485	0.7316	129.81	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.5180	0.6357	62.81	< 0.01
			Robertson *et al.* MGM PC Strength	1.1869	0.5882	38.70	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.9103	0.5800	31.77	< 0.01
			Simis *et al.* 2005 Band Ratio	1.4933	0.5249	77.14	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	0.8623	0.2792	80.22	< 0.02
		Chlorophyll *a*	Mittenzwey *et al.* 1991 Band Ratio	0.6994	0.9166	38.99	< 0.01
			Gitelson *et al.* 1986 Band Ratio	0.6832	0.8624	41.76	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	0.7401	0.8515	36.16	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.7802	0.8490	32.51	< 0.01
			Robertson *et al.* MGM Chl Strength	0.7784	0.7632	34.50	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.2104	0.6029	99.01	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.3874	0.5807	83.48	< 0.01

The validation on the Morse 2005 ASD data set indicates that the best performing model based on the correlation coefficient was the curve fitting model ($R^2 = 0.93$). All other models had strong correlation coefficients ($0.78 < R^2 < 0.84$). However, all models over predicted phycocyanin ($1.43 < \text{slope} < 2.08$). The validation on the Morse 2005 AISA Eagle data set indicates that the best two performing algorithms according to correlation coefficients were the band tuning models ($0.76 < R^2 < 0.80$). While the semi-empirical and curve fitting models had weaker correlation coefficients ($0.51 < R^2 < 0.64$), the slope for these two models ($1.14 < \text{slope} < 1.22$) were much closer to 1 than the remaining models ($1.61 < \text{slope} < 1.74$). The validation on the Morse 2007 OO USB4000 data set shows that the semi-empirical model had the strongest correlation coefficient ($R^2 = 0.80$); however, it also grossly over predicted phycocyanin (slope = 2.23). The curve fitting and band tuning models had similar correlation coefficients ($0.60 < R^2 < 0.69$) and yielded slopes much closer to one ($0.75 < \text{slope} < 1.34$). The validation on the Eagle Creek 2007 OO USB4000 data set indicates that the best two performing models according to their correlation coefficient were based off of the semi-empirical and Dall'Olmo *et al.* (2003) 3-band tuning model ($0.63 < R^2 < 0.70$); however, both over estimated phycocyanin ($1.66 < \text{slope} < 2.63$). The two models with slopes close to one ($1.04 < \text{slope} < 1.10$) were the Gitelson *et al.* (2005) 3-band tuning model and the curve fitting model. Both models performed with weaker correlation coefficients ($0.54 < R^2 < 0.61$).

The validation on the Eagle Creek and Morse 2007 OO USB4000 data set resulted in a similar trend as the validation with the Eagle Creek 2007 OO USB4000 data set. The two best performing models according to the correlation coefficient were based

on the semi-empirical and Dall'Olmo (2003) 3-band tuning algorithms ($0.64 < R^2 < 0.73$), but over estimated phycocyanin ($1.52 < \text{slope} < 2.45$). Models that predicted with slopes closer to one ($0.91 < \text{slope} < 1.19$) and reasonable correlation coefficients ($0.58 < R^2 < 0.59$) were based on the Gitelson *et al.* (2005) 3-band tuning and curve fitting algorithms.

The Calibration Model Based on the Randomly Selected Eagle Creek and Morse 2005, 2006, and 2007 Data Set

The validation of the models based on the randomly selected Eagle Creek and Morse data set is presented in Table 3-7. When the models were validated with the remaining data points from the random Eagle Creek and Morse 2005-2007 data set, the best performing model according to the correlation coefficient was based on the curve fitting model ($R^2 = 0.74$). The two remaining models with reasonable correlation coefficients were based on the Gitelson *et al.* (2005) 3-band tuning and modified band ratio algorithms. Since data points were randomly assigned between calibration and validation data sets the slopes from all models were near one ($0.89 < \text{slope} < 1.01$).

Validation of the Calibration Models for Estimating Chlorophyll a

Calibration Models Based on the Morse 2005 ASD Field Spec Data Set

The best performing model demonstrated through the validation on the Morse 2005 AISA Eagle data set was the Mittenzwey *et al.* band ratio algorithm resulting in a correlation coefficient of 0.89. Other models that had strong correlation coefficients ($0.76 < R^2 < 0.83$) include those based on the Gitelson *et al.* (1986) band ratio, modified band ratio, and semi-empirical algorithms. The model based on the curve fitting algorithm had a reasonable correlation coefficient ($R^2 = 0.61$). Only the model based on

the semi-empirical algorithm had a near linear slope (slope = 1.04). The model based on the Gitelson *et al.* (1986) band ratio algorithm grossly under predicted (slope = 0.47). The remaining models with reasonable correlation coefficients under predicted slightly ($0.71 < \text{slope} < 0.79$).

The validation on the Morse 2007 OO USB4000 data set indicates that based on correlation coefficients ($0.86 < R^2 < 0.94$) the band ratio, Gitelson *et al.* (2005) 3-band tuning, and curve fitting using strength algorithms performed best. However if the slope ($0.89 < \text{slope} < 1.12$) and rmse ($17.94 < \text{RMSE} < 19.39$) are considered, the best performing models are those using Gitelson *et al* (2005) 3-band tuning, curve fitting using strength, and semi-empirical algorithms.

Table 3 - 7. Applying models calibrated with the Randomly Selected Data Set: Models are sorted with descending correlation coefficient (R2). Selected points come from Eagle Creek and Morse data sets from 2005, 2006, and 2007. Intercept was forced through zero, therefore ideal models will have a 1) slope near one 2) correlation coefficient (R^2) near one and 3) low RMSE.

Data Set	Sensor Used	Pigment	Model	Slope	R^2	RMSE (µg/L)	p
Eagle Creek & Morse 2005, 2006, 2007 Validation	All Sensors	Phycocyanin	Robertson *et al.* MGM PC Strength	1.0133	0.7380	42.92	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	0.9080	0.6314	45.00	< 0.01
			Simis *et al.* 2005 Modified Band Ratio	0.9269	0.5267	49.88	< 0.01
			Simis *et al.* 2005 Semi-Empirical	0.8890	0.4606	62.36	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	0.8947	0.4586	59.83	< 0.01
			Simis *et al.* 2005 Band Ratio	0.8969	0.4350	56.58	< 0.01
		Chlorophyll *a*	Mittenzwey *et al.* 1991 Band Ratio	1.0354	0.7883	21.39	< 0.01
			Gitelson *et al.* 1986 Modified Band Ratio	0.9949	0.7703	20.61	< 0.01
			Dall'Olmo *et al.* 2003 3 Band Tuning	1.0220	0.7146	21.09	< 0.01
			Gitelson *et al.* 2005 3 Band Tuning	1.0220	0.7146	21.09	< 0.01
			Gitelson *et al.* 1986 Band Ratio	1.0055	0.6963	24.90	< 0.01
			Robertson *et al.* MGM Chl Strength	0.9979	0.6862	25.84	< 0.01
			Simis *et al.* 2005 Semi-Empirical	1.0641	0.3873	37.83	< 0.01

Calibration Models Based on the Eagle Creek and Morse 2006 ASD Field Spec Data Set

Models validated with the models calibrated with the Eagle Creek and Morse 2006 ASD Field Spec data set performed poorly according to the correlation coefficients ($R^2 < 0.50$) on all Geist data sets for predicting chlorophyll *a* with the exception of the

OO USB4000 data set. While model built using the Mittenzwey *et al.* (1991) band ratio algorithm shows a strong correlation ($R^2 = 0.88$), chlorophyll *a* is grossly under predicted (slope = 0.54). This under prediction occurs for all models when applied to the Geist 2007 OO USB4000 data set ($0.18 < \text{slope} < 0.56$).

The semi-empirical algorithm resulted in the weakest correlation coefficient ($R^2 = 0.59$), when validated on the Morse 2005 ASD Field Spec data set. All other models had strong correlation coefficients ($0.70 < R^2 < 0.81$). All models had near linear correlations ($1.06 < \text{slope} < 1.19$). Similar results were obtained when models were applied to the Morse 2005 AISA Eagle data set. All models had a strong correlation coefficient ($0.73 < R^2 < 0.87$) with the exception of the semi-empirical model ($R^2 = 0.39$). However, the best models, based on both slope and correlation coefficient ($0.85 < \text{slope} < 0.93$), were the Mittenzwey *et al.* (1991) band ratio, Dall'Olmo *et al.* (2003) 3-band tuning, and the curve fitting algorithms. When applied to the Morse 2007 OO USB4000 data set, five models based on the Dall'Olmo *et al.* (2003) 3-band turning, band ratios, modified band ratio, and curve fitting algorithms had strong correlation coefficients ($0.82 < R^2 < 0.94$). Of these models the band turning and curve fitting algorithms had the closest slopes to one ($0.94 < \text{slope} < 0.98$). These same five models had the highest correlation coefficients ($0.89 < R^2 < 0.91$) and closest slopes to one ($0.64 < \text{slope} < 0.73$) when applied to the Eagle Creek 2007 OO USB4000 data set.

Not surprisingly, when validated with the Eagle Creek and Morse 2007 OO USB4000 data set, the same models validated with the Morse 2007 OO USB4000 data set had the highest correlation coefficients ($0.76 < R^2 < 0.92$) and similar slopes ($0.68 < \text{slope} < 0.78$).

Calibration Models Based on the Randomly Selected Eagle Creek and Morse 2005, 2006, and 2007 Data Set

When the models were validated with the models calibrated with the random Eagle Creek and Morse 2005-2007 data set, the worst performing model according to the correlation coefficient was based on the semi-empirical algorithm ($R^2 = 0.39$). All other models performed nearly equally according to correlation coefficient ($0.69 < R^2 < 0.79$), slope ($0.99 < \text{slope} < 1.04$) and error ($20.61 < \text{RMSE} < 25.84$).

Discussion

Performance of Band Ratio/Modified Band Ratio Algorithms

For chlorophyll *a* prediction all band ratio/modified band ratio algorithms perform nearly equally with all calibration and validation data sets. Only one algorithm, based on Gitelson *et al.* (1986), performed poorly when calibrated with the ASD Field Spec data sets and applied to the AISA-Eagle data sets resulting in under estimation of chlorophyll *a* and an increase in RMSE. In both cases, chlorophyll *a* was greatly under predicted when compared to other band ratios. This is likely due to calibration of the AISA-Eagle data set.

The performance of the band ratio/modified band ratio algorithms for predicting phycocyanin are not nearly as good as for predicting chlorophyll *a* estimation. There are several possibilities why the accuracy of phycocyanin estimation is reduced with band ratio/modified band ratio algorithms. The first possibility is spectral overlapping of phycocyanin with chlorophyll *a*. An increase in phycocyanin does not necessarily correspond to a linear increase in other potentially interfering constituents, therefore reducing temporal and spatial transferability. While we have shown that phycocyanin

concentration can shift the reflectance trough near 620 nm (Figure 3-1), it is possible that different water constituents (i.e. chlorophyll *a*) also influence the position of the reflectance trough. Secondly, spectral saturation could occur at higher concentrations, reducing correlation between the reflectance at these wavelengths and phycocyanin concentration. Lastly, the wavelength chosen by most algorithms to reduce the influence of CDOM (~700 nm) is relatively far from the primary wavelength related to phycocyanin absorption (~628 nm), when compared to the primary wavelength used by many chlorophyll *a* algorithms (~670 nm). Since these equations make the assumption that the CDOM absorption value at 620 nm is near the absorption value of CDOM at 700 nm, differences in the CDOM absorption value between these two wavelengths will reduce the accuracy of the models built with these algorithms. In a similar way, tripton also can affect the algorithm.

Performance of Band Tuning Algorithms

In this paper we found that the Dall'Olmo *et al.* (2003) stepwise regression method consistently worked better than or as well as the Gitelson *et al.* (2005) method for predicting chlorophyll *a*. Although the Dall'Olmo *et al.* (2003) method did not perform well when the model was calibrated with the Morse 2005 ASD Field Spec and then applied to the Morse 2007 OO USB4000 data set. This discrepancy is likely due to the calibration wavelengths utilized by the Morse 2005 ASD Field Spec models. Very different $R_{\lambda 2}$ wavelengths were determined (Dall'Olmo = 721 nm; Gitelson = 697 nm), and previous papers have found performance of algorithms to be sensitive to the NIR wavelength (Gitelson 1992; Schalles *et al.* 1998; Yacobi *et al.* 1995).

