1-1-2012

Remote Sensing of Cyanobacteria in Turbid Productive Waters

Sachidananda Mishra

Remote sensing of cyanobacteria in turbid productive waters

By

Sachidananda Mishra

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Earth and Atmospheric Science
in the Department of Geosciences

Mississippi State, Mississippi

August 2012

Remote sensing of cyanobacteria in turbid productive waters

By

Sachidananda Mishra

Approved:

Deepak R. Mishra Assistant Professor of Geosciences (Major Advisor)	John Rodgers Associate Professor of Geosciences (Committee Member)
Karen McNeal Assistant Professor of Geosciences (Committee Member)	Zhongping Lee Professor of Environmental, Earth and Ocean Sciences University of Massachusetts at Boston (Committee Member)
Mike Brown Associate Professor of Geosciences (Graduate Coordinator)	Gary L. Myers Professor and Dean of the College of Arts & Sciences

Name: Sachidananda Mishra

Date of Degree: August 11, 2012

Institution: Mississippi State University

Major Field: Earth and Atmospheric Science

Major Professor: Dr. Deepak R. Mishra

Title of Study: Remote sensing of cyanobacteria in turbid productive waters

Pages in Study: 186

Candidate for Degree of Doctor of Philosophy

Cyanobacterial algal bloom is a major water quality issue in inland lakes, reservoirs, and estuarine environments because of its scum and bad odor forming and toxin producing abilities. Health risks from cyanobacterial toxin can vary from skin irritations to fever, intestinal problems, and neurological disorders. Terminations of blooms also cause oxygen depletion leading to hypoxia and widespread fish kills. Adding to the problem, many species of cyanobacteria produce odorous compounds such as geosmin and 2-methylisoborneol (MIB) that cause “earthy-muddy” and “musty” odor in drinking water, which is also a serious issue in aquaculture and drinking water industry. Therefore continuous monitoring of cyanobacterial presence in recreational water bodies, surface drinking water sources, and water bodies dedicated for aquaculture is highly required for their early detection and subsequent issuance of a health warning and reducing the economic loss.

Remote sensing techniques offers the capability of identifying and monitoring cyanobacterial blooms in a synoptic scale. Over the years, the scientific community has focused on developing methods to quantify cyanobacterial biomass using phycocyanin,

an accessory photosynthetic pigment, as a marker pigment. However, because of the confounding influence of chlorophyll-*a* and other photo pigments, remote retrieval of phycocyanin signal from turbid productive water has been a difficult task. This dissertation analyzes the potential of remote sensing techniques and develops empirical and quasi-analytical algorithms to isolate the phycocyanin signal from the remote sensing reflectance data using a set of radiative transfer equations and retrieves phycocyanin concentration in the water bodies. An extensive dataset, consisting of *in situ* radiometric measurements, absorption measurements of phytoplankton, colored dissolved organic matter, detritus, and pigment concentration, was used to optimize the algorithms. Validations of all algorithms were also performed using an independent dataset and errors and uncertainties from the algorithms were discussed. Despite the simplicity, an empirical model produced highest accuracy of phycocyanin retrieval, whereas, the newly developed quasi-analytical phycocyanin algorithm performed better than the existing semi-analytical algorithm. Results show that remote sensing techniques can be used to quantify cyanobacterial phycocyanin abundance in turbid and hypereutrophic waters.

DEDICATION

To my parents

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor, Dr. Deepak Mishra, for his insightful guidance and technical advice throughout the tenure of this dissertation. Being a terrific mentor, he made my doctoral program experience very exciting. I also thank him for teaching me the art and science of research as well as scientific writing, and providing me the insight to the overall functioning of research and academia. Kind acknowledgements are also due to the members in my dissertation committee, Dr. John Rodgers, Dr. Karen McNeal, and Dr. Zhongping Lee, for their technical guidance and constructive criticism that significantly improved the quality of this dissertation. Special thanks are due to Dr. Lee for first of all agreeing to be in my dissertation committee, answering all my questions, and critically reviewing my research work. His contribution to this dissertation has been immense.

I express my thanks to Dr. Craig Tucker, director of the Warm Water Aquaculture Center, for allowing me to use his lab space and helping me out in the planning of field sampling at Delta Research Extension Center, Stoneville, Mississippi. Kind acknowledgements are also due to Dr. Matt Ross and Dr Sam from the College of Veterinary Medicine for teaching me and helping me out with HPLC analysis of phytoplanktonic pigments. I also thank Margaret E. Dennis from Warm Water Aquaculture Center for her help with algae identification and organism counting.

I would like to thank everyone in Dr. Mishra's lab, especially Shubho and Chris for their support and always keeping the lab environment light and pleasant with their witty humor. I also thank them for their help during field works and logistics. I am thankful to all my colleagues and friends at the Department of Geosciences, especially Mark Baldwin, Rob Thornton, Patricia Kambesis, Amelia Fox, and Louis Infante.

I would like to thank my beautiful wife, Lipi, for her unconditional love and support, and listening to all my frustrations. She has been my ardent critique, source of inspiration and motivation throughout the course of this dissertation. Finally, I thank my parents, parents-in-law, brothers and sisters for their continuous love, support, and encouragement.

I am thankful to Geological Society of America for financially supporting my research in form of a graduate student research grant, and European Space Agency for providing MERIS data for free of cost through a category-1 proposal. I will always be indebted to the Department of Geosciences, Mississippi State University for funding my research through a financial assistantship without which this dissertation would have never been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Background

Cyanobacterial harmful algal bloom is a major water quality problem in inland lakes, reservoirs, estuarine, and coastal waters. Many genera of cyanobacteria (*Anabaena spp*., *Microcystis aeruginosa, Cylindrospermopsis raciborskii*) produce several types of toxins such as: neurotoxins (anatoxin-a), hepatotoxins (microcistin), and cytotoxins (cylindrospermopsins) (Skulberg *et al*., 1992; Metcalf *et al.*, 2008). Their toxin and bad odor causing capabilities have drawn attention for early detection by frequent monitoring of water quality. Traditional laboratory analysis based monitoring of water quality is time consuming, expensive, and often a very difficult task. However, unique optical properties of phycocyanin (PC), a characteristic cyanobacterial photosynthetic pigment, in visible and near-Infrared wavelength range can be used to develop algorithms/techniques that can detect and quantify cyanobacterial biomass in natural water bodies. Cyanobacterial light harvesting pigments are composed of chlorophylls, carotinoids, and an accessory photosynthetic phycobilin pigment, PC. It is widely accepted that PC can be considered as the diagnostic photosysnthetic pigment of cyanobacteria and hence an indicator of the presence of cyanobacterial biomass. Fortunately, PC has very distinct absorption characteristics (A_{max} at 620 nm) that is prominent in reflectance spectra acquired from cyanobacteria dominated water bodies (Glazer 1989; Richardson 1996). When the

cyanobacterial biomass dominate the water body, the reflectance spectra shows an enhanced absorption feature between 600 to 625 nm and a reflectance maximum at around 650 nm (Dekker *et al.*, 1993; Kutser *et al.*, 2006; Mishra *et al.*, 2009) which can be used to fingerprint cyanobacteria in remotely sensed data.

Over the past years, researchers have been reasonably successful in exploiting PC absorption feature at 620 nm to develop empirical, semi-empirical, and semi-analytical models. Most research pertaining to the detection and mapping of cyanobacteria from *in situ* remote sensing data have used the absorption and reflectance features at 620 and 650 nm to develop relationships between R_{rs} and PC concentrations. From a comprehensive literature review it was found that most of the researchers have focused on four broad types of algorithms to quantify PC based on its absorption feature at 620 nm. The four algorithms include: 1) a single reflectance band ratio empirical algorithm (Schalles and Yacobi, 2000); 2) Semi-empirical baseline algorithm (Dekker, 1993); 3) Multiple Band Linear Regression Algorithm (Vincent *et al.*, 2004); and 4) a nested semi-analytical band ratio algorithm (Simis *et al.*, 2005). Each algorithm has been further explained and their pros and cons have been discussed below.

1.1.1 Band ratio empirical algorithms

Schalles and Yacobi (2000) developed a reflectance band ratio model using maximum reflectance between 640 and 660 nm and minimum reflectance between 615 and 635 nm to detect PC. The radiometric quantities that were used to develop the algorithm included ratios of radiance measured just beneath the water surface,$L_u(0^-, \lambda)$, and the radiance of a lambertian Spectralon panel, $E_d(0^+, \lambda)/\pi$ denoted as $R(\lambda)$.

$$PC = \left[\frac{R(650)}{R(625)} - 0.97\right] / 0.000912 \quad (1.1)$$

The main advantage of this algorithm is its simplicity to use. Being an empirical relationship between a simple band ratio and the measured PC concentration, by definition, this algorithm is not sensitive to types of radiometric measurements and radiometers used. Ruiz-Verdu *et al.* (2007) present an in-depth evaluation of this algorithm using two datasets acquired by two different radiometers including: 1) PR-650 (Photo Research Inc.), and 2) ASD-FR (Analytical Spectral devices, Inc.). They reported that the above mentioned band ratio algorithm was least sensitive to the radiometer types and showed minimal difference in the estimated PC concentration between the two datasets. However, this algorithm is prone to negative predictions because of its band architecture. By definition, this model will predict negative PC concentration when the reflectance ratio is less than 0.97 which is very common in turbid waters with low cyanobacteria concentration (Ruiz-Verdu *et al.*, 2008). In their validation study, Ruiz-Verdu *et al.* (2008) showed that the single band ratio algorithm predicted 36.4% of the data negative. Mishra *et al*. (2009) have shown that $R_{rs}(650)$ is very sensitive to optical properties of other photosynthetic pigments such as chlorophyll-*a*, (chl-*a*), and chl-*b* absorption, and allo-phycocyanin fluorescence. Any variations in these pigment concentration also affect the band ratio and the prediction accuracy. In addition, the required bands are only available on airborne and space-borne hyperspectral sensors. Therefore, it is not feasible to use this algorithm for regular PC monitoring activities in a large spatial scale using conventional multispectral satellite data.

1.1.2 Semi-empirical baseline algorithm

Semi-empirical baseline algorithm proposed by Dekker (1993) exploits the relative magnitude of the reflectance trough at 620 nm with respect to the center of a baseline between 600 nm and 648 nm. The algorithm was developed using data from shallow eutrophic lakes in Netherlands. Dekker (1993) proposed the original algorithm as:

$$PC = -24.6 + 13686 \times \{0.5 \times [R^{0-}(600) + R^{0-}(648)] - R^{0-}(624)\} \quad (1.2)$$

where, R^{0-} is the sub-surface irradiance reflectance measured just below the air-water interface. The main advantage of this algorithm is its minimal sensitivity to uncertainties in radiometric measurements. Band differencing associated with this algorithm offsets the measurement uncertainties (Ruiz-Verdu *et al.*, 2008). Similar to the band ratio empirical algorithm, the major drawback of this model is its sensitivity to confounding photo pigments other than PC because of the inclusion of the band at 648 nm (Mishra *et al.*, 2009). It predicts negative values when the difference between mid-point of the baseline and $R_{rs}^{0-}(648)$ falls below 0.0018 that is also common in turbid waters with low cyanobacterial abundance. This makes the algorithm vulnerable to negative predictions with higher relative error for PC concentrations lower than 10 mg m^{-3}. Also because of the required spectral resolution of 2 nm, the algorithm is only suitable for hyperspectral sensors and presently not feasible for regular monitoring activities.

1.1.3 Multiple band linear regression algorithm

Vincent *et al.* (2004) developed a spectral band ratio model to quantify PC concentrations in Lake Erie, OH, USA using Landsat 7 ETM+ data and were successful

in detecting and mapping PC in the western basin of Lake Erie. The proposed multiple linear regression model using multiple band ratios was as follows:

$$PC = 47.7 - 9.21 \times \left(\frac{B3}{B1}\right) + 29.7 \times \left(\frac{B4}{B1}\right) - 118 \times \left(\frac{B4}{B3}\right) - 6.81 \times \left(\frac{B5}{B3}\right) + 41.9 \times \left(\frac{B7}{B3}\right) \times -14.7 \left(\frac{B7}{B4}\right) \quad (1.3)$$

where, B*n* represents the dark-object subtracted digital numbers in the *n*th Landsat spectral band. It was the first model to be developed for Landsat data to monitor cyanobacterial blooms in inland lakes. The main advantage of this model is the high spatial resolution of Landsat data (30 × 30 m) as compared to other dedicated ocean color sensors. However, the poor temporal resolution was the major hurdle for regular monitoring of cyanobacterial harmful algal bloom in inland lakes. Based on the fact that Landsat bands do not have the capability to spectrally distinguish the cyanobacteria from other type of algae, this multiple band ratio algorithm with bands in Visible and Mid-Infrared wavelength range might simply be detecting any species of phytoplankton in the water (Kutser, 2006). Hence, this algorithm requires further validation to prove its sensitivity to phycocyanin concentration in cyanobacteria.

1.1.4 Nested band semi-analytical algorithm

Nested band semi-analytical algorithm (Simis *et al.*, 2005) estimates absorption coefficient of PC at 620 nm, $a_{PC}(620)$, and relates to PC concentration through the PC specific absorption coefficient, $a^*_{PC}(620)$. It uses reflectance measurements at 620, 665, 709, and 779 nm matching the Medium Range Imaging Spectroradiometer (MERIS) channels 6, 7, 9, and 12 respectively. Simis *et al.* (2005) algorithm follows three steps to quantify PC concentration. First, it estimates chl-*a* concentration using a semi-analytical

approach as in Gons *et al.* (1999); second, it derives the chl-*a* absorption at the phycocyanin absorption band (620 nm), subtracts the chl-*a* contribution at 620 nm, and thus theoretically isolates PC absorption at 620 nm; third, using specific absorption coefficient of phycocyanin at 620 nm, it estimates PC concentration. To achieve the above mentioned steps, back-scattering coefficient at 779 nm (b_b^{779}) is derived and then chl-*a* absorption coefficient at 665 nm and PC absorption at 620 nm are derived as given below:

$$a_{chl}(665) = \left(\left\{\left[\frac{R_{rs}(708)}{R_{rs}(665)}\right][a_w(708) + b_b]\right\} - a_w(665) - b_b\right)\gamma^{-1} \quad (1.4)$$

where, $a_{chl}(665)$ is the absorption by chl-*a* at 665 nm. b_b is the back-scattering coefficient derived, after assuming spectral homogeneity, as below:

$$b_b^{779} = \frac{1.61 \times (779)}{0.082 - (0.6 \times R_{rs}(779))} \quad (1.5)$$

$$a_{PC}(620) = \left(\left\{\left[\frac{R_{rs}(708)}{R_{rs}(620)}\right][a_w(708) + b_b]\right\} - a_w(620) - b_b\right)\delta^{-1} - \varepsilon a_{chl}(665) \quad (1.6)$$

Model parameters γ and δ empirically relates the model retrieved $a_\varphi(665)$ and $a_\varphi(620)$ to corresponding pad-measured values. Note that the algorithm retrieves non-water absorption, $a_{t\text{-}w}(665)$ and $a_{t\text{-}w}(620)$ in Eq. 1.4 and 1.6; but treats them as $a_{chl}(665)$ and $a_{chl}(620)$+$a_{PC}(620)$ by assuming $a_{CDM}(665)$ and $a_{CDM}(620)$ is negligible. Finally, PC concentration is derived as:

$$PC = \frac{a_{PC}(620)}{a^*_{PC}(620)} \quad (1.7)$$

Randolph *et al.* (2008) validated the semi-empirical algorithm developed by Simis *et al.* (2005) and reported that the performance of the algorithm was promising in

retrieving PC concentrations and estimating cyanobacterial abundance in Geist and Morse reservoirs, IN, USA. However, further validation and fine-tuning of Simis *et al.* (2005) algorithm using spectroscopic studies was suggested. Similarly, Hunter *et al.* (2010) evaluated the performance of Simis *et al.*, (2005) algorithm and documented its promising performance as compared to other existing PC algorithms. It was able to retrieve PC concentration within 10-200 mg m^{-3} with better accuracy. However, the model significantly over and underestimated the PC concentration below 10 mg m^{-3} and above 200 mg m^{-3} respectively.

In another study, Simis *et al.* (2007) documented the influence of phytoplankton pigments other than chl-*a* on the remote estimation of cyanobacterial biomass. They concluded that the presence of chl-*a*, *b* and *c*, and pheophytin tend to overestimate the PC concentrations in predictive models, and the estimation errors tend to be significant at low PC concentrations. They also concluded that even if a semi-empirical model to predict PC concentrations is corrected for chl-*a*, the absorption by the above mentioned pigments still influences the 620 nm absorption thereby affecting the estimation accuracy of PC retrieval.

1.2 Research question

Studies have shown that the nested semi-analytical algorithm (Simis *et al.,* 2005) outperforms all other existing PC algorithms. The very feature of removal of chl-*a* effect at the absorption band of PC distinguishes it from the rest of the algorithms. In case-2 type inland lakes and coastal waters, where optical properties are significantly influenced by CDOM, detritus, mineral particles that do not co-vary with the phytoplankton pigment concentration, water quality remote sensing is even more complex. In these types of

optically complex waters, the assumptions considered in Simis *et al.* (2005) algorithm such as: 1) the spectrally neutral back-scattering coefficient, and 2) influence of absorption by CDOM and detritus at 665 nm is negligible might fail. If the inversion algorithm overestimates absorption by chl-*a* concentration at 665 nm, the same error propagates to the estimation of $a_{PC}(620)$. Besides, the model parameters γ and δ play critical role in retrieving $a_{chl}(665)$ and $a_{PC}(620)$ from $R_{rs}(\lambda)$. Any uncertainty with these parameters will directly affect the algorithm's accuracy. Selecting the appropriate γ and δ is a big challenge as they tend to vary with geographic regions, or even in the same region but at a different time, and with optical condition of the waters.

Understanding the optical complexities of case-2 water types, I propose to use a quasi-analytical algorithm (QAA) (Lee *et al.,* 2002) to derive $a_{\varphi}(\lambda)$. QAA is an algebraic inversion algorithm that was designed to convert multiband R_{rs} data into inherent optical properties such as total absorption and scattering coefficients, $a_t(\lambda)$ and $b_b(\lambda)$, in the visible wavelength range. Further, based on the spectral properties, it decomposes $a_t(\lambda)$ into two absorption components including 1) $a_{\varphi}(\lambda)$,and 2) a combined absorption by detritus and colored dissolved organic matter (CDOM), $a_{CDM}(\lambda)$. In this research, $a_{\varphi}(620)$ was further decomposed to retrieve $a_{PC}(620)$ and use it for retrieval of PC.

As outlined above, chl-*a* also absorbs light energy at PC absorption maximum, ~620 nm. Therefore, accurate retrieval of PC concentration depends on the degree of accuracy of $a_{\varphi}(620)$ retrieval from $R_{rs}(\lambda)$ as well as the separation of chl-*a* effect from the PC signal. The questions this dissertation addresses are:

1) Will a better chl-*a* correction scheme increase the accuracy of PC retrieval?
2) Will quasi-analytical approach reduce the PC estimation error as compared to empirical and semi-analytical approach?

1.3 Objectives

This dissertation focuses on developing empirical, and quasi-analytical algorithms to retrieve PC concentration in cyanobacteria using remotely sensed data. The models were developed using radiative transfer theory and *in situ* data including hyperspectral $R_{rs}(\lambda)$ spectra, absorption coefficients of phytoplankton, CDOM, detritus, and pigment measurements. The specific objectives are:

1) Study the spectral properties of cyanobacteria and the interaction between chl-*a* on PC optical features.
2) Develop a reflectance band ratio empirical model to quantify cyanobacteria. Compare its performance with existing empirical and semi-analytical algorithms.
3) Parameterize the QAA algorithm using *in situ* bio-optical dataset collected from several inland aquaculture ponds with highly turbid and productive waters; validate the algorithm using an independent dataset collected from the same area in a different year.
4) Develop a QAA for PC using hyperspectral data acquired using a field spectroradiometer.
5) Develop an index using R_{rs} data measured at red and NIR wavelength region to quantify chl-*a* concentration in turbid and productive water.

1.4 Outline of the dissertation

This dissertation consists of 7 chapters. Chapter 1 provides the introduction to the nature of problem; outlines the state-of-the art methods and techniques in the literature; and lists the research objectives. Chapter 2 deals with a comprehensive analysis of spectral properties of cyanobacteria and its interaction with other photopigments. Chapter 3 deals with the development of a single reflectance band ratio empirical model to quantify cyanobacterial PC concentration and its validation using an independent dataset for accuracy assessment. Performance of the empirical model was also compared with a semi-analytical model. Addressing the limitations and challenges of empirical and semi-analytical algorithm, a quasi-analytical algorithm was also proposed in Chapter 4 and 5. Chapter 4 focuses on paramaterizing the quasi-analytical algorithm for bio-optical inversion in turbid and hypereutrophic waters encountered in our study region. Validation and accuracy assessment of the algorithm was also carried out and discussed in this chapter. In Chapter 5, quasi-analytical algorithm for PC was conceptualized, calibrated, and validated using *in situ* data suite. In Chapter 6, a normalized band difference index is presented to quantify chl-*a* concentration in turbid productive waters. Chapter 7 compares the performance of all PC models, outlines the summary of the dissertation research, and highlights the future direction.

1.5 References

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

Gons, H. (1999). Optical teledetection of chlorophyll *a* in turbid inland waters, *Environmental Science Technology*, 33: 1127–1132

Glazer, A. N. (1989). Light guides. Directional energy transfer in a photosynthetic antenna. *Journal of Biological Chemistry, 264*, 1-4.

Hunter, P.D., Tyler, A.N., Carvalho, L., Codd, G.A. & Maberly, S.C. (2010) Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes. *Remote Sensing of Environment,* 114 (11): 2705-2718.

Kutser, T., Metasamma, L., Strombeck, N., & Vahtmae, E. (2006). Monitoring cyanobacterial blooms by satellite remote sensing. *Estuarine Coastal and Shelf Science, 67*, 303-312.

Lee, Z.P., Carder, K.L., R.A., & Arnone, R. (2002). Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters, *Applied Optics,* 41: 5755-5772.

Metcalf, J. S., Banack, S. A., Lindsay, J., Morrison, L. F., Cox, P. A., and G.A. Codd. (2008). Cooccurrence of beta-N-methylamino-L-alanine, a neurotoxic amino acid with other cyanobacterial toxins in British waterbodies, 1990–2004. *Environ. Microbiol., 10,* 702−708.

Mishra, S., Mishra, D.R., & Schluchter. W.M. (2009). A Novel Algorithm for Predicting PC Concentrations in Cyanobacteria: A Proximal Hyperspectral Remote Sensing Approach. *Remote Sensing*, *1*: 758-775.

Randolph, K., Wilson, J., Tedesco, L., Li, L., Pascual, D., & Soyeux, M. (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.

Ruiz-Verdu, A., Simis, S. G. H., de Hoyos, C., Gons, H. J., & R. Pena-Martinez. (2008). An evaluation of algorithms for the remote sensing of cyanobacterial biomass. *Remote Sensing of Environment,* 112: 3996−4008.

Schalles, J., & Yacobi, Y. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid and Chl-a pigments in eutrophic waters. *Archive Hydrobiological Special Issues Advance Limnological, 55,* 153-168.

Skulberg, O.M., Carmichael, W.W., Anderson, R.A., Matsunaga, S., Moore, R.E. and Skulberg, R. (1992). Investigations of a neurotoxic Oscillatorialean strain (cyanophyceae) and its toxin. Isolation and characterization of homoanatoxin-a, *Environmental Toxicology and Chemistry,* 11: 321-329.

Simis, S., Peters, S., & Gons, H. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *American Society of Limnology and Oceanography, 50,* 237–245.

Simis, S., Ruiz-Verdú, A., Domínguez-Gómez, J., Peña-Martinez, R., Peters, S., & Gons, H. (2007). Influence of phytoplankton pigment composition on remote sensing cyanobacterial biomass. *Remote Sensing of Environment*, *106*, 414−427.

Vincent, R. K., Qin, X. M., McKay, R. M. L., Miner, J., Czajkowski, K., Savino, J., & Bridgeman, T. (2004). Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment*, *89*, 381-392.

CHAPTER II

SPECTRAL REFLECTANCE PROPERTIES OF CYANOBACTERIA

2.1 Overview

Understanding spectral reflectance properties of cyanobacteria is essential for remote sensing algorithm development. In this experimental study two species of cyanobacteria and one species of green algae were cultured in laboratory and reflectance spectra of algae were collected in controlled experiments. Results show that cyanobacterial reflectance peak located at ~650 nm is extremely sensitive to chlorophyll-*a* absorption. Magnitude of the peak decreased and location of the peak moved towards longer wavelength with increasing chlorophyll-*a* concentration. On the other hand, the chlorophyll-*a* absorption feature moved towards shorter wavelength with increasing chlorophyll-*a* concentration. Therefore, inclusion of 650 nm reflectance feature in empirical phycocyanin models will produce erroneous model retrievals.

2.2 Introduction

Remote sensing offers the capability of identifying and mapping biophysical parameters in water bodies by observing subtle variations in the water color in temporal and spatial scale. As remote sensing reflectance $R_{rs}(\lambda)$, the ratio of water leaving radiance to the downwelling solar irradiance, is primarily a function of absorption, $a(\lambda)$, and back-scattering, $b_b(\lambda)$ of incident light in the water column (Gordon *et al*., 1975), absorption of

a primary optically active component can be analytically obtained from $R_{rs}(\lambda)$. However, prior knowledge of absorption spectrum of the target constituent is essential to select the wavelength region where the constituent is having the highest sensitivity and dominates the total absorption budget. Therefore, the success of an algorithm for remote detection of cyanobacteria depends on our understanding of its optical signature. In this experimental study, characteristic spectral reflectance properties of cyanobacteria have been investigated by measuring reflectance (%) and pigment data using laboratory cultured cyanobacteria.

2.3 Methods

2.3.1 Laboratory culture of cyanobacteria and green algae

Two species of cyanobacteria, *Synechocystis sp*. (PCC 6803) and *Anabena* (also known as Nostoc) *sp*. PCC 7120 and one species of green algae (*Ankistrodesmus falcatus*) were grown in BG11 medium (Stanier *et al*., 1971) supplemented with 10 mM Hepes-KOH, pH 8.0. Small cultures of 100 mL were grown in 500 mL Erlenmeyer flasks with shaking under constant illumination of cool white light at 40-80 μmol photons $m^{-1}s^{-1}$ to the mid- to- late exponential phase of growth (OD_{730} = 0.8-1.5). Reflectance measurements were taken using these samples of cyanobacteria and green algae.

2.3.2 Reflectance measurements

Four experiments were conducted using a hyperspectral USB 4000 radiometer (Ocean Optics, Inc., Dunedin, Florida) for the spectroscopic analysis of the cyanobacterial species. The objective of these experiments was to study the spectral reflectance properties of cyanobacteria with changing pigment concentration.

Synechocystis was used in the first three experiments whereas *Anabaena* was studied in the fourth experiment. In all experiments, percent reflectance (ρ_λ) data of the two cyanobacterial species were acquired at varying concentrations and also in association with different concentrations of green algae. In the experiments, water samples that were scanned to acquire ρ_λ data were prepared as follows. First, 100 mL of an exponential phase culture of a particular cyanobacteria species was mixed with 500 mL of tap water for dilution and the dilution sequence was continued by adding 100, 200, and 500 mL of tap water subsequently.

Samples were placed in containers (20 L) painted black and scanned under controlled light from two 500 watt halogen lamps. Calibrated USB 4000 spectroradiometer with a 25^o field of view (FOV) optical fiber was used to acquire the above-surface upwelling radiance ($L_{\lambda,target}$) data of the water samples. The spectroadiometer was calibrated by measuring the upwelling radiance ($L_{\lambda,cal}$) of a Spectralon reflectance standard with 99% reflectance (Labsphere, Inc., North Sutton, New Hampshire). ρ_λ was computed using the calibration panel coefficient (cal_{coeff}) (available in the CALMIT Data Acquisition Program (CDAP;CALMIT, University of Nebraska-Lincoln). ρ_λ data were collected within a range of 400-900 nm with a spectral resolution of 1 nm. The equation used to compute ρ_λ is presented below.

$$\rho_{\lambda} = \left(\frac{L_{\lambda,target}}{L_{\lambda,cal}}\right) \times cal_{coeff} \tag{2.1}$$

Phycocyanin (PC) and chlorophyll-*a* (chl-*a*) concentrations were measured after each dilution step by using a chlorophyll sensor (part no. 6025) and cyanobacteria sensor measuring PC abundance (part no. 6131) attached to a YSI 6600 Multi-parameter water

quality sonde (YSI Inc., Yellow Springs, Ohio). Chl-*a* sensor measures in a range 0 to 500 µg L^{-1} with a resolution of 0.1 µg L^{-1} and linearity with an R^2 of 0.9999 relative to serial dilution of rhodamine with solution in the range from 0 to 500 µg L^{-1}. Cyanobacteria sensor measures in a range 0:280,000 cell mL^{-1}. The resolution of the cyanobacteria sensor is 220 cells mL^{-1}. Linearity of the cyanobacteria (PC) sensor has R^2 of 0.9999 for serial dilution of rhodamine with solution from 0 to 400 µg L^{-1} (YSI user's manual). Sensors were calibrated before use as per the instructions in the YSI user manual. In the first experiment (Exp I), measured ranges of Chl-*a* and PC were 0.7 to 7.8 µg/L and 7,050 to .247,960 cells mL^{-1}, respectively. The same procedure and dilution sequence was repeated in the second experiment (Exp II) on a different day. The measured Chl-*a* and PC ranges for Exp II were varied from 1.8 to 3.7 µg/L and from 506 to 126,570 cells mL^{-1} respectively.

In experiment III and IV (Exp III and Exp IV), measurements were started with low concentrations of pigment in the water samples. Cyanobacteria cells from a dense culture were added step-wise to increase concentrations subsequently. The minimum and maximum concentration of recorded Chl-*a* and PC in Exp III were 2.1 and 21.9 µg L^{-1}, and 4095 and 273,883 cells mL^{-1} respectively. In Exp IV, *Anabaena* was studied with cell densities ranging from 4550 to 244,500 cells mL^{-1}. The descriptive statistics of all experimental data are summarized in Table 2.1.

Two additional datasets were acquired during Exp III and IV in order to study the reflectance spectra of *Synechocystis and Anabaena* in association with varying concentration of green algae (*Ankistrodesmus falcatus*). The objective was to study the influence of high Chl-*a* concentration on reflectance spectra of PC and the dynamics of

various optical features in PC reflectance spectra. The PC concentration was kept constant at 180,000 cells mL^{-1} for *Synechocystis* and 220,000 cells mL^{-1} for *Anabaena*; whereas the green algae concentration was increased in sequence for both of the experiments. This was achieved by keeping the *Synechocystis* or *Anabaena* at the above described concentration level, while continuing to add *Ankistrodesmus* in order to increase the non-cyanobacterial Chl-*a* concentration.

2.4 Results and discussions

2.4.1 Analysis of reflectance spectra

Spectral characteristic features of the pigments Chl-*a* and PC are prominent in the reflectance spectra of *Synechocystis* (Exp I-III) and *Anabaena* (Exp IV) (Fig. 2.1). The green peak approximately at 550 nm is because of the scattering from algal cells in the water and also due to the relatively low absorption by Chl-*a* and carotenoids (Gitelson *et al*., 1999; Gitelson *et al*., 2000) (Fig. 2.1). The spectral troughs near 617 nm and 680 nm appear because of strong absorption by PC and Chl-*a*, respectively (Limasson *et al*., 1973; Gitelson *et al*., 2000; Schalles and Yacobi, 2000), whereas the peak at 654 nm appears because of the prominent absorption on both sides at 617 nm and 680 nm and also because of the phycocyanin fluorescence maximum at 650 nm (Rowan, 1989). Similarly, a peak near 700 nm appears because of two strong absorption features on either side, one by Chl-*a* at 680 nm and the other by water itself at 750 nm (Gitelson *et al*., 1999). The spectral characteristics of *Anabaena* (the position of absorption minima and reflectance maxima) appeared similar to those found for *Synechocystis* PCC 6803 (Fig. 2.1. D).

2.4.2 Usefulness of 654 nm peak in PC band ratio models

It was observed from the experimental data that the peak at 654 nm, which appears because of the absorption by PC and Chl-*a* on both side at 617 and 680 nm respectively, is very dynamic in nature. Its appearance and magnitude depends on PC, Chl-*a* and *b* concentrations. Metasamma *et al*. (2006) reported that R_{rs} peak at 650 nm only appears and can be detected by remote sensing instruments (with 10 nm spectral resolution and 1000:1 signal-to-noise ratio) when Chl-*a* concentration reaches 8-10 mg m^{-3}. However, in a few experiments it was observed that the 654 nm peak appeared on the reflectance spectra acquired by USB4000 (sampled at 10 nm spectral resolution with 250:1 signal-to-noise ratio) even when the Chl-*a* and PC concentration were 0.7 µg L^{-1} (or 0.7mg m^{-3}) and 7050 cells mL^{-1}, respectively (Fig. 2.2). On the other hand, in another experiment, the peak at 654 nm was not observed even when the Chl-*a* and PC concentrations were 7.4 µg L^{-1} and 21,050 cells mL^{-1} , respectively. That proved our intitial conclusion that the proportion of the concentration of PC to Chl-*a* controls the appearance of the 654 nm peak.

Similarly in Exp III, when the Chl-*a* concentration was 2.1-7.4 µg L^{-1} and PC concentration was 4095-21050 cells mL^{-1}, the high Chl-*a* concentration strongly absorbed light at 654 nm, thereby lowering the $(\rho_{617}^{-1}\rho_{654})$ ratio. Consequently, no peak appeared at 654 nm even if the PC concentration was sufficient enough to form the peak. Hence, two conclusions that can be drawn from the experimental data: (1) when Chl-*a* concentration is $\geq$ 2.1 µgL^{-1}, the 654 nm peak does not appear on the reflectance spectra of BG even with cell concentrations up to 21,050 cells mL^{-1}, whereas the PC present in 7050 cells mL^{-1} appears as a reflectance peak at 654 nm in the presence of 0.5µg L^{-1} of

Chl-*a*; (2) in the presence of high Chl-*a* concentration, absorption by Chl-*a* lowers ρ_{617} and form 654 nm peak.

Therefore, it is clear from the experiments that the appearance of the 654 nm peak depends on both Chl-*a* and PC concentration, and also on the 654 to 617 nm reflectance ratio. The specific absorption spectrum of PC shows that the absorption at 617 nm is three-fold higher than at 654 nm (Glazer, 1988), however, if Chl-*a* is also present in the water along with PC, the absorption at 654 nm would increase. Hence in the presence of high Chl-*a*, higher concentrations of PC are also required to be present in the water in order to form the prominent 654 nm peak. This suggests that Chl-*a* is the major contributing pigment that affects the magnitude of the 654 nm peak. The predictive ability of the spectral band ratio algorithms containing ρ_{654} therefore depends on the concentration of Chl-*a*, and because of this, the 654 nm peak cannot be accurately used to quantify PC efficiently in the case of variable PC:Chl-*a* ratios in water.

In the two additional experiments, the PC concentration was kept constant and the Chl-*a* concentration was increased in sequence by adding cultured *Ankistodesmus* to explore the dependence of the peak at 654 nm on Chl-*a* and Chl-*b.* The objective of the experiments was to study precisely the behavior of the 654 nm peak in the prescence of green algae. As the Chl-*a* concentration increased, the 654 nm peak shifted to 660 nm, and another reflectance peak appeared at 640 nm when the Chl-*a* concentration reached 122.8 µg L^{-1} (Fig. 2.3. A, B). Movement of the 654 nm peak to 660 nm might be due to absorption by Chl-*b* which is a major accessory photopigment in green algae. On the other hand, the Chl-*a* absorption feature blue-shifted from 680 nm to 670 nm at 310 µg L^{-1} of Chl-*a*. This blue-shift may be explained by the dominance of scattering by algal cells

and fluoresence by Chl-*a* at 680 nm over absorption at the same wavelength. Thus the $\rho_{678-682}$ started increasing and the absorption by Chl-*a* became prominent at 670 nm. This instability of the 654 nm peak with increasing Chl-*a* and *b* also points to the lack of precision and utility of this peak in empirical models to quantify PC.

2.5 Conclusions

Experimental data obtained in this study confirm that a distinct trough observed around 620 nm in the reflectance spectrum of cyanobacteria can be attributed to the assessory photopigment phycocyain and used as an optical marker of cyanobacteria. However, non-cyanobacterial chl-*a* from other chlorophyll bearing organisms can significantly affect the reflectance signal at 620 nm and cause the retrieval of phycocyanin erroneous. Therefore, contribution of chl-*a* should be accurately quantified and subtracted from the phycocyanin signal at 620 nm to reduce the chl-a sensitivity of phycocyanin algorithms. Similarly, reflectance peak around 654 nm blue-shifted with increase in chl-*a* concentration. Because of its high sensitivity to chl-*a*, reflectance peak at 654 nm should not be used in reflectance band ratio algorithm to quantify cyanobacterial biomass.

Table 2.1 Descriptive statistics of pigment measurements in the four proximal sensing experiments..

Exp.	Pigment	Mean	Std. Dev.	Min	Max	n
I	PC (cells mL^{-1})	85529.82	75586.96	7050.00	247960.00	11
	Chl-*a* ($\mu g\ L^{-1}$)	3.40	2.39	0.70	7.80	8
II	PC (cells mL^{-1})	50409.40	41727.17	506.00	126570.00	20
	Chl-*a* ($\mu g\ L^{-1}$)	2.48	0.78	1.80	3.70	5
III	PC (cells mL^{-1})	118360.00	100771.72	4095.00	273883.00	12
	Chl-*a* ($\mu g\ L^{-1}$)	13.32	7.12	2.10	21.90	12
IV	PC (cells mL^{-1})	94137.09	76424.23	4550.00	244500.00	11

N refers to the total number of readings acquired. Chl-a data was not acquired in Exp. IV

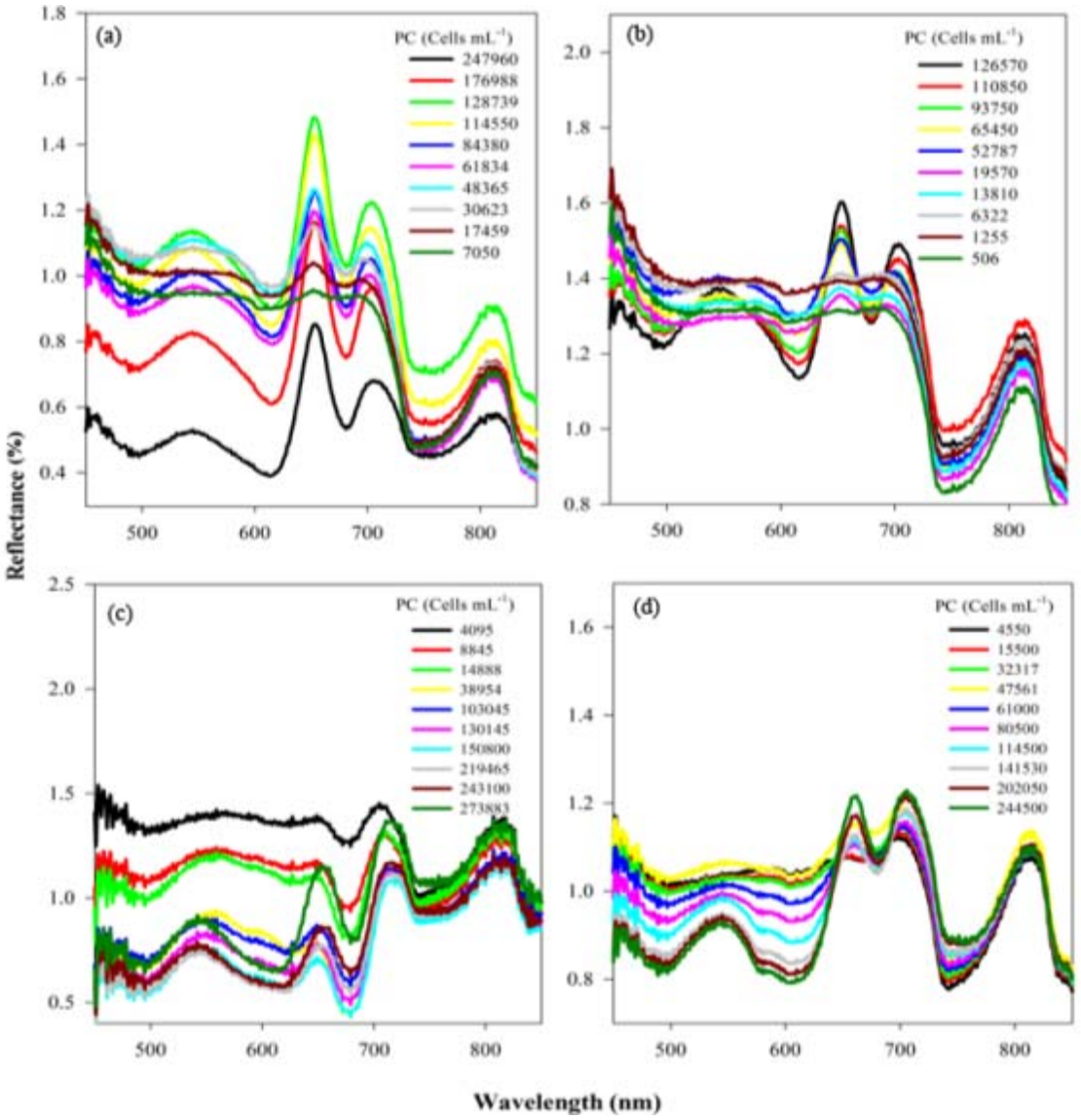

Figure 2.1 Percent reflectance spectra of *Synechocystis* PCC 6803 from Exp I, II, III respectively (a, b, c).

Percent reflectance spectra of *Anabaena* from Exp IV(d).

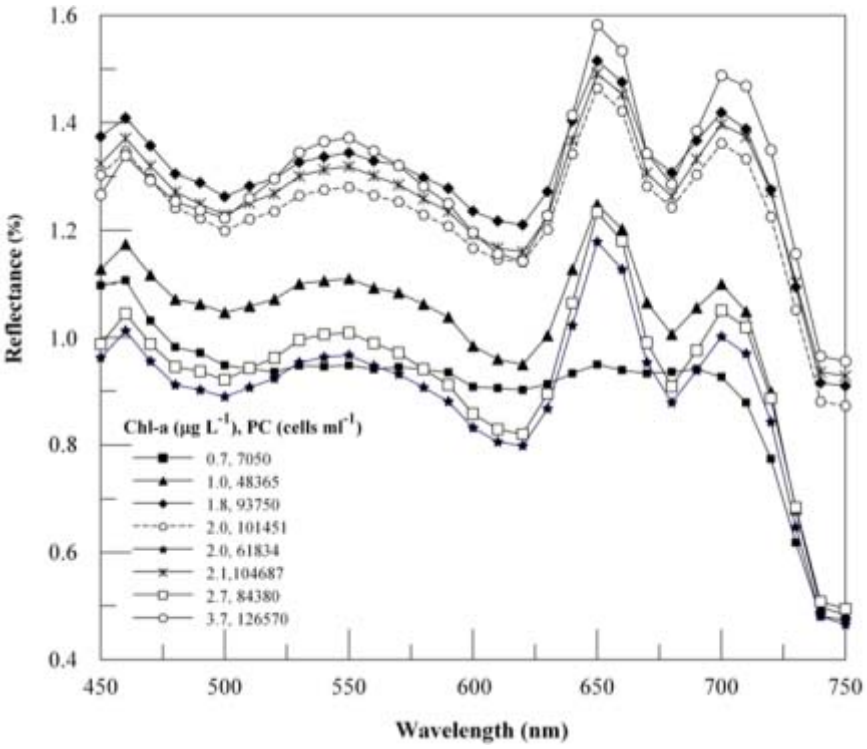

Figure 2.2 Percent reflectance spectra of *Synechocystis* and *Anabaena* showing appearance and dynamics of 650 nm peak at different Chl-*a* concentrations.

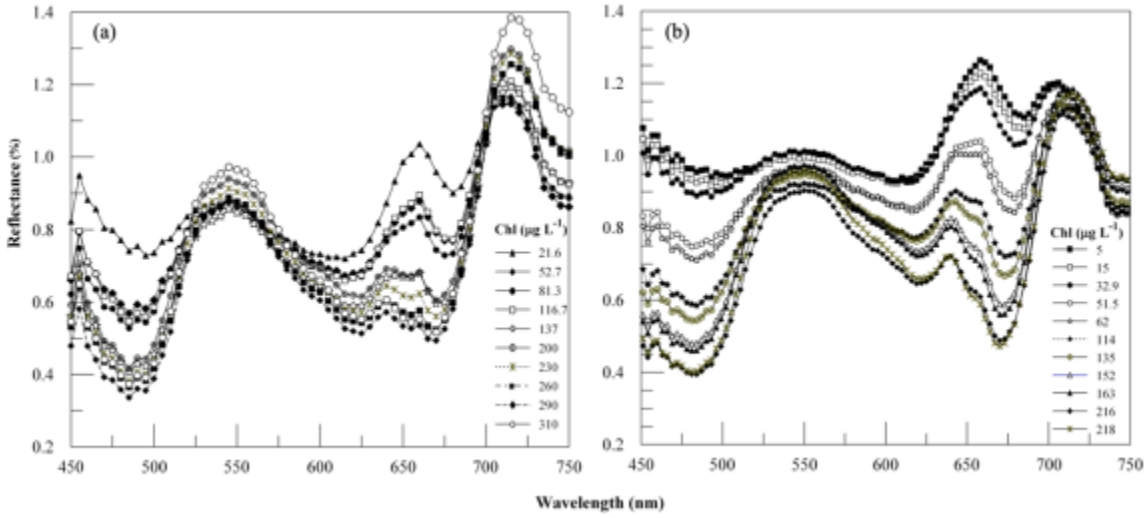

Figure 2.3 Effect of varying green algae concentrations on (A) *Synechocystis* and (B) *Anabaena* reflectance spectra.

2.6 References

Gitelson, A.A., Schalles, J.F., Rundquist, D.C., Schiebe, F.R., & Yacobi, Y.Z. (1999). Comparative reflectance properties of algal cultures with manipulated densities. *Journal of Applied Phycology.,* 11: 345-354.

Gitelson, A.A., Yacobi, Y.Z., Schalles, J.F., Rundquist, D.C., Han, L., Stark, R., & Etzion, D. (2000). Remote estimation of phytoplankton density in productive waters. *Archive Hydrobiol. Sp. Issues Advances in Limnology*, 55: 121-136.

Glazer, A. N. (1988). Phycobiliproteins. *Methods in Enzymology,* 167: 291-303.

Gordon, H.R., O.B. Brown, & M.M. Jacobs. (1975). Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean. *Applied Optics*, 14: 417-427.

Lemasson, C., Marsac, N.T.D., & Cohen-Bazire, G. (1973). Role of allophycocyanin as a light-harvesting pigment in cyanobacteria, *P. Natl. Acad. Sci. USA*., *70*: 3130-3133.

Metsamaa, L., Kutser, T. and Strömbeck, N. (2006). Recognizing cyanobacterial blooms based on their optical signature: a modeling study. *Boreal Environment Research*, 11: 493-506.

Rowan, K.S. *Photosynthetic Pigments of Algae*, 1st ed.; Cambridge University Press: Cambridge, UK, 1989.

Schalles, J.F., and Yacobi, Y.Z. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll pigments in eutrophic waters. *Archiv für Hydrobiologie Special Issues Advances in Limnology*, 55: 153-168.

Stanier, R.Y., Kunisawa, R., & Mandel, M. (1971). Cohen-Bazire, G. Purification and properties of unicellular blue-green algae (order Chroococcales) *Bacteriol. Rev.*, 35: 171-205.

CHAPTER III

EMPIRICAL MODELS FOR REMOTE ESTIMATION OF CYANOBACTERIAL PHYCOCYANIN

3.1 Overview

Cyanobacterial harmful algal bloom (CHAB) is a major water quality and public health issue in inland and coastal environments as it can cause threats to aquatic life by producing various toxins and depleting dissolved oxygen concentrations, alter the food-web dynamics and the overall ecosystem functioning (Mishra *et al*., 2009). From a commercial point of view, some species of cyanobacteria produce earthy/musty off-flavoring compounds in farm-raised channel catfish ponds that reduce the demand of the production causing huge economic loss for the aquaculture industry. Frequent monitoring of water quality in a synoptic scale has been possible by the virtue of remote sensing techniques. In this research, *in situ* data were collected from several highly turbid and hyper-eutrophic aquaculture ponds in the southeastern United States including hyperspectral remote sensing reflectance (R_{rs}), chl-*a* ,chl-*b*, and Phycocyanin (PC) pigment concentrations. An empirical algorithm (Mishra *et al.*, 2009), which was developed using laboratory cultured data to quantify cyanobacterial biomass, and a semi-analytical algorithm (Simis *et al.*, 2005) were validated by means of the measured field dataset. Error analysis showed that the semi-analytical algorithm produced the largest estimation error (mean RE=62% and median RE=56%) as compared to the empirical

models. Mean relative error of $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$ models was 25.4 and 21.6% respectively. The median relative error of $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$ was ~20% and 22%. Overall performance of both empirical models was similar to each other.

3.2 Introduction

Cyanobacteria are photosynthetic microorganisms commonly found in eutrophic inland lakes and reservoirs often dominating the phytoplankton community in summer months (Gibson and Smith, 1982). Cyanobacteria are the largest, most diverse group of prokaryotes that can multiply very rapidly in the summer when temperature, light, and nutrient runoff from fertilizers increase. Some of the genera of cyanobacteria produce various types of neurotoxins and hepatotoxins that adversely affect animals and humans (Carmichael, 1997; Iwasa *et al.*, 2002; Ballot *et al.*, 2003). Hence cyanobacteria detection and monitoring is critical in lake, coastal, and estuarine environments. Traditional sampling techniques for this process can be time-consuming and expensive, and real time spatial monitoring of cyanobacteria in large lakes, estuaries and coastal waters has been extremely difficult (Backer, 2002). Cyanobacteria have certain photosensitive pigments with distinct optical characteristics, allowing their detection and mapping by airborne and space-borne optical sensors.

Cyanobacteria have photosynthetic reaction centers that are structurally and functionally similar to those found in eukaryotic chloroplasts, but their light-harvesting pigments are composed of Chlorophyll-*a* (chl-*a*) and the phycobiliproteins (PBP) (Golbeck and Bryant, 1991; Richardson, 1996). Cyanobacteria minimally contain at least three different spectrally detectable PBPs including phycocyanin (PC) (A_{max}=620 nm),

allophycocyanin (AP) (A_{max}=650 nm), and allophycocyanin B (APB) (A_{max}=670 nm) (Glazer, 1989; Richardson, 1996). Some cyanobacteria may also contain phycoerythrin (PE) (A_{max}=565 nm). These PBPs form a large protein complex which transfers energy primarily to Photosystem II reaction centers. Photosystem I reaction centers bind chl-*a*, the majority of which serve as antenna pigments for light absorption (Golbeck and Bryant, 1991).

Hyperspectral remote sensing reflectance (R_{rs}) spectra for natural water bodies are generally complex because of the presence of many optically active constituents such as chl-*a*, carotenoids, total suspended solids (TSS), and Colored Dissolved Organic Matter (CDOM). However, researchers have been reasonably successful in exploiting PC 620 nm absorption feature to develop empirical and semi-empirical models to detect PC as a marker for cyanobacteria from water bodies (Simis *et al.*, 2005). Most research pertaining to the detection and mapping of cyanobacteria from *in situ* remote sensing spectra have used the absorption and reflectance features from 620 and 650 nm to develop a relationship between R_{rs} and PC concentrations. To date, three algorithms have been proposed to quantify PC based on its absorption feature at 620 nm: a single band ratio algorithm (Schalles and Yacobi, 2000), a semi-empirical algorithm (Dekker, 1993) and a nested semi-empirical band ratio algorithm (Simis *et al.*, 2005).

Dekker (1993) used an empirical baseline algorithm to quantify PC concentration from remote sensing reflectance while Schalles and Yacobi (2000) developed a reflectance band ratio model using maximum reflectance between 640 and 660 nm to minimum reflectance between 615 and 635 nm to detect PC. The major drawbacks of these empirical models are that they do not address the effect of chl-*a* concentrations on

the estimation accuracy of the PC models (Simis *et al.*, 2007). Although the reflectance peak between 640 and 660 nm has generally been used to develop empirical relationships to quantify PC, these empirical models for PC detection need narrower spectral resolution than is provided by commonly used ocean color satellite sensors (Kutser *et al.*, 2006). In a separate study, Vincent *et al.* (2004) developed a spectral band ratio model to quantify PC concentrations using Landsat 7 ETM+ data and were successful in detecting and mapping PC in the western basin of Lake Erie. They used all bands of Landsat 7 sensor except the band 6 (thermal infrared) in the model including the near- and middle infrared bands (i.e., band 4:760-900 nm; band 5:1550-1750 nm; and band 7: 2080-2350 nm). However, radiance (L_w) in these infrared bands is greatly reduced due to water absorption and modeling reflectance spectra beyond 750 nm could be accounting for turbidity in the lake caused by algal biomass instead of phycocyanin (Kutser *et al.*, 2006). Simis *et al.* (2005) developed a semi-empirical algorithm that used the band ratio from 709 nm to 620 nm for PC estimation. They discussed the influence of chl-*a* absorption at 620 nm and also included the impact of the variable PC: chl-*a* ratio on the performance of the algorithm. They concluded that the model error significantly increased as the PC:chl-*a* ratio decreased, providing evidence for the effect of chl-*a* absorption at 620 nm.

Simis *et al.* (2007) documented the influence of phytoplankton pigments other than chl-*a* on the remote estimation of cyanobacterial biomass. They concluded that the presence of chl-*a*, *b* and *c*, and pheophytin tend to overestimate the PC concentrations in predictive models, and that estimation errors tend to be significant at low PC concentrations. Therefore, even if a semi-empirical model to predict PC concentrations is corrected for chl-*a*, the absorption by the above mentioned pigments still influences the

620 nm band, affecting the accuracy of PC estimation. Randolph *et al.* (2008) validated the semi-analytical algorithm developed by Simis *et al.* (2005) and reported that the performance of the algorithm was promising in retrieving PC concentrations and estimating cyanobacteria abundance in Geist and Morse reservoirs, Indiana. However, further validation and fine-tuning of Simis *et al.* (2005) algorithm using spectroscopic studies was suggested.

In this research, PC concentrations in cyanobacteria have been quantified using empirical (Mishra *et al.*, 2009) and semi-analytical (Simis *et al.*, 2005) approaches and performances of the algorithms have been compared using a field radiometric and pigment dataset collected from highly turbid and productive waters dominated by cyanobacteria. The specific objectives of this research include: (1) calibrating the algorithms using an *in situ* dataset and (2) validating the algorithms for assessing estimation accuracy. *In situ* data were collected from Delta Research Extension Center aquaculture ponds, a Mississippi State University research facility located near Stoneville, MS, USA during 13-16 July, 2010 and 28-29 April, 2011. Average depth of aquaculture ponds is about 1.0 m with surface area ranging from 0.4-1.62 ha. Ponds do not have natural watersheds and solely depend on rain water and irrigation as their water source. The aquaculture ponds were selected for this study because of their eutrophic to hypereutrophic status (chl-*a* concentration exceeding 1000 mg m^{-3} in summer months) with cyanobacteria dominating the phytoplankton community during summer and occasionally during spring months (Tucker and Boyd, 1985).

3.3 Data and methods

3.3.1 Pigment measurements

3.3.1.1 Chlorophyll-*a*

Water samples for chl-*a* and chl-*b* analysis were simultaneously collected in 1L Niskin bottles and immediately filtered onto GF/F filters (Whatman, 0.7 μm pore size) under low vacuum (<5 inch of Mercury). Samples were extracted in triplicates using 90% acetone and concentrations were measured using HPLC following Environmental Protection Agency method 447 (Arar, 1997).

3.3.1.2 Phycocyanin

Water samples for PC analysis were filtered immediately after collection through a 0.2 μm nucleopore membrane filters (Milipore) under low vacuum. Filters were placed into a 15 mL falcon tube then frozen at -80^{o}C until analysis. Prior to analysis, filters were transferred to a 50 mL polycarbonate centrifuge tube, allowed to reach ambient room temperature, and then suspended in 5 mL of 50 mM phosphate buffer. Samples were homogenized following Sarada *et al.* (2009) using a sonicator. The tip of the sonicator was rinsed twice with 5 mL of 50 mM phosphate buffer each time and collected in the centrifuge tube. Samples were centrifuged at 5^{o} C, 27200 g for 25 minutes. Samples were again homogenized and the tip of the sonicator was rinsed with 5 mL of buffer and collected in the centrifuge tube and again centrifuged in the same settings. Finally, supernatant was collected and absorbance was measured using a Perkin Elmer lambda 850 spectrophotometer. Concentration of PC was calculated using the equation from Bennett and Bogorad (1973).

3.3.2 Spectral measurements

A dual sensor-system with two inter-calibrated Ocean Optics spectroradiometers was used to collect remote sensing reflectance data in the range 400-900 nm with a sampling interval of 0.3 nm following DallO'lmo *et al.* (2005). Radiometer 1, equipped with a 25° field-of-view optical fiber measured the sub-surface upwelling radiance just below the air-water interface, expressed in digital numbers as $DN_{Lu}(\lambda)$; whereas radiometer 2, equipped with an optical fiber and cosine diffuser (yielding a hemispherical field of view) acquired above surface down welling irradiance, expressed in digital numbers as $DN_{Ed}(\lambda)$. To match their transfer functions, inter-calibration of the radiometers was accomplished by measuring the upwelling radiance of a white Spectralon reflectance standard (Labsphere, Inc., North Sutton, NH) simultaneously with incident irradiance. The two radiometers were inter-calibrated immediately before and after measurements in each field site. After the data acquisition, R_{rs} was calculated as follows:

$$R_{rs}(\lambda) = \frac{t}{n^2}\frac{DN_{Lu}(\lambda)}{DN_{Ed}(\lambda)}\frac{DN_{Ed,ref}(\lambda)}{DN_{Lu,ref}(\lambda)}\frac{\rho_{ref}(\lambda)}{\pi}F_i(\lambda) \qquad (3.1)$$

Where, t is the transmittance at the air-water interface (0.98); n is the refractive index of water (1.34); $DN_{Lu,ref}$ and $DN_{Ed,ref}$ are digital numbers representing upwelling radiance and downwelling irradiance over the white Spectralon panel; ρ_{ref} is the irradiance reflectance of the Spectralon panel; $F_i(\lambda)$ is the spectral immersion factor (Ohde and Siegel, 2003). For each station 6 consecutive scans were recorded and further averaged to calculate a representative $R_{rs}(\lambda)$ spectrum (Fig. 3.1).

3.3.3 Phycocyanin algorithm

3.3.3.1 Empirical PC algorithm

Mishra *et al.* (2009) developed a reflectance band ratio algorithm to quantify phycocyanin in cyanobacteria using a dataset collected from laboratory experiments on cultured cyanobacteria. The desired band ratio required two wavelengths those that are maximally and minimally sensitive to PC concentration. Absorption spectra of chl-*a,* chl-*b*, PC, PE, and xanthophylls show that pigment absorption significantly decrease after 700 nm (Raven *et al.*, 1976). Therefore, 700 nm was selected as one spectral band with minimal sensitivity to PC and other pigments. Similarly, another band at 600 nm was selected representing PC absorption. Although 617 nm is the location of PC absorption maximum, 600 nm was chosen to avoid the influence of chl-*a* absorption at 617 nm. Emerson and Lewis's (1942) findings on the photosynthetic efficiency of PC in *Chroococcus*, a cyanobacterium, shows that the total absorption of light by pigments at 600 nm is because of 89% absorption by PC and 11% absorption by chlorophylls, whereas at 617 nm, the chlorophyll absorption increases to approximately 18%. Therefore 600 nm was selected as the second band with maximum sensitivity to change in PC concentration. Mishra *et al.* (2009) used $R_{rs}(700)/R_{rs}(600)$ as the reflectance band ratio on the assumption that it is most sensitive to a change in PC concentration and the least sensitive to change in any other pigment concentration. However, in hypereutropic waters, 700 nm could be contaminated by chl-*a* fluorescence, and pigment absorption at this band center could still be significant (Mishra *et al*., 2012). Therefore, in this study 708 nm was selected instead of 700 nm for the band ratio. $R_{rs}(620)$ was also used because of the availability of this band in Medium Resolution Imaging Spectrometer

(MERIS) sensor, to form another band ratio $R_{rs}(708)/ R_{rs}(620)$ to retrieve PC in cyanobacteria dominated waters.

3.3.3.2 Semi-analytical algorithm

Simis *et al.* (2005) developed a semi-analytical algorithm to retrieve cyanobacterial PC concentration in highly turbid and productive waters. Salient features of the algorithm are as follows. Remote sensing reflectance just below the air-water interface, $R_{rs}(0^-,\lambda)$, can be related to the inherent optical properties of the water such as absorption coefficient (a), back scattering coefficient (b_b) through a factor f that is dependent on the light field geometry (Morel and Gentili,1993). Based on Gordon *et al.* (1975) widely accepted relation, inherent optical properties and apparent optical properties can be related by the following relation.

$$R_{rs}(0-,\lambda) = f\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)} \quad (3.2)$$

Where, f is a scaling factor dependent on the light field geometry and volume scattering function (Morel and Gentili, 1993). $a(\lambda)$and $b_b(\lambda)$ are total absorption and total back scattering coefficients. Considering R_{rs} at another wavelength, where, $a(\lambda)$ and $b_b(\lambda)$ is known or at least they can be approximated, and $a(\lambda)$ at one wavelength can be solved analytically. Therefore, absorption by chl-*a* and PC can be analytically estimated from two R_{rs} bands and by solving two equations generated from Eq. 3.2. In this case, three assumptions have been made including 1) $a(620)$ is the sum of absorption by chl-*a*, PC, and water, 2) $a(665)$ is primarily because of chl-*a* and water absorption, and 3) $a(708)$ is only because of water absorption. In addition, this algorithm assumes that $b_b(\lambda)$ is spectrally neutral and it can be estimated from R_{rs} at a NIR wavelength, $R_{rs}(779)$.

Based on the assumptions, $R_{rs}(665)$ and $R_{rs}(708)$ can be used to retrieve chl-*a* absorption. Similarly $R_{rs}(620)$ and $R_{rs}(708)$ can be used to retrieve a sum of chl-*a* and PC absorption. Further, absorption by PC at 620 nm can be estimated by subtracting chl-*a* absorption at 620 nm from the combined absorption by chl-*a* and PC at 620 nm. Analytical equation to retrieve chl-*a* and PC absorption are provided below.

$$a_{chl}(665) = \left(\left\{\left[\frac{R_{rs}(708)}{R_{rs}(665)}\right][a_w(708) + b_b]\right\} - a_w(665) - b_b\right)\gamma^{-1} \quad (3.3)$$

Similarly absorption by PC at 620 nm can be written as:

$$a_{PC}(620) = \left(\left\{\left[\frac{R_{rs}(708)}{R_{rs}(620)}\right][a_w(708) + b_b]\right\} - a_w(620) - b_b\right)\delta^{-1} - \varepsilon a_{chl}(665) \quad (3.4)$$

Where, the model coefficients γ and δ empirically relates the model retrieved phytoplankton absorption at 665nm and 620 nm to the pad-measured ones; whereas, ε relates the chl-*a* absorption at 620 nm and 665 nm. The backscattering coefficient, $b_b(\lambda)$ can be estimated as (Gordon *et al.*, 1988; Gons,1999):

$$b_b(\lambda) = \frac{a_w(779)\alpha R_{rs}(779)}{0.82 - \alpha R_{rs}(779)}, \quad (3.5)$$

where, α accounts for refraction at the water surface (=0.68), and the factor 0.82 accounts for the average cosine of downward irradiance (Gordon *et al.*, 1988).

3.4 Results and discussions

3.4.1 Water quality

Summary of the measured pigment concentrations is presented in Table 3.1. Pigment concentration showed a wide variation in the entire dataset. PC concentration varied from 68.13 to 3032.47 mg m^{-3} with an average of 418.76 mg m^{-3}. Chl-*a*

concentration varied within 59.4-1376.6 mg m^{-3} with an average of 302.06 mg m^{-3}. Out of 25 samples, Chl-*b* concentration was below the detection limit of the instrument in 7 samples in the 2010 dataset. In the remaining samples, chl-*b* values varied within 1.56-13.71 mg m^{-3}. The average value of chl-*b* (4.35 mg m^{-3}) was lower than the standard deviation of chl-*a* replicates; which implies that chl-*b* values are within the uncertainty range of chl-*a* and therefore effect of chl-*b* on the remote estimation of PC should be negligible or minimal. Timing of our field sampling coincided with the peak of the summer bloom period in 2010 and the early bloom period in 2011. Therefore, the average chl-*a* concentration in 2010 and 2011 was ~833 and 185 mg m^{-3}. Similarly, the average PC concentration in 2010 and 2011 was ~570 and 142 mg m^{-3}. PC: chl-*a* ratio varied within 0.3-3.29 (mean=1.23) indicating cyanobacterial dominance in the phytoplankton community structure (Table 3.1). These values are in the similar range of previously reported values from the Morse and the Geist reservoir, IN, USA (Randolph *et al.*, 2008) and lakes and reservoirs from Spain and Netherlands (Simis *et al.*, 2005; Ruiz-Verdu *et al.*, 2008). Strong dependence was found between chl-*a* and PC concentration in the entire dataset (r=0.91). Similarly, a strong positive correlation was found between chl-*a* and chl-*b* (r=0.79).

3.4.2 Empirical PC algorithm

Empirical PC algorithm was calibrated using least-square regression between the measured PC concentration and the reflectance band ratios, $R_{rs}(708)/R_{rs}(600)$ and $(708)/R_{rs}(620)$. The dataset was randomly divided into two subsets, one for model calibration (n=16) and another for model validation (n=9). Regression analysis was performed on log-transformed PC concentration and reflectance band ratios. As shown in

Fig. 3.2a, $R_{rs}(708)/R_{rs}(600)$ ratio showed a linear relationship with PC concentration producing an R^2 of 0.77 and standard error of the estimate (STE) of 0.23. Similarly, $R_{rs}(708)/R_{rs}(620)$ also showed a linear relationship with measured PC concentration in the same dataset producing a STE of 0.204 and R^2 of 0.82. Further, empirically developed PC models were validated for accuracy assessment by applying them on the independent validation dataset. PC concentrations in the validation dataset were retrieved using the empirical equations and compared with measured PC values (Fig. 3.3a). Model errors were expressed as relative errors (RE):

$$Relative\ Error(\%) = \frac{(PC_{ref} - PC_{model})}{PC_{ref}} \times 100, \quad (3.6)$$

where, PC_{ref} and PC_{model} are measured and model retrieved PC concentration. For samples with high PC concentration (PC>200 mg m^{-3}), PC retrievals by $R_{rs}(708)/R_{rs}(600)$ model were close to the measured values. The maximum and minimum relative error of PC retrieval was -8% (over prediction) and -43% (over prediction); whereas, the average and median relative error were -8.73% and -16.6% respectively. Similarly, average and median relative error of the $R_{rs}(708)/R_{rs}(620)$ model were -3.31% and -16.63%. In both cases, median relative errors (~16%) show that performance of both models was very similar (Table 3.2, Fig 3a). Regression between measured PC and model retrieved PC produced high R^2 for both models (=0.99). Linear regression coefficient, slope of the best-fit line, was found to be 1.17 and 1.22 and intercepts were -14.7 and -45.2 for $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$ respectively. Ruiz-Verdu *et al.* (2008) have validated the three basic types of existing PC algorithm using an extensive dataset and reported that single reflectance ratio (Schalles and Yacobi, 2000) and the empirical

baseline algorithm (Dekker, 1993) produced 36.4 and 20.9% of negative predictions of PC. In contrast, because of the log-transformation of dependent and independent variables before regression, models don't produce any negative predictions in this study.

Analysis of relative error shows that maximum errors occurred with the samples with PC concentration less than ~200mg m^{-3}. Both of the models produced similar relative error for all samples excluding two. Relative error was less than 25% for samples with PC concentration within the range of 203-1881 mg m^{-3}(Fig. 3.3b). Model errors did not show any trend with PC: chl-*a* ratio. Unlike findings from semi-analytical PC model (Simis *et al.*, 2005) relative errors did not increase with a decrease in PC: chl-*a* ratio (Fig. 3.3c).

3.4.3 Semi-analytical PC algorithm

Semi-analytical algorithm was used to retrieve $a_{PC}(620)$ from R_{rs} data and PC concentrations were estimated by using $a^*_{PC}(620)$. Simis *et al.* (2005) optimized the model parameters (γ=0.68, δ=0.84, and ε=0.24) using a dataset collected from lakes in Netherland. In this study, same values of model parameters γ and δ produced severe underestimation of $a_\varphi(620)$ and $a_\varphi(665)$. The values reported by Simis *et al.* (2005) did not hold true in our dataset because of optically different waters in our study area. Therefore, values of γ and δ for our dataset were estimated by least-square regression technique (Fig. 3.4). The new values of γ and δ were found to be 0.5 and 0.67 and used in the model. The mean value of $a^*_{PC}(620)$ for various lakes in Spain and Netherland (=0.007 m^2 mg^{-1} of PC) and ε reported by Simis *et al.* (2006) were used to retrieve PC. Further, model retrieved PC concentrations were compared with measured PC values for accuracy assessment. The mean relative error of model estimation was 73%, whereas, the

median error was ~46%. Simis *et al.* (2005) documented that their PC model error significantly increased with a decrease in PC:chl-*a* ratio which is consistent with our observation. Maximum error occurred for the samples with PC concentration < 200 mg m^{-3} and a PC:chl-*a* ratio of <1 (Fig. 3. 5).

To understand the sources of error, model derived $a_{PC}(620)$ values were comparing with spectrophotometrically measured $a_{PC}(620)$ values. As shown in Fig. 3.6, model retrieved $a_{PC}(620)$ values of all samples, excluding one, were overestimated (Fig. 3.6). Least-square regression analysis showed a strong linear relation between the modeled and measured values (R^2=0.84, p<0.0001) but produced an intercept value of 0.99 (significantly more than the ideal, 0). This could be interpreted as systematic overestimation of retrieved values either because of: 1) absorption by accessory pigments other than chl-*a* and colored dissolved and detrital matter at 620 nm, or 2) use of a smaller value for the parameter ε while correcting for chl-*a* contribution. Simis *et al.* (2005) reported that the presence of chl-*b* could overestimate $a_{PC}(620)$ absorption and eventually contribute to the model error. In this dataset, the pigment concentration is predominantly dominated by chl-*a* and PC. Out of 25 samples, chl-*b* concentration was below the detection limit in 7 samples. In addition, average chl-*b* concentration in this dataset is just 1.47% of the average chl-*a* concentration and ~0.6% of the average chl-*a* and PC concentration combined. Therefore, effect of accessory pigment absorption on the overestimation of $a_{PC}(620)$ may be considered negligible at least in this dataset. The next possible reason behind the overestimation of $a_{PC}(620)$ could be from the model parameter ε. In this study, ε was used as a constant with a value of 0.24 as reported in Simis *et al.* (2005). However, it is known that ε could vary among different species of phytoplankton

functional types, degree of diversity in phytoplankton community, as well as pigment packaging. The value of ε increases when the dominant phytoplankton community changes from cyanobacteria to green algae. Also in case of pigment packaging, $a^*_{chl}(665)$ decreases with the degree of packaging and produces a ε value>0.24 (Simis *et al.*, 2005). In this study, the values of ε are not known and hence it is not possible to investigate its effect further. However, it can be generalized that the overestimation could have come from using a smaller value for ε that causes inefficient correction of chl-*a* absorption at 620 nm.

Another important source of model estimation error could be from the use of inaccurate $a^*_{PC}(620)$. In this study, an average value of $a^*_{PC}(620)$ (=0.007 $m^2\,mg^{-1}$ of PC) reported from the lakes in Spain and Netherland (Ruiz-Verdu *et al.*, 2008) was used to retrieve PC concentration. However, $a^*_{PC}(620)$ considerably varies within a wide range because of change in environmental and light conditions. $a^*_{chl-a}(665)$ significantly varies because of cell morphology and photo adaptation (Sathyendranath *et al.*, 1987; Bricaud *et al.*, 1995). Similar reasons may cause variability in $a^*_{PC}(620)$ as well. Therefore, use of a fixed $a^*_{PC}(620)$ value will cause considerable inaccurate PC retrievals in the final step.

It is also believed that that the biggest challenge of the semi-analytical algorithm is to find out the appropriate model parameters when applied in a new geographic region. Model parameters γ and δ play a critical role to retrieve $a_\varphi(665)$ and $a_\varphi(620)$ from $R_{rs}(\lambda)$. Values of γ and δ optimized in a particular water type cannot be used in other waters with different optical conditions. For example, optimized values of γ and δ in lakes from Netherland decreased 26.4 and 20.2% in our study region because of different optical

conditions. Values of γ and δ could also be affected by uncertainties caused from the model assumptions/simplifications such as: 1) pigment absorption at 708 nm is negligible, and 2) $b_b(\lambda)$ spectrum is spectrally neutral. It should be noted that in hyper-eutrophic waters such as our study region, pigment absorption at 708 nm is still considerable and in many cases total absorption at 708 nm is dominated by pigment absorption. Average pigment absorption at 708 nm was about 48% of the average total absorption at that wavelength (Mishra *et al.*, 2012). Therefore, the first assumption does not hold true in this study and produce an erroneous $a_\varphi(620)$ and $a_\varphi(665)$. Similarly, assumption of spectral neutrality of $b_b(\lambda)$ spectrum might fail in waters with high mineral content. If this simplification is not considered, Eqs. 3.3 and 3.4 will have a multiplication factor, $b_b(665)/b_b(709)$ and $b_b(620)/b_b(709)$, and spectrally neutral b_b will be replaced by $b_b(665)$ and $b_b(620)$. Although this cannot be investigated further due to lack of $b_b(\lambda)$ data to investigate this further, a simple analysis shows that for a given $b_b(779)$, if a standard $b_b(\lambda)$ model is considered (Gorden and Morel, 1983), $b_b(665)/b_b(709)$ and $b_b(620)/ b_b(709)$ may vary from 1 to 1.13 and 1 to 1.3 respectively for a spectral power (η) varying between 0 to 2. Similarly, if the spectrally neutral b_b value is used (as in Eqs. 3.3 and 3.4) in place of $b_b(665)$ and $b_b(620)$, it will underestimate the $b_b(665)$ and $b_b(620)$ within a range of 0 to 27% and 0 to 36% for a spectral power (η) varying between 0 to 2. As the parameters γ and δ address all these uncertainties such as proportion of pigment absorption in the total absorption budget and errors from assumptions, values of these model parameters cannot be approximated as constants and should only be used with caution.

3.5 Conclusions

In this research both empirical and semi-analytical approach was used to quantify phycocyanin concentration in cyanobacteria. Two reflectance band ratios such as $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$ were used to retrieve PC concentration using least-square regression. Model parameters γ and δ in the semi-analytical algorithm (Simis *et al.*, 2005) were optimized for our study region and used in the present study. Performances of all three models were analyzed by comparing estimation errors of the common samples (samples used in the empirical model validation). Results showed that the semi-analytical algorithm produced the highest estimation error among all (mean RE=62% and median RE=56%) (Fig.3.7). It should be noted that γ and δ should not be used as constants and transferred to other study regions as in Randolph *et al*. (2008) and Hunter *et al*. (2010) as they depend on the optical conditions of the water body as well as the validity of assumptions or simplifications of the algorithm.

Despite the simplicity, empirical models performed better than the semi-analytical model. Mean relative error of $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$ model was 25.4% and 21.6% respectively. However, the median relative error of $R_{rs}(708)/R_{rs}(600)$ (~20%) was 2% lower than the $R_{rs}(708)/R_{rs}(620)$ model. As $a^*_{chl}(620)$ is greater than $a^*_{chl}(600)$, Mishra *et al.* (2009) suggested to use 600 nm instead of 620 nm in the reflectance band ratio to reduce the effect of chl-*a* dependence of the empirical model. However in this study, use of $R_{rs}(708)/R_{rs}(600)$ band ratio did not outperform $R_{rs}(708)/R_{rs}(600)$ rather the overall performance of both models was somewhat similar to each other. Even though the reason is not evident, it is assumed that this could be explained by pigment packaging effect. Samples were collected during algal bloom periods when the algal density as well

as the intracellular pigment concentration was very high. In this scenario, difference between $a^*_{chl}(620)$ and $a^*_{chl}(600)$ is possibly insignificant and therefore, $R_{rs}(600)$ and $R_{rs}(620)$ in the reflectance band ratio produces similar results in the final output. Because of lack of sufficient data, it has not been studied how the empirical models will perform in a study area with mixed phytoplankton community structure or when cyanobacteria co-occur with green algae and dominate the phytoplankton community. Model comparison and validation analysis was performed on a small dataset collected during two different season of the year 2010 and 2011. In the future, a large dataset should be used for comparison of model performance and accuracy assessment.

Table 3.1 Descriptive statistics of pigment and absorption coefficient measurements.

Parameters	Mean	Std	Min	Max	*N*
Pigment Concentrations					
PC (ug/L)	418.76	669.75	68.13	3032.47	25
Chl-*a* (µg l^{-1})	295.96	302.06	59.40	1376.60	25
Chl-*b* (µg l^{-1})	4.36	3.15	1.57	13.71	18
PC:Chl-*a*	1.23	0.73	0.30	3.29	25

Table 3.2 Model Parameters from regression analysis and results from the model validation: slope (*a*) and Intercept (*b*), coefficient of determination (R^2); adjusted coefficient of determination; standard errors of the estimates (STE) for both models are

	Calibration (*n*=16)						Validation (*n*=9)	
Indices	*a*	*b*	R^2	*Adj.* R^2	STE	*p*	Mean RE (%)	Median RE(%)
$R_{rs}(708)/R_{rs}(600)$	2.78	1.44	0.77	0.75	0.23	<0.0001	-8.73	-16.60
$R_{rs}(708)/R_{rs}(620)$	2.34	1.39	0.82	0.8	0.204	<0.0001	-3.31	-16.73

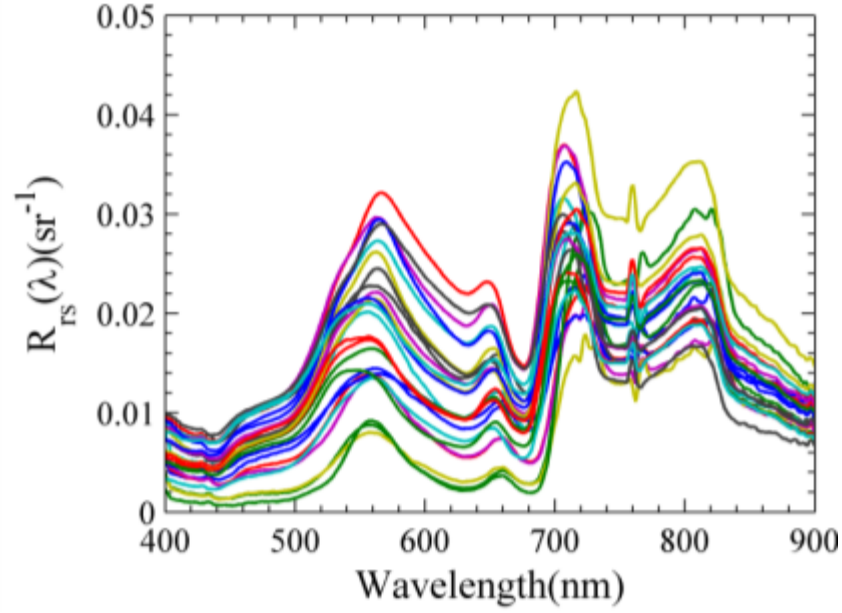

Figure 3.1 Remote sensing reflectance, $R_{rs}(sr^{-1})$ measured at study sites.

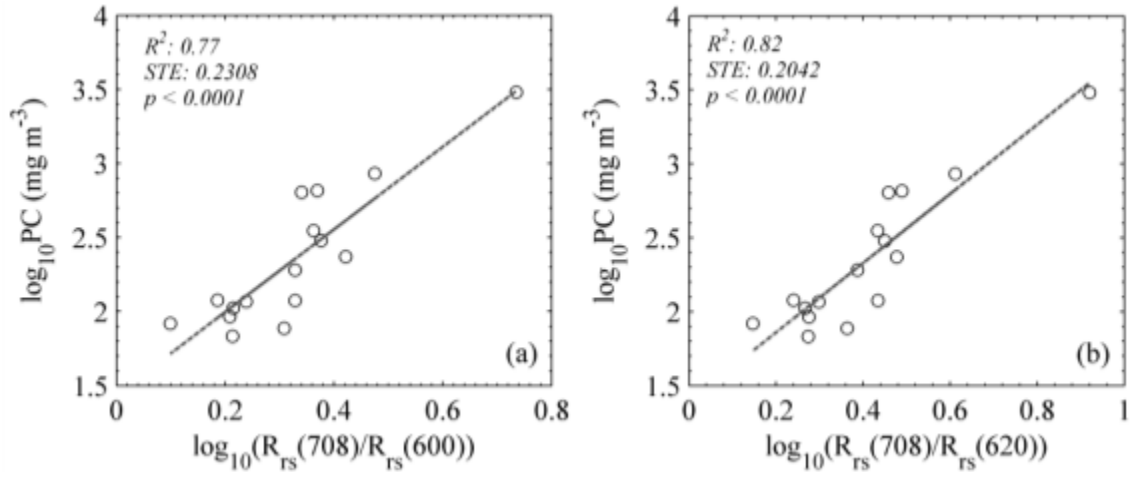

Figure 3.2 Calibration of empirical PC algorithm using (a) $R_{rs}(708)/R_{rs}(600)$ and (b) $R_{rs}(708)/R_{rs}(620)$ ratios.

Regression results are provided in Table 3.2.

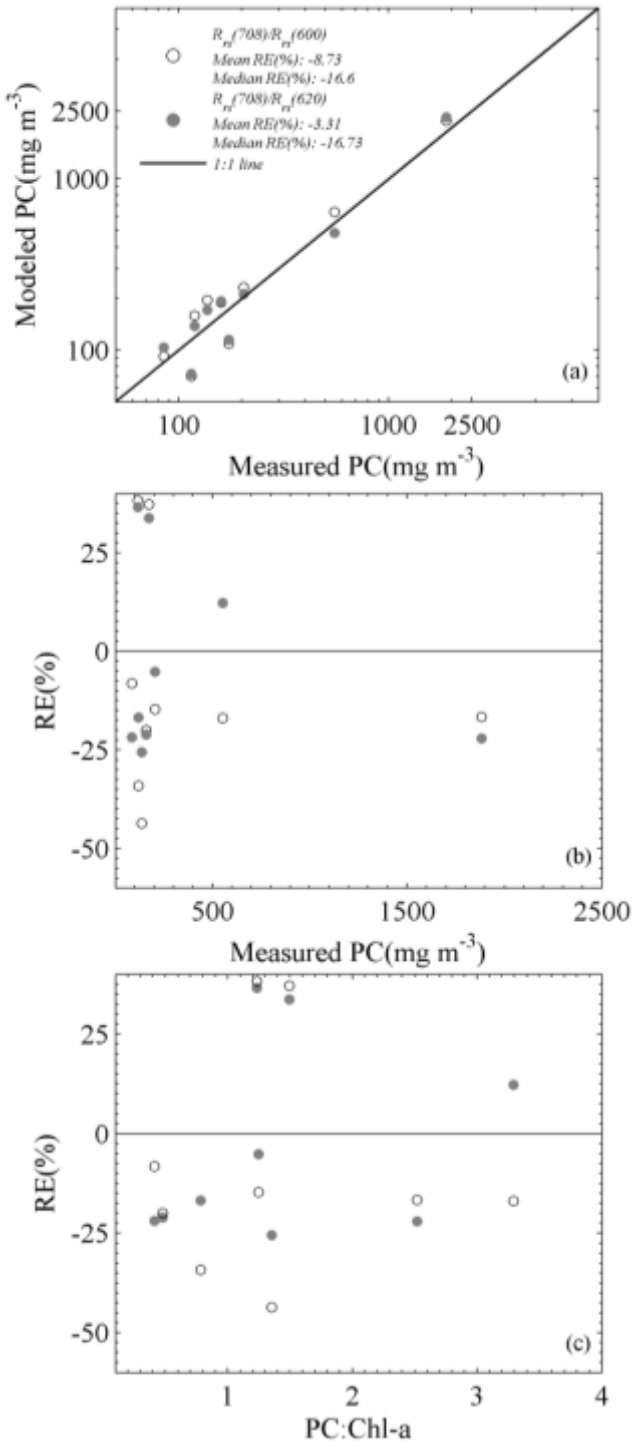

Figure 3.3 (a) Comparison of model retrieved PC concentration with measured values.

Solid line is the 1:1 line, (b) Scatter plot between relative error (%) and measured PC concentration, and (c) Scatter plot between relative error (%) and PC: chl-*a* pigment ratio.

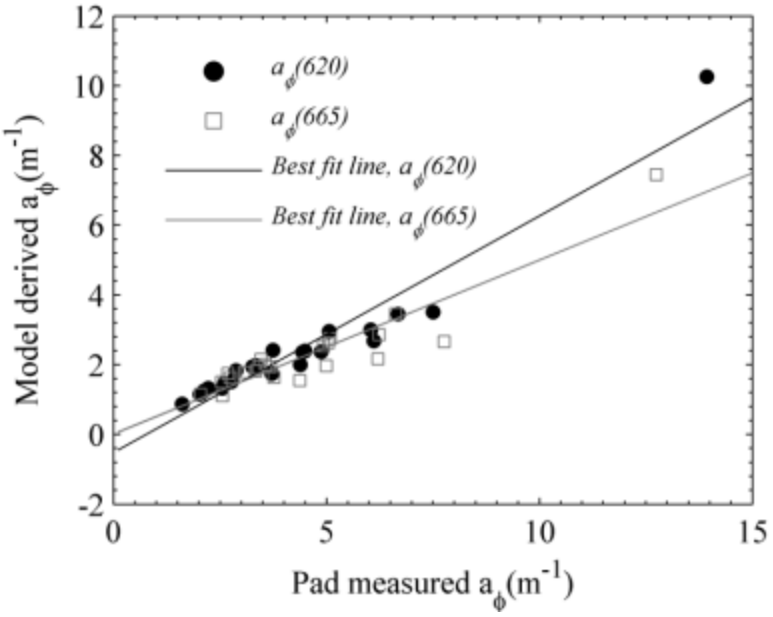

Figure 3.4 Relationship between pad-measured $a_\varphi(665,620)$ (x axis) and model retrieved uncorrected $a_\varphi(665, 620)$ (y axis).

Solid black and gray lines are the best-fit lines with regression equations- $a_\varphi(665)$: y = 0.678x - 0.5129, R^2=0.92, $a_\varphi(665)$: y = 0.4989x + 0.0155, R^2=0.85.

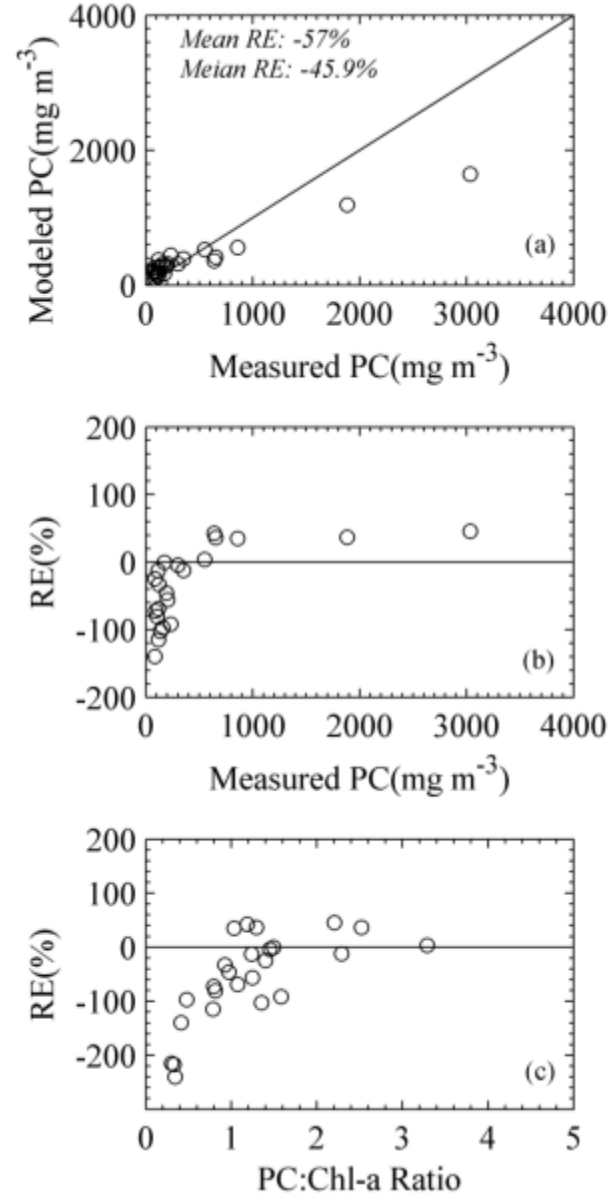

Figure 3.5 Validation of retrieved PC values from semi-analytical algorithm

(a) comparison of modeled PC values with measured ones. Solid line is the 1:1 line, (b) Scatter plot between Relative error, (RE)(%) and measured PC concentration, and (c) scatter plot between relative error, (RE)(%) and PC: chl-*a* pigment ratio.

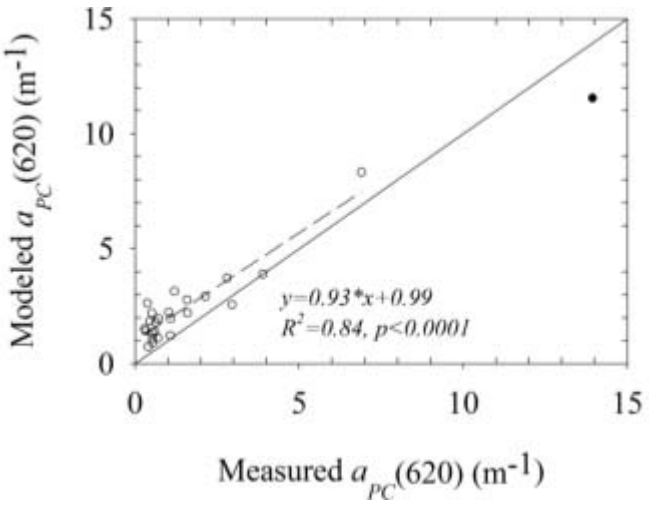

Figure 3.6 Comparison of aPC(620) with spectrophotometrically measured values.

Solid line is the 1:1 line and the dashed line is the least-square fit line. Sample represented with a filled circle was excluded from regression analysis.

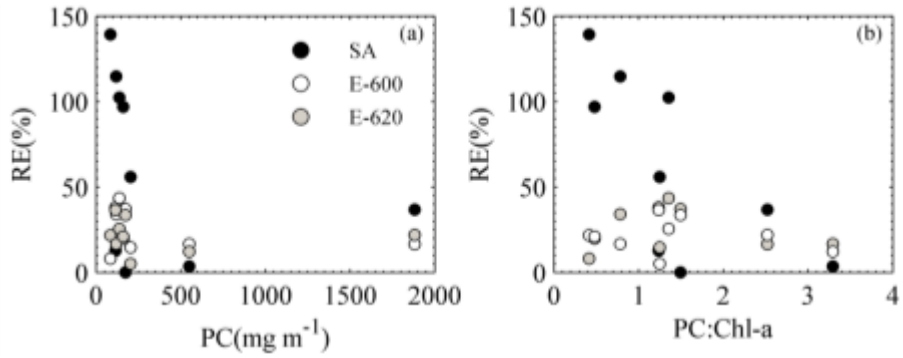

Figure 3.7 (a) Comparison of model estimation error from the semi-analytical and empirical models, and (b) scatter plot between RE(%) and PC: chl-*a* ratio.

Solid black, empty, and solid gray circles represent the semi-analytical algorithm, $R_{rs}(708)/R_{rs}(600)$, and $R_{rs}(708)/R_{rs}(620)$ empirical models. Note that relative errors are high for samples with PC:chl-*a* ratio of about less than 1.5.

3.6 References

Arar E.J. (1997). In vitro determination of chlorophylls a, b, c1 + c2 and pheopigments in marine and freshwater algae by visible spectrophotometry. USEPA Method 446-0.

Backer, L. C. (2002). Cyanobacterial harmful algal blooms: developing a public health response. *Lake and Reservoir Management, 18*, 20−31.

Ballot, A., Stephan, P., Wiegand, C., Kotut, K., & Krienitz, L. (2003). Cyanobacterial toxins in Lake Baringo, Kenya. *Limnologica, 33,* 2-9.

Bennett, A., & Bogorad, L. (1973). Complementary chromatic adaptation in a filamentous blue-green alga. *Journal of Cell Biology*, 58: 419–435.

Bricaud, A., Babin, M., Morel, A., & Claustre, H. (1995). Variability in the Chl-a-specific absorption-coefficients of natural phytoplankton—analysis and parameterization. *Journal of Geophysical Research-Oceans, 100,* 13321–13332.

Carmichael, W. W. (1997). The cyanotoxins. *Advances in Botanical Research, 27,* 211–256.

Dall'Olmo, G. & Gitelson, A.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(3): 412-422.