In contrast, the Gitelson *et al.* (2005) method worked best for predicting phycocyanin. The only time the Gitelson *et al.* (2005) 3-band tuning algorithm does not work satisfactorily for predicting phycocyanin is when models were calibrated with the Morse 2005 ASD data set and applied to the Morse 2007 OO USB4000 data set. This is likely due to either the extrapolation of the models or the relatively short $R_{\lambda 1}$ wavelength (615 nm) calculated using the step-wise regression proposed by Gitelson *et al.* (2005) on a small calibration data set (n = 16). Since the Dall'Olmo *et al.* (2003) method works with extrapolation, the short wavelength selected by the Gitelson *et al.* (2005) method is likely to be the main cause of the performance of this model. All of the other calibration data sets resulted in a wavelength of 628 nm or greater (Table 3-2). The wavelength selected using this calibration data set does not best represent the temporal differences in Morse Reservoir. Increasing the sample size and temporal variability, as in the Eagle Creek and Morse 2006 ASD Field Spec calibration data set, would likely overcome aberrant wavelengths selected with these methods.

The band tuning method works well for estimating pigments. One reason this method outperforms traditional band ratio methods is that wavelengths used in the algorithm can be adjusted for correlation to phycocyanin concentrations with respect to conditions in the calibration reservoir(s). While the maximum peak/trough method does adjust the wavelength utilized, the ratio of peaks and troughs do not necessarily correlate to phycocyanin concentration as strongly as the spectral parameters derived by band tuning.

Performance of Semi-Empirical Algorithms

For estimating chlorophyll *a,* the Simis *et al.* (2005) algorithm performed well when transferred temporally when calibrated with the Morse 2005 ASD Field Spec and validated with the Morse 2007 OO USB4000 data set. However, this algorithm performed poorly when calibrated with the Eagle Creek and Morse 2006 ASD Field Spec, and Eagle Creek and Morse Random data sets. The failure of this algorithm is likely due to its inability to transfer across the two reservoirs due to slight differences in the inherent optical properties (IOP) (i.e. absorption and backscattering coefficients) for estimating chlorophyll *a* in the two reservoirs. While most algorithms tested in this paper show that Eagle Creek and Morse Reservoirs are similar and can be combined in one data set, there is incompatibility between Eagle Creek and Morse reservoirs for estimating chlorophyll *a* using the Simis *et al.* (2005) semi-empirical approach.

In contrast, the semi-empirical approach shows more promise for phycocyanin prediction across both Eagle Creek and Morse Reservoirs. While the coefficients of determination are high in most of the validation data sets, the models under/over predict phycocyanin. It was stated as a problem in the Simis *et al.* (2005) paper to use fixed absorption coefficients. The semi-empirical algorithm utilized in this study lacks an effective method for removing the influence of suspended sediments which likely vary both temporally and spatially. Incorporating these parameters may increase the performance of this algorithm.

Performance of Curve-Fitting Algorithms

According to this study, models built on MGM output parameters for FWHM and area do not transfer across instruments. This is possibly due to the differences in spectral

resolution. FWHM and area, a function of FWHM, is likely more sensitive to spectral resolution since the curve fitting algorithm adjusts the width of the Gaussian curve based on the spectral curve supplied. If models are built on data sets with different spectral resolutions, the differences in spectral resolution may affect transferability of the models based on MGM output parameters FWHM and area. One way to test this is to average spectra to different spectral resolutions and apply the MGM models built from one data set to another.

Curve fitting models using the MGM output parameters for strength successfully predicted both chlorophyll *a* and phycocyanin. It is likely that MGM can remove the some of the effects caused by minor differences in the inherent optical properties between reservoirs during the continuum removal processes. Although models built from MGM strength were able to accurately estimate chlorophyll *a*; these models did not outperform band ratio/modified band ratio models. As stated in Robertson *et al.* (in review), the MGM output parameter for chlorophyll *a* likely includes absorption features caused by chlorophylls *b* and/or *c*. This could reduce the overall performance of predicting chlorophyll *a.*

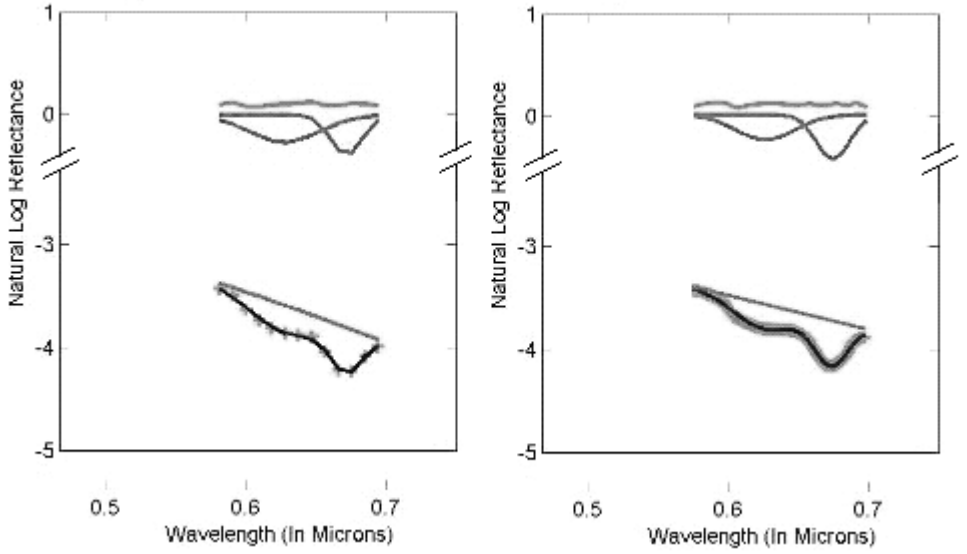
In contrast the models built from MGM strength were able to perform as well as or better than other models in estimating phycocyanin concentration. One reason that these models outperform other algorithms is the ability of the MGM software to minimize the influence of overlapping chlorophyll *a* and phycocyanin absorptions.

One exception to the performance of the MGM models based on the output parameter for strength is when the models were applied to the AISA-Eagle data set. The poorer performance with these data sets is likely due to reduced spectral resolution. With

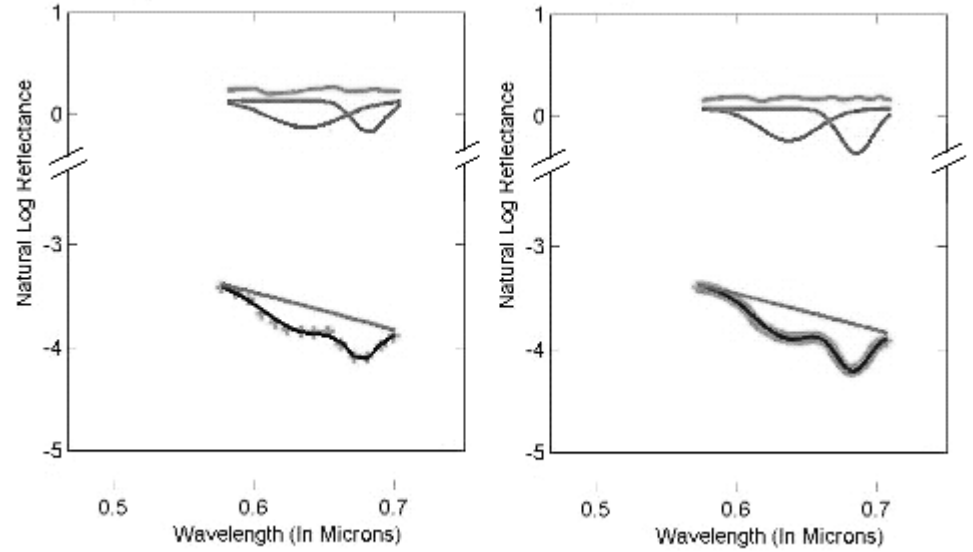
fewer wavelengths, it is possible that the true absorption peak is not represented in all spectra, causing flattened peaks and reduced correlation (Figure 3-4).

Figure 3 - 4. Differences in MGM curve fitting between ASIA and ASD Field Spec data collection: a) sample #1 b) sample #15

a)



b)



Issues involving Transferability of Models to Geist Reservoir

As stated in Robertson *et al.* (in review), there was difficulty in transferring traditional band ratio and MGM based models to Geist Reservoir, likely caused by dredging and mining operations that stir up sediment and provide a direct input of ground water into the reservoir. A local mining company routinely dredges for sands and silts in the northeastern basin of the reservoir. Vallely (2008) found sediments to be one of the major confounding factors in influencing algorithms tested in her study. Additionally, the same company has a gravel pit adjacent to the reservoir. In order to maintain a water level to extract the sand and gravel from this pit, water is pumped directly into the reservoir. This direct input of ground water alters the water chemistry constantly, making it difficult to correlate ground truth data to aerial/satellite based sensors. In this study we confirmed several additional methods, such as the semi-empirical and band tuning methods are also incapable of addressing this issue. As a result of our study, these algorithms work best on reservoirs with minimal anthropogenic impacts and are likely only transferable to other reservoirs with similar conditions.

Conclusion

For predicting chlorophyll *a*, no algorithm greatly outperformed other algorithms. Because band ratio and modified band ratio models are easy to use, we recommend using these approaches to estimate chlorophyll *a* of the Central Indiana Reservoirs over other methods tested in this study (semi-empirical, 3 band tuning, and MGM). The result indicates that the band tuning method works equally well for estimating chlorophyll *a*; however, this method will require additional time in developing calibration wavelengths. In addition, the spectral resolution of the AISA data affected the effectiveness of MGM

for estimating chlorophyll *a*, while the MGM with higher spectral resolution datasets yielded chlorophyll estimates at high accuracy. In contrast the semi-empirical method is affected by spatial transferability, not spectral resolution. The semi-empirical model can work temporally on the same reservoir; however, due to differences in inherent optical properties (IOP) across reservoirs, these models are less effective.

For predicting phycocyanin, only two algorithms performed consistently well with different calibration and validation data sets: the models built with the MGM phycocyanin strength and the Gitelson *et al.* (2005) 3-band tuning algorithm. These sophisticated models are an improvement to the traditional band ratios in estimating phycocyanin. They take into account the inherent complex spectral overlapping of phycocyanin and other pigments near 620 nm. MGM can separate spectral overlapping by spectral deconvolution. Band tuning methods select wavelengths that correlate to either direct pigment estimation phycocyanin or water constituents that interfere with this signal. The wavelengths correlating to other water constituents can then be used to remove the effects of these water constituents on estimating phycocyanin concentration. While the semi-empirical model should be able to resolve spectral overlapping, this model does not perform well in estimating phycocyanin when calibrated with multiple reservoirs. The reduced effectiveness of this model is related to the same issues when estimating chlorophyll *a* and is likely due to differences in the inherent optical properties between the reservoirs tested.

Acknowledgements

We would like to thank Veolia Water Indianapolis and the Central Indiana Water Resources Partnership for financial support. We would also like to thank Lani Pascual,

Kate Randolph, Becky Sengpiel, Lara Vallely, Bob Hall, Vince Hernly, and other staff, graduate students and interns involved with the collection and processing of the data sets.

References

Arar, E.J., & Collins, G.B. (1997). In Vitro Determination of Chlorophyll *a* and Pheophytin a in Marine and Freshwater Algae by Fluorescence. In. Cincinnati, OH.

Bold, H.C., & Wynne, M.J. (1985). *Introduction to the Algae: structure and reproduction*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Chorus, I., & Bartram, J. (1999). *Toxic Cyanobacteria in Water: A guide to their public health consequences, monitoring and management*. Bury St Edmunds, Suffolk: St. Edmundsbury Press.

Christensen, V.G., Graham, J.L., Milligan, C.R., Pope, L.M., & Ziegler, A.C. (2006). Water quality and relation to taste and odor compounds in the North Fork Ninnescah River, and Cheney Reservoir, south-central Kansas. In, *U.S. Geologcial Survey Scientific Investigations Report 2006.*

Cole, T.M., & Tillman, D.H. (1999). Water Quality Modeling of Lake Monroe Using CE-QUAL-W2. In, *Waterways Experiment Station*. Vicksburg, MS: US Army Corps of Engineers Dall'Olmo, G., & Gitelson, A. (2005). Effect of bio-optical parameter variability on the remote estimation of chlorophyll *a* concentration in turbid productive waters: experimental results. *Applied Optics, 44*, 412-422.

Combe, J.-P., Launeau, P., Carrere, V., Despan, D., Meleder, V., Barille, L., & Sotin, C. (2005). Mapping microphytobenthos biomass by non-linear inversion of visible-infrared hyperspectral images. *Remote Sensing of Environment, 98*, 371-387.

Dall'Olmo, G., & Gitelson, A. (2005). Effect of bio-optical parameter variability on the remote estimation of chlorophyll *a* concentration in turbid productive waters: experimental results. *Applied Optics, 44*, 412-422.