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

Emerson, R. & Lewis, C. M. (1942). The photosynthetic efficiency of phycocyanin in Chroococcus, and the problem of carotenoid participation in photosynthesis. *Journal of General Physiology, 25,* 579-595.

Gons, H. (1999). Optical teledetection of chlorophyll *a* in turbid inland waters, *Environmental Science Technology*, 33: 1127–1132

Gordon, H., Brown, O., Evans, R., Brown, J., Smith, R., Baker, K., & Clark, D. K. (1988). A semianalytical radiance model of ocean color, *Journal of Geophysical Research*, 93: 10909–10924

Gibson, C.E. &, Smith, R.V. (1982). Freshwater plankton., In: N. G. Carr and Whitton, B. N. (eds), The biology of cyanobacteria. Blackwell Scientific Publication Ltd. Oxford. pp. 463-489.

Glazer, A. N. (1989). Light guides. Directional energy transfer in a photosynthetic antenna. *Journal of Biological Chemistry, 264*, 1-4.

Golbeck, J. H., & Bryant, D. A. (1991). Photosytem I , In C. P. Lee (ed.), *Current Topics In Bioenergetics, Light Driven Reactions In Bioenergetics* (pp. 83-177). New York: Academic Press Inc.

Gordon, H. R., and Morel, A. (1983). Remote assessment of ocean color for interpretation of satellite visible imagery-A review, 44 pp., Springer-Verlag, New York.

Iwasa, M., Yamamoto, M., Tanaka, Y., Kaito, M., & Adachi, Y. (2002). Spirulina-associated hepatotoxicity. *American Journal of Gastroenterology, 97,* 3212–3213.

Kutser, T., Metasamma, L., Strombeck, N., & Vahtmae, E. (2006). Monitoring cyanobacterial blooms by satellite remote sensing. *Estuarine Coastal and Shelf Science, 67*, 303-312.

Mishra, S., Mishra, D.R., & Schluchter. W.M. (2009). A Novel Algorithm for Predicting PC Concentrations in Cyanobacteria: A Proximal Hyperspectral Remote Sensing Approach. *Remote Sensing*, *1*: 758-775.

Mishra, S., Mishra, D. R., & Lee, Z.P. (2012). Bio-Optical Inversion in Highly Turbid and Cyanobacteria Dominated Waters, *IEEE Transactions on Geoscience and Remote Sensing* (Under Review)

Morel, A. & Gentili, B. (1993). Diffuse reflectance of oceanic waters. 2. Bidirectional aspects., *Applied Optics,* 32, 6864-6872.

Ohde, T. & Siegel, H. (2003). Derivation of immersion factors for the hyperspectral TriOS radiance sensor, *Journal of Optics A: Pure and Applied Optics,* 5**:** L12–L14.

Randolph, K., Wilson, J., Tedesco, L., Li, L., Pascual, D., & Soyeux, M. (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.

Raven, P. H., Evert R. F., & Curtis, H. (1976). *Biology of Plants*. New York: Worth Publishers.

Richardson, L. L. (1996). Remote sensing of algal bloom dynamics. *BioScience, 46,* 492-501.

Rowan, K. S. (1989). *Photosynthetic Pigments of Algae*. Cambridge: Cambridge University Press.

Ruiz-Verdu, A., Simis, S. G. H., de Hoyos, C., Gons, H. J., & R. Pena-Martinez. (2008). An evaluation of algorithms for the remote sensing of cyanobacterial biomass. *Remote Sensing of Environment,* 112: 3996−4008.

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

Sathyendranath, S., Lazzara, L., & Prieur, L. (1987). Variations in the spectral values of specific absorption of phytoplankton. *American Society of Limnology and Oceanography, 32,* 403−415.

Schalles, J., & Yacobi, Y. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid and Chl-a pigments in eutrophic waters. *Archive Hydrobiological Special Issues Advance Limnological, 55,* 153-168.

Simis, S., Peters, S., & Gons, H. (2005). Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *American Society of Limnology and Oceanography, 50,* 237–245.

Simis, S., Ruiz-Verdú, A., Domínguez-Gómez, J., Peña-Martinez, R., Peters, S., & Gons, H. (2007). Influence of phytoplankton pigment composition on remote sensing cyanobacterial biomass. *Remote Sensing of Environment*, *106*, 414−427.

Tucker, C.S., & Boyd, C.E. (1985). Water quality. In Channel catfish culture. Ed. C.S. Tucker. Elsevier Science Publishers. Amsterdam. p. 135-227.

Vincent, R. K., Qin, X. M., McKay, R. M. L., Miner, J., Czajkowski, K., Savino, J., & Bridgeman, T. (2004). Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment*, *89*, 381-392.

CHAPTER IV

A QUASI-ANALYTICAL ALGORITHM TO QUANTIFY CYANOBACTERIAL PHYCOCYANIN: MODEL PARAMATERIZATION

4.1 Overview

Phytoplankton pigment absorption data from algal bloom dominated waters is highly desirable to better understand the primary productivity and carbon uptake by algal biomass at a regional scale. However, retrieving phytoplankton pigment absorption coefficients, in turbid and hyper-eutrophic waters, from remote sensing reflectance is often challenging because of the optical complexity of the water body. In this paper, a quasi analytical algorithm (QAA) has been parameterized using *in situ* data to retrieve inherent optical properties from above surface remote sensing reflectance in highly turbid and cyanobacteria dominated aquaculture ponds. New empirical relationships were developed to retrieve absorption by colored detrital matter in order to decompose the total absorption coefficient to derive phytoplankton absorption coefficient. The model was validated using a separate dataset by comparing the model derived optical parameters with *in situ* measured values. Percentage error of the estimated total absorption coefficient, $a_t(\lambda)$, values varied from 15.22-24.13 % within 413-665 nm and the average error was 19.87%. Maximum and minimum error occurred at 443 and 665 nm respectively. Similarly percentage error for phytoplankton absorption coefficient, $a_\varphi(\lambda)$, varied from 16.94-34.15 % within 413-665 nm range and the average error was 23.93%.

Spectral shape of modeled $a_{\varphi}(\lambda)$ matched very well with the measured $a_{\varphi}(\lambda)$ spectral shape. In this current form, it can be applied to retrieve $a_{\varphi}(\lambda)$ in turbid productive waters as well as cyanobacterial algal bloom (CAB) dominated waters in inland lakes, coastal estuarine environments, and marginal seas using MERIS and other hyperspecral sensors.

4.2 Introduction

Phytoplankton pigment concentration is often considered as a proxy for algal biomass and an indicator of ecological health of inland, coastal, and open ocean pelagic waters. Phytoplankters are responsible for half of the global primary productivity (Falkowski *et al*., 2004; Field *et al*., 1998), therefore, a better understanding of their phenology is essential to study the global carbon cycle and its effect on changing climate (Brewin *et al*., 2010). Ocean color remote sensing offers great potential for quantifying phytoplankton biomass and primary productivity in global ocean. There exist a large number of methods and algorithms of empirical and semi-analytical nature for remote estimation of phytoplankton pigment concentration using ocean color data (Gordon and Morel, 1983; Morel and Prieur, 1977; O'Reilly *et al*., 1998; Gons *et al*., 2000; Maritorena *et al*., 2002). Empirical algorithms normally rely on statistical relationship between band ratios of reflectance and bio-optical parameters. Therefore, these algorithms often produce accurate retrievals if the concentration range and proportion of optically active constituents are similar or close to the ones used in the calibration dataset.

In contrast, semi-analytical algorithms (Gordon *et al*., 1988; Roesler and Perry, 1995; Garver and Siegel, 1997; Carder *et al*., 1999; Maritorena *et al*., 2002) are based on the inversion of optical properties of water using radiative transfer models, where, $R_{rs}(\lambda)$ (abbreviations are in Table 1) is linked with absorption and scattering budget of light by

optically active constituents and water itself. Semi-analytical algorithms are comparatively more robust than empirical algorithms and can retrieve optical parameters accurately given that the inputs, such as the spectral shape of phytoplankton absorption coefficient and the spectral slope of gelbstoff absorption, are accurate. These input parameters are often pragmatically assumed to be the global mean (e.g., Maritorena *et al*., 2002); if the local values significantly deviate from the global mean, the accuracy of the model retrieval reduces accordingly.

To address this issue, Lee *et al*. (2002) developed a multiband quasi-analytical algorithm to retrieve absorption and backscattering coefficients from $R_{rs}(\lambda)$ in optically deep waters. The advantage and uniqueness of this algorithm is that, unlike other semi-analytical algorithms, it does not require spectral models of the absorption coefficient of phytoplankton, non-algal particulates (NAP) and colored dissolved organic matter (CDOM) for the derivation of total absorption coefficient, $a_t(\lambda)$. In addition, other semi-analytical models derive component contributions first, and then estimate $a_t(\lambda)$; while QAA derives $a_t(\lambda)$ first, then estimates the contributions of components to the absorption budget. The model has been extensively validated using simulated and field datasets from different geographic regions (Lee *et al*., 2002; Lee *et al*., 2004; Le *et al*., 2009, Craig *et al*., 2006; Zhu *et al*., 2010). In brief, the algorithm empirically estimates a_t at a reference wavelength (λ_0), where, the $a_t(\lambda_0)$ is dominated by $a_w(\lambda_0)$ then it estimates particulate back-scattering coefficient at a reference wavelength, $b_{bp}(\lambda_0)$, by using relationships between $r_{rs}(\lambda)$ and IOPs derived from radiative transfer equations. Further, it estimates $b_{bp}(\lambda)$ and then $a_t(\lambda)$ spectrum from measured $R_{rs}(\lambda)$. The derived $a_t(\lambda)$ is further decomposed into $a_\varphi(\lambda)$ and $a_{CDM}(\lambda)$ taking advantage of the exponential function of

$a_{CDM}(\lambda)$ spectrum, which can be modeled from $r_{rs}(\lambda)$. Le *et al*. (2009) recently used this algorithm to retrieve $a_t(\lambda)$ in highly turbid waters of Meiliang Bay in Taihu Lake, China and reported that the algorithm successfully retrieved $a_t(\lambda)$ with a difference of less than 20% between the measured and modeled data. $a_t(\lambda)$ spectrum was not further decomposed in that study and therefore the accuracy of $a_\varphi(\lambda)$ retrieval cannot be investigated.

Further assessment of the QAA performance was carried out in this study by using a dataset with very high chlorophyll-*a* (chl-*a*) concentration (up to 1376 mg m^{-3}) and extremely high $a_t(443)$(47.21 m^{-1}) that were acquired from turbid and eutrophic waters, such as aquaculture ponds in Stoneville, MS. Upon application of the QAA to the R_{rs} data to retrieve $a_t(\lambda)$, the model underestimated the $a_t(\lambda)$ values up to a factor of 5-10 at all wavelengths within 400-700 nm range (Fig. 1a). This underestimation can be explained by the inaccurate retrieval of a_t at the reference wavelength, 560 nm. In addition, existing steps to separate $a_{CDM}(\lambda)$ from $a_t(\lambda)$ to retrieve $a_\varphi(\lambda)$ produced many negative retrievals of $a_{CDM}(\lambda)$. Similarly, the model retrieved $a_\varphi(\lambda)$ values with twofold underestimation as compared to the pad-measured values (Fig 2a and 2b). QAA in the native form is able to retrieve $a_t(\lambda)$ and $a_\varphi(\lambda)$ successfully in waters where $a_t(443)$ is less than ~0.5 m^{-1}. However, the empirical scheme to retrieve the $a_t(\lambda_o)$ at a reference wavelength located around 700 nm needs to be parameterized, instead of simply using the value of pure water (Doron *et al*., 2007), for successful retrieval of $a_t(\lambda)$ and $a_\varphi(\lambda)$ in turbid and productive waters and algal bloom dominated waters with $a_t(443)$ value often exceeding 0.5 m^{-1}.

The main objective of this research was to parameterize the QAA algorithm for accurate retrieval of $a_\varphi(\lambda)$ in highly turbid and productive waters. The specific objectives are: (1) parameterize the QAA algorithm using *in situ* bio-optical dataset collected from several inland ponds with highly turbid and productive waters; (2) validate the algorithm using an independent dataset collected from the same area in a different year; and (3) evaluate the accuracy of the retrieval by comparing the model derived $a_\varphi(\lambda)$ with pad-measured $a_\varphi(\lambda)$ values.

4.3 Data and methods

In situ data were collected from aquaculture ponds at Delta Research Extension Center, Mississippi State University research facility located near Stoneville, MS, USA during multiple field campaigns in 2010 (13-16 July) and 2011 (28-29 April). Depth of the ponds varied from 1.2-1.8 meters and the area ranged from 4-8 ha. The pond management practices in Mississippi include a high feeding rate, which results in a very high concentration of nitrogen and phosphorous (Tucker and Boyd, 1985). High concentrations of nutrients (mean nitrite concentration = 0.12 mg L^{-1} and mean total ammonia concentration = 0.48 mg L^{-1}) (Tucker *et al*., 2009) along with high temperature and intense sunlight cause the development of phytoplankton blooms in these impoundments during summer months. Chl-*a* concentration goes up to as high as 1000 mg m^{-3} and cyanobacteria are the abundant species of phytoplankton during these summer blooms (Tucker and Boyd, 1985). Impoundments are typically turbid because of wind induced re-suspension of minerals and clay particles in the water column.

4.3.1 Remote sensing reflectance (R_{rs})

Hyperspectral R_{rs} data were collected using two calibrated Ocean Optics spectroradiometer (Ocean Optics Inc., Dunedin, FL, USA) as given in Dall'Olmo *et al*. (2005). The dual-fiber system, with two inter-calibrated Ocean Optics radiometers mounted on a platform, acquired reflectance data in the range of 400-900 nm with a sampling interval of 0.3 nm. One radiometer, equipped with a 25° field-of-view optical fiber pointed downward to measure the upwelling radiance just below the air-water interface, expressed in digital numbers as $DN_{Lu}(\lambda)$; while the second radiometer, equipped with an optical fiber and cosine diffuser (yielding a hemispherical field of view), pointed upward to acquire above surface downwelling irradiance just above the surface , expressed in digital numbers as $DN_{Ed}(\lambda)$. To match their transfer functions, inter-calibration of the radiometers was accomplished by measuring the upwelling radiance of a white Spectralon reflectance standard (Labsphere, Inc., North Sutton, NH, USA) simultaneously with incident irradiance. To mitigate the impact of solar elevation on radiometer inter-calibration, the anisotropic reflectance from the calibration target was corrected in accordance with Jackson *et al*. (1992). The two radiometers were inter-calibrated immediately before and after measurements in each field site. After the data acquisition, $R_{rs}(\lambda)$ was calculated as follow:

$$R_{rs}(\lambda) = \frac{DN_{Lu}(\lambda)}{DN_{Ed}(\lambda)} \frac{DN_{Ed,ref}(\lambda)}{DN_{Lu,ref}(\lambda)} \frac{\rho_{ref}(\lambda)}{\pi} F_i(\lambda) \frac{t}{n^2} \quad (4.1)$$

where, t is the transmittance at the air-water interface (0.98); n is the refractive index of water (1.34); $DN_{Lu,ref}$ and $DN_{Ed,ref}$ are digital numbers representing upwelling radiance and downwelling irradiance over the white Spectralon panel; ρ_{ref} is the irradiance

reflectance of the Spectralon panel; F_i (λ) is the spectral immersion coefficient (Ohde and Siegel, 2003). For each station, a representative $R_{rs}(\lambda)$ spectrum was calculated by averaging six consecutive scanned spectra (Fig. 3).

4.3.2 Absorption measurements-$a_d(\lambda)$, $a_\varphi(\lambda)$, and $a_{CDOM}(\lambda)$

Surface water samples were collected in 1 L Niskin bottles and immediately filtered onto 0.7 µm Whatman GF/F filters under low vacuum (<5 inch of mercury). The volume of water filtered varied from 50-100 mL depending on the load of particulate matter in the sample. Particulate absorption coefficient, $a_p(\lambda)$, was determined using standard quantitative filtration technique (QFT) procedure as described in Fargion and Muller (2000). A Perkin Elmer lambda-850 spectrophotometer (Perkin Elmer Inc., Waltham, MA, USA) with an integrating sphere was used to scan the samples within a spectral range from 400 to 800 nm. Absorption coefficient of detrital matter, $a_d(\lambda)$, was measured after bleaching the filters in 0.2% active chlorine aqueous solutions to remove the phytoplankton pigments (Tassan and Ferrari, 1995). $a_p(\lambda)$ and $a_d(\lambda)$ were calculated after correction for optical path-length elongation due to multiple scattering using a 'β factor' for highly turbid and productive waters as in Dall'Olmo (2006). Further, $a_\varphi(\lambda)$ was computed by subtracting $a_d(\lambda)$ from $a_p(\lambda)$. Finally, $a_\varphi(\lambda)$ was corrected for residual scattering by subtracting $a_\varphi(800)$ from all wavelengths.

Water samples for colored dissolved organic matter (CDOM) analysis were filtered immediately after collection through a 0.2 µm nucleopore membrane filters under low vacuum. Filtered samples were stored at 4°C in amber color glass bottles untill analysis. Filtered samples were allowed to reach ambient room temperature to minimize temperature bias between samples and blank (Milli-Q water) before absorption

measurements. Absorbance, $A_{CDOM}(\lambda)$, of discrete samples were measured using the high performance Perkin Elmer Lambda 850 spectrophotometer within 350 and 850 nm range with 1 nm spectral resolution. The absorbance data were corrected for baseline fluctuations by subtracting the mean value between 780 and 784 nm from the absorbance measured at each wavelength. The specified wavelength was chosen for baseline correction because $a_{CDOM}(782)$ is negligible and it is minimally affected by temperature dependent variations in water absorption (Buiteveld *et al*., 1994). $a_{CDOM}(\lambda)$ (m^{-1}) for path-length, l (m) was calculated as:

$$a_{CDOM}(\lambda) = \frac{2.303[A_{CDOM}(\lambda)]}{l} \quad (4.2)$$

4.3.3 Water quality parameters

Water samples for chl-*a* and chl-*b* analysis were collected in 1 L Niskin bottles and immediately filtered onto GF/F filters (Whatman, 0.7 µm pore size) under low vacuum (<5 inches of mercury). Chl-*a* and chl-*b* were extracted in triplicates using acetone extraction procedure and concentrations were measured using HPLC as in Environmental Protection Agency (EPA) method 447 (Arar, 1997). Only for the 2010 data set, the concentrations of total suspended solids (TSS), organic suspended solids (OSS), and inorganic suspended solids (ISS) were measured gravimetrically (ASTM, 1999).

4.3.4 Accuracy assessment methods

The inversion model was parameterized using the data set collected in July 2010 and the performance of the model was evaluated using an independent data set collected in April 2011. Model derived $a_{CDM}(\lambda)$ and $a_{\varphi}(\lambda)$ were compared with pad-measured

$a_{CDM}(\lambda)$ and $a_{\varphi}(\lambda)$ and the percentage error (ε) was calculated for all samples to access the accuracy of the inversion method (Lee *et al*., 2002).

$$RMSE_{log10} = \left\{ \frac{\sum_{i=1}^{n} \left[log_{10}\left(q_i^{model}\right) - log_{10}\left(q_i^{measure}\right) \right]^2}{n} \right\}^{1/2} \quad (4.3)$$

Linear percentage error:

$$\varepsilon = (10^{RMSE_{log10}} - 1) \times 100\%, \quad (4.4)$$

where, q represents optical parameters such as $a_t(\lambda)$, $a_{CDM}(\lambda)$, and $a_{\varphi}(\lambda)$; and the superscripts "model" and "measure" represent model derived and laboratory measured parameters respectively. In addition, the root mean square error (RMSE) and slope, intercept, and R^2 of regression relations between measured and modeled parameters were also used to access the accuracy whenever needed.

4.4 Parameterization of the QAA inversion algorithm

4.4.1 Derivation of total absorption coefficients, $a_t(\lambda)$

One of our main objectives of our research was to reparameterize the QAA algorithm to make it suitable to work particularly in highly absorbing waters and algal bloom scenarios in inland ponds, lakes, and coastal and estuarine environments, where, chl-*a* concentration reaches as high as 1000 mg m^{-3}. As discussed earlier, QAA algorithm with reference wavelength at 560 nm underestimated $a_t(\lambda)$ values in our study area up to a factor of 5-10 at all wavelengths within 400-700 nm range (Fig. 1a). This underestimation can be explained by the inaccurate retrieval of a_t at the reference wavelength, 560 nm. In the entire dataset $a_t(560)$ varied from 1.14 to 10.38 m^{-1}, which is a factor of 18-167 of $a_w(560)$ value. At 560 nm, the total absorption budget is dominated

by non-water absorbing constituents that make the empirical estimation of $a_t(560)$ unreliable. An underestimated $a_t(560)$ value causes an underestimation of $b_{bp}(560)$ and propagates the error to a_t estimation at other wavelengths. Lee *et al*. (2002) shifted the reference wavelength to 640 nm for accurate retrieval of $a_t(\lambda)$ in strongly absorbing (eutrophic) waters ($a_t(440)>0.3$ m^{-1}). In studies (e.g. Doron *et al*., 2007; Le *et al*., 2009) the reference wavelength was shifted to 708 nm and $a_t(708)$ was assumed equal to $a_w(708)$, and $a_t(\lambda)$ were successfully retrieved with an average percentage difference of less than 15% between the measured and predicted $a_t(\lambda)$ values. However, retrieval accuracy did not improve drastically when the reference wavelength was moved to 708 nm in this dataset. The model still underestimated $a_t(\lambda)$ values up to about two fold at all wavelengths (Fig. 1b). This is because of the fact that in this present scenario, $a_t(708)$ values range from 1.06-4.1 m^{-1} which is a factor of 1.28-4.95 of $a_w(708)$ values in the entire dataset. Therefore, $a_t(708)$ cannot be assumed to be pure water absorption coefficient at 708 nm in highly eutrophic waters. Attempts were made to find an empirical relationship between χ formulated in the Lee *et al*. (2002) and $a_t(708)$; however, the output was not satisfactory. I formulated χ using $r_{rs}(\lambda)$ data at 443, 620, and 708 nm from the current dataset. The mathematical formulation of χ is expressed as:

$$\chi = log_{10}\left(\frac{0.01*r_{rs}(443)+r_{rs}(620)}{r_{rs}(708)+0.005*\frac{r_{rs}(620)}{r_{rs}(443)}*r_{rs}(620)}\right) \quad (4.5)$$

4.5 Results and discussions

4.5.1 Water quality parameters

The aquaculture ponds were highly turbid and productive during the sampling time with an average chl-*a* of 532 µg l^{-1} and average TSS of 177 mg l^{-1} in the dataset collected in 2010. The proportions of suspended solids in those waters were equally dominated by minerals and organic particles. ISS/TSS ratio varied from 16-68% indicating the water to be highly turbid and similarly high value of OSS/TSS (32-84%) indicates their organic origin (Table 3). In general, TSS values were positively correlated with that of chl-*a* (R^2=0.23), which corroborates that the algal biomass contributed towards the TSS concentrations in the impoundments. TSS and ISS showed a strong positive correlation indicating minerals and clay particles as a primary source of TSS (R^2=0.93, p<0.0001, Fig.4a). OSS and chl-*a* values were also correlated very well (R^2=0.79, p<0.0001) indicating chl-*a* is one of the primary source of OSS (Fig. 4b).

Results from CDOM analysis showed that the spectral slope of S_{CDOM} varied from 0.012 to 0.017 nm^{-1}. The average value of S_{CDOM} was 0.015 nm^{-1}, which was in accordance with observations in other studies with similar waters (Dall'Olmo *et al*., 2005; Twardowski *et al*., 2004; Babin *et al*., 2003). Values of $a_{CDOM}(443)$ ranged from 0.66 to 2.59 m^{-1} with a mean of 1.25 m^{-1} and they were positively correlated with chl-*a* (R^2=0.39, figure not shown) and OSS (R^2=0.43, figure not shown) suggesting that the degradation of algal biomass as an important source of CDOM in the aquaculture ponds. In general, the decomposition of algal biomass and mineralization of CDOM controlled the magnitude of a_{CDOM} as the ponds are closed systems.

4.5.2 Deriving total absorption coefficients, $a_t(\lambda)$

Phytoplankton pigment concentrations in the samples collected during the 2010 field season were very high ($a_\varphi(443)_{max} = 37.77 m^{-1}$). Therefore, the $a_t(708)$ was still dominated by pigment absorption, for example, $a_\varphi(708)$ contribution to the $a_t(708)$ varied from 20 to 79% in the dataset. Also, magnitudes of $a_t(708)$ ranged from 1.06 to 4.1 m^{-1} with a standard deviation of 0.74 m^{-1}, which is a factor of 1-4 of the $a_w(708)$ values. Therefore, $a_t(708)$ cannot be assumed to be $a_w(708)$ in these kinds of waters unlike Le *et al*. (2009) assumed in Meiliang Bay, China. Similar to Lee *et al*. (2002), $a_{t-w}(708)$ was parameterized by finding an empirical relationship with χ described in the method section (Fig. 5, Table 2). The regression analysis showed strong sensitivity between $a_{t-w}(708)$ and χ (R^2=0.84, n=20, p<0.0001). To evaluate the accuracy of the empirical relation, it was applied to the 2011 data set and validation analysis showed an excellent agreement between the measured and modeled $a_t(708)$ by producing a percent linear error of 10.2%, and slope and R^2 between measured and modeled values were 0.96 and 0.72 respectively (figure not shown). Accuracy of the developed model was comparable with the accuracy of the Lee *et al*. (2002) inversion model at blue-green spectral region for open ocean waters and Le *et al*. (2009) model for turbid case 2 waters. It is anticipated that even higher accuracy can be achieved if the reference wavelength can be even moved further towards longer wavelengths around 753 nm; however, accurate retrieval of $R_{rs}(\lambda)$ around that spectral region from a remote sensor is a challenging task in case 2 waters.

4.5.3 Validation of the inversion model

Accuracy of the newly parameterized inversion model was assessed using an independent dataset acquired during a different year and season (sampling time: April

2011). Outputs of the inversion model such as $a_t(\lambda)$, $a_{CDM}(\lambda)$, and $a_\varphi(\lambda)$ were compared with the corresponding measured parameters within a wavelength range of 413-665 nm and the error analysis was performed. Following the steps described above and summarized in Table 2, $a_t(\lambda)$ for each sample was derived from $R_{rs}(\lambda)$ measurements. $b_{bp}(\lambda_0)$ at the reference wavelength (708 nm) was derived using step 3 in Table 2, and $b_{bp}(\lambda)$ at each wavelength was modeled using the spectrally dependent hyperbolic function (Gordon and Morel, 1983) summarized in steps 4 and 5 (Table 2). Modeled $b_{bp}(\lambda)$ spectra showed considerable variations both in magnitude and spectral shape (Fig. 6). $b_{bp}(560)$ varied from 0.96 m^{-1} to 1.87 m^{-1} with an average value of 1.23 m^{-1} in the validation dataset. Accuracy of the $b_{bp}(\lambda)$ and the exponent of the hyperbolic function, η, were not studied because of unavailability of the measured values in this study. Further, $a_t(\lambda)$ values were derived from $u(\lambda)$ as described in step 6 (Table 2).

The comparison between modeled $a_t(\lambda)$ and measured $a_t(\lambda)$ is shown in Fig. 7. The magnitude and spectral shape of the modeled $a_t(\lambda)$ spectra matches well with the measured ones. In order to quantify the deviation, percentage errors were calculated as a function of wavelength (MERIS band centers), $\varepsilon(\lambda)$, and at each station, $\varepsilon(n)$, within a wavelength range of 413-665 nm (MERIS band1-band7). No comparisons were made at 681 nm as there was no effort to correct for the effect of chl-*a* fluorescence at that wavelength. The $\varepsilon(n)$ values varied from 6.3-57.6% with an average of 17.15% for all sampling stations (Table 4). Although $\varepsilon(n)$ values were positively correlated with $a_d(443)$ ($r = 0.38$) and $a_d(443)/a_p(443)$ ratio ($r = 0.28$), high retrieval errors at sampling stations S-2 and S-20 with $\varepsilon(n)$ values of 57.6% and 41.3% respectively could not be explained by the $a_d(443)/a_p(443)$ ratio, where a ~2% difference in $a_d(443)/a_p(443)$ ratio between

stations caused an increase in $\varepsilon(n)$ values by ~50%. Further analysis showed that the retrieval error of $a_{t\text{-}w}(708)$, empirically modeled from χ, at S-2 and S-20 stations were 40% and 23% respectively. Therefore, it can be concluded that modeling error associated with $a_{t\text{-}w}(708)$ at those locations propagated to the final step and increased the $a_t(\lambda)$ retrieval error. After excluding those two sampling stations, the error range and average reduced to 6.3-28.8% and 13.75%, which is very encouraging. Similarly, the error for $\varepsilon(\lambda)$ varied from 15.22-24.13 % and the average $\varepsilon(\lambda)$ was 19.87%. Maximum and minimum error occurred at 443 and 665 nm respectively (Table 5).

$a_{CDM}(\lambda)$ were retrieved using steps 7-10 (Table 2). The model retrieved $a_{CDM}(\lambda)$ with an average of 2.44 fold underestimation at 413 nm. The maximum underestimation at 413 nm was observed to be ~8 fold at station 9. Underestimation of $a_{CDM}(\lambda)$ values in these turbid productive waters was probably because of a combination of following three reasons: 1) modeled $a_t(413)$ and $a_t(443)$ values were underestimated, 2) $a_{CDM}(443)$ makes just about 24% of $a_t(443)$, i.e. close to negligible information to be sensed by a passive remote sensor, consequently analytically resolved a_{CDM} could even be negative (see Lee *et al.,* 2010); and 3) the modeled ξ and ζ values did not match the field values. In our dataset, measured values of ξ and ζ varied between 1.45-1.75 and 0.86-1.1 respectively; whereas, the corresponding QAA v5 derived ξ and ζ values ranged from 1.72-1.77 and 0.91-0.96 (Fig. 2c and 2d). An overestimation of ξ and underestimation of ζ can also lead to the underestimation of $a_{CDM}(443)$. Because of the same reasons, the model also produced a few negative $a_{CDM}(443)$ retrievals. Wherever the model retrieved negative values of $a_{CDM}(443)$, $a_{CDM}(\lambda)$ was assumed negligible for the retrieval of $a_\varphi(\lambda)$ from $a_t(\lambda)$.

Finally, $a_\varphi(\lambda)$ were calculated by subtracting $a_{CDM}(\lambda)$ and water absorption coefficient, $a_w(\lambda)$, from $a_t(\lambda)$. The model retrieved $a_\varphi(\lambda)$ values were compared with measured $a_\varphi(\lambda)$ at each station within 413-665 nm spectral range and for all samples at each MERIS band centers. Even though the retrieved $a_{CDM}(\lambda)$ were underestimated, retrieved $a_\varphi(\lambda)$ within 413 nm to 665 nm wavelength range were quite good (Fig. 8). For all sampling stations, the $\varepsilon(n)$ varied within 7.09 -57.36% with an average of 26.04%. The maximum error occurred at station 6. The model produced high error at four stations with $\varepsilon(n)$ exceeding 30% because of combination of two reasons: 1) either retrieved $a_{CDM}(\lambda)$ was negative, and therefore not subtracted from $a_t(\lambda)$, or 2) the retrieved $a_{CDM}(\lambda)$ was severely underestimated at that station. Similarly, the $\varepsilon(\lambda)$ varied from 15.9-41.27 % within 413-665 nm and the average error was 27.24%. Maximum and minimum error occurred at 413 and 665 nm respectively (Table 5).

Spectral shape of both measured and modeled $a_\varphi(\lambda)$ were also compared as another way of accuracy evaluation (Fig. 9). All measured and modeled $a_\varphi(\lambda)$ spectra were normalized to $a_\varphi(665)$ and the average spectral shape was computed for all samples at each wavelength. Spectral shape of modeled $a_\varphi(\lambda)$ matched very well (R^2=0.97) with the measured $a_\varphi(\lambda)$ within 510-665 nm wavelength range. The maximum deviation was found at blue spectral region, 413 nm (~83%) and 443 nm (~95%).

4.5.4 First order empirical approach for $a_{CDM}(\lambda)$ estimation

As shown earlier, due to a small contribution of CDM absorption to the total, the analytical scheme retrieved either very low or negative CDM absorption coefficient. To put a remedy for this outcome for such waters, a supplementary method has also been provided for the first order estimation of $a_{CDM}(\lambda)$ using empirical approach specifically

for these type of turbid productive waters. Existing CDOM empirical algorithms in the literature (e.g. D'Sa *et al*., 2006; Mannino *et al*., 2008) exploited the blue wavelength regions by selecting $R_{rs}(\lambda)$ at a single band or a R_{rs} band ratio to model a_{CDOM}. $a_{CDOM}(443)$ can be empirically modeled from a reflectance band ratio (Kowalczuk *et al*., 2005; Shanmugam, 2011). However, in algal laden highly absorbing waters, total absorption is dominated by a_{φ} in the blue spectral region. Therefore, the uncertainty of retrieving a_{CDM} at blue spectral region increases with higher a_{φ} to a_{CDM} ratio. In this research, $a_{CDM}(381)$, $a_{CDM}(443)$, and $a_{CDM}(510)$ values were derived using empirical approach. Linear regression equations were established between a log transformed reflectance band ratio, $[\log_{10}(r_{rs}(490)/r_{rs}(510))]$, denoted as $log_{10}(r_{510}^{490})$, and measured $a_{CDM}(381)$, $a_{CDM}(443)$ and $a_{CDM}(510)$. Selection of r_{rs} bands at 490 and 510 nm is based on the assumption that a_{φ} contribution at those bands will be comparatively less than 413-443 spectral region. A straight line fit explained the relations (Fig. 10a, b, and c). $a_{CDM}(510)$ showed a stronger empirical relationship with the band ratio by producing the highest R^2. Also, $\log_{10}(r_{rs}(490)/r_{rs}(665))$ showed its sensitivity to $a_{CDM}(381)$, $a_{CDM}(443)$ and $a_{CDM}(510)$ producing similar statistics as the previous band ratio. For comparison, scatter plot between $\log_{10}(r_{rs}(490)/r_{rs}(665))$ and $a_{CDM}(510)$ (Eq. 4.8) is shown in Fig. 10c.The resulting linear-fit regression equations were established as:

$$a_{CDM}(381) = -76.907[log_{10}(r_{510}^{490})] + 1.41, \quad R^2 = 0.45, n = 20 \qquad (4.6)$$

$$a_{CDM}(443) = -40.023[log_{10}(r_{510}^{490})] - 0.82, \quad R^2 = 0.47, n = 20 \qquad (4.7)$$

$$a_{CDM}(510) = -18.72[log_{10}(r_{510}^{490})] - 0.73, \quad R^2 = 0.5, n = 20 \qquad (4.8)$$

$$a_{CDM}(510) = -6.15[log_{10}(r_{665}^{490})] + 0.15, \quad R^2 = 0.5, n = 20 \qquad (4.9)$$

Similarly, S_{CDOM} (nm^{-1}) is often empirically modeled as a function of a_{CDOM} at ultraviolet and blue spectral regions (Kowalczuk *et al*., 2006; Schwarz *et al*., 2002) or an absorption ratio such as $a_{CDOM}(413)/a_{CDOM}(350)$ (Shanmugam, 2011). For model calibration, two empirical relationships were established between measured S_{CDM} and a_{CDM} using $a_{CDM}(381)/ a_{CDM}(510)$ and $a_{CDM}(443)/a_{CDM}(510)$ (Fig. 10d, and e). The empirical relationships are presented as:

$$S_{CDM-381} = 0.0057 \left[\frac{a_{CDM}(381)}{a_{CDM}(510)}\right]^{0.4884}, \qquad R^2 = 0.94, n = 20 \tag{4.10}$$

$$S_{CDM-443} = 0.005 \left[\frac{a_{CDM}(443)}{a_{CDM}(510)}\right]^{1.1327}, \qquad R^2 = 0.95, n = 20 \tag{4.11}$$

To confirm the accuracy of the S_{CDM} models, Eqs. 4.9 and 4.10 were applied to the validation dataset to retrieve $S_{CDM\text{-}381}$ and $S_{CDM\text{-}443}$ for each sample. Note that in validation, a_{CDM} ratios were modeled from $r_{rs}(490)$ and $r_{rs}(510)$ using Eq. 4.5-4.7 and used to retrieve $S_{CDM\text{-}381}$ and $S_{CDM\text{-}443}$. The residuals [measured S_{CDM} –modeled S_{CDM}] were calculated and analyzed to study the departures of the model predicted values from the measured S_{CDM} values. The residuals did not reveal any trend of over or under estimation (Fig. 10f). Most of the residuals are within $\pm$ 0.002 nm^{-1} with the exception of one sample up to -0.005 nm^{-1}. Overall, $S_{CDM\text{-}381}$ offered better retrievals compared to $S_{CDM\text{-}443}$ (Fig. 10f). Accuracy of the modeled $a_{CDM}(510)$ was better than $a_{CDM}(381)$ and $a_{CDM}(443)$. Therefore, $a_{CDM}(510)$ was selected as reference a_{CDM} in the exponential model to retrieve $a_{CDM}(\lambda)$. $a_{CDM}(\lambda)$ was modeled using the classic exponential model given by Bricaud *et al*. (1981):

$$a_{CDM}(\lambda) = a_{CDM}(510)\, e^{-S_{CDM}(\lambda - 510)}, \tag{4.12}$$

where, either $S_{CDM-381}$ or $S_{CDM-443}$ can be used in place of S_{CDM}. For comparison, both $S_{CDM-381}$ and $S_{CDM-443}$ were used to model $a_{CDOM}(\lambda)$ and termed as $a_{CDM-381}(\lambda)$ and $a_{CDM-443}(\lambda)$ respectively.

4.5.4.1 Limitation of the empirical retrieval of $a_{CDM}(\lambda)$

The empirical steps to retrieve $a_{CDM}(\lambda)$ in turbid productive waters were developed using a dataset with low a_{CDM} contribution to the $a_t(\lambda)$ (~8-45% with a mean of ~24% at 413 nm). Therefore, it should be used for first order estimation of $a_{CDM}(\lambda)$ only in low CDM (relative to $a_\varphi(\lambda)$) waters and existing steps in QAAv5 should be used otherwise. Because of overlapping absorption peak of chl-*a* and $a_{CDM}(\lambda)$, two band ratio of R_{rs} such as $R_{rs}(490)/R_{rs}(560)$ and $R_{rs}(413)/R_{rs}(443)$, denoted as R_{560}^{490} and R_{443}^{413}, were used to characterize the waters those have very high $a_t(443)$ and low $a_{CDM}(443)/a_t(443)$. In case 1 waters, R_{560}^{490} ratio increases with any decrease in chl-*a*. Similarly any increase in R_{443}^{413} can be interpreted as decrease in CDOM concentration in the water. Even though both R_{560}^{490} and R_{443}^{413} ratios are also sensitive to CDOM and chl-*a* relative concentration, the effect is minimal (Morel and Gentili, 2009). However, in case 2 waters the effect is considerable and therefore it is not feasible to use a single R_{rs} ratio to characterize either chl-*a* or CDM relative contribution. In our dataset, R_{560}^{490} ratio varies within 0.12 to 0.54 whereas R_{443}^{413} varies within 0.89 to 1.61. These values were compared with corresponding values in International Ocean Color Coordination Group (IOCCG) synthetic dataset and found out the range of R_{rs} ratios that best characterize the waters under study (Fig. 11). Data from this study formed a distinct cluster of points in the three dimensional space formed by R_{560}^{490}, R_{443}^{413}, and $a_{CDM}(\lambda)/a_t(443)$; where $a_t(443)$ is very

high and contribution of $a_{CDM}(443)$ to $a_t(443)$ is very low (~8-45%). The space could be represented by using thresholds for R_{560}^{490} and R_{443}^{413} ratios ($R_{560}^{490} < 0.54$ and $R_{443}^{413} > 0.89$). Based on the observations from this data, it is suggested that the proposed a_{CDM} empirical scheme should be used in low relative CDM waters when R_{560}^{490} is less than 0.54 and R_{443}^{413} is greater than 0.89.

It should also be noted that the above discussed empirical schemes to model $a_{CDM}(\lambda)$ were derived from a small dataset but with a wide range of optical parameters (n=20) and the relations can be re-parameterized using a large dataset for even better accuracy and predictive ability. Nevertheless, results show that a_{CDM} at 381, 443, and 510 nm, and S_{CDM} can be empirically modeled from $R_{rs}(\lambda)$ data in turbid and productive waters with low relative CDM contribution.

4.5.4.2 Validation of the first order empirical estimates of $a_{CDM}(\lambda)$

Modeled $a_{CDM\text{-}381}(\lambda)$ and $a_{CDM\text{-}443}(\lambda)$ values were compared with the measured $a_{CDM}(\lambda)$ values (Fig.12a and Fig. 13a, b, and c). For all sampling stations, the $\varepsilon(n)$ values varied within 11.36-133% with an average of 47.38% within 413-665 nm wavelength range. Out of 21 samples, errors of predictions for 5 samples were very high exceeding 84%. However, the maximum errors occurred at higher wavelengths starting from 560 nm where magnitude of a_{CDM} is less. Similarly, $\varepsilon(\lambda)$ varied from 38.85-79.9% with an average of ~52% within 413-665 nm wavelength range. The minimum and maximum ε occurred at 413 and 665 nm. The error at shorter wavelengths (413-510 nm) varied from 38.85 to 46.72%. It should also be noted that the high ε values for $a_{CDM\text{-}381}(\lambda)$ at longer wavelengths has very little effect on the decomposition of $a_t(\lambda)$ signal in that region

because of low magnitude of a_{CDM} at longer wavelengths. For example, an error of 80% in $a_{CDM}(665)$ at station-20 can yield a deviation of 0.192 m^{-1} from the measured value, which is only ~4% of a_t value at that wavelength [$a_t(665)$=4.43 m^{-1}]. Modeled $a_{CDM\text{-}443}(\lambda)$ values also showed magnitude of ε as similar to $a_{CDM\text{-}381}(\lambda)$. In the blue spectral region, $a_{CDM\text{-}381}(\lambda)$ produced comparatively less error (~1%) than $a_{CDM\text{-}443}(\lambda)$.

When first order estimates of $a_{CDM}(\lambda)$ were used, retrieval accuracy of $a_\varphi(\lambda)$ somewhat improved at least in this dataset. $a_\varphi(\lambda)$ were calculated by subtracting $a_{CDM\text{-}381}(\lambda)$ and $a_w(\lambda)$ from $a_t(\lambda)$. Two sets of $a_\varphi(\lambda)$ such as $a_{\varphi\text{-}381}(\lambda)$ and $a_{\varphi\text{-}443}(\lambda)$ were calculated by subtracting $a_{CDM\text{-}381}(\lambda)$ and $a_{CDM\text{-}443}(\lambda)$ from $a_t(\lambda)$ respectively. The modeled $a_\varphi(\lambda)$ values were compared with measured values within 413-665 nm spectral range. For all sampling stations, the $\varepsilon(n)$ varied within 7.8 -76% with an average of 21.46%. The maximum error occurred at station-2 (76%) and after excluding station-2, the error range reduced to 7.8-43.12% and the average error reduced to 18.73%. This implies that the retrieved values are in close agreement with the measured ones (Fig. 12b and Fig. 13d, e, and f). Similarly, the $\varepsilon(\lambda)$ varied from 16.94-34.15 % within 413-708 nm and the average error was 23.93%. Maximum and minimum error occurred at 413 and 665 nm respectively (Table 5). $a_{\varphi\text{-}443}(\lambda)$ errors were ~2% more than $a_{\varphi\text{-}381}(\lambda)$ in this study. Overall, both produced similar accuracy. Also, results showed that the spectral shape of modeled $a_\varphi(\lambda)$ matched very well with the measured $a_\varphi(\lambda)$ (Fig. 14). The maximum deviation was found at blue spectral region, 413 and 443 nm (~30%), and the deviation was less than 15% elsewhere within the visible region.

4.5.5 Significance and applicability of the model

Accurate estimation of $a_\varphi(\lambda)$ data from turbid productive waters is highly desirable to better understand the primary productivity and carbon uptake by algal biomass in a regional scale. To our knowledge, there is a lack of literature that focuses on retrieval of IOPs in highly turbid and productive waters. High turbidity and hyper-eutrophic condition is not just restricted to aquaculture ponds, rather it frequently occurs in inland lakes, estuaries, and marginal seas. Cyanobacterial algal blooms (CAB) often occur in inland lakes and estuaries and critically affect the annual carbon and nutrient cycle. CAB is a serious water quality problem in the Baltic Sea that covered approximately an area of 100,000 km^2 in 1994 and the blooming period lasted about a month (Kahru, 1997). CABs introduce new productions of nitrogen in the ecosystem through nitrogen fixation and further increase the trophic status of the system. Chl-*a* concentration during CABs often exceed 1000 mg m^{-3} in those hyper-eutrophic waters; e.g. during CAB in the Baltic Sea in summer 2002, chl-*a* concentrations were recorded as high as 1024 mg m^{-3} (Kutser, 2004). Hyper-eutrophic status of waters and associated algal blooms have also been reported in Patos Lagoon, Brazil (Yunes *et al*., 1996), Gulf of Finland (Bianchi *et al*., 2000), Lake Erie, USA (Vincent *et al*., 2004), Lake Pontchartrain, USA (Mishra and Mishra, 2010) and Taihu Lake, China (Hu *et al*., 2010). The inversion algorithm presented in this study has been calibrated using a dataset with very high chl-*a* (95-1376 mg m^{-3}) and TSS (69-401 mg l^{-1}) range. Therefore, it could be applied to retrieve $a_\varphi(\lambda)$ in turbid productive waters as well as CAB dominated waters in inland lakes, coastal estuarine environments, and marginal seas. In addition, spectral shape of $a_\varphi(\lambda)$ spectrum could be obtained using $R_{rs}(\lambda)$ data from hyperspectral sensors

such as Hyperion, the Hyperspectral Imager for the Coastal Ocean (*HICO*), and the future hyperspectral sensor, the Hyperspectral Infrared Imager (HyspIRI), that may provide more information on phytoplankton functional types.

4.6 Conclusion

A quasi-analytical inversion algorithm was parameterized to retrieve $a_\varphi(\lambda)$ spectrum from above surface remote sensing reflectance data in highly turbid and productive waters. $a_\varphi(\lambda)$ retrieval accuracy of the algorithm significantly improved after parameterization. Even though the algorithm underestimated $a_{CDM}(\lambda)$ values and produced a few negative retrievals, predicted $a_\varphi(\lambda)$ matched very well with the measured $a_\varphi(\lambda)$. An empirical scheme to retrieve $a_t(708)$ was also developed using a small dataset. Parameterization of the empirical algorithm using a large dataset acquired from turbid productive waters could potentially increase the retrieval accuracy. A supplementary method has also been provided for first order estimation of $a_{CDM}(\lambda)$ using novel empirical techniques. a_{CDM} at 381, 443, and 510 nm and S_{CDM} were empirically modeled from $r_{rs}(490)/r_{rs}(510)$ and used in the exponential model to derive $a_{CDM}(\lambda)$. Limitations of the empirical scheme for $a_{CDM}(\lambda)$ retrieval were discussed; and it was suggest that proposed a_{CDM} empirical scheme should only be used when $R_{560}^{490} < 0.54$ and $R_{443}^{413} > 0.89$) (in low relative a_{CDM} waters). The analytical decomposition steps from QAAv5 should be used elsewhere. Retrieval accuracy of $a_\varphi(\lambda)$ somewhat improved after incorporating the first order estimates of $a_{CDM}(\lambda)$ in the algorithm. In the present form, the algorithm can be applied to $R_{rs}(\lambda)$ data from multispectral sensor like MERIS and hyperspectral sensors to retrieve $a_\varphi(\lambda)$ in highly turbid and productive waters as well as in algal bloom (no scum forming) dominated hyper-eutrophic waters.

Table 4.1 Symbols and abbreviations

Symbol	Description	Unit
$a_p(\lambda)$	Absorption coefficient of particulate matter	m^{-1}
$a_d(\lambda)$	Absorption coefficient of detrital matter	m^{-1}
$A_{CDOM}(\lambda)$	Absorbance of colored dissolved organic matter	
$a_{CDOM}(\lambda)$	Absorption coefficient of colored dissolved organic matter	m^{-1}
$a_{CDM}(\lambda)$	Absorption coefficient of CDOM and detritus	m^{-1}
$a_w(\lambda)$	Absorption coefficient of pure sea water	m^{-1}
$a_\varphi(\lambda)$	Absorption coefficient of phytoplankton pigments	m^{-1}
$a_t(\lambda)$	Total absorption coefficients, $a_w(\lambda)+ a_\varphi(\lambda)+ a_{CDM}(\lambda)$	m^{-1}
$a_{t\text{-}w}(\lambda)$	$a_t(\lambda) - a_w(\lambda)$	m^{-1}
$b_{bp}(\lambda)$	Backscattering coefficients of particulate matter	m^{-1}
$b_{bw}(\lambda)$	Backscattering coefficients of pure sea water	m^{-1}
$b_b(\lambda)$	Backscattering coefficients of the total, $b_{bw}(\lambda)+b_{bp}(\lambda)$	m^{-1}
η	Spectral power for backscattering coefficient	
$R_{rs}(\lambda)$	Above-surface Remote sensing reflectance	sr^{-1}
$r_{rs}(\lambda)$	Subsurface Remote sensing reflectance	sr^{-1}
S_{CDM}	Spectral slope of CDOM and detritus absorption coefficient	nm^{-1}
TSS	Total suspended solids	mg l^{-1}
ISS	Inorganic suspended solids	mg l^{-1}
OSS	Organic suspended solids	mg l^{-1}
IOP	Inherent optical properties	
u	Ratio of backscattering coefficient to the sum of absorption and backscattering coefficient	

Table 4.2 Steps of the inversion model and their mathematical derivations

Steps	Property	Derivation
Step 0	r_{rs}	$= R_{rs}/(0.52 + 1.7R_{rs})$
Step 1	$u(\lambda) = \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}$	$= \frac{-0.0895+\sqrt{(g_0)^2+4g_1*r_{rs}(\lambda)}}{2*g_1}$; $g_0 = 0.089, g_1 = 0.125$
Step 2	$a(\lambda_0)$ where $\lambda_0 = 708\ nm$	$a(\lambda_0) = a_w(708) + 10^{-0.7153-2.054\chi-1.047\chi^2}$ $where\ \chi$ $= log_{10}\left(\frac{0.01 * r_{rs}(443) + r_{rs}(620)}{r_{rs}(708) + 0.005 * \frac{r_{rs}(620)}{r_{rs}(443)} * r_{rs}(620)}\right)$
Step 3	$b_{bp}(\lambda_0)$	$= \frac{u(\lambda_0)a(\lambda_0)}{1-u(\lambda_0)} - b_{bw}(\lambda_0)$
Step 4	η	$= 2.0\left\{1 - 1.2\ exp\left(-0.9\frac{r_{rs}(443)}{r_{rs}(555)}\right)\right\}$
Step 5	$b_{bp}(\lambda)$	$= b_{bp}(\lambda_0)\left(\frac{\lambda_0}{\lambda}\right)^{\eta}$
Step 6	$a(\lambda)$	$= \frac{(1-u(\lambda))\left(b_{bw}(\lambda) + b_{bp}(\lambda)\right)}{u(\lambda)}$
Step 7	$\zeta = \frac{a_\varphi(411)}{a_\varphi(443)}$	$= 0.74 + \frac{0.2}{0.8 + r_{rs}(443)/r_{rs}(555)}$
Step 8	$\xi = \frac{a_{CDM}(411)}{a_{CDM}(443)}$	$= e^{S(443-411)}$ $where\ s = 0.015 + \frac{0.002}{0.6 + r_{rs}(443)/r_{rs}(560)}$
Step 9	$a_{CDM}(443)$	$= \frac{[a(411) - \zeta a(443)] - [a_w(411) - \zeta a_w(411)]}{\xi - \zeta}$
Step 10	$a_{CDM}(\lambda)$	$= a_{CDM}(443)\ e^{-S(\lambda-443)}$
Step 11	$a_\phi(\lambda)$	$= a_t(\lambda) - a_w(\lambda) - a_{CDM}(\lambda)$

Table 4.3 Descriptive statistics of water quality parameters.

	Average	St. dev.	Minimum	Maximum	*n*
July 2010, Calibration Dataset					
Chl-*a* (µg l^{-1})	532.58	331.82	95.68	1376.57	20
Chl-*b* (µg l^{-1})	14.07	10.29	5.94	38.82	20
$a_{\varphi}(443)$ (m^{-1})	15.55	9.00	3.40	37.77	20
$a_{\varphi}(665)$ (m^{-1})	6.23	3.81	1.21	16.19	20
$a_{d}(443)$ (m^{-1})	2.90	2.07	0.79	8.01	20
$a_{CDM}(443)$ (m^{-1})	4.33	2.06	1.56	9.45	20
TSS (mg l^{-1})	177.33	80.20	69.80	401.20	20
OSS (mg l^{-1})	102.30	25.65	60.00	162.00	20
ISS (mg l^{-1})	85.41	65.40	14.60	247.20	20
Chl-*a*/TSS (mg/g)	3.22	1.84	0.81	7.70	20
ISS/TSS	0.41	0.16	0.16	0.68	20
OSS/TSS	0.59	0.16	0.32	0.84	20
April 2011, Validation Dataset					
$a_{\varphi}(443)$ (m^{-1})	10.27	2.87	6.43	18.04	21
$a_{\varphi}(665)$ (m^{-1})	3.96	1.32	2.08	6.73	21
$a_{d}(443)$ (m^{-1})	2.14	0.77	0.98	3.91	21
$a_{CDM}(443)$ (m^{-1})	3.20	0.85	1.93	5.19	21

'St. dev.' is the standard deviation, *n* is the total number of samples considered for the analysis.

Table 4.4 Percentage error, ε, between measured and model derived optical parameters for different samples

Sample	$a_t(\lambda)$	$a_\varphi(\lambda)$	$a_{CDM\text{-}381}(\lambda)$	$a_{CDM\text{-}443}(\lambda)$	$a_{\varphi\text{-}381}(\lambda)$	$a_{\varphi\text{-}443}(\lambda)$
1	6.32	29.28	41.92	40.74	15.30	15.54
2	57.62	31.92	42.60	44.04	76.09	76.84
3	10.10	18.66	111.59	109.82	11.04	11.56
4	7.69	16.16	23.06	22.32	14.07	14.08
5	21.57	14.58	33.25	33.00	34.80	34.98
6	27.88	57.36	88.68	88.07	20.76	20.44
7	13.11	26.57	84.32	83.56	13.40	14.18
8	9.92	27.11	133.07	131.56	17.44	17.68
9	9.57	15.73	22.79	24.48	14.03	14.27
10	14.52	26.92	44.54	45.09	11.53	11.56
11	22.32	55.05	48.78	47.94	22.24	21.98
12	6.55	19.32	61.68	62.03	16.30	16.78
13	12.22	23.85	15.97	16.17	13.06	13.08
14	9.90	11.85	23.61	23.74	26.83	28.17
15	7.93	24.31	28.99	30.95	7.81	7.82
16	21.46	38.29	22.99	23.58	22.04	21.89
17	12.08	9.30	16.21	17.71	14.55	14.85
18	12.62	7.09	15.38	17.02	15.24	15.45
19	6.72	14.21	11.36	13.04	8.44	8.56
20	41.34	26.71	104.18	109.16	43.17	44.66
21	28.82	52.46	20.17	21.00	32.61	32.36
Min	6.32	7.09	11.36	13.04	7.81	7.82
Max	57.62	57.36	133.07	131.56	76.09	76.84
Average	17.16	26.04	47.38	47.86	21.46	21.75

. $a_\varphi(\lambda)$ values were derived using QAA v5 with reference wavelength at 708 nm and parameterized $a_t(708)$. Other parameters were derived implementing first order estimates of $a_{CDM}(\lambda)$ in the QAA v5. Min, max, and average values of the retrieved parameters are also provided.

Table 4.5 Percentage error, ε, between measured and model derived optical parameters at different wavelengths

Wavelength (nm)	$a_t(\lambda)$	$a_\varphi(\lambda)$	$a_{CDM\text{-}381}(\lambda)$	$a_{CDM\text{-}443}(\lambda)$	$a_{\varphi\text{-}381}(\lambda)$	$a_{\varphi\text{-}443}(\lambda)$
413	21.73	40.10	38.86	40.37	34.15	36.03
443	24.13	41.27	42.03	43.11	23.66	23.72
490	21.57	31.58	45.14	45.43	22.49	22.55
510	21.43	22.74	46.72	46.72	26.97	26.97
560	18.10	18.70	51.91	51.31	25.43	25.10
619	16.96	20.42	59.78	59.50	17.93	17.95
665	15.22	15.90	79.91	80.11	16.94	16.88
Min	15.22	15.90	38.85	40.37	16.94	16.88
Max	24.13	41.27	79.91	80.11	34.15	36.03
Average	19.87	27.24	52.05	52.36	23.93	24.17

. $a_\varphi(\lambda)$ values were derived using QAA v5 with reference wavelength at 708 nm and parameterized $a_t(708)$. Other parameters were derived implementing first order estimates of $a_{CDM}(\lambda)$ in the QAA v5. Min, max, and average values are calculated within 413-665 nm wavelength range.

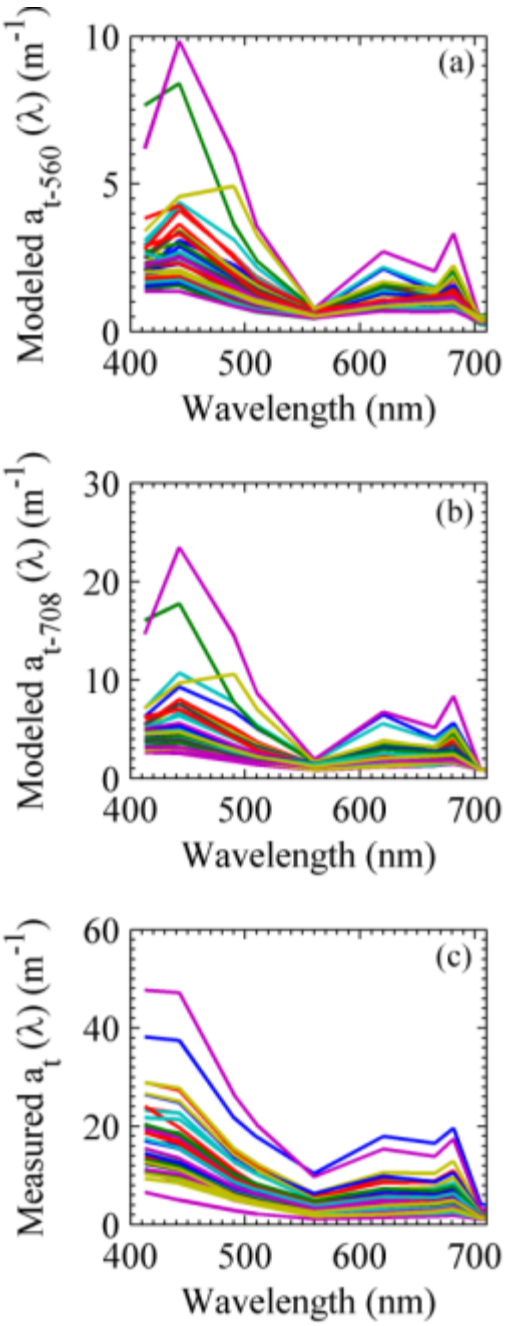

Figure 4.1 (a) QAA derived $a_t(\lambda)$ values after taking 560 nm as the reference wavelength, (b) QAA derived $a_t(\lambda)$ values after shifting the reference wavelength to 708 nm and assuming $a_t(708)$ is equal to $a_w(708)$, and (c) the measured $a_t(\lambda)$ values.

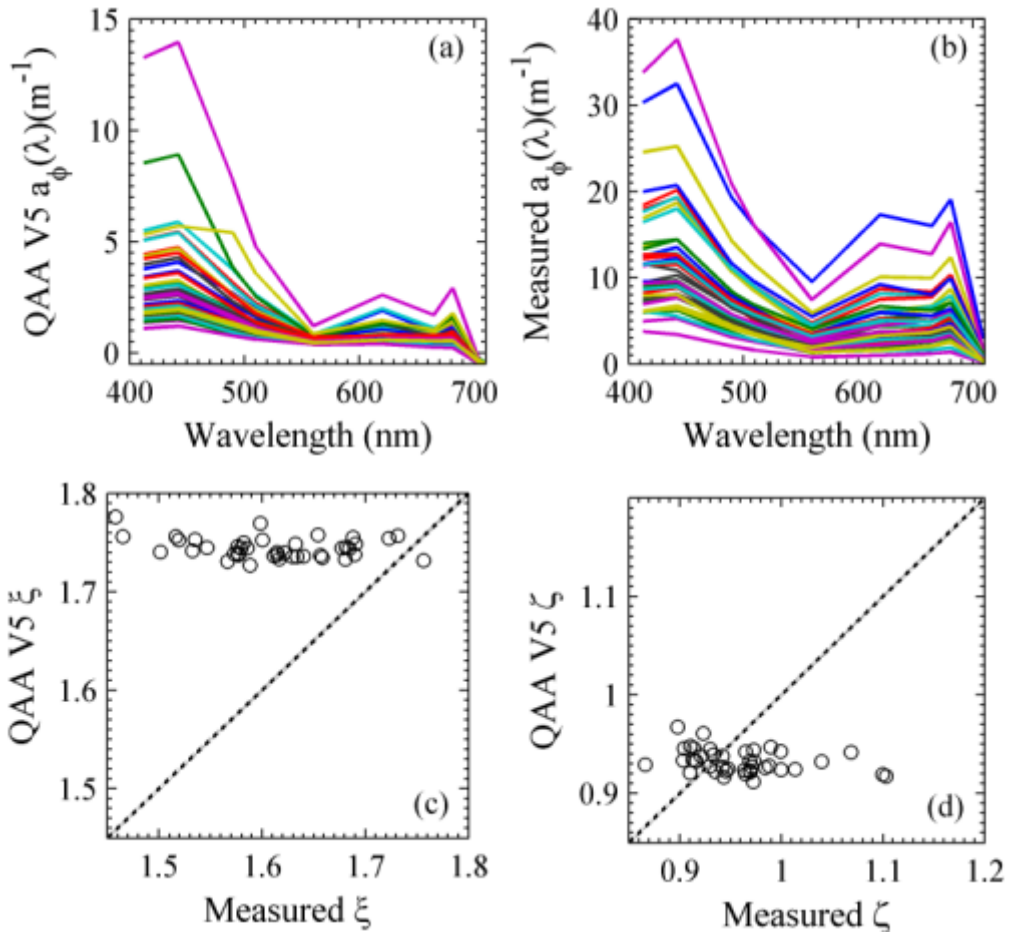

Figure 4.2 (Comparison of QAA v5 derived $a_\varphi(\lambda)$ (a) with pad-measured $a_\varphi(\lambda)$ (b), ξ (c) and ζ (d) with the corresponding measured values.

Dotted lines are 1/1 lines.

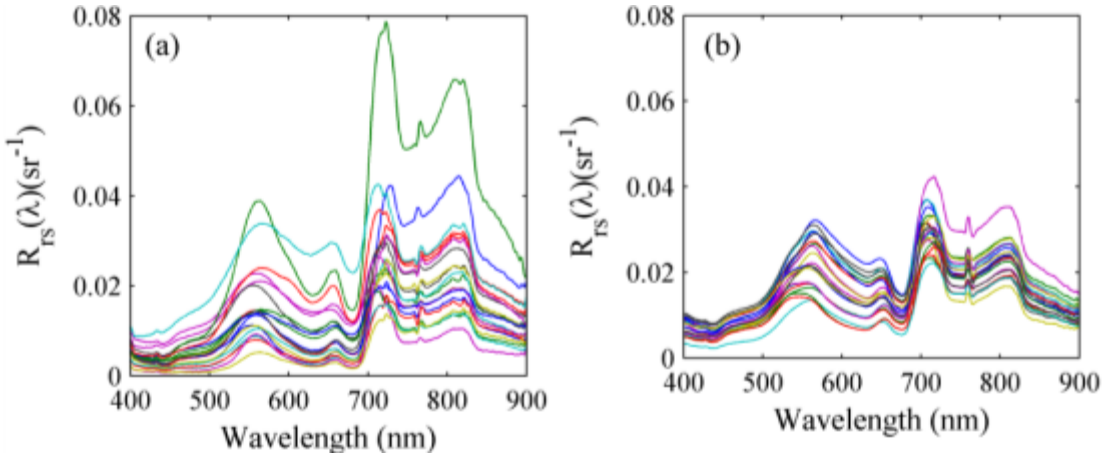

Figure 4.3 (Measured $R_{rs}(\lambda)$ spectra, (a) dataset collected in July 2010 (n=20), and (b) dataset collected in April 2011(n=21).

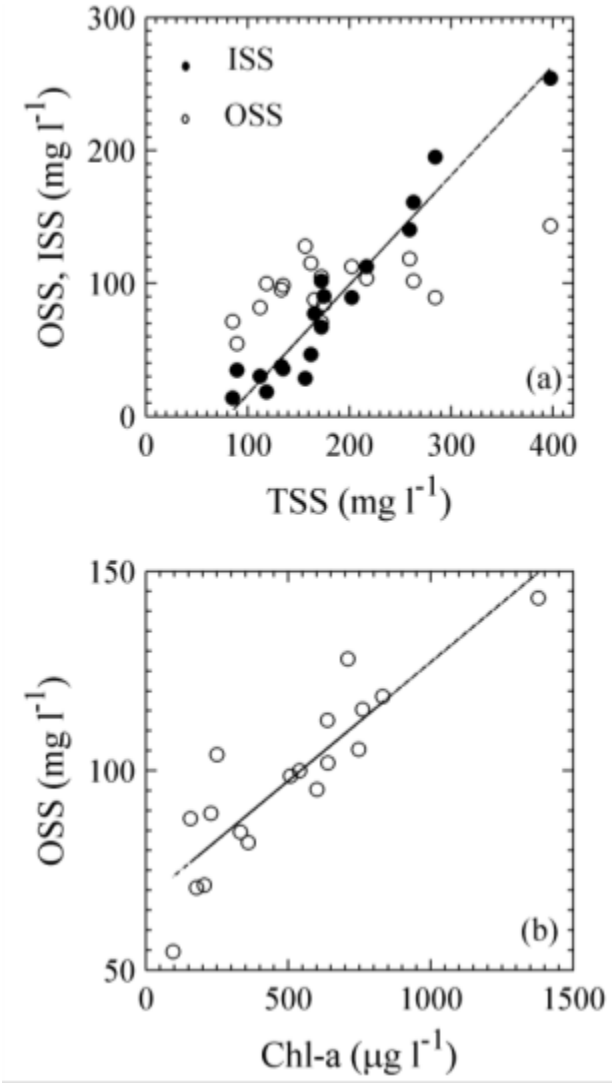

Figure 4.4 (a) Scatter plot between OSS, ISS and TSS in the dataset collected in July 2010.

Open circles: organic suspended solids; and filled circles: inorganic suspended solids. Linear least square fit shows the strong relationship between TSS and ISS (*ISS = 0.5402*TSS - 42.314,* R^2=0.93). (b) Scatter plot of OSS and chl-*a*. Linear least square fit shows the strong relationship between the two parameters (*OSS = 0.0413*chl-a + 44.41,* R^2 = 0.79).

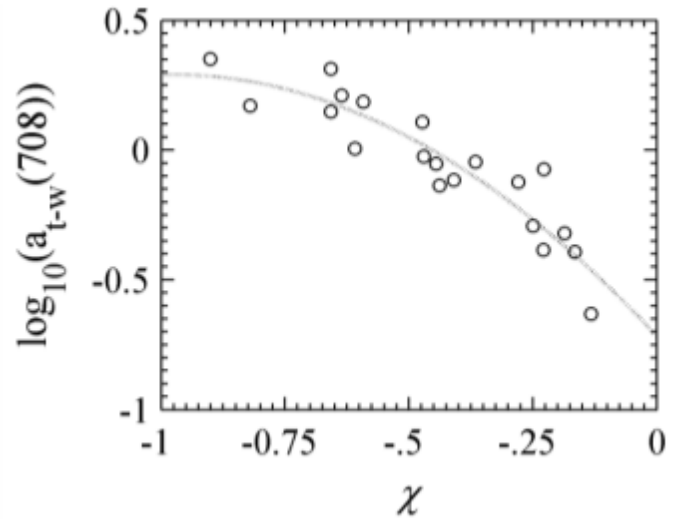

Figure 4.5 Calibration of $a_{t\text{-}w}(708)$ using the 2010 dataset.

Dotted line is the least-square fit second order polynomial trend line.

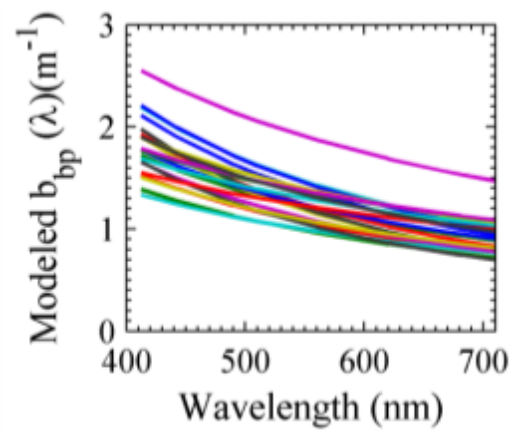

Figure 4.6 Modeled $b_{bp}(\lambda)$ spectra from $R_{rs}(\lambda)$

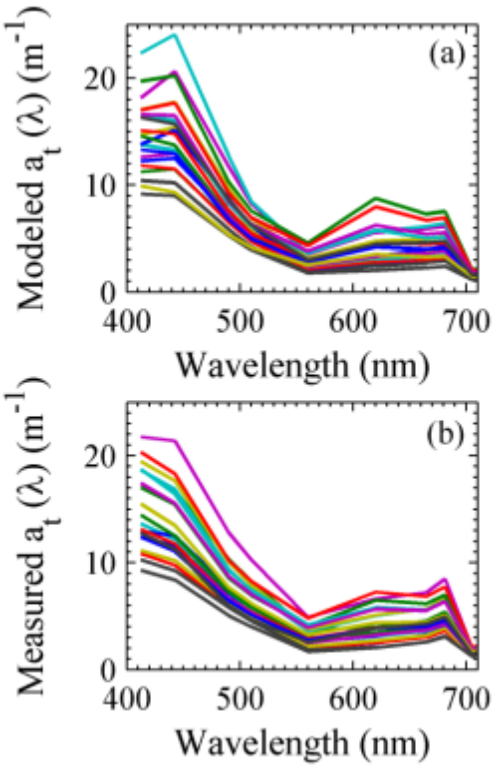

Figure 4.7 (a)Modeled $a_t(\lambda)$ from $R_{rs}(\lambda)$ data using the inversion model after parameterization, and (b) Measured $a_t(\lambda)$ using filtered-pad method.

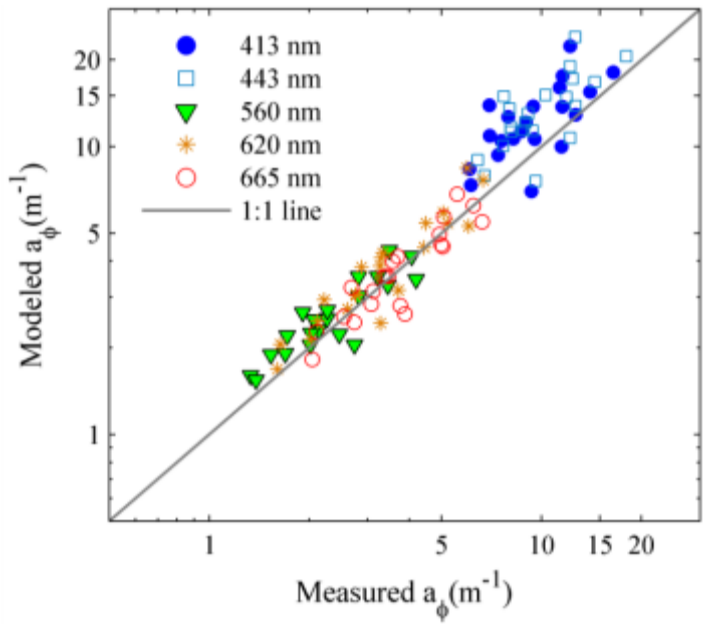

Figure 4.8 Scatter plot between measured $a_\varphi(\lambda)$ and modeled $a_\varphi(\lambda)$ at MERIS band centers.

The solid gray line is the 1:1 line.

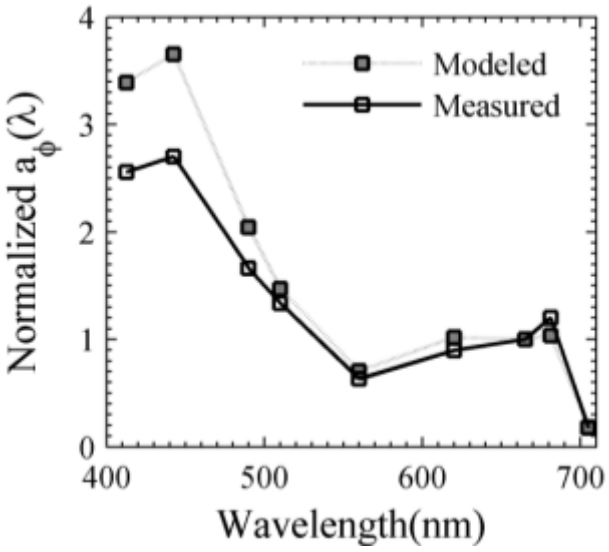

Figure 4.9 Comparison of spectral shape of measured and modeled $a_\varphi(\lambda)$ spectrum.

Spectra were normalized to $a_\varphi(665)$ and average of normalized spectra of all samples are plotted for comparison.

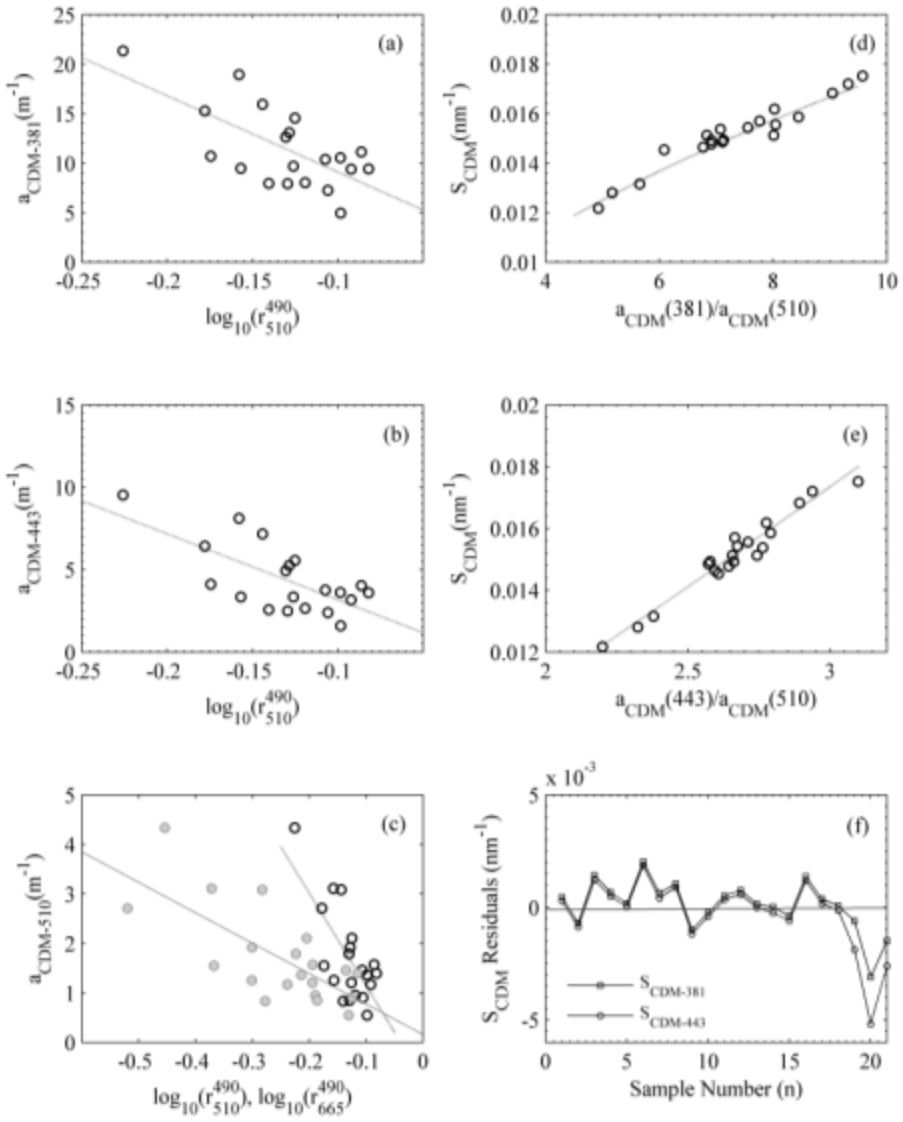

Figure 4.10 Calibration and validation of a_{CDM} and S_{CDM}. Empirical relationship between $\log_{10}(r_{rs}(490)/r_{rs}(510))$ ($log_{10}(r_{510}^{490})$ in the figure) and (a) $a_{CDM}(381)$, (b) $a_{CDM}(443)$, (c) $a_{CDM}(510)$.

Filled scatter points shows the empirical relationship between $\log_{10}(r_{rs}(490)/r_{rs}(665))$ ($log_{10}(r_{665}^{490})$ in the figure) and $a_{CDM}(510)$. Solid gray line is the least-square fit line. (d) Power fit between measured $a_{CDM}(381)/a_{CDM}(510)$ and S_{CDM}, (e) Power fit between measured $a_{CDM}(443)/a_{CDM}(510)$ and S_{CDM}. The dotted lines are the least-square fit lines and (f) S_{CDM} residuals from the validation analysis where $S_{CDM-381}$ and $S_{CDM-443}$ were modeled from corresponding a_{CDM} ratios derived from rrs data collected in April 2011.

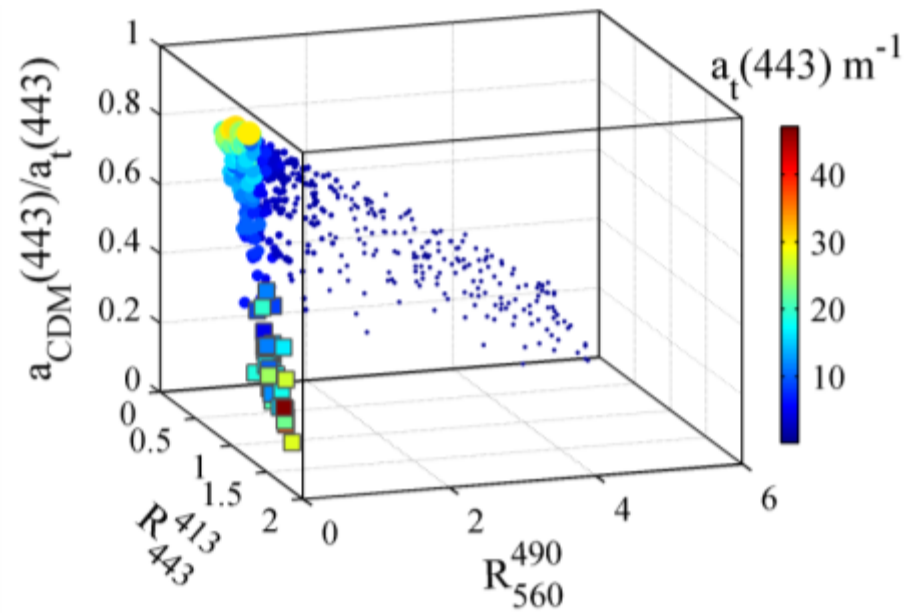

Figure 4.11 Scatter plot of $a_{CDM}(443)/a_t(443)$ as a function of $R_{rs}(490)/R_{rs}$ (560) and R_{rs} (413)/R_{rs} (443) denoted as R^{490}_{560} and R^{413}_{443} in the figure.

Data from this study (squares) are displayed in comparison with IOCCG synthetic dataset (circles). Size and color of the scatter points correspond to the magnitude of $a_t(443)$. Because of the wide range of $a_t(443)$ among all data, $a_t(443)$ in IOCCG data have been multiplied by 10 for better visual comparison. Note that the data points from this study fall in the three dimensional space where $a_t(443)$ is very high and $a_{CDM}(443)$ contribution to $a_t(443)$ is very low (~8-45%).

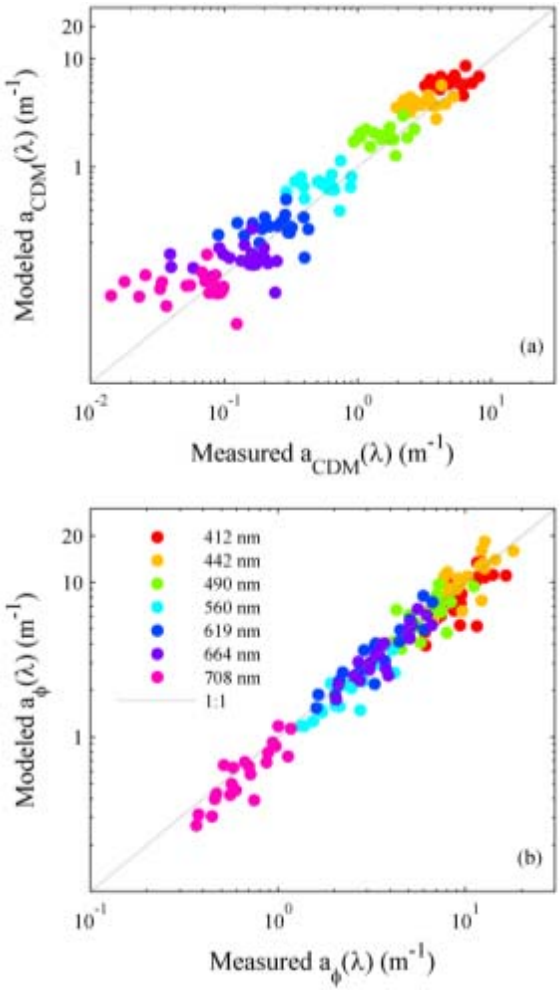

Figure 4.12 (a) Scatter plot between measured $a_{CDM}(\lambda)$ and modeled $a_{CDM}(\lambda)$ at MERIS band centers, and (b) Similar plot for $a_{\varphi}(\lambda)$.

The dotted line is the 1:1 line.

4.7 References

Arar, E.J. 1997. Determination of Chlorophyll a, b, c1 and c2, and Pheophytin a in marine and freshwater phytoplankton by High Permormance Liquid Cromatography. EPA ,(EPA.metod 447.0)

Babin, M., Stramski, D., Ferrari, G.M., Claustre, H., Bricaud, A., Obolensky, G. & Hoepffner, N. (2003). Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *Journal of Geophysical Research-Oceans*, *108*(C7), 3211, doi/10.1029/2001JC000882.

Bianchi, T. S., Engelhaupt, E., Westman, P., Andren, T., Rolff, C., & elmgren, R. (2000). Cyanobacterial blooms in the Baltic Sea/ Natural or human-induced?, *Limnology and Oceanography,* 45, 716-726.

Brewin, R.J.W., Sathyendranath, S., Hirata, T., Lavender, S.J., Barciela, R.M., & Hardman-Mountford, N.J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecological Modeling*, 221, 1472-1483.

Bricaud, A., Morel, A., & Prieur, L. (1981). Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains, *Limnology and Oceanography*, *26*, 671 43-53.

Buiteveld, H., Hakvoort, J.H.M., and Donze, M. (1994). The optical properties of pure water. *Ocean Optics XII*.

Carder, K. L., Chen, F. R., Lee, Z. P., Hawes, S. K. & Kamykowski, D. (1999). Semianalytic Moderate-Resolution Imaging Spectrometer algorithms for chl-*a* and absorption with bio-optical domains based on nitrate-depletion temperatures, *Journal of Geophysical Research*, 104, 5403–5421.

Carder, K.L., Hawes, S.K., Baker, K.A., Smith, R.C., Steward, R.G., & Mitchell, B.G. (1991). Reflectance model for quantifying chlorophyll a in the presence of productivity degradation products. *Journal of Geophysical Research*, 96, 20599-20611

Craig, S. E.; Lohrenz, S. E.; Lee, Z.; Mahoney, K. L.; Kirkpatrick, G. J.; Schofield, O. M. & Steward, R. G. (2006). Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga, Karenia brevis. *Applied Optics*, 45, 5414-5425.

Doron, M. , Babin, M., Mangin, A. &Fanton d'Andon, O. (2006). Estimation of light penetration, and horizontal and vertical visibility in oceanic and coastal waters from surface reflectance. *Journal of Geophysical Research*, 112, C06003, doi: 10.1029/2006JC004007.

D'Sa, E.J., Miller, R.L.,& Del Castillo, C. (2006). Bio-optical properties and ocean color algorithms for coastal waters influenced by the Mississippi River during a cold front. *Applied Optics*, 45:7410-7428.

Dall'Olmo, G. & Gitelson, A. A. (2006).Effect of bio-optical parameter variability and uncertainties in reflectance measurements on the remote estimation of chlorophyll-*a* concentration in turbid productive waters/ Modeling results, *Applied Optics*, 45, 3577–3592.

Falkowski, P.G., Katz, M.E., Knoll, A.H., Quigg, A., Raven, J.A., Schofield, O., & Taylor, F.J. (2004). The evolution of modern eukaryotic phytoplankton. *Science*, 305, 354-360.

Fargion, G. S. and J. L. Mueller. (2000). Ocean optics protocols for satellite ocean colour sensor validation. Revision 2," NASA Tech. Memo. 209966, *SeaWiFS Technical Report Server, (NASA Goddard Space Flight Center*).

Field, C.B., Behrenfeld, M.J., Randerson, J.T., & Falkowski, P. (1998). Primary production of the biosphere/ integrating terrestrial and oceanic components. *Science*, 281, 237-240.

Garver, S. A., & Siegel,D. (1997).Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation. 1. Time series from the Sargasso Sea, *Journal of Geophysical Research.* 102, 18607–18625.

Gons, H.J., Rijkeboer, M., Bagheri, S., & Ruddick, K.G. (2000). Optical teledetection of chlorophyll a in estuarine and coastal waters. *Environmental Science and Technology*, 34, 5189-5192

Gordon, H. R., and Morel, A. (1983). *Remote assessment of ocean color for interpretation of satellite visible imagery/ A review*, 44 pp., Springer-Verlag, New York.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., & Clark D. K. (1988). A Semianalytic Radiance Model of Ocean Color, Journal of Geophysical Research., 93(D9), 10,909–10,924, doi/10.1029/JD093iD09p10909.

Hu, C., Lee, Z.P., Ma, Ronghua, Yu, K., Li, D. (2010). Moderate resolution imaging spectroradiometer (MODIS).observation of cyanobacteria blooms in Taihu Lake, Chia, *Journal of Geophysical research*, 115, C04002, doi/10.1029/2009JC005511.

Kahru, M. (1997). Using satellites to monitor large-scale environmental change in the Baltic Sea, p. 43–61. *In* M. Kahru and C. W. Brown [eds.], Monitoring algal blooms/ New techniques for detecting large-scale environmental change. Springer-Verlag.

Kowalczuk, P., J. Olszewski, M. Darecki & S. Kaczmarek (2005)/ Empirical relationships between coloured dissolved organic matter (CDOM) absorption and apparent optical properties in Baltic Sea waters, *International Journal of Remote Sensing*, 26/2, 345-370

Kowalczuk, P., Stedmon, C., & Markager, A. S. (2006). Modelling absorption by CDOM in the Baltic Sea from season, salinity and chlorophyll. *Marine Chemistry*, 101, 1−11.

Kutser, T. (2004). Quantitative detection of chlorophyll in cyanobacterial blooms by the satellite remote sensing, *Limnology and Oceanography*, 49, 2179-2189.

Le, C.F., Li, Y.M., Zha, Y., Sun, D., & Yin, B. (2009). Validation of a Quasi-Analytical Algorithm for Highly Turbid Eutrophic Water of Meiliang Bay in Taihu Lake, China, *IEEE Transactions on Geosciences and Remote Sensing*, 47 (8), 2492-2500.

Lee, Z.P., Carder, K.L., R.A., and R. Arnone. (2002). Deriving inherent optical properties from water color/ A multi-band quasi-analytical algorithm for optically deep waters, *Applied Optics*, *41*, 5755-5772.

Lee,Z.P., Carder,K.L. (2004).Absorption spectrum of phytoplankton pigments derived from hyperspectral remote-sensing reflectance. *Remote Sensing of Environment* 89 p. 361–368

Mannino, A., Russ, M. E., & Hooker, S. B. (2008). Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the U.S. Middle Atlantic Bight, *Journal of Geophysical Research*, 113, C07051, doi:10.1029/2007JC004493.

Maritorena, S., Siegel, D.A., & Peterson, A.R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, 41, 2705-2714

Mishra, D. R. & Mishra, S. (2010). Plume and bloom/ effect of the Mississippi River diversion on the water quality of Lake Pontchartrain, *Geocarto International*, 25, 555-568.

Morel, A. &Gentili, B. (2009). A simple band ratio technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data. *Remote Sensing of Environement,* 113, 998-1011.

Morel, A. & Prieur, L. (1977).Analysis of variations in ocean color," *Limnololy and Oceanography*, 22, 709–722.

Ohde, T. & Siegel, H. (2003). Derivation of immersion factors for the hyperspectral Trios radiance sensor. *Journal of Optics a-Pure and Applied Optics*, 5*,* 12-14.

O'Reilly, J., Maritorena, S., Mitchell, B., Siegel, D., Carder, K., Garver, S., Kahru, M., & McClain, C. (1998).Ocean color chlorophyll algorithms for Sea-WiFS, *Journal of Geophysical Research,* 103, 24 937–24 953.

Pope, R. & Fry, E. (1997).Absorption spectrum (380 - 700 nm) of pure waters/ II. Integrating cavity measurements, *Applied Optics*, 36, 8710-8723.

Roesler, C. S., & Perry, M. J. (1995).In situ phytoplankton absorption, fluorescence emission, and particulate backscattering spectra determined from reflectance, *Journal of Geophysical Research*, 100, 13279–13294.

Saba, V.S., Friedrichs, M.A.M., Antoine, D., Armstrong, R.A., Asanuma, I., Aumont, O., Behrenfeld, M.J., Ciotti, A.M., Dowell, M., Hoepffner, N., Hyde, K.J.W., Ishizaka, J., Kameda, T., Marra, J., Mélin, F., Moore, J.K., Morel, A., O'Reilly, J., Scardi, M., Smith Jr., W.O., Smyth, T.J., Tang, S., Uitz, J., Waters, K., Westberry, T.K. (2011). An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. *Biogeosciences*, 8, 489-503.

Schwarz, J. N., Kowalczuk, P., Cota, G. F., Mitchell, B. G., Kahru, M., Chavez, F. P., Cunningham, A., McKee, D., Gege, P., Kishino, M., Phinney, D. A., & Raine, R. (2002). Two models for absorption by colored dissolved organic matter (CDOM). *Oceanologia*, 44, 209−241.

Shanmugam, P. (2011). New models for retrieving and partitioning coloured dissolved organic matter in the global oceans. Implications for remote sensing. *Remote Sensing of Environment*, 115, 1501–1521.

Tassan, S. & Ferrari, G.M. (2003). Variability of Light Absorption by Aquatic Particles in the Near-Infrared Spectral Region. *Applied Optics*, 42, 4802-4810.

Tucker, C.S., & Boyd, C.E (1985). Watr quality. *In* Channel catfish culture. Ed. C.S. Tucker. Elsevier Science Publishers. Amsterdam. Pp. 135-227.

Tucker, C. S., Kingsbury, S. K., & Mischke, C. C. (2009): Bacterial Bioaugmentation of Channel Catfish Ponds, North American Journal of Aquaculture, 71:4, 315-319

Twardowski, M.S., Boss, E., Sullivan, J.M. & Donaghay, P.L. (2004). Modeling the spectral shape of absorption by chromophoric dissolved organic matter. *Marine Chemistry*, 89, 69- 88.

Yunes, J.S., Salomon, P.S., Matthiensen, A., Beattie, K.A., Raggett, S.L., & Codd, G.A. (1996). Toxic blooms of cyanobacteria in the Patos Lagoon Estuary, southern Brazil, *Journal of Aquatic Ecosystem Health*, 5, 223-229.

Vincent, R.K., Qin, X.M., McKay, R.M.L., Miner, J., Czajkowski, K., Savino, J., & Bridgeman, T. (2004). Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment,* 89, 381-392.

Zhu, W., Yu, Q., Tian, Y. Q., Chen, R. F., & Gardner, G. B. (2011). Estimation of chromophoric dissolved organic matter in the Mississippi and Atchafalaya river plume regions using above-surface hyperspectral remote sensing, *Journal of Geophysical Research,* 116, C02011, doi:10.1029/2010JC006523.