Dall'Olmo, G., Gitelson, A., & Rundquist, D. (2003). Towards a unified approach for remote estimation of chorophyll *a* in both terrestrial vegetation and turbid productive waters. *Geophysical Research Letters, 30*, 1938-1942.

Dekker, A.G. (1993). Detection of the optical water quality parameters for eutrophic waters by high resolution remote sensing. In. Ansterdam: Free University

Falconer, I.R. (2005). *Cyanobacterial Toxins of Drinking Water Supplies*. Boca Raton, FL: CRC Press.

Gitelson, A.A. (1992). The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. *International Journal of Remote Sensing, 13*, 3367 – 3373.

Gitelson, A., Dall'Olmo, G., Moses, M., Rundquist, D., Barrow, T., Fisher, T.R., Gurlin, D., & Holz, J. (2008). A simple semi-analytical model for remote estimation of chlorophyll *a* in turbid waters: Validation. *Remote Sensing of Environment, 112.*

Gitelson, A.A., Laorawat, S., Keydan, G.P., & Vonshak, A. (1995). Optical properties of dense algal cultures outdoors and their application to remote estimation of biomass and pigment concentration in *Spirulina Platensis* (cyanobacteria). *Journal of Phycology, 31*, 828-834.

Gitelson, A., Nikanorov, A., Sabo, G., & Szilagyi, F. (1986). Etude de la qualite des eaux de surface par teledetection. *IAHS Publications, 157*, 111-121.

Gitelson, A.A., Vina, A., Ciganda, V., Rundquist, D.C., & Arkebauer, T.J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters, 32*, 1-4.

Han, L. (1997). Spectral Reflectance with Varying Suspended Sediment Concentrations in Clear and Algae-Laden Waters. *Photogrammetric Engineering and Remote Sensing, 63*, 701-705.

Han, L., & Rundquist, D. (1994). The Response of both Surface Reflectance and the Underwater Light Field to Various Levels of Suspended Sediments: Preliminary Results. *Photogrammetric Engineering and Remote Sensing, 60*, 1463-1471.

IDEM (2006). Indiana Integrated Water Monitoring and Assessment Report 2006. In. Indianapolis: Indiana Department of Environmental Management.

Jeffrey, S.W., Mantoura, R.F.C., & Wright, S.W. (1997). *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods*. Paris: UNSECO.

Jeffrey, S.W., & Wright, S.W. (2006). *Photosynthetic Pigments in Marine Microalgae: Insights from Cultures and the Sea*. Enfield, NH: Science Publishers.

Li, L., Pascual, D.L., Tedesco, L.P., Randolph, K.L., Sengpiel, R.E., Hall, B.E., & Wilson, J.S. (2006). Developing a survey tool for the rapid assessment of blue-green algae in central Indiana's reservoirs. In, *2005 LARE Grant Report*: Indiana Department of Natural Resources.

Li, L., Sengpiel, R.E., Pascual, D.L., Tedesco, L.P., Wilson, J.S., & Soyeux, E. (2009). Hyperspectral Remote Sensing to Estimate Chlorophyll *a* and Phycocyanin in a Mesotrophic Reservoir. *International Journal of Remote Sensing*, in press.

Lohrenz,S. E., Weidemann, A.D., & Tuel, M. (2003). Phytoplankton spectral absorption as influenced by community size structure and pigment composition. *Journal of Plankton Research, 25*, 35-61.

Mittenzwey, K.-H., Breitwieser, S., Penig, J., Gitelson, A.A., Dubovitzkii, G., Garbusov, G., Ullrich, S., Vobach, V., & Müller, A. (1991). Fluorescence and Reflectance for the *in-situ* Determination of Some Quality Parameters of Surface Waters. *Acta Hydrochimica et Hydrobiologica, 19*, 1-15.

Randolph K.L. (2008). *Remote sensing of cyanobacteria in case II waters using optically active pigments, chlorophyll a and phycocyanin* Indianapolis: Indiana University Purdue University Indianapolis M.S. Thesis.

Randolph, K.L., Wilson, J.S., Tedesco, L.P., Lin, L., Pascual, D.L., & Soyeux, E. (2008). Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll *a* and phycocyanin *Remote Sensing of Environment, 112*, 4009-4019.

Richardson, L.L. (1996). Remote Sensing of Algal Bloom Dynamics. *BioScience, 46*, 492-501.

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2009). Using a modified Gaussian model to predict cyanobacterial pigment abundance in central Indiana reservoirs. *IEEE Journal of Selected Topics in Earth Observation and Remote Sensing,* in review.

Rowan, K.S. (1989). *Photosynthetic pigments of algae*. New York, New York: Cambridge University Press.

Sarada, R.M., Pillai, G., & Ravishankar, G.A. (1999). Phycocyanin from *Spirulina sp.*: influence of processing of biomass on phycocyanin yield, analysis of efficacy of extraction methods and stability studies on phycocyanin. *Process Biochemistry, 34*, 795-801.

Schalles, J. (2006). Optical Remote Sensing Techniques to Estimate Phytoplankton Chlorophyll *a* Concentrations in Costal Waters with Varying Suspended Matter and CDOM Concentrations. In L.L. Richardson & L. E.F. (Eds.), *Remote Sensing of Aquatic Costal Ecosystem Processes: Science and Management Applications* (pp. 27-75). Dordrecht, The Netherlands: Springer.

Schalles, J.F., Gitelson, A., Yacobi, Y.Z., & Kroenke, A. (1998). Estimation of chlorophyll *a* from time series measurements of high spectral resolution reflectance in a eutrophic lake. *Journal of Phycology, 34*, 383-390.

Schalles, J.F., & Yacobi, Y.Z. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid, and chlorophyll pigments in eutrophic waters. *Archives Hydrobiologica Special Issues on Advanced Limnology, 55*, 153-168.

Sengpiel R.E.(2007) *Using airborne hyperspectral imagery to estimate chlorophyll a and phycocyanin in three central Indiana mesotrophic to eutrophic reservoirs*. Indianapolis, Indiana University Purdue University Indianapolis M.S. Thesis.

Simis, S.G.H., Peters, S.W.M., & Gons, H.J. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography, 50*, 237-245.

Simis, S.G.H., Ruiz-Verdu, A., Dominguez, J.A., Pena-Martinez, R., Peters, S.W.M., & Gons, H.J. (2007). Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sensing of Environment, 106*, 414-427.

Strombeck, N., & Pierson, D.C. (2001). The effects of variability in the inherent optical properties on estimations of chlorophyll *a* by remote sensing in Swedish freshwaters. *The Science of the Total Environment, 268*, 123-137.

Sunshine, J.M., C.M. Pieters, S.F. Pratt and K.S. McNaron-Brown (1999). Absorption Band Modeling in Reflectance Spectra: Availability of the Modified Gaussian Model. *30th Annual Lunar and Planetary Science Conference. Abstract 1306*. Houston, TX.

Sunshine, J.M., Pieters, C.M., & Pratt, S.F. (1990). Deconvolution of Mineral Absorption Bands: An Improved Approach. *Journal of Geophysical Research, 95*, 6955-6966.

Vallely, L.A. (2008). *Confounding Constituents in Remote Sensing of Phycocyanin*. Indianapolis: Indiana University Purdue University Indianapolis M.S. Thesis.

Yacobi, Y., Gitelson, A., & Mayo, M. (1995). Remote sensing of chlorophyll in Lake Kinneret using high-spectral-resolution radiometer and Landsat TM: spectral features of reflectance and algorithm development. *Journal of Plankton Research, 17*, 2155-2173.

IV: USING A PARTIAL LEAST SQUARES (PLS) METHOD FOR ESTIMATING CYANOBACTERIAL PIGMENTS IN EUTROPHIC INLAND WATERS

Abstract

Midwestern lakes and reservoirs are commonly exposed to anthropogenic eutrophication. Cyanobacteria thrive in these nutrient rich-waters and some species pose a threat to humans through the production of toxins and other compounds that degrade the water quality. Managers for drinking water production are interested in the rapid identification of cyanobacterial blooms to minimize effects caused by harmful cyanobacteria. There is potential to monitor cyanobacteria through the remote sensing of two algal pigments: chlorophyll *a* (CHL) and phycocyanin (PC). Several empirical methods that develop spectral parameters (e.g., simple band ratio, band-tuning, and band absorption) sensitive to these two pigments and map reflectance to the pigment concentration have been used in a number of investigations using field-based spectroradiometers. This study tests a multivariate analysis approach partial least squares (PLS) regression for the estimation of CHL and PC. PLS models were trained with 35 spectra collected from three central Indiana reservoirs during a 2007 field campaign with dual-headed Ocean Optics USB4000 field spectroradiometers (355 – 802 nm, nominal 1.0 nm intervals), and CHL and PC concentrations of the corresponding water samples analyzed at Indiana University-Purdue University at Indianapolis. Validation of these models with 19 remaining spectra show that PLS (CHL $R^2 = 0.90$, slope = 0.91, RMSE = 20.61 μg/L; PC $R^2 = 0.65$, slope = 1.15, RMSE = 23.04. μg/L) performed equally well to the band tuning model based on Gitelson *et al.* 2005 (CHL: $R^2 = 0.75$, slope = 0.84, RMSE = 40.16 μg/L; PC: $R^2 = 0.59$, slope = 1.14, RMSE = 20.24 μg/L).

Introduction

Since cyanobacteria have been known to reduce surface water quality through the production of toxins (Bold and Wynne 1985; Chorus and Bartram 1999) and taste-altering chemicals (Chorus and Bartram 1999; Christensen *et al.* 2006), surface water managers are interested in rapid identification of cyanobacterial blooms. Current approach to estimating cyanobacterial abundance is to estimate cyanobacterial pigments chlorophyll *a* and phycocyanin from *in situ* reflectance (Dekker 1993; Gitelson *et al*. 1995; Randolph 2007; Randolph *et al.* 2008; Richardson 1996; Schalles 2006; Schalles and Yacobi 2000; Sengpiel 2007; Simis *et al*. 2005; Simis *et al*. 2007), and relate chlorophyll *a* to total algal biomass, phycocyanin to cyanobacteria abundance.

Most remote sensing models estimating chlorophyll *a* use the strong absorption at 660-666 nm due to the minimum influence of other pigments (Jeffrey *et al.* 1997; Jeffrey and Wright 2006; Rowan 1989). However, chlorophyll *a* does have additional absorption features. Chlorophyll *a* strongest absorption feature is in the spectral range of 428-432 nm; however, this range includes overlapping absorptions of carotenoids in the 400-500 nm range (Jeffrey *et al.* 1997; Rowan 1989). Chlorophyll *a* absorbs less strongly around 382.7 nm, 409-417.6 nm, 530-535.5 nm, 575-580.3 nm and 614-618.2 nm (Jeffrey *et al.* 1997). Cyanobacterial phycocyanin has strong absorption in the region of 612-628 nm along with a florescence maximum in the region of 632-651 nm (Rowan 1989). Cyanobacteria have also been shown to contain the related biliprotein phycoerythrin which absorbs 550-572 nm and fluoresces around 570-581 nm (Rowan 1989).

Traditionally the aquatic remote sensing community uses band ratios to estimate cyanobacterial pigments. Although the band ratio method may work on one reservoir, it

is not always transferable spatially due to differences in inherent optical properties (i.e. absorption and backscattering coefficients) (Dekker 1993; Randolph *et al* 2008). Dall'Olmo *et al*. (2003) attempted to address this issue by creating an algorithm that uses a step-wise regression method to determine the wavelengths of highest correlation to chlorophyll *a*. Gitelson *et al.* (2005) proposed a similar three-band tuning model with a slightly different starting strategy for the stepwise regression. The band tuning approach was validated in Dall'Olmo *et al*. (2003) and Gitelson *et al.* (2005), Gitelson *et al.* (2008) for estimating chlorophyll *a*. In all three studies a strong correlation between total and estimated chlorophyll content ($R^2 > 0.9$). Robertson (2009) proved that this approach is also valid for phycocyanin estimation.

Robertson (2009) recently compared several remote sensing methods for estimating chlorophyll *a* and phycocyanin, including empirical band ratios, band tuning, semi-empirical, and curve fitting models. Robertson (2009) found that empirical band ratios were the most effective and time saving method for estimating chlorophyll *a* from three eutrophic reservoirs located in central Indiana. This study did show that curve fitting (Robertson, 2009), 3 band tuning (Dall'Olmo *et al.,* 2003; Gitelson *et al.*, 2005; Gitelson *et al.*, 2008) also effective. Only semi-empirical models (Simis *et al.* 2005) did not perform equally well due to differences in inherent optical properties between reservoirs. The 3 band tuning method was the best performing method for estimating phycocyanin from these same reservoirs. The curve fitting (Robertson, 2009) method was also effective in estimating phycocyanin when spectral resolution was high (≤ 1 nm). Robertson (2009) also confirmed earlier studies that band ratios were less effective due to spectral overlapping of absorption features from different pigments near 620 nm.

This study seeks to test an alternative approach partial least squares (PLS) regression for estimating cyanobacterial pigments chlorophyll *a* and phycocyanin. In order to determine if PLS is an effective alternative to commonly used methods, this study will compare the results from PLS to two other methods: traditional band ratios, and a 3 band tuning method. These two methods were selected for the comparison because the band ratio proves an effective method for estimating chlorophyll *a* and the band tuning for phycocyanin.

Methods

Study Site

Three reservoirs located in central Indiana were investigated: Eagle Creek, Geist, and Morse. Figure 4-1 shows the location of each reservoir. Eagle Creek, Geist, and Morse Reservoirs provide water for > 800,000 residents of the greater Indianapolis area. All three reservoirs have similar characteristics including depth (3.2-4.7 m), surface area (5-7.5 km^2), volume (21-28 million m^3), watershed area (420-590 km^2) and residence time (55-70 days) (Li *et al.* 2006). They are also impaired by high nutrient loads (mean total P = 94-100 $\mu g*L^{-1}$, mean total N = 2-4.1 $\mu g*L^{-1}$) (Li *et al.* 2006) which promote the growth of nuisance algae, including cyanobacteria. All three reservoirs have reported taste and odor issues (IDEM 2006; Li *et al.* 2006).

In situ Reflectance Measurement

Number of samples for water samples and spectra signatures, and the pigment ranges for these samples are listed in Table 4-1. Samples were collected on each reservoir only once and were not taken with any other reservoirs concurrently. Spectra were collected facing the sun on a boat away from shore.

Figure 4 - 1. Central Indiana reservoirs: Map of the three study locations; Eagle Creek, Geist, and Morse Reservoirs. Indianapolis is indicated for reference.

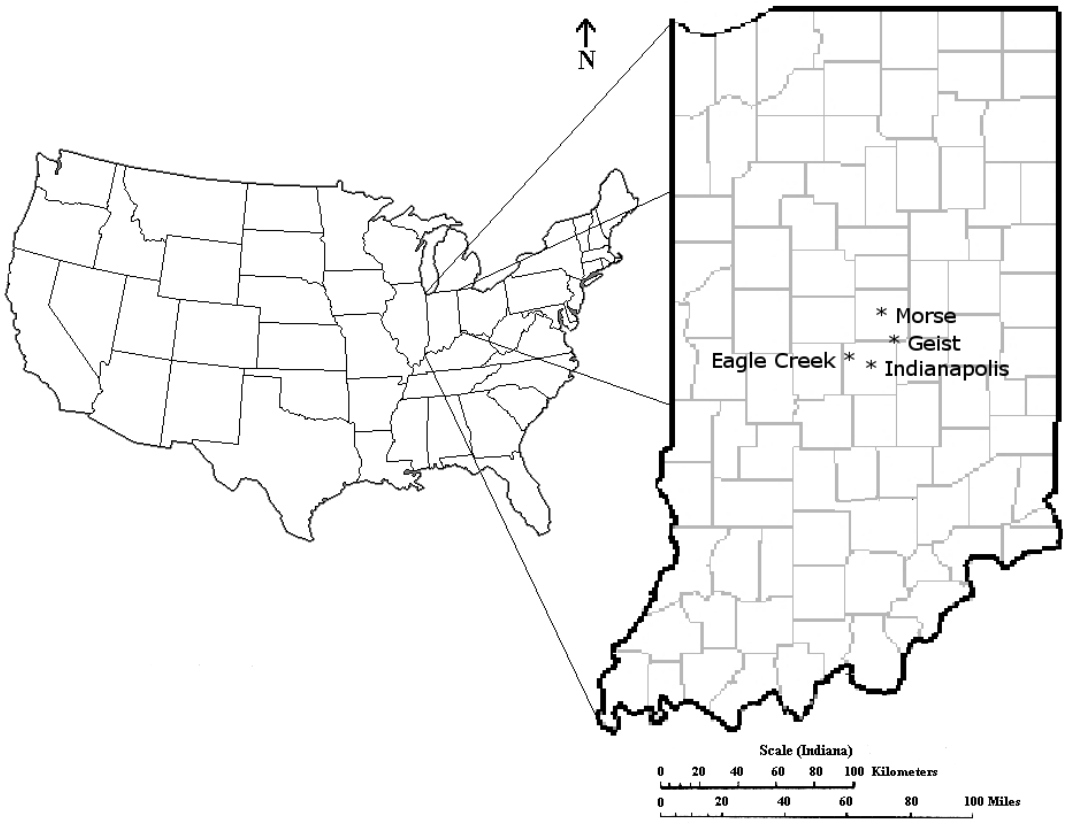


Table 4 - 1. Summary of sample collection: Includes the reservoir, number of samples (n) and pigment ranges.

Pigment	Reservoir								
	Eagle Creek			Geist			Morse		
	Sample Number (n)	Min. (µg/L)	Max. (µg/L)	Sample Number (n)	Min. (µg/L)	Max. (µg/L)	Sample Number (n)	Min. (µg/L)	Max. (µg/L)
Chl	22	16.9	255	16	14.45	193.18	14	26.9	203.8
PC	22	30.9	114.1	16	0.91	148.95	15	41.2	136.3

Water samples for pigment analysis were collected from surface water with 1 L amber HDPE bottles and stored on ice prior to filtration. The boat-based sensor used in this study includes a pair of Ocean Optics (OO) USB4000 visible and near infrared (V/NIR) spectroradiometers (Ocean Optics, Inc., Dunedin, FL, USA) in a dual head system. One USB 4000 spectroradiometer was set with a cosine corrector to determine down-welling radiance, thus reducing atmospheric effects. The second USB 4000

spectroradiometer was used to measure up-welling radiance from the target of interest. The USB 4000 spectroradiometer has a spectral resolution of approximately 0.2 nm and a spectra range of 351-1047 nm for a total of 3645 bands

The fiber optic cable for the USB 4000 spectroradiometers has an IFOV of 0.14 rad. The cable was mounted on an extendable pole with a nadir viewing angle and held approximately 0.5 m above the water surface. This produces a measured water surface area with a diameter of 0.07 m. Calibration for upwelling irradiance for the USB 4000 spectroradiometers was completed by using a 20% Spectrolon reflectance panel. In order to reduce noise in the spectra, the reflectance spectrum at each site was averaged over 8 readings. These spectra were averaged to 56 spectral bands ranging from 355 nm - 802 nm with an average spectral resolution of 7.8 nm.

Lab Analysis of Water Samples

Steps for pigment analysis were completed under subdued light conditions. Water samples were filtered in duplicate within 24 hours of the sample collection for later extraction of pigments. Samples for the extraction of chlorophyll *a* were prepared on 0.47 µm acetate filters, and those for the extraction of phycocyanin on 0.47 µm glass fiber filters (GFF). The samples were stored at -20°C for no longer than 6 months before analysis.

Chlorophyll *a* was extracted following the extraction method described in Environmental Protection Agency (EPA) 445 (Arar and Collins 1997). The concentration of chlorophyll *a* was corrected for pheophytin and measured fluormetrically using a TD-700 Fluorometer (Turner Designs, Inc.) fitted with a Daylight White Lamp and Chlorophyll Optical Kit (340-500 nm excitation filter and emission

filter > 665 nm). For chlorophyll *a* analysis the fluorometer was calibrated using chlorophyll *a* from a spinach standard (Sigma-Aldrich 10865).

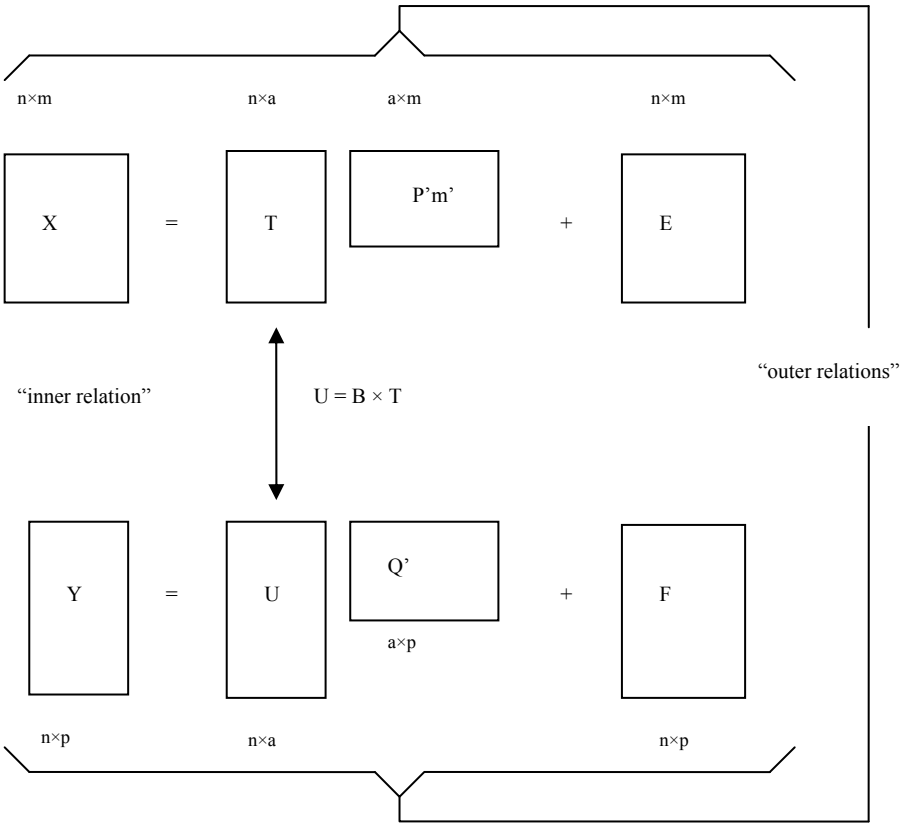
Phycocyanin was extracted based on the method described in Sarada (1999) and Randolph *et al.* (2008). Phycocyanin was measured fluormetrically using a TD-700 Fluorometer (Turner Designs, Inc.) fitted with a Cool White Mercury Vapor Lamp and a Phycocyanin Optical Kit (630 nm excitation and 660 nm emission filters). For phycocyanin analysis the fluorometer was calibrated using C-phycocyanin from *Spirulina sp.* (Sigma-Aldrich P6161).

If the percent error calculated between replicates in pigment analysis was larger than 20%, the sample was not used in data analysis.

Estimation of Pigment Abundance with PLS

Partial least squares (PLS) regression is a full spectrum multivariate statistical analysis tool developed by Wold (1966). In the 1980's, PLS gained popularity in the field of chemistry for spectral analysis (Geladi and Kowalski 1986; Haaland and Thomas 1988). The underlying assumption of a PLS model is that the reflectance spectra are driven by components or factors that are linear combinations of observed explanatory variables. PLS is designed to find as few eigenvectors of the explanatory variables as possible. These eigenvectors should produce score values that summarize the variance of the explanatory variables and are highly correlated with the response variables. PLS determines a few eigenvectors of the explanatory variables such that the corresponding scores not only explain the variance of the explanatory variables but also have high correlation to the response variables. A simplified PLS model (Figure 4-2) consists of two outer relations resulting from the eigenstructure decomposition of both the matrix

Figure 4 - 2. A diagram shows two outer relations and one inner relation of a PLS model



containing explanatory variables (i.e., spectral bands) and the matrix containing response variables (i.e., pigments or water content), and an inner relation that links the resultant score matrices from these two eigenstructure decompositions (Geladi and Kowalski, 1986). The goal of PLS modeling is to minimize the norm of *F* while maximizing the covariance between *X* and *Y* by the inner relation. This inner relation is a multiple linear regression between the score matrices *U* and *T* in which *B* is an $n \times n$ regression coefficient matrix determined via least square minimization. Haaland and Thomas (1988) provided a detailed description of PLS and Geladi and Kowalski (1986) published a tutorial for using PLS. Application of PLS has been utilized in a wide variety of remote sensing applications including but not limited to mapping/estimating soil components

(Chang and Laird 2002; Kooistra *et al*. 2001), lunar mineral composition (Li 2006, 2008), and plant/crop characteristics (Borregaard *et al*. 2000; Cho *et al*. 2007; Hansen and Schjoerring 2003; Nguyen and Lee 2006).