CHAPTER V

A QUASI-ANALYTICAL ALGORITHM TO QUANTIFY CYANOBACTERIAL PHYCOCYANIN: MODEL CALIBRATION AND VALIDATION

5.1 Overview

In this research, a novel technique to monitor cyanobacterial algal bloom (CHAB) using remote sensing reflectance measurements is presented. A multi-band quasi analytical algorithm that determines phytoplankton absorption coefficients, $a_\varphi(\lambda)$, from above surface remote sensing reflectance, $R_{rs}(\lambda)$ measurements using inversion method was used. *In situ* data including remote sensing reflectance, phytoplankton pigment concentration, and absorption coefficients of optically active constituents in the water were collected from highly turbid and productive aquaculture ponds. A novel technique was developed to further decompose the a_φ to obtain phycocyanin absorption coefficient, a_{PC} at 620 nm, a primary peak of phycocyanin absorption spectrum. Validation of the model produced a median percentage relative error of ~22%. Strong empirical relationship between phycocyanin absorption coefficients at 620 nm and measured phycocyanin concentrations (R^2=0.96, p<0.0001) also shows its potential for empirical retrieval of PC in optically complex waters. Results demonstrate that the new approach will be suitable for quantifying phycocyanin concentration in cyanobacteria dominated turbid productive waters. Band architecture of the model matches with the band configuration of the Medium Resolution Imaging Spectrometer (MERIS), European

space born remote sensor, and assures that MERIS reflectance products can be used to quantify cyanobacterial phycocyanin in optically complex waters.

5.2 Introduction

Cyanobacteria, also known as blue-green algae, dominate the phytoplankton community in inland lakes, estuaries and coastal waters during summer months, form algal blooms, and degrade the water quality. Many genera of cyanobacteria are known to produce toxins such as: neurotoxins (anatoxin-a), hepatotoxins (microcistin), and cytotoxins (cylindrospermopsins) (Metcalf *et al.*, 2008, Boyer, 2008). In addition to toxicity, many species of cyanobacteria produce odorous compounds such as geosmin and 2-methylisoborneol (MIB) that cause "earthy-muddy" and "musty" odor in drinking water, which is also a serious issue in aquaculture industry. Fish stock absorbs the musty odor causing compound which eventually cause severe economic loss either because of increasing the production cost or by decreasing the demand (Hanson, 2003). Because of the potential health risk and economic loss to fishery, aquaculture, and tourism industry, cost-effective monitoring solutions should be developed for early warning in the affected regions. Remote sensing is a more viable option for this kind of environmental monitoring because of low cost, synoptic, and temporally repetitive nature of monitoring capabilities. Optical properties of phycocyanin (PC), the characteristic cyanobacterial photosynthetic pigment, in visible and near-infrared (NIR) wavelength range can be used to develop algorithms to detect and quantify cyanobacterial biomass in natural waters. PC has very distinct absorption characteristics (absorption maximum at 620 nm) which is prominent in the reflectance spectra acquired from cyanobacteria dominated water bodies (Glazer 1989; Richardson 1996). When the cyanobacterial biomass dominates the water

body, the reflectance spectrum shows an enhanced absorption feature at around 600-625 nm and a reflectance maximum at 650 nm (Mishra *et al*., 2009; Kutser *et al.*, 2006; Dekker *et al.*, 1993) which can be used to fingerprint cyanobacteria in remotely sensed data.

Most research pertaining to the detection and mapping of cyanobacteria from *in situ* remote sensing spectra have used the absorption and reflectance features from 620 and 650 nm to develop relationships between R_{rs} and PC concentrations. To date, four broad types of algorithms have been proposed to quantify PC based on its absorption feature at 620 nm: 1) single band ratio empirical algorithm (Mishra *et al*., 2009; Schalles and Yacobi, 2000); 2) semi-empirical baseline algorithm (Dekker, 1993); 3) multiple band linear regression algorithm (Vincent *et al.*, 2004); and 4) a nested semi-analytical band ratio algorithm (Simis *et al.*, 2005). Ruiz-Verdu *et al.* (2008) evaluated the performances of all existing algorithms and reported that different algorithms perform differently based on the strengths and weaknesses of the algorithms, such as, ease of use, spectral band availability, and other methodological biases (Ruiz-Verdu *et al.*, 2008).

It is known that chlorophyll-a (chl-*a*) and other accessory photosynthetic pigments act as confounding constituents while retrieving PC concentration from R_{rs} data (Mishra *et al*., 2009; Simis *et al*., 2005). In addition to PC, chl-*a* and other accessory pigments also absorb light at the PC absorption maximum (around 620 nm) and thus affect the accuracy of PC prediction algorithms. In order to minimize the chl-*a* influence, Simis *et al*. (2005) semi-analytically subtracted the chl-*a* absorption component from the phytoplankton absorption at 620 nm, $a_{\varphi}(620)$. It should be noted that Simis *et al*. (2005) used two empirically derived coefficients, relating model derived $a_{\varphi}(665)$ and $a_{\varphi}(620)$

with measured ones, to estimate the proxies for $a_{\varphi}(665)$ and $a_{\varphi}(620)$ from model derived non-water absorption coefficients, $a_{t\text{-}w}(665)$ and $a_{t\text{-}w}(620)$. If the relative proportion of a_{φ} and absorption by colored detrital matter (colored dissolved organic matter +detritus), a_{CDM} varies outside the optimized range, the model could estimate erroneous PC values especially in turbid waters. In addition, the model does not consider $a_{pc}(665)$ before subtracting the chl-*a* contribution at 620 nm which may also increase the estimation error. Similarly, Mishra *et al*. (2009) suggested to use the spectral information at 600 nm in simple band ratio algorithms even though the PC absorption maximum is centered around 620 nm, because at 600 nm chl-*a* has comparatively less absorption than at 620 nm. However, 600 nm could also be contaminated by $a_{CDM}(600)$ and add uncertainties to PC estimation in highly turbid and productive waters.

In this study, a conceptual quasi analytical algorithm was developed to retrieve cyanobacterial PC concentration in turbid and productive waters. The specific objectives of this research are: 1) develop the conceptual algorithm to retrieve PC concentration in cyanobacteria, and 2) optimize the model parameters using optical and pigment measurements from turbid productive waters dominated by cyanobacterial biomass, and validate the algorithm for accuracy assessment. Two field campaigns were carried out at Delta Research Extension Center aquaculture ponds, Mississippi State University research facility located near Stoneville, MS, USA during 13-16 July, 2010 and 28-29 April, 2011. These aquaculture ponds are shallow water systems with surface area ranging from 0.4-1.62 ha and average depth of 1.0 m and have no watersheds. The aquaculture ponds were sampled for this study because of their eutrophic to hypereutrophic status (chl-*a* concentration exceeding 1000 mg m^{-3} in summer months)

with cyanobacteria dominating the phytoplankton community during summer and occasionally during spring months (Tucker and Boyd, 1985).

5.3 Development of the conceptual algorithm

5.3.1 The inversion method

The quasi analytical algorithm (QAA) (Lee *et al*., 2002) is a multi band inversion algorithm that inverts total absorption coefficients, $a_t(\lambda)$, and particulate backscattering coefficients, $b_{bp}(\lambda)$, from $R_{rs}(\lambda)$ measurements. Unlike other semi analytical algorithms (e.g., Hoge and Lyon, 1996; Maritorena et al., 2002), QAA retrieves $a_t(\lambda)$ first and then decomposes it into individual absorption components. In addition, it does not require spectral models for $a_\varphi(\lambda)$ and retrieves it independently from $R_{rs}(\lambda)$. The model has been extensively validated using simulated and field datasets from different geographic regions (Lee *et al*., 2002; Lee *et al*., 2004; Le *et al*., 2009, Craig *et al*., 2006; Zhu *et al*., 2010) and shown that the model was able to retrieve $a_t(\lambda)$ with a percentage difference of <20% between the measured and modeled data within 413-665 nm range. Recently the algorithm was parameterized to retrieve $a_\varphi(\lambda)$ in extremely turbid and cyanobacteria dominated hypereutrophic waters, where, $a_\varphi(443)$ (3.44-37.67 m^{-1}) contributes >54% of the $a_t(443)$ (4.99-47.21 m^{-1}) (Mishra *et al*., 2012). The model was able to retrieve $a_\varphi(\lambda)$ with an average percentage error of 27.24% within 413-665 nm range (Mishra *et al*., 2012). A brief description of the parameterized QAA algorithm including the mathematical steps is provided in Appendix A.

5.3.2 Decomposition of $a_\varphi(620)$ for PC retrieval

In this research, the QAA derived $a_\varphi(\lambda)$ was further decomposed to retrieve PC absorption at 620 nm. $a_\varphi(\lambda)$ provides the information about the absorption by phytoplankton pigments as a whole. It should be noted that absorption by all intracellular, both major and accessory pigments contribute to the $a_\varphi(\lambda)$ spectrum. The proposed decomposition method is based on the simple assumption that $a_\varphi(620)$ and $a_\varphi(665)$ are approximately equal to the sum of chl-*a* and PC contribution at those wavelengths. In another words, contribution of pigment absorption other than chl-*a* and PC to the $a_\varphi(620)$ and $a_\varphi(665)$ was considered negligible.

$$a_\varphi(665) = a_{chl}(665) + a_{pc}(665) \tag{5.1}$$

$$a_\varphi(620) = a_{chl}(620) + a_{pc}(620) \tag{5.2}$$

where, $a_{chl}(\lambda)$ and $a_{pc}(\lambda)$ are the absorption coefficients of chl-*a* and PC respectively. These two simple algebraic equations can be solved to retrieve $\mathrm{a_{pc}}(620)$ as:

$$a_{PC}(620) = \frac{\psi_1 \mathrm{a}_\varphi(620) - \mathrm{a}_\varphi(665)}{\psi_1 - \psi_2}, \tag{5.3}$$

where, $\psi_1 = \mathrm{a_{chl}}(665)/\mathrm{a_{chl}}(620)$ and $\psi_2 = \mathrm{a_{PC}}(665)/\mathrm{a_{PC}}(620)$ and their values can be empirically retrieved from R_{rs} band ratios. Further, if the specific absorption coefficient of PC at 620 nm, $\mathrm{a^*_{pc}}(620)$, is known, concentration of PC can be estimated as:

$$\mathrm{PC\ (mg\ m^{-3})} = \frac{\mathrm{a_{pc}}(620)(\mathrm{m^{-1}})}{\mathrm{a^*_{pc}}(620)(\mathrm{m^2 mg^{-1}})} \tag{5.4}$$

However, a^*_{pc} significantly varies with species type, season, and geographic region. Any uncertainties associated with $a^*_{pc}(620)$ estimation will introduce error in

predicting PC concentration. Because of the difficulties associated with retrieving $a_{pc}^{*}(620)$ sample by sample and especially its inapplicability in remote sensing research, an empirical approach has been used to retrieve $a_{pc}^{*}(620)$ from a reflectance band ratio.

The major advantage of this algorithm is that unlike the existing semi analytical method (Simis *et al*., 2005) this method does not neglect the $a_{CDM}(620)$. It also does not assume the $b_{bp}(\lambda)$ to be spectrally neutral. In addition, it considers the PC absorption at 665 nm and incorporates that information to algebraically retrieve$a_{PC}(620)$.

5.4 Data and methods

5.4.1 Remote sensing reflectance (R_{rs})

A dual sensor-system with two inter-calibrated Ocean Optics spectroradiometers was used to collect remote sensing reflectance data in the range 400-900 nm with a sampling interval of 0.3 nm as in DallO'lmo, *et al*., (2005). Radiometer 1, equipped with a 25° field-of-view optical fiber measured the upwelling radiance just below the air water interface, expressed in digital numbers as $DN_{Lu}(\lambda)$; whereas, radiometer 2, equipped with an optical fiber and cosine diffuser (yielding a hemispherical field of view) acquired above surface down welling irradiance, expressed in digital numbers as $DN_{Ed}(\lambda)$. To match their transfer functions, inter-calibration of the radiometers was accomplished by measuring the upwelling radiance of a white Spectralon reflectance standard (Labsphere, Inc., North Sutton, NH) simultaneously with incident irradiance. The two radiometers were inter-calibrated immediately before and after measurements in each field site. After the data acquisition, R_{rs} was calculated as follows:

$$R_{rs}(\lambda) = \frac{t}{n^2}\frac{DN_{Lu}(\lambda)}{DN_{Ed}(\lambda)}\frac{DN_{Ed,ref}(\lambda)}{DN_{Lu,ref}(\lambda)}\frac{\rho_{ref}(\lambda)}{\pi}F_i(\lambda) \qquad (5.5)$$

Where, t is the transmittance at the air-water interface (0.98); n is the refractive index of water (1.33); $DN_{Lu,ref}$ and $DN_{Ed,ref}$ are digital numbers representing upwelling radiance and downwelling irradiance over the white Spectralon panel; ρ_{ref} is the irradiance reflectance of the Spectralon panel; F_i (λ) is the spectral immersion factor (Ohde and Siegel, 2003). For each station 6 consecutive scans were recorded and further averaged to calculate a representative R_{rs}(λ) spectrum (Fig. 5.1a).

5.4.2 Water quality parameters

5.4.2.1 Chlorophylls

Water samples for chl-*a* and chl-*b* analysis were simultaneously collected in 1L Niskin bottles and immediately filtered onto GF/F filters (Whatman, 0.7 µm pore size) under low vacuum (<5 inch of Mercury). Sample were extracted in triplicates using acetone extraction procedure and concentrations were measured using HPLC as in Environmental Protection Agency method 447 (Arar, 1997).

5.4.2.2 Phycocyanin

Water samples for PC analysis were filtered immediately after collection through a 0.2 µm nucleopore membrane filters (Milipore) under low vacuum. Filters were placed into a 15 mL falcon tube then frozen at -80^{o}C until analysis. Prior to analysis, filters were transferred to a 50 mL polycarbonate centrifuge tubes, allowed to reach ambient room temperature, and then suspended in 5 mL of 50 mM phosphate buffer. Samples were homogenized as in Sarada *et al.* (2009) using a sonicator. Tip of the sonicator was rinsed twice with 5 mL of 50 mM phosphate buffer each time and collected in the centrifuge tube. Samples were centrifuged at 5^{o} C, 27,200g for 25 minutes. Samples were again

homogenized and the tip of the sonicator was rinsed with 5 mL of buffer and collected in the centrifuge tube and again centrifuged in the same settings. Finally, supernatant was collected and absorbance was measured using a Perkin Elmer lambda 850 spectrophotometer (Perkin Elmer Inc., Waltham, MA, USA) (Fig. 5.2). Concentration of PC was calculated using the equation from Bennett and Bogorad (1973).

5.4.2.3 Measurement of absorption coefficients

Surface water samples were collected in 1 L Niskin bottles and immediately filtered onto 0.7 µm Whatman GF/F filters under low vacuum (<5 inch of mercury). The volume of water filtered varied from 50-100 mL depending on the load of particulate matter in the sample. Particulate absorption coefficient, $a_p(\lambda)$, and absorption coefficient of detrital matter, $a_d(\lambda)$, were determined using standard quantitative filtration technique (QFT) as described in Fargion and Muller (2000). A Perkin Elmer lambda-850 spectrophotometer with an integrating sphere was used to measure absorbance of the samples within a spectral range from 400 to 800 nm. Further, $a_\varphi(\lambda)$ was computed by subtracting $a_d(\lambda)$ from $a_p(\lambda)$. Finally, $a_\varphi(\lambda)$ was corrected for residual scattering by subtracting $a_\varphi(800)$ from all wavelengths (Tassan and Ferrari, 2003). Water samples for colored dissolved organic matter (CDOM) analysis were filtered immediately after collection through a 0.2 µm nucleopore membrane filters under low vacuum. A detailed description of the methods of absorption measurements from the aquaculture ponds were provided in Mishra *et al*. (2012). Absorbance of CDOM, $A_{CDOM}(\lambda)$ was measured following the standard protocol (Fargion and Muller, 2000) as described in Mishra *et al* (2012). The $a_{CDOM}(\lambda)$ (m^{-1}) for path length, l (m^{-1}) was calculated as:

$$a_{CDOM}(\lambda) = \frac{2.303[A_{CDOM}(\lambda)]}{l} \quad (5.6)$$

5.4.3 Error analysis

PC concentration varied within about 2-3 order of magnitude in the current dataset. Therefore, it is more meaningful to assess the uncertainty of model retrieved PC by providing mean and median relative error (%) instead of root mean square error (RMSE). The errors were calculated as provided in Campbell and O'Reilly (2006).

Mean error or bias is calculated as:

$$bias(m) = \frac{1}{N}\sum_{i=1}^{N}(log\hat{C}_i - logC_i) \quad (5.7)$$

Where, $\hat{C}_i$ is model derived parameter and C_i is the measured parameter.

Mean Squared Error (MSE) is calculated as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(log\hat{C}_i - logC_i)^2 \quad (5.8)$$

Root Mean Squared Error (RMSE) is then calculated as:

$$RMSE = \sqrt[2]{MSE} \quad (5.9)$$

$$s = \sqrt{\frac{N(RMSE^2 - m^2)}{N-1}} \quad (5.10)$$

Assuming normality of log error, $log\hat{C}_i - logC_i$, the relative error statistics can be estimated as:

$$Mean\ Relative\ Error = \left\{exp\left(M + \frac{S^2}{2}\right) - 1\right\} \times 100 \quad (5.11)$$

$$Median\ Relative\ Error = \{exp(M) - 1\} \times 100 \quad (5.12)$$

$$Std\ dev\ of\ Relative\ Error = \left\{Mean\sqrt{\exp(S^2)}\right\} \times 100 \qquad (5.13)$$

where $M = m\ ln(10)$ and $S = s\ ln(10)$ are the mean and standard deviation of $ln(\hat{C}/C)$. Slope and R^2 of least-square fit between measured and model retrieved parameters were also used to report the consistency of the algorithm retrievals whenever needed.

5.5 Results and discussions

5.5.1 Water quality parameters

Analysis of water samples from the aquaculture ponds showed a wide range of pigment concentration. PC concentration varied from 68.13 to 3032.47 mg m^{-3} with an average value of 418.76 mg m^{-3}. The high values indicate the abundance of PC containing cyanobacteria in the waters during the sampling period. Chl-*a* concentration varied within 59.4-1376.6 mg m^{-3} with an average of 302.06 mg m^{-3}. PC: chl-*a* ratio varied within 0.3-3.29 (mean=1.23) indicating cyanobacterial dominance in the phytoplankton community structure (Table 5.1). Strong dependence was found between chl-*a* and PC concentration in the entire dataset (r=0.91). Similarly, a strong positive correlation was found between chl-*a* and PC in the 2010 dataset (r=0.94, Fig. 5.3) further corroborating the fact that algal community was mostly dominated by cyanobacterial biomass. However, a weak relationship was found between those pigments in the 2011 dataset indicating a mixed algal community. Chl-*b* concentration varied from 1.57-13.71 with a mean of 4.36 mg m-3. Strong positive correlation was found between chl-*a* and chl-*b* (r=0.79).

Absorption measurements show that the water samples were collected from highly turbid and productive waters where $a_\varphi(443)$(4.24-37.67 m^{-1}) contributes >59% towards the $a_t(443)$ (8.85-47.21 m^{-1}) (Table 5.1). $a_{CDM}(620)$ varied within 0.09-1.12 m^{-1} with a mean value of 0.32 m^{-1} supporting the fact that its effects cannot be assumed negligible in semi- and quasi analytical algorithms to retrieve PC in turbid and productive waters. In our dataset, $a_{CDM}(620)$ was ~3-11% of $a_\varphi(620)$. An uncertainty of this magnitude may introduce considerable amount of error in the analytical estimation of PC especially in waters with low PC concentration. Mean relative contribution of a_{CDM} to $a_t(665)$ and $a_t(708)$ was found to be 3.4 and 3.8% respectively.

5.5.2 Retrieval of $a_\varphi(\lambda)$ spectrum from $R_{rs}(\lambda)$

QAA retrieved $a_\varphi(\lambda)$ values were compared with filter-pad measured values for accuracy assessment (Fig. 4). In this research, the retrieval accuracy of $a_\varphi(620)$ and $a_\varphi(665)$ was shown because the proposed conceptual model uses retrieved a_φ values only at those band centers. A detailed parameterization and validation of the QAA to retrieve $a_\varphi(\lambda)$ in highly turbid and productive water can be found in Mishra *et al*. (2012). Mean relative error (%) for $a_\varphi(620)$ and $a_\varphi(665)$ of all samples was 16.15 and 14.72% respectively; whereas the median relative errors of $a_\varphi(620)$ and $a_\varphi(665)$ retrieval were 11.6 and 12.74% respectively. One samples showed relatively high error for both $a_\varphi(620)$ (~60%) and $a_\varphi(665)$ (30%). Overall, the relative error of retrieved $a_\varphi(620)$ and $a_\varphi(665)$ was about 30%.

5.5.3 Retrieval of $a_{PC}(620)$ from $a_{\varphi}(620)$

5.5.3.1 Using ψ_1 and ψ_2 measured from experimental data

Newly proposed method (Eq. 5.3) was used to retrieve $a_{PC}(620)$ from $a_{\varphi}(620)$ and $a_{\varphi}(665)$. It should be noted that the biggest challenge of this method was to tune the values of ψ_1 and ψ_2 from R_{rs} band ratios because they vary considerably due to package effect and therefore cannot be considered as constants. Value of ψ_1, $[a_{chl}(665)/a_{chl}(620)]$, and ψ_2, $[a_{PC}(665)/a_{PC}(620)]$, were measured *in vitro* after extracting pigments and measuring their absorption coefficients in a spectrophotometer. Chl-*a* was extracted in acetone from 7 samples in 2010 dataset and the data were used to calculate ψ_1, whereas, the same absorption data those were used to quantify PC concentration was also used to calculate ψ_2.

Strong dependence (power fit) was found between ψ_1 and chl-*a* concentration in this dataset (R^2=0.93, Fig. 5.5a). . After including ψ_1 values for the lower range of chl-*a* concentration from literature (Bidigare *et al*., 1990; Bricaud *et al*., 2004), R^2 of the power fit increased to 0.96 indicating its validity even in the lower chl-*a* (5-50 mg m^{-3}) range (relative to very high concentration in this study) (Fig. 5.5a). Considering $r_{rs}(560)/r_{rs}(665)$ as a proxy of chl-*a* concentration, an empirical model was established between the band ratio and ψ_1. A logarithmic fit showed strong relationship between the two (R^2=0.89, Fig. 5.5b). Empirical model was used to retrieve ψ_1 for all samples for further analysis. Similarly, ψ_2 was empirically modeled from $r_{rs}(620)/r_{rs}(665)$ band ratio. A power fit explained about 50% variance in the dataset (R^2=0.5, Fig. 5.5c).

Using the empirically retrieved ψ_1 and ψ_2 in Eq. 5.3, $a_{PC}(620)$ values were calculated from $a_{\varphi}(620)$ and $a_{\varphi}(665)$. For accuracy assessment, retrieved $a_{PC}(620)$

values were compared with $a_{PC}(620)$ measured *in vitro* (Fig. 5.6a). It should be noted that there was a systematic overestimation of model retrieved $a_{PC}(620)$, however, given that the retrieved $a_{PC}(620)$ were close to the measured values, intercept of the best-fit line between PC and $a_{PC}(620)$ should be close to zero. In this case, a linear fit produced an intercept of 1.76 (Fig. 5.6a) which may indicate that experimentally measured (in vitro) ψ_1 values were larger than the corresponding ψ_1 values if measured *in vivo* because of no package effect.

5.5.3.2 Using ψ_1 and ψ_2 derived by optimization

Understanding the difficulty of *in vivo* measurements, ψ_1 values were retrieved by optimization. For each sample, ψ_1 was solved using *in vitro* measured $a_{PC}(620)$ and ψ_2, and filter-pad measured $a_\varphi(620)$ and $a_\varphi(665)$ (Eq. 5.3). ψ_1 was varied iteratively within 0.8-2 with a 0.001 increment and the ψ_1 value that returns the most accurate $a_{PC}(620)$ value was selected. To assess the accuracy of the decomposition, the method was evaluated in an ideal case where the parameters ψ_1 and ψ_2 are known. $a_{PC}(620)$ was estimated from the optimized ψ_1, measured ψ_2 values. The model retrieved $a_{PC}(620)$ values were compared with measured $a_{PC}(620)$ (Fig. 5.7a). Retrieved $a_{PC}(620)$ matched very well with the measured values with an average relative error of ~10%. A linear fit between model retrieved and measured $a_{PC}(620)$ produced a slope of 1.02 and an intercept of 0.0059 (R^2=0.99, Fig. 5.6a).

The newly proposed method produced very high retrieval accuracy when sample by sample optimized ψ_1 and measured ψ_2 values were used; however, this method is not practical for a remote sensing mapping protocol. Therefore, another empirical model was

developed by establishing a power fit between optimized ψ_1 values and $r_{rs}(560)/r_{rs}(665)$ (R^2=0.39, Fig. 5.8). Later, modeled ψ_1 and ψ_2 values were used to retrieve $a_{PC}(620)$. Fig. 5.9(a) shows the comparison of model retrieved $a_{PC}(620)$ with measured $a_{PC}(620)$. Excluding 2 samples, all other samples fell close to the 1:1 line. The average relative error of the model retrieval was ~28%. After excluding the points with high $a_{\varphi}(620)$ estimation error, the average relative error decreased to ~22%.

5.5.4 Retrieving PC concentration from $a_{PC}(620)$

PC concentration can be calculated if $a_{PC}(620)$ and $a^*_{PC}(620)$ are known (Eq. 5.4). However, $a^*_{PC}(620)$ values widely vary among different species of cyanobacteria as well as with varying nutrient and light conditions (Simis *et al*., 2005; Tandeau, 1977). Because of this variability, Simis *et al*. (2005) used a sample-by-sample $a^*_{PC}(620)$ information to retrieve PC concentration in lake Loosdrecht, The Netherland and was able to retrieve PC with a better accuracy (RMSE=6.5 mg m^{-3} or 19%; R^2=0.94). However, the accuracy considerably decreased when a constant $a^*_{PC}(620)$ was used to retrieve the PC concentration. In this study, PC concentration was retrieved in three different ways, such that the effects of $a^*_{PC}(620)$ on final PC retrieval can be analyzed, including: 1) using known $a^*_{PC}(620)$ values, 2) using the mean $a^*_{PC}(620)$ value, and 3) using modeled $a^*_{PC}(620)$ values. The three different cases were analyzed using $a_{PC}(620)$ retrieved from: a) optimized ψ_1 and known ψ_2, and b) modeled ψ_1 and ψ_2.

5.5.4.1 PC from optimized ψ_1 and measured ψ_2

PC concentrations were estimated from $a_{PC}(620)$ derived by using optimized ψ_1 and measured ψ_2 values. Therefore, final estimation errors of PC can only be attributed to

$a^*_{PC}(620)$. First, sample-by-sample measured values of $a^*_{PC}(620)$ were used in an ideal case and compared with measured PC values (Fig.5.7b). The mean relative error of PC estimation was 10.06%. Slope and R^2 of the straight line fit between measured and retrieved PC concentration were ~1 and 0.99 highlighting very high retrieval accuracy. However, using sample-by-sample $a^*_{PC}(620)$ is not feasible in remote sensing. Therefore, the mean value of $a^*_{PC}(620)$ of all samples was used in this study (0.0048 m^2 mg^{-1}) to estimate PC. When compared, estimated PC values still matched very well with the measured (R^2=0.98, slope=0.96, Fig. 5.6b). However, the mean relative error increased to 22.32% (Table 5.2). It was also noticed that $a^*_{PC}(620)$ linearly increased with $r_{rs}(620)/r_{rs}(665)$ (R^2=0.4) and explained the variability to some extent. Use of modeled $a^*_{PC}(620)$ brought the mean relative error down to 20.36%. In all three cases, mean relative error of PC estimation varied from ~10 to 22%.

5.5.4.2 PC from modeled ψ_1 and ψ_2

PC estimation error was minimal when it was calculated using the above mentioned optimization process. However, this is impractical to be used in remote sensing, therefore, PC concentration was estimated from $a_{PC}(620)$ that was retrieved from modeled ψ_1 and ψ_2. Sample-by-sample measured $a^*_{PC}(620)$ were used for PC estimation (Fig. 5.9b). The mean and median relative error of PC estimation was 37.18% and 27.2% respectively. When the mean $a^*_{PC}(620)$ was used, the mean and median relative error decreased to ~34.86% and 25.96% respectively. Similarly, mean relative error of PC estimation using modeled $a^*_{PC}(620)$ was 36.23% and the median relative error decreased to ~22% (Table 5.2).

5.5.5 Empirical model for PC retrieval

An empirical model was also developed for potential use to retrieve PC with an expectation that it will reduce the uncertainties in PC prediction because of the variability in $a_{PC}^{*}(620)$. A straight line fit explained the strong relationship and variance between $a_{PC}(620)$ and PC concentration very well (R^2=0.96, p<0.0001) (Fig. 5.10). The functional relationship between $a_{PC}(620)$ and PC concentration can be explained using parameters of the straight line fit where the slope is interpreted as the mean $a_{PC}^{*}(620)$. In this dataset, the slope was close to the mean value of $a_{PC}^{*}(620)$.

5.5.6 Algorithm Uncertainties

Even if we assume that the *in situ* R_{rs} measurements were completely error free; there are still multiple sources of uncertainties at different steps of the above developed procedure. Lee *et al*. (2010) have systematically studied the sources of uncertainties of inverted inherent optical property products from QAA. For coastal waters, errors from the modeled $a_t(\lambda_0)$ at the reference wavelength will have a larger effect, as compared to $b_{bp}(\lambda)$ modeling errors on the $a_t(\lambda)$ estimation. It was found that for strongly absorbing waters, when $a_t(440)$ approaches 0.5 m^{-1}, uncertainty in retrieving $a_t(440)$ could be up to $\pm$ 37% (Lee *et al*, 2010). There could be even considerably higher uncertainties in a_{φ} and $a_{CDM}(\lambda)$ estimation because of the algebraic nature of the decomposition of $a_t(\lambda)$ signal using empirically derived spectral shape parameter. In this research, even another step has been added to decompose the $a_{\varphi}(620)$ for retrieving $a_{PC}(620)$ using similar algebraic methods, where the coefficients, ψ_1 and ψ_2, were modeled from r_{rs} band ratios. Therefore, uncertainties with those values will introduce errors in the final estimation of $a_{PC}(620)$ as well.

Another critical source of error in PC modeling is from the variability of $a^*_{PC}(620)$. As discussed earlier, $a^*_{PC}(620)$ can vary based on the response of the cyanobacteria species to environmental and optical conditions (Simis *et al*., 2005). Grossman *et al*. (1993) reported that a number of phycoerythrin (PE) and PC hexamers in phycobilisome structure (PBS) change based on the quality of light availability. More PC pigments are synthesized in cyanobacteria cells when exposed to red light, whereas, more PE pigments are synthesized in cyanobacteria cells when exposed to green light. Hence, the intra-cellular PC and PE concentrations are sensitive to the quality of light availability and can have corresponding effect on a^*_{PC}. It has also been reported that the cell morphology and photo adaptation may cause variability in $a^*_{chla}(\lambda)$ (Satyendranath *et al*., 1987; Bricaud *et al*., 1995). Similar reasons may cause variability in $a^*_{PC}(620)$ as well. Also, being an accessory photosynthetic pigment that harvests light in the "green gap" of chl-*a*, the cellular PC is more likely to vary because of changing nutrient level and light condition and consequently affecting the $a^*_{PC}(\lambda)$ (Tandeau, 1977). In this study, $a^*_{PC}(620)$ varied from 0.003 to 0.006 $m^2\ mg^{-1}$ with an average value of 0.0048 $m^2\ mg^{-1}$. If the mean $a^*_{PC}(620)$ is used as a constant, the model will have 25% underestimation for the sample with highest $a^*_{PC}(620)$ and 37.5% overestimation for the sample with lowest $a^*_{PC}(620)$. However, use of the the modeled $a^*_{PC}(620)$ values addresses the variability to some extent and somewhat reduces the estimation error. Model uncertainties also depend on the relative contribution of cyanobacteria in the water supporting the findings in Simis *et al* (2005). Results show that the relative errors were higher for samples with PC:Chl-*a* > 0.5 (Fig. 5.11). Samples with lower a_{PC}

(620)/a_{φ}(620) and a_{PC}(620)/a_{CDM}(620) values also showed higher relative errors (Fig. 5.12).

Another source of error could have been from the analytical measurement of PC concentration. Sarada *et al*. (1999) reported that extraction procedures can affect the quality of extracted PC and hence the final accuracy of PC estimation. The samples extracted by homogenization using a sonicator showed a minor second peak at 678 nm, in addition to the major peak at 620 nm, indicating some chlorophyll contamination while disintegrating the cells (Sarada *et al*., 1999). In this study, homogenization procedure was used for the extraction of PC. Hence samples with high chl-*a* and chl-*b* concentration will have higher chance of inaccuracy. Similarly, extraction efficiency of PC samples also affects the final PC concentration as well as the $a_{PC}^{*}(620)$. For example, lower (higher) extraction efficiency could over (under) estimate $a_{PC}^{*}(620)$ and will eventually affect the final estimation accuracy.

5.6 Conclusions

A novel method has been developed to quantify PC using the multiband quasi-analytical algorithm (Lee *et al*., 2002). In a parallel study, QAA was parameterized to retrieve $a_{\varphi}(\lambda)$ in extremely turbid and hypertrophic waters (Mishra *et al*., 2012). The newly developed method was successful in decomposing the a_{φ}(620) to $a_{PC}(620)$ with a mean relative error of ~28%. Magnitude of $a_{PC}^{*}(620)$ varied widely (0.003-0.006 m^2 mg^{-1}) in this dataset. To reduce the uncertainty in PC estimation because of $a_{PC}^{*}(620)$ variability, an empirical model was developed to retrieve $a_{PC}^{*}(620)$ from r_{rs}(620)/r_{rs}(665). PC concentration was calculated by dividing $a_{PC}(620)$ values by

modeled $a^*_{PC}(620)$. The mean and median relative error of PC estimation was 36.23% and ~22 % respectively. An empirical relationship was also established between model derived $a_{PC}(620)$ and measured PC concentration. A straight-line fit to the data explained strong dependence between the variables (R^2=0.96, p=0.0001). The newly developed method shows a strong potential of quantifying and mapping PC in optically complex turbid and productive waters. Findings from this research are based on a small dataset with a wide range and because of the same reason the robustness of the new empirical steps to retrieve ψ_1 and ψ_2 have not been fully assessed. Future work will focus on using a large dataset for a robust optimization and validation of the proposed model. Results demonstrate that the new approach will be suitable for quantifying PC concentration in cyanobacteria dominated turbid productive waters such as inland lakes and estuaries. Band architecture of the model matches with the spectral channels of the Medium Resolution Imaging Spectrometer (MERIS) and assures that MERIS reflectance products can be used to quantify cyanobacterial harmful algal blooms in optically complex waters.

Table 5.1 Descriptive statistics of pigment and absorption coefficient measurements

Parameters	Mean	Std	Min	Max	*N*
Pigment Concentrations					
PC (ug/L)	418.76	669.75	68.13	3032.47	24
Chl-*a* (µg l^{-1})	295.96	302.06	59.40	1376.60	24
Chl-*b* (µg l^{-1})	4.36	3.15	1.57	13.71	18
PC:Chl-*a*	1.23	0.73	0.30	3.29	24
Total Absorption Coefficients, a_t(m^{-1})					
a_t(443) (m^{-1})	16.21	8.15	8.85	47.21	24
a_t(620) (m^{-1})	5.25	2.97	2.07	15.33	24
a_t(665) (m^{-1})	5.42	2.54	2.55	13.84	24
a_t(708) (m^{-1})	1.76	0.44	1.23	3.07	24
Phytoplankton Absorption Coefficients, a_φ (m^{-1})					
a_φ(443) (m^{-1})	12.54	6.77	5.24	37.67	24
a_φ(620) (m^{-1})	4.66	2.80	1.61	13.93	24
a_φ(665) (m^{-1})	4.80	2.44	2.04	12.75	24
a_φ(708) (m^{-1})	0.86	0.40	0.37	1.94	24
a_φ(443)/ a_t(443)	0.76	0.06	0.59	0.85	24
a_φ(620)/ a_t(620)	0.87	0.04	0.78	0.94	24
Absorption Coefficients of CDOM and Detritus, a_{CDM} (m^{-1})					
a_{CDM}(443) (m^{-1})	3.66	1.63	2.08	9.54	24
a_{CDM}(620) (m^{-1})	0.32	0.21	0.09	1.12	24
a_{CDM}(665) (m^{-1})	0.19	0.13	0.04	0.67	24
a_{CDM}(709) (m^{-1})	0.07	0.06	0.01	0.30	24
a_{CDM}(620)/a_φ(620)	0.07	0.02	0.03	0.11	24
PC Absorption Coefficients					
$\boldsymbol{a_{PC}(620)}$ (m^{-1})	1.87	2.92	0.30	13.94	24

Table 5.2 Mean, median, standard deviation of relative error from models with different ψ_1 and $a_{PC}^*(620)$.

Rel. Error (%)	Mean	Median	Std	Slope	R^2
Optimized ψ_1 and Measured ψ_2					
Known $\boldsymbol{a_{PC}^*}$	10.06	6.19	8.03	~1	0.99
Mean $\boldsymbol{a_{PC}^*}$	22.32	15.93	16.55	0.96	0.98
Modeled $\boldsymbol{a_{PC}^*}$	20.36	21.61	12.94	1.15	0.99
Modeled ψ_1 and ψ_2					
Known $\boldsymbol{a_{PC}^*}$	37.18	27.20	34.99	~1	0.98
Mean $\boldsymbol{a_{PC}^*}$	34.86	25.96	33.14	0.97	0.98
Modeled $\boldsymbol{a_{PC}^*}$	36.23	22.08	31.31	1.16	0.98

Slope and R^2 from least-square regression between measured and modeled PC concentrations were also provided.

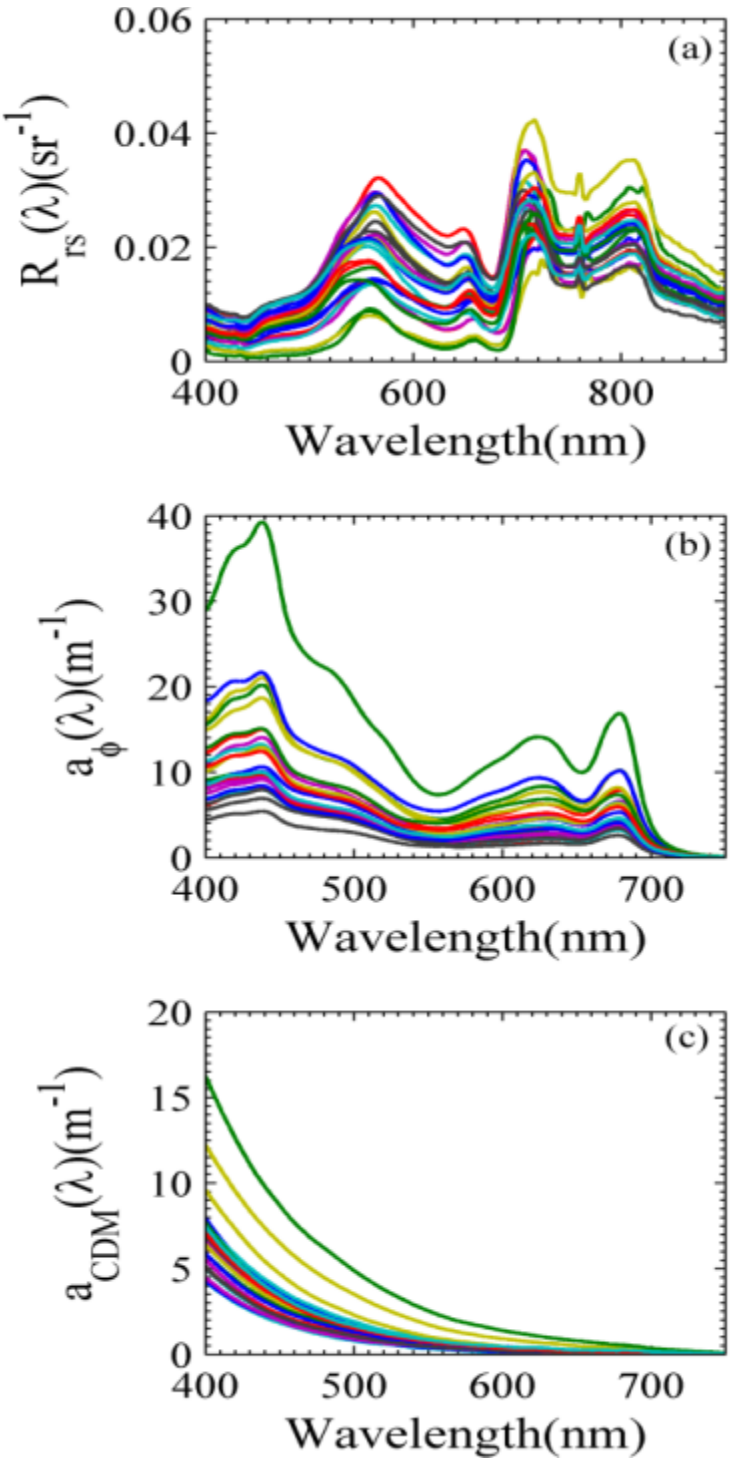

Figure 5.1 (a) $R_{rs}(\lambda)$ spectra collected from aquaculture ponds in July 2010 and April 2011, (b) measured phytoplankton absorption coefficients, $a_{\varphi}(\lambda)$, using filter-pad technique, and (c) measured $a_{CDM}(\lambda)$.

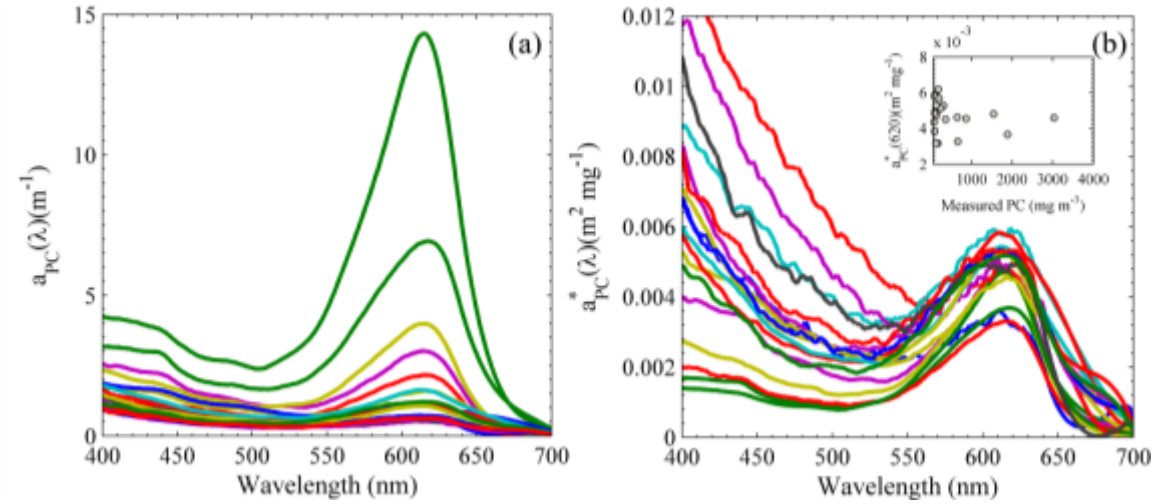

Figure 5.2 (a) $a_{PC}(\lambda)$ measured using spectrophotometric method, and (b) phycocyanin specific absorption coefficients, $a^*_{PC}(\lambda)$, of samples analyzed in this study.

Scatter plot between PC concentration and $a^*_{PC}(620)$ is also shown in the inset.

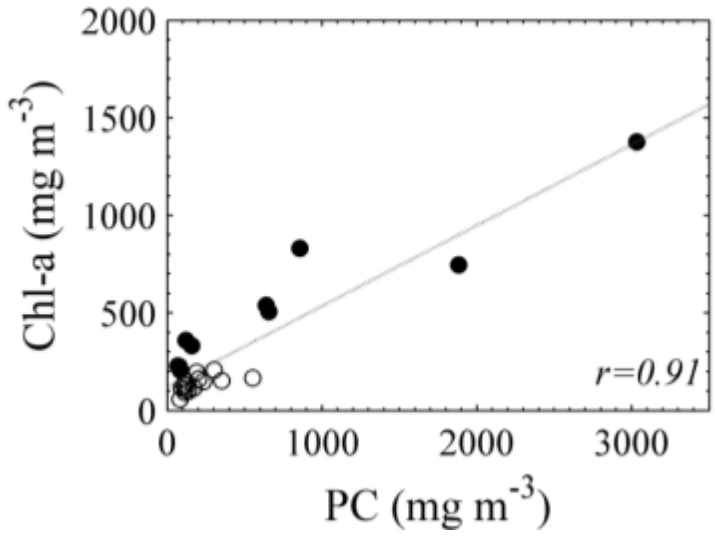

Figure 5.3 Relationship between measured PC and chl-*a* concentration in the samples.

Solid and empty circles represent data from 2010 and 2011 respectively.

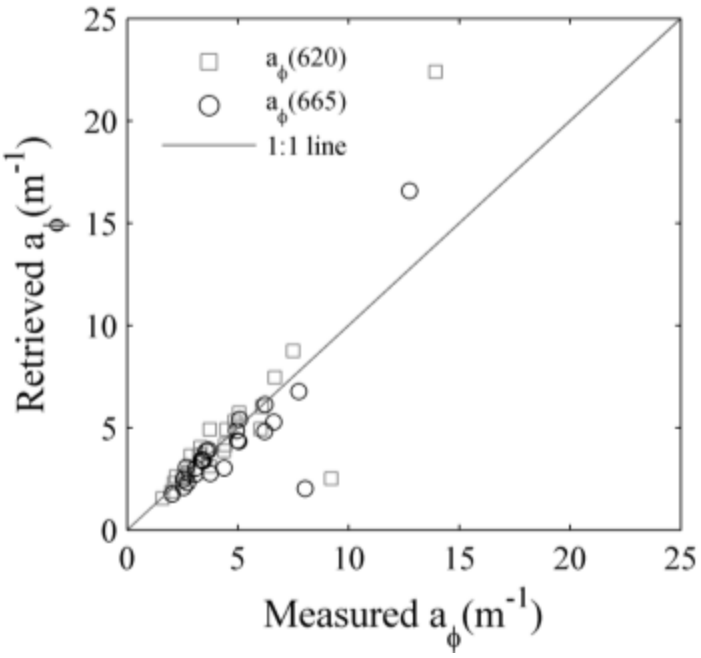

Figure 5.4 Comparison of model retrieved $a_\varphi(\lambda)$ with pad-measured values at 560, 620, and 665 nm wavelengths.

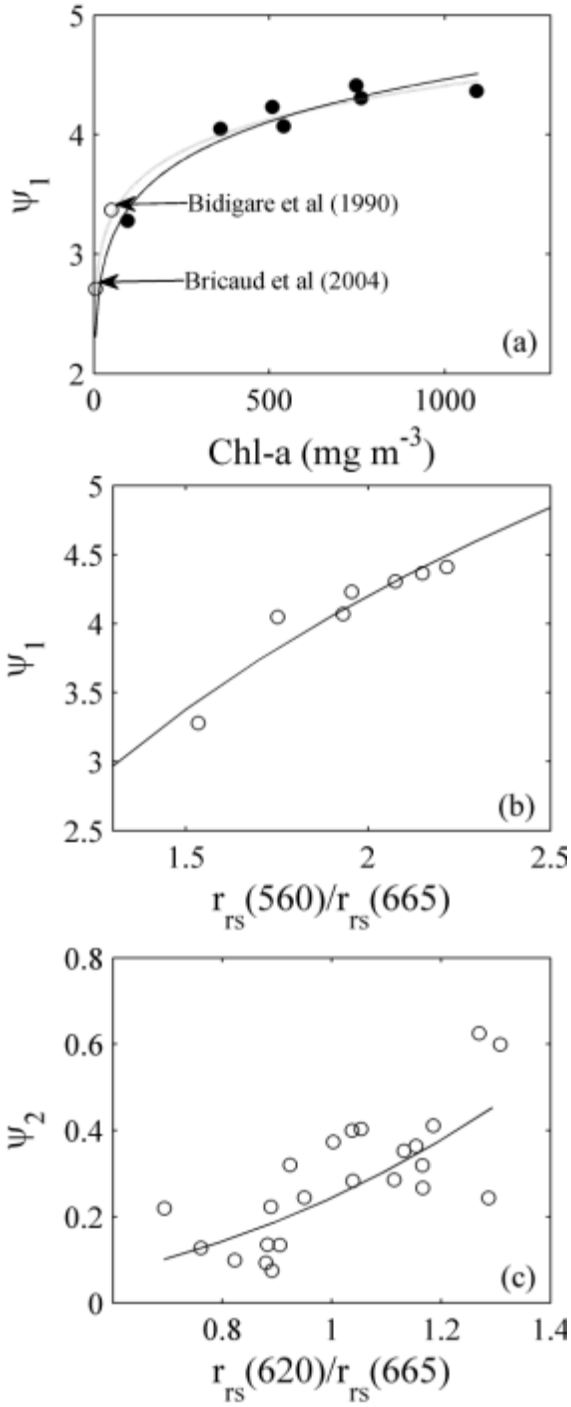

Figure 5.5 (a) Scatter plot showing empirical relationship between ψ_1, $[a_{chl}(665)/a_{chl}(620)]$, chl-a concentration.

Filled circles are data points from this study; (b) Empirical relationship between $\psi_1$and $r_{rs}(560)/r_{rs}(665)$, and (c) empirical relationship between ψ_2, $[a_{PC}(665)/a_{PC}(620)]$, and $r_{rs}(620)/r_{rs}(665)$.

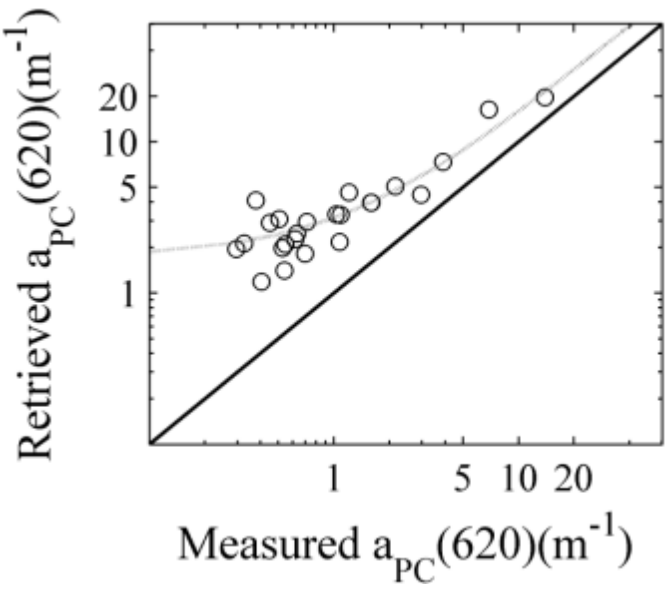

Figure 5.6 (a) Comparison of model retrieved $a_{PC}(620)$ with spectrophotometrically measured$a_{PC}(620)$.

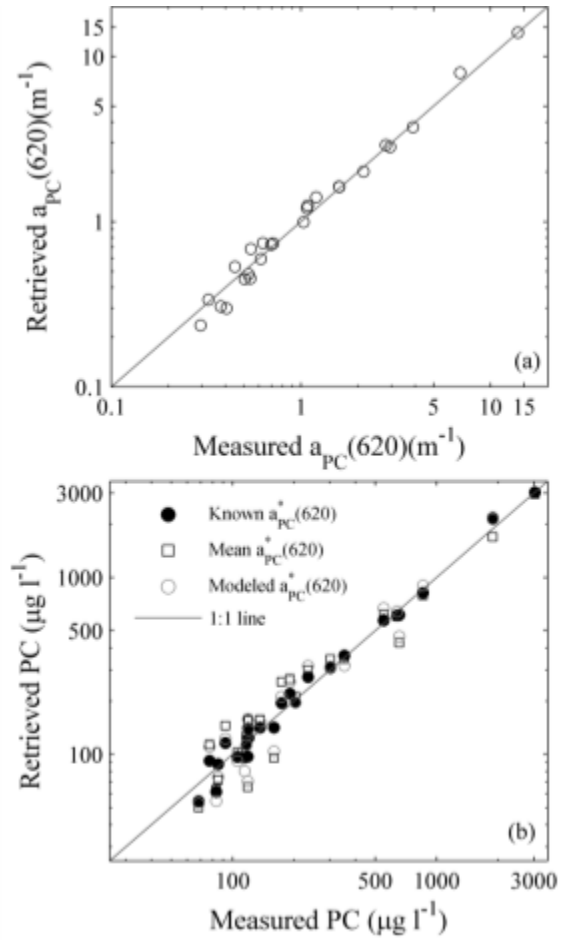

Figure 5.7 Comparison of a) Model retrieved $a_{PC}(620)$ with measured$a_{PC}(620)$; b) Estimated PC concentration with measured ones. Solid lines are 1:1 lines.

Filled circles, squares and empty circles represent estimated PC values using known $a^*_{PC}(620)$, mean $a^*_{PC}(620)$, and empirically modeled $a^*_{PC}(620)$.

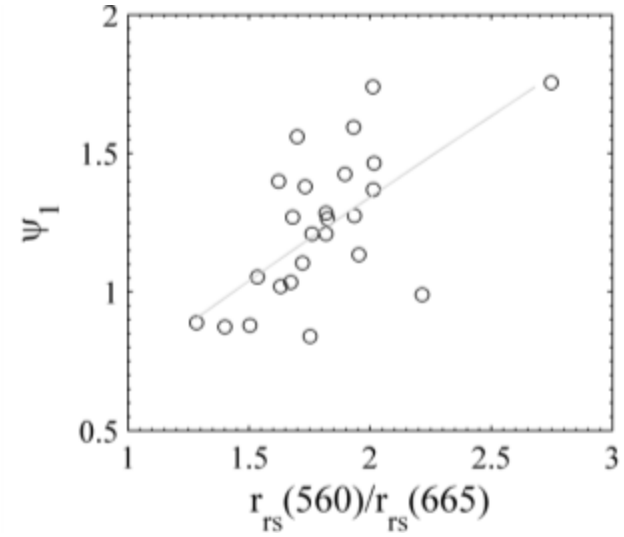

Figure 5.8 Scatter plot between ψ_1 and $r_{rs}(560)/r_{rs}(665)$.

Dotted line is the least square fit line.

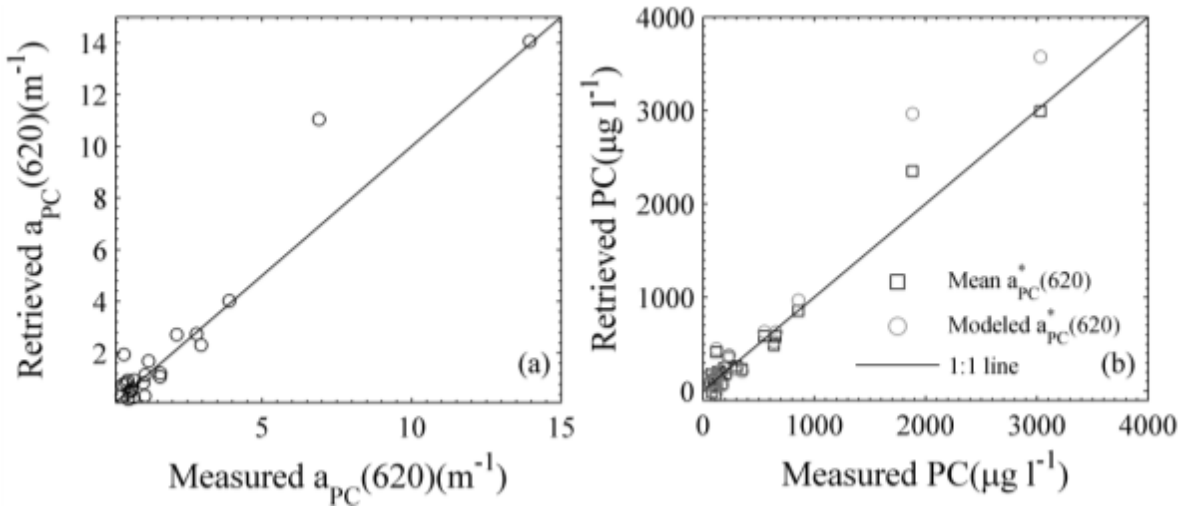

Figure 5.9 Comparison of a) Model retrieved $a_{PC}(620)$ with measured$a_{PC}(620)$; b) Estimated PC concentration with measured ones.

Solid lines are 1:1 lines. Squares and empty circles represent estimated PC values using mean $a^*_{PC}(620)$, and empirically modeled $a^*_{PC}(620)$.

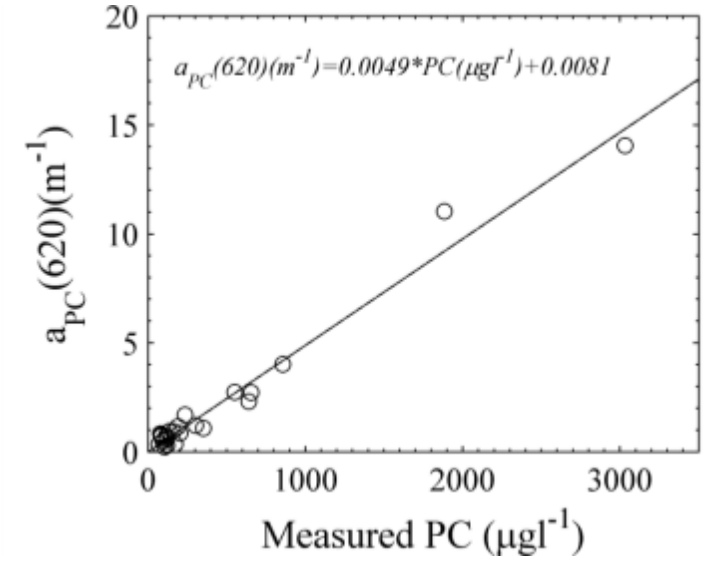

Figure 5.10 Scatter plot between measured PC concentration and from modeled ψ_1 and ψ_2.

Note that intercept of the least square fit line is close to zero unlike the case in Fig. 6b.

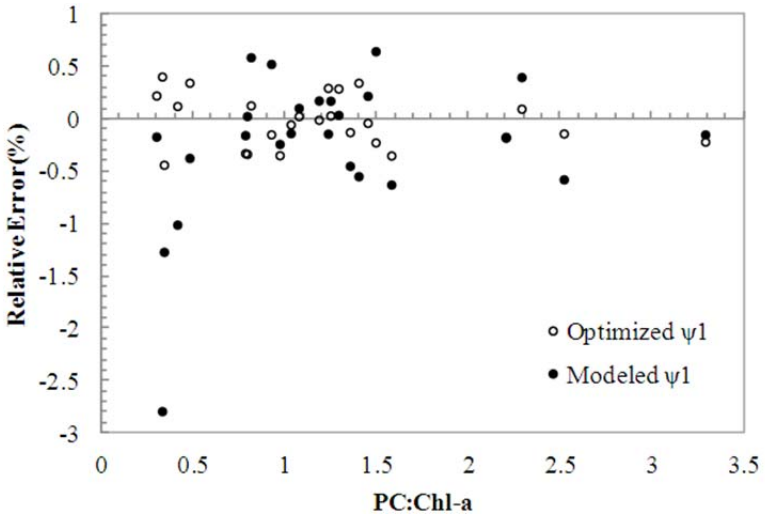

Figure 5.11 Relationship between PC:chl-a and relative error of PC estimation.

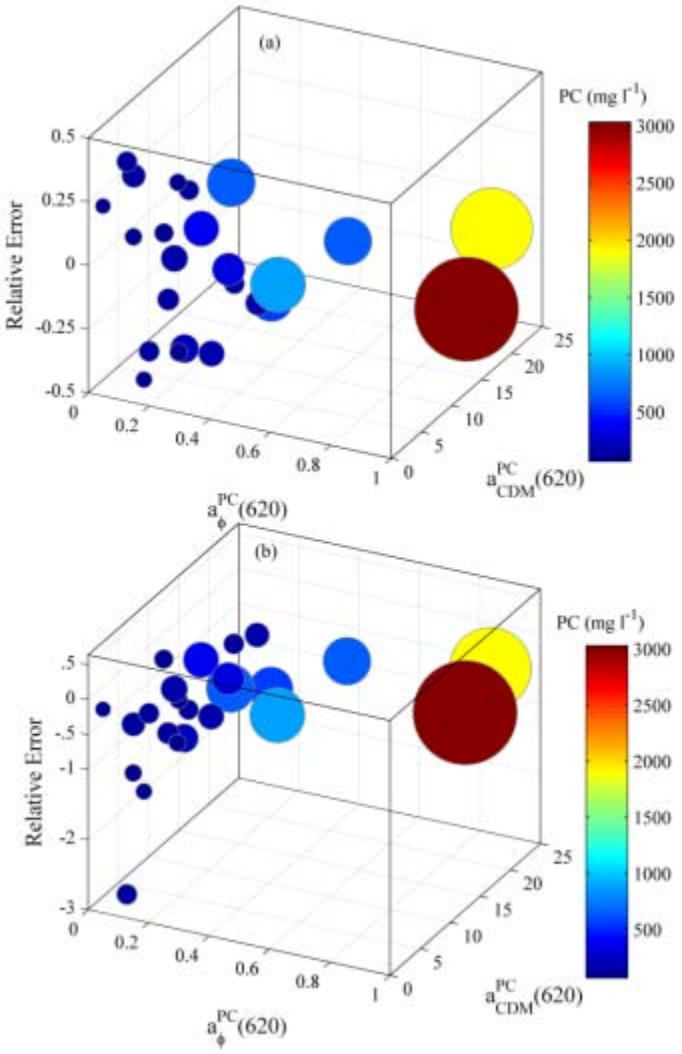

Figure 5.12 Relationship between relative error of PC estimation and relative abundance of PC absorption: a) errors from the model using optimized ψ_1 and measured ψ_2, and b) errors from the model using modeled ψ_1 and ψ_2.

Colors and size of the bubbles indicate the PC concentration.

5.7 References

Arar E.J. (1997). In vitro determination of chlorophylls a, b, c1 + c2 and pheopigments in marine and freshwater algae by visible spectrophotometry. USEPA Method 446-0.

Bennett, A., & Bogorad, L. (1973). Complementary chromatic adaptation in a filamentous blue-green alga. *Journal of Cell Biology*, 58: 419–435.

Bidigare, R. R., Ondrusek, M. E., Morrow, J. H. & Kiefer, D. A. (1990). In vivo absorption properties of algal pigments. Ocean Optics X, SPIE Proceedings, Society of Photo-Optical Instrumentation Engineers, 1302: 290–302.

Boyer, G. L. (2008) Cyanobacterial Toxins in New York and the Lower Great Lakes Ecosystems. In: "Proceedings of the Interagency International Symposium on Cyanobacterial Harmful Algal Blooms" H. K. Hudnell, Ed., *Advances in Experimental Medical Biology*, 619: 151-163.

Bricaud, A., Claustre, H., Ras, J. & Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. *Journal of Geophysical Research 109(C11): doi: 10.1029/2004JC002419. issn: 0148-0227.*

Bricaud, A., Babin, M., Morel, A. & Claustre, H. (1995). Variability in the chlorophyllspecific absorption-coefficients of natural phytoplankton - Analysis and parameterization. *Journal of Geophysical Research-Oceans,* 100(C7): 13321-13332.

Campbell, J. W. & O'Reilly, J. E. (2006). Metrics for Quantifying the Uncertainty in a Chlorophyll Algorithm: Explicit equations and examples using the OC4.v4 algorithm and NOMAD data, Ocean Color Bio-optical Algorithm Mini-workshop, 7 - 29 September 2005, University of New Hampshire, Durham, New Hampshire.

Craig, S. E., Lohrenz, S. E., Lee, Z. P. Mahoney, K. L., Kirkpatrick, G. J., Schofield, O. M., & Steward, R. G. (2006). Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga, Karenia brevis, *Applied Optics*, 45 (21): 5414-5425.

Dall'Olmo, G. & Gitelson, A.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(3): 412-422.

Dekker, A. (1993). Detection of the optical water quality parameters for eutrophic waters by high resolution remote sensing, Ph.D. thesis, Free University, Amsterdam.

Fargion, G. S. & J. L. Mueller. (2000). Ocean optics protocols for satellite ocean colour sensor validation. Revision 2," NASA Tech. Memo. 209966, *SeaWiFS Technical Report Server, (NASA Goddard Space Flight Center*).

Glazer, A.N. (1989). Light guides: directional energy transfer in a photosynthetic antenna. *Journal of Biological Chemistry,* 264: 1-4.

Hanson T.R. (2003) Economic impact of off-flavour to the U.S. catfish industry. In: Off-Flavors in Aquaculture (ed. By A.M. Rimando & K.K. Schrader), pp. 13-29. *American Chemical Society*, Washington, DC.

Hoge, Frank E. & Paul E. Lyon, Satellite retrieval of inherent optical properties by linear matrix inversion ofoceanic radiance models: An analysis of model and radiance measurement errors, *Journal of Geophysical Research, 101,*16,631-16,648,1996.

Kutser, T., Metsamaa, L., Strombeck, N. & Vahtmae, E. (2006). Monitoring cyanobacterial blooms by satellite remote sensing. *Estuarine Coastal and Shelf Science*, 67: 303-312.

Le, C.F., Li, Y.M., Zha, Y., Sun, D., & Yin, B. (2009). Validation of a Quasi-Analytical Algorithm for Highly Turbid Eutrophic Water of Meiliang Bay in Taihu Lake, China, *IEEE Transactions on Geosciences and Remote Sensing*, 47 (8): 2492-2500.

Lee, Z.P., Carder, K.L., R.A., & Arnone, R. (2002). Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters, *Applied Optics,* 41: 5755-5772.

Lee, Z.P., & Carder, K.L. (2004). Absorption spectrum of phytoplankton pigments derived from hyperspectralremote-sensing refl ectance. *Remote Sensing of Environment,* 89: 361-368.

Lee, Z.P., Arnone, R., Hu, C., Werdell, J., & Lubac, B. (2010). Uncertainties of optical parameters and their propagation in an analytical ocean color inversion algorithm, *Applied Optics*, 49(3): 369-381.

Maritorena, S., Siegel, D.A., & Peterson, A.R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, 41: 2705-2714

Metcalf, J. S., Banack, S. A., Lindsay, J., Morrison, L. F., Cox, P. A., & G.A. Codd. (2008). Cooccurrence of beta-N-methylamino-L-alanine, a neurotoxic amino acid with other cyanobacterial toxins in British waterbodies, 1990–2004. *Environmental Microbiology*, 10: 702−708.

Mishra, S., Mishra, D. R., & Lee, Z.P. (2012). Bio-Optical Inversion in Highly Turbid and Cyanobacteria Dominated Waters, *IEEE Transactions on Geoscience and Remote Sensing* (Under Review)

Mishra, S., Mishra, D.R., & Schluchter. W.M. (2009). A Novel Algorithm for Predicting PC Concentrations in Cyanobacteria: A Proximal Hyperspectral Remote Sensing Approach. *Remote Sensing*, *1*: 758-775.

Ohde, T. & Siegel, H. (2003). Derivation of immersion factors for the hyperspectral TriOS radiance sensor, *Journal of Optics A: Pure and Applied Optics,* 5**:** L12–L14.

Richardson, L.L. (1996). Remote sensing of algal bloom dynamics. *BioScience*., 46: 492-501.

Ruiz-Verdu, A., Simis, S. G. H., de Hoyos, C., Gons, H. J., & R. Pena-Martinez. (2008). An evaluation of algorithms for the remote sensing of cyanobacterial biomass. *Remote Sensing of Environment,* 112: 3996−4008.

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

Sathyendranath, S., Lazzara, L. & Prieur, L. (1987). Variations in the spectral values of specific absorption of phytoplankton. *Limnology & Oceanography,* 32**:** 403–415.

Schalles, J., & Yacobi, Y. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll-a pigments in eutrophic waters, *Archives fur Hydrobiologia –Sp. Issues Advances in Limnology*, 55: 153-168.

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

Tandeau de Marsac N. (1977). Occurrence and nature of chromatic adaptation in cyanobacteria. *Journal of Bacteriology,* 130**:** 82–91.

Tucker, C.S., & Boyd, C.E. (1985). Water quality. In Channel catfish culture. Ed. C.S. Tucker. Elsevier Science Publishers. Amsterdam. p. 135-227.

Vincent, R. K., Qin, X. M., McKay, R. M. L. , Miner, J., Czajkowski, K., Savino, J., & Bridgeman, T. (2004). Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie*, Remote Sensing of Environment,* 89: 381-392.

Zhu, W., Yu, Q., Tian, Y. Q., Chen, R. F., & Gardner, G. B. (2011). Estimation of chromophoric dissolved organic matter in the Mississippi and Atchafalaya river plume regions using above-surface hyperspectral remote sensing, *Journal of Geophysical Research,* 116, no. C02011, 1-22.

CHAPTER VI

NORMALIZED DIFFERENCE CHLOROPHYLL INDEX: A NOVEL MODEL FOR REMOTE ESTIMATION OF CHLOROPHYLL-A

6.1 Overview

In this research, a Normalized Difference Chlorophyll Index (NDCI) has been proposed to predict chlorophyll-a (chl-*a*) concentration from remote sensing data in estuarine and coastal turbid productive (case 2) waters. NDCI calibration and validation results derived from simulated and MEdium Resolution Imaging Spectrometer (MERIS) datasets show its potential application to widely varying water types and geographic regions. A quadratic function (R^2 = 0.95, p<0.0001) accurately explained the variance in the simulated data for a chl-*a* range of 1-60 mg m^{-3}. Similarly a twofold calibration and validation of chl-*a* models using MERIS dataset, (chl-*a* range: 0.9-28.1 mg m^{-3}) yielded R^2 of 0.9, and RMSE of ~2 mg m^{-3} respectively. NDCI was applied on images over the Chesapeake Bay and Delaware Bay, the Mobile Bay, and the Mississippi River delta region in the northern Gulf of Mexico. The newly developed algorithm was successful in predicting chl-*a* concentration with approximately 12% overall bias for all above study regions. In case of remote coastal waters with no ground truth data, NDCI can be used to detect algal bloom and qualitatively infer chl-*a* concentration ranges very similar to NDVI's application in terrestrial vegetation studies.

6.2 Introduction

Accurate remote estimation of biophysical parameters such as chlorophyll-*a* (chl-*a*) and phytoplankton biomass in turbid productive waters is essential for large-scale and multi-temporal studies related to primary production, carbon cycle, biogeochemical cycles, and overall inland and coastal water quality. However, it is still a challenge because of the presence of non-covarying optically active constituents whose absorption features overlap with chl-*a*. Spectral channels in the blue-green part of the electromagnetic spectrum are heavily affected by the presence of constituents such as Colored Dissolved Organic Matter (CDOM), detritus, and tripton. Empirical algorithms (e.g., OC4v4) that use blue and green spectral channels often provide a relatively accurate estimate of chl-*a* in case 1 waters where the total non-water absorption is dominated by phytoplankton, however, do not provide reasonable estimates of chl-*a* in turbid productive waters (O'Reilly *et al*., 1998). In order to reduce the estimation error of chl-*a* in turbid productive waters, semi-analytical models have also been proposed (Maritorina *et al*., 2002, Gons *et al*., 2000). However, the success of these semi-analytical models depends on the accurate parameterization of the inherent optical properties of the medium that often poses a considerable challenge. Magnuson *et al*. (2004) re-parameterized the original semi-analytical model proposed by Maritorina *et al*. (2002) to make it suitable for the Chesapeake Bay and Mid-Atlantic Bight region. They reported that the re-parameterized model was successful in attributing CDOM absorption in the total absorption budget and offered accurate estimation when compared to OC4v4 model and the accuracy of chl-*a* estimation was within 30-50% of the *in situ* measured values. Even though the re-parameterized semi analytical model produced better accuracy than OC4v4,

the uncertainty was still very high. Because of the difficulties in obtaining the information for re-parameterization, such as specific absorption coefficient of phytoplankton, $a_{ph}^{*}(\lambda)$, spectral slope of colored dissolved organic matter, S_{CDOM}, and spectral slope of detritus, $S_{detritus}$, alternative approaches have been encouraged to improve chl-*a* estimation in turbid productive waters. Over the past years, numerous algorithms have been proposed to quantify chl-*a* in turbid productive waters using red-near infrared (NIR) bands and these algorithms can be classified into three primary groups such as: 1) two-band empirical (Moses *et al*. 2009; Tzortziou *et al*. 2007), 2) three or four-band empirical (Dall'Olmo & Gitelson, 2005; Le *et al*., 2009), and 3) three-band semi-analytical (Gons *et al*., 1999, 2008) models. In this study, accuracy of red and NIR based empirical and semi-analytical algorithms has been assessed in geographically diverse water bodies and a novel band difference algorithm has been proposed for accurately mapping chl-*a* concentration in turbid productive waters.

Four algorithms from the list above were selected for further validation and examination of their performance and transferability to different water bodies. The selected algorithms discussed below are widely applied and unique in their band architecture.

Moses *et al*. (2009) presented a two-band model (hereafter M09) using red and NIR bands to quantify chl-*a* in turbid productive waters. To match the band configuration of MERIS sensor, the conceptual model was designed as:

$$C_{chl-a} \propto R_{rs}^{-1}(665) \times R_{rs}(708) \qquad (6.1)$$

They applied M09 on MERIS images over Azov Sea, Russia and reported the high accuracy potential of the model to estimate chl-*a* in turbid productive waters. Similarly, Tzortziou *et al*. (2007) collected an extensive bio-optical dataset to examine the relationship between inherent and apparent optical properties in the mid Chesapeake Bay, USA. They observed a better relationship (R^2=0.54) between remote sensing reflectance (R_{rs}) ratio at 677 and 554 nm, $R_{rs}(677)/R_{rs}(554)$, and chl-*a* concentration in the bay compared to blue-green spectral band ratios. In this study, this ratio has been modified based on the MERIS band configuration, $R_{rs}(665)/R_{rs}(559)$ and named it T07 for further reference.

$$C_{chl-a} \propto R_{rs}^{-1}(559) \times R_{rs}(665) \tag{6.2}$$

Dall'Olmo & Gitelson (2005) (hereafter D05) presented a three-band model using red and NIR bands. The three-band model architecture was as follows:

$$C_{chl-a} \propto [R_{rs}^{-1}(665) - R_{rs}^{-1}(708)] \times R_{rs}(753) \tag{6.3}$$

The three band algorithm was based on several assumptions including, (i) the absorption by suspended solids and CDOM beyond 700 nm is approximately equal to that at 665-675 nm and the difference between them is very small and can be neglected, (ii) the total chl-*a*, CDOM, and total suspended sediment (TSS) absorption beyond 730 nm is nearly zero, and (iii) back-scattering coefficient of chl-*a* is spectrally invariant. They reported that D05 was successful to predict accurate estimate of chl-*a* in turbid productive water bodies with wide range of optical complexity. Moses *et al*.(2009) further validated D05 using MERIS data from the Azov Sea and documented that D05 was able to retrieve chl-*a* concentration with a RMSE of 5.02 mg m^{-3} (for a chl-*a* range:

18.37-47.86 mg m^{-3}). Gons *et al*. (2008) presented a semi-analytical algorithm (hereafter G08) for chl-*a* retrieval using MERIS data which was a modification of the parent algorithm (Gons *et al*., 1999). G08 uses the relationship between inherent optical properties and the R_{rs} at three wavelengths, solves for chl-*a* absorption at 665 nm, and estimates chl-*a* by dividing $a_{chl}(665)$ by the specific absorption coefficient of chl-*a,* $a^*_{chl-a}(665)$.

$$[C_{chl-a}] = \left\{\left[\frac{R_{rs}(708.75)}{R_{rs}(665)}\right] * (0.70 + b_b) - 0.40 - b_b^{1.06}\right\}/0.016 \quad (6.4)$$

where b_b is the back-scattering coefficient and was expressed as:

$$b_b = \frac{1.61 * R_{rs}(775)}{0.82 - 0.6 * R_{rs}(775)} \quad (6.5)$$

Gons *et al*. (2008) reported that G08 successfully retrieved chl-*a* concentration in the Laurentian great lakes producing residuals less than 35% of the measured values. They also reported that G08 did not perform well in areas with chl-*a* less than 5 mg m^{-3} and even produced some negative values in oligotrophic waters.

Although, the three-band algorithms, D05 and G08, have excellent predictive ability, the biggest challenge of these models is that they require R_{rs} measurements at 753 and 775 nm. Based on existing atmospheric correction schemes for turbid productive waters, getting reliable estimates of R_{rs} at these wavelengths is a difficult task. In addition, another inherent difficulty of semi-analytical models such as G08 is the use of $a^*_{chl-a}(665)$. Any uncertainty associated with $a^*_{chl-a}(665)$ can contribute to inaccurate estimates of chl-*a*.

After carefully examining the strengths and weaknesses of the above models, our goal in this research was to develop an algorithm that can perform better than the existing

algorithms. Two of the most important criteria for a successful spectral algorithm development include: (1) applicability to satellite data, and (2) transferability to widely varying geographic regions without producing significant uncertainties. Both criteria have been tested and analyzed in this study as part of the model validation. A novel index, Normalized Difference Chlorophyll Index (NDCI) has been proposed, and its sensitivity to chl-*a* concentration has been demonstrated in turbid productive waters. A chl-*a* model using NDCI has been calibrated and validated by analyzing four datasets (one simulated and three field datasets) representing unique turbid productive water bodies and presented its potential use for chl-*a* estimation in optically complex waters. The purpose of using a simulated data was to test the model performance and sensitivity to a wide range of optical parameters in the water.

NDCI uses R_{rs} at 665 nm, $R_{rs}(665)$, and 708 nm, $R_{rs}(708)$, emulating the Medium Resolution Imaging Spectrometer (MERIS) channels. Similar to other turbid productive chl-*a* algorithms, this index uses the information from the reflectance peak centered at 700 nm which is maximally sensitive to the variations in chl-*a* concentration in water. Similarly, a wide spectral absorption peak between 665 nm and 675 nm is generally assigned to the absorption by chl-*a* pigments. Those two spectral features centered at 665 nm and 708 nm were selected to develop NDCI and to avoid the confounding influence of CDOM and TSS on the water reflectance spectra at shorter wavelengths. Also, as both bands are closely located, it is assumed that the CDOM and TSS absorption is similar in magnitude. Based on the results from bio-optical modeling in this study, the combined range of CDOM and TSS absorptions at 665 nm (0.0193-0.1899 m^{-1}) and 708 nm (0.015-0.1603 m^{-1}) in the study regions are approximately equal and the difference between them

can be assumed as negligible. Further, following the simplistic concept of normalized difference vegetation index (NDVI) applied in vegetation status monitoring; NDCI was developed by taking the spectral band difference at 708 nm and 665 nm and normalizing by the sum of their reflectance to eliminate any uncertainties in the estimation of R_{rs}, seasonal solar azimuth differences, and atmospheric contributions at those wavelengths. NDCI is formulated as:

$$C_{chl-a} \propto \frac{[R_{rs}(708) - R_{rs}(665)]}{[R_{rs}(708) + R_{rs}(665)]} \tag{6.6}$$

The overarching objective of this research was to improve the accuracy of chl-*a* retrieval in turbid productive waters using a simple, easy to implement, intuitive (such as NDVI for vegetation), and universal model. Throughout this paper, these criteria have been tested using several steps including: (1) developing a dataset simulating a wide range of bio-optical parameters to examine the conceptual model, (2) testing the model using an *in situ* dataset collected from a global bio-optical data archive and corresponding MERIS data, (3) evaluating the performance of several existing chl-*a* algorithms for turbid productive waters using the simulated and remotely sensed datasets, (4) applying the model in three unique study regions, such as Chesapeake-Delaware Bay, the Mississippi River Delta, and the Mobile Bay, and (5) finally and most importantly, developing a generalized but practical relationship between NDCI values and chl-*a* range in an attempt to make NDCI intuitive and applicable when/where ground truth data is not available.

6.3 Data and methods

6.3.1 Bio-optical modeling

The simulated R_{rs} spectra used for the model conception, calibration, and validation were approximated by the following method. Irradiance reflectance, $R(z,\lambda)$, is defined as (Morel & Prieur, 1977):

$$R(z,\lambda) = \frac{E_u(\lambda,z)}{E_d(\lambda,z)}, \tag{6.7}$$

where, $E_u(\lambda,z)$ and $E_d(\lambda,z)$ are upwelling irradiance (W m^{-2} nm^{-1}) and downwelling irradiance (W m^{-2} nm^{-1}) at depth (z) in the water column. Gordon *et al*. (1975) have approximated $R(0,\lambda)$ at the water surface as a function of the in-water absorption and scattering coefficients as:

$$R(0,\lambda) = f\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)} \tag{6.8}$$

where, $b_b(\lambda)$ is the total backscattering coefficient (m^{-1}); $a(\lambda)$ is the total absorption coefficient (m^{-1}); and f is the proportionality factor that depends on the solar zenith angle and light field geometry. Further, Kirk (1984) modeled f as a function of the cosine of the solar zenith angle (μ_0) of the refracted photons as:

$$f(\mu_0) = -0.629\mu_0 + 0.975 \tag{6.9}$$

$R(0,\lambda)$ is closely related to the spectral remote sensing reflectance, $R_{rs}(\lambda)$ as (Carder & Steward, 1985):

$$R_{rs}(\lambda,0) = \frac{L_w(\lambda,0)}{E_d(\lambda,0)} \tag{6.10}$$

where, $L_w(\lambda,0)$ is the water leaving radiance (W m^{-2} sr^{-1} nm^{-1}); and $E_d(\lambda,0)$ is the downwelling irradiance (W m^{-2} nm^{-1}) above the air-water interface. For a nadir looking sensor, L_w can be estimated from L_u as in Mobley (1999):

$$L_w(\lambda,0) = \tau L_u(\lambda,0) \tag{6.11}$$

where, τ is a non-dimensional proportionality factor that relates upwelling radiance measured just below the water surface to water leaving radiance. For most remote sensing applications, τ can be reasonably approximated as 0.54 (Mobley, 1999). $L_u(\lambda, 0)$ is the subsurface upwelling radiance and for uniform angular distribution, L_u can be formulated as (Jerlov, 1968):

$$L_u(\lambda,0) = E_u(\lambda,0)/Q \tag{6.15}$$

where, $E_u(\lambda,0)$ is upwelling irradiance (Wm^{-2} sr^{-1} nm^{-1}) and Q is the angular distribution factor of spectral radiance and assumed to be 4 based on the average solar zenith angle in our study area (Morel & Gentili,1996). Finally, combining Eqs. 5.6-5.11, R_{rs} can be written as follows:

$$R_{rs}(\lambda) = 0.0448\frac{b_b(\lambda)}{(a(\lambda)+b_b(\lambda))} \tag{6.13}$$

To model the fluorescence component of R_{rs}, water leaving radiance due to chlorophyll fluorescence at 685 nm ($L_{w,fl}(685)$) was modeled as in Gilerson *et al*. (2007). $R_{rs,fl}(685)$ was estimated as:

$$R_{rs,fl}(685) = \frac{L_{w,fl}(685)}{E_d(685)} \tag{6.14}$$

Based on the Hydrolight (Mobley & Sundman, 2001) simulation as in Gilerson *et al*. (2007) for a clear sky condition, E_d was considered as 1.1 Wm^{-2} nm^{-1}. Finally, $R_{rs,fl}$ was modeled at every wavelength using a Gaussian peak centered at 685 nm with a standard deviation (STD) of 10.6 nm (Mobley,1994). The final $R_{rs}(\lambda)$ values were modeled as a sum of R_{rs} estimations from Eq. 6.12 and Eq. 6.13.

6.3.2 Simulation of R_{rs} data

Using the above bio-optical model, two R_{rs} datasets representing four case 2 water bodies such as Mississippi Delta region, Mobile Bay, Chesapeake Bay, and Delaware Bay were selected. Further, assuming that there are three optically active constituents in the water (i.e., phytoplankton, suspended matter, and CDOM), the total absorption, a (λ) coefficients can be written as follows:

$$a(\lambda) = \sum_{i=1}^{n} a_i(\lambda), \tag{6.15}$$

where, a_i are the absorbing components such as water (a_w), chl-*a* (a_{chl}), CDOM (a_{CDOM}), and non-algal particle (a_{nap}). Similarly, total back-scattering coefficient can be written as:

$$b_b(\lambda) = \sum_{i=1}^{n} b_{b,i}(\lambda) \tag{6.16}$$

where, $b_{b,i}$ are the back-scattering components such as water (b_{bw}) and total particulate matter (b_{bp}). Phytoplankton absorption coefficient was modeled as the product of specific absorption coefficient (a_{φ}^{*}) and the chl-*a* concentration ($C_{chl\text{-}a}$). a_{φ}^{*} values were taken from Ciotti *et al*. (2002) as a sum of specific absorption coefficient of micro-

planktons and pico-planktons with different weighting factor, S_f, ranging from 0.1 to 0.5 that corresponds to typical turbid productive coastal waters (Ciotti *et al*., 2002).

$$a_{\varphi}^{*}(\lambda) = S_f \cdot a_{pico}^{*}(\lambda) + \left(1 - S_f\right) \cdot a_{micro}^{*}(\lambda) \tag{6.17}$$

$a_w(\lambda)$ and $b_{b,w}(\lambda)$ values were taken from Pope & Fry (1997). CDOM absorption was expressed as a function of the absorption coefficient at 440 nm, a_{CDOM} (440), and a slope factor, S_{CDOM}, as follows (Bricaud *et al*., 1981):

$$a_{CDOM}(\lambda) = a_{CDOM}(440)e^{-S_{CDOM}(\lambda - 440)} \tag{6.18}$$

Similar to CDOM absorption, a_{nap} was expressed as an exponentially decaying function with respect to the wavelength as:

$$a_{nap}(\lambda) = a_{nap}(443)e^{-S_{nap}(\lambda - 443)} \tag{6.19}$$

where, S_{nap} = 0.0123 nm^{-1} was taken from Babin *et al*. (2003). $a_{nap}(443)$ was expressed as:

$$a_{nap}(443) = (0.031)(0.81)C_{chl-a} \tag{6.20}$$

where, 0.031 and 0.81 are the mass-specific a_{nap} coefficient and C_{nap}:$C_{chl\text{-}a}$ ratio respectively. This relationship explains the observed co-variation between a_{nap} and $C_{chl\text{-}a}$ (Babin *et al*. 2003). Particulate back-scattering coefficient was expressed as:

$$b_{bp}(\lambda) = b_{b,p}^{*}(550)\left(\frac{550}{\lambda}\right)^{y} C_p \tag{6.21}$$

where, $b_{b,p}^{*}(550)$ is the specific back-scattering coefficient of particulate matter (0.0086 m^{2}g^{-1}) (Kiefer & Reynolds, 1992), y is the spectral slope of $b_{b,p}$ (set to 0 as in Dall'Olmo *et al*., 2006), and C_p is the concentration of particulate matter.

Previously published ranges of $a_{CDOM}(440)$ and concentrations of inorganic suspended sediment (ISS) resembling each study region were used in the bio-optical model (Kutser *et al*., 2009; Miller *et al*., 2002; Tzortziou *et al*., 2007) (Table 6.1). Assuming optical similarity, published values from Mississippi Sound have been considered for the Mobile Bay. The range of chl-*a* concentration used for the simulation widely varied from 1 to 60 mg m^{-3}. Concentrations and values of model parameters, such as chl-*a*, ISS, $a_{CDOM}(440)$, and S_{CDOM} were randomly varied at each iteration step to mimic the natural variability in the study regions. R_{rs} data were simulated at 1 nm interval from 400 nm to 760 nm (n=200). R_{rs} measurements were further simulated at each MERIS band centers λ_i, for band 1 to 10, by taking the weighted average of each R_{rs} spectra using the spectral response function (SRF(λ)) of MERIS (http://earth.eo.esa.int/pub/ESA_DOC/MERIS_Wavelengths_and_Irradiances_Model2004.xls) as weights (Eq. 6.20). Finally, the simulated MERIS spectra were used for further analysis.

$$R_{rs}(\lambda_i) = \frac{\sum_\lambda R_{rs}(\lambda) SRF(\lambda)}{\sum_\lambda SRF(\lambda)} \tag{6.22}$$

6.3.3 Field data

The concept of NDCI was developed using simulated data, however, field datasets acquired from the same geographic regions were used to further validate the NDCI and chl-*a* relationship. The field datasets consisted of chl-*a* concentrations from Chesapeake Bay, Delaware Bay, the river Mississippi Delta region, and the Mobile Bay (Fig. 6.1). However, the field data were not used in the simulation because of the absence of a wide chl-*a* range, and the unavailability of other data such as inorganic suspended

solid concentration (ISS), $a_{CDOM}(\lambda)$, and S_{CDOM} that are required as input to the bio-optical model. *In situ* chl-*a* data from Chesapeake Bay and Delaware Bay (n = 38; collection dates: April, 15, 16, and 18, 2008, May 14, 2008) were downloaded from SeaWiFS Bio-Optical Archive and Storage System (SeaBASS) archives. Similarly, chl-*a* data for Mississippi Delta region (n=10; collection date: May 19, 2007) were acquired from the NASA bio-Optical Marine Algorithm Data (NOMAD) archive. Both SeaBASS and NOMAD archives contain high quality bio-optical global datasets which are suitable for calibrating and validating ocean color algorithms (Werdell & Baily, 2005). *In situ* data from Mobile Bay were collected and analyzed using high performance liquid chromatography (HPLC) (by Hugh MacIntyre, Dolphin Island Sea Lab) (n=8; collection date: Nov 07, 2007)**.** The frequency plot of chl-*a* observation in all field sites is presented in Fig. 6.2 and the summary of chl-*a* measurements, solar zenith and solar azimuth angles of the study sites is also summarized in Table 6.2.

6.3.4 Satellite Data

Simultaneous observations of full resolution level 2 data acquired by MERIS sensor onboard ENVISAT were downloaded using European Space Agency's client EOLI (Earth Observation Link) for NDCI model calibration and validation. Beam 3.6 software (Brockmann Consult, Geesthacht, Germany) was used to process and analyze the MERIS images. Image data acquired on April 15 2008 (Universal Time Co-ordinates-15:43:39.476) over Chesapeake Bay had some cloud cover and was masked for land and cloud. However, cloud cover over the Mississippi River Delta region and the Mobile Bay was minimal. Corresponding MERIS images over Chesapeake and Delaware Bay was not available on August 16, 2008 and therefore, image acquired on August 15, 2008 was

used. Images acquired on April 18, 2008 (UTC- 15:49:21.537) and May 14, 2008 (UTC-15:32:21.458) over Chesapeake Bay were also used in this study. Similarly, images for the Mississippi Delta and the Mobile Bay were acquired on May 19, 2007(UTC-16:20:52.509) and November 07, 2007 (UTC-16:15:06.954).

MERIS level 2 products are atmospherically corrected for normalized water leaving reflectance. Corresponding R_{rs} spectra were extracted from MERIS images for the *in situ* sampling locations and are presented (Fig. 6.3B). R_{rs} spectra showed high variability in magnitude in the visible spectral domain. Maximum values of R_{rs} were observed at the green channel (559 nm) and maximum variability was also occurred at 559 nm and 620 nm. As expected, R_{rs} in the blue spectral region showed lower reflectance because of high absorption by chlorophylls, CDOM, and non-algal particulate matters in the water. A large number of pixels showed negative R_{rs} at 412.5 and 442.5 nm possibly because of over correction for the atmospheric scattering. The spectral shape of the average R_{rs} spectrum were similar to the simulated data (Fig. 6.3A)

6.3.5 Model calibration and validation

NDCI model was calibrated and validated using a simulated and a field dataset. A one-fold calibration and validation was performed using the simulated dataset, whereas, a three-fold calibration and validation was performed using the field dataset. The three-fold calibration and validation to the field data was based on three varying parameters including (a) solar zenith angle (θ_s), (b) solar azimuth angle (φ_s), and (c) geographic region. In the first two calibration and validation, the field data sorted out based on solar angle parameters (θ_s , φ_s) in order to test the robustness of the model to variations observed in satellite data because of atmospheric interferences and seasonal changes, and

transferability of the model to other similar coast water bodies. In the third calibration and validation, the data were sorted out based on geographic region in order to maintain the independence aspect of the field data.

The first calibration and validation dataset was sampled after sorting based on descending θ_s and further dividing the data into two subsamples, one for calibration (n=29) and the other for validation (n=20). Similarly, the second calibration and validation was performed after sorting the field data based on descending φ_s and further dividing into a calibration dataset (n=29) and a validation dataset (n=20). Remote sensing reflectance (R_{rs}), being an apparent optical property, is prone to vary with any change in the light field geometry and atmospheric conditions even though the concentrations of all the biophysical variables (TSS, chl-*a*, CDOM) remain constant in the water column. While formulating NDCI, it was hypothesized that a normalized band difference index will be less sensitive to any uncertainties because of variations in light field geometry, atmospheric effects, and radiometric calibration differences by virtue of its spectral band selection and model architecture. In the current scenario, solar zenith (θ_s) and solar azimuth angles (φ_s) have been used as a measure to represent seasonal and spatial change in light field as these solar angles change with geographic regions, and with season in a geographic region.

Finally, the third set of calibration and validation was performed by dividing the dataset from all study regions into two subsets based on geographic regions. Calibration dataset contained all sample points from the Chesapeake Bay and Delaware Bay (n=35), whereas, sample points from the Mississippi Delta and the Mobile Bay were used as the validation dataset (n=14). In this way, the model calibrated for a region with a specific

set of bio-optical and physical characteristics will be validated for a different region with a different set of bio-optical and physical parameters demonstrating the strength and the transferability of the algorithm to other coastal regions.