In this study we combined the data from all three reservoirs and randomly divided the data set into a calibration (n = 35) and validation (n = 19) data set. In the case of processing spectral data using PLS, the PLS_Toolbox ver. 3.5 (Eigenvector Research, Inc.) for MATLAB (The MathWorks, Inc.) was utilized. Reflectance spectra were converted to absorbance and preprocessed using mean center. To select the optimal number of PLS factors, this study used the 'leave one out' method as described in Haaland and Thomas (1988).

Estimation of Pigment Abundance with Band Ratios and Band Tuning Methods

Robertson (2009) found that band ratio algorithms represented by equations 4-1 and 4-2 performed better than other three other band ratio algorithms in a comparative study.

Chlorophyll *a*: $(R_{705} - R_{670})/ R_{670}$ (Mittenzwey *et al.* 1991) (eq. 4-1)

Phycocyanin: R_{709} / R_{620} (modified from Simis *et al.* 2005) (eq. 4-2)

Wavelengths utilized in the two 3-band tuning methods were calculated using a step-wise regression program written in MATLAB that follows the methods outlined in Dall'Olmo *et al*. (2003), Gitelson *et al.* (2005), and Robertson (2009), a summary of final wavelengths is located in Table 4-2. These wavelengths were then used to calibrate the model found in equation 4-3 for estimating pigment concentrations

$$(1/R_{\lambda 1} - 1/R_{\lambda 2}) * R_{\lambda 3} \quad \text{(eq. 4-3)}$$

Table 4 - 2. Wavelengths determined using band tuning methods: Final wavelengths as determined by stepwise regression methods described in Dall'Olmo *et al*. 2003 and Gitelson *et al.* 2005. Wavelength ranges are restrained as follows: $R_{\lambda 1(Chlorophyll)}$ = 660-690 nm, $R_{\lambda 1(Phycocyanin)}$ = 610-640, $R_{\lambda 2}$ = 690-730 nm, and $R_{\lambda 3}$ = 740-800 nm.

	Dall'Olmo *et al.*2003		Gitelson *et al.*2005	
Wavelengths	**Chl**	**PC**	**Chl**	**PC**
R1	664.1	632.4	695.5	632.4
R2	718.7	703.2	734	703.2
R3	764.4	764.4	764.4	764.4

Results and Discussion

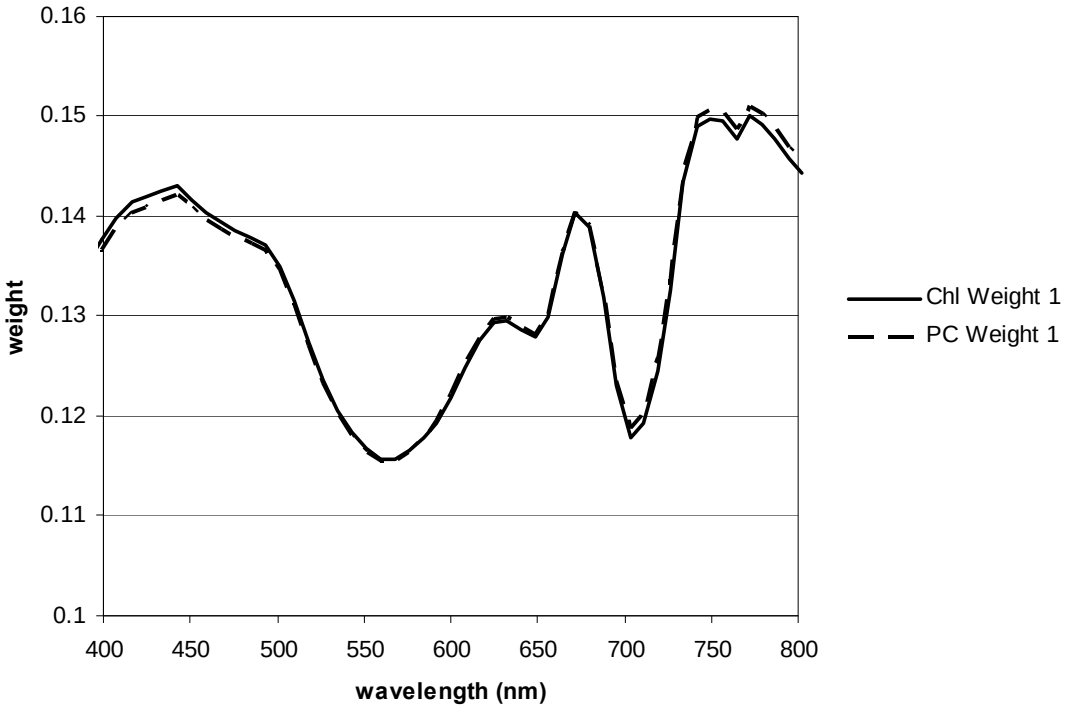
Calibration of the models

A summary of the calibration models for the band ratio and band tuning algorithms are in Table 4-3. All calibration models have similar correlation coefficients ($0.78 < R^2 < 0.87$). For the two PLS models, examining the weight loading vectors can provide insight to how PLS is working and determine which regions are more sensitive to pigment concentration. Figure 4-3 displays the first three weight loading vectors. For chlorophyll *a* estimation the first weight explain 87.19% of the total variation with only one additional weight explaining greater than 5% of the variation (9.90%). For phycocyanin estimation the first weight explained 99.84% of the total variation. No other weight explained greater than 5% of the variation.

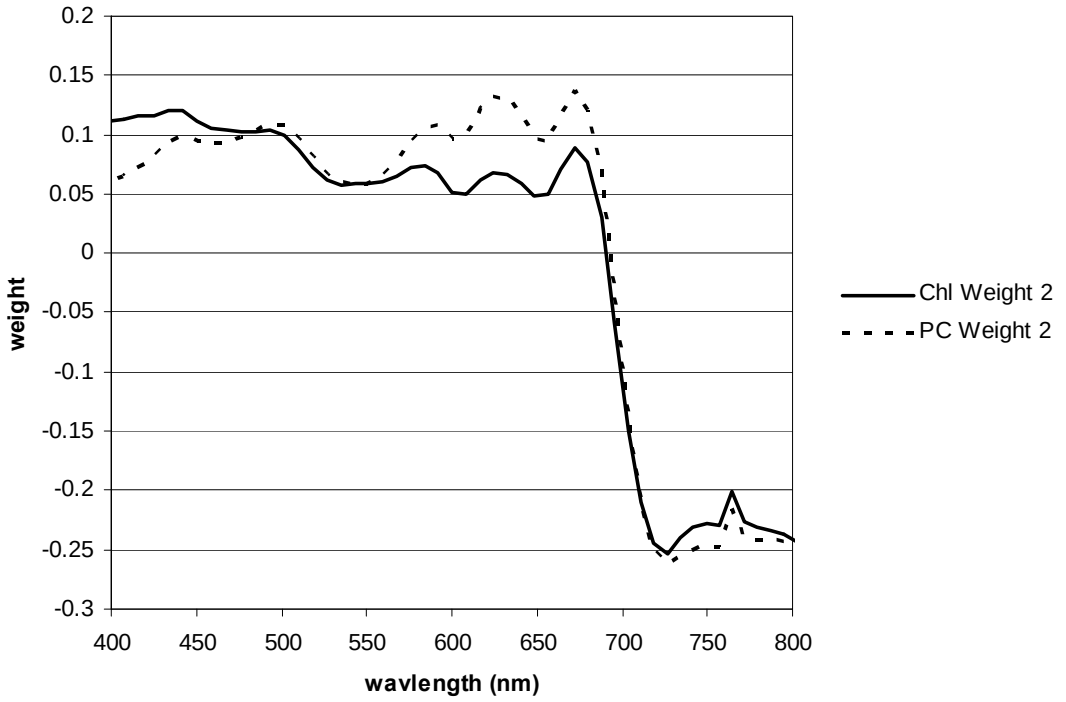
Table 4 - 3. Building models with the calibration data set: (n = 35) Models based on correlation between pigments and spectral indices/parameters. Where x = spectral index and y = pigment concentration.

Pigment	Spectral Parameter	R2	Model
Phycocyanin	Simis *et al*. 2005 Band Ratio	0.8703	y = 158.73x - 106.60
	Dall'Olmo *et al*. 2003 3 Band Tuning	0.8037	y = 400.00x + 30.60
	Gitelson *et al*. 2005 3 Band Tuning	0.7903	y = 400.00x + 30.60
Chlorophyll *a*	Mittenzwey *et al*. 1991 Band Ratio	0.7283	y = 357.14x - 20.14
	Dall'Olmo *et al*. 2003 3 Band Tuning	0.7618	y = 344.83x + 70.38
	Gitelson *et al*. 2005 3 Band Tuning	0.7795	y = 625x + 412.63

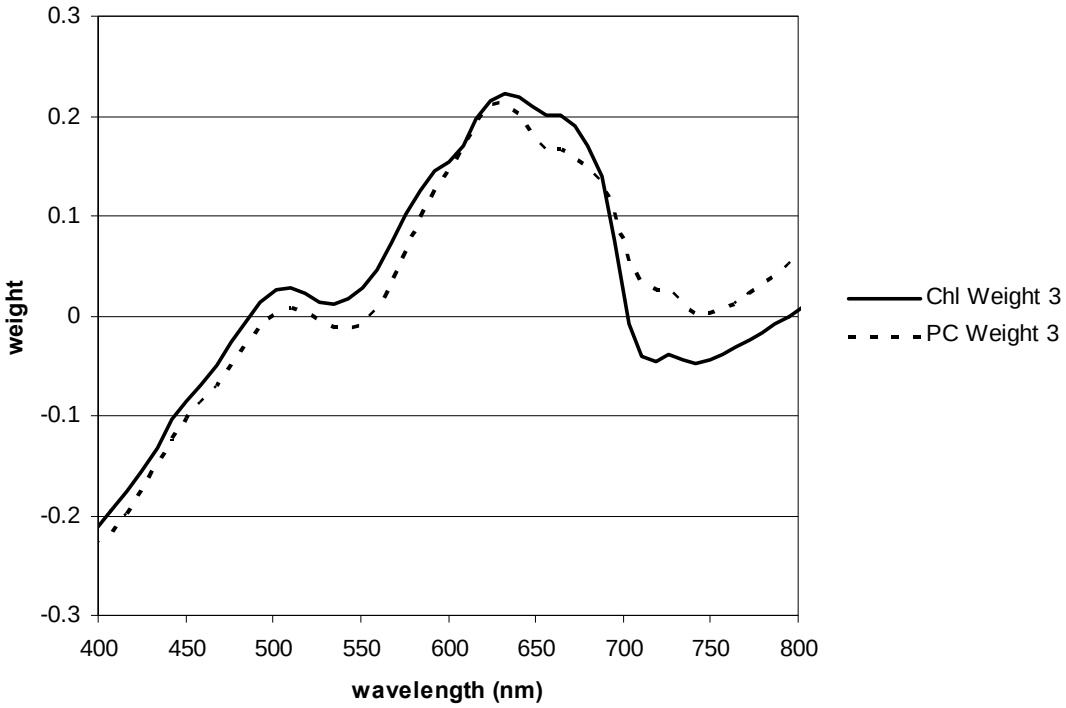
Figure 4 - 3. Weights for PLS models: a) weight 1, b) weight 2, and c) weight 3. For chlorophyll *a* estimation the weights explain 87.19%, 9.90%, and 1.84% of the variation respectively. For phycocyanin estimation the weights explained 99.84%, 0.11%, and 0.04% of the variation respectively.



a)



b)



c)

The shapes of the first weight for both chlorophyll *a* and phycocyanin estimation are nearly identical (Figure 4-3a). Haaland and Thomas (1986) describe the first weight as an average absorption spectra. The regions that most strongly correlate to pigment absorption are the regions where carotenoid/chlorophyll (400-500 nm), phycocyanin (610-640 nm), chlorophyll (660-680 nm), and water (730-780 nm) absorptions occur. Areas in the spectrum that have weaker correlation to pigment absorption include the 'green peak' (~550 nm) and the reflectance peak caused by cell scattering (~700 nm).

While the remaining two weights for pigment estimation are not as significant as the first, they are worth examining. In weight two (Figure 4-3b), there are still similarities between the chlorophyll *a* and phycocyanin weights; however, there are a few regions of deviation. For both pigments water absorption (730-800 nm) has a negative weight. This negative weight in the water absorption region is likely due to

compensation for the strong positive weight in the weight 1 that does not correlated to pigment estimation. Greater weight is attributed to estimating chlorophyll *a* in the 400-500 nm range with smaller peaks in weights at higher wavelengths. While the 400-500 nm region has strong chlorophyll absorptions, accessory pigments such as the carotenoids do influence reflectance spectra in this range. Without including carotenoid data, we cannot determine if the weight is more closely related to chlorophyll estimation or these accessory pigments. In contrast, the weight for phycocyanin estimation shows the highest values at the region of phycocyanin absorption (610-640 nm) and chlorophyll absorption. This consistence can be attributed to co-linearity of the pigments. This co-linearity originates from the cyanobacterial dominance in the waters investigated in this study where an increase in cyanobacterial abundance correlates to increases in both chlorophyll *a* and phycocyanin. In waters not dominated by cyanobacteria, it is likely the co-linearity between both pigments will not be as strong and the weights would likely change. The minor peak in weights near 580 nm is likely due to absorption and fluorescence of phycoerythrin that normally masked in the reflectance spectra but also increases with increased algal abundance.