For the model calibration, best fit functions were calculated based on least-squares regression analysis. Using the calibrated equations, chl-*a* concentrations were predicted for the validation dataset. The accuracy of the model prediction was assessed by comparing the predicted chl-*a* and the measured chl-*a* concentration. The comparison was expressed in terms of root mean squared error (RMSE), coefficient of determination (R^2) between measured and predicted chl-*a*, the slope of the best fit line (m), and the mean ratio that was calculated as the average of the ratios of the predicted chl-*a* values to the measured chl-*a* values. RMSE was computed as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2}{n}} \tag{6.23}$$

where n is the number of observations, $\hat{Y}$ is the predicted value of chl-*a*, Y is the observed or measured value of chl-*a*, and k is the number of predictors

6.4 Results

6.4.1 Model calibration and validation using simulated data

Several existing blue-green band ratio algorithms such as OC4 (O'Reilly *et al*., 1998, 2000) and two-band red-NIR models (T07, M09) were compared with NDCI during model calibration. Three-band models that use R_{rs} values at 753 nm and beyond were not considered for comparison because the absorption coefficients of water and phytoplankton available in literature do not produce reliable estimates of R_{rs} beyond 750 nm (Kutser, 2006). Model calibration results using the blue-green band ratio showed high

sensitivity to the changes in a_{CDOM} and $a_{nap.}$. Therefore, developing a single empirical relationship using blue-green band ratios during model calibration was not possible because of the absence of a common trend line for all study areas (not shown). In contrast, red-NIR based models showed a single trend for all study areas indicating their insignificant sensitivity to the changes in a_{CDOM} and a_{nap}.

Various linear and nonlinear trend lines were fitted to the data and best outputs were finalized and reported for all models (Table 6.3). Relationship between NDCI and chl-*a* was essentially nonlinear. However, logarithmic and power trend lines could not be fitted to the data because of the negative NDCI values. An exponential function with two-parameters explained 93% of variance in the data (R^2=0.93, standard error of the estimate (STE) =4.38 mg m^{-3} of chl-*a*, p<0.0001) where as a second order polynomial produced the highest R^2 and lowest estimation error (R^2=0.95, STE=3.62 mg m^{-3} of chl-*a*, p<0.0001) in the simulated dataset (Table 6.3). Quadratic polynomial equation was finally selected as the final NDCI-chl-*a* equation because of higher R^2 and lower STE. Unlike NDCI, linear relationships were found between the two-band models and chl-*a* during model calibration. STE and R^2 for M09 were 3.76 mg m^{-3} and 0.95 (p<0.0001) respectively. T07 showed a weak relationship with chl-*a* in the model calibration stage with STE and R^2 of 10.76 mg m^{-3} and 0.61 (p<0.0001) respectively.

All calibrated models were examined for their predictive ability and transferability to other geographic regions (Table 6.4). Using the calibration equation, chl-*a* values were predicted for an independently modeled dataset and were compared with actual chl-*a* concentrations. NDCI produced a root mean squared error (RMSE) of 4.83 mg m^{-3}, whereas, M09 and T07 produced RMSE of 5.26 and 21.78 mg m^{-3}

respectively. NDCI produced the most accurate prediction showing highest coefficient of determination (R^2=0.93) between actual and predicted chl-*a* and the slope of the regression line was close to 1 (m=1.05). On the simulated dataset, M09 performance was similar to NDCI with R^2=0.92 and m=1.07 and T07 model predicted significantly inaccurate values. The residuals from NDCI and M09 validation did not reveal a trend of over or under estimation; however, residuals from T07 had a clear trend of over estimation (not shown).

6.4.2 Model calibration and validation using field data

Two sets of calibration/validation were performed on selected models based on the solar zenith and solar azimuth angles on MERIS data from all study regions (Table 6.3). Similar to the simulated data calibration results, in the field data calibration based on solar zenith angle, NDCI showed a strong relationship with chl-*a* concentration (quadratic function) producing R^2 and STE of 0.90 and 2.49 mg m^{-3} respectively (p<0.0001) (Fig. 6.4). A two parameter-exponential trend line was also fitted that explained 87% of variance in the data (R^2=0.87, STE=2.7 mg m^{-3} of chl-*a*, p<0.0001). However, quadratic polynomial equation was finally selected because of higher R^2 and lower STE. M09 performed as the second best model with R^2 of 0.82 and STE of 3.24 mg m^{-3} (p<0.0001), however, the performance gap between NDCI and M09 was found to be higher in field data compared to the simulated data. Unlike the expected linear relationship between D05 and chl-*a* (Dall'Olmo & Gitelson, 2005), a quadratic function explained the maximum variance in the data (R^2 =0.56, p<0.0001) producing a STE of 5.22 mg m^{-3}. T07 model produced the least R^2 (0.48) and highest STE value (5.53 mg m^{-3}, p<0.0001).

All four models plus G08 and MERIS level-2 chl-*a* product (Algal-2) were validated using the corresponding validation dataset (Fig. 6.5; Table 6.4). G08 was not parameterized and the chlorophyll equation from Gons *et al*. (2008) was used for validation. NDCI performed the best with the least RMSE (1.87 mg m^{-3}), high R^2, $m \approx 1$, and no clear residual trend of over or under-estimation. T07 produced the highest RMSE (10.01 mg m^{-3}) and the least R^2. Performance of the 3-band model, D05, was close to NDCI in the validation stage producing RMSE of 1.97 mg $m^{-3}$and M09 showed a clear trend of over estimation (Fig. 6.5). G08 had a RMSE of 4.06 mg m^{-3} and showed a trend of overestimation for all predictions greater than 7 mg m^{-3}. MERIS product (Algal-2) showed excellent prediction for chl-*a* concentration below 13 mg m^{-3} with a few exceptions (Fig. 6.5).

Very similar trend in results were observed with the second calibration dataset based on solar azimuth angle (Table 6.3). NDCI showed the strongest relationship with chl-*a* (R^2=0.90, STE= 2.11 mg m^{-3}, p<0.0001), whereas, M09 explained 82% variation in the data and produced a STE of 2.61 mg m^{-3} (p<0.0001). Neither a linear nor a quadratic function explained the relationship between D05 and chl-*a* very well, therefore the R^2 was low and the STE was high. Similarly, T07 model showed weakest relationship with chl-*a* producing lowest R^2 and highest STE values (p=0.0134). Validation results showed that the MERIS Algal-2 chl-*a* product produced the least RMSE (1.69 mg m^{-3}), however, the R^2 between predicted and actual was low and the regression slope was significantly lower than 1 (Table 6.4). In contrast, NDCI produced a RMSE of 2.04 mg m^{-3} with the regression slope near 1. D05 and G08 produced a RMSE of 2.46 and 3.56 mg m^{-3} respectively. Also, G08 produced high estimation errors for low chl-*a* values

(oligotrophic waters) with five negative prediction. Similar to the previous validation, T07 model produced the highest RMSE and proved to have the least predictive ability.

Finally, the last set of calibration and validation was performed on the datasets those were sampled based on geographic regions. Similar to the previous calibration results, NDCI showed the strongest relationship with measured chl-*a* (R^2= 0.72, STE = 2.15, p <0.0001) and the weakest relationship was found between T07 and chl-*a* (R^2=0.01, p=0.439) (Table 6.3, Fig. 6.6). Note that the third validation dataset contained four points with chl-*a* concentration higher than the chl-*a* concentration range used in the calibration. Therefore, the validation results for all models are shown in two ways including (a) data points within the calibration range (values outside the brackets) and (b) the entire dataset (values inside the bracket) (Table 6.4, Fig. 6.7). Validation results show that NDCI was successful predicting chl-*a* concentration with highest accuracy producing a RMSE of 1.43 (2.37) mg m^{-3}. The R^2 between the measured and predicted chl-*a* was 0.94 with a regression slope of 0.88 (0.91). Algal-2 showed maximum prediction error with two-fold over estimation. It is also evident that, unlike other calibrated models, predictions by the NDCI equation outside its calibration range was also very accurate and the equation can be reliably used to predict chl-*a* up to 28.17 mg m^{-3} (Fig. 6.7).

It should be noted that in the simulated and the first field dataset, NDCI showed the most accurate predictive ability overall with no trend in residuals (Fig. 6.8). However, in the second validation dataset three different models performed best in three different categories. G08 produced the highest R^2, Algal-2 produced the least RMSE, and NDCI produced a regression slope closest to 1. Similarly, NDCI showed excellent prediction in

the third validation dataset producing lowest RMSE, highest R^2 between predicted and actual chl-*a* and *m* close to 1. However, if all the three validation parameters are considered simultaneously, NDCI showed overall better predictive ability (Table 6.4, Fig. 6.8). Thus, the three-fold calibration resulted three sets of chl-*a* equations with coefficients very close to each other. The difference among them causes minimal impact on the prediction accuracy of the models. It has been verified by applying all three NDCI equations on an independent validation dataset (validation dataset 3) and the results shown are very encouraging. All three equations produced a RMSE very close to 2.37 mg m^{-3} with a STD of 0.127 mg m^{-3} mg m^{-3} of chl-*a* (not shown).

6.4.3 Chlorophyll mapping using MERIS data

MERIS images acquired on April 15, 2008 were selected for the Chesapeake and Delaware Bay for mapping spatial distribution of chl-*a.* Image preprocessing steps included georeferencing and land and cloud masking. Pixels with inaccurate reflectance values (mostly negative values) and Open Ocean or case 1 areas were masked out . NDCI chl-*a* equation generated from the first calibration dataset (based on solar zenith) was applied on MERIS images from Chesapeake and Delaware Bay, Mississippi Delta region, and the Mobile Bay to prepare chl-*a* distribution maps. As shown in validation results, NDCI was successful in predicting chl-*a* concentration with a 12% overall bias for all regions (Table 6.5). chl-*a* estimation accuracy for Chesapeake Bay and the Mississippi Delta were within 5-7 % of the measured *in situ* values, whereas, the bias in Mobile Bay was 32% (Table 6.5). Overall, MERIS derived chl-*a* distribution maps were consistent with the published chl-*a* levels for the Chesapeake Bay and the Mississippi Delta. The maximum chl-*a* concentration value mapped on the images was 59.08 mg m^{-3} (Figs. 9

and 10). For further analysis, all case 2 water pixels were sampled and chl-*a* values were analyzed. Frequency plot of the mapped chl-*a* in Chesapeake and Delaware Bay shows that the chl-*a* concentration in the majority of the bay was within 7-25 mg m^{-3} on July 15, 2008 (Figs. 9, 10C). In the upper Chesapeake Bay and along the shorelines, chl-*a* concentration was comparatively higher. All tributary rivers in the middle Bay region also showed higher chl-*a* concentration (Fig. 6.9A).

Significant portions of the MERIS image acquired on May 19, 2007 covering Mississippi Delta were of poor quality. Most of the pixels were flagged by MERIS atmospheric correction scheme with inaccurate reflectance measurements and therefore masked out. The maximum chl-*a* concentration observed in the map was 58.9 mg m^{-3} (Fig. 6.9B) and the frequency plot revealed that most of the pixels had chl-*a* values within 10-30 mg m^{-3} (Fig. 6.10B). High concentration of chl-*a* was observed in the eastern boundary of the delta.

The highest level and the range of frequently occurring chl-*a* concentration recorded on the Mobile Bay map derived from the cloud free MERIS image were 58.46 mg m^{-3} and within 7-20 mg m^{-3} respectively (Fig. 6.10C). As expected, north-shore of the bay showed the highest chl-*a* concentration. Because the Mobile River and the Tensaw River drains nutrient-rich water which stimulates the growth of phytoplankton and primary production making it the most productive part of the estuary.

6.5 Discussion

6.5.1 Algorithm performance

Accurate quantification and mapping of chl-*a* concentration in turbid productive waters using remote sensing data can create enormous opportunities for biogeochemists

and climate scientists to understand the functioning of global nutrient cycles. As discussed earlier, chl-*a* mapping in turbid productive waters is often challenging because of the effect of other optically active constituents such as CDOM, detritus, and mineral particles whose concentration do not co-vary with chl-*a* (Morel & Prieur, 1977). Performances of NDCI and M09 model on the simulated calibration dataset were very similar although the standard error of the estimation for NDCI was less than the M09 model. However, NDCI produced higher R^2 and less STE than the M09 in the field dataset. Both algorithms outperformed T07 at the calibration stage. The simulated dataset was produced by varying all possible bio-optical parameters mimicking the natural variations in the Chesapeake Bay, Mississippi Delta, and the Mobile Bay. The basic difference between the two datasets was the possible existence of the remnant atmospheric contamination in the field dataset. This implies that M09 is probably more sensitive to the atmospheric parameters than NDCI and thus produced lower R^2 and higher STE with the field dataset. Since T07 uses a green channel at 560 nm, it is highly sensitive to CDOM and detritus in the water and thus produced the highest STE and RMSE in all stages of calibration and validation (Table 6.3 and 6.4).

Consistent performance of NDCI among all models in all study regions was evident from the lowest STD of the mean ratio of 0.29 (Table 6.5). In contrast, MERIS Level2 chlorophyll product (Algal-2) performed poorly in all study regions with highest STD of 0.65. NDCI also produced the least bias, e.g., 7% and 15% in the Cheasepeake-Delaware Bay and in the Mississippi Delta respectively. Similarly, NDCI produced a mean ratio of 1.32 with the lowest STD of 0.28 in the Mobile Bay (Table 6.5).

6.5.2 Chlorophyll mapping using MERIS data

R_{rs} spectra sampled from the MERIS images were very similar in shape but higher in magnitude when compared to some of the published spectra from the same regions (Darecki *et al*., 2003; Dall'Olmo & Gitelson, 2005, Moses *et al*., 2009). This is because of the differences in the interpretation of R_{rs}. For example, Simis (2006) considered normalized water leaving reflectance, the radiometric quantity available from MERIS Level 2 products, $[\rho_w]_N$, as an equivalent of R_{rs}. However, if we notice the formulation of both radiometric measurements, $[\rho_w]_N$ is a product of R_{rs} and a constant, π, which could be the reason of higher magnitude of sampled spectra from MERIS images. Most of the spectra from the Chesapeake Bay exhibited an absorption feature centered at 620 nm (MERIS channel 6) which implies the presence of phycocyanin in the bay waters. Phycocyanin is widely accepted as a characteristic photopigment in cyanobacteria. Thus, the above mentioned distinct optical signature implies abundance of cyanobacteria in the Chesapeake Bay during summer months. Spectra from the river Mississippi Delta region and the Mobile Bay lacked this optical feature. In all study regions, some pixels produced extremely inaccurate estimation of chl-*a* (extreme negative and positive values). Reflectance spectra were randomly extracted from those extreme pixels and analyzed. All spectra showed unusual spectral shape and negative reflectance values at several spectral channels (Fig. 6.11). Those contaminated pixel values are believed to be originated from inaccurate atmospheric correction scheme. Those contaminated pixels were masked out from the final chlorophyll map products.

6.5.2.1 Chl-*a* and NDCI relationship

One of the biggest advantages of NDCI is that its range varies between -1 to +1 for areas with no cloud cover and adjacency and bottom effects. Therefore, qualitative chl-*a* mapping (such as NDVI for vegetation) and bloom detection using satellite data is possible for remote areas where field data is unavailable or unusable. Further analysis of the relationship between NDCI and chl-*a* helped us to associate an approximate chl-*a* range with certain NDCI values which has tremendous application and will make mapping of chl-*a* more accurate for remote areas. Based on the absorption properties and the spectral band structure of NDCI, in optically clear water bodies NDCI is expected to hold values closer to -1. NDCI values in water bodies with moderate to high algal biomass are expected to vary in the range of -0.3 to close to 1. In case of algal blooms with surface scum on water bodies, NDCI values would vary within a range of 0.5 to 1 (Table 6.6). To reinforce the validity of the NDCI and chl-*a* range relationship, MERIS images of two entirely different areas: 1) a mesotrophic water body (Lake Pontchartrain, LA, USA), and 2) an eutrophic inland water body (Lake Apopka, FL, USA) were considered. Lake Pontchartrain is a large, shallow, oligohaline, and semienclosed estuary located in southeastern Louisiana, USA. MERIS image acquired on October 14, 2010 is shown in true color composite and corresponding NDCI map is also shown (Fig. 6.12A and B). NDCI values ranged between 0.4 to around 0 qualitatively suggesting chl-*a* concentration below ~20 mg m^{-3} which is evident from the true color composite (Fig. 6.12A and B). On the other hand, lake Apopka is a large (surface area = 124 km^2), shallow (mean depth=1.7m), hyper eutrophic inland lake located in central Florida, USA and notorious for high nutrient concentrations and phytoplankton biomass (Carrick &

Schelske, 1997). During summer months (April-July), chlorophyll concentration in the lake Apopka reaches as high as 105 mg m^{-3} (Carrick & Schelske, 1997). MERIS image of the lake acquired on April 29, 2010 is shown in true color and along with the corresponding NDCI map (Fig. 6.12C and D). The true color composite shows a severe phytoplankton algal bloom in the lake that gives the lake water a deep green hue such as terrestrial vegetation. Most frequently occurring NDCI values in the lake varies from 0.4 to 0.6 implying a severe bloom condition in the lake corresponding very well with data from literatures. Similarly, severe algal bloom was evident in the Lake Harris and the Lake Eustis in the same MERIS image (Fig. 6.12C).

6.5.2.2 Possible sources of estimation error

Errors associated with atmospheric correction of the MERIS imageries produced extreme negative and positive estimations of chl-*a* over a few pixels in all study regions. In ocean color remote sensing studies, accuracy of the atmospheric correction scheme controls the accuracy of the mapped biophysical variable because the errors from the atmospheric correction stage are propagated to the final product. Based on the spectral band architecture of NDCI, it is clear that it would predict higher chl-*a* values upon increase in the difference between $R_{rs}(708)$ and $R_{rs}(665)$. Any disproportionate increase or decrease in reflectance at 665 and 708 nm creates inaccurate estimations of chl-*a*. For example, pixels with disproportionate increase in reflectance at 665 nm as compared to 708 nm will produce extreme negative values ; similarly disproportionate increase in reflectance at 708 nm will estimate extreme positive values of chl-*a*.

Seven *in situ* data points sampled from the Chesapeake and Delaware Bay on July 16, 2008 did not have corresponding MERIS image and used MERIS reflectance product from July 15, 2008. Next day image was used based on the assumption that the overall concentration and spatial distribution of biomass does not change within a day. A similar assumption has also been made while using MERIS data previously (Gons *et al*., 2008; Moses *et al*., 2009). In reality, this assumption might not hold true as the water circulation in estuarine and coastal environment is very dynamic in nature. This might have caused some uncertainties in model calibration and validation.

6.6 Conclusion

A new index, NDCI has been proposed to predict chl-*a* concentrations in optically complex turbid productive waters from remotely sensed data. Accuracy and potential applicability of the model has been assessed by extensively calibrating and validating it on: (1) simulated, and (2) MERIS data representing three unique study areas with a wide range of trophic status and optical complexity. The results presented in this research illustrate the potential of NDCI to quantify chl-*a* concentration when used with remote sensing reflectance data from MERIS sensor. A quantitative interpretation of NDCI values has also been analyzed and presented that could be potentially useful for remote areas with no field data availability. This makes NDCI widely applicable to coastal waters as NDVI to terrestrial vegetation.

Table 6.1 Ranges of aCDOM (440) (m^{-1}), S_{CDOM} and ISS (mg l^{-1}) from all study regions used in the bio-optical modeling of $R_{rs}(\lambda)$.

Study Regions	a_{CDOM} (440) (m^{-1})	S_{CDOM}	ISS (mg l^{-1})
Mississippi River Delta	0.05-0.07	0.016	2-5
Chesapeake Bay and Delaware Bay	2.0-5.0	0.016	5-10
Mississippi Sound/Mobile Bay	3.13-4.27	0.016	2-5

Table 6.2 Descriptive statistics of chl-*a* (mg m^{-3}), solar zenith and solar azimuth angles in the study regions.

Parameters	Min	Max	Average
Mississippi Delta, 19th May 2007 (n = 6)			
Chl-*a* (mg m^{-3})	14.351	28.175	21.042
Solar Zenith Angle	23.303	23.393	23.349
Solar Azimuth Angle	108.306	108.709	108.495
Chesapeake Bay and Delaware Bay (n = 35)			
Chl-*a* (mg m^{-3})	0.903	16.061	7.255
Solar Zenith Angle	26.958	34.196	31.492
Solar Azimuth Angle	131.943	144.597	140.914
Mobile Bay (n = 8)			
Chl-*a* (mg m^{-3})	4.208	6.395	5.256
Solar Zenith Angle	50.353	50.577	50.471
Solar Azimuth Angle	154.587	154.915	154.807

All angular measurements are in degree.

Table 6.3 Model calibration: all model parameters including a_0, a_1, and a_2 with corresponding standard error of estimate (STE) are provided; In case of linear regression, a_0 and a_1 correspond to intercept and slope of the fitted equation.

Indices	a_0	a_1	a_2	R^2	Adj. R^2	STE of estimate	*P*
Simulated Dataset (n = 100)							
NDCI	42.197	236.5	314.97	0.95	0.95	3.62	<0.0001
M09	-64.055	106.335	*	0.95	0.95	3.76	<0.0001
T07	-39.739	102.717	*	0.61	0.61	10.76	<0.0001
Field Dataset (Solar Zenith Angle, n = 29)							
NDCI	14.039	86.115	194.325	0.90	0.89	2.49	<0.0001
M09	-15.617	31.133	*	0.82	0.81	3.24	<0.0001
D05	14.07	177.56	808.03	0.56	0.52	5.22	<0.0001
T07	-1.832	26.56	*	0.48	0.47	5.53	<0.0001
Field Dataset (Solar Azimuth Angle, n = 29)							
NDCI	14.279	79.607	181.45	0.90	0.90	2.11	<0.0001
M09	-15.992	31.196	*	0.85	0.84	2.61	<0.0001
D05	14.15	156.88	769.86	0.49	0.45	4.93	0.001
T07	4.643	15.473	*	0.20	0.17	6.08	0.0134
Field Dataset (Chesapeake Bay and Delaware Bay, n=35)							
NDCI	13.55	87.99	212.6	0.72	0.7	2.15	<0.0001
M09	-8.88	20.96	*	0.59	0.58	2.57	<0.0001
D05	11.52	136.13	666.46	0.43	0.4	3.09	<0.0001
T07	6.0	3.164	*	0.01	0.0	4.02	0.4394

* Not Applicable

Coefficient of determination (R^2), adjusted R^2 and p values of the regression models are also provided.

Table 6.4 Three-fold model validation results: root-mean-square-error (RMSE) in mg m^{-3}, coefficient of determination (R^2), and the slope of the regression line (m) are reported for all models.

Indices	RMSE (mg m^{-3})	R^2	m
Simulated Dataset (n = 100)			
NDCI	4.83	0.93	1.05
M09	5.26	0.92	1.07
T07	21.78	0.26	0.57
Field Dataset (Solar Zenith Angle, n = 20)			
NDCI	1.89	0.80	1.005
M09	3.27	0.80	1.115
D05	1.97	0.81	0.795
T07	10.013	0.54	-0.544
G08	4.066	0.83	0.574
MERIS	5.856	0.74	0.403
Field Dataset (Solar Azimuth Angle, n =16)			
NDCI	2.04	0.34	0.64
M09	3.08	0.31	0.22
D05	2.46	0.06	0.34
T07	8.01	0.07	0.09
G08	3.56	0.48	0.22
MERIS	1.69	0.33	0.44
Field Dataset (Mobile Bay and Mississippi Delta, n =14)			
NDCI	1.43 (2.37)	0.94 (0.92)	0.88 (0.91)
M09	2.82 (5.17)	0.92 (0.94)	0.49 (0.45)
D05	2.69 (6.49)	0.91 (0.62)	0.51 (0.45)
T07	5.49 (10.08)	0.53 (0.04)	-0.06 (-0.01)
G08	4.95	0.92	1.43
MERIS	7.3	0.94	2.29

The RMSE, R^2, and m values of the predictions within the calibrated chl-*a* range (values outside the brackets) and the RMSE, R^2, and m values of the entire dataset (values inside the bracket) are provided.

Table 6.5 Comparison of mean ratio (STANDARD DEVIATION) of modeled and measured chl-*a* for all field regions combined, Chesapeake Bay, Mobile Bay, and Mississippi Delta region.

Models	All Regions	Chesapeake and Delaware Bay	Mobile Bay	Mississippi Delta
NDCI	1.12 (0.29)	1.07 (0.28)	1.32 (0.28)	0.95 (0.12)
$R_{rs}^{-1}(665)R_{rs}(708)$	1.47 (0.42)	1.31 (0.38)	1.76 (0.36)	0.93 (0.11)
3-Band	1.17 (0.31)	1.09 (0.29)	1.33 (0.30)	0.81 (0.22)
$R_{rs}^{-1}(560)R_{rs}(665)$	2.20 (1.60)	1.48 (1.52)	3.52 (0.62)	0.70 (0.15)
G08	1.26 (0.43)	1.30 (0.38)	1.17 (0.54)	1.26 (0.16)
MERIS	1.11 (0.65)	1.31 (0.70)	0.74 (0.30)	1.46 (0.43)

Mean ratio= average of the ratios of the predicted chl-*a* values to the measured chl-*a* values.

Table 6.6 Qualitative comparison between NDCI and chl-*a* concentration from all study regions.

NDCI range	Chl-*a* range (mg m^{-3})
< -0.1	< 7.5
-0.1 to 0	7.5-16
0 to 0.1	16-25
0.1 to 0.2	25-33
0.2 to 0.4	33-50
0.4 to -0.5	> 50
0.5 to 1	Severe bloom

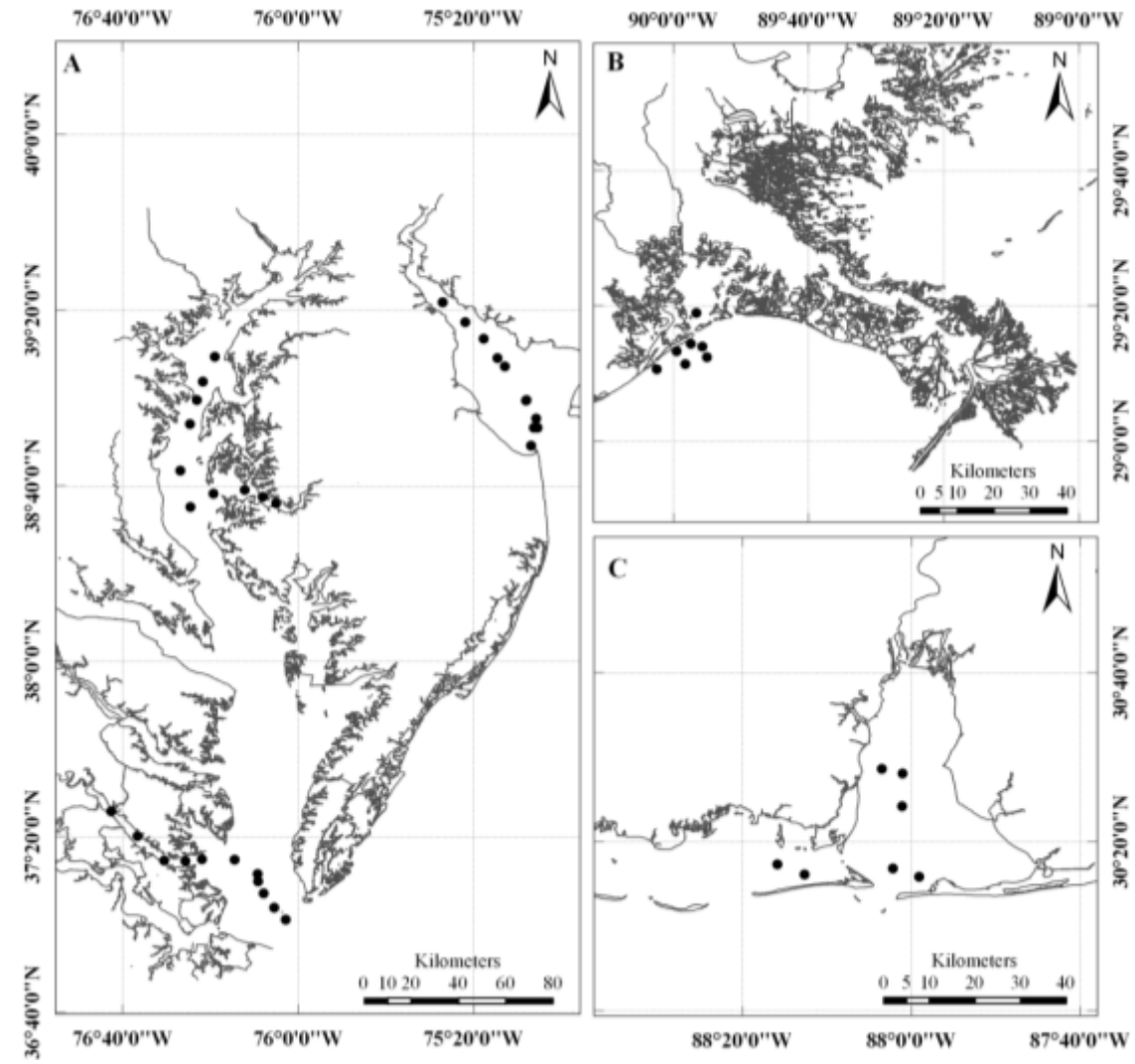

Figure 6.1 Location map of data points used in the present study.

(A) Chesapeake Bay and Delaware Bay, (B) River Mississippi Delta region in the northern Gulf of Mexico, and (C) Mobile Bay.

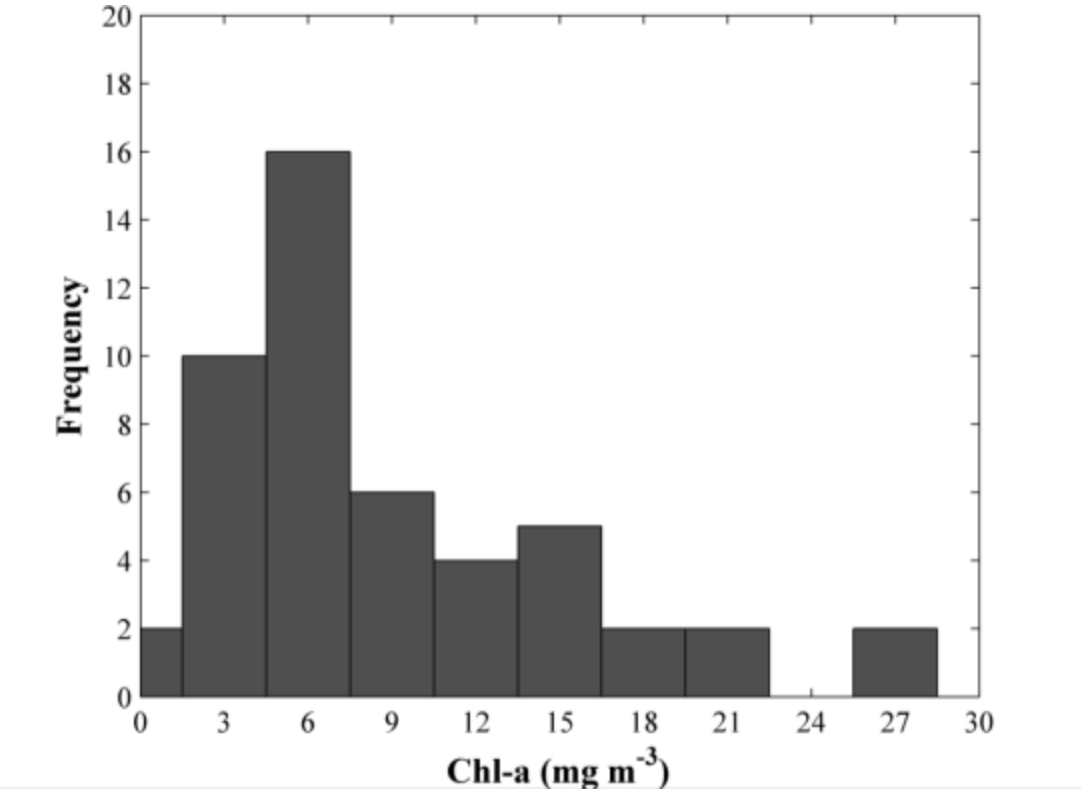

Figure 6.2 Frequency histogram of surface chl-*a* concentration (mg m^{-3}) measured from Chesapeake and Delaware Bay, Mississippi Delta region, and Mobile Bay.

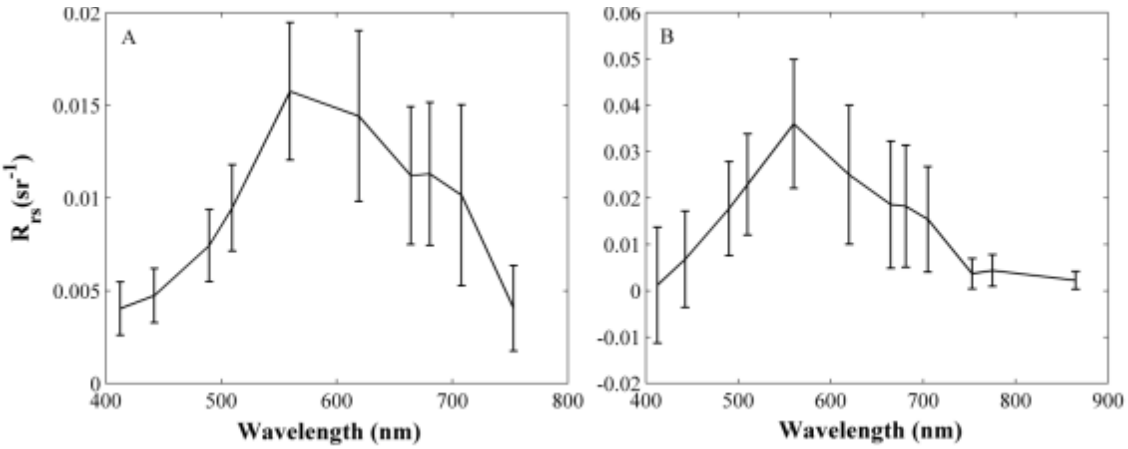

Figure 6.3 Average R_{rs} spectra derived from (A) the simulated dataset (n=200), (B) MERIS images of the study regions (n=49).

Y-error bars are the one STD of R_{rs} at MERIS band centers.

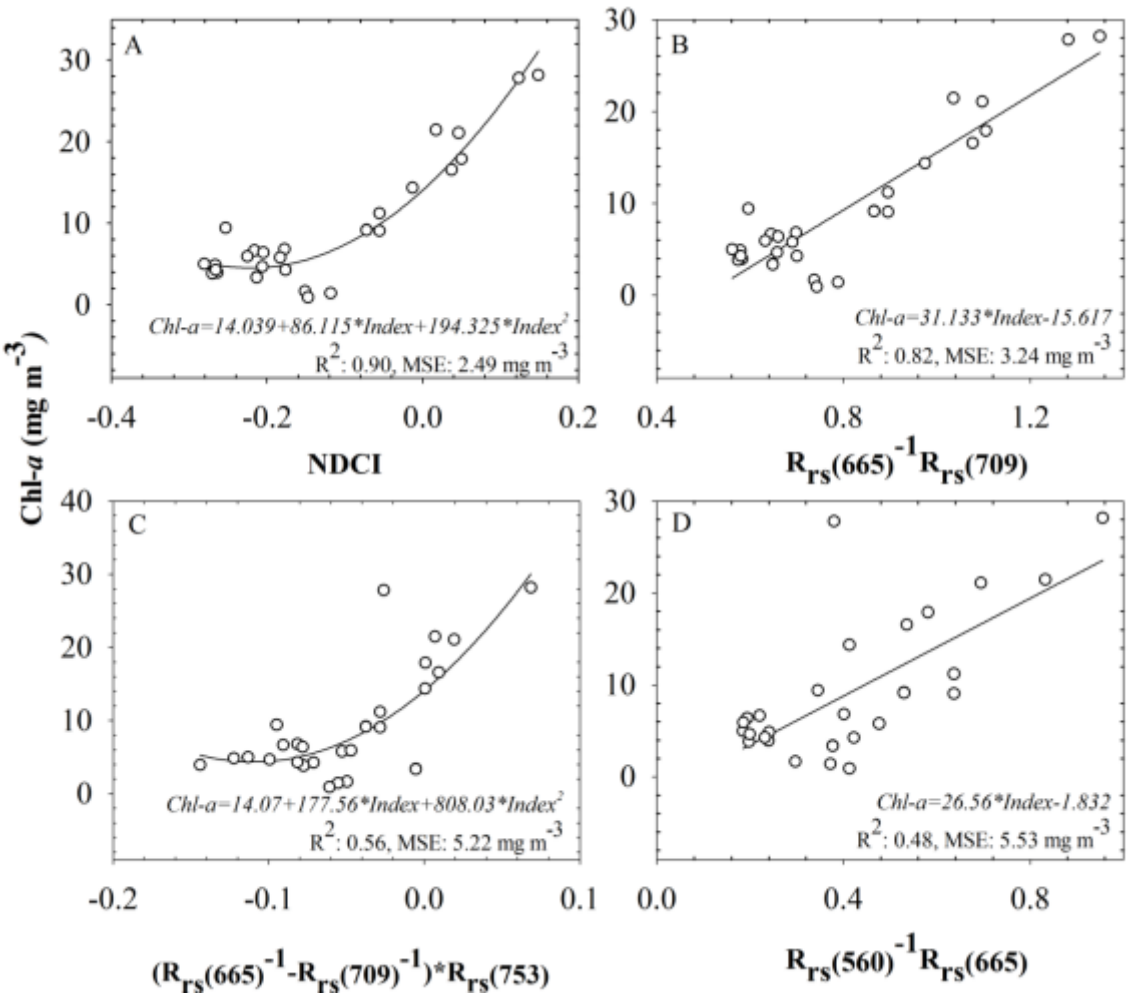

Figure 6.4 Calibration plots from the first calibration dataset that was sampled based on solar zenith angle:

(A) NDCI, (B) M09, (C) D05, and (D) T07.

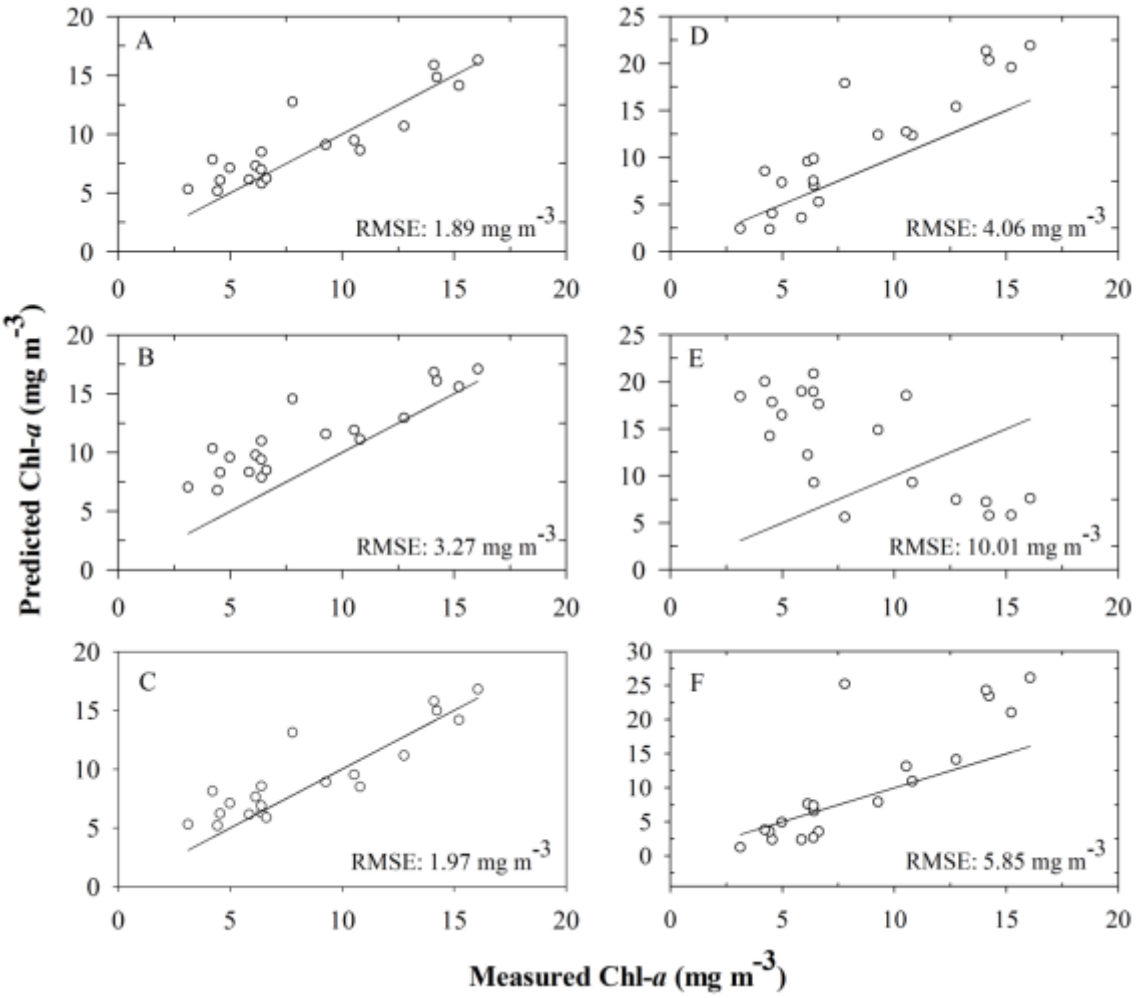

Figure 6.5 Validation plots from the first validation dataset that was sampled based on solar zenith angle: (A) NDCI (B) M09, (C) D05, (D) G08, (E) T07, and (F) MERIS case 2 chl-*a* product (Algal-2).

Straight lines on the plots are the 1-to-1 lines.

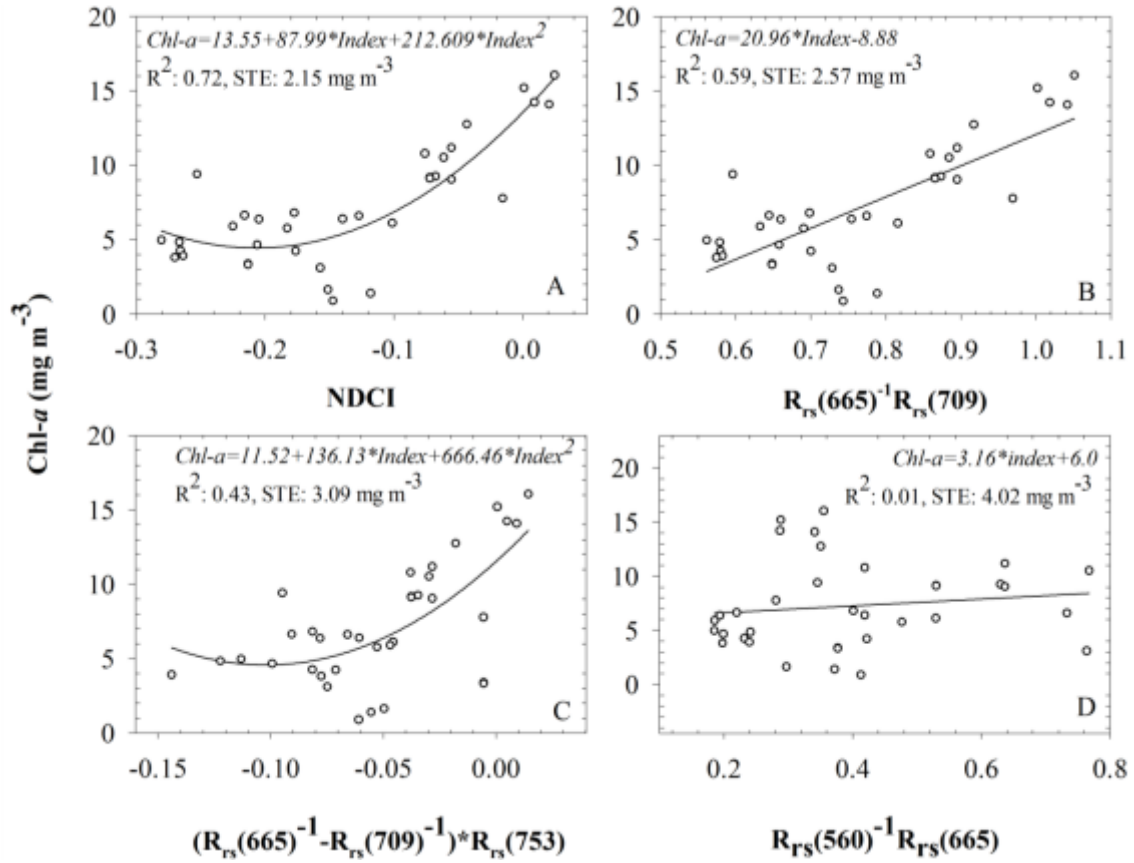

Figure 6.6 Calibration plots from the third calibration dataset that was sampled based on geographic region:

(A) NDCI, (B) M09, (C) D05, and (D) T07. The calibration dataset contained data points from the Chesapeake Bay and Delaware Bay.

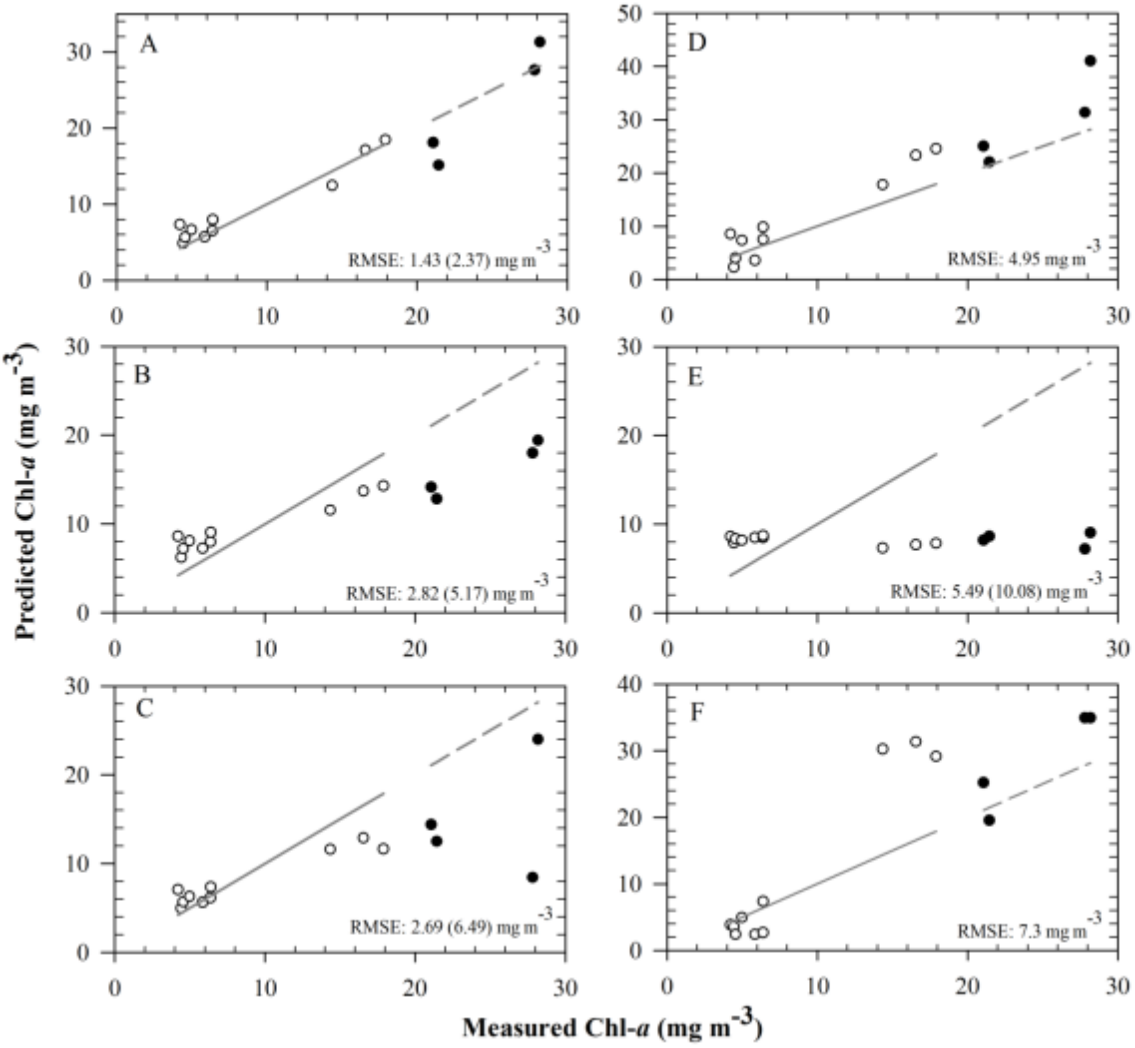

Figure 6.7 Validation plots from the third validation dataset that was sampled based on geographic region:

(A) NDCI (B) M09, (C) D05, (D) G08, (E) T07, and (F) MERIS case 2 chl-*a* product (Algal-2). Validation dataset contained data points from the Mobile Bay and the River Mississippi Delta. Solid circles represent the predictions outside of the calibrated chl-*a* range. Solid and dashed lines are the 1:1 lines for the predictions inside and outside of the calibrated chl-*a* range respectively. RMSE of the predictions within the calibrated chl-*a* range (values outside the brackets) and the RSME of the entire dataset (values inside the bracket) are provided.