The third weight explains very little of the PLS model; however, upon examining the shape of the third weight (Figure 4-3c), there are only minor differences between both chlorophyll *a* and phycocyanin estimation. The only region with a strong positive weight in both pigment estimations is between 550-700 nm. This region again is related to absorptions dominated by phycocyanin and chlorophyll absorption. In weight 3, the phycocyanin peak (610-640 nm) dominates both weights in the positive direction. While a near linear slope from 400 nm to 500 nm dominates in the negative direction. This

negative weight in this region is likely due to the correction of the positive weight for carotenoids in the previous two weights.

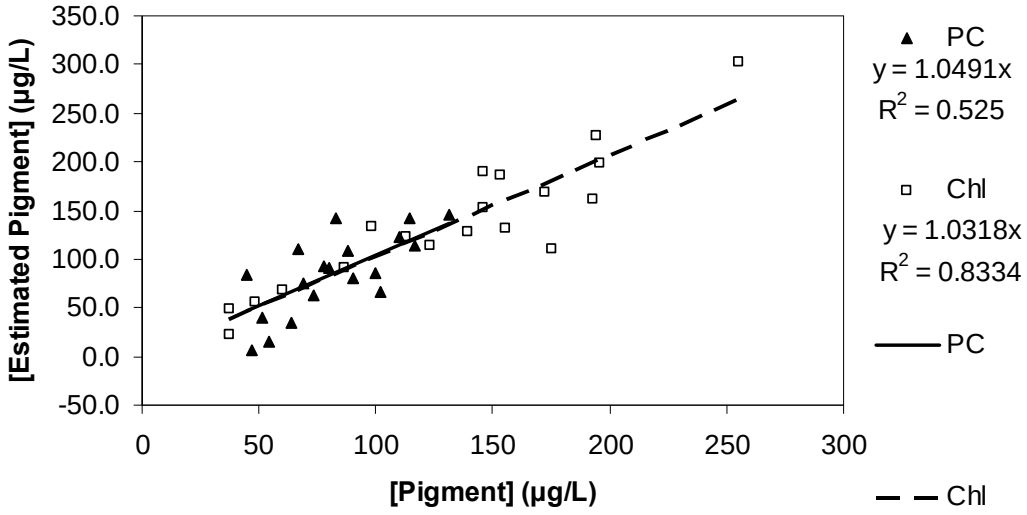
Validation of the models

A summary of the validation of the models can be found in Table 4-4. When the calibration models were applied to the validation data set, all models had high statistical significance ($p < 0.01$). As shown in Figure 4-4a, one exception was the band ratio based on the algorithm presented in Simis *et al.* (2005) ($p < 0.05$). All models predicted with a near linear relationship ($0.84 < slope < 1.15$) (Figure 4-4). For chlorophyll *a* estimation, the two best performing models were the 3 band tuning method from Dall'Olmo *et al.* 2003 and the PLS method and both resulted in nearly identical coefficients of determination ($R^2 = 0.90$) and root mean square error (RMSE ≈ 20 μg/L). . For phycocyanin estimation the PLS method had a slightly higher coefficient of determination ($R^2 = 0.65$) and lower error (RMSE = 23.04) as compared to the two 3 band tuning methods that yielded the coefficient of determination ($R^2 = 0.59$) and error (RMSE = 31.2 μg/L).

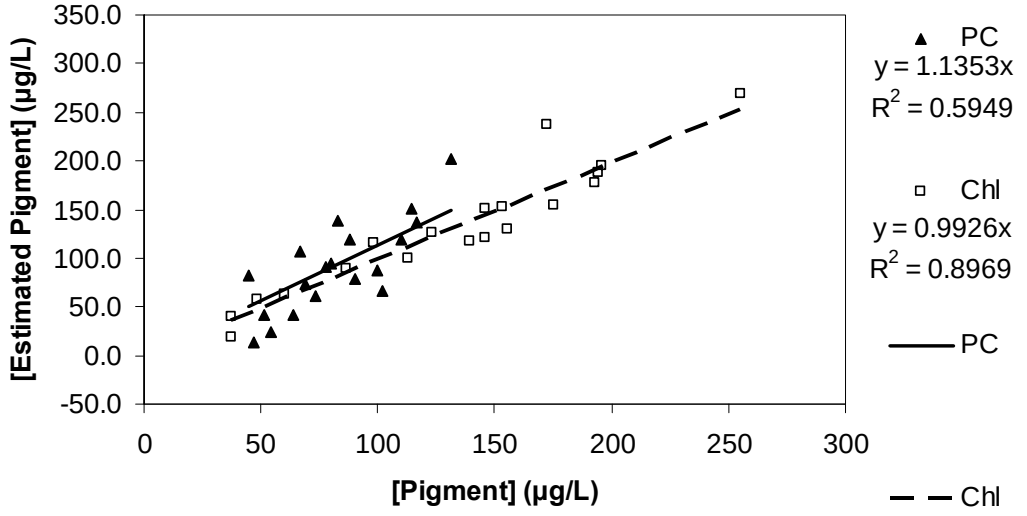
Table 4 - 4. Validation of the models: (n = 19) Intercept was forced through zero, therefore ideal models will have a 1) slope near one 2) correlation coefficient (R^2) near one and 3) low RMSE.

Pigment	Model	Slope	R^2	RMSE (ppb)	p
Phycocyanin	Simis *et al*. 2005 Band Ratio	1.0491	0.5250	27.74	< 0.05
	Dall'Olmo *et al*. 2003 3 Band Tuning	1.1353	0.5949	31.20	< 0.01
	Gitelson *et al*. 2005 3 Band Tuning	1.1353	0.5949	31.20	< 0.01
	PLS	1.1518	0.6514	23.04	< 0.01
Chlorophyll *a*	Mittenzwey *et al*. 1991 Band Ratio	1.0318	0.8334	27.22	< 0.01
	Dall'Olmo *et al*. 2003 3 Band Tuning	0.9926	0.8969	20.24	< 0.01
	Gitelson *et al*. 2005 3 Band Tuning	0.8359	0.7517	40.16	< 0.01
	PLS	0.9102	0.8955	20.61	< 0.01

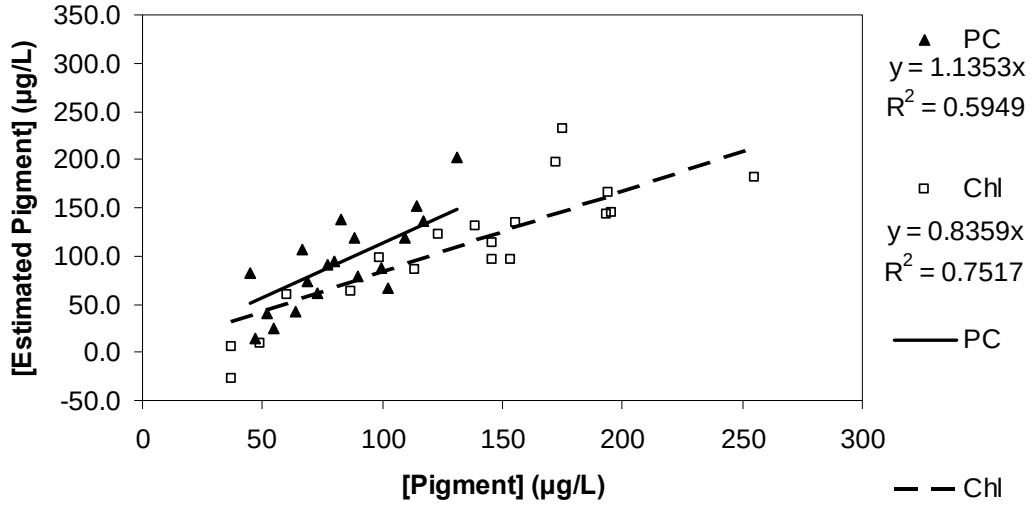
Figure 4 - 4. Estimating pigments with the validation data set: Correlations between actual and estimated pigment concentration (µg/L) from a) band ratio algorithms, b) Dall'Olmo *et al.* (2003) 3-band tuning, c) Gitelson *et al.* (2005) 3-band tuning, d) Partial Least Squares (PLS).



a)

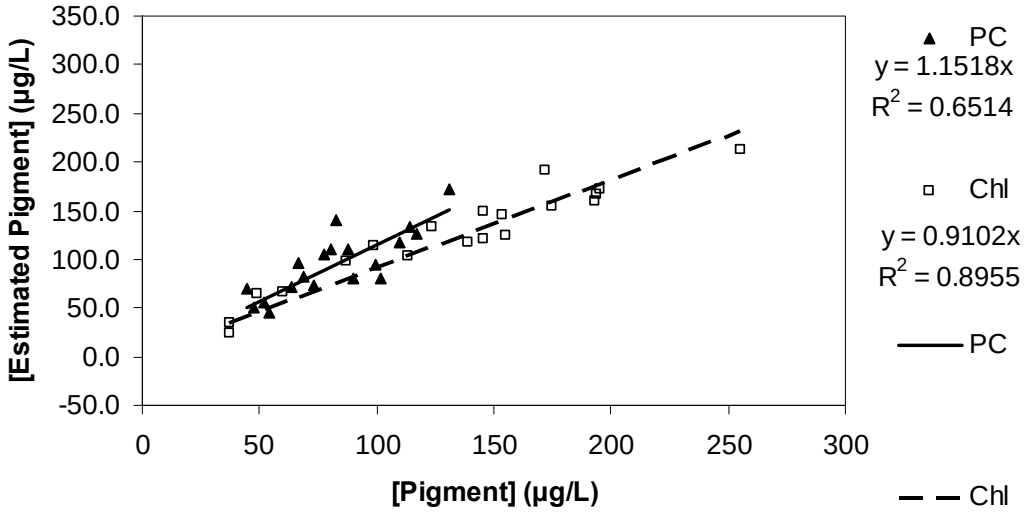


b)



c)

d)



Both the band tuning method and PLS for chlorophyll *a* estimation performed equally well in estimating chlorophyll *a*. Since chlorophyll *a* has a region of strong absorption with little overlap of accessory pigments, fewer wavelengths are necessary for its estimation. In contrast, the slight improvement in estimating phycocyanin with PLS

may occur because of spectral overlapping in the region where phycocyanin absorbs the strongest (~620 nm). Since PLS incorporates wavelengths from the entire spectrum, it can better estimate phycocyanin. While the band tuning method is limited to three bands, it performs nearly as well, indicating that this model explains nearly as much of the variation in phycocyanin estimation as the PLS model.

While the three band tuning model works well in this study, Le *et al*. (2009) has shown that an additional band is necessary for modeling on highly turbid lakes. Since PLS is a full spectrum model, additional wavelengths should not be necessary to build a model in these conditions. In contrast, PLS contains redundant and unnecessary information because it does incorporate the entire spectrum into the model. If it is determined that a 3 band model works well, it is possible to use a smaller spectral dataset to build these models.

Conclusion

In this study the 3 band tuning and PLS methods are preferred over traditional band ratios for estimating phycocyanin and chlorophyll *a*. When comparing between the two complex models tested in this study neither the 3 band tuning nor the PLS method outperformed the other. Both methods perform nearly equal in this study for predicting cyanobacterial pigments. Both models are recommended for future development

Acknowledgements

We would like to thank Veolia Water Indianapolis and the Central Indiana Water Resources Partnership for financial support. We would also like to thank Lani Pascual, Kate Randolph, Becky Sengpiel, Lara Vallely, Bob Hall, Vince Hernly, and other staff, graduate students and interns involved with the collection and processing of the data sets.

References

Arar, E.J., & Collins, G.B. (1997). In Vitro Determination of Chlorophyll *a* and Pheophytin a in Marine and Freshwater Algae by Fluorescence. In. Cincinnati, OH.

Bold, H.C., & Wynne, M.J. (1985). *Introduction to the Algae: structure and reproduction*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Borregaard, T., Nielsen, H., Norgaard, L., & Have, H. (2000). Crop-weed Discrimination by Line Imaging Spectroscopy. *Journal of Agricultural Engineering Research, 75*, 389-400.