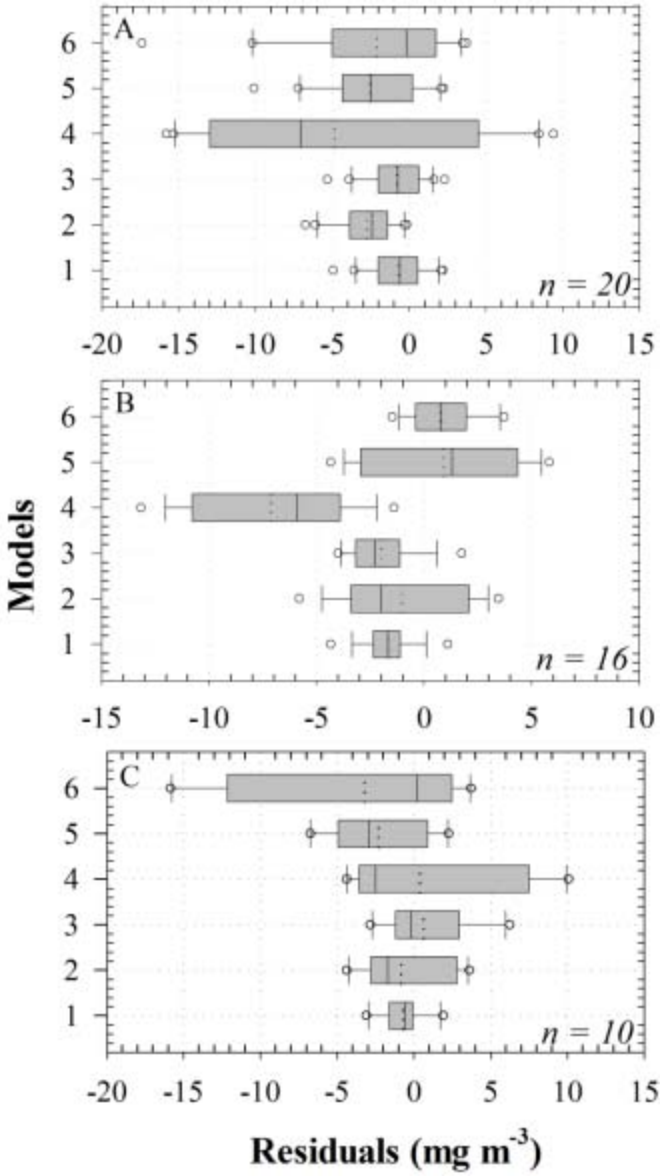

Figure 6.8 Box plots showing residuals from three fold validation, (A) dataset based on solar zenith angle, (B) dataset based on solar azimuth angle, and (C) dataset based on geographic regions of all models such as (1) NDCI, (2) M09, (3) D05, (4) T07, (5) G08, and (6).

MERIS chl-*a* product (Algal-2). The dotted lines inside boxes are the mean lines.

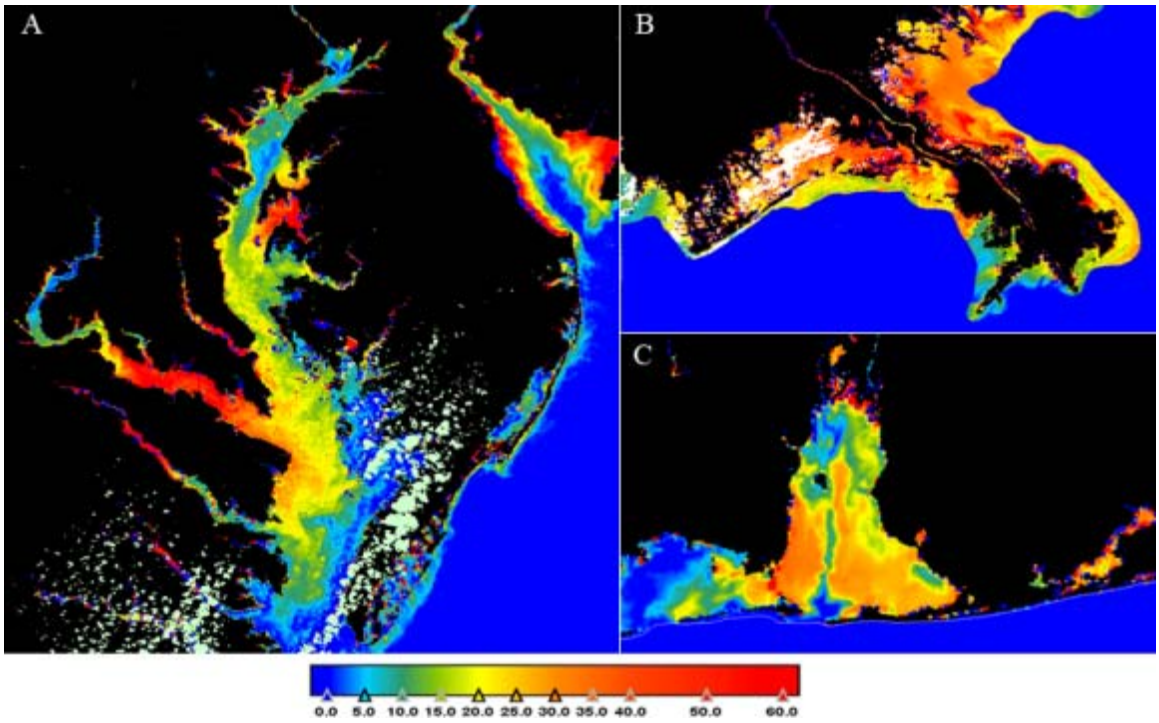

Figure 6.9 **(**A) Spatial distribution of chl-*a* map in Chesapeake Bay, upper Bay: north of 39^{o}N, mid-Bay (37.5^{o}N -39^{o}N), and lower-Bay (south of 37.5^{o}N).

Delaware Bay is located in the north-east corner, (B) Spatial distribution of chl-*a* in Mississippi delta region (pixels with no data were shown in white), and (C) in the Mobile Bay. Light green patches in the figure are clouds and the black and blue color represents land and open oceans respectively. Concentration of chl-*a* are in mg m^{-3}.

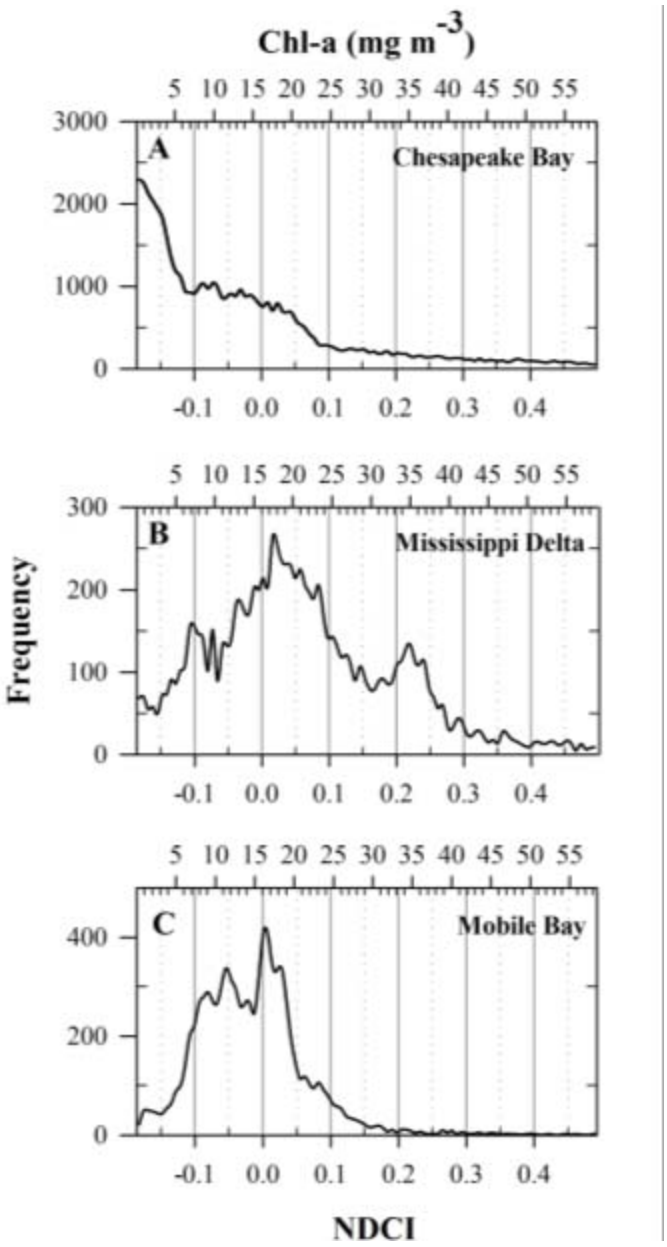

Figure 6.10 Relationship between NDCI range and chl-*a* concentration at three unique case 2 water bodies. (A) Chesapeake Bay, (B) Mississippi Delta, and (C) Mobile Bay.

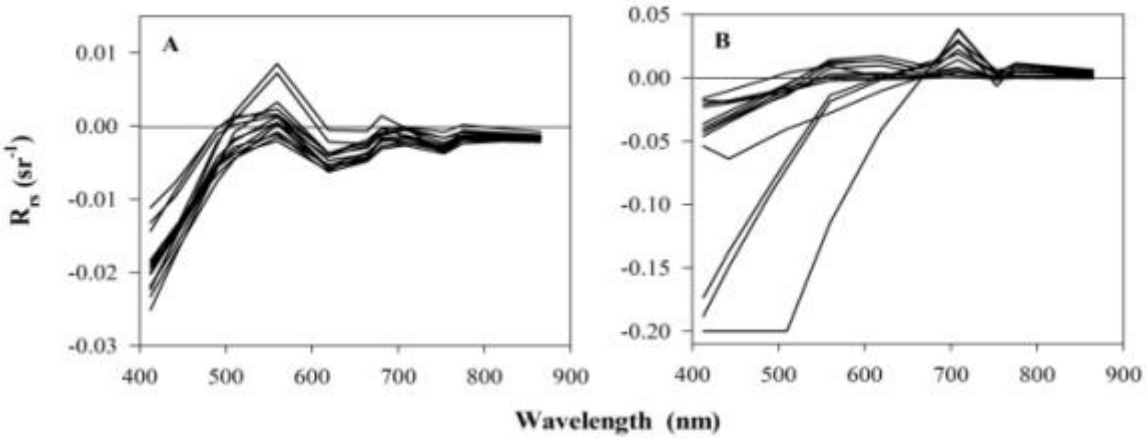

Figure 6.11 Typical reflectance spectra of contaminated pixels in MERIS Level 2 reflectance product.

(A) Spectra that produced extreme negative chl-*a* concentration, and (B) Spectra that produced extreme positive predictions.

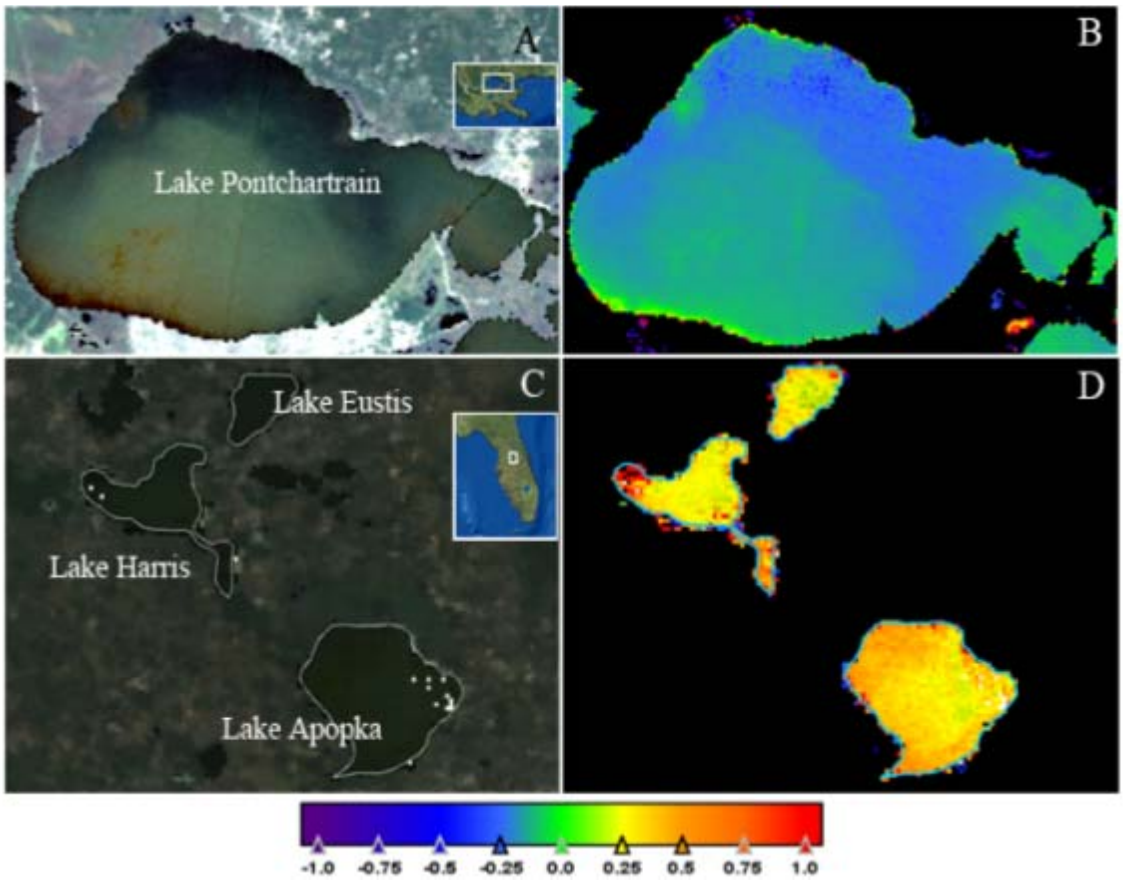

Figure 6.12 (A) True color MERIS image of Lake Pontchartrain, LA, USA acquired on October 14, 2010, (B) Corresponding NDCI image, (C) True color MERIS image of Lake Apopka, Lake Harris, and Lake Eustis, FL, USA, acquired in April 29, 2010 (White color represents pixels with no data), and (D) Corresponding NDCI image.

Location maps are provided in insets.

6.7 References

Babin, M., Stramski, D., Ferrari, G. M., Claustre, H., Bricaud, A., Obolensky, G., *et al*. (2003). Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *Journal of Geophysical Research-Oceans*, *108*(C7), 3211.

Bricaud, A., Morel, A., & Prieur, L. (1981). Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains, *Limnology and Oceanography*, *26*, 43-53.

Carder, K. L., & Steward, R. G. (1985). A remote-sensing reflectance model of a red-tide dinoflagellate off west Florida, *Limnology and Oceanography*, *30*, 286-298.

Carrick, H. J., & Schelske, C. L. (1997). Have we overlooked the importance of small phytoplankton in productive waters?, *Limnology and Oceanography*, *42*, 1613-1621.

Ciotti, A. M., Lewis, M. R., & Cullen, J. J. (2002). Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnology and Oceanography*, *47*, 404-417.

Dall'Olmo, G., Gitelson, A., Rundquist, D., Leavitt, B., Barrow, T., & Holz, J. (2005). Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and nearinfrared bands, *Remote Sensing of Environment*, *96*, 176–187.

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. A. (2006). Effect of bio-optical parameter variability and uncertainties in reflectance measurements on the remote estimation of chlorophyll-*a* concentration in turbid productive waters: Modeling results, *Applied Optics*, 45, 3577–3592.

Darecki, M., Weeks, A., Sagan, S., Kowalczuk, P., & Kaczmarek, S. (2003). Optical characteristics of two contrasting case 2 waters and their influence on remote sensing algorithms, *Continental Shelf Research*, *23*, 237-250.

Gilerson, A., Zhou, J., Hlaing, S., Schalles, J., Gross, B., Moshary, F., & Ahmed, S. (2007). Fluorescence component in the reflectance spectra from coastal waters-Dependence on water composition, *Optics Express*, *15*, 15702-15721.

Gons, H. J.,Auer, M. T., & Effler, S. W. (2008). Meris satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes, *Remote Sensing of Environment*, *112*, 4098-4106.

Gons, H. J., Burger-Wiersma, T., Otten, J. H., & Rijkeboer, M. (1992). Coupling of phytoplankton and detritus in a shallow, eutrophic lake (Lake Loosdrecht, The Netherlands). *Hydrobiology*, *233*, 51–59.

Gons, H. J., Rijkeboer, M. & Ruddick, K. G. (2002). A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *Journal of Plankton Research*, *24*, 947–951.

Gordon, H. R., Brown, O. B., & Jacobs, M. M. (1975). Computed relationships between the inherent and apparent optical properties of a flat, homogeneous ocean, *Applied Optics*, *14*, 417–427.

Jerlov, N. G. (1968). *Optical Oceanography*, New York: Elsevier, 1968

Kiefer, D. A., & Reynolds, R. A. (1992). Advances in understanding phytoplankton fluorescence and photosynthesis. In. P. G. Falkowsky, & A. D. Woodhead (Eds.), Primary productivity and biogeochemical cycles in the sea (pp. 155-174). Plenum Press.

Kirk, J. T. O. (1984). Dependence of relationships between inherent and apparent optical properties of water on solar altitude, *Limnology and Oceanography, 29*, 350–356.

Kutser, T., Hiire, M., Metsamaa, L., Vahtmae, E., Paavel, B., & Aps, R. (2009). Field measurements of spectral backscattering coefficient of the Baltic Sea and boreal lakes, *Boreal Environment Research*, *14*, 305-312.

Le, C., Li, Y., Zha, Y., Sun, D., Huang, C., & 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 Environment* , *113*, 1175-1182.

Magnuson, A., Hardings, L.W., Jr., Mallonee, M.E., & Adolf, J.E. (2004). Bio-optical model for Chesapeake Bay and the Middle Atlantic Bay, *Estuarine Coastal and Shelf Sciences*, *61*, 403-424.

Maritorena, S., Siegel, D.A., & Peterson, A.R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, *41*, 2705–2714.

Miller, R. L., Belz, M., Castillo, C. D., & Trzaska, R. (2002). Determining CDOM absorption spectra in diverse coastal environments using a multiple pathlength, liquid core waveguide system, *Continental Shelf Research*, *22*, 1301–1310.

Mobley, C. D. (1994). Light and water-Radiative transfer in natural waters, San Diago: Academic Press, (Chapter 5).

Mobley,C.D. (1999). Estimation of the Remote-Sensing Reflectance from Above-Surface Measurements, *Applied Optics*, *38*, 7442-7455.

Mobley, C.D., & Sundman, L.K. (2001). Hydrolight 4.2, Sequoia Scientific, Inc.

Morel, A., & Gentili, B. (1996). Diffuse reflectance of oceanic waters . remote-sensing problem, *Applied Optics*, *35*, 4850-4862.

Morel, A. & Prieur, L. (1977). Analysis of variations in ocean color, *Limnology and Oceanography*, *22(4)*, 709–722.

Moses, W. J., A. A. Gitelson, S. Berdnikov, & Povazhnyy, V. (2009). Satellite Estimation of Chlorophyll-a Concentration Using the Red and NIR Bands of MERIS—The Azov Sea Case Study. *IEEE Geoscience and Remote Sensing Letters, 6 (4),* 845-849.

O'Reilly, J., Maritorena, S., Mitchell, B., Siegel, D., Carder, K., Garver, S., Kahru, M., & McClain, C. (1998). Ocean color chlorophyll algorithms for Sea-WiFS, *Journal of Geophysical Research*, 24 937–24 953.

O'Reilly, J., Maritorena, S., O'Brien, M. C., Siegel, D., Toole, D., Menzies, D., Smith, R. C., Mueller, J. L., Mitchell, B. G., Kahru, M., Chavez, F. P., Strutton, P. G., Cota, G. F., Hooker, S. B., McClain, C. R., Carder, K. L., Muller-Karger, F., Harding, L., Magnuson, A., Phinney, D., Moore, G. F., Aiken, J., Arrigo, K. R., Letelier, R. M., & Culver, M. (2000). SeaWiFS postlaunch calibration and validation analyses, Part 3, NASA Goddard Space Flight Center, Greenbelt, MD, NASA Tech. Memo. 2000- 206892, *11*.

Pope, R., & Fry, E. (1997). Absorption spectrum (380 - 700 nm) of pure waters: II. Integrating cavity measurements, *Applied Optics*, *36*, 8710-8723.

Simis, S. G. H. (2006). Blue-Green Catastrophe: remote sensing of mass viral lysis of cyanobacteria, PhD dissertation, Vriji University, Amsterdam.

Tzortziou, M., Subramaniam, A., Herman, J. R., Gallegos, C. L., Neale, P. J., & Harding, L. W., Jr. (2007). Remote sensing reflectance and inherent optical properties in the mid Chesapeake Bay, *Estuarine Coastal and Shelf Science*, *72, 16*-32.

Werdell, P. J., & Baily, S.W. (2005). An improved In situ bio-optical data set for ocean color algorithm development and satellite data product validation, *Remote Sensing of Environment*, *98*, 122-140.

CHAPTER VII

CONCLUSIONS AND FUTURE RECOMMENDATIONS

7.1 Conclusions

This dissertation investigates the potential of remote sensing techniques to quantify cyanobacterial pigment concentration, e.g. chlorophyll-*a* (chl-a) and phycocyanin (PC), and develops empirical and quasi analytical algorithms for effective monitoring of cyanobacterial algal blooms using remotely sensed data. The major contributions from the dissertation are provided below.

Empirical algorithms using reflectance band ratios, $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$, were developed to quantify phycocyanin (PC) concentration in cyanobacteria. Model outputs were compared with the performance of a semi-analytical PC algorithm (Simis *et al*., 2005). Empirical models produced lower estimation errors as compared to the semi-analytical model in our dataset. Mean relative errors (RE) of $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$ models were 25.4% and 21.6% respectively. However, the median RE of $R_{rs}(708)/R_{rs}(600)$ (~20%) was 2% lower than the $R_{rs}(708)/R_{rs}(620)$ model. Values of model parameters γ and δ used in the semi-analytical algorithm (Simis *et al*., 2005) drastically changed in our study region. Since the model parameters hold a high order of importance in the algorithm because of their critical role in retrieving $a_{\varphi}(620)$ and $a_{\varphi}(665)$ values, they should not be treated as constants, as in Randolph *et al*. (2008) and Hunter *et al*. (2010), and transferred to other study regions

with considerably different optical conditions. The biggest challenges of the semi-analytical algorithm is to find the appropriate values of γ and δ for new study regions or even in the same region but at a different time period. In cyanobacteria dominated hypereutrophic waters, variability in $a_{PC}^{*}(620)$ and uncertainty in the value of ε, that relates $a_{chl}(665)$ with $a_{chl}(620)$, could cause erroneous retrieval of PC concentrations.

In order to address the limitations/challenges with empirical and semi-analytical approach, quasi-analytical approach was used to retrieve $a_{\varphi}(620)$ and $a_{\varphi}(665)$ from $R_{rs}(\lambda)$. The QAA algorithm has been extensively calibrated and validated using simulated and field data in different geographic regions (Lee *et al*., 2002; Lee *et al*., 2004; Le *et al*., 2009, Craig *et al*., 2006; Zhu *et al*., 2011). Accuracy of QAA inversion is quite promising in clear waters with $a_t(440)<0.3 m^{-1}$. In one of the previous studies carried out using field data, percent error of $a_t(\lambda)$ retrieval was 6.9% in clearer waters and the retrieval error increased to 10.3% for less clear waters with $a_t(440)$ ranging from 0.025-2.3 m^{-1}(Lee *et al*., 2002). In another study, QAA algorithm successfully retrieved $a_t(\lambda)$ in turbid productive waters of Taihu Lake, China (where $a_t(440)$ approaches ~12.0 m^{-1}) with an average error of 8.95% within 400 to 700 nm wavelength range (Le *et al*., 2009). However, algorithm performance was not satisfactory in highly turbid and productive waters encountered in our study region where $a_t(443)$ approaches ~47.0 $m^{-1}$and $a_{\varphi}(443)$ contributes >54% of $a_t(443)$. Therefore, the algorithm was parameterized using an extensive dataset comprising *in situ* radiometric measurements, absorption coefficients of phytoplankton, CDOM, detritus, and pigment data from highly turbid and cyanobacteria dominated aquaculture ponds. After parameterization, the algorithm retrieved $a_t(\lambda)$ values with a percent error of ~15-24 % and an average error of 19.87% within a wavelength

range of 413-665 nm. Similarly, percentage error for phytoplankton absorption coefficient, $a_\varphi(\lambda)$, varied from 16.94-34.15 % within 413-665 nm range and the average error was 23.93%. Results from this research show that the reference wavelength in the QAA should be moved to 708 nm for retrieving $a_t(\lambda_0)$. In eutrophic to hypereutrophic waters, $a_t(\lambda_0)$ should not be assumed as $a_w(\lambda_0)$; rather $a_t(\lambda_0)$ should be retrieved as shown in the newly parameterized model (Mishra *et al*., 2012). For the first time, this research documents bio-optical inversion in such cyanobacteria dominated hypereutrophic waters.

The newly parameterized QAA was used to retrieve $a_\varphi(665)$ and $a_\varphi(620)$ and further decomposed the $a_\varphi(620)$ to obtain PC absorption at that wavelength, $a_{PC}(620)$. Two new coefficients ψ_1 [$a_{chl}(665)/a_{chl}(620)$] and ψ_2, [$a_{PC}(665)/a_{PC}(620)$] were introduced for the decomposition of $a_\varphi(620)$ using an algebraic method. Mean RE of $a_{PC}(620)$ retrieval was ~10% when optimized values of ψ_1 and measured values of ψ_2 were used (representing an ideal case). Similarly, mean RE of PC prediction was ~10% and 22.32% when sample specific $a^*_{620}(620)$ and the mean $a^*_{620}(620)$ were used. From a practical point of view, when modeled values of ψ_1 and of ψ_2 were used, mean RE of $a_{PC}(620)$ and PC retrieval increased to ~28% and 36% respectively. It should be noted that the retrieval accuracy obtained in this study using the quasi-analytical PC algorithm is far superior to the semi-analytical PC algorithm accuracy reported in Simis *et al*. (2005). The semi-analytical algorithm produced an average error of 19.7% when sample specific model coefficients (γ, δ, and ε) and sample by sample $a^*_{620}(620)$ was used (Simis *et al*., 2005). However, the errors were extremely high with increasing overestimation for the samples with lower PC concentration when fixed values of γ, δ, ε, and $a^*_{620}(620)$ were used (not reported).

Finally, the absolute relative error (ARE) (%) of PC estimation from all models including empirical models, semi-analytical, and quasi-analytical algorithms were compared (Table 7.1, Fig. 7.1). Errors from all models were relatively higher for samples with PC concentration < 250 mg m^{-3} as compared to samples with higher PC values. Overall, empirical models produced relatively better accuracy, where mean and median ARE were within ~20 and 25%. However, errors from the empirical models should not be directly compared with errors from other models because the validation analysis for the empirical models used only 9 points, whereas, all data points were used in validation of other algorithms. Semi-analytical algorithm produced the largest amount of error among all. The mean and median errors for the semi-analytical algorithm were 45% and 73% respectively. Median ARE from QAA decreased from ~26% to 24% when modeled values of $a_{PC}^{*}(620)$ were used instead of a mean $a_{PC}^{*}(620)$ (Table 7.1).

A chlorophyll-*a* (chl-*a*) threshold is often used to monitor cyanobacteria whenever remote estimation of PC is not feasible. World Health Organization uses a guideline of approximately 50 mg m^{-3} of chl-*a* in cyanobacteria as a moderate health warning level; which is equivalent to a PC concentration of about 50-100 mg m^{-3} based on an average intercellular PC:chl-*a* ratio of 1-2 (WHO, 2003). PC:chl-a ratio can also be used as an indicator of cyanobacterial abundance and a threshold of this ratio may also be used to decide when PC retrieval may be erroneous. Therefore, an index, normalized difference chlorophyll index (NDCI), was proposed to predict chl-*a* concentration from remote sensing data in estuarine and coastal turbid productive (case 2) waters. The model was conceptualized using a large simulated dataset. Further, it was calibrated and validated using field data from three unique study regions such as: Chesapeake/Delaware

Bay, Mobile Bay, and the river Mississippi delta region in the Northern Gulf of Mexico. The newly developed index showed strong relationship with chl-*a* concentration (R^2= 0.9, p<0.001) and yielded lower estimation error (RMSE~2 mg m^{-3}, chl-*a* range: 0.9-28.1 mg m^{-3}) as compared to many existing chl-*a* empirical algorithms for turbid productive waters. NDCI chl-*a* algorithm was applied to MERIS images over the Chesapeake Bay and Delaware Bay, the Mobile Bay, and the Mississippi River delta region and spatial distribution of surface chl-a was mapped. NDCI was successful in predicting chl-*a* concentration with approximately 12% overall bias for all above study regions.

7.2 Future research

In the past, most of the remote sensing study of water quality and clarity have addressed algorithm development as well as issues of temporal and spatial variation of chl-*a*, (Carder *et al*., 1999, O'Reilly *et al*, 1998; O'Reilly *et al*, 2000, Gons *et al*., 2002, Dallolmo *et al*, 2005; Gons *et al*., 2008; Mishra and Mishra, 2012), suspended particulate matter (Miller and Mckee, 2002; D'Sa *et al*., 2007), and CDOM (Bricaud *et al*., 1981; Carder *et al*, 1999, Kutser *et al*., 2005; D'Sa *et al*., 2006; Ahn *et al*., 2008). However, very limited numbers of research have focused on remote sensing of PC (Dekker, 1993; Schalles and Yacobi, 2000; Vincent *et al*., 2004; Simis *et al*., 2005; Mishra *et al*., 2009). Most of the cyanobacteria remote sensing researches have been carried out in very selective inland lakes, estuaries and coastal waters (e.g., Lake Erie, Morse and Geist Reservoirs, IN, USA, Baltic Sea, Taihu Lake, China, and lakes are reservoirs in Spain and Netherlands). Future research should focus extensively on collection of cyanobacterial pigment data including PC, as well as inherent optical properties from geographically diverse and large water bodies for algorithm calibration/validation.

Especially, the newly developed algorithms in this dissertation are calibrated/optimized using a small dataset, but with a wide range of pigment concentration, collected from eutrophic to hypereutrophic waters. Many empirical models developed in this dissertation should be further calibrated and validated using a large dataset collected from different water bodies for accuracy assessment as well as to increase their transferability to other geographic regions.

More importantly, very few researches have utilized remotely sensed satellite data products for calibration and validation of existing PC algorithms for monitoring cyanobacteria. Future effort should focus on application of algorithms to satellite data and develop cyanobacterial algal bloom related bio-physical data products for dissemination to the science community as well as to general public. At present, all PC algorithms use radiometric information measured at 620 nm. Fortunately a few existing and future ocean color sensors such as (1) Medium Resolution Imaging Spectrometer (MERIS), a European space born remote sensor, (2) Hyperspectral Imager for the Coastal Ocean (HICO), and (3)The Hyperspectral Infrared Imager (HyspIRI) offers radiometric data at that band center. Data products from these sensors can be used to develop tools to be used in an early warning system for detecting toxic algal blooms, and mapping the spatial and temporal distribution of toxic cyanobacteria in inland and coastal waters.

Table 7.1 Statistics of absolute relative error (%) of PC retrieval from all models.

	Min	Max	Median	Mean
Empirical, $R_{rs}(708)/R_{rs}(600)$				
PC<250 mg m^{-3}	8.16	43.62	34.15	28
PC >250 mg m^{-3}	16.6	16.87	16.74	16.74
All Data	8.16	43.62	19.96	25.5
Empirical, $R_{rs}(708)/R_{rs}(620)$				
PC <250 mg m^{-3}	5.15	36.58	21.86	22.95
PC >250 mg m^{-3}	12.31	22.09	17.2	17.2
All Data	5.15	36.58	21.86	21.67
Semi-Analytical				
PC <250 mg m^{-3}	0.02	239.69	80.32	94.7
PC >250 mg m^{-3}	3.64	45.64	35.79	27.13
All Data	0.02	239.69	45.96	73.08
Quasi-Analytical, mean $a^*_{PC}(620)$				
PC <250 mg m^{-3}	0.42	250.62	46.52	56.58
PC >250 mg m^{-3}	0.14	34.6	12.72	14.42
All Data	0.14	250.62	25.96	43.09
Quasi-Analytical, Modeled $a^*_{PC}(620)$				
PC <250 mg m^{-3}	2.83	278.8	44.74	57.79
PC >250 mg m^{-3}	3.87	57.6	17.76	23.42
All Data	2.83	278.8	23.82	46.79

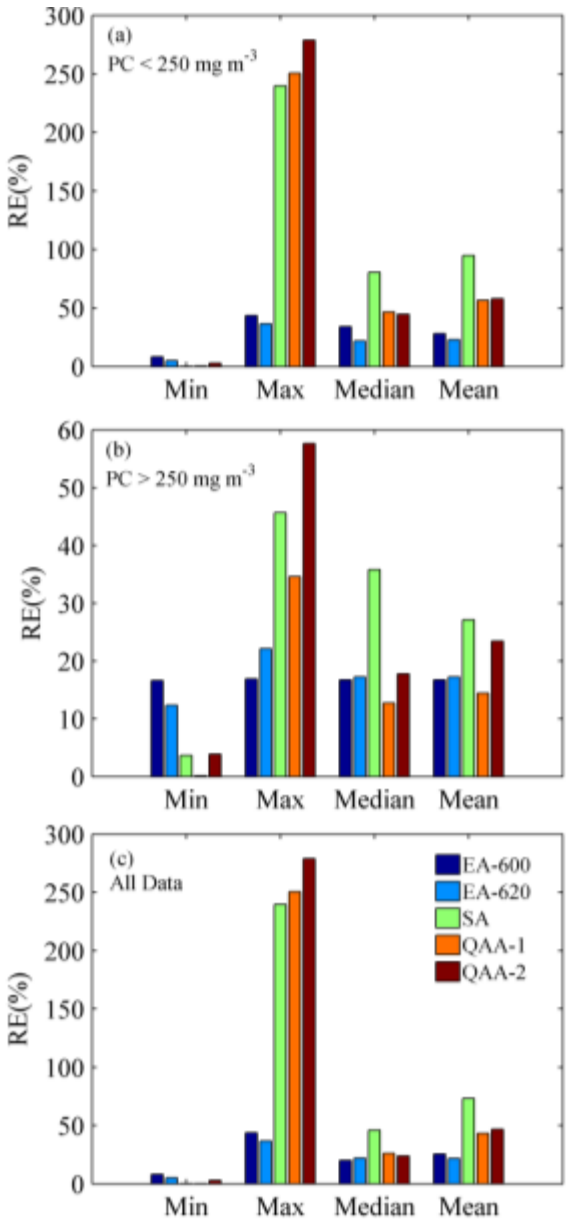

Figure 7.1 Comparison of absolute relative errors of PC retrieval from all models; in the legend, EA-600, EA-620 represents empirical models using reflectance ratios $R_{rs}(708)/R_{rs}(600)$ and $R_{rs}(708)/R_{rs}(620)$ respectively; SA is the semi-analytical algorithm; QAA-1 and QAA-2 represents QAA with mean $a^*_{PC}(620)$ and modeled $a^*_{PC}(620)$.

Min, max, median, and mean error of (a) samples with PC concentration <250 mg m^{-3}, (b) samples with PC concentration >250 mg m^{-3}, and (c) all data points.

7.3 References

Ahn, Y. H., Shanmugam, P., Moon, J. E., & Ryu, J. H. (2008). Satellite remote sensing of a low-salinity water plume in the East China Sea. Annales Geophysicae, 6, 2019−2035.

Bricaud, A., Morel, A., Prieur, L. (1981). Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domain. *Limnology & Oceanography*, 26, 43-53.

Carder, K. L., Chen, F. R., Lee, Z. P., Hawes, S. K. & Kamykowski, D. (1999). Semianalytic Moderate-Resolution Imaging Spectrometer algorithms for chl-*a* and absorption with bio-optical domains based on nitrate-depletion temperatures, *Journal of Geophysical Research*, 104, 5403–5421.

Craig, S. E., Lohrenz, S. E., Lee, Z. P., Mahoney, K. L., Kirkpatrick, G. J., Schofield, O. M., & Steward, R. G. (2006). Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga, Karenia brevis, *Applied Optics*, 45 (21): 5414-5425.

Dall'Olmo, G. & Gitelson, A.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(3): 412-422.

Dekker, A. (1993). Detection of the optical water quality parameters for eutrophic waters by high resolution remote sensing, Ph.D. thesis, Free University, Amsterdam.

D'Sa, E.J., Miller, R.L. & Del Castillo, C. (2006). Bio-optical properties and ocean color algorithms for coastal waters influenced by the Mississippi River during a cold front. *Applied Optics*, 45:7410-7428.

D'Sa, E.J., Miller, R.L. & B.A. McKee. 2007. Suspended particulate matter dynamics in coastal waters from ocean color: applications to the northern Gulf of Mexico. *Geophysical Research Letters*, doi:10.1029/2007GL031192.

Gons, H. J.,Auer, M. T., & Effler, S. W. (2008). Meris satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes, *Remote Sensing of Environment*, *112*, 4098-4106.

Gons, H. J., Rijkeboer, M. & Ruddick, K. G. (2002). A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *Journal of Plankton Research*, *24*, 947–951.

Hunter, P.D., Tyler, A.N., Carvalho, L., Codd, G.A. & Maberly, S.C. (2010) Hyperspectral remote sensing of cyanobacterial pigments as indicators for cell populations and toxins in eutrophic lakes. *Remote Sensing of Environment,* 114 (11): 2705-2718.

Le, C.F., Li, Y.M., Zha, Y., Sun, D., & Yin, B. (2009). Validation of a Quasi-Analytical Algorithm for Highly Turbid Eutrophic Water of Meiliang Bay in Taihu Lake, China, *IEEE Transactions on Geosciences and Remote Sensing*, 47 (8), 2492-2500.

Lee, Z.P., Carder, K.L., R.A., & Arnone, R. (2002). Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters, *Applied Optics,* 41: 5755-5772.

Lee, Z.P., & Carder, K.L. (2004). Absorption spectrum of phytoplankton pigments derived from hyperspectralremote-sensing refl ectance. *Remote Sensing of Environment,* 89: 361-368.

Miller, R.L. & Mckee, B.A. (2004). Using MODIS Terra 250 m imagery to map concentrations of total suspended matter in coastal waters, *Remote Sensing of Environment*, 93, 259-266.

Mishra, S. & Mishra, D. R. (2012). Normalized Difference Chlorophyll Index: A Novel Model for Remote Estimation of Chlorophyll-*a* Concentration in Turbid Productive Waters, *Remote Sensing of Environment*, 117, 394-406.

Mishra, S., Mishra, D.R., & Schluchter, W.M. (2009). A Novel Algorithm for Predicting PC Concentrations in Cyanobacteria: A Proximal Hyperspectral Remote Sensing Approach. *Remote Sensing*, *1*: 758-775.

Mishra, S., Mishra, D. R., & Lee, Z.P. (2012). Bio-Optical Inversion in Highly Turbid and Cyanobacteria Dominated Waters, *IEEE Transactions on Geoscience and Remote Sensing* (Under Review)

O'Reilly, J., Maritorena, S., Mitchell, B., Siegel, D., Carder, K., Garver, S., Kahru, M., & McClain, C. (1998).Ocean color chlorophyll algorithms for Sea-WiFS, *Journal of Geophysical Research,* 103, 24 937–24 953.

O'Reilly, J., Maritorena, S., O'Brien, M. C., Siegel, D., Toole, D., Menzies, D., Smith, R. C., Mueller, J. L., Mitchell, B. G., Kahru, M., Chavez, F. P., Strutton, P. G., Cota, G. F., Hooker, S. B., McClain, C. R., Carder, K. L., Muller-Karger, F., Harding, L., Magnuson, A., Phinney, D., Moore, G. F., Aiken, J., Arrigo, K. R., Letelier, R. M., & Culver, M. (2000). SeaWiFS postlaunch calibration and validation analyses, Part 3, NASA Goddard Space Flight Center, Greenbelt, MD, NASA Tech. Memo. 2000- 206892, *11*.

Randolph, K., Wilson, J., Tedesco, L., Li, L., Pascual, D., & Soyeux, M. (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.

Schalles, J., & Yacobi, Y. (2000). Remote detection and seasonal patterns of phycocyanin, carotenoid and chlorophyll-a pigments in eutrophic waters, *Archives fur Hydrobiologia –Sp. Issues Advances in Limnology,* 55:153-168.

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

Vincent, R. K., Qin, X. M., McKay, R. M. L. , Miner, J., Czajkowski, K., Savino, J., & Bridgeman, T. (2004). Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie*, Remote Sensing of Environment,* 89: 381-392.

Zhu, W., Yu, Q., Tian, Y. Q., Chen, R. F., & Gardner, G. B. (2011). Estimation of chromophoric dissolved organic matter in the Mississippi and Atchafalaya river plume regions using above-surface hyperspectral remote sensing, *Journal of Geophysical Research,* 116, no. C02011, 1-22.

APPENDIX A

STEPS OF QUASI-ANALYTICAL ALGORITHM (QAA)

QAA takes above surface remote sensing reflectance as input and first calculates sub-surface remote sensing reflectance r_{rs}, as:

$$r_{rs} = R_{rs}/(0.52 + 1.7R_{rs}) \quad \text{(A.1)}$$

r_{rs} is a function of u, the ratio of back scattering coefficient bb, to the sum of tatal absorption coefficients and backscattering coefficients.

$$u(\lambda) = \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \quad \text{(A.2)}$$

u can be empirically derived from r_{rs} as in Gordon *et al.*(1988):

$$u(\lambda) = \frac{-g0+\sqrt{(g_0)^2+4g_1*r_{rs}(\lambda)}}{2*g_1} \quad \text{(A.3)}$$

where $g_0 = 0.089$ and $g_1 = 0.125$. Further, the QAA estimates total absorption coefficients at a reference wavelength, λ_0. In the native form, QAA uses 555 or 650 nm as λ_0 based on the level of turbidity. However, in another study, Mishra *et al.*, (2012) have parameterized the QAA for extremely turbid and productive water where the λ_0 was moved to 708 nm because of strong absorption by non-water optical constituents even at 650 nm.

$$a(\lambda_0) = a_w(709) + 10^{-0.8125-2.3404\chi+1.24\chi^2} \quad \text{(A.4)}$$

$$where\ \chi = log_{10}\left(\frac{0.01*r_{rs}(442)+r_{rs}(620)}{r_{rs}(708)+0.005*\frac{r_{rs}(620)}{r_{rs}(442)}*r_{rs}(620)}\right) \quad \text{(A.5)}$$

The QAA then analytically retrieves particulate backscattering coefficients at the reference wavelength, $b_{bp}(\lambda_0)$.

$$b_{bp}(\lambda_0) = \frac{u(\lambda_0)a(\lambda_0)}{1-u(\lambda_0)} - b_{bw}(\lambda_0) \quad \text{(A.6)}$$

The spectral power, η, is empirically estimated from:

$$\eta = 2.0\left\{1 - 1.2\ exp\left(-0.9\frac{r_{rs}(443)}{r_{rs}(555)}\right)\right\} \quad \text{(A.7)}$$

Knowing $b_{bp}(\lambda_0)$ and η , the QAA estimates $b_{bp}(\lambda)$ at other wavelengths from:

$$b_{bp}(\lambda) = b_{bp}(\lambda_0)\left(\frac{\lambda_0}{\lambda}\right)^{\eta} \quad \text{(A.8)}$$

The total absorption coefficient is then calculated as:

$$a(\lambda) = \frac{(1-u(\lambda))\left(b_{bw}(\lambda)+b_{bp}(\lambda)\right)}{u(\lambda)} \quad \text{(A.9)}$$

The QAA further decomposed the total absorption spectrum into 1) a_{CDM}, a combined absorption by colored dissolved organic matter (CDOM) and detrital matter, and 2) $a_{\varphi}(\lambda)$.

$$a_{CDM}(443) = \frac{[a(411)-\zeta a(443)]-[a_w(411)-\zeta a_w(411)]}{\xi-\zeta} \quad \text{(A.10)}$$

Where $\zeta = \frac{a_{\varphi}(411)}{a_{\varphi}(443)} = 0.74 + \frac{0.2}{0.8+r_{rs}(443)/r_{rs}(555)}$ and $\xi = \frac{a_{CDM}(411)}{a_{CDM}(443)} = e^{S(443-411)}$

$a_{CDM}(\lambda)$ is then calculated using the exponential function as:

$$a_{CDM}(\lambda) = a_{CDM}(443)\ e^{-S(\lambda-443)} \quad \text{(A.11)}$$

Absorption coefficients by phytoplankton is then calculated as:

$$a_{\phi}(\lambda) = a_t(\lambda) - a_w(\lambda) - a_{CDM}(\lambda) \quad \text{(A.12)}$$