Chang, C.-W., & Laird, D.A. (2002). Near-infrared reflectance spectroscopic analysis of soil C and N. *Soil Science, 167*, 110-116.

Cho, M.A., Skidmore, A., Corsi, F., Van Wieren, S.E., & Sobhan, I. (2007). Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation, 9*, 414-424.

Chorus, I., & Bartram, J. (1999). *Toxic Cyanobacteria in Water: A guide to their public health consequences, monitoring and management*. Bury St Edmunds, Suffolk: St. Edmundsbury Press.

Christensen, V.G., Graham, J.L., Milligan, C.R., Pope, L.M., & Ziegler, A.C. (2006). Water quality and relation to taste and odor compounds in the North Fork Ninnescah River, and Cheney Reservoir, south-central Kansas. In, *U.S. Geologcial Survey Scientific Investigations Report 2006.*

Dall'Olmo, G., & Gitelson, A. (2005). Effect of bio-optical parameter variability on the remote estimation of chlorophyll *a* concentration in turbid productive waters: experimental results. *Applied Optics, 44*, 412-422.

Dall'Olmo, G., Gitelson, A., & Rundquist, D. (2003). Towards a unified approach for remote estimation of chorophyll *a* in both terrestrial vegetation and turbid productive waters. *Geophysical Research Letters, 30*, 1938-1942.

Dekker, A.G. (1993). Detection of the optical water quality parameters for eutrophic waters by high resolution remote sensing. In. Ansterdam: Free University

Geladi, P., & Kowalski, B.R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta, 185*, 1-17.

Gitelson, A., Dall'Olmo, G., Moses, M., Rundquist, D., Barrow, T., Fisher, T.R., Gurlin, D., & Holz, J. (2008). A simple semi-analytical model for remote estimation of chlorophyll *a* in turbid waters: Validation. *Remote Sensing of Environment, 112.* 3582-3593.

Gitelson, A.A., Laorawat, S., Keydan, G.P., & Vonshak, A. (1995). Optical properties of dense algal cultures outdoors and their application to remote estimation of biomass and pigment concentration in *Spirulina Platensis* (cyanobacteria). *Journal of Phycology, 31*, 828-834.

Gitelson, A.A., Vina, A., Ciganda, V., Rundquist, D.C., & Arkebauer, T.J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters, 32*, 1-4.

Haaland, D.M., & Thomas, E.V. (1988). Partial least-squares methods for spepctral analyses. 1. relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry, 60*, 1193-1202.

Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment, 86*, 542-553.

IDEM (2006). Indiana Integrated Water Monitoring and Assessment Report 2006. In. Indianapolis: Indiana Department of Environmental Management.

Kooistra, L., Wehrens, R., Leuven, R.S.E.W., & Buydens, L.M.C. (2001). Possibilities of visible–near-infrared spectroscopy for the assessment of soil contamination in river floodplains. *Analytica Chimica Acta, 446*, 97-105.

Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., and Lu, H. (2009) A four-band semi-analytical model for estimating chlorophyll *a* in highly turbid lakes : The case of Taihu Lake, China. *Remote Sensing of the Environment, 113,* 1175-1182.

Li, L. (2006). Partial least squares modeling to quantify lunar soil composition with hyperspectral reflectance measurements. *Journal of Geophysical Research, 111*, 1-13 E04002.

Li, L. (2008). Partial least squares methods for spectrally estimating lunar soil FeO abundance: A stratified approach to revealing nonlinear effect and qualitative interpretation. *Journal of Geophysical Research, 113,* E12013.

Li, L., Pascual, D.L., Tedesco, L.P., Randolph, K.L., Sengpiel, R.E., Hall, B.E., & Wilson, J.S. (2006). Developing a survey tool for the rapid assessment of blue-green algae in central Indiana's reservoirs. In, *2005 LARE Grant Report*: Indiana Department of Natural Resources.

Nguyen, H.T., & Lee, B.-W. (2006). Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression. *European Journal of Agronomy, 24*, 349-356.

Randolph K.L. (2008). *Remote sensing of cyanobacteria in case II waters using optically active pigments, chlorophyll a and phycocyanin* Indianapolis: Indiana University Purdue University Indianapolis M.S. Thesis.

Randolph, K.L., Wilson, J.S., Tedesco, L.P., Lin, L., Pascual, D.L., & Soyeux, E. (2008). Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll *a* and phycocyanin *Remote Sensing of Environment, 112*, 4009-4019.

Richardson, L.L. (1996). Remote Sensing of Algal Bloom Dynamics. *BioScience, 46*, 492-501.

Robertson, A.L. (2009) *Using band ratio, semi-empirical, curve fitting, and partial least squares (PLS) models to estimate cyanobacterial pigment concentration from hyperspectral reflectance* Indianapolis: Indiana University Purdue University Indianapolis M.S. Thesis.

Sarada, R.M., Pillai, G., & Ravishankar, G.A. (1999). Phycocyanin from *Spirulina sp.*: influence of processing of biomass on phycocyanin yield, analysis of efficacy of extraction methods and stability studies on phycocyanin. *Process Biochemistry, 34*, 795-801.

Schalles, J. (2006). Optical Remote Sensing Techniques to Estimate Phytoplankton Chlorophyll *a* Concentrations in Costal Waters with Varying Suspended Matter and CDOM Concentrations. In L.L. Richardson & L. E.F. (Eds.), *Remote Sensing of Aquatic Costal Ecosystem Processes: Science and Management Applications* (pp. 27-75). Dordrecht, The Netherlands: Springer.

Schalles, J.F., Gitelson, A., Yacobi, Y.Z., & Kroenke, A. (1998). Estimation of chlorophyll *a* from time series measurements of high spectral resolution reflectance in a eutrophic lake. *Journal of Phycology, 34*, 383-390.

Schalles, J.F., & Yacobi, Y.Z. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid, and chlorophyll pigments in eutrophic waters. *Archives Hydrobiologica Special Issues on Advanced Limnology, 55*, 153-168.

Sengpiel R.E.(2007) *Using airborne hyperspectral imagery to estimate chlorophyll a and phycocyanin in three central Indiana mesotrophic to eutrophic reservoirs*. Indianapolis, Indiana University Purdue University Indianapolis M.S. Thesis.

Simis, S.G.H., Peters, S.W.M., & Gons, H.J. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography, 50*, 237-245.

Simis, S.G.H., Ruiz-Verdu, A., Dominguez, J.A., Pena-Martinez, R., Peters, S.W.M., & Gons, H.J. (2007). Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sensing of Environment, 106*, 414-427.

Wold, H. (1966). Nonlinear estimation by iterative least squares procedure. In F. David (Ed.), *Research Papers in Statistics* (pp. 441-444). New York: John Wiley.

V: CONCLUSION

The results from studies presented in chapters 2-4 confirm that hyperspectral remote sensing of cyanobacterial pigments is feasible in multiple inland water bodies. Both chlorophyll *a* and phycocyanin showed strong correlation between the estimated and actual pigment values with several of the algorithms.

In all three studies, chlorophyll *a* estimation was similar with all tested algorithms. While complex models may perform slightly better in some cases, the benefit of using traditional band ratios is ease and quickness in applying the models to new data sets in estimating pigment concentrations. Therefore, these studies suggest using traditional band ratios for predicting chlorophyll *a* concentrations in eutrophic inland waters.

While minor differences in methods make it difficult to directly compare the results from the curve fitting and PLS models, these studies show that phycocyanin is better estimated using complex models such as the curve fitting, PLS, and band tuning. The curve fitting models are limited by spectral resolution and are best applied in cases where very high spectral resolution is available (≈1 nm). PLS and band tuning methods perform equally in similar conditions. The band ratio algorithm has an advantage over PLS in that once the calibration wavelengths are determined; pigment estimations can be derived quickly and easily using traditional band ratios. Either model is recommended for estimating phycocyanin. Situational needs should be considered in determining the method used by water managers.

These algorithms performed well when applied to Eagle Creek and Morse Reservoirs; however, issues arose when some of these models were applied to Geist Reservoir. No algorithm was able to overcome issues attributed to the anthropogenic disturbances in Geist Reservoir that affect water characteristics. In order to address this

issue, current algorithms will have to be improved or a different approach, such as bio-optical modeling, should be investigated.

VI: REFERENCES

Arar, E.J., & Collins, G.B. (1997). In Vitro Determination of Chlorophyll *a* and Pheophytin a in Marine and Freshwater Algae by Fluorescence. In. Cincinnati, OH.

Bold, H.C., & Wynne, M.J. (1985). *Introduction to the Algae: structure and reproduction*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Borregaard, T., Nielsen, H., Norgaard, L., & Have, H. (2000). Crop-weed Discrimination by Line Imaging Spectroscopy. *Journal of Agricultural Engineering Research, 75*, 389-400.

Chang, C.-W., & Laird, D.A. (2002). Near-infrared reflectance spectroscopic analysis of soil C and N. *Soil Science, 167*, 110-116.

Cho, M.A., Skidmore, A., Corsi, F., Van Wieren, S.E., & Sobhan, I. (2007). Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation, 9*, 414-424.

Chorus, I., & Bartram, J. (1999). *Toxic Cyanobacteria in Water: A guide to their public health consequences, monitoring and management*. Bury St Edmunds, Suffolk: St. Edmundsbury Press.

Christensen, V.G., Graham, J.L., Milligan, C.R., Pope, L.M., & Ziegler, A.C. (2006). Water quality and relation to taste and odor compounds in the North Fork Ninnescah River, and Cheney Reservoir, south-central Kansas. In, *U.S. Geologcial Survey Scientific Investigations Report 2006.*

Cole, T.M., & Tillman, D.H. (1999). Water Quality Modeling of Lake Monroe Using CE-QUAL-W2. In, *Waterways Experiment Station*. Vicksburg, MS: US Army Corps of Engineers Dall'Olmo, G., & Gitelson, A. (2005). Effect of bio-optical parameter variability on the remote estimation of chlorophyll *a* concentration in turbid productive waters: experimental results. *Applied Optics, 44*, 412-422.

Combe, J.-P., Launeau, P., Carrere, V., Despan, D., Meleder, V., Barille, L., & Sotin, C. (2005). Mapping microphytobenthos biomass by non-linear inversion of visible-infrared hyperspectral images. *Remote Sensing of Environment, 98*, 371-387.

Constantino (1999). In the Lap of Luxury. *Pit and Quarry* Cleveland, OH.

Dall'Olmo, G., & Gitelson, A. (2005). Effect of bio-optical parameter variability on the remote estimation of chlorophyll *a* concentration in turbid productive waters: experimental results. *Applied Optics, 44*, 412-422.

Dall'Olmo, G., Gitelson, A., & Rundquist, D. (2003). Towards a unified approach for remote estimation of chorophyll *a* in both terrestrial vegetation and turbid productive waters. *Geophysical Research Letters, 30*, 1938-1942.

Dekker, A.G. (1993). Detection of the optical water quality parameters for eutrophic waters by high resolution remote sensing. In. Ansterdam: Free University

Delwiche, C.F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *The American Naturalist, 154*, S164-177.

Falconer, I.R. (2005). *Cyanobacterial Toxins of Drinking Water Supplies*. Boca Raton, FL: CRC Press.

Geladi, P., & Kowalski, B.R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta, 185*, 1-17.

Gitelson, A.A. (1992). The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. *International Journal of Remote Sensing, 13*, 3367 – 3373.

Gitelson, A., Dall'Olmo, G., Moses, M., Rundquist, D., Barrow, T., Fisher, T.R., Gurlin, D., & Holz, J. (2008). A simple semi-analytical model for remote estimation of chlorophyll *a* in turbid waters: Validation. *Remote Sensing of Environment, 112.*

Gitelson, A.A., Laorawat, S., Keydan, G.P., & Vonshak, A. (1995). Optical properties of dense algal cultures outdoors and their application to remote estimation of biomass and pigment concentration in *Spirulina Platensis* (cyanobacteria). *Journal of Phycology, 31*, 828-834.

Gitelson, A., Nikanorov, A., Sabo, G., & Szilagyi, F. (1986). Etude de la qualite des eaux de surface par teledetection. *IAHS Publications, 157*, 111-121.

Gitelson, A.A., Vina, A., Ciganda, V., Rundquist, D.C., & Arkebauer, T.J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters, 32*, 1-4.

Haaland, D.M., & Thomas, E.V. (1988). Partial least-squares methods for spepctral analyses. 1. relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry, 60*, 1193-1202.

Han, L. (1997). Spectral Reflectance with Varying Suspended Sediment Concentrations in Clear and Algae-Laden Waters. *Photogrammetric Engineering and Remote Sensing, 63*, 701-705.

Han, L., & Rundquist, D. (1994). The Response of both Surface Reflectance and the Underwater Light Field to Various Levels of Suspended Sediments: Preliminary Results. *Photogrammetric Engineering and Remote Sensing, 60*, 1463-1471.

Hansen, P.M., & Schjoerring, J.K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment, 86*, 542-553.

Hiroi, T., Taylor, L.A., Morris, R.V., Keller, L.P., McKay, D.S., & Wentworth, S. (2000). Initial Results of MGM Analysis on Apollo 17 Soil Suite. In, *Lunar and Planetary Science XXXI(CD_ROM), abstract 1880, Houston, Texas.*

IDEM (2006). Indiana Integrated Water Monitoring and Assessment Report 2006. In. Indianapolis: Indiana Department of Environmental Management.

Jeffrey, S.W., Mantoura, R.F.C., & Wright, S.W. (1997). *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods*. Paris: UNSECO.

Jeffrey, S.W., & Wright, S.W. (2006). *Photosynthetic Pigments in Marine Microalgae: Insights from Cultures and the Sea*. Enfield, NH: Science Publishers.

Kaper, H.G., Smits, D.W., Schwarz, U., Takakuro, K., & Woerden, H.V. (1966). Computer Analysis of Observed Distributions into Gaussian Components. *Bull. Astr. Inst. Netherlands, 18*, 465-487.

Kooistra, L., Wehrens, R., Leuven, R.S.E.W., & Buydens, L.M.C. (2001). Possibilities of visible–near-infrared spectroscopy for the assessment of soil contamination in river floodplains. *Analytica Chimica Acta, 446*, 97-105.

Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., and Lu, H. (2009) A four-band semi-analytical model for estimating chlorophyll *a* in highly turbid lakes : The case of Taihu Lake, China. *Remote Sensing of the Environment, 113,* 1175-1182.

Li, L. (2006). Partial least squares modeling to quantify lunar soil composition with hyperspectral reflectance measurements. *Journal of Geophysical Research, 111*, 1-13 E04002.

Li, L. (2008). Partial least squares methods for spectrally estimating lunar soil FeO abundance: A stratified approach to revealing nonlinear effect and qualitative interpretation. *Journal of Geophysical Research, 113,* E12013.

Li, L., Pascual, D.L., Tedesco, L.P., Randolph, K.L., Sengpiel, R.E., Hall, B.E., & Wilson, J.S. (2006). Developing a survey tool for the rapid assessment of blue-green algae in central Indiana's reservoirs. In, *2005 LARE Grant Report*: Indiana Department of Natural Resources.

Li, L., Sengpiel, R.E., Pascual, D.L., Tedesco, L.P., Wilson, J.S., & Soyeux, E. (2009). Hyperspectral Remote Sensing to Estimate Chlorophyll *a* and Phycocyanin in a Mesotrophic Reservoir. *International Journal of Remote Sensing*, in press.

Lohrenz,S. E., Weidemann, A.D., & Tuel, M. (2003). Phytoplankton spectral absorption as influenced by community size structure and pigment composition. *Journal of Plankton Research, 25*, 35-61.

Mittenzwey, K.-H., Breitwieser, S., Penig, J., Gitelson, A.A., Dubovitzkii, G., Garbusov, G., Ullrich, S., Vobach, V., & Müller, A. (1991). Fluorescence and Reflectance for the *in-situ* Determination of Some Quality Parameters of Surface Waters. *Acta Hydrochimica et Hydrobiologica, 19*, 1-15.

Mustard, J.F., Poulet, F., Gendrin, A., Bibring, J.-P., Langevin, Y., Gondet, B., Mangold, N., Bellucci, G., & Altieri, F. (2005). Olivine and Pyroxene Diversity in the Crust of Mars. *Science, 307*, 1594-1597.

Nguyen, H.T., & Lee, B.-W. (2006). Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression. *European Journal of Agronomy, 24*, 349-356.

Pieters, C.M., & McFadden, L.A. (1994). Meteorite and Asteroid Reflectance Spectroscopy: Clues to Early Solar System Processes. *Annual Review of Earth and Planetary Sciences, 22*, 457-497.

Randolph K.L. (2008). *Remote sensing of cyanobacteria in case II waters using optically active pigments, chlorophyll a and phycocyanin* Indianapolis: Indiana University Purdue University Indianapolis M.S. Thesis.

Randolph, K.L., Wilson, J.S., Tedesco, L.P., Lin, L., Pascual, D.L., & Soyeux, E. (2008). Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll *a* and phycocyanin *Remote Sensing of Environment, 112*, 4009-4019.

Richardson, L.L. (1996). Remote Sensing of Algal Bloom Dynamics. *BioScience, 46*, 492-501.

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2009). Using a modified Gaussian model to predict cyanobacterial pigment abundance in central Indiana reservoirs. *IEEE Journal of Selected Topics in Earth Observation and Remote Sensing,* in review.

Rowan, K.S. (1989). *Photosynthetic pigments of algae*. New York, New York: Cambridge University Press.

Sarada, R.M., Pillai, G., & Ravishankar, G.A. (1999). Phycocyanin from *Spirulina sp.*: influence of processing of biomass on phycocyanin yield, analysis of efficacy of extraction methods and stability studies on phycocyanin. *Process Biochemistry, 34*, 795-801.

Schalles, J. (2006). Optical Remote Sensing Techniques to Estimate Phytoplankton Chlorophyll *a* Concentrations in Costal Waters with Varying Suspended Matter and CDOM Concentrations. In L.L. Richardson & L. E.F. (Eds.), *Remote Sensing of Aquatic Costal Ecosystem Processes: Science and Management Applications* (pp. 27-75). Dordrecht, The Netherlands: Springer.

Schalles, J.F., Gitelson, A., Yacobi, Y.Z., & Kroenke, A. (1998). Estimation of chlorophyll *a* from time series measurements of high spectral resolution reflectance in a eutrophic lake. *Journal of Phycology, 34*, 383-390.

Schalles, J.F., & Yacobi, Y.Z. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid, and chlorophyll pigments in eutrophic waters. *Archives Hydrobiologica Special Issues on Advanced Limnology, 55*, 153-168.

Sengpiel R.E.(2007) *Using airborne hyperspectral imagery to estimate chlorophyll a and phycocyanin in three central Indiana mesotrophic to eutrophic reservoirs*. Indianapolis, Indiana University Purdue University Indianapolis M.S. Thesis.

Simis, S.G.H., Peters, S.W.M., & Gons, H.J. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography, 50*, 237-245.

Simis, S.G.H., Ruiz-Verdu, A., Dominguez, J.A., Pena-Martinez, R., Peters, S.W.M., & Gons, H.J. (2007). Influence of phytoplankton pigment composition on remote sensing of cyanobacterial biomass. *Remote Sensing of Environment, 106*, 414-427.

Strombeck, N., & Pierson, D.C. (2001). The effects of variability in the inherent optical properties on estimations of chlorophyll *a* by remote sensing in Swedish freshwaters. *The Science of the Total Environment, 268*, 123-137.

Sunshine, J.M., C.M. Pieters, S.F. Pratt and K.S. McNaron-Brown (1999). Absorption Band Modeling in Reflectance Spectra: Availability of the Modified Gaussian Model. *30th Annual Lunar and Planetary Science Conference. Abstract 1306*. Houston, TX.

Sunshine, J.M., Pieters, C.M., & Pratt, S.F. (1990). Deconvolution of Mineral Absorption Bands: An Improved Approach. *Journal of Geophysical Research, 95*, 6955-6966.

Vallely, L.A. (2008). *Confounding Constituents in Remote Sensing of Phycocyanin*. Indianapolis: Indiana University Purdue University Indianapolis M.S. Thesis.

Wold, H. (1966). Nonlinear estimation by iterative least squares procedure. In F. David (Ed.), *Research Papers in Statistics* (pp. 441-444). New York: John Wiley.

Yacobi, Y., Gitelson, A., & Mayo, M. (1995). Remote sensing of chlorophyll in Lake Kinneret using high-spectral-resolution radiometer and Landsat TM: spectral features of reflectance and algorithm development. *Journal of Plankton Research, 17*, 2155-2173.

CURRICULUM VITAE

Anthony Lawrence Robertson

Academic Background:

Indiana University ~ Purdue University Indianapolis, Indianapolis, IN
Master of Science in Earth Science July 2009
Thesis: Using band ratio, semi-empirical, curve fitting and partial least squares (PLS) models to estimate cyanobacterial pigment concentration from hyperspectral reflectance

- Veolia Water Indianapolis Research Assistantship
- Department of Earth Science Travel Award
- Graduate Student Office Travel Fellowship
- Graduate Student Organization Education Enhancement Grant
- School of Science Graduate Student Council Travel Award

Indiana University Northwest, Gary, IN
Teaching Certification May 2005

Purdue University, West Lafayette, IN
Bachelor of Science in Biochemistry, Minor in Biology May 2003

- Officer and member of Biochemistry Club:
  Elementary Outreach Coordinator

Research:

IUPUI, Indianapolis, IN Summer 2006, May 2007-July 2009

- Remote sensing of cyanobacteria in Case II waters
- Train and coordinate seven undergraduate interns
- Mentor undergraduate senior research project
- Techniques: field water collection and treatment, boat-based spectra collection with multiple instruments, field backscattering measurements, pigment quantification using a fluorometer, spectrophotometeric measurements with cuvettes and integrating sphere, determination of organic/inorganic percentage, high performance liquid chromatography, and algal cell culturing

Purdue University, West Lafayette, IN May 2002-May 2003

- Proteomic study of protein phosphatase 5
- Maintained laboratory stocks of chemical reagents and disposable equipment
- Techniques: polymerase chain reaction, molecular cloning, tissue culturing, agarose electrophoresis, polyacrylamide gel electrophoresis, western blotting, bacterial culturing, DNA purification, transformation, and protein assays: radioactive and spectrometric assays

Research (cont.):

Purdue University, West Lafayette, IN January 2001-May 2002

- Sexual selection using *Nauphoeta cinerea* (lobster cockroach) and *Ambystoma texanum* (small mouth salamander)
- Designed and implemented a research project with *N. cinerea*
- Techniques: laboratory animal maintenance, animal capture, behavioral analysis, and statistical analysis

Professional Memberships:

American Society of Photometry and Remote Sensing (ASPRS)	2007-present
American Geophysical Union (AGU)	2008-present
International Society of Optical Engineering (SPIE)	2009-present
Indiana Geographic Information Council, Inc. (IGIC)	2009-present

Participation in Professional Meetings and Presentations:

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2009) Using a Partial Least Squares (PLS) Method in Predicting Cyanobacterial Pigments in Eutrophic Inland Waters. SPIE Optics and Photonics. San Diego, CA August (Oral Presentation). Abstract 7454-7

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2009) Comparing band ratio, semi-empirical, modified Gaussian models in predicting cyanobacterial pigments in eutrophic inland waters. UCOWR/NIWR Conference on Urban Water Management: Issues and Opportunities. Chicago, IL July (Oral Presentation). Abstract 15-3

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2008) Comparing band ratio, semi-empirical, modified Gaussian models in predicting cyanobacterial pigments in eutrophic inland waters. American Geophysical Union. San Francisco, CA December (Oral Presentation). Abstract H43K-02

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2008) Using a modified Gaussian model to predict concentrations of blue-green algal pigments in eutrophic Indiana reservoirs. American Society for Photogrammetry and Remote Sensing. Portland, OR. April (Oral Presentation). Abstract 6-1

Robertson, A.L. (2003) Characterization of metal ion effects on protein phosphatase 5. Purdue University School of Agriculture Undergraduate Poster Contest and Exhibition. West Lafayette, IN. April (Poster Presentation)

Robertson, A.L. (2002) Sperm allocation in a cockroach, *Nauphoeta cinerea.* Midwest Ecology and Evolution Conference. Bowling Green, OH. March (Poster Presentation)

Work in Review/Prep:

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2009) Using a modified Gaussian model to predict cyanobacterial pigment abundance in central Indiana reservoirs. *IEEE Journal of Selected Topics in Earth Observation and Remote Sensing.* in review

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (2009) Comparison between spectral indices, curve-fitting, and semi-empirical models in predicting algal pigments in Case II waters. *Remote Sensing of the Environment.* in review

Robertson, A.L., Li, L., Tedesco, L., Wilson, J. and Soyeux, E. (in prep) Using a Partial Least Squares (PLS) Method for Estimating Cyanobacterial Pigments in Eutrophic Inland Waters.

Teaching Experience:

High School Teacher:

MSD of Warren Township August 2005-August 2007

- Taught Biology, Chemistry, and Integrated Chemistry-Physics

Teaching Assistant:

Purdue University January 2003-May 2003

- Co-taught entry-level biology lab for non-majors (BIOL111